



Universidad Autónoma de Chiapas

Memorias del Congreso

ISBN:

968-607-477-588-4

Octubre 2011

Editor

Universidad Autónoma Metropolitana-Azcapotzalco
Av. San Pablo 180 Col. Reynosa Tamaulipas
Azcapotzalco, México D.F.

Coeditor

Universidad Autónoma de Chiapas
Carretera a Puerto Madero km. 1.5,
Tapachula, Chiapas. México

Autor

Varios

Compiladores

Luis Antonio Álvarez Oval
Miguel Magos Rivera
Ramiro Velázquez Guerrero

VII Semana Nacional de Ingeniería Electrónica
Memorias del Congreso

ISBN:

968-607-477-588-4

México D.F. 2011

Comité Organizador

Mtra. Paloma Ibáñez Villalobos

Rectora

Universidad Autónoma Metropolitana – Azcapotzalco

Mtro. Jaime Valls Esponda

Rector

Universidad Autónoma de Chiapas

Dr. Emilio Sordo Zabay

Director de la División de Ciencias Básicas e Ingeniería

Universidad Autónoma Metropolitana – Azcapotzalco

Mtro. Gonzalo Vázquez Nataren

Secretario Académico

Universidad Autónoma de Chiapas

Dr. Andrés Ferreyra Ramírez

Jefe del Departamento de Electrónica

Universidad Autónoma Metropolitana – Azcapotzalco

Mtra. Susana Patricia García Sampedro

Encargada de la Dirección de la Facultad de Contaduría Pública

Universidad Autónoma de Chiapas

Dr. Miguel Magos Rivera

Presidente Comité Organizador

Universidad Autónoma Metropolitana - Azcapotzalco

M. en C. Luis Antonio Álvarez Oval

Presidente Comité Local

Universidad Autónoma de Chiapas

Lic. Christian Mauricio Castillo Estrada y Mtro. Aron de la Cruz Vázquez

Coordinadores Locales

Universidad Autónoma de Chiapas

Dr. Ramiro Velázquez Guerrero

Coordinador Logístico

Universidad Panamericana - Aguascalientes

Comité Técnico

Dra. Abascal Mena Ma. Del Rocío	UAM-Cuajimalpa
Dr. Acosta Lúa Cuauhtémoc	Centro Universitario de la Ciénega-UDG
Dr. Adam Medina Manuel	CENIDET-Cuernavaca
Dra. Alcaraz Mejía Mildreth Isadora	Centro Universitario de la Ciénega-UDG
M.C. Alfaro Ruiz Victor M.	Universidad de Costa Rica
M.C. Álvarez Oval Luis Antonio	Universidad Autónoma de Chiapas
Dr. Anzures Marin Juan	Universidad Michoacana de San Nicolás de Hidalgo
Dr. Arechiga Martínez René	New Mexico Tech Electrical Engineering Department, U.S.A.
Dr. Astorga Zaragoza Carlos	CENIDET-Cuernavaca
Dr. Aviles Cruz Carlos	UAM-Azcapotzalco
Dr. Campos Rodríguez Raúl	Centro Universitario de la Ciénega-UDG
Dr. Carlos Hernández Salvador	Cinvestav-Salttillo
Dr. Contreras Ordaz Marco Antonio	Universidad Tecnológica de la Mixteca
MCC Cruz Gálvez Jorge Elpidio	Universidad Autónoma de Chiapas
Dr. Delgado Mata Carlos	Universidad Panamericana-Aguascalientes
Dr. Gutiérrez Díaz de León Carlos A.	Universidad Panamericana-Aguascalientes
Dr. Incera Diéguez José Alberto D.	Instituto Tecnológico Autónomo de México
Dr. Jiménez Alaniz Juan Ramón	UAM-Iztapalapa
Dr. Llanes Santiago Orestes	CUJAE-Habana, Cuba
Dr. López Ornelas Erick	UAM-Cuajimalpa
Dr. López Pacheco Dino Martin	Université de Nice, Francia
Dr. Magos Rivera Miguel	UAM-Azcapotzalco
Dr. Ramírez Treviño Alberto	Instituto Tecnológico Superior de Cajeme
Dr. Reyes Reyes Juan	Instituto Tecnológico de Zacatepec
Dr. Rodríguez Segura Elías José J.	Instituto Tecnológico de Celaya
Dra. Ruiz Beltrán Elvia	Instituto Tecnológico de Aguascalientes
M. C. Santiago Espinosa Felipe	Universidad Tecnológica de la Mixteca
Dr. Sumali Anton	Sandia National Laboratories, U.S.A.
M. Sc. Tafur Sotelo Julio Cesar	Pontificia Universidad Católica del Perú
Dr. Varona Salazar Jorge	SnowBush México
Dr. Vázquez Álvarez Iván	UAM-Azcapotzalco
Dr. Velázquez Guerrero Ramiro	Universidad Panamericana-Aguascalientes

Presentación

Con la realización de la VII Semana Nacional de Ingeniería Electrónica, el Departamento de Electrónica de la Universidad Autónoma Metropolitana Azcapotzalco y la Facultad de Contaduría Pública de la Universidad Autónoma de Chiapas cumplen con el objetivo de proporcionar un foro para la discusión de ideas, la comparación e intercambio de resultados y la divulgación de los trabajos de investigación relacionados con la Ingeniería Electrónica y sus diversas ramas.

En esta ocasión el evento se lleva a cabo en las instalaciones de la Universidad Autónoma de Chiapas, institución de educación pública de reconocido prestigio ubicada en la Ciudad de Tapachula, Chiapas.

Las memorias incluyen 73 artículos en extenso, los cuales fueron seleccionados por el Comité Técnico de entre más de 100 trabajos enviados por autores de todo el país. Cabe señalar que se realizó un esfuerzo importante en el arbitraje de los trabajos con el objetivo de asegurar la calidad de los mismos. Nuestro reconocimiento a los miembros del Comité Técnico por la labor realizada.

Finalmente, agradecemos a los miembros del Comité Organizador Local por su esfuerzo para lograr que este evento se realice en las mejores condiciones.

Tapachula, Chiapas, octubre de 2011.

Comité Organizador de la VII Semana Nacional de Ingeniería Electrónica.

ÍNDICE

Título de la Ponencia	Página
Simulación de un Comportamiento Caótico en un Circuito no Lineal con Control para Supresión de Caos	1
Uso de la Lógica Difusa en el Control de un Reactor Nuclear	8
Control de Velocidad por Pasividad de un Motor de Imanes Permanentes y su Simulación en Simnon	15
Supervisión y vigilancia de una planta de tratamiento de agua potable	25
Sistema Concentrador de Comunicaciones X-10	32
Implementación de una Red Neuronal Artificial en un FPGA	41
Sistema de Identificación y Control del Servo Motor Amira para Uso Didáctico Desarrollado en Instrumentación Programable	51
The RFID Tag Performance Applied to the Vehicular Identification	60
Diseño de Antena para Etiqueta en aplicaciones en RFID	67
Sistema de localización y seguimiento de unidades móviles en el Distrito Federal	73
Comparación de Mediciones de Atenuación en los Diversos Entornos de Terreno de la Ciudad de México y el Modelo Hata	81
Optimización de un Módulo Peltier de Dos Etapas en Régimen Eléctrico Oscilante	89
Diseño de Circuitos Integrados CMOS: Una Aproximación desde la Perspectiva de la Ley de Ohm	95
FPGA Based Parametric Platform for Evolvable Hardware	104
Characteristics Study of an Electronic Systems Design Strategy that Includes the Traditional Strategy and the Modified Traditional Strategy.	111
Formación del educador universitario desde la visión del educando Diagnóstico Aplicado en el Área de Electrónica y Computación	121
Cómo Capturar Alumnos Usando Videojuegos.	128
IMS-Learning Design: Retos de Adaptabilidad evaluada y metamodelos de personalización en las UoL	136

Título de la Ponencia	Página
Simulación en SIMNON del Control Directo de Par de un Motor de Imanes Permanentes	142
Electronic Ballast Based on a Series Active Input Current Shaper	154
Covertidor Push-Pull Alimentado en Corriente para Baja Tensión de Salida	161
A Three Port Converter for Renewable Energy Applications	169
A Six-Switch Three-level Current Source Inverter	178
A Single Stage Electronic Ballast Based on the 'Z' Source Inverter	184
Balastro Electrónico para Lámparas de HID MH con Control de Intensidad Luminosa	193
Análisis y Reducción de Ruido en una Máquina Cortadora de Plasma XY	201
Sistema de Adquisición de Datos Automatizado para Pruebas Mecánicas Uniaxial y Triaxial	209
Diseño y Calibración de un Fotómetro con Interfaz Virtual	219
Desarrollo de un paquete tecnológico para mantenimiento predictivo industrial	229
Propuesta de diseño de un sensor para determinar la concentración de LiBr mediante las técnicas ópticas de Transmitancia (T) e Interferencia Multimodal (MMI)	235
Aplicación de la Domótica en el monitoreo de sensores electrónicos a través de una interfaz de comunicación con Dispositivos móviles vía Web Services (SOAP-XML)	245
Cálculo de Bioimpedancia mediante Detección Síncrona de Fase basado en Hardware Reconfigurable	253
Método Geométrico para la Orientación Ortogonal en Luces Robóticas	261
Modelado y Análisis de los Circuitos Serie RC y LC Descritos por una Ecuación Diferencial de Orden Fraccionario	268
Automatización de una Prueba Triaxial Rápida "LG-TRIAXIAL"	275
Ecuaciones de Maxwell de Orden Fraccionario	284
Transmisor Indicador de Temperatura (TIT) con Interfaz para LABVIEW.	292
Sistema Domótico de Red Inalámbrica para Monitoreo y Control del Gasto Electrodoméstico utilizando FPGA y LabVIEW.	301
Tarjeta de adquisición de datos para el registro de señales mioeléctricas	310
Construcción de un Sistema RFID para Aplicaciones en Control de Procesos	318
Mejora del Desempeño de un Equipo de Corte a través del Análisis de Vibración	327
Análisis del Ensamble de una Máquina Cortadora de Plasma Utilizando la Metodología de Diseño para Ensamble	335

Título de la Ponencia	Página
Robot Humanoide ZERO: Diseño, Análisis y Prototipo	342
Manipulación de una Brazo Robótico por medio de una tarjeta FPGA	351
Aplicación de Robots Industriales Coordinados	358
Simulación de un Sistema Robótico 3 DOF Orientado a la Rehabilitación Neuromuscular del Brazo	364
Sistema de ojo robótico y aplicación al seguimiento de objetos	371
Diseño de un Control Óptimo para un Robot Móvil	380
Estudio de Viabilidad de un Nuevo Sistema de Puertas para el Transporte Colectivo Metro	389
Clasificación de Piezas Metálicas mediante Perfilometría Láser	396
Reconocimiento del Lenguaje de Señas Mexicano para su Interpretación y uso en Sistemas Robóticos	404
Herramienta de clasificación de información basada en métodos de clusering: Aplicación a la segmentación de imágenes	413
Reconocimiento de Voz para Comunicación y Diagnóstico de Personas con Disartria en México	422
Entrenamiento de Perceptrón Unicapa para la Detección Aproximada de Bordes en una Imagen Digital.	432
Reconocimiento Automático del Habla para Pronunciación de la Lengua Mixteca	440
Estimación Espectral de Potencia de Señales EEG con Métodos No Paramétricos	449
Generación de Indicadores de Productividad Mediante la Implementación de un Datawarehouse en SQL Server Enterprise Aplicado a Hospitales de Alta Especialidad	458
Pantalla Interactiva para Acceso a Material Didáctico Multimedia utilizando Mapas Conceptuales	466
La Háptica como Complemento a la Realidad Virtual para el Aprendizaje de un Idioma Extranjero	473
OkAR: Generador de cargas académicas para la Ingeniería en Tecnologías de la Información de la Universidad Tecnológica de Izúcar de Matamoros	481
Desarrollo de un Sistema para la Contabilidad de Costos basado en Tecnología Web (ExtJs-PHP), con la implementación de un DataWarehouse (MySQL)	488
Consideraciones en la Migración de Datos Locales hacia Entornos Distribuidos	496
Sistema de Detección de Intrusiones basado en el Método de los Espacios de Entropía	503

Título de la Ponencia	Página
Representación formal del conocimiento para apoyar la adopción de sistemas de información clínica	513
Sistemas de Rastreabilidad en el Piso de Fabricación Basados en SOA (ServiceOrientedArchitecture)	520
Buscador Web móvil semántico	529
Software para el Cálculo de los Parámetros de Resistencia del Suelo	537
Simulación de Vehículos Reactivos con Ruedas Diferenciales en Webots	544
Interfaz Gráfica y Sistema de Mando Inalámbrico para la Enseñanza de Niños con Problemas Neuromotores.	550
Propuesta de Sistema de Conteo Electrónico de Personas que Utilizan el Servicio de Sanitarios de la Central Internacional de Autobuses de Culiacán	561
Control de una Red de Iluminación a través de un sistema CAN embebido	567
Paint con FPGA: Sistema para la Evaluación de las Interfaces VGA y PS/2 de la Tarjeta de Desarrollo Nexys2	575
Descripción de un Sistema de Adquisición Controlado por FPGA para Procesamiento de Pulsos Provenientes de PMTs	585

Simulación de un Comportamiento Caótico y Control para Supresión de Caos en un Circuito No Lineal

Gilberto Sánchez Mares, Edgar Iván Muñoz Guevara, *Jorge Alberto Padilla García

*Universidad Politécnica de Aguascalientes, Prolongación Mahatma Gandhi Km 2, Aguascalientes, Ags., México, Tel:(449) 4421400, gilberto_gsm@hotmail.com.

Resumen

El caos es un comportamiento estacionario de un sistema dinámico y tiene como característica principal la sensibilidad a condiciones iniciales. Éste tipo de sistemas es tema de estudio para varias áreas del conocimiento tales como las Matemáticas, la Física, Telecomunicaciones, Medicina, entre otras más. Como resultados de investigación se han encontrado aplicaciones relevantes en el área de la Electrónica. Dentro de las más importantes destaca el diseño de un circuito electrónico generador de caos, cuyo descubridor Julien Clinton Sprott demuestra que el sistema presenta un comportamiento no lineal y cuya construcción está conformada por resistores, capacitores y amplificadores operacionales inversores.

El presente trabajo muestra una simulación del circuito de Sprott cuya finalidad es probar su eficiencia como atractor caótico. Además, se realizó su modelado matemático obteniendo una serie de funciones elementales (sistema no-lineal) teniendo como base los parámetros adimensionales $B=1.2$ y $C=2.0$ para $G(x)$. Por último, se realizó la simulación del oscilador en el software en LabVIEW™ de National Instruments a la cual se integró un control no lineal para la supresión de caos.

Los resultados obtenidos muestran un control eficiente para la supresión del caos el cual fue comprobado de manera satisfactoria en el software de LabVIEW™.

Palabras Clave: Atractor caótico, Circuito caótico de Sprott, Supresión de caos, Sistema no lineal,

I. Introducción

El caos se puede definir como un comportamiento estacionario de un sistema dinámico, que no es un punto de equilibrio ni es periódico, además de que una de sus principales características es la de que muestran una sensible dependencia a las condiciones iniciales [1]. Después de tres décadas de estudio, las condiciones suficientes del caos en un sistema de ecuaciones diferenciales ordinarias (EDOs) siguen sin conocerse.

Un ejemplo de sistema no lineal que representa un comportamiento de tipo caótico es el de J. C. Sprott [2] el cual se puede implementar en un circuito electrónico como el que se muestra en la Fig. 1.

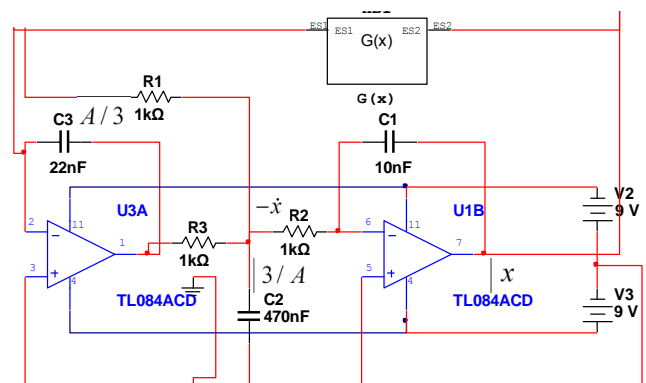


Fig. 1. Circuito caótico de Sprott que resuelve la Ecuación (1) usando uno de los elementos no lineales para realimentarlo.

Este sistema está modelado por

$$\ddot{x} + A\dot{x} + \dot{x} = G_n(x) \quad (1)$$

donde los puntos encima de la variable x significan el tiempo de derivación (primero, segundo, tercero), A es una constante adimensional mayor a 0, y $G_n(x)$ puede tomar una de las siguientes formas no lineales [2,3]

$$G_1(x) = B|x| - C$$

$$G_2(x) = -B \max(x, 0) + C$$

$$G_3(x) = Bx - C \operatorname{sig}(x)$$

$$G_4(x) = -Bx + C \operatorname{sig}(x)$$

donde B y C son constantes adimensionales ideales seleccionadas para producir caos.

Por otra parte, el control de caos consiste en el diseño de estrategias que permiten asignar dinámicas deseadas a sistemas caóticos.

En el control del caos uno de los problemas básicos es el de la supresión de caos [4].

La supresión de caos consiste en estabilizar las trayectorias de un sistema caótico alrededor de algún punto de equilibrio y su importancia se basa en que las oscilaciones de éste tipo de sistemas son impredecibles y pueden producir daños.

II. Desarrollo

Se simuló el circuito de Sprott utilizando la forma de $G_4(x)$ la cual se muestra en la Fig. 2; ésta función fue seleccionada debido al atractor caótico que produce, ya que tiene una la forma peculiar mostrada en la Fig. 3

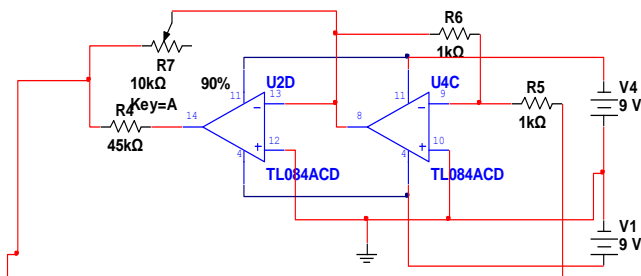


Fig. 2. Circuito electrónico que genera el caos para $G(x)$.

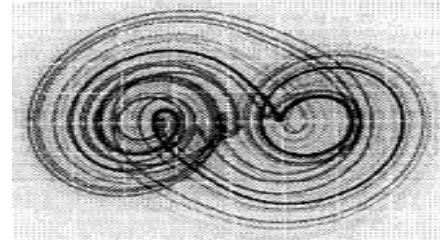


Fig. 3. Atractor caótico producido por la función $G_4(x)$.

En la Ecuación (1) se muestra la función que se introdujo en el programa implementado en el lenguaje de programación de LabVIEW en el cual se simuló el comportamiento del oscilador para así poder diseñar un control no lineal con la finalidad de suprimir el caos generado por este sistema.

Para iniciar con el control se debe tomar en cuenta el sistema afín

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x) \end{aligned} \quad (2)$$

donde $f(x)$ y $g(x)$ son los campos a ser evaluados y $h(x)$ es el vector de salidas que se pueden medir del sistema a controlar.

De las consideraciones mostradas en la igualdad (2), se debe tomar en cuenta que para el diseño de nuestra estrategia de control debemos seleccionar la variable que habrá de inducir la dinámica deseada y con ello definir el campo vectorial $g(x)$, con la finalidad de ver si dicho campo podrá influir en el resto de la planta dado por el campo vectorial $f(x)$ y si la salida elegida de $h(x)$ como realimentación al sistema permite lograr el objetivo de control deseado.

Para ello es necesario obtener un valor escalar de la proyección de la salida de $h(x)$ sobre $f(x)$ para así poder obtener la influencia de dicha salida sobre la parte señalada de la planta, para después obtener la proyección del campo vectorial $g(x)$ y así verificar si se puede llevar el control de la planta [5].

Dichas proyecciones se efectúan por medio de la aplicación de gradientes $(\partial h(x)/\partial x)$ y derivadas direccionales de las funciones $h(x)$ a lo largo del

campo vectorial $f(x)$ bajo las reglas de algebra de Lie [6]

Una vez definido lo anterior se pasa a hacer la prueba de controlabilidad de grado relativo (r), la cual puede ser efectuada de manera gráfica y analítica.

a) Prueba gráfica

Para la prueba gráfica se debe considerar, que nuestro sistema puede verse como un dibujo dado por un conjunto de nodos y un conjunto de conexiones tal y como se muestra en la Fig. 4

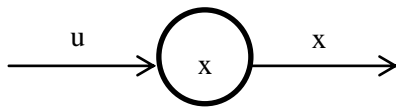


Fig. 4. Diagrama representativo de la prueba gráfica

Este tipo de representación sigue la equivalencia (3) para así indicar la longitud de la ruta (n), la cual es el número mínimo de nodos que la ruta tiene.

$$r = n - 1 \tag{3}$$

b) Prueba analítica

Para ésta prueba se tiene que r es el número entero más pequeño tal que el efecto de salida se ve reflejado en los campos vectoriales $f(x)$ y $g(x)$ [6] y se expresa tal y como se ve en (4).

$$\begin{aligned} L_g L_f^k h(x) &= 0 \\ L_g L_f^{n-1} h(x) &= 0 \end{aligned} \tag{4}$$

donde n es el número de estados en el vector L_f .

Bajo ésta consideración, el sistema afín cambia en (5):

$$\begin{aligned} \dot{z}_1 &= L_f h(x), \\ &\cdot \\ &\cdot \\ &\cdot \\ z^{(n-1)} &= L_g L_f^{n-1} h(x) \\ z_n &= L_f^n h(x) + L_g L_f^{n-1} h(x)u \end{aligned} \tag{5}$$

Despejando u de (5), obtenemos que nuestra ley de control se define como se muestra en la Ecuación (6)

$$u = \frac{-L_f^n h(x) + \sum_{i=0}^{n-1} k_i [L_f^i h(x)]}{L_g L_f^{n-1} h(x)} \tag{6}$$

III. Resultados

1. Simulación del Circuito

Tomando como base el valor práctico de $A=0.6$, se producen salidas similares de los amplificadores operacionales a bifurcaciones y caos; además $B=1.2$ y $C=2$ son los valores óptimos para poder realizar el atractor caótico.

De forma física podemos decir que A es el valor utilizado para obtener la capacitancia de C_3 utilizando la Igualación (7) y de C_2 usando (8)

$$C_3 = A / 3 \tag{7}$$

$$C_2 = 3 / A \tag{8}$$

Obteniendo así que $C_3=0.02F$ y $C_2=0.5F$ como se muestra en la Fig. 5, sin embargo para fines prácticos C_2 se sustituye por un capacitor de $0.47F$ dando así los mismos resultados del atractor.

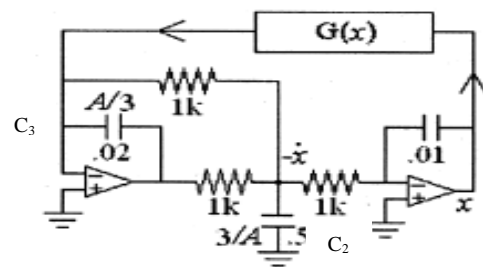


Fig. 5. Valores de capacitores C_3 y C_2 para el oscilador caótico de Sprott.

Al igual que A ; B y C pueden ser vistos como los valores óptimos para obtener la resistencia necesaria para producir el atractor deseado; además se debe considerar la entrada de voltaje (V_s) como 9 volts, teniendo en cuenta estas consideraciones los valores de los resistores R_7 y R_4 se calculan como se muestra en (9) y (10)

$$R_7 = 1/B \quad (9)$$

$$R_4 = V_s / C \quad (10)$$

Teniendo como resultado que $R_7=8333\Omega$ y $R_4=45k\Omega$ tal y como se representa en la Fig. 6, para el caso de R_7 se utilizó un potenciómetro de $10k\Omega$ el cual se posicionó en el valor.

En la Fig. 7 se muestra la simulación del circuito de Sprott así como los puntos a medir en el osciloscopio.

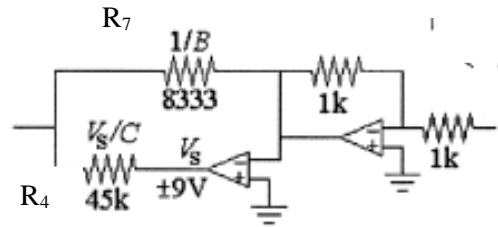


Fig. 6. Valores de resistores R_7 y R_2 para el atractor caótico de Sprott.

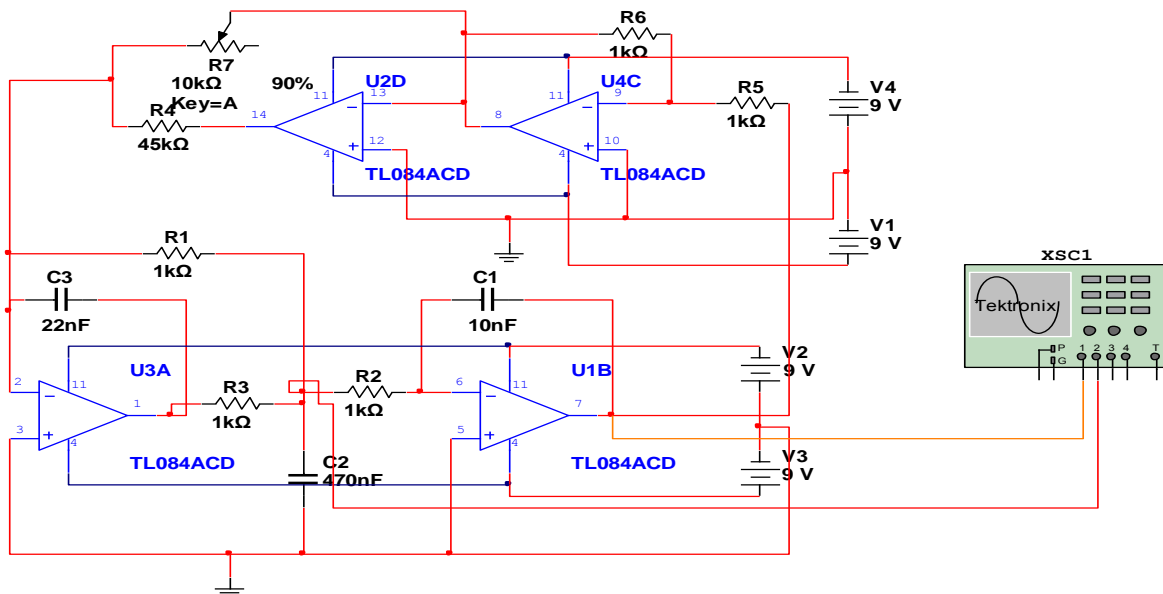


Fig. 7. Circuito caótico de Sprott simulado con la $G(x)$ elegida para ver el funcionamiento del atractor.

En la Fig. 8 se muestra la respuesta del circuito en el osciloscopio, demostrando así el funcionamiento y su similitud con la Fig. 3; en a) tenemos la forma que se despliega en XY y en b) la que se obtiene en el formato YT.

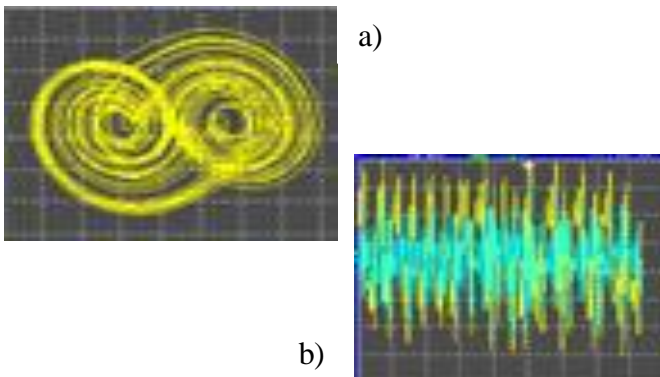


Fig. 8. Respuesta del atractor caótico.

2. Control no lineal

Definiremos el estado que queremos controlar como x_1 , teniendo en cuenta esto, decimos que $f(x)$ y $g(x)$ son equivalentes a (12)

$$f(x) = \begin{pmatrix} x_2 \\ x_3 \\ C \text{sig}(x_1) - x_2 - Ax_3 \end{pmatrix}$$

$$g(x) = \begin{pmatrix} 0 \\ 0 \\ -x_1 \end{pmatrix} \quad (12)$$

Iniciaremos usando la prueba de control gráfica, nuestro grafo se puede ver en la Fig. 9, donde nuestra longitud de ruta está indicada de color rojo.

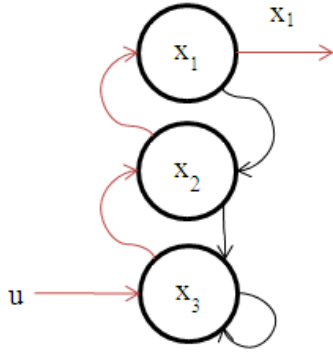


Fig. 9. Prueba de controlabilidad gráfica para nuestra función afín [1].

Utilizando (3) y tomando en consideración que $n=4$ (número de flechas indicadas en rojo), ya que es la ruta más corta para llegar a controlar x_1 decimos que $r=3$; por lo tanto tenemos tres estados para analizar.

Una vez definidos los estados pasamos a la prueba analítica y procedemos a calcular cada uno.

Cuando $n=0$

$$L_f^0 h(x) = x_1$$

$$L_g L_f^0 h(x) = (1, 0, 0) \begin{pmatrix} 0 \\ 0 \\ -x_1 \end{pmatrix} = 0$$

Para $n=1$

$$L_f^1 h(x) = (1, 0, 0) \begin{pmatrix} x_2 \\ x_3 \\ Csig(x_1) - x_2 - Ax_3 \end{pmatrix} = x_2$$

$$L_g L_f^1 h(x) = (1, 0, 0) \begin{pmatrix} 0 \\ 0 \\ -x_1 \end{pmatrix} = 0$$

Si $n=2$

$$L_f^2 h(x) = (0, 1, 0) \begin{pmatrix} x_2 \\ x_3 \\ Csig(x_1) - x_2 - Ax_3 \end{pmatrix} = x_3$$

$$L_g L_f^2 h(x) = (0, 0, 1) \begin{pmatrix} 0 \\ 0 \\ -x_1 \end{pmatrix} = -x_1$$

Cuando $n=3$

$$L_f^3 h(x) = (0, 0, 1) \begin{pmatrix} x_2 \\ x_3 \\ Csig(x_1) - x_2 - Ax_3 \end{pmatrix}$$

$$L_f^3 h(x) = Csig(x_1) - x_2 - Ax_3$$

Una vez que se obtuvieron los estados proseguimos a obtener nuestra ley de control como se define en (6), resultando como nuestro control no lineal

$$u = \frac{-(Csig(x_1) - Ax_3 - x_2) + k_0(x_1 - x_{ref}) + k_1(x_2) + k_2(x_3)}{-x_1}$$

donde k_0 , k_1 y k_2 son constantes a determinar.

En la Fig. 10 se muestra el programa con un nodo de fórmula en el cual se ingresa nuestra función de control u y en la Fig. 11 se observa la interfaz gráfica de usuario donde se obtiene la respuesta del sistema y de la acción de control, la cual se puede ver que no es muy brusca y por lo tanto se puede implementar de manera física sin ningún problema.

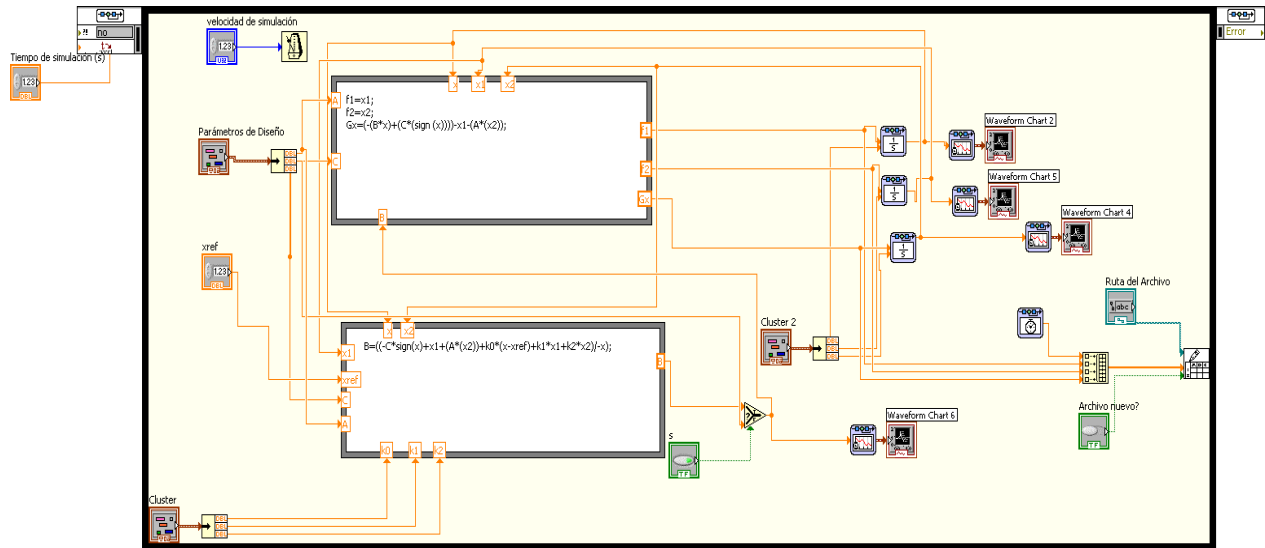


Fig. 10. Programa en LabVIEW para simular la función (1) con su control no lineal.

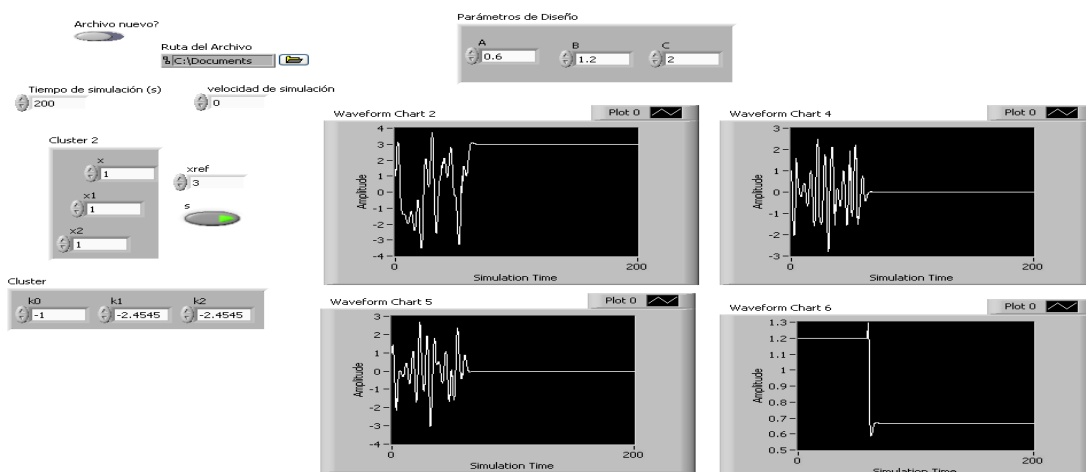


Fig. 11. Interfaz gráfica para simular el sistema caótico y el control no lineal.

IV. Conclusiones

En este trabajo se presentó, cómo mediante una función no lineal se puede diseñar un circuito caótico de aplicación experimental así como una simulación óptima que tomaba el comportamiento deseado. Además de diseñarse una estrategia de control de caos (mediante control no lineal).

La importancia del estudio de la supresión de caos, radica en que es de gran utilidad en tratamientos de desórdenes cardiacos y epilepsias, en sistemas laser, entre otros.

V. Referencias

- [1] J. G. Fernández, “Caracterización y filtrado de señales inmersas en ruido caótico y estocástico”, DR.I. Tesis, Universidad Nacional del Mar de Plata, Mar de Plata, Argentina 2005.
- [2] J. C. Sprot, “A new Class of Chaotic Circuit”, ELSEVIER, 2000, pp.19-23.
- [3] X. Xiao-dong, Y, Wan-li, Z, Su-wen, “Finite Time synchronization of Sprott Circuits whit Uncertain Parameters”,

- International Conference on Advanced Computer Control, 2009, pp. 693-696.
- [4] J. A. Rodríguez, J De León, “Control de Sistemas Caóticos”, Ciencia UANL, vol. XI, número 001, 2008, pp. 29-35.
- [5] J. F. Briones de la Torre, M. A. Paz Ramos, J. I. O. Rodríguez Martínez, M. O. Aguilar, Justo, “Tutorial para el Diseño y Aplicación de Estrategias de Control No Lineal”, Cuarto Congreso Nacional de Ingeniería Mecatrónica, Aguascalientes, México, 2010, pp. 134-147.
- [6] A. Isidori, “NonLinear Control Systems: An Introduction”, Springer, Germany, 1985.

VI. Autores

Gilberto Sánchez Mares es actualmente estudiante de la carrera de Ingeniería Electrónica por la Universidad Politécnica de Aguascalientes (UPA).

Edgar Iván Muñoz Guevara es actualmente estudiante de la carrera de Ingeniería Electrónica por la Universidad Politécnica de Aguascalientes (UPA).

Jorge Alberto Padilla García es actualmente estudiante de la carrera de Ingeniería Electrónica por la Universidad Politécnica de Aguascalientes (UPA).

Uso de la Lógica Difusa en el Control de un Reactor Nuclear

***Marco Antonio Márquez Vera, **Filiberto Muñoz Palacios, ***Blanca Diana Balderrama Hernández**

*Universidad Politécnica de Pachuca, carretera Pachuca-Cd. Sahagún, Km 20, Rancho Luna, Ex Hacienda de Sta. Bárbara, Zempoala, Hgo. 771-215-48-73, marquez@upp.edu.mx

**Universidad Politécnica de Pachuca, carretera Pachuca-Cd. Sahagún, Km 20, Rancho Luna, Ex Hacienda de Sta. Bárbara, Zempoala, Hgo. 771-161-85-29, mupafi@upp.edu.mx

***Secretaría de Educación Pachuca, carretera México-Pachuca Km 84.5, Colonia Venta Prieta, Pachuca, Hgo. 717-35-00, dianabh@hotmail.com

Resumen

En el presente trabajo se muestran los resultados en simulación sobre el control de un reactor nuclear. Dado que en el modelo de la planta el crecimiento en la cantidad de neutrones retardados es muy rápido, la acción de control propuesta es adaptable, esto con el fin de estabilizar el sistema ya que se asume que no es posible realizar la medición de neutrones en línea. El control propuesto es difuso con adaptación en los consecuentes de sus reglas que son del tipo Takagi-Sugeno. Se muestra también el uso de la lógica difusa tipo II, que en este caso promedia a dos sistemas tipo I similares al anterior, finalmente se hace una comparación entre ambas técnicas de control.

Palabras Clave: Control, lógica difusa, reactor nuclear.

I. Introducción

El control de las plantas nucleares es crítico dadas las terribles consecuencias que puede tener una mala acción de control, por lo cual se toman muchas medidas de seguridad ante posibles fallas [1]. Un ejemplo claro de estas consecuencias son las que sufre Japón por el terremoto del pasado 11 de marzo de este año.

Un reactor nuclear es una instalación en donde mediante fisión nuclear [2] se producen neutrones de manera autosostenida liberando grandes cantidades de calor. El calor generado por un reactor nuclear puede ser usado para calentar agua, y el vapor producido, al mover unas turbinas, genera electricidad; a dicho proceso se le conoce como planta nucleoelectrónica [4].

Para la fisión nuclear se usa un combustible como el óxido de uranio UO_2 , y si un neutrón golpea el núcleo de un átomo de uranio a 2 Km/s lo parte, los neutrones liberados de la fisión pueden viajar a 20,000 Km/s y se seguir el proceso se tiene una reacción en cadena que

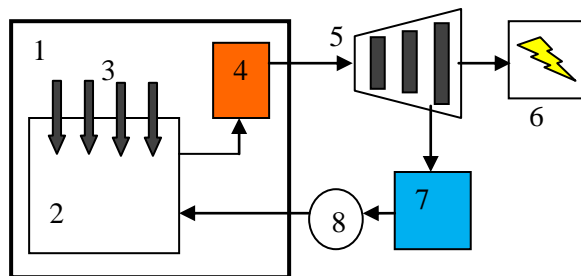
desemboca en una explosión nuclear; para evitar esto es necesario un moderador que frene a los neutrones liberados de la fractura del núcleo. El moderador necesita ser de peso atómico pequeño y no debe absorber neutrones [2,3], los moderadores más usados son el grafito, el agua ligera y pesada, y algunos líquidos orgánicos.

Es necesario también un refrigerante que debe circular por el núcleo, suele ser agua y es similar a una piscina de 10m de profundidad. Como medio de control y protección se tienen barras de control hechas de cadmio, boro o hafnio, que absorben neutrones [2], los fluidos absorbentes también ayudan a apagar al reactor, y el edificio contenedor puede ser de plomo o cemento de alta densidad para retener la radiación [4], en la Fig. 1 se muestra un esquema general de la planta nucleoelectrónica.

El modelo matemático del reactor está basado en los existentes en el Instituto Nacional de Investigaciones Nucleares (ININ) y es del tipo TRIGA (por sus siglas Training Research Isotope General Atomics) [17]. Este reactor tiene un

núcleo de Uranio 235 con hidruro de zirconio, este último funciona como moderador para evitar que la reacción crezca si las barras de control no se pueden introducir rápidamente.

Dado que el modelo es no lineal se propone usar lógica difusa para el control del reactor, existen varios simuladores [5] incluso en internet se pueden probar algunos [4,5], si el conocimiento de un experto se puede representar en reglas, es posible hacer una máquina de inferencia con estos resultados [6].



1. Edificio de concreto
2. Núcleo
3. Barras de control
4. Generador de vapor
5. Turbina
6. Alternador
7. Condensador
8. Bomba de recirculación

Fig. 1. Esquema de una planta nucleoelectrica

La lógica difusa fue presentada por Lofti Zadeh en 1965 para describir clases u objetos que no tienen un criterio bien definido de pertenencia. El uso de la lógica difusa en sistemas de control tuvo su auge a partir de que Mamdani realizara en control de una máquina de vapor con estas técnicas [6], desde entonces han surgido muchísimas aplicaciones como en modelado, identificación de parámetros, control, procesamiento de imágenes, entre otras [6,10,13].

Una de las técnicas de control para sistemas no lineales se presenta en [7], en donde se pueden variar los consecuentes de las reglas en base al desempeño del control respecto a un modelo de referencia. Debido a que el trabajar con información incompleta o imprecisa tiene desventajas, el uso de la lógica difusa simplifica el trabajo, e incluso el conocimiento de un

experto se puede imitar, por lo que existen críticas al uso de estas técnicas [9].

En 1975 Lofti Zadeh definió la lógica difusa tipo II, la cual es una generalización de la ahora llamada lógica difusa tipo I. En la tipo II se tiene que la parte difusa (incertidumbre) es a su vez difusa [14], es decir, al no poder estar seguros de que el grado de pertenencia de un objeto a un conjunto está dado por una función en particular, lo cual es un tema muy debatido y por ello también atacada la lógica difusa [9], se propone que haya cierta incertidumbre en la función de pertenencia, teniendo con ello dos grados de pertenencia, ver Fig. 2.

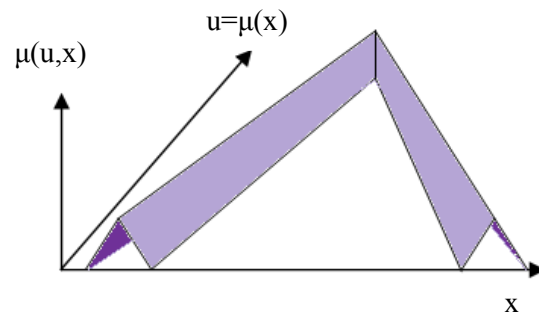


Fig. 2. Función de pertenencia con lógica difusa tipo II

Una vez evaluado el sistema difuso tipo II, hay que simplificarlo a un sistema tipo I mediante un reductor [8]. En la literatura hay trabajos en donde la parte difusa de la pertenencia, llamada huella de incertidumbre, es modelada por dos conjuntos difusos tipo I, y el promedio de ambos es la salida del sistema difuso [8], en este trabajo se usan estas ideas para controlar el reactor.

II. Desarrollo

El modelo del reactor empleado es suponiendo una cinética puntual [3], lo cual es válido si el reactor está cerca de ser crítico, lo cual significa que el factor de generación es uno, y que no se presenta una perturbación grande.

Para entender mejor el modelo del reactor, se definen algunos conceptos. El factor de generación K representa los neutrones generados y si $K=1$ se dice que el reactor es crítico y se establece la reacción nuclear. La reactividad es la

desviación relativa del factor de generación, si es negativa se dice que el reactor está subcrítico. Esta reactividad está compuesta por la suma de la reactividad interna y la externa del reactor, siendo esta última dependiente de las barras de control y será la señal de control en el sistema.

En (1) presenta el modelo puntual del reactor TRIGA usado

$$\begin{aligned} \dot{x}_1 &= \frac{u + x_3 - \beta}{\Lambda} x_1 + \lambda x_2 \\ \dot{x}_2 &= \frac{\beta}{\Lambda} x_1 + \lambda x_2 \\ \dot{x}_3 &= -\alpha k x_1 - \gamma x_3 + \alpha k \eta_0 \end{aligned} \quad (1)$$

donde x_1 es la potencia neutrónica, x_2 es la concentración de precursores de neutrones retardados, estos son los neutrones liberados por los fragmentos de la fisión que suelen ser inestables, x_3 es la reactividad interna del reactor, la cual es la desviación relativa del factor de generación, siendo este último es el número de neutrones en una generación, dividido entre el número generado en la fisión inmediata anterior [2]. u es la reactividad externa del reactor, Λ es el tiempo de generación de neutrones instantáneos, β es la fracción de neutrones retardados, λ es una constante de decaimiento, α es el negativo del coeficiente de reactividad por temperatura, η_0 es una potencia asumida como estacionaria, k es el recíproco de la cantidad calorífica del reactor en $^{\circ}C/Ws$ y γ es el inverso del tiempo promedio de transferencia de calor del combustible al refrigerante en $1/s$ [3], en la tabla 1 se muestran los valores de los parámetros empleados para realizar las simulaciones.

Parámetro	Valor
β	6.433×10^{-3}
α	0.01358
k	0.1919×10^{-4}
γ	0.2
λ	38×10^{-6}
Λ	0.4024

Tabla 1. Valores de simulación

El control basado en lógica difusa tipo I es un sistema tipo Sugeno, se trata de un PI difuso basado en un PD, en el que posteriormente se integra la señal de control para hacer el PI en (2) se muestra el cálculo de la variación de la señal de control, donde $\mu_i(x_i(k))$ es el grado de pertenencia de la variable x_i a la función de pertenencia i . El controlador, así como el sistema de adaptación tienen 11 funciones de pertenencia trapezoidales para el error y la derivada del error, dando 121 reglas en la máquina de inferencia. El consecuente de las reglas es constante siendo un sistema tipo Singleton [7].

$$\Delta u_i(k) = \frac{\sum_{i=1}^n \mu(x_i(k)) u_i(k)}{\sum_{i=1}^n \mu(x_i(k))}. \quad (2)$$

En la Tabla 2 se muestra un ejemplo de la memoria asociativa difusa [13] que puede ser usada para el controlador, ésta muestra sólo cinco funciones de pertenencia por cada variable. Un sistema difuso similar es el encargado de modificar los consecuentes del controlador difuso, dependiendo del grado de activación de las reglas [7]. En este trabajo se mantiene fija en cero a la regla central, como en [15], para evitar oscilaciones si es que la salida ya sigue muy de cerca a la referencia. Según la regla que se haya activado será la regla a modificar, por ello es que el adaptador de las reglas tiene la misma memoria asociativa difusa, y es posible modificar las ganancias de entrada para variar la velocidad de adaptación [16].

$\frac{de(t)}{dt}$ \ / $e(t)$	Muy Neg.	Neg.	Cero	Pos.	Muy Pos.
Muy Neg.	-1	-0.75	-0.5	-0.25	0
Neg.	-0.75	-0.5	-0.25	0	0.25
Cero	-0.5	-0.25	0	.025	0.5
Pos.	-0.25	0	0.25	0.5	0.75
Muy Pos.	0	0.25	0.5	0.75	1

Tabla 2. Ejemplo de memoria asociativa difusa.

En la literatura existen diferentes métodos de adaptación, algunos ajustan las ganancias de entrada y salida en el sistema difuso [13], y otros

ajustan los antecedentes y/o los consecuentes de las reglas [7,11].

El este trabajo al ajustar los consecuentes es posible ver qué reglas se activan y con qué pertenencia lo hacen. De este modo es posible aprender más sobre el sistema en particular que se está controlando y así poder hacer trabajos con control iterativo si fuera el caso. Un esquema del control empleado se presenta en la Fig. 3.

Otra técnica con muchas expectativas por sus resultados es la lógica difusa tipo II. En algunos trabajos se emplea el algoritmo de evolución humana para ajustar la huella de incertidumbre, y con esto se reduce el error de seguimiento [8].

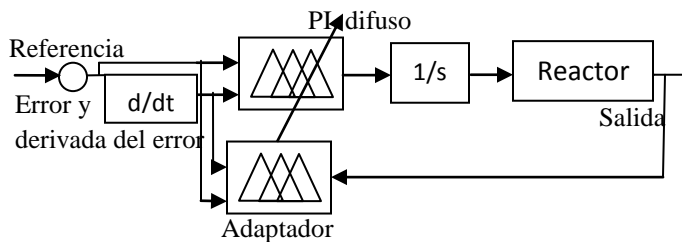


Fig. 3. Esquema del PI difuso

Se propuso el empleo de funciones de pertenencia similares a las del PI difuso, pero en este caso se agregaron otras reglas definidas de la misma forma pero con un modificador lingüístico para intensificarlas [13]. Si se mostraran todas las funciones de pertenencia, por ejemplo para el error, se verían como en la Fig. 4. Para cada una de las funciones se tiene un consecuente, el cual es promediado para cada función que tiene el mismo centro.

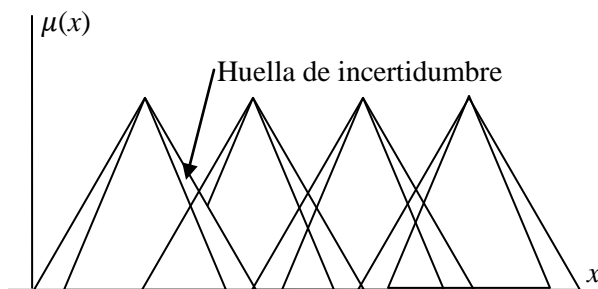


Fig. 4 Funciones de pertenencia con lógica difusa tipo II

Con una ganancia para el adaptador diferente para cada sistema, se obtiene el PI difuso tipo II, al promediarse ambos sistemas se tiene un filtro que responde mejor al ruido y a las perturbaciones externas. En el modelo del reactor es posible apreciar cómo la reactividad interna aumenta en forma considerable, lo cual hace crecer la potencia neutrónica. Con la lógica difusa tipo II se tienen más variables para poder disminuir la potencia; dado que se asume que no se mide la reactividad, la compensación puede ser en base al tiempo, así, para cuando ha transcurrido cierto lapso, se puede restringir la señal de control para estabilizar el sistema y evitar un crecimiento muy rápido en el número de neutrones generados.

III. Resultados

Al usar lógica difusa se puede filtrar el ruido de las mediciones, ya que es una especie de interpolación [8] entre las funciones de pertenencia que se activan. Se obtuvo que para el controlador difuso adaptable, se activan las reglas de las orillas, esto se puede mejorar, pero al mover la ganancia para que se active la mayoría de las reglas, el ajuste de sus consecuentes es más lento y se presentan picos en la potencia neutrónica, aunque grandes, de poca duración según la simulación. Existen trabajos en los que se demuestra que sin aplicar todas las reglas posibles, según la partición difusa de las variables [12], es posible controlar al sistema. En este caso esto provocaba mayor tiempo para que la salida esté dentro de cierta cota alrededor de la referencia. En las Fig. 5 y 6 se presentan las simulaciones del control con lógica difusa tipo I y II, es posible apreciar cómo la respuesta mejora con el avance del tiempo, mientras que al inicio hay muchas oscilaciones.

IV. Discusión

Haciendo un análisis sobre el error con ambas técnicas se observa que el error al final de la simulación es pequeño, y por tanto, al usar el criterio de ITAE (integral del absoluto del error por el tiempo), ambas técnicas tienen un desempeño similar, mientras que al usar la ISE

(integral del error al cuadrado) [16] se ve la mejora del uso de la lógica difusa tipo II, ya que se comporta de manera similar a un filtro, dado que al inicio de la simulación se presentan oscilaciones alrededor de la referencia, la tipo II disminuye este efecto. Se muestra la forma de calcular la ISE y la ITAE en (3) y (4). En la Tabla 3 se presentan los resultados de las integrales del error para comparar los resultados, donde se aprecia que el error usando este algoritmo de control con lógica difusa tipo II es 18% mayor con la ITAE, pero en cuanto a las oscilaciones al inicio, se tiene que usando el ISE, el error con el tipo I es más de 65% mayor que con el otro método.

$$ITAE = \int_0^{\infty} t|e(t)|dt \quad (3)$$

$$ISE = \int_0^{\infty} e^2(t)dt \quad (4)$$

Error \ Lógica difusa	Tipo I	Tipo II
ITAE	1.0171x10 ¹⁰	1.205x10 ¹⁰
ISE	1.255x10 ¹⁵	7.511x10 ¹⁴

Tabla 3. Análisis de errores en la simulación.

V. Conclusiones

Se obtuvo que la lógica difusa es un buen método para controlar un reactor nuclear, dados los inconvenientes de este tipo de plantas, se hacen muchas pruebas antes de aplicar las leyes de control. Como no se tenía una idea clara de que consecuentes poner en las reglas del controlador, se propuso el método de adaptación con lógica difusa tipo I, el error al final de la simulación es pequeño pero mientras se ajustan las reglas se tienen picos de potencia elevados. Una vez que se tuvo experiencia sobre el control del reactor, se propuso usar lógica difusa tipo II y con ello se redujeron las oscilaciones del inicio. Al analizar el error respecto a la referencia se notó como con

la ISE, el error es menos de la mitad que con el primer método.

Lo que resta averiguar es si los actuadores (barras de control) pueden reaccionar a la velocidad y con la magnitud requeridas para el control del reactor.

VI. Referencias

- [1] J. R. Lamarsh. Introduction to Nuclear Theory. Addison Wesley. 1975. Pp. 201-210.
- [2] P. Reuss. Exercices de Neutronique. EDP Sciences. 2004. Pp. 152-160.
- [3] J. M. Martínez. Reactores Nucleares. Ediciones ESTII Madrid. 2002. Pp. 18-35.
- [4] C. Ahnert. Tecnología Nuclear, Centrales Nucleares. ETSIN. 2007. Pp. 15-31.
- [5] Simulador Conceptual de una Central Nuclear. www.sen.upc.es/Situacion/Sim_Con_CenNu.htm. Abril, 2011.
- [6] E. Mamdani, S. Assilian, "An experiment in linguistic systesis with a fuzzy logic controller". Int. Journal of man-Machine Studies. Vol. 7. 1975. Pp. 1-13.
- [7] J. R. Layne, K. M. Passino, "Fuzzy Model Reference Learning Control". Journal of Intelligent and Fuzzy Systems. Vol. 4. 1996. Pp 33-47.
- [8] O. Castillo, P. Melin, Type 2 Fuzzy Logic, theory and applications. Studies in fuzziness and soft computing. Springer Verlag. Vol. 223. 2008. Pp. 243.
- [9] A. Aceves. "Usos y abusos de la lógica difusa para el control de procesos: Una alternativa para modelar lo incompleto de la información y lo impreciso de una información". Mantenimiento Productivo. Vol. 2(5). 2001. Pp. 12-17.
- [10] X. Luo, Z. Sun y F. Sun, "A new Approach to Fuzzy modeling and Control for nonlinear Dynamics Systems: Neuro-

- Fuzzy dynamics Characteristic Modeling and Adaptive Control Mechanism”. Int. Journal of Control, Automation and Systems. Vol. 7(1). 2009. Pp. 123-132.
- [11] R. Babuska. Fuzzy Modeling for Control. Kluwer Academic Publishers. 1998. Pp. 128-143.
- [12] P. J. MacVicar, “Fuzzy Sets for Man-Machine Interaction”. Int. Journal of Man-Machine Studies. Vol. 8. 1997. Pp. 687-697.
- [13] T. J. Ross. Fuzzy Logic with Engineering Applications. John Wiley & Sons. 2008. Pp. 254-263.
- [14] Q. Liang, J. Mendel. “Interval Type II Fuzzy Logic Systems: Theory and Design”. Transactions on Fuzzy Systems. Vol. 8(5). 2000. Pp. 535-550.
- [15] M. Márquez, J. Waissman, O. Gutiérrez. “Fuzzy Model Based Iterative Learning Control for Phenol Biodegradation”. Foundations of Fuzzy Logic and Soft Computing. 2007. Pp. 328-337.
- [16] K. Aström, T. Hägglund. Advanced PID Control. Instrumentation, Systems and Automation Society. 2006. Pp. 128-129.
- [17] Acerca del ININ. www.inin.gob.mx/plantillas/. Febrero, 2011.

VII. Autores

M. en C. Marco Antonio Márquez Vera obtuvo su título de Maestría en Ciencias con especialidad en automatización y control por la Universidad Autónoma del Estado de Hidalgo en el 2007.

M. en C. Filiberto Muñoz Palacios obtuvo su título de Maestría en Ciencias con especialidad en automatización y control por la Universidad Autónoma del Estado de Hidalgo en el 2006.

Lic. Blanca Diana Balderrama Hernández es Licenciada en Educación por la Normal Superior en el 2006.

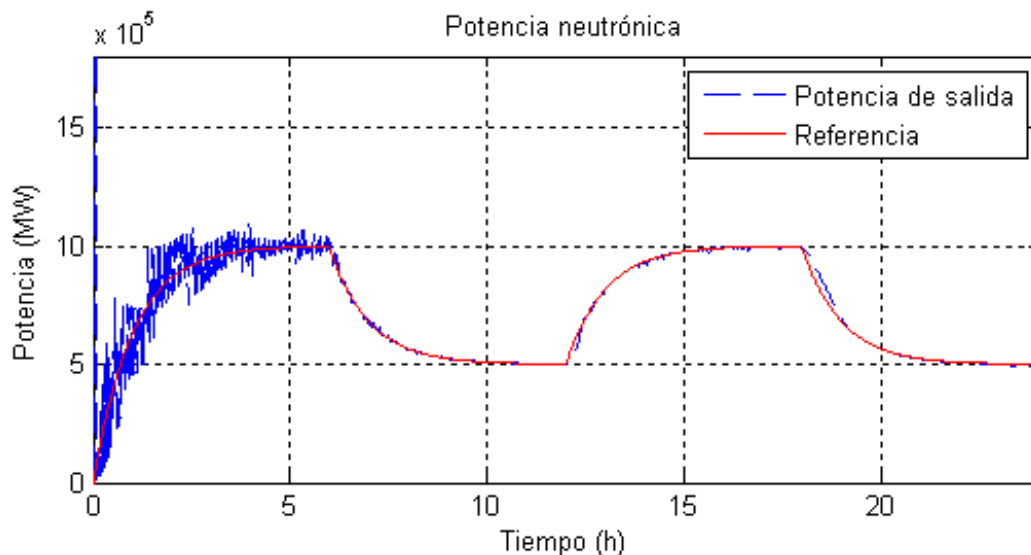


Fig. 5 Simulación del control con lógica difusa tipo I

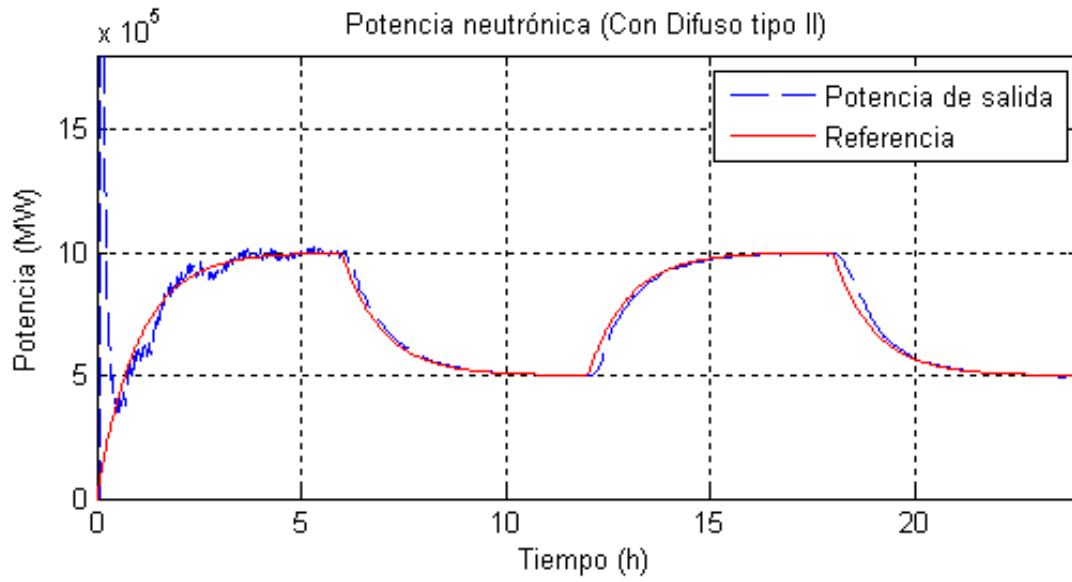


Fig. 6 Simulación del control con lógica difusa tipo II

Control de Velocidad por Pasividad de un Motor de Imanes Permanentes y su Simulación en Simnon

*,**F.H. Ramírez-Leyva, **J. J. Vásquez-Sanjuan, **E. Peralta-Sánchez

*Departamento de Ingenierías, Universidad Popular Autónoma del Estado de Puebla (UPAEP)

**Universidad Tecnológica de la Mixteca (UTM)

**Instituto de Electrónica y Mecatrónica, Carretera a Acatlima Km. 2.5 Huajuapán de León, Oaxaca, C.P. 69000, Tel. (01953) 5320399 ext. (555); hugo@mixteco.utm.mx

Resumen

Los motores de imanes permanentes con sus características de alta densidad de potencia y eficiencia dominan la mayoría de aplicaciones cuando se tienen requerimientos de control de alto desempeño de posición, velocidad y par. En este trabajo se muestra la forma en que se realiza control de velocidad del motor por pasividad así como la demostración de estabilidad de éste. Se comprueba con simulaciones que el controlador funciona adecuadamente. Las simulaciones se hicieron en Simnon y se incluye el código desarrollado.

Palabras clave: Pasividad, Control de Campo orientado, Motor de imanes permanentes

I. Introducción

Los motores de imanes permanentes de corriente directa (CD) o alterna (CA) han incrementado su uso en diversas aplicaciones industriales, especialmente para las de baja potencia. Sus principales características son: alta densidad de potencia y alta eficiencia. El motor de imanes permanentes de CD sin escobillas se utiliza en aplicaciones en el que su alto contenido de rizado de par no impacte, se utiliza principalmente por tener un control simple. El motor de CA sin escobillas, conocido como Motor de Imanes Permanentes (PMSM por sus siglas en inglés) se usa principalmente en aplicaciones que requieren un alto desempeño de control. Actualmente el precio de este tipo de motores, al igual que los procesadores digitales de señales (DSP) de alta eficiencia es competitivo, lo cual ha incentivado su uso, especialmente cuando se requiere de una alta eficiencia [1].

Los métodos de control más comunes para el control del PMSM es el de Campo Orientado (FOC) y el Control Directo de Par (DTC). En el control FOC las ecuaciones del motor son transformadas un sistema de coordenadas que rota en sincronía con el flujo de los imanes

permanentes. Esto permite separar las ecuaciones del motor y con ello controlar indirectamente el flujo y torque mediante un controlador PI de la corriente, de forma similar al muy conocido control de CD de una máquina [2]. El principio de operación del DTC se basa en seleccionar el vector de voltaje que se le aplica al estator en función de las diferencias entre el par de referencia y el actual, así como el flujo magnetizante [3], tiene la ventaja de que el único parámetro que se requiere del motor es la resistencia del estator.

La pasividad es una metodología en la cual se define un controlador cuyo principal objetivo es mantener la respuesta en lazo cerrado en forma pasiva, es decir que disipe energía.

Este trabajo propone un control de fácil implementación para regular la velocidad de salida del motor de imanes permanentes. La propuesta de control usa la retroalimentación de la salida pasiva del error. Es un controlador lineal para un sistema no lineal, el cual se basa en el almacenamiento total de la energía del sistema para después disiparla mediante la retroalimentación de la salida pasiva, que hace perder energía al sistema y con ello consigue

regular las variables físicas del motor a los valores de referencia deseados. La técnica aplicada se encuentra documentada en [4-8].

SIMNON es un programa que resuelve ecuaciones diferenciales y en diferencias en forma numérica. Por tanto es capaz de simular el comportamiento de sistemas continuos y sistemas discretos representados mediante ecuaciones en diferencias. Ha sido utilizado en muchas Universidades del mundo para apoyo en educación e investigación en disciplinas como el Control Automático, Biología, Ingeniería Química, Economía, Ingeniería Eléctrica, etc. También SIMNON ha sido utilizado en la industria para la Simulación de Sistemas de Control [5].

El trabajo se divide en 4 secciones. En la segunda se describe el modelo del motor PMSM así como la demostración de la estabilidad del controlador por pasividad. En la tercera sección se describe la forma en que se realizó el modelo en SIMNON, en la cuarta sección se presentan las simulaciones y en la sección 5 se muestran las conclusiones.

II. Modelo del Motor

El motores de imanes permanentes esta compuesto por el rotor y el estator al igual que el de inducción o el de CD, la diferencia es que usa imanes permanentes para producir el campo magnético en el entrehierro del rotor y el estator. En la figura 1(a) se muestra el rotor y como se colocan los imanes en forma superficial, en la figura 1(b) la fotografía de un motor con un rotor. Para un motor PMSM de AC, el embobinado es similar al de inducción en la parte del estator (al cual se le llama armadura) [2].

De acuerdo a la forma en cómo se colocan las barras de magnetos en el rotor, se clasifica en dos tipos: Motores Síncronos con Magnetos Permanentes en el Interior (IPMSM) y Motores Síncronos con Magnetos Permanentes Superficiales (SPMSM). Los primeros tienen una mayor robustez mecánica, por lo cual se pueden usar en aplicaciones de alta velocidad.

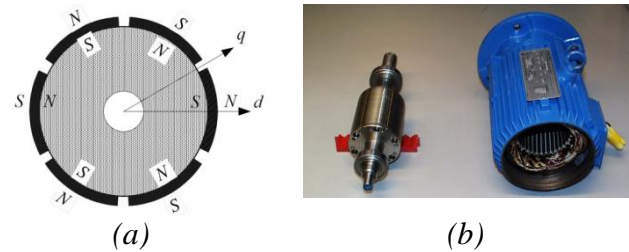


Fig. 1. (a) Diagrama del rotor de un motor de imanes permanentes. (b) Fotografía del motor de imanes y su rotor

En el estator se montan 3 bobinas que tienen una separación espacial de 120° , las cuales se alimentan con voltaje trifásico. La corriente que consumen es función de la inductancia propia y la resistencia de cada bobina, y los flujos acoplados entre cada bobina y su interacción con el flujo que genera el imán permanente.

El modelo del motor de imanes permanentes es muy complicado de analizar en un sistema de referencia trifásico estático, sin embargo éste se simplifica mucho cuando se realiza la transformación de Clarke a un nuevo sistema de referencia de dos ejes (Ecuación 1). Los dos tipos de transformaciones usadas son $\alpha \beta$, que es un sistema estacionario con α y β ortogonales. La transformación de Park asociada al rotor se le llama dq , y es un sistema de referencia de dos ejes que giran a la velocidad del rotor (Ecuación 2). En la figura 2 se muestra la relación de ejes en dq .

$$\begin{bmatrix} V_\alpha \\ V_\beta \end{bmatrix} = \frac{2}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} V_d \\ V_q \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} V_\alpha \\ V_\beta \end{bmatrix} \quad (2)$$

El motor PMSM en el sistema de referencia dq se modela con dos ecuaciones diferenciales (Ecuación 3), donde V_d , V_q , i_d e i_q son los voltajes y corrientes en el eje de referencia dq , L_d y L_q son las inductancias, R la resistencia del estator, Ψ_{PM} es la magnitud del flujo de los imanes permanentes y ω_r la velocidad eléctrica del rotor.

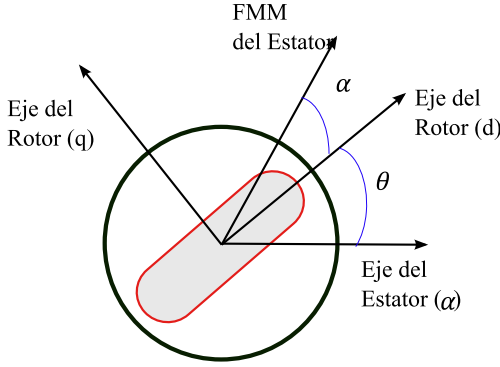


Fig. 2. Representación del motor en los ejes dq

$$\begin{aligned} V_d &= L_d \frac{di_d}{dt} - \omega_r L_q i_q + R i_d \\ V_q &= L_q \frac{di_q}{dt} + \omega_r (L_d i_d + \Psi_{PM}) + R i_q \end{aligned} \quad (3)$$

La parte mecánica del motor se modela con la ecuación diferencial (4), donde τ_e es el par del motor, T_L el par de la carga, T_f el par de la fricción, J el momento de inercia del rotor y P el número de par de polos. τ_e es función de las corrientes en dq , como se muestra en la misma ecuación.

$$\begin{aligned} \tau_e - T_L - T_f &= \frac{J}{P} \frac{d\omega_r}{dt} \\ \tau_e &= \frac{3}{2} P [i_d i_q (L_d - L_q) + i_q \Psi_{PM}] \end{aligned} \quad (4)$$

La velocidad mecánica en revoluciones por minuto (ω_{rpm}) está relacionada con la velocidad eléctrica y el par de polos. La ecuación (5) es la que se utiliza para hacer esta conversión.

$$\omega_{rpm} = \frac{30\omega_r}{P\pi} \quad (5)$$

Para motores con imanes superficiales $L_d=L_q=L_0$, lo cual implica que el par de carga es sólo función de la corriente i_q . Por lo cual la ecuación (4) se transforma de la siguiente manera.

$$\tau_e = \frac{3}{2} P \Psi_{PM} i_q \quad (6)$$

$$\frac{2}{3} \frac{J}{P^2} \frac{d\omega_r}{dt} = \Psi_{PM} i_q - \frac{2}{3P} \left(T_L + \frac{b\omega_r}{P} \right) \quad (7)$$

Si el sistema se representa por ecuaciones diferenciales pone en variables de estado, este queda como:

$$\begin{aligned} & \begin{bmatrix} L_0 \frac{di_d}{dt} \\ L_0 \frac{di_q}{dt} \\ \frac{2}{3} \frac{J}{P^2} \frac{d\omega_r}{dt} \end{bmatrix} \\ &= \begin{bmatrix} \omega_r L_0 i_q - R i_d + V_d \\ -\omega_r (L_0 i_d + \Psi_{PM}) - R i_q + V_q \\ \Psi_{PM} i_q - \frac{2}{3} \frac{T_L}{P} - \frac{2}{3} \frac{b\omega_r}{P^2} \end{bmatrix} \end{aligned} \quad (8)$$

Para poder usar pasividad es necesario representar al sistema de la forma expresada en la ecuación (9), donde $\mathbf{X} = [i_d \ i_q \ \omega_r]^T$ es el vector de estados, \mathbf{A} es una matriz diagonal constante, $\mathbf{u} = [V_d \ V_q]^T$ la entrada de control, $\mathbf{J}(\mathbf{u})$ matriz de interconexión interna, \mathbf{R} matriz de pérdidas y $\boldsymbol{\varepsilon}$ vector de perturbaciones.

$$\mathbf{A}\dot{\mathbf{X}} = \mathbf{J}(\mathbf{u})\mathbf{X} - \mathbf{R}\mathbf{X} + \mathbf{B}\mathbf{u} + \boldsymbol{\varepsilon} \quad (9)$$

Las propiedades que deben de cumplir las matrices $\mathbf{J}(\mathbf{u})$ y \mathbf{R} para que el sistema sea controlable por pasividad son:

$$\mathbf{J}^T(\mathbf{u}) = -\mathbf{J}(\mathbf{u}) \quad (10)$$

$$\mathbf{R}^T(\mathbf{u}) = \mathbf{R} \quad (11)$$

Por inspección se puede demostrar que el sistema en variables de estado (8) representado de acuerdo con la ecuación (9) es la ecuación (12).

A simple vista se observa que la matriz $\mathbf{J}(\mathbf{u})$ cumple con la propiedad (10). La matriz \mathbf{R} es diagonal, por lo cual su transpuesta es igual a ella misma. Con lo cual se cumple con la premisa de que el sistema puede ser controlado por pasividad.

$$\begin{aligned}
& \begin{pmatrix} L_o & 0 & 0 \\ 0 & L_o & 0 \\ 0 & 0 & \frac{2}{3} \frac{J}{P^2} \end{pmatrix} \begin{bmatrix} \frac{di_d}{dt} \\ \frac{di_q}{dt} \\ \frac{d\omega_r}{dt} \end{bmatrix} \\
& = \begin{pmatrix} 0 & \omega_r L_o & 0 \\ -\omega_r L_o & 0 & -\Psi_{PM} \\ 0 & \Psi_{PM} & 0 \end{pmatrix} \begin{bmatrix} i_d \\ i_q \\ \omega_r \end{bmatrix} \\
& - \begin{pmatrix} R & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & \frac{2b}{3P^2} \end{pmatrix} \begin{bmatrix} i_d \\ i_q \\ \omega_r \end{bmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{bmatrix} V_d \\ V_q \end{bmatrix} \\
& + \begin{pmatrix} 0 \\ 0 \\ \frac{2}{3} \frac{1}{P} \end{pmatrix} T_L
\end{aligned} \quad (12)$$

Para aplicar la técnica de pasividad se necesita definir un sistema que es copia del sistema original, como se muestra en la ecuación (13), donde \mathbf{X}^* son las variables de estado deseadas, y \mathbf{u}^* la entrada deseada.

$$\mathbf{A}\dot{\mathbf{X}}^* = \mathbf{J}(\mathbf{u}^*)\mathbf{X}^* - \mathbf{R}\mathbf{X}^* + \mathbf{B}\mathbf{u}^* + \boldsymbol{\varepsilon} \quad (13)$$

Restando la ecuación (12) a la (13) queda:

$$\mathbf{A}\dot{\mathbf{X}} - \mathbf{A}\dot{\mathbf{X}}^* = \mathbf{J}(\mathbf{u})\mathbf{X} - \mathbf{R}\mathbf{X} + \mathbf{B}\mathbf{u} + \boldsymbol{\varepsilon} - \mathbf{J}(\mathbf{u}^*)\mathbf{X}^* + \mathbf{R}\mathbf{X}^* - \mathbf{B}\mathbf{u}^* - \boldsymbol{\varepsilon} \quad (14)$$

$$\mathbf{A}\dot{\mathbf{e}} = \mathbf{J}(\mathbf{u})\mathbf{X} - \mathbf{R}\mathbf{e} + \mathbf{B}\mathbf{e}_u - \mathbf{J}(\mathbf{u}^*)\mathbf{X}^* \quad (15)$$

Donde $\mathbf{e} = \mathbf{X} - \mathbf{X}^*$ y $\mathbf{e}_u = \mathbf{u} - \mathbf{u}^*$. Sumándole y restándole el término $\mathbf{J}(\mathbf{u}^*)\mathbf{X}$, la ecuación (15) se modifica a

$$\mathbf{A}\dot{\mathbf{e}} = \mathbf{J}(\mathbf{u})\mathbf{X} - \mathbf{R}\mathbf{e} + \mathbf{B}\mathbf{e}_u - \mathbf{J}(\mathbf{u}^*)\mathbf{X}^* + \mathbf{J}(\mathbf{u}^*)\mathbf{X} - \mathbf{J}(\mathbf{u}^*)\mathbf{X} \quad (16)$$

$$\mathbf{A}\dot{\mathbf{e}} = [\mathbf{J}(\mathbf{u}) - \mathbf{J}(\mathbf{u}^*)]\mathbf{X} - \mathbf{R}\mathbf{e} + \mathbf{B}\mathbf{e}_u + \mathbf{J}(\mathbf{u}^*)\mathbf{e}$$

Esta ecuación se linealiza a $\mathbf{J}(\mathbf{u})$ por series de Taylor, alrededor de \mathbf{u}^* , quedando:

$$\mathbf{J}(\mathbf{u}) = \mathbf{J}(\mathbf{u}^*) + \left. \frac{\partial \mathbf{J}(\mathbf{u})}{\partial \mathbf{u}} \right|_{\mathbf{u}^*} \mathbf{e}_u \quad (17)$$

Despejando la derivada parcial de la ecuación (17) y como la matriz \mathbf{J} es constante, entonces la derivada parcial es 0.

$$\mathbf{J}(\mathbf{u}) - \mathbf{J}(\mathbf{u}^*) = \left. \frac{\partial \mathbf{J}(\mathbf{u})}{\partial \mathbf{u}} \right|_{\mathbf{u}^*} \mathbf{e}_u = 0 \quad (18)$$

Sustituyendo la ecuación (18) en la (16) se obtiene

$$\mathbf{A}\dot{\mathbf{e}} = -\mathbf{R}\mathbf{e} + \mathbf{B}\mathbf{e}_u + \mathbf{J}(\mathbf{u}^*)\mathbf{e} \quad (19)$$

Para demostrar que el sistema es estable, se emplea el segundo principio de Lyapunov. Para ello se propone la función candidata de Lyapunov definida positiva.

$$V(\mathbf{e}) = \frac{1}{2} \mathbf{e}^T \mathbf{A} \mathbf{e} \quad (20)$$

$V(\mathbf{e})$ es definida positiva, ya que la matriz \mathbf{A} es diagonal y los elementos de la diagonal son mayores a cero. Derivando con respecto al tiempo se obtiene la ecuación (21).

$$\dot{V}(\mathbf{e}) = \mathbf{e}^T \mathbf{A} \dot{\mathbf{e}} = -\mathbf{e}^T \mathbf{R} \mathbf{e} + \mathbf{e}^T \mathbf{B} \mathbf{e}_u + \mathbf{e}^T \mathbf{J}(\mathbf{u}^*) \mathbf{e} \quad (21)$$

Sea $\mathbf{e}_u = -\boldsymbol{\gamma} \mathbf{B}^T \mathbf{e}$ y por ser $\mathbf{e}^T \mathbf{J}(\mathbf{u}^*)$ asimétrica, entonces $\mathbf{e}^T \mathbf{J}(\mathbf{u}^*) \mathbf{e} = 0$, sustituyendo esto en la ecuación (21) queda

$$\dot{V}(\mathbf{e}) = -\mathbf{e}^T [\mathbf{R} + \mathbf{B} \boldsymbol{\gamma} \mathbf{B}^T] \mathbf{e} \quad (22)$$

Para que \dot{V} sea definida negativa $[\mathbf{R} + \mathbf{B} \boldsymbol{\gamma} \mathbf{B}^T] \geq 0$. Con lo cual queda demostrada la estabilidad del controlador.

Para encontrar la ley de control se parte del hecho que $\mathbf{e}_u = \mathbf{u} - \mathbf{u}^* = -\boldsymbol{\gamma} \mathbf{B}^T \mathbf{e}$, de donde se obtiene la ecuación (23)

De donde se obtiene la ley de control $\mathbf{u} = \mathbf{u}^* - \boldsymbol{\gamma} \mathbf{B}^T \mathbf{e}$, quedando como se muestra en la ecuación (24)

$$\mathbf{e}_u = - \begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{bmatrix} i_d - i_d^* \\ i_q - i_q^* \\ \omega_r - \omega_r^* \end{bmatrix} \quad (23)$$

$$= - \begin{bmatrix} \gamma_1(i_d - i_d^*) \\ \gamma_2(i_q - i_q^*) \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} V_d \\ V_q \end{bmatrix} = \begin{bmatrix} V_d^* - \gamma_1(i_d - i_d^*) \\ V_q^* - \gamma_2(i_q - i_q^*) \end{bmatrix} \quad (24)$$

El valor deseado V_d^*, V_q^*, i_d^* e i_q^* se obtienen a partir del punto de equilibrio del sistema descrito por la ecuación (13), el cual se obtiene cuando $\dot{\mathbf{X}}^* = 0$. Esto implica que:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \omega_r^* L_q i_q^* - R i_d^* + V_d^* \\ -\omega_r^* (L_d i_q^* + \Psi_{PM}) - R i_q^* + V_q^* \\ \Psi_{PM} i_q^* - \frac{2}{3P} \left(T_L + \frac{b\omega_r^*}{P} \right) \end{bmatrix} \quad (25)$$

Despejando se llega a

$$\begin{aligned} V_d^* &= R i_d^* - \omega_r^* L_q i_q^* \\ V_q^* &= R i_q^* + \omega_r^* (L_d i_q^* + \Psi_{PM}) \\ i_q^* &= \frac{1}{\Psi_{PM}} \frac{2}{3P} \left(T_L + \frac{b\omega_r^*}{P} \right) \end{aligned} \quad (26)$$

Para tener el máximo torque $i_d^* = 0$, sustituyendo esto en la ecuación (26), se obtiene el valor deseado.

$$\begin{aligned} V_d^* &= -\omega_r^* L_q i_q^* \\ V_q^* &= \omega_r^* (L_d i_q^* + \Psi_{PM}) \\ i_q^* &= \frac{1}{\Psi_{PM}} \frac{2}{3P} \left(T_L + \frac{b\omega_r^*}{P} \right) \end{aligned} \quad (26)$$

Sustituyendo la ecuación (26) en la ecuación (24), se obtiene la ley de control para el motor

III. Implementación en Simnon

Para comprobar al controlador por pasividad, se usó el simulador matemático Simnon. En forma conceptual el diagrama a bloques del controlador se muestra en la figura 3. Las entradas al controlador son la velocidad deseada y el par de carga, con base en estos calcula las corrientes y voltajes deseados, la sintonización del

controlador se hace variando las ganancias γ_1 y γ_2 hasta obtener la mejor respuesta. Del controlador se determina el voltaje V_d y V_q , que el inversor se encarga de convertir a su voltaje equivalente en las 3 fases (V_a, V_b y V_c). En el motor se le miden las corrientes de la fase a y b, y se estima la corriente de la fase c partiendo del hecho que es un sistema balanceado. Se requiere conocer la posición del roto ya que la transformación de Park las usa.

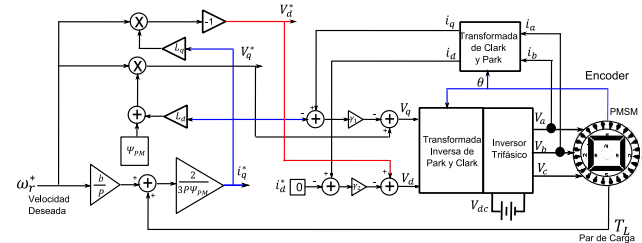


Fig. 3. Diagrama a bloques del control por pasividad

Los parámetros del modelo del motor se muestran en la tabla 1. Este tipo de motor es de imanes superficiales, por esa razón las inductancias son iguales.

Voltaje del bus V_{DC}	300V
Par de polos P	3
Magnetización de los Imanes Ψ_{PM}	0.2547 Wb
Momento de Inercia J	0.000141 kg/m ²
Resistencia R	0.86 Ω
Inductancia en d L_d	0.00665 H
Inductancia en q L_q	0.00665 H
Fricción Viscosa B	0.00038 Ns

Tabla 1. Parámetros del motor

En la implementación en Simnon, se decidió usar un archivo de tiempo continuo en donde se definió la planta, este archivo se llama ModPMSM.T. Este modelo recibe los voltajes de cada fase y genera las corrientes de fase, la velocidad y el par de carga. Mediante un archivo de interconexión (Conex.T) se declara la interconexión las señales. El control por pasividad se hizo en un archivo de tiempo discreto ConPas1k.T, el cual recibe las variables del modelo del motor y genera los voltajes requeridos, el tiempo de muestreo de sistema de tiempo discreto fue de 100 μ s. En la figura 4 se

muestra las relaciones de señales entre los diferentes archivos. En el apéndice 1 se muestra el código de cada uno de los archivos.

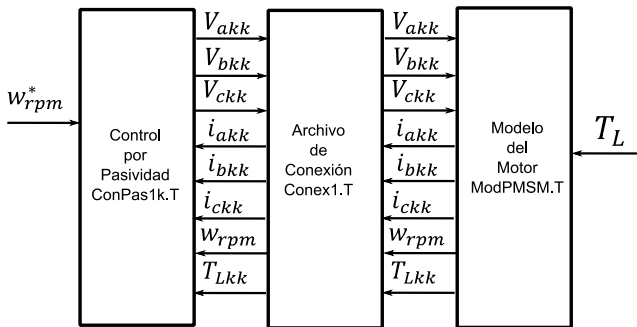


Fig. 4. Representación a bloques del motor con controlador en Simnon.

Una vez capturado el modelo del sistema en Simnon, se procedió a sintonizar las ganancias del controlador. De acuerdo a la ecuación (22), cualquier valor positivo de ellas puede ser usado. Sin embargo al correr las simulaciones se observó que a medida que el valor de γ_2 se hacía más grande aumentaba el rizo en el par. El valor de las ganancias que mejores resultados dieron son $\gamma_1=10$ y $\gamma_2=1$.

IV. Pruebas y Resultados

Para comprobar la efectividad del control por pasividad, se simuló considerando el par de carga y la velocidad variables. En la figura 5 se muestran los resultados de la simulación de la velocidad. La velocidad deseada es de 500rpm que está de color verde, la respuesta simulada (de color negro) se aproxima con muy poco error en este rango. Cuando la velocidad deseada es -500rpm el error es menor al 10%. Cuando nuevamente la velocidad es positiva el error es pequeño.

En la figura 6 se muestra el par de carga y el par generado por el motor. De 0 a $t=1.2s$ el de carga es 0Nm, de 1.2s a 2.5s $T_L=5Nm$, de 2.5 a 3s $T_L=-5Nm$, de 3s a 4 s 5Nm y de 4s a 6s $T_L=4Nm$. Como se puede ver, el par generado por el motor es el mismo que el de carga, excepto cuando hay cambios de velocidad.

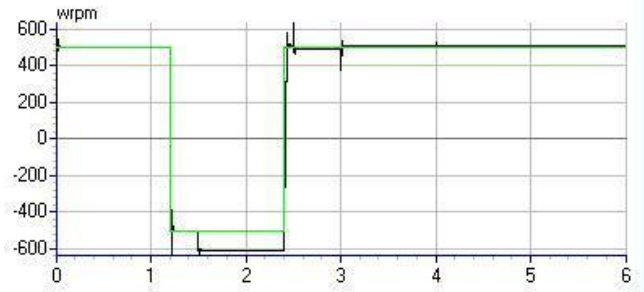


Fig 5. Velocidad deseada y su respuesta

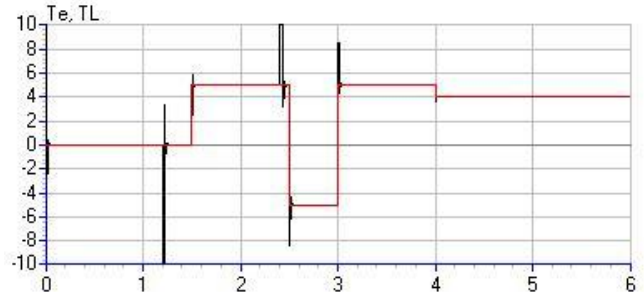


Fig. 6. Par de carga y par producido por el motor

En la figura 7 se muestran las gráficas de valfa y vbeta, y en la figura 8 se muestran los voltajes de las fases en a, b y c, en ambos casos los voltajes están acotados entre $\pm 50V$ y tienen patrones muy similares. La figura 9 muestra el voltaje de beta vs voltaje en alfa, como se puede ver todos los voltajes se encuentran dentro de un círculo de aproximadamente 50V como lo predice la teoría.

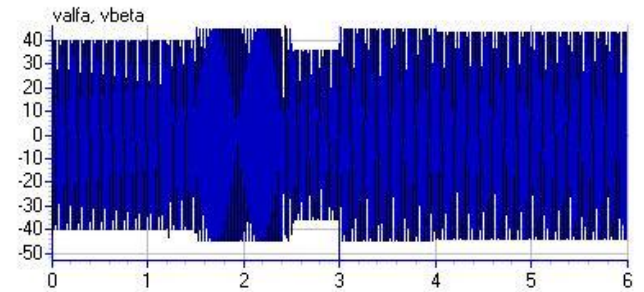


Fig. 7. Gráfica del Voltaje alfa y beta vs tiempo

La figura 10 muestra la corriente en alfa y beta, la mayor parte del tiempo se mantiene acotado en el rango de $\pm 5A$, excepto cuando ocurren cambios bruscos de velocidad, en esos momentos se tienen picos de corriente de 10A y -35A. La figura 11 muestra las corrientes de las fases a, b y c, comparándola con la figura 10, ambas gráficas tienen perfiles muy parecidos, ya que la

transformación entre uno y otro sistema es una constante.

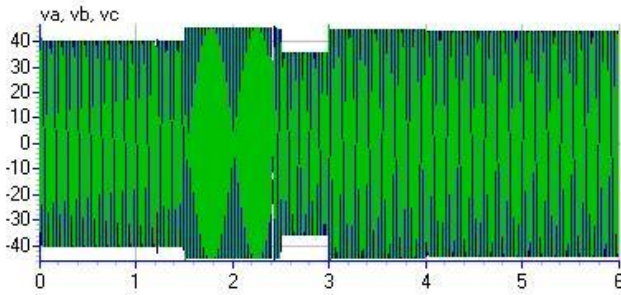


Fig. 8. Gráfica de voltaje de cada fase

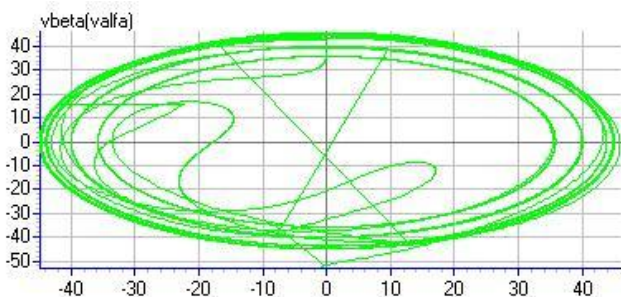


Fig. 9. Voltaje en beta vs voltaje alfa

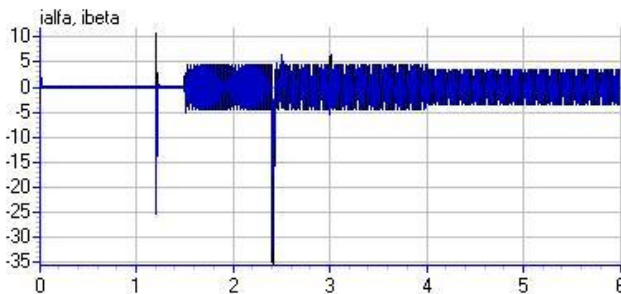


Fig. 10. Gráfica de la corriente el alfa y beta vs tiempo

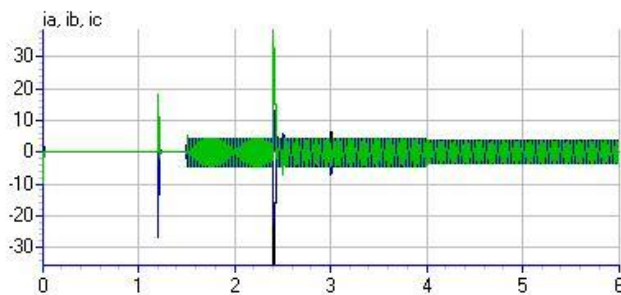


Fig. 11. Gráfica de la corriente de cada fase

VI. Conclusiones

En este trabajo se mostró que la forma en que se realiza el control de un motor de imanes permanentes por la técnica de pasividad, con su respectiva demostración de estabilidad. Para su realización se requiere tener las mediciones del par de carga y la posición del rotor.

Para su comprobación se realizó el modelo del motor en Simnon, por ser muy flexible y fácil de usar, y se comprobó que para velocidades positivas se tiene una muy buena respuesta tanto en par como en velocidad. Para velocidades negativas se tiene un error menor al 10%.

Como trabajos futuros se pretende trabajar en un estimado de carga y observar la respuesta del controlador. También se pretende realizar la comprobación experimental del controlador y comparar las ventajas de este controlador contra los usados comúnmente como son el de campo orientado y el directo de par.

VII. Autores

F. Hugo Ramírez Leyva obtuvo su título de Maestría en Ciencias con especialidad en electrónica por el Instituto Nacional de Astrofísica Óptica y Electrónica. Actualmente es profesor investigador, adscrito al Instituto de Electrónica y Mecatrónica de la Universidad Tecnológica de la Mixteca. Actualmente es estudiante de doctorado en Ingeniería Mecatrónica en la Universidad Popular Autónoma del Estado de Puebla.

Jacob J. Vásquez Sanjuan. Es Ingeniero en electrónica egresado de la BUAP, obtuvo su título de Maestría en Ciencias con especialidad en Electrónica en el Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), realizó estudios en el doctorado en Electrónica y Computación en la Universidad Tecnológica de la Mixteca y actualmente realiza su tesis doctoral en modelado de flujo vehicular

Edgar Peralta Sánchez. Licenciado en Electrónica por la BUAP, Puebla, México

(1995). Maestro en ciencias con especialidad en Ingeniería Eléctrica por LA UDLAP, Puebla, (1999). PhD en Ingeniería Eléctrica por la Universidad de Manchester, Reino Unido. (2006). Actualmente es Catedrático en el Departamento de Ingenierías de la Universidad Popular Autónoma del Estado de Puebla (UPAEP).

VIII. Referencias

- [1] Zeroug. H and Hocine. L, “Sensorless Control of a Surface Mounted Permanent Magnet Motor by Signal Injection Associated with Space Vector Modulation Technique”, pag 1573-1580, Electric Machines and Drives Conference, 2009. IEMDC '09.
- [2] Dariusz Świerczyński, “Direct Torque Control with Space Vector Modulation (DTC-SVM) of Inverter-Fed Permanent Magnet Synchronous Motor Drive”, Doctoral Thesis, Warszawa, 2005, Warsaw University Of Technology.
- [3] L. Zhong, M. F. Rahman, W.Y. Hu & K. W. Lim. “Analysis of Direct Torque Control in Permanent Magnet Synchronous Motro Drives”. IEEE Transactions on Power Electrónica, Vol. 12, No. 3, May 1997.
- [4] Linares-Flores, J., Sira-Ramírez, H., Yescas-Mendoza, E. and Vásquez-Sanjuan, J. J. (2011), “A comparison between the algebraic and the reduced order observer approaches for on-line load torque estimation in a unit power factor rectifier-DC motor system”. Asian Journal of Control, 13: n/a. doi: 10.1002/asjc.234.
- [5] H. Sira-Ramírez, R. Silva-Ortigoza, “Control Design Techniques in Power Electronics Devices”, Springer-Verlag, Power Systems Series, London, 2006.
- [6] J. Linares-Flores, J. Reger, H Sira-Ramírez, “A time-varying linear state feedback tracking controller for a Boost-converter driven DC machine”, in Proc. 4th IFAC-Symposium on Mechatronic Systems,

Heidelberg, Germany, pp. 926-931, Sep. 2006.

- [7] J. Linares-Flores, H. Sira-Ramírez, J. Reger, and R. Silva-Ortigoza, “An exact tracking error dynamics passive output feedback controller for a buck-Boost-converter driven DC motor”, In Proc. IEEE International Power Electronics Congress, Cholula, Pue., México, pp. 1-5, October 2006.
- [8] J. Linares-Flores, J. Reger, and H. Sira-Ramírez, “Sensorless tracking control of two DC-drives via a double Buck-converter”, In Proc. 45Th IEEE Conference on Decision and Control, San Diego, CA, USA, pp. 6229-6234, December 2006.
- [9] Simnon, “<http://dctrl.fi-b.unam.mx/tutoriales/digital/simnon21.html> . Fecha de último acceso 22/Ocubre/2010.
- [10] Vladan Petrovic, Romeo Ortega, and Aleksandar M. Stankovic; Interconnection and Damping Assignment Approach to Control of PM Synchronous Motors; IEEE Transactions on Control Systems Technology, Vol. 9, No. 6, November 2001.

Modelo de la planta ModPMSM.T

CONTINUOUS SYSTEM ModPMSM

"Este archivo simula el motor e inversor

"Las entradas son los voltajes de fase.

"Las salidas son las corrientes de fase y la velocidad

"Created: 13/03/2011

" Inputs and outputs:

INPUT Vakk Vbkk Vckk

OUTPUT iakk ibkk ickk wrpmkk TLkk

STATE id iq w teta

DER did diq dw dteta

TIME t

"Envia y recibe variables

va=Vakk

vb=Vbkk

vc=Vckk

iakk=ia

ibkk=ib

```

ickk=ic
wrpmkk=wrpm

"Modelo del motor
did=(Vd+w*Lq*iq-R*id)/Ld
diq=(Vq-w*(Ld*id+Ppm)-R*iq)/Lq
Te1=(3/2)*P*( id*iq*(Ld-Lq)+iq*Ppm )
Te2=if Te1<Temax then Te1 else Temax
Te= if Te2>-Temax then Te2 else -Temax

```

```
dw=P*(Te-TL)/J - B*w/J
```

```

utao1=if t<1.5 then 0 else 5
utao2=if t<2.5 then 0 else -10
utao3=if t<3.0 then 0 else 10
utao4=if t<4.0 then 0 else -1
TL2=utao1+utao2+utao3+utao4+0.0001
TL=TL2
TLkk=TL

```

```

dteta=w
wmec=w/P
wrpm=wmec*30/pi "Velocidad en rpm
wkk=wrpm

```

```

"Transformacion alfa beta
Valfa=(2/3)*(va-vb/2-vc/2)
Vbeta=(2/3)*(sqrt(3)/2)*(vb-vc)
Vo=va+vb+vc

```

```

Vd=Valfa*cos(teta)+Vbeta*sin(teta)
Vq=-Valfa*sin(teta)+Vbeta*cos(teta)

```

```

ialfa=id*cos(teta)-iq*sin(teta)
ibeta=id*sin(teta)+iq*cos(teta)

```

```
"CAMBIO
```

```

ia=id*cos(teta) - iq*sin(teta)
ib=id*cos(teta-2*pi/3) - iq*sin(teta-2*pi/3)
ic=id*cos(teta+2*pi/3) - iq*sin(teta+2*pi/3)

```

```
" Initializations:
```

```

pi=3.1416
Vdc=300
P=3
Ppm=0.2547
J=14.1e-4

```

```

R=0.86
Lq=0.00665
Ld=0.00665
Tf=0
B=0.00038
Temax=10
END

```

Modelo del controlador ConPas1k.T

```

DISCRETE SYSTEM ConPas1k
" Este archivo implementa el control por
pasividad en un motor PMSM
"Created: 3/05/2011
"Sistema de control en tiempo discreto
"Se uno con ModPMSM y Conex1

INPUT iakk ibkk ickk wrpmkk TLkk
OUTPUT Vakk Vbkk Vcck

" States and time variables:
STATE inwrpmka
NEW inwrpmk
TIME t
TSAMP ts
ts=t+h
h=10e-5

"Envia y recibí Variables
Vakk=vak2
Vbkk=vbk2
Vcck=vck2
iak=iakk
ibk=ibkk
ick=ickk
TLk=TLkk
wrpmk=wrpmkk
inwrpmk=inwrpmka+(h*wrpmk*P*pi)/30
arotk=inwrpmk

ialfak=(2/3)*(iak-ibk/2-ick/2)
ibetak=(2/3)*(sqrt(3)/2)*(ibk-ick)
iok=iak+ibk+ick
idk=ialfak*cos(arotk) + ibetak*sin(arotk)
iqk=-ialfak*sin(arotk) + ibetak*cos(arotk)

```

```
"CONTROLADOR DE VELOCIDAD
```

```

es1=if t>1.2 then -2 else 0
es2=if t>2.4 then 2 else 0
wref=500*(es1+es2+1) "En RPM
wref2=wref*pi*P/(30) "Eléctrican con factor
de corrección
ewk=wref-wrpmk

idbk=0

iqbk=(TLk+B*wref2/P)*2/(3*Ppm*P)*(B*wref)/
Kt
vdbk=R*idbk-wref2*Lq*iqbk
vqbk=(R*idbk)+wref2*(Ld*iqbk+Ppm)

lambda1=10
lambda2=1
Vdslk=vdbk-lambda1*(idk-idbk)
Vqslk=vqbk-lambda2*(iqk-iqbk)

Vd1k=if Vdslk>Vdc then Vdc else Vdslk
Vdk=if Vd1k<-Vdc then -Vdc else Vd1k
Vq1k=if Vqslk>Vdc then Vdc else Vqslk
Vqk=if Vq1k<-Vdc then -Vdc else Vq1k

vak2=vdk*cos(arotk) - vqk*sin(arotk)
vbk2=vdk*cos(arotk-2*pi/3) - vqk*sin(arotk-
2*pi/3)
vck2=vdk*cos(arotk+2*pi/3) -
vqk*sin(arotk+2*pi/3)

valfak=vdk*cos(arotk) - vqk*sin(arotk)
vbetak=vdk*sin(arotk) + vqk*cos(arotk)

" Parameter values:
pi=3.1416
Vdc=120
P=3
Ppm=0.2547
J=14.1e-4
R=0.86
Lq=0.00665
Ld=0.00665
Tf=0
B=0.00038
Kt=0.272
END

```

Archivo de Interconexión Conex1.T

```

CONNECTING SYSTEM Conex1
"Archivo de interconexion
"Created: 13/03/2011

" Time, if needed:
TIME t
Vakk[ModPMSM]=Vakk[ConPas1k]
Vbkk[ModPMSM]=Vbkk[ConPas1k]
Vckk[ModPMSM]=Vckk[ConPas1k]
iakk[ConPas1k]=iakk[ModPMSM]
ibkk[ConPas1k]=ibkk[ModPMSM]
ickk[ConPas1k]=ickk[ModPMSM]

wrpmkk[ConPas1k]=wrpmkk[ModPMSM]
TLkk[ConPas1k]=TLkk[ModPMSM]

" Connections:

END

```

Macro Macro1.T

```

MACRO Macro1
"Macro para la simulacion del control por
pasividad
"Created: 13/03/2011

" Enter commands here:
syst ModPMSM ConPas1k Conex1
store Te TL wrpm teta arotk valfak vbetak Wref
ewk w
simu 0 6 10e-5
split 2 1
ashow wrpm wref
text 'wrpm'
ashow Te TL
text 'thk'

END

```

Apéndice 1. Modelo en Simnon

Supervisión y vigilancia de una planta de tratamiento de agua potable

*Héctor Ricardo Hernández De León, **Jorge Camas Anzueto, ***Madaín Pérez Patricio

Instituto Tecnológico de Tuxtla Gutiérrez, Carretera Panamericana Km. 1080, Col. Terán, CP 29050-Tuxtla Gtz., Chiapas, (961)6150380 ext 325.

* hhernandezd@ittg.edu.mx, **jcamas@ittg.edu.mx, ***mperez@ittg.edu.mx.

Resumen

La industria del agua potable tiene una presión creciente para producir un agua tratada de mayor calidad a costos más reducidos. Los esquemas de supervisión y vigilancia para el presente proceso biológico se relacionan con el comportamiento de un sistema dinámico condicionado por las acciones que se ejerzan sobre el mismo. Esas acciones pueden ser ejercidas como acciones deseadas, a través de variables manipuladas o a través de variables no manipuladas directamente, generalmente llamadas perturbaciones. Los efectos de estas acciones se pueden ver reflejados en una o más variables del sistema que bajo ciertas condiciones se desea mantener en un valor determinado (variables controladas). El proceso de clarificación es una de las operaciones necesarias para tratar el agua en una planta potabilizadora típica. Esta operación consiste principalmente de la unidad Coagulación-Floculación el cual permite eliminar las partículas coloidales y fuentes de contaminación. Su vigilancia, en la mayor parte de las instalaciones es manual y requiere de análisis de laboratorio largos y costosos (fenómeno biológico). El aspecto innovador de la presente investigación reside en la aplicación de técnicas de control inteligente y multivariable en el análisis de información de la planta potabilizadora (lógica difusa para el clasificador).

Palabras Clave: Supervisión, vigilancia, planta de tratamiento de agua potable, clasificador difuso, sistema de adquisición de datos.

I. Introducción

Las herramientas clásicas de supervisión que se aplican a un proceso industrial, como lo es el tratamiento del agua potable, deben ser complementadas por herramientas de vigilancia, de diagnóstico y de un apoyo a la decisión de tal forma que se integren globalmente a la supervisión [1] (figura 1).

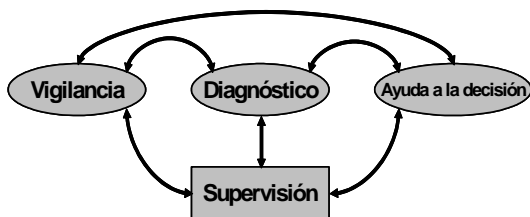


Fig. 1. Vigilancia, diagnóstico y ayuda a la decisión a nivel de la supervisión.

La supervisión consiste principalmente en vigilar la ejecución de una operación, en seguida,

proponer acciones correctivas en caso de ser necesarias. Así, la vigilancia utiliza los datos provenientes de la planta, los cuales son procesados con el objetivo de generar estados funcionales de comportamiento del proceso y establecer un diagnóstico real interpretadas por el experto.

El control en un proceso industrial debe permitir la estabilidad y precisión dentro de límites aceptables de funcionamiento. Este es el objetivo del proyecto de supervisión y diagnóstico en plantas de tratamiento de agua potable, dentro del cual se sitúa el presente trabajo de investigación. La aplicación a una planta real particulariza las especificaciones del problema, los tipos de señales y las fallas posibles. Esencialmente el sistema de supervisión ha de determinar si el funcionamiento se encuentra en una fase normal o anormal, y en consecuencia decidir, o ayudar a

la toma de decisiones. La presente investigación se articula en dos partes principalmente: (1) el sistema de adquisición de datos de las principales variables físicas del proceso y (2) el análisis de la información proporcionado por las variables a la entrada de la planta potabilizadora, con la finalidad de establecer un esquema de vigilancia de la planta.

Muchos esquemas generales de supervisión y diagnóstico aplicados a diferentes dominios han sido propuestos [2,3]. Desde un punto de vista general las diferentes arquitecturas presentadas son similares y pueden ser descritas tomando en cuenta las siguientes funciones generales: la detección de fallas, el diagnóstico, la reconfiguración del proceso y el mantenimiento [4,5]. En general, el diagnóstico de un proceso industrial puede ser separado en dos categorías tomando en cuenta si ellos necesitan explícitamente un modelo de proceso o si ellos están basados en la disposición de datos históricos de funcionamiento del proceso. La primera categoría reposa en un conocimiento a profundidad del sistema incluyendo las relaciones causales entre los diferentes elementos mientras que la segunda, en el conocimiento a partir de experiencias pasadas [6]. La inclusión de una aproximación difusa para implementar una estrategia de diagnóstico del proceso en cuestión es aplicada en el presente artículo, por medio de la utilización de conjuntos difusos de las diferentes variables a tomarse en cuenta.

Un sistema de inferencias difusas se utiliza con la finalidad de que sea capaz de imitar la toma de decisión de un operador humano a partir de reglas verbales traduciendo sus conocimientos relativos a un proceso en particular. La relación matemática entre una falla y sus síntomas es la más difícil de obtener. Por tanto, los operadores humanos responsables del mantenimiento y de la operación del sistema son a menudo capaces, por su experiencia, de determinar, en una base de observaciones, el elemento que falla y que es el origen del comportamiento anormal. Este tipo de conocimiento puede representarse, con la ayuda de reglas de la forma:

IF condición THEN conclusion

donde condición son los síntomas observados y conclusión es el elemento que falla. Este tipo de conocimientos puede ser utilizado para construir un sistema de ayuda al diagnóstico de una instalación industrial. La noción de subagrupamiento difuso fue introducida por Zadeh en 1965.

Una breve descripción de una planta típica de tratamiento de agua potable y sus unidades implicadas se presenta en la siguiente sección. Enseguida, se presenta la estrategia de clasificación usando conjuntos difusos, después, la metodología usada para el diseño y para terminar, los resultados experimentales y conclusiones.

El tratamiento de agua implica procesos químicos y biológicos que transforman a este en agua potable que satisface un conjunto de estándares de calidad en un precio razonable, siempre respetando el cuidado del medio ambiente.

El objetivo del presente artículo es el análisis de las variables más importantes a la entrada del proceso de potabilización en un esquema de diagnóstico en tiempo real a partir de la clasificación difusa y la relación existente entre estados funcionales de comportamiento de la planta.

II. Planta potabilizadora bajo estudio

La planta de tratamiento de agua potable "SMAPA" [7] (ciudad de Tuxtla, México), que fue utilizada como sitio de prueba para este estudio, proporciona el agua a más de 800 mil habitantes y tiene una capacidad nominal de 800 l/s de agua. La figura 2 presenta la descripción esquemática de las operaciones necesarias para tratar el agua en una planta típica. Generalmente abarca 5 unidades: pretratamiento, preoxidación, clarificación, desinfección y refinación.

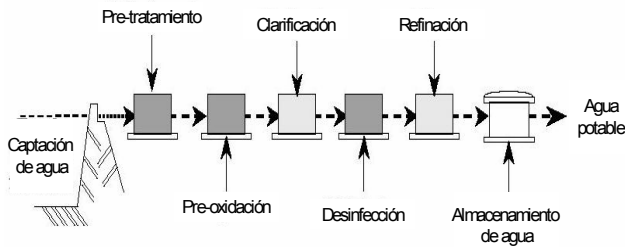


Fig. 2. Planta típica de tratamiento de agua potable.

El proceso de la coagulación (etapa de clarificación) es causado agregando al agua un compuesto altamente iónico (sulfato de aluminio) [8,9].

La reacción forma un precipitado que atrae los sólidos y las partículas coloidales (partículas negativas). Este precipitado, en forma de lodo, se posiciona en el fondo del tanque por gravedad (figura 3). La etapa siguiente es la filtración, donde partículas existentes aun en el agua quedan atrapadas por camas de arena.

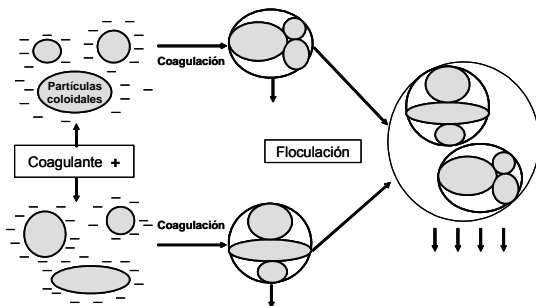


Fig. 3. Proceso coagulación-floculación.

La desinfección utiliza cloro. El agua después se almacena en un tanque y está lista para ser transportada a través de la red de distribución. En este proceso la supervisión y control son esenciales.

III. Desarrollo. El método de clasificación usando lógica difusa.

A partir de la disponibilidad de los datos de la planta de tratamiento de agua potable, se establecen los conjuntos difusos mediante la obtención del grado de adecuación de un objeto a una clase, denominada adecuación marginal. Es importante tomar en cuenta separadamente los datos de cada uno de los descriptores a clasificar.

Una de las principales características del presente esquema de clasificación basado en conjuntos difusos es el análisis de los datos, en el cual se introduce el concepto de la indistinguibilidad total (clase no-informativa NIC). Esta clase acepta todos los objetos con la misma adecuación. Esta adecuación actúa como un umbral mínimo que debe asumir un elemento de una clase significativa. También, el modelo difuso permite efectuar clasificaciones supervisadas por un experto.

Las funciones de aprendizaje y de reconocimiento son simples y rápidas. El ajuste del parámetro de asociación de características permite la obtención de familias de particiones ordenadas con respecto al concepto de "exigencia". La descripción del algoritmo es el siguiente: Se considera una colección de objetos o situaciones X y un conjunto finito de n descriptores cualitativos o cuantitativos A . Un objeto viene representado por un vector con n componentes del vector x donde el componente j^a es el valor tomado por el descriptor j^o del objeto. La información que cada descriptor aporta contribuye al grado de pertenencia de un elemento a una clase mediante el grado de adecuación marginal (MAD) como se muestra en la figura 4.

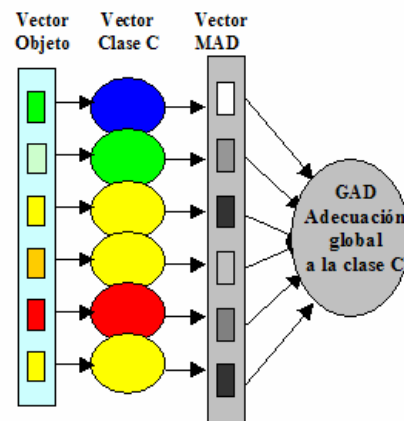


Fig. 4. Método de clasificación empleando lógica difusa: el grado de adecuación marginal.

El grado de adecuación global (GAD) a cada clase se obtiene combinando los MAD mediante un operador lógico difuso (fuzzy) como se indica en la misma figura. El uso de una compensación

lineal entre un operador de intersección y su correspondiente operador de unión y demuestran que un incremento del parámetro grado de exigencia conduce a una clasificación más restrictiva. El cálculo de los parámetros de la clase se hace independientemente para cada componente, es decir usando los datos de cada sensor. Las funciones necesarias para estos cálculos son particularmente sencillas pues sólo utilizan cálculos de promedios, lo cual puede realizarse secuencialmente. Cada vector representativo de un elemento, o situación del vector x se verá atribuido a una de las clases pre-existentes a condición de que su adecuación global (GAD) sobrepase la de la clase no-informativa NIC, en caso contrario el elemento se atribuye a la clase NIC, si no se efectúa auto-aprendizaje, o se crea una nueva clase inicializada por dicho vector x . Es así innecesario el conocimiento previo del número de clases total de la clasificación. En la figura 5 se representa la estructura de la clasificación con 2 clases, una de ellas pudiendo ser la clase NIC, como una red neuronal de dos capas, la primera capa corresponde a los valores de los atributos del objeto a clasificar, y la unidad de salida consiste en un elemento de búsqueda del máximo de adecuación.

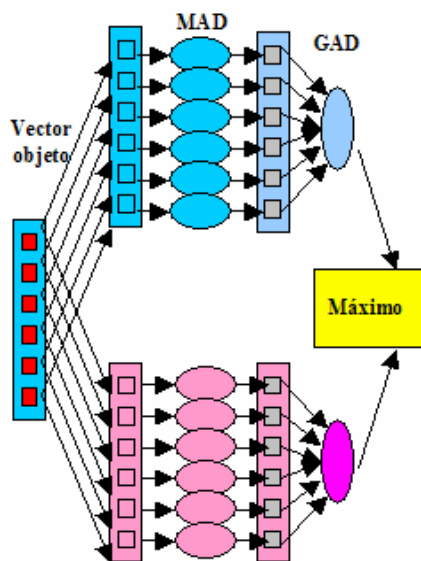


Fig. 5. Método de clasificación difuso: estructura de la clasificación con 2 clases.

Con esta estructura se resuelve el problema del reconocimiento pasivo de objetos y su atribución a la clase de máxima adecuación y aunque su respuesta es una clasificación estricta, se construye una clasificación difusa.

Ahora bien, en relación a la supervisión y vigilancia del proceso, a continuación se presenta una estrategia que se articula alrededor de dos desarrollos: (1) un sistema de adquisición de datos tomando en cuenta las variables más importantes del proceso y (2) una estrategia general para la vigilancia de dicha planta, aplicando el método de reconocimiento de formas basado en clasificación difusa.

La primera parte se compone de la parte electrónica de adquisición inalámbrica de datos (figura 6). Un sistema de adquisición de datos (SAD) es un sistema que puede ser utilizado para el control automático de un proceso o simplemente para el monitoreo de las variables del proceso. En un sistema de adquisición de datos se capturan señales eléctricas asociadas con las variables físicas del proceso. Normalmente, los puertos de entrada de un sistema de adquisición de datos tienen ciertas limitaciones en cuanto al tipo de señales que pueden interpretar.

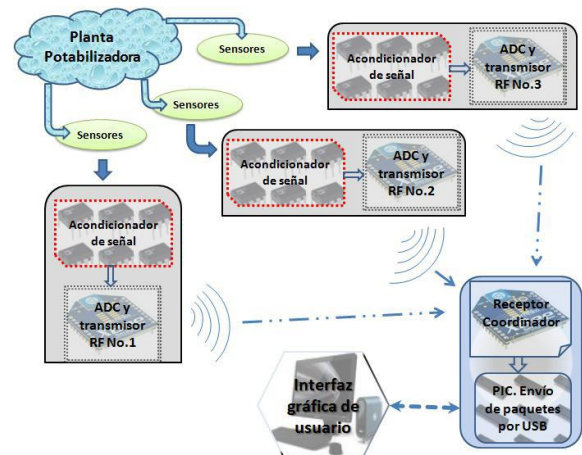


Fig. 6. Sistema inalámbrico de adquisición de datos para la planta potabilizadora

El acondicionamiento de señales es parte fundamental en el sistema de adquisición de datos. Por ejemplo, se pueden tener señales de voltaje o señales de corriente, o bien, la información puede estar contenida sobre la

amplitud de la señal o puede estar asociada a las variaciones de frecuencia. Entonces, es importante mencionar la importancia de las variaciones en amplitud o frecuencia de los datos que contienen la información. Por otra parte, existe una gran variedad de sensores que transforman una variación de una cantidad física en una señal eléctrica análoga. Es necesario que los sensores se comuniquen sin problemas con los sistemas de adquisición de datos ya que las señales de provenientes de un sensor son normalmente señales de bajo nivel no apropiadas para su captura inmediata en un SAD. La interface que resuelve los conflictos y permite la comunicación del sensor con el SAD es lo que se conoce como etapa de acondicionamiento de señal. Se tienen algunas señales más adecuadas que otra para la transmisión, por ejemplo, una señal de corriente no sufre la atenuación que una línea de transmisión le impone a una señal de voltaje; por otra parte, la información contenida en las variaciones de frecuencia de una señal viaja más segura a través de la línea mientras que la información contenida en la amplitud de otra señal representa una víctima fácil para el temible ruido aditivo. Las transformaciones de señal de un tipo a otro antes y después de la transmisión se realizan por medio de bloques de acondicionamiento de señal.

La segunda parte consiste en desarrollar el método general para el diagnóstico multivariable de la planta. Comprende las siguientes partes: el pretratamiento de los datos, el método de reconocimiento de formas (clasificador difuso) que permite generar la asociación de clases a estados funcionales del proceso (figura 7), y finalmente la etapa de reconocimiento en tiempo real (figura 8) [10,11]. En esta última etapa se incluye la participación activa del experto, tanto en la fase de aprendizaje como en la etapa de reconocimiento en línea. El operador de la planta puede observar en todo momento el estado actual del proceso en forma de estados funcionales de comportamiento en relación a los datos leídos en tiempo real. En caso de existir una falla o un funcionamiento anormal, el sistema lo indicará por medio de su sistema de alarmas.

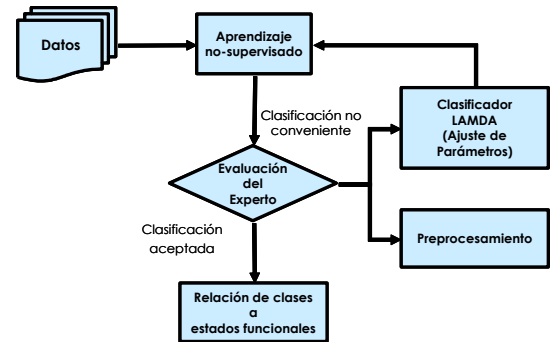


Fig. 7. Asociación de clases a estados funcionales (análisis de datos históricos-aprendizaje no-supervisado).

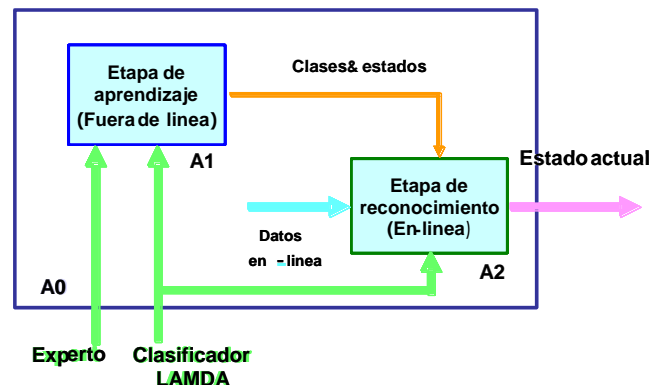


Fig. 8. Etapa de reconocimiento en línea de estados funcionales (uso del clasificador difuso).

IV. Resultados

Con la aplicación de la metodología descrita anteriormente se obtienen los siguientes resultados. Se cuenta con una base de datos de 364 mediciones de 9 variables más la dosis de coagulante aplicada –prueba de jarras-, durante un período de 12 meses (enero a diciembre de 2008). Para la generación de los estados funcionales del proceso de potabilización y de acuerdo al método propuesto en la sección anterior, se utiliza el algoritmo de clasificación difusa como estrategia general, tanto para la obtención del modelo de la planta (utilizando los datos históricos), como en el reconocimiento de los estados funcionales en tiempo real de acuerdo a la medición en línea de las variables características del agua a la entrada de la planta.

La figura 9 muestra la gráfica de resultados correspondiente a 5 clases, resultado de la aplicación del método de clasificación difuso. Estas 5 clases son asociadas a 5 estados funcionales (mostrados en la figura), con la asistencia del experto de la estación de tratamiento de agua potable.

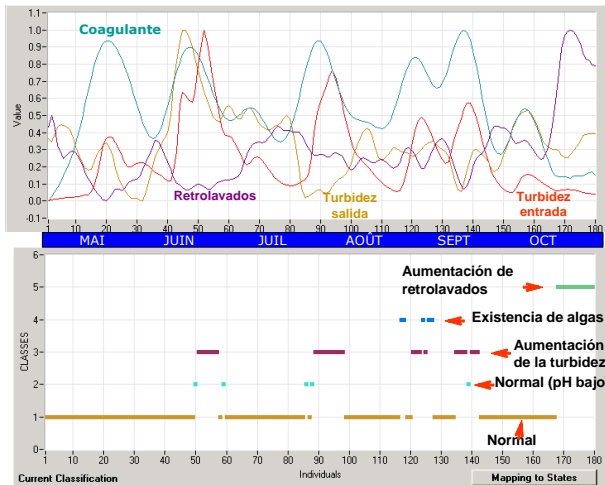


Fig. 9. Detección de 5 estados funcionales durante el período lluvias de mayo a octubre del 2008 de la planta SMAPA de la ciudad de Tuxtla, México.

V. Discusión

Los datos históricos de la planta potabilizadora resultan muy importantes para poder establecer el modelo de comportamiento de la misma. En el análisis de los estados de comportamiento se hizo participar de manera activa al experto de tal forma que los estados generados medidos en tiempo real fueran los más cercanos a los requeridos por los responsables de la planta potabilizadora.

VI. Conclusiones

Se ha presentado una herramienta de supervisión y vigilancia de la planta potabilizadora SMAPA de Tuxtla Gutiérrez. Estos trabajos son basados en la explotación de los datos del agua a la entrada de la planta, con la finalidad de realizar una vigilancia durante el proceso de potabilización. Para llevarlo a cabo, se ha propuesto utilizar el método de clasificación difusa, como base para establecer el modelo de comportamiento de la planta, de tal forma que se

adapte a las exigencias de la supervisión automática.

De manera muy general, se ha considerado que el método utilizado puede ser complementado con el uso de otros algoritmos en paralelo de tal manera que permitan el mejoramiento de los resultados (por ejemplo, la optimización de la partición de clases). Además, la formalización de un método para la determinación de la dinámica entre estados funcionales es una tarea indispensable para que la aproximación pueda evolucionar en la construcción de modelos en el dominio del diagnóstico de procesos complejos.

La relación costo-beneficio no es determinada aún con precisión, ya que la actividad inmediata a realizar es la implementación del modelo directamente en una planta potabilizadora real (datos históricos ya conocidos), en conjunción con los valores de sensores en tiempo real. Es necesario comentar que las plantas potabilizadoras del país y de Latinoamérica tienen un comportamiento similar, en cuanto a su operación y manejo de variables características.

Finalmente, en esta investigación se han tomado en cuenta los conceptos de cuidados del medio ambiente, mediante la observación de la reducción del uso de compuestos químicos (dosis de coagulante) en la etapa de clarificación del agua.

VII. Referencias

- [1] L. Trabé-Massuyès, P. Dague, F. Guerrin, *Le raisonnement qualitatif*: Hermes, France, 1997.
- [2] J. Aguilar-Martin, "Knowledge-based real time supervision. Tempos-Modify workshop. Budapest, Hongrie. 1996.
- [3] M. Dojat, N. Ramaux, D. Fontaine. Scenario recognition for temporal reasoning in medical domains. *Artificial Intelligence in Medicine*, tome 14, pp 139-155, 1998.
- [4] R. Iserman. Supervisión, fault detection and fault diagnosis methods. *Control Eng practice*, vol 5, No. 5, pp 639-652, Elsevier, France, 1997.

[5] J. Colomer, J. Meléndez, J. Ayza. Sistemas de supervisión: introducción a la monitorización y supervisión experta de procesos: métodos y herramientas. Barcelona, Petisa, 2000.

[6] S. Dash, V. Venkatasubramanian. Challenges in the industrial applications of fault diagnostic systems. Computers and chemical eng 24, pp 785-791, Elsevier, France, 1997.

[7] SMAPA, Sistema Municipal de Agua Potable y Alcantarillado de Tuxtla. Manual de procedimientos, Tuxtla, Gutiérrez, Chiapas, México, 2010.

[8] C. Lind. Coagulation Control and Optimization: Part One, Public Works, Oct 1994, 56-57

[9] C. Lind. Coagulation Control and Optimization: Part Two. Public Works, Nov 1994, 32-33

[10] H. Hernández De León: Supervision et diagnostic des procédés de production d'eau potable. Thèse de doctorat, INSA Toulouse, France/CNRS (2006).

[11] Kempowsky, T.: SALSA(Situation Assessment using LAMDA Classification Algorithm). User's Manual. Rapport LAAS/CNRS No. 04160 (2010).

Dr. Madain Pérez Patricio es Ingeniero en electrónica por el Instituto Tecnológico de Tuxtla Gtz., Chiapas y Doctor en Ciencias por la Universidad de Lille, Francia.

VIII. Autores

Dr. Héctor Ricardo Hernández De León es Ingeniero en Electrónica del Instituto Politécnico Nacional en 1982. Maestro en Ciencias en Ingeniería Electrónica del CENIDET, en 1990. Doctor en Ciencias en Sistemas Automáticos del Instituto Nacional de Ciencias Aplicadas de Toulouse, Francia, del 2002 al 2006. Su línea de investigación es el control y automatización de procesos. Actualmente es profesor investigador del posgrado en Mecatrónica y líder del cuerpo académico "Automatización de Procesos".

Dr. Jorge Camas Anzueto es Ingeniero en electrónica por el Instituto Tecnológico de Tuxtla Gtz., Chiapas y Doctor en Ciencias por el INAOE de Puebla.

Sistema Concentrador de Comunicaciones X-10

***José Miguel Torres Garfias, ** José Antonio Moreno Espinosa**

Universidad Tecnológica de la Mixteca, Carretera a Acatlima km 2.5, Huajuapán de León, Oaxaca, México.

*jmtg139@hotmail.com. **jamoreno@mixteco.utm.mx

Resumen

El acceso a las tecnologías PLC permite la creación de redes de comunicación utilizando el cableado eléctrico doméstico. El protocolo X-10 permite la transmisión y recepción de mensajes de esta manera, y con ello lograr un sistema domótico.

En este documento se describe el diseño y construcción de un “Sistema Concentrador de Comunicaciones X-10”, el cual cuenta con hasta 10 direcciones distintas y puede controlar igual número de dispositivos. Para la verificación del correcto funcionamiento de este sistema, se utilizaron módulos X-10 con direcciones individuales, con lo que se podía verificar que el sistema tuviera el mismo funcionamiento que los módulos.

La etapa de control de este módulo se logró utilizando un dsPIC30F4013, el software para el manejo del sistema y de los módulos individuales fue realizado en el lenguaje Microsoft Visual C# 2008.

Palabras Clave: Concentrador, dsPIC30F4013, PLC, X-10.

I. Introducción

X-10 (o también referenciado como X10) es uno de los primeros protocolos usados para sistemas domóticos y es el más extendido tanto en Europa como en Estados Unidos.

X-10 implementa la tecnología de corrientes portadoras *-carriers-* y utiliza como medio de transmisión la red eléctrica [1]. Por lo que X-10 no requiere de cableado adicional al eléctrico y los sistemas que aplican este protocolo pueden ser utilizados tanto en viviendas ya construidas como en las de nueva construcción.

Dado que se tiene el conocimiento suficiente del protocolo X-10, y se dispone de herramientas (como los PICs, dsPICs y la PC) que permitan el desarrollo de sistemas basados en la funcionalidad del protocolo, se desarrollaron módulos para el control de aparatos conectados a la red eléctrica doméstica y depuración de las comunicaciones X-10. Cada uno de estos módulos (receptores) posee una única dirección y permite el control de un solo aparato.

El manejo de estos módulos de control se logra con un módulo interfaz (transmisor) que recibe de una PC los comandos y direcciones del aparato a controlar. El transmisor se encarga de convertir los datos recibidos de la PC al formato X-10 e inyectarlo a la red eléctrica para lograr la comunicación con el módulo receptor X-10 al que se envió el comando.

Cuando se requiere del control de dispositivos que se hallen juntos, o que su funcionamiento tenga relación entre cada uno de ellos, o que formen parte de una red de control, sería ideal contar con un único módulo que permita el control de todos ellos, es decir un concentrador que aglomere las comunicaciones relacionadas con esos dispositivos.

A lo largo de este documento se describe principalmente el diseño, construcción y pruebas de un “Sistema Concentrador de Comunicaciones X-10”. Igualmente se tiene una descripción del funcionamiento del protocolo X-10.

II. El protocolo X-10

Como cualquier otro protocolo, el protocolo X-10 posee formas de asignar direcciones a sus elementos, de sincronizar las comunicaciones, de representar la información digital y de mecanismos de operación para lograr una exitosa transmisión. A continuación se describen estos puntos de manera general.

Direccionamiento de dispositivos

El protocolo X-10 soporta 16 grupos de direcciones denominados *House code* o códigos de casa, (A-P), y otras 16 direcciones por cada código de casa denominadas *Unit code* o códigos de unidad (1-16). El protocolo X-10 permite de esta manera manejar hasta 256 direcciones distintas. La comunicación se realiza mediante cadenas de control que son sucesiones de unos y ceros que completan los comandos [2]. Las direcciones, que están compuestas del *House code* y el *Unit code* son definidas por el usuario.

Tipos de dispositivos

Existen cuatro tipos de dispositivos:

Transmisores X-10: también conocidos como controladores, envían la señal X-10 a los módulos a controlar inyectándola a la red eléctrica.

Receptores X-10: reciben señales X-10 de la red eléctrica doméstica que fueron enviadas por algún transmisor y realizan la operación contenida en el mensaje X-10. Tienen una dirección definida por el usuario (compuesta de *House code* y *Unit code*). Múltiples dispositivos pueden poseer un mismo código.

Bidireccionales X-10: tienen la capacidad de responder y confirmar la realización de un comando.

Inalámbricos: utilizan señales de radio para aumentar el rango de cobertura de las comunicaciones X-10.

Sincronización de señales

Las transmisiones y recepciones X-10 son sincronizadas con el cruce por cero de la línea de AC [3]. De esta manera los transmisores saben cuándo transmitir y los receptores cuándo recibir un mensaje. Cada ciclo de la señal de 60 Hz de AC tiene dos cruces por ceros, uno en el semiciclo positivo y otro en el semiciclo negativo.

Representación de la información digital

Los datos son aceptados bit por bit, y se representan como la presencia o ausencia de una señal portadora de 120 kHz (*carrier*) que ocurre después del cruce por cero del ciclo positivo o negativo. Un '1' binario es representado como la presencia del *carrier* de 120 kHz y un '0' binario se representa por la ausencia de ese *carrier*. La duración del *carrier* es de 1ms +/- 50µs y se toleran retardos de hasta 200µs [3]. En la figura 1 se muestra cómo el *carrier* se transmite en un cruce por cero.

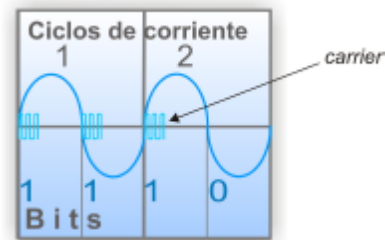


Fig. 1. Transmisión del *carrier* en un cruce por cero.

Mensaje X-10 básico

Un mensaje básico se compone de tres códigos: *Start code* de 4 bits, *House code* de 4 bits y *Key code* de 5 bits. El *Key code* a su vez puede ser un *Unit code* (cuando se transmite una dirección) o un *Function code* (cuando se transmite un comando) [4].

La tabla 1 muestra las direcciones y comandos del protocolo X-10, en ella se puede ver que los *Function code* tienen sólo cuatro bits y no cinco como se acaba de mencionar, pero cuando se envía un *Function code* simplemente se agrega un quinto bit con valor '1'.

Configuración	H1	H2	H4	H8
A	0	1	1	0
B	1	1	1	0
C	0	0	1	0
D	1	0	1	0
E	0	0	0	1
F	1	0	0	1
G	0	1	0	1
H	1	1	0	1
I	0	1	1	1
J	1	1	1	1
K	0	0	1	1
L	1	0	1	1
M	0	0	0	0
N	1	0	0	0
O	0	1	0	0
P	1	1	0	0
All Units Off	0	0	0	1
All Units On	0	0	0	1
On	0	0	1	1
Off	0	0	1	1
Dim	0	1	0	1
Bright	0	1	0	1
All Lights Off	0	1	1	1
Extended code	0	1	1	1
Hail Request	1	0	0	1
Hail Acknowledge	1	0	0	0
Pre-set Dim	1	0	1	1
Extended Data (analog)	1	1	0	1
Status=On	1	1	0	1
Status=Off	1	1	1	1
Status Request	1	1	1	1

Tabla 1. Direcciones y comandos X-10.

Transmisión completa de un mensaje X-10

A excepción de los bits del *Start code*, los bits del *House code* y del *Key code* se transmiten en forma de complemento real en ciclos alternos de corriente. Debido al medio de transmisión utilizado, cada bloque de información se transmite dos veces para que el sistema aumente su fiabilidad. Entre cada par de bloques de información *Start code*–*House code*–*Key code* se espera tres ciclos de corriente de AC completos (seis cruces por cero) donde no se envía nada. Para esta regla existe una excepción (comandos *Dim* y *Bright*). En la primera transmisión del par de bloques completo *Start code*–*House code*–*Key code*, en el *Key code* se envía la dirección *Unit code*, y en la segunda transmisión se envía el *Function code* como comando [4]. La

figura 2 muestra cómo se realiza una transmisión completa X-10.

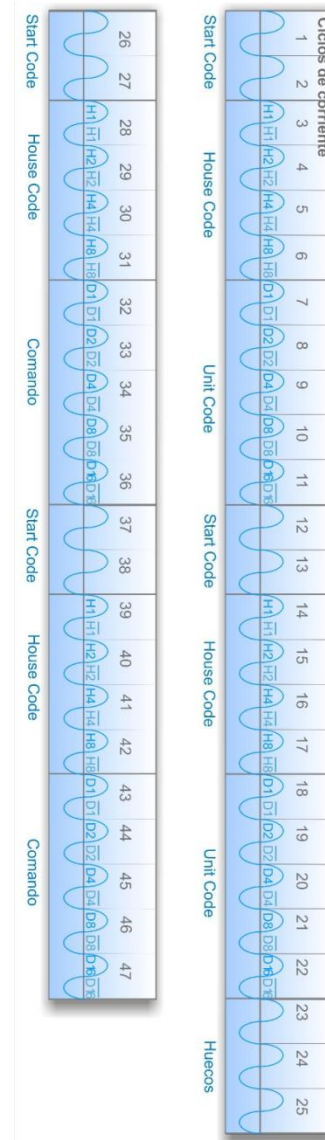


Fig. 2. Transmisión completa de un mensaje X-10.

III. Implementación de X-10 en sistemas de diseño propio

Si desea llevarse a cabo la implementación de un sistema de comunicación basado en el protocolo X-10, es posible hacerlo al obtener dispositivos comerciales X-10, tanto transmisores como receptores. Estos dispositivos son de tipo *plug-in*, basta con conectarlos para comenzar a usarlos. Sin embargo los manuales de estos dispositivos

no proveen la información necesaria si se desea un sistema por diseño propio.

En 1997 la patente de X-10 venció, por lo que se convirtió en un protocolo estándar de facto abierto [5]. Desde entonces, numerosas compañías han desarrollado sistemas y dispositivos basados en X-10, y se han vuelto empresas fabricantes de este tipo de aparatos.

Pero si desea implementarse un sistema de automatización con características especiales o distintas a las provistas por productos comerciales, este debe ser diseñado por un desarrollador, como en este caso. A raíz de la posibilidad de desarrollar sistemas de este tipo han sido publicados artículos y documentos que exponen ideas y/o técnicas para desarrollar sistemas de automatización basados en X-10 [6].

La nota de aplicación AN236 [7] de Microchip discute la implementación de X-10 en un PICmicro MCU para crear un “Controlador de Hogar” que pueda enviar y recibir señales X-10. Algunas ideas para el desarrollo de los módulos individuales X-10 y del “Sistema Concentrador de Comunicaciones X-10” surgieron a partir de esta nota de aplicación.

IV. Diseño y construcción de módulos individuales X-10

Para la creación del sistema de comunicación X-10 se requiere de la parte de *hardware* y la parte de *software*. La integración del *hardware* y el *software* en un sistema único funciona de la siguiente manera:

El *software*

Se desarrolló un *software* que se ejecuta en una PC, el cual fue diseñado bajo lenguaje Visual C#. El *software* envía información digital que contiene una dirección y un comando (seleccionados por el usuario) hacia un módulo transmisor. Este *software* cuenta además con varias herramientas, como asignar nombres a

direcciones y programar tareas. La ventana del principal del *software* se muestra en la figura 3.



Fig. 3. Ventana principal del *software* para el control de dispositivos X-10.

El transmisor X-10

El transmisor convierte la señal digital recibida de la PC desde el puerto serial a una señal analógica (compuesta por *carriers*) y la inyecta a la red eléctrica. La figura 4 muestra el diagrama a bloques del transmisor X-10 desarrollado.

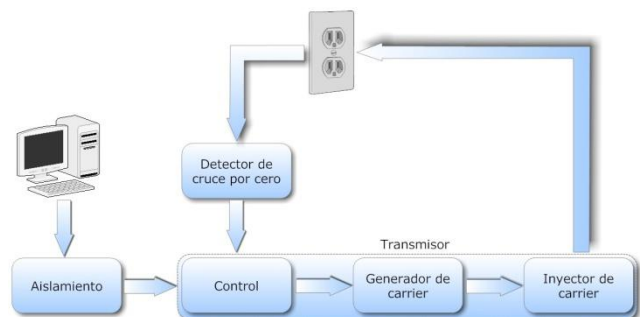


Fig. 4. Diagrama a bloques del Transmisor X-10.

El receptor X-10

El receptor posee una dirección única y obtendrá la información de la red eléctrica y si la dirección corresponde a la suya, se encargará de que una carga conectada a él ejecute el comando recibido. El diagrama a bloques del receptor X-10 es mostrado en la figura 5.

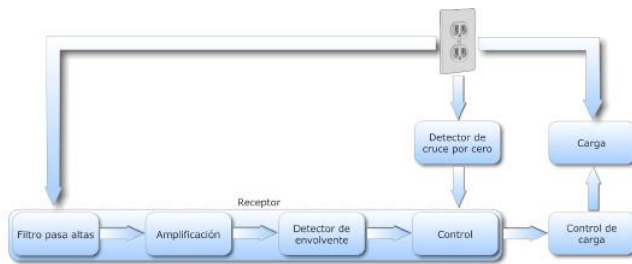


Fig. 5. Diagrama a bloques del receptor X-10.

Construcción de los módulos

Las etapas de control del módulo Transmisor y Receptor X-10 se obtuvieron con un PIC16F84. El diseño del detector de cruce por cero se basó en la nota de aplicación AN251 [8] de Microchip, aprovechando la estructura interna de los PICmicro MCU. Las demás etapas se obtuvieron con circuitos analógicos, pero dado el fin del documento no son descritas. Los módulos Transmisor y Receptor X-10 terminados se muestran en la figura 6 y 7 respectivamente.



Fig. 6. Módulo Transmisor X-10 terminado



Fig. 7. Módulo Receptor X-10 terminado.

V. Sistema Concentrador de Comunicaciones X-10

El receptor X-10 que ya ha sido diseñado permite que le sea asignada una única dirección. Sin embargo si se desea conectar equipos juntos y diferenciables a la red de comunicaciones X-10 es necesario contar con un módulo que concentre las comunicaciones X-10 de esos equipos.

Este módulo ofrecería entonces la capacidad de depurar la red de los dispositivos X-10 conectados al cableado eléctrico doméstico. El uso de este módulo permitiría también reducir la cantidad de módulos Receptores X-10. El módulo concentrador disminuiría los errores en la recepción de los datos, dado que en caso de que los módulos controlados fueran bidireccionales, se tendría un solo módulo contestando los mensajes.

Por lo tanto se diseñó un “Sistema Concentrador de Comunicaciones X-10” al que se le pueden asignar hasta 10 direcciones distintas e igual número de dispositivos. Este nuevo módulo permite además visualizar en un LCD los datos X-10 extraídos de la red eléctrica y muestra los estados de los dispositivos controlados, con lo que se puede verificar si los mensajes enviados por el Transmisor X-10 fueron recibidos y de qué comandos y direcciones se trataron.

La etapa de control del “Sistema Concentrador de Comunicaciones X-10” se logró utilizando un dsPIC30F4013 del fabricante Microchip. El código del dsPIC30F4013 fue realizado en el entorno de desarrollo “MikroC compiler for dsPIC30/33 and PIC 24” bajo lenguaje C.

Un dsPIC cuenta con los recursos típicos de un PIC, además de soportar las funciones clásicas de los DSP[9]. El dsPIC30F4013 cuenta con dos módulos UART, que fueron utilizados para la comunicación serial con la PC. El dsPIC30F4013 se encarga también de mostrar los datos recibidos en un LCD. La figura 8 muestra el encapsulado del dsPIC30F4013.

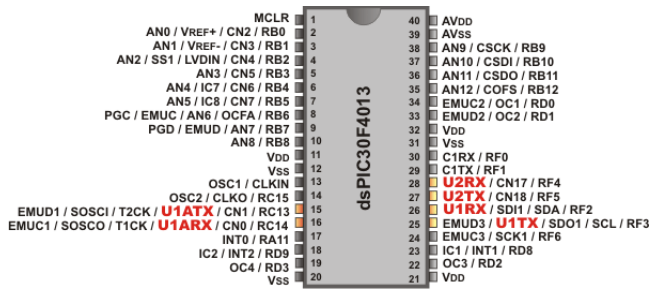


Fig. 8. Encapsulado PDIP del dsPIC30F4013

Las etapas detector de cruce por cero y detector de *carrier* fueron implementadas de igual forma que en el Receptor X-10 ya diseñado. Las 10 direcciones asignadas al “Sistema Concentrador de Comunicaciones X-10” son: B9, C3, E5, G14, I2, H13, J11, L15, M8 y O6. Cuando todos los dispositivos conectados al “Sistema Concentrador de Comunicaciones X-10” han sido encendidos y se desea que ahora todos sean apagados habría que mandar el comando OFF a cada uno de ellos, para evitar esto se dotó a este módulo de un *switch* tipo *push-button* para dar un *reset* (reiniciar) al estado de los dispositivos. Por lo tanto el *reset* lleva los dispositivos con estado de encendido al estado de apagado. El “Sistema Concentrador de Comunicaciones X-10” terminado es mostrado en la figura 9.

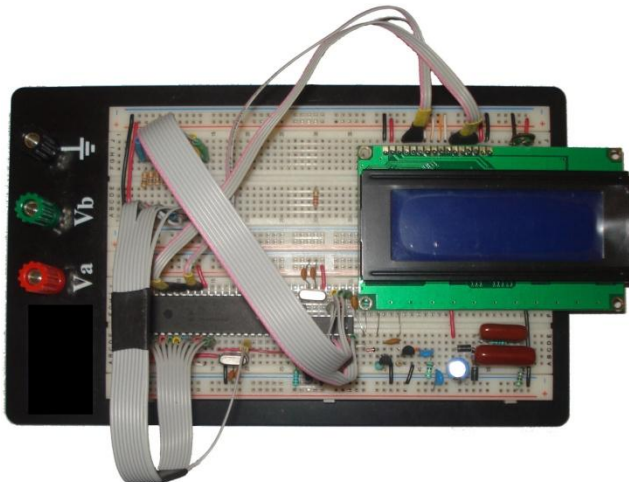


Fig. 9. Sistema Concentrador de Comunicaciones X-10.

El transmisor X-10 ya diseñado se comunica con el puerto serie de la PC para recibir los comandos

y direcciones, convierte la información digital a *carriers* y los inyecta a la red de la línea eléctrica. Sin embargo, si la transmisión de datos X-10 falló, una posible causa es un error en la comunicación serial al recibir los datos de la PC. Por lo que se diseñó un “Monitor X-10 para la transmisión serial”, que además de tener la funcionalidad del transmisor X-10 ya diseñado muestra en un LCD los datos recibidos de la PC para monitorear la comunicación serial. Para la etapa de control del “Monitor X-10 para la transmisión serial” se utilizó también un dsPIC30F4013.

El “Monitor X-10 para la transmisión serial” complementa la funcionalidad del “Sistema Concentrador de Comunicaciones X-10”, permitiendo no solo la verificación de los datos recibidos, sino también de los transmitidos. El “Monitor X-10 para la transmisión serial” es mostrado en la figura 10.

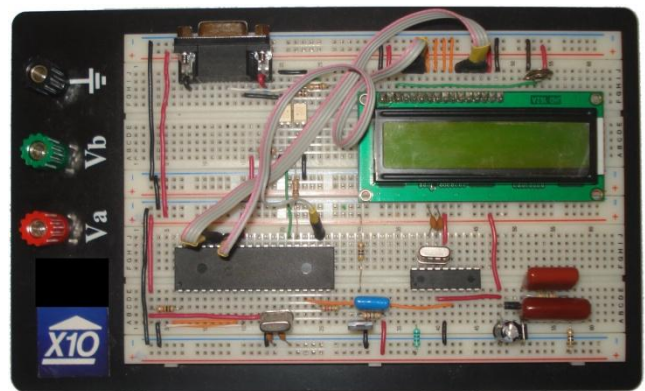


Fig. 10. Monitor X-10 para la transmisión serial.

VI. Pruebas y Resultados

Mediciones

Las etapas principales en un sistema X-10 son la Detección del Cruce por Cero, la Generación de *carrier* y Detección de *carrier*. Para comprobar que la circuitería de estas etapas funciona correctamente se realizaron las mediciones de las señales correspondientes.

La señal de cruce por cero generada por los módulos se muestra en la figura 11, donde se puede observar que su frecuencia es de casi 120Hz.

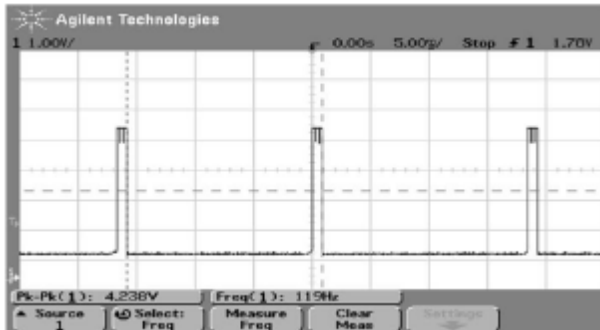


Fig. 11. Señal de cruce por cero.

Para obtener el *carrier* de 120 kHz, primero el transmisor genera una señal cuadrada de igual frecuencia, la cual es mostrada en la figura 12.

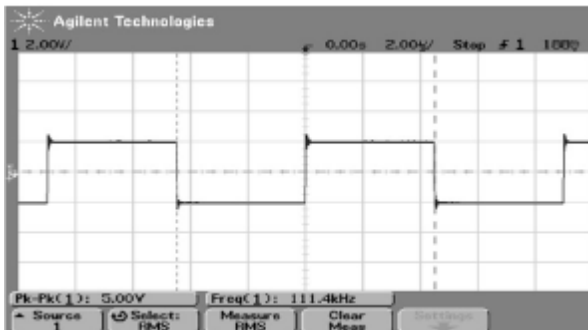


Fig. 12. Señal cuadrada de 120 kHz .

Después esa señal pasa por otra etapa para ser inyectada a la línea de AC. La señal de *carrier* montada en la línea puede verse en la figura 13.

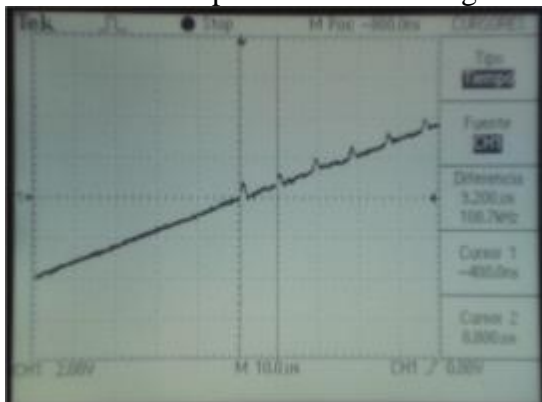


Fig. 13. Señal de *carrier* montada en la de AC.

El *carrier* después del filtro en el receptor se muestra en la figura 14, donde se puede apreciar que la duración de la señal es de 1ms.

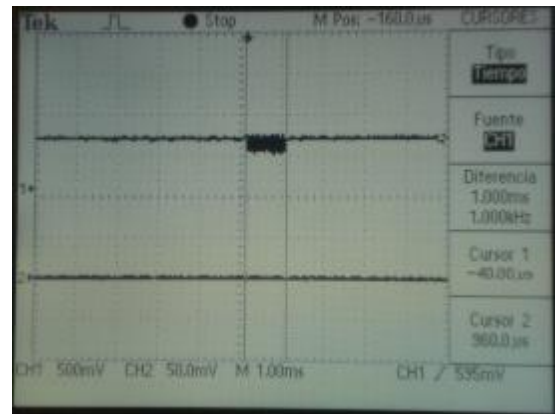


Fig. 14. *Carrier* después del filtro.

Pruebas de funcionamiento

Las primeras pruebas consistieron en evaluar el funcionamiento de los módulos X-10 individuales, para luego hacer una comparación con el “Sistema Concentrador de Comunicaciones X-10.” Desde el *software* creado se envió el comando ON a la dirección A1, como resultado se obtuvo que un foco que estaba conectado al módulo Receptor X-10 con dirección A1 se encendiera estando previamente apagado (ver Fig. 15).

Habiendo comprobado el funcionamiento de los módulos individuales se continuó con el del “Sistema Concentrador de Comunicaciones X-10”, para ello se envió el comando ON a la dirección B9. El módulo recibió correctamente el mensaje y este fue mostrado en el LCD

Una vez probado el “Sistema Concentrador de Comunicaciones X-10”, se continuó a enviar comandos a las 10 direcciones, alternando el orden de los comandos y direcciones para comprobar que todo el sistema completo funcionara de manera correcta y que en efecto el sistema concentrara las comunicaciones X-10 de las direcciones que posee. El sistema respondió de forma exitosa a todos los comandos enviados

a sus 10 direcciones, por lo que muestra igual funcionamiento que los módulos individuales, con la diferencia que se tiene un solo módulo receptor pero 10 direcciones distintas e igual número de dispositivos.

Con el fin de monitorear los datos que se reciben de la PC y que se inyectan a la red eléctrica, se utilizó el “Monitor X-10 para la transmisión serial”, cuando se envió el mensaje ON a la dirección B9, el mensaje fue desplegado en el LCD del módulo

VII. Conclusiones

Se creó un sistema controlado por la PC mediante un software hecho bajo Visual C# que envía órdenes de un módulo transmisor hacia módulos receptores X-10 (módulos individuales).

Las redes PLC basadas en el protocolo X-10 resultan atractivas para crear soluciones rápidas, baratas y sencillas para crear sistemas de automatización de dispositivos.

El “Sistema Concentrador de Comunicaciones X-10” ofrece la ventaja de reducir *hardware* para el control y automatización de dispositivos electrónicos.

El “Sistema Concentrador de Comunicaciones X-10” permite también depurar sistemas X-10 y evaluar la calidad de las comunicaciones por el cableado eléctrico doméstico.

Gracias a este último módulo es posible conectar equipos que por su naturaleza se encuentren juntos y que sean diferenciables.

Disminuir la cantidad de *hardware* necesario al utilizar el “Sistema Concentrador de Comunicaciones X-10” conlleva el beneficio de tener una red de control ordenada y además ahorrar recursos.

El módulo “Monitor X-10 para la transmisión serial” complementa el funcionamiento del “Sistema Concentrador de Comunicaciones X-10”, verificando que los datos recibidos de la PC

e inyectados a la red eléctrica sean los que se reciben por el concentrador.

VIII. Referencias

- [1] G. O’ Driscoll, “The Essential Guide to Home Networking Technologies”, Prentice Hall PTR, 2001.
- [2] C. Romero, F. Vásquez, C. de Castro, “Domótica e Inmótica. Viviendas y edificios inteligentes”, Alfaomega-RaMa, 2008.
- [3] D. Rye, “The X-10 POWERHOUSE Power Line Interface Model #PL513 and Two-Way Power Line Interface Model #TW523”, Revision 2.4, X-10 (USA) Inc.
- [4] X10 Ltd., “X10; "Standard" and "Extended" X-10 Code Formats”; September 22, 1993. Rev. 11/20/94, 07/12/96, 12/18/96, 04/02/97, 07/22/97, 05/07/98”, 1993.
- [5] Technical Pacifica, “Easy X10 Projects for Creating a Smart Home”, Sams Technical Publishing, LLC, 2005.
- [6] J.M. Torres Garfias, “Sistema de comunicación usando como soporte físico la red eléctrica mediante el protocolo X10”, 2011, Universidad Tecnológica de la Mixteca.
- [7] J. Burroughs, “X-10® Home Automation Using the PIC16F877A”, Microchip Technology Inc., 2002.
- [8] D. Cox, “Interfacing to AC Power Lines”, Microchip Technology Inc., 1997.
- [9] J. Ma. Angulo, A. Extebarría, I. Angulo, “dsPIC Diseño Práctico de aplicaciones”, McGraw Hill, 2008.

IX. Autores

José Miguel Torres Garfias obtuvo su título de Ingeniero en Electrónica en la Universidad Tecnológica de la Mixteca (UTM) en Huajuapán de León, Oaxaca, México.

José Antonio Moreno Espinosa obtuvo el título de Ingeniero en Electrónica por la Universidad Autónoma Metropolitana y el título de Master en Sistemas y Redes de Comunicaciones por la Escuela Superior de Ingenieros de Telecomunicación de la Universidad Politécnica de Madrid. Actualmente es profesor adscrito al Instituto de Electrónica y Mecatrónica en la Universidad Tecnológica de la Mixteca, Huajuapán de León, Oaxaca, México.

Implementación de una Red Neuronal Artificial en un FPGA

*Juan José Martínez Nolasco, **Francisco Gutiérrez Vera, ***J. Guadalupe Zavala Villalpando,
****Víctor Cacique Borrego, *****Alonso Jiménez Garibay

*Instituto Tecnológico de Celaya, juanjose@itc.mx.
** Instituto Tecnológico de Celaya, pancho@itc.mx.
*** Instituto Tecnológico de Celaya, jgzavala@itc.mx.
**** Instituto Tecnológico de Celaya, victor@itc.mx.
***** Instituto Tecnológico de Celaya, alonso@itc.mx.

Resumen

En el presente trabajo se describe una propuesta de diseño de una red neuronal artificial (RNA) implementada en un dispositivo lógico programable (PLD). El entrenamiento de la RNA se realizó utilizando el toolbox de Matlab. Antes de implementar la RNA en el PLD se verificó su correcto funcionamiento implementando un instrumento virtual en LabVIEW.

El lenguaje empleado para la descripción de la RNA implementada en hardware es el VHDL y los sistemas de desarrollo utilizados para realizar la implementación de los controladores son de Altera y el software Quartus II es el utilizado para realizar la descripción, simulación, síntesis e implementación del modelo.

También se presentan los resultados del control neuronal embebido en línea con el proceso. La propuesta se desarrolló en base a las variables disponibles en el módulo didáctico *armfield PCT-9*. Además, se presenta una forma de implementar un control neuronal interno utilizando dos FPGA y sincronizando el funcionamiento de los mismos obteniendo resultados favorables.

Palabras Clave: Control, Labview, FPGA, Matlab y Quartus II.

I. Introducción

Existen sistemas cuyo análisis por métodos convencionales es poco eficiente, ya que estos sistemas presentan características como incertidumbre y alinealidades. Cuando la complejidad de un proceso aumenta puede ser muy difícil encontrar un modelo matemático basado en leyes físicas y esto puede dificultar el diseño de un controlador con métodos clásicos. Otra opción es el uso de control inteligente. En general, es posible identificar la conveniencia de estrategias de control inteligente cuando los parámetros de la planta pueden cambiar significativamente, los controladores tradicionales deben ser constantemente ajustados,

es difícil modelar matemáticamente el sistema, la inexactitud del sensor introduce niveles de incertidumbre no aceptables, se requiere un procesamiento adaptativo de señal para cubrir cambios en el ambiente o condiciones del proceso.

Las RNA se han establecido como una metodología para el modelado y control de sistemas no lineales, permitiendo la solución de algunos problemas en ingeniería. En particular, el uso de RNA recurrentes para modelado y aprendizaje se ha incrementado rápidamente en años recientes. Para proveer una evaluación racional de los nuevos métodos es importante compararlos y contrastarlos con las metodologías

tradicionales. Algunas de las características relevantes de las RNA en este contexto son: su capacidad para aproximar transformaciones no lineales, alto grado de paralelismo; gran robustez a fallas de componentes, pueden ser adaptadas en línea y en tiempo real, pueden procesar simultáneamente múltiples entradas y salidas.

Otra técnica vanguardista es el manejo de dispositivos lógicos programables aplicables al control de procesos industriales. La combinación de ambas técnicas nos permite el diseño de sistemas de automatización empleando redes neuronales.

Hasta estos momentos se han realizado diversos trabajos de control neuronal utilizando diversos software y también se han utilizado una infinidad de técnicas de implementación de RNA en FPGA [1 al 8]. En este artículo se presenta una técnica muy sencilla de implementar una RNA en un FPGA y su aplicación en el control de un proceso de nivel de agua. En la implementación del control neuronal interno se utilizan dos FPGA con una RNA cada uno.

II. Control Neuronal

Desde el punto de vista de la teoría de control, la habilidad de las RNA para manejar sistemas no lineales es significativa. El Control Neuronal es una de las alternativas que ofrece el control inteligente para resolver el problema de controlar un sistema con niveles de “inteligencia”, inherentes a las propiedades de las RNA, que emulan ciertas características propias del cerebro humano.

II.1. Control Directo Por Modelo Inverso

El control por modelo inverso fue uno de los primeros métodos propuestos y consiste en entrenar una RNA que actúa como la función de transferencia inversa del sistema y se utiliza como un controlador. El modelo inverso se coloca en cascada con el sistema controlado, como se ilustra en la figura 1, de tal manera que

resulte la identidad entre la respuesta deseada y la salida del sistema controlado.

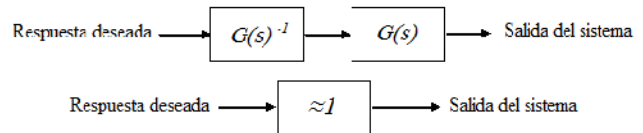


Fig. 1.- Control por modelo inverso.

II.2. Control Por Modelo Interno

El control por modelo interno está muy relacionado con el control inverso. Un controlador por modelo interno requiere tanto de un modelo directo como de un modelo inverso del sistema a controlar. Su esquema de control es mostrado en la figura 2 [9].

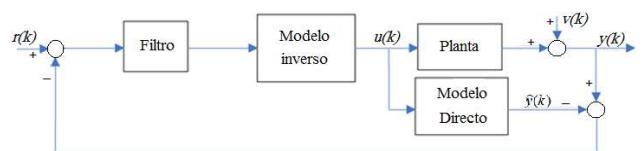


Fig. 2.- Esquema del control por modelo interno.

III. Desarrollo

III.1. Descripción Del Módulo ARMFIELD

En la figura 3 se ilustra el Proceso de Nivel utilizado para realizar todos los experimentos.

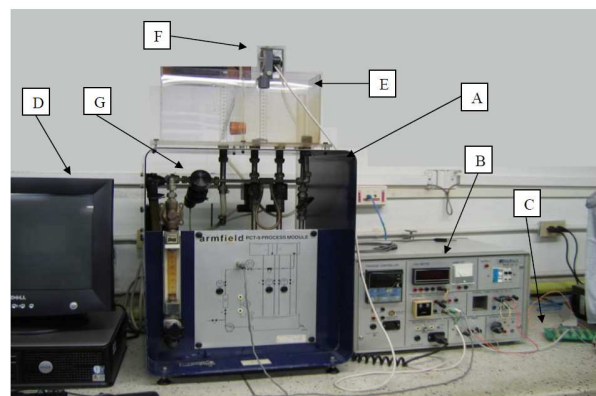


Fig. 3.- Sistema de desarrollo.

La variable controlada es el nivel de agua que se deposita dentro del tanque de almacenamiento mostrado en “E” de la figura 3. El rango de operación del tanque es de 0 a 2843 cm^3 . La

medición del nivel de agua en el tanque de almacenamiento se realiza con el sensor ilustrado en “F” de la figura 3. El sensor cuenta con: un flotador que varía su posición dependiendo de la variable del proceso, una cuerda que en un extremo tiene sujeto el flotador y del otro extremo tiene la referencia de movimiento, la polea que gira con el movimiento de la cuerda y un potenciómetro cuya perilla es el eje de la polea. Para mayor información del funcionamiento del sensor revisar [10].

La variación de resistencia del sensor es recibida por el módulo acondicionador de señales. El módulo acondicionador de señales mostrado en “B” de la figura 3 realiza la conversión para tener una señal de salida de voltaje.

III.2. Kit de Desarrollo.

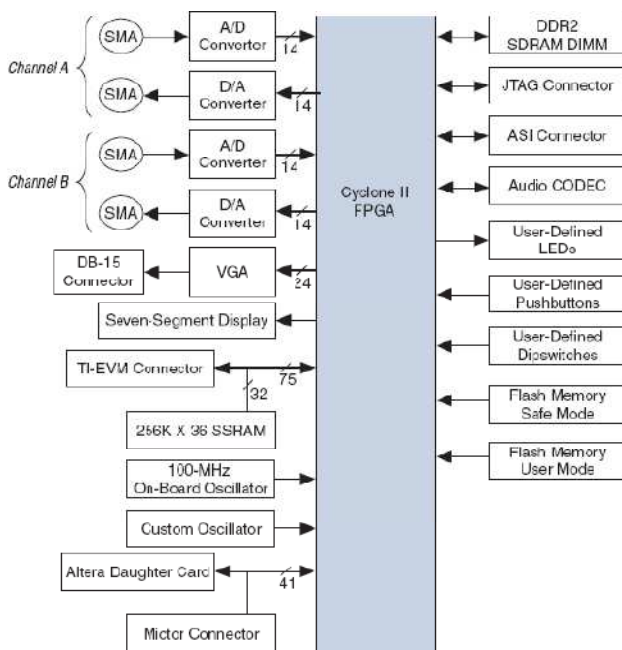


Fig. 4.- Diagrama del sistema de desarrollo.

El sistema de desarrollo utilizado para la implementación de la RNA del sistema inverso fue el sistema de desarrollo *DSP Development Kit, Cyclone II Edition* desarrollado por Altera y contiene el FPGA *Cyclone II EP2C70F672C6*. En la figura 4 se ilustra el diagrama funcional del sistema de desarrollo. La

descripción de los componentes del sistema de desarrollo se pueden observar en [11].

El sistema de desarrollo donde se implementó el modelo directo del sistema para realizar el control por modelo interno fue el *TREX-S2-60-4* de Altera.

III.3. Modelo Inverso.

La ecuación 1 modela el vector de regresión de la RNA del modelo inverso, la señal $u(k)$ es la acción de control que realimenta a las entradas de la RNA, la señal $r(k)$ es la referencia del controlador y la señal $y(k)$ representa la salida del proceso.

$$\begin{aligned} u(k) &= [u(k-1), u(k-2), u(k-3), u(k-4), r(k), \\ y(k), y(k-1), y(k-2), y(k-3)]. \end{aligned} \quad (1)$$

La RNA empleada para implementar el modelo inverso del sistema es un MLP. Para el aprendizaje se utilizó el algoritmo de entrenamiento de *Levenberg-Marquart* ya que con el algoritmo bayesiano la RNA no converge a un resultado eficaz. En la estructura de la RNA se tiene una capa de entrada con 9 neuronas, cuatro capas ocultas de 5 neuronas cada una de ellas y cada neurona con función de activación sigmoideal y la capa de salida con una neurona con función de activación lineal.

III.4. Modelo Interno.

Para corregir la desventaja del control neuronal por modelo inverso ante las perturbaciones se implementó el control neuronal por modelo interno. Para implementar el modelo interno se necesita el modelo directo del sistema y su estructura cuenta con una capa de entrada de 10 elementos, una capa oculta con diez neuronas y cada una de ellas con función de activación sigmoideal y una capa de salida con una neurona con función de activación lineal.

III.5. Modelo Inverso en FPGA.

En la figura 5 se ilustra la estructura global del controlador. Este cuenta con los siguientes bloques:

Estructura externa del controlador. Trabaja con la señal de control, referencia y nivel, estas señales trabajan con un rango analógico de [0, 1] Volts y se representa con una precisión de 10 bits cada una.

Contador. Se encarga de sincronizar las señales de entrada de la RNA y las conexiones entre cada capa oculta. Tiene como entrada la señal de reloj proveniente del oscilador del FPGA con una frecuencia de 14.74 MHz y a su salida cuenta con dos señales de reloj una de 1.67 Hz (señal que le proporciona el periodo de muestreo de 1.67 muestras/segundo al controlador) y la otra de 1.85 KHz (señal que va a los registros de las funciones de activación de cada neurona).

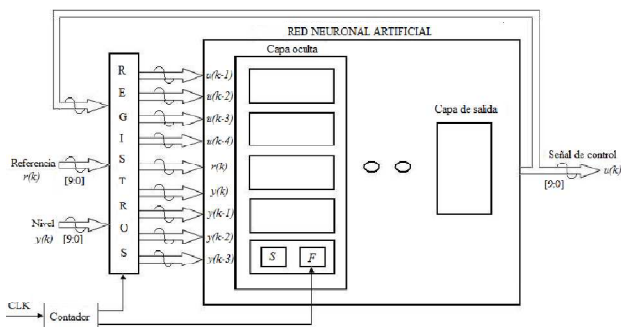


Fig. 5.- Diagrama del control por modelo inverso.

Registros. Se encargan de proporcionar los retardos correspondientes a cada señal y de sincronizar la salida de cada una de estas. La sincronización de los registros se realiza con la señal de reloj con frecuencia de 1.67 Hz proveniente del contador.

La RNA. Cada neurona está dividida en dos bloques el S y el F. El bloque S sirve para obtener la Neta de cada neurona, esta se obtiene por medio de la ecuación 2 en donde S_n es la Neta de la neurona n, w_{nj} son los pesos, b_n es el umbral y x_j es la entrada de cada neurona.

$$S_n = \sum_{j=1}^m w_{nj} + b_n \tag{2}$$

Los valores de los pesos obtenidos con Matlab son del tipo double y se encuentran en el rango de [-10, 10]. Se realizó un escalamiento en Matlab aplicando la ecuación 3. Los pesos obtenidos son del tipo punto fijo y se representan como números signados con precisión de 15 bits y se utilizaron para la implementación en hardware.

$$pesos_{(Hardware)} = \text{redondeo}[(pesos_{(Matlab)}) * 1024] \tag{3}$$

Para los umbrales se necesitó realizar el escalamiento ilustrado en la ecuación 4 para sumar en el mismo orden de los productos $w_{nj}x_j$. Los umbrales obtenidos son del tipo punto fijo y se representan como números signados con precisión de 23 bits.

$$umbral_{(Hardware)} = \text{redondeo}[(umbral_{(Matlab)}) * 1,048,576] \tag{4}$$

Para el diseño de la Neta de cada neurona se utilizó la descripción flujo de datos, empleando todos los datos tipo "integer" obteniendo la estructura en hardware ilustrada en la figura 6.

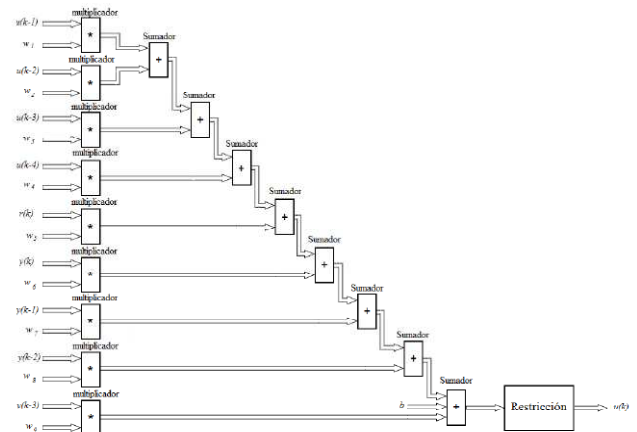


Fig. 6.- Estructura general de la Neta.

Esta estructura es totalmente combinacional. Para que el resultado correcto esté a la salida del último sumador de la Neta se necesita primero realizar la multiplicación de todos los pesos con

su respectiva entrada, el siguiente paso es realizar la suma entre las dos primeras multiplicaciones y después de obtener el resultado de esta esperar a realizar la siguiente suma y así sucesivamente hasta llegar al último sumador que realiza la suma de todas las multiplicaciones con el *umbral* para obtener a la salida el valor de la *Neta*. La salida del último sumador tiene un valor signado con una precisión de 23 *bits*.

El bloque de restricción a la salida del último sumador se implementó para eliminar los 13 *bits* menos significativos dejando la salida con un valor signado con una precisión de 10 *bits*. Como la salida de la *Neta* es la entrada de la función de activación y la entrada de esta es representada por 1024 direcciones es necesario sumar 512 a la salida de esta última operación obteniendo a la salida de la *Neta* un valor sin signo con una precisión de 10 *bits*.

El bloque *F* se utiliza para obtener la *función de activación* sigmoideal representada en la ecuación 5, donde *S* es la *Neta* de cada neurona y y_n es la salida de la neurona *n* que va hacia la entrada de otra neurona.

$$y_n = \frac{2}{(1 + e^{-2s}) - 1} \quad (5)$$

Las funciones de activación *sigmoideales* se implementaron con *megafunciones* de Altera en memoria *ROM* interna, donde las entradas de la función son las direcciones de la memoria y la salida de la función es la palabra guardada en cada dirección. El rango de entrada de la función de activación está comprendido entre el rango de [-4, 4]. Para calcular el valor correspondiente a cada dirección de memoria se trabajó con la ecuación 6 obteniendo 1024 direcciones de memoria.

$$\text{Direccion} = \text{redondeo}[(Neta + 4) * 128] \quad (6)$$

La salida de la función sigmoideal opera en un rango de [-1, 1], almacenados en cada dirección de memoria como números signados con precisión de 11 *bits*. La estructura de la función

implementada en la memoria se muestra en la figura 7.

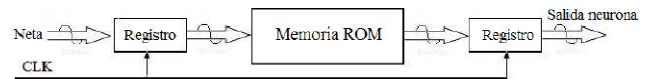


Fig. 7.- Estructura de la función de activación.

Los ciclos de reloj requeridos para procesar un dato desde la entrada de la RNA hasta su salida se llama cadencia y está definida por el número de registros que existen desde la entrada de la RNA hasta su salida. La RNA cuenta con cinco capas ocultas, cada capa oculta cuenta con dos registros, puesto que a la salida de la *Neta* está conectada a la función de activación y esta tiene dos registros. La capa de salida es de una sola neurona con función de activación lineal por lo tanto no cuenta con ningún registro. Por lo tanto la cadencia de la RNA es de 10. Para verificar el funcionamiento del controlador y adquirir los datos del comportamiento se utilizaron dos tarjetas de adquisición de datos y el software LabVIEW.

III.6. Modelo Directo en FPGA.

El sistema de desarrollo utilizado para implementar este controlador es el *TREX-S2-60-4*. En la figura 8 se ilustra el diagrama a bloques del modelo directo implementado. La implementación de cada uno de los bloques sigue la misma lógica que la utilizada para el control por modelo inverso.

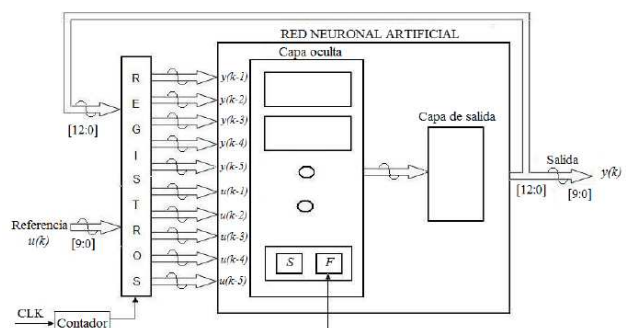


Fig. 8.- Modelo directo en hardware.

Después de verificar el funcionamiento de la RNA con el modelo directo mediante simulación

se implementó la estructura de la figura 9. En donde las señales de reloj se generan en el contador del modelo directo y se envían al controlador por modelo inverso.

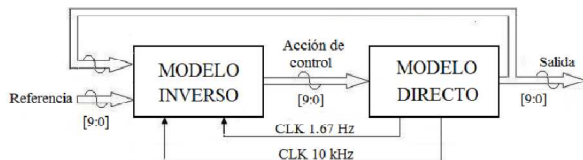


Fig. 9.- Modelos en Hardware.

III.7. Modelo Interno en FPGA.

Para la implementación del control por modelo interno se utilizaron: el controlador por modelo inverso y el modelo directo implementados en hardware. En la figura 10 se ilustra el diagrama a bloques de la implementación del controlador por modelo interno. En el sistema de desarrollo *Cyclone II* se implementó la RNA con el modelo inverso del sistema incluyendo el sumador *A* de la figura 10 y en el sistema de desarrollo *TREX-S2-60-4* se implementó la RNA con el modelo directo del sistema incluyendo el sumador y el selector *B* de la figura 10. Las señales referencia, control, realimentación y nivel se manipularon con una precisión de 8 bits. Las señales de reloj necesarias para el funcionamiento de la RNA son generadas por la RNA con el modelo directo del sistema.

Las herramientas empleadas para verificar el funcionamiento del controlador son: dos tarjetas de adquisición de datos para realizar la conversión analógico-digital y digital-analógico además de utilizarse para visualizar gráficamente el funcionamiento del sistema en LabVIEW, dos computadores personales para monitorear las señales provenientes de los FPGAs con la herramienta "Signal Tap" de Quartus. Aunque se utilizaron todas estas herramientas con el propósito de observar el correcto funcionamiento del sistema, este tiene la capacidad de trabajar de forma autónoma.

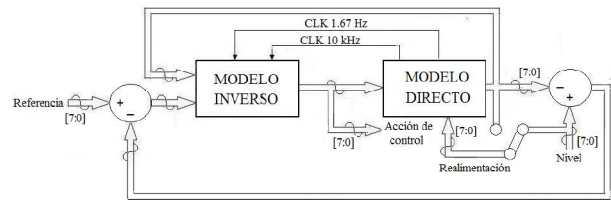


Fig. 10.- Modelo interno en hardware.

IV. Resultados

IV.1. Modelo Inverso en LabVIEW.

El funcionamiento del control por modelo inverso en línea con el proceso utilizando LabVIEW se presenta a continuación. En la figura 11 se muestra el seguimiento de la variable del proceso a la referencia, colocando valores en todo el rango de operación, con dificultades para controlar en los valores pequeños de la referencia (0.05) y en el rango superior (0.85). En la figura 12 se observa el comportamiento de la acción de control ante los cambios de referencia ilustrados en la figura 11.

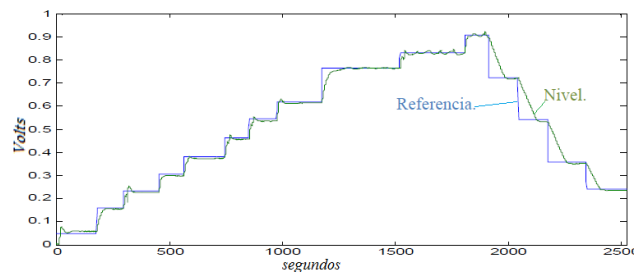


Fig. 11.- Control por modelo inverso en el proceso.

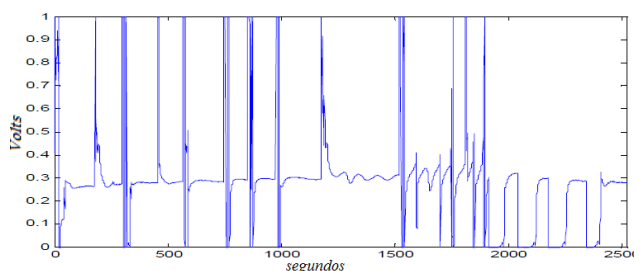


Fig. 12.- Acción de control.

Una vez analizado el comportamiento del control por modelo inverso ante cambios de referencia, se analizó su comportamiento ante perturbaciones. En la figura 13 se muestra la referencia y la variable del proceso, y en la figura

14 se ilustra la acción de control. Las perturbaciones constantes y pequeñas aplicadas al proceso se tratan de corregir por el controlador (A de la figura 13) pero cuando se le aplica una perturbación grande (aproximadamente 3 veces el desague normal del tanque) el controlador pierde estabilidad y comienza a oscilar (B de la figura 13). Las perturbaciones momentáneas se observa que son corregidas adecuadamente por el controlador (C de la figura 13).

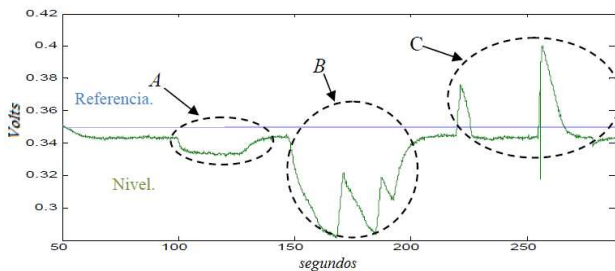


Fig. 13.- Variable del proceso ante perturbaciones.

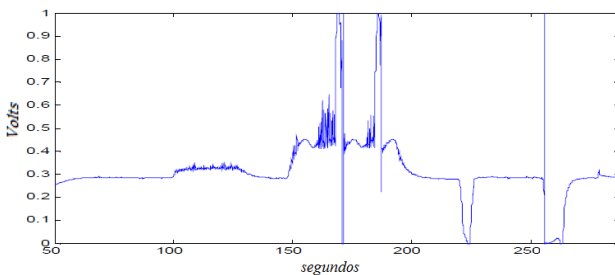


Fig. 14.- Acción de control ante perturbaciones.

IV.2. Modelo Interno en LabVIEW.

El control por modelo interno funciona igual que el control por modelo inverso para cambios de referencia, pero su comportamiento ante perturbación se corrigió notablemente. En la figura 15 se muestra la referencia y la variable del proceso para el control por modelo interno, y en la figura 16 se ilustra el comportamiento de la acción de control.

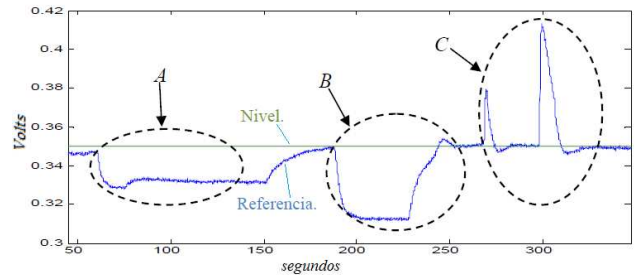


Fig. 15.- Control interno ante perturbaciones.

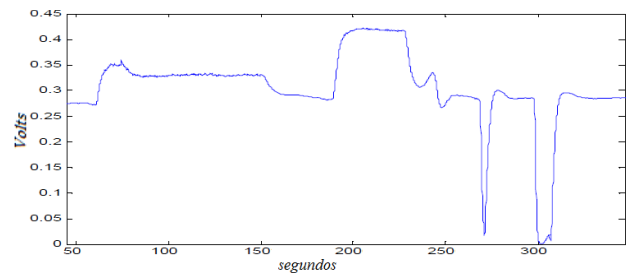


Fig. 16.- Acción de control.

Las perturbaciones aplicadas al proceso con este controlador son aproximadamente iguales a las del control por modelo inverso. Para las perturbaciones grandes la respuesta del sistema presenta errores en estado estacionario (A y B de la figura 15 son perturbaciones permanentes, $A < B$); mientras mayor sea la perturbación, mayor será el error en estado estacionario, pero la acción de control en ningún momento hace oscilatorio el sistema. La corrección de perturbaciones instantáneas es mejor que el control por modelo inverso (C de la figura 15).

IV.3. Modelo Inverso en FPGA.

Al implementar el control neuronal directo por modelo inverso sobre el sistema de desarrollo se obtuvo el reporte de compilación ilustrado en la figura 17.

Revision Name	inverso
Top-level Entity Name	inverso
Family	Cyclone II
Device	EP2C70F672C6
Timing Models	Final
Met timing requirements	Yes
Total logic elements	11,563 / 68,416 (17 %)
Total combinational functions	11,245 / 68,416 (16 %)
Dedicated logic registers	715 / 68,416 (1 %)
Total registers	715
Total pins	43 / 422 (10 %)
Total virtual pins	29
Total memory bits	285,184 / 1,152,000 (25 %)
Embedded Multiplier 9-bit elements	0 / 300 (0 %)
Total PLLs	0 / 4 (0 %)

Fig. 17.- Reporte de compilación.

La frecuencia máxima de operación del controlador embebido en hardware es de 44.61 *MHz*. Como la cadencia del controlador es 10, el controlador puede procesar 4.47 datos por microsegundo.

En la prueba experimental con el controlador interactuando con el proceso se obtuvieron los resultados ilustrados en las figuras 18 y 19. En la figura 18 se muestra el comportamiento de la variable del proceso (línea en color azul) ante cambios de referencia (línea en color verde). En los cambios en la señal de referencia se trata de cubrir todo el rango de operación obteniendo un error en estado estacionario significativo para valores menores a 0.1 *Volts* de referencia (A de la figura 18) y se pierde el control para valores superiores a 0.9 *Volts* (B de la figura 18). El controlador en hardware aumentó el rango de operación estable comparándolo con el implementado en LabVIEW. En la figura 19 se ilustra el comportamiento de la señal de control para los cambios de referencia mencionados anteriormente. El controlador es muy sensible ante perturbaciones. En A de la figura 19 se observa el comportamiento de la señal de control ante dos perturbaciones momentáneas (el comportamiento de la variable del proceso se muestra en C de la figura 19) y en B de la figura 19 se muestra el comportamiento ante una perturbación constante (el comportamiento de la variable del proceso se muestra en D de la figura 18).

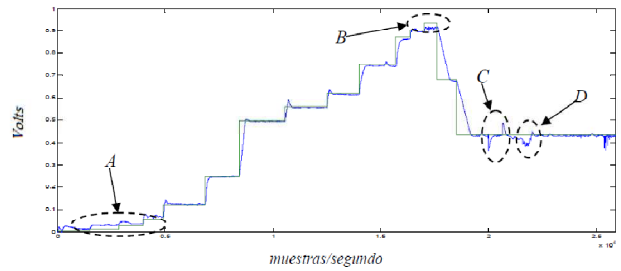


Fig. 18.- Comportamiento del controlador.

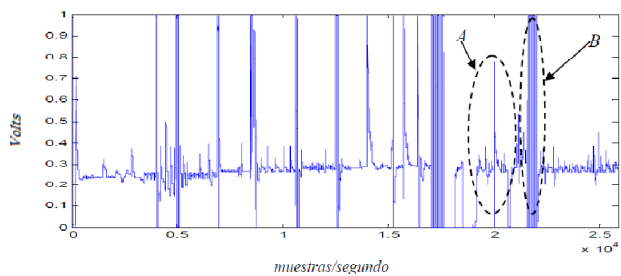


Fig. 19.- Señal de control.

IV.4. Modelo Directo en FPGA.

Al implementar el modelo directo del proceso sobre el sistema de desarrollo se obtuvo el reporte de compilación ilustrado en la figura 20. La frecuencia máxima de operación del controlador es de 38.28 *MHz*. Como la cadencia del controlador es 2, el controlador puede procesar 19.14 datos por microsegundo.

Revision Name	directo
Top-level Entity Name	directo
Family	Stratix II
Device	EP2S60F1020C4ES
Timing Models	Final
Met timing requirements	Yes
Logic utilization	23 %
Combinational ALUTs	10,552 / 48,352 (22 %)
Dedicated logic registers	755 / 48,352 (2 %)
Total registers	755
Total pins	28 / 719 (4 %)
Total virtual pins	21
Total block memory bits	981,873 / 2,544,192 (39 %)
DSP block 9-bit elements	20 / 288 (7 %)
Total PLLs	0 / 12 (0 %)
Total DLLs	0 / 2 (0 %)

Fig. 20.- Reporte de compilación.

Se realizó el experimento ilustrado en la figura 9 implementando los dos modelos neuronales, uno en cada sistema de desarrollo con el propósito de verificar que el modelo inverso controlara el modelo directo obteniendo los resultados

mostrados en la tabla 1. Se realizan tres pruebas cada una con un valor de referencia diferente y se muestra el valor que tomara la salida del modelo directo (nivel). En la prueba 1 existe un error en estado estacionario de 3 unidades (el rango de operación es de 1,024 unidades) por lo tanto el control se puede considerar eficiente.

Tabla 1.- Implementación en hardware.

PRUEBA	SEÑAL	VALOR
1	REFERENCIA	160
	NIVEL	157
	CONTROL	271
2	REFERENCIA	252
	NIVEL	247
	CONTROL	279
3	REFERENCIA	380
	NIVEL	373
	CONTROL	290

IV.5. Modelo Interno en FPGA.

Después de realizar las modificaciones de la estructura ilustrada en la figura 10 los reportes de compilación de cada sistema de desarrollo no presentan gran variación. La frecuencia máxima de operación es de 48.99 MHz.

En el sistema de desarrollo donde se implementó el modelo directo la frecuencia máxima de operación es de 42.7 MHz. Por lo tanto el control por modelo interno opera con la frecuencia menor de los dos sistemas implementado 42.7 MHz. En la figura 21 se muestra el comportamiento del control por modelo interno, la línea roja representa la referencia, la línea verde la variable del proceso y la línea azul la señal de control. El comportamiento para cambios de referencia es igual a los del controlador por modelo inverso.

En A de la figura 21 se muestra el comportamiento del controlador para el proceso en estado estable. En B de la misma figura se aplican perturbaciones momentáneas a las cuales el controlador responde correctamente corrigiendo el error y en C se aplica una

perturbación constante la cual el controlador trata de corregir teniendo un error en estado estacionario como se había observado en la implementación con LabVIEW.

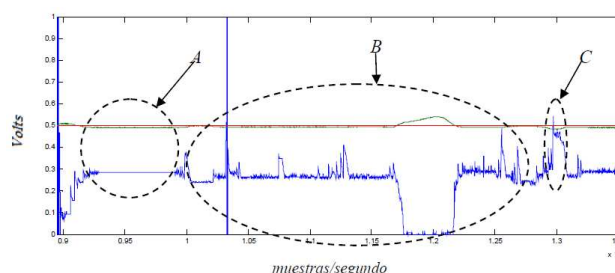


Fig. 21.- Comportamiento del control interno.

V. Conclusiones

La velocidad de operación de la RNA implementada en hardware es grande, por lo tanto se puede implementar un controlador neuronal en un FPGA para variables rápidas.

El uso de elementos del FPGA fue mínimo, lo cual permite implementar una RNA de mayor tamaño o una RNA del mismo tamaño pero utilizando mayor número de bits.

Al realizar el entrenamiento de las RNA en Matlab se observó que el aumento en el número de capas tiene menor gasto computacional que el aumento de número de neuronas por capa y por lo tanto un menor número de componentes al implementar la RNA.

En cuanto a la comparación del funcionamiento de los controladores en LabVIEW y los implementados en FPGA, los controladores en FPGA presentan la ventaja de poder utilizar mayor rango de operación y la desventaja de que la acción de control tiene cierta cantidad de ruido que en este caso fue despreciable por los convertidores utilizados y la reacción del actuador. Además, los controladores realizados en el FPGA tienen menor error que los hechos en LabVIEW.

VI. Referencias

- [1] L. Zhang, M. Pan, S. Quan. Multiple Model Predictive Control for Water Management in PEMFC Based on Recurrent Neural Network Optimization. *Networking, Sensing and Control*. 2008. 853–858.
- [2] Bastos, J.L. Figueroa, FPGA implementation of neural network-based controllers for power electronics applications, 2006.
- [3] Chao-Ming Chung, Chih-Min Lin, Hardware Implementation of CMAC Neural Network using FPGA Approach, 2007.
- [4] Seul Jung and Sung su Kim, Hardware Implementation of a Real-Time Neural Network Controller With a DSP and an FPGA for Nonlinear Systems, 265-271, 2007.
- [5] Himavathi S., Muthuramalingam A. Feedforward Neural Network Implementation in FPGA Using Layer Multiplexing for Effective Resource Utilization, 880-888, 2007.
- [6] D. Peng, H. Zhang, Adaptive Inverse Control for Water Level of Boiler Drum Based on Neural Network. *Intelligent Control and Automation*, 2911-2915, 2006.
- [7] P Yang, D. Peng, Neural networks internal model control for water level of boiler drum in power station. *Machine Learning and Cybernetics*. 5. 3300-3303.
- [8] J. J. Martínez, Estudio Comparativo De Esquemas De Control Aplicado A Un Modulo Didáctico De Nivel De Agua, VI Semana Nacional de Ingeniería Electrónica, 13 al 15 de Octubre 2010.
- [9] E. N. Sánchez Camperos, Redes neuronales, conceptos fundamentales y aplicaciones a control automático. 2006. Prentice-Hall. 165-175.
- [10] B. Kuo, *Sistemas de control automático*. Séptima Edición. 2006. Prentice Hall. 160-166.
- [11] Altera, Cyclone II DSP Development Board, 2006.

Sistema de Identificación y Control del Servo Motor Amira para uso Didáctico Desarrollado en Instrumentación Programable

*Wilebaldo Martínez-Velazco, *Felix Emilio Luis-Pérez, **Fermín Hugo Ramírez-Leyva

*División de Estudios de Posgrado

**Instituto de Electrónica y Mecatrónica, Universidad Tecnológica de la Mixteca (UTM) Carretera a Acatlima
Km. 2.5 Huajuapán de León, Oaxaca, C.P. 69000, Tel. (01953) 5320399 ext. (555)
{wmartinez, eluis, hugo}@mixteco.utm.mx

Resumen

En este trabajo se muestra una forma de implementar la identificación de parámetros y control de un sistema servo motor Amira DR300. El sistema de identificación de parámetros es implementado en el lenguaje de programación gráfica LabVIEW, las señales de voltaje, velocidad y corriente del motor analizado se toman por medio de una tarjeta de adquisición de datos USB1208FS y para la potencia se usa una fuente de voltaje programable de la firma HP. El módulo de identificación de parámetros es programado en un script de Matlab el cual es ejecutado desde la interfaz de LabVIEW. Con este sistema se pueden hacer experimentos de control de servomotores y facilitar el entendimiento de los conceptos clásicos de control.

Palabras clave: Instrumentación programable, Control, Identificación de parámetros, LabVIEW, Amira DR300.

I. Introducción

La ingeniería de control ha tenido un fuerte impacto en el desarrollo tecnológico. Al inicio de la era industrial, el control de los procesos se llevó a cabo basándose en la intuición y la experiencia acumulada. Más tarde, el mercado exigió mayor calidad en las piezas fabricadas, esto condujo al desarrollo de teorías para explicar el funcionamiento de los procesos. Estas teorías derivaron en estudios analíticos que permitieron realizar el control de la mayor parte de las variables de interés en los procesos industriales [1].

Para reforzar los conocimientos vistos en clase sobre la teoría de control, es importante llevar a la práctica los conceptos teóricos. Los laboratorios convencionales, han sido tradicionalmente el único sitio para desarrollar prácticas experimentales. Posiblemente una de las principales ventajas que ofrece un laboratorio convencional es su interactividad, al permitir que el estudiante tenga contacto con una planta real. Al poder observar lo que sucede en sus

experimentos, el alumno comienza a desarrollar habilidades cognitivas [2].

La rápida adopción de la PC en los últimos 20 años generó una revolución en la instrumentación de ensayos, mediciones y automatización de procesos. Un importante desarrollo resultante de la ubicuidad de la PC es el concepto de instrumentación virtual, el cual ofrece variados beneficios a ingenieros y científicos que requieren de mayor productividad, precisión y rendimiento [3].

Un instrumento virtual consiste de una computadora equipada con programas (software) y hardware que cumplen, en conjunto, las funciones correspondiente a los instrumentos tradicionales. Los instrumentos virtuales representan un apartado fundamental de los sistemas de instrumentación basados en el hardware y sistemas centrados en el software que aprovechan la potencia de cálculo, productividad y capacidad de conexión de las populares computadoras de escritorio y estaciones de

trabajo. Con los instrumentos virtuales, los ingenieros y científicos construyen sistemas de medición y automatización que se ajustan exactamente a sus necesidades (definidos por el usuario) en lugar de estar limitados por los instrumentos tradicionales de funciones fijas (definidos por el fabricante) [3].

Cuando se tiene un sistema y éste puede ser controlable, generalmente se le desarrolla un modelo matemático con base en ecuaciones diferenciales, con el cual se describe su correspondiente comportamiento. Para sistemas lineales la ecuación diferencial se transforma en una función de transferencia, y con base en ésta se determina la estabilidad del sistema. El modelo matemático puede ser determinado por experimentación, o con la medición de la respuesta del sistema ante entradas conocidas [4].

La identificación de sistemas consiste en determinar el modelo dinámico de un sistema con base en la medición de los estímulos y la respuesta, los datos muestreados se pueden usar en un amplio rango de aplicaciones como son ingeniería mecánica, biología e ingeniería eléctrica entre otros [5].

Existen varias técnicas para la estimación de parámetros de un sistema, entre las que se encuentran: las técnicas paramétricas, las no paramétricas, aquellas que tienen conocimiento parcial del modelo, con base al modelo, las de respuesta en lazo cerrado y los estimadores recursivos. El mejor método a usar depende de la naturaleza del sistema que se desea caracterizar [5]. Los filtros de Kalman es un método de tipo recursivo el cual es ampliamente usado para identificar sistemas cuando se tienen señales con ruido, con base en las mediciones de la entrada y salida se estiman los parámetros del modelo del sistema.

En los últimos años, LabVIEW se ha posicionado como uno de los ambientes de desarrollo más usados para aplicaciones con instrumentación programable. Posee varios “Toolbox” que facilitan su uso en aplicaciones específicas, uno

de ellos es el de identificación de sistemas y control, en el cual se tienen programados diferentes técnicas de estimación de sistemas y control. Sin embargo hay que hacer un pago extra para poder usarlo. Una alternativa, es desarrollar aplicaciones propias tomando el mismo LabVIEW como plataforma de desarrollo, al mismo tiempo, el desarrollador adquiere la experiencia de saber cómo se realizan los procedimientos de estimación.

En la Universidad Tecnológica de la Mixteca (UTM), en el laboratorio de Control se cuenta con varias plantas de la firma Amira, entre ellas el Laboratorio de experimentación de control de velocidad con carga variable DR300. Este equipo es usado para probar prácticas de control de velocidad. Posee una plataforma desde la cual se pueden probar controladores PID, o controles por lógica difusa [6]. Sin embargo tiene una arquitectura cerrada y no se puede acceder a los métodos que realizan el control. El motor también permite ser controlado externamente.

En este trabajo se muestra la forma de realizar un software para identificación del sistema, usando el algoritmo de identificación por mínimos cuadrados aplicado al sistema DR300 de Amira. También se implementa un controlador PI para el control de velocidad del sistema y se muestra el correcto funcionamiento del mismo variando la velocidad deseada y el par de carga. El ambiente de desarrollo es LabVIEW, para la adquisición de datos se usa una tarjeta de adquisición de datos USB y la etapa de potencia es implementada con una fuente de voltaje programable. El propósito de este sistema es mostrar a los alumnos la forma en que se obtienen los parámetros del sistema y así comprobar cómo funciona este controlador en un sistema real.

II. Sistema Amira

El control de motores es una de las principales aplicaciones en la industria, ya que la mayoría de los sistemas que tienen movimiento utilizan uno para generarlo. El sistema de entrenamiento Amira Modelo DR300, es un servo mecanismo

que consiste de dos motores idénticos, los cuales son conectados por un embrague mecánico [6]. El primer motor es usado para controlar la velocidad de rotación. El segundo motor es comúnmente llamado generador, y es usado para la simulación de la carga de torque [9]. Con este sistema se puede observar el comportamiento del motor ante condiciones de carga constante o variable. Para controlar el par de carga generado por el motor o el generador, en su panel posterior cuenta con 2 conectores BNC donde se puede aplicar el correspondiente voltaje de control (ver Fig. 1).

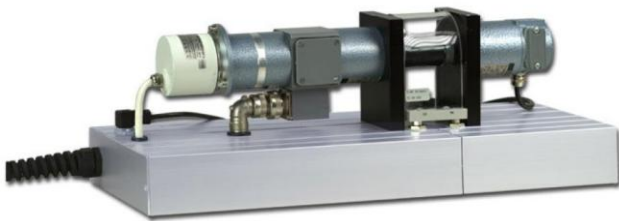


Fig. 1. Servo motor Amira DR300 y su generador.

El sistema Amira se conecta a un panel frontal donde se tienen señales de voltaje proporcionales a la velocidad, corriente del motor y corriente del generador. También es posible probar un controlador PID o cambiar el control hacia una computadora personal (PC) en la cual se ejecuta un programa propietario del fabricante para probar diferentes esquemas de control.

Eléctricamente el motor se modela como se muestra en la Fig. 2. Dicho circuito es regido por la ecuación 1[6].

$$L \frac{di(t)}{dt} + Ri(t) + k_e \omega(t) = u(t)$$

$$J \frac{d\omega(t)}{dt} + k_m i(t) + b\omega(t) + f_c \text{sign}(\omega(t)) = m_z \quad (1)$$

Donde: $i(t)$ [A], $\omega(t)$ [rad/s], $u(t)$ [V], m_z [N m] son la corriente del motor, la velocidad de rotación, voltaje de entrada y el par de carga respectivamente. R , L , k_e , k_m , J , b y f_c son la resistencia, la inductancia, la constante eléctrica, constante mecánica del motor, el momento de

inercia, la fricción viscosa y la fricción de Coulomb.

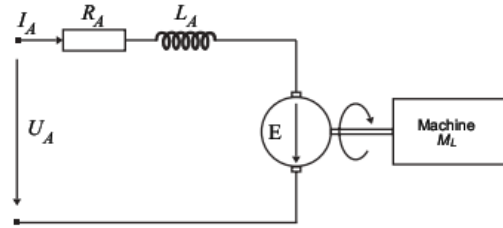


Fig. 2. Diagrama eléctrico del motor.

El filtro de Kalman es un conjunto de ecuaciones matemáticas que proveen una solución recursiva eficiente del método de mínimos cuadrados. Esta solución permite calcular un estimador lineal, insesgado y óptimo del estado de un proceso en el tiempo actual (t_k), con base en la información disponible en el tiempo anterior (t_{k-1}) [10].

El filtro de Kalman tiene como objetivo resolver el problema general de estimar el estado $X \in R^n$ de un proceso controlado en tiempo discreto, el cual es dominado por una ecuación lineal en diferencia estocástica. Las variables aleatorias representan el error del proceso y de la medida respectivamente. Se asume que son independientes entre ellas, que son ruido blanco y con distribución de probabilidad normal, ambas con media cero.

En la práctica las matrices de covarianza de la perturbación del proceso y de la perturbación de la medida, podrían cambiar en el tiempo, por simplicidad en general se asumen que son constantes. Para el cálculo de los parámetros del motor en función de las mediciones se propone el sistema de la siguiente forma.

$$u(t) = \Psi_1^T \theta_1 \quad (2)$$

$$i(t) = \Psi_2^T \theta_2$$

Donde $u(t)$ e $i(t)$ son el voltaje y la corriente en el motor, $\Psi_1^T = \left[\frac{di}{dt} \quad i(t) \quad \omega(t) \right]$ y $\Psi_2^T = \left[\frac{d\omega}{dt} \quad \omega(t) \quad m_z \quad \text{sign}(\omega(t)) \right]$, son los vectores de las mediciones, finalmente $\theta_1^T = [L \quad R \quad k_e]$ y

$\theta_2^T = \left[-J/k_m \quad -b/k_m \quad 1/k_m \quad -f_c/k_m \right]$ son los parámetros a identificar.

Tomando como base el modelo (2) se propone un observador del tipo $\hat{u} = \Psi_1^T \theta_1$. La función de error es $e = u - \hat{u}$. Después de aplicar un método de minimización de error (optimización), se obtienen las siguientes ecuaciones con las cuales se hace la identificación de los parámetros en forma recursiva.

$$\begin{aligned}
 e_k &= \tau_k - \Psi_k^T \hat{\theta}_{k-1} \\
 P_k &= P_{k-1} - P_{k-1} \Psi_k \left[I + \Psi_k^T P_{k-1} \Psi_k \right]^{-1} \Psi_k^T P_{k-1} \\
 \hat{\theta}_k &= \hat{\theta}_{k-1} + P_{k-1} \Psi_k \left[I + \Psi_k^T P_{k-1} \Psi_k \right]^{-1} e_k
 \end{aligned}
 \tag{3}$$

Donde: $e_k \in n \times 1$ es el error, $\tau \in n \times 1$ son las entradas, $P_k \in n \times n$ la matriz de covarianza e $I \in n \times n$ la matriz identidad. En nuestro caso, se propusieron como matrices iniciales para P, una matriz diagonal con valor de 1000 y para θ un vector de ceros.

III. Control PI de velocidad

Un controlador PI reacciona al error con la suma de la parte proporcional y la parte integral. La parte proporcional es el producto del error por la ganancia proporcional (k_p) como se muestra en la ecuación (4). En motores una variable a controlar es la velocidad deseada (ω_d), por lo cual el error se define como la diferencia entre la velocidad deseada y la medida (ω), $e = \omega_d - \omega$.

$$P = k_p e \tag{4}$$

Por otro lado la acción integral representa una reacción al acumulado del área bajo la curva del error. La ecuación (5) describe la implementación de la parte integral en forma discreta, donde T es el período de muestreo, $e(kT) = \omega_d - \omega$, es el error actual, A_k es el área del error y k_i la ganancia integral. En la Fig. 3 se muestra el

diagrama a bloques del controlador PI en forma discreta.

$$\begin{aligned}
 u(k) &= k_p e(kT) + k_i A_k \\
 A_k &= A_{k-1} + e(kT)T
 \end{aligned}
 \tag{5}$$

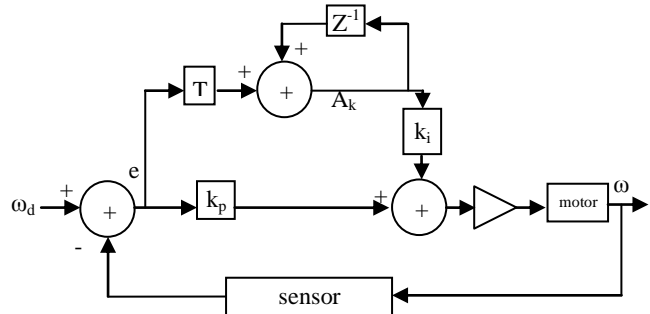


Fig. 3. Diagrama a bloques del control PI.

Para realizar la identificación del sistema y hacer el control PI se necesita hacer un sistema de lazo cerrado desde el cual mida la velocidad, compare con la velocidad de referencia y aplique el voltaje de control necesario para reducir el error. Para el caso de la identificación de sistema, se necesita aplicar una señal de excitación y medir la velocidad, corriente, voltaje y con base en la ecuaciones (3), estimar los parámetros del modelo.

La Fig. 4 muestra la forma en que se realiza el control del motor Amira. Consta de dos fuentes de voltaje programables, una tarjeta de adquisición de datos por USB para la medición de los voltajes proporcionales a la velocidad, corriente y voltaje aplicado. La tarjeta de adquisición de datos es conectada a una laptop para el procesamiento de los datos.

La fuente de voltaje Agilent E3631A, es una fuente de poder de alto rendimiento de 80W y tres salidas de voltaje, equipado con las interfaces GPIB y RS-232 [7]. Tiene tres salidas de voltaje, dos de ellas, trabajan en el rango de 0 a $\pm 25V$ con una corriente entre 0 y 1 Ampere. La tercer salida puede proporcionar voltajes entre 0 y 6V con una corriente de 0 – 5A. El voltaje y corriente para

cada una de las salidas puede ser ajustada independientemente desde el panel de control o bien, puede ser programada con las interfaces GPIB y RS-232.

En este trabajo se utilizó el puerto RS-232 para enviar los comandos de control y fijar el voltaje de la fuente de 6V que es la que se le conecta al motor. Los comandos de comunicación SCPI (*Standard Commands for Programmable Instruments*) se usan para configurar la fuente. El par de carga se fijó con una fuente externa.

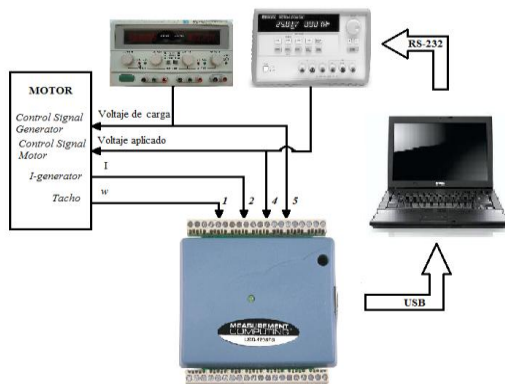


Fig. 4. Ciclo de control del sistema.

La tarjeta de adquisición usada fue la USB-1208FS de la firma Measurement Computing [8]. Esta tarjeta tiene 8 canales de entrada para el convertidor Analógico Digital en el rango de $\pm 10V$, la frecuencia de conversión es de hasta 50k muestras/s, 2 salidas analógicas y 16 puertos de entrada/salida digitales con niveles TTL. Se usaron 3 canales analógicos para la medición de la corriente y voltaje del motor y uno para la velocidad.

La computadora portátil tiene un puerto RS-232 y 3 puertos USB uno de los cuales se usa para conectar la USB-1208FS. En esta computadora se cargó el programa LabVIEW desde el cual se hizo la programación del algoritmo de identificación de sistemas y el controlador PI. Para facilitar la programación se usó el toolbox de LabVIEW para la ejecución de scripts Matlab.

IV. Desarrollo del software

El sistema realizado para la estimación de parámetros fue implementado en la plataforma de desarrollo LabVIEW en su versión estudiantil 8.2, debido a su facilidad de uso en protocolos de comunicación vía serial y la vasta librería de instrumentos de medición con la que cuenta.

En la Fig. 5 se muestra un diagrama a bloques del sistema que se implementó para la estimación de parámetros. Cada bloque representa una función específica del sistema.

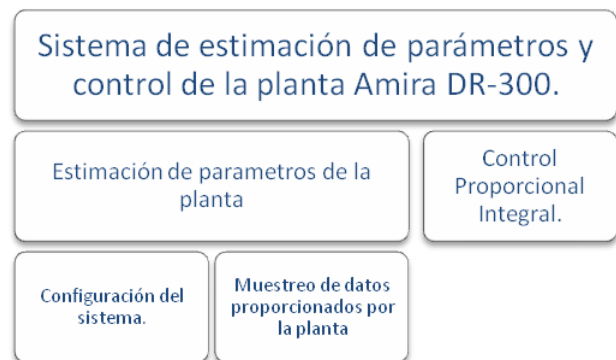


Fig. 5. Modelado del software desarrollado.

El sistema para identificación consta básicamente de tres etapas, en la primera de ellas se realiza la configuración de las fuentes de voltaje, en la segunda etapa se obtienen las señales de los sensores de voltaje y corriente de la planta, en la etapa final, se calculan los parámetros del motor con ayuda del filtrado de Kalman. Además es posible realizar el control de velocidad de los motores de la planta con un controlador de tipo Proporcional Integral.

La pantalla principal del sistema es la que se muestra en la Fig. 6. En la pestaña de configuración se solicita al usuario los correspondientes voltajes que se aplicarán a la planta por medio de la fuente programable. En ella se solicita al usuario el tipo de señal que se aplicará al sistema, esta señal puede variar de un valor constante, una señal tipo diente de sierra, un voltaje aleatorio, o bien un voltaje de señal cuadrada. También se solicita al usuario el voltaje

correspondiente al par de carga, el cual siempre es un valor constante (ver Fig. 7).



Fig. 6. Ventana principal del sistema de identificación de parámetros.



Fig. 7. Pantalla de configuración del sistema.

En la siguiente ventana, la de mediciones, se toman en forma discreta las señales producidas por la planta por medio de la tarjeta de adquisición de datos USB, los datos tomados por la tarjeta son graficados en pantalla y pueden ser direccionados a un archivo de texto plano para su posterior utilización. En esta ventana se puede configurar la frecuencia de muestreo y el número de muestras que se toman de las señales para un correcto muestreo de las mismas. Dicha ventana es mostrada en la Fig. 8.

En la cuarta ventana, correspondiente a la identificación paramétrica, se hace uso del

archivo generado por las mediciones para ejecutar el algoritmo de Kalman y estimar los correspondientes parámetros de la planta.

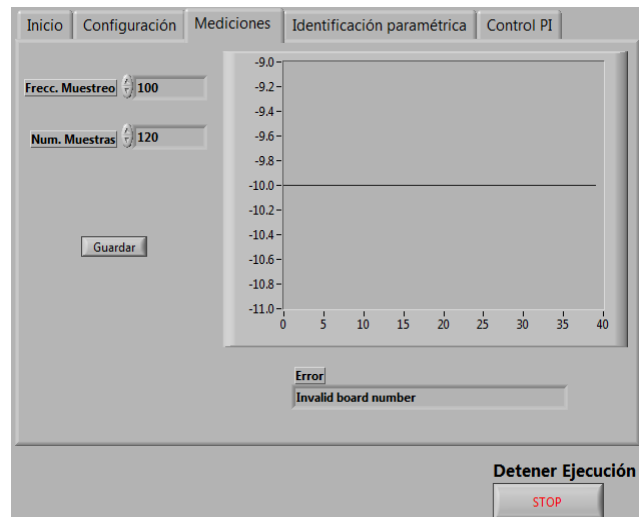


Fig. 8. Pantalla de muestreo de datos.

En la ventana correspondiente al controlador PI, se puede realizar el control de velocidad del motor a través de la programación de un controlador PI. Para este control es necesario proporcionar la constante de proporcionalidad k_p , la constante de integración k_i y el tiempo de muestreo T_s , parámetros necesarios para implementar esta técnica tal como se explico en la sección II.

Al ejecutar el programa el usuario proporciona la velocidad a la que se desea que trabaje el servomotor, k_p , k_i y T_s . La respuesta del controlador se puede observar en la gráfica de la Fig. 9, donde se aprecian tres señales, la velocidad deseada, la velocidad del motor y el voltaje aplicado a la planta para controlar el sistema. También se pueden observar los mensajes de error que se pudieran producir en el envío de datos hacia la fuente de voltaje (error IO) y en la comunicación con la tarjeta de adquisición de datos (ErrMsg), el puerto serial que se comunica con la fuente de voltaje (VISA resource name) y la cadena de datos aplicada a esta.

Al Motor Amira se le aplica carga mecánica mediante el generador que tiene acoplado, éste se

controla aplicando un voltaje con una fuente de voltaje externa, la cual no se controla desde este programa.

b/k_m	0.0527	-
l/k_m	0	-
f_c/k_m	0.2479	-

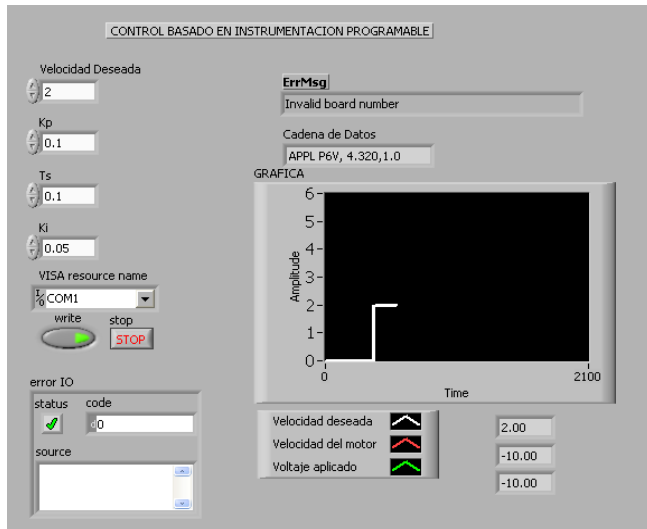


Fig. 9. Control Proporcional Integral.

V. Pruebas y resultados

Con los datos obtenidos de la velocidad, corriente y voltaje aplicados a la planta, se realizó la estimación de los correspondientes parámetros de Inductancia, Resistencia y Fuerza electromotriz inducida. En la Tabla 1 se muestran los datos obtenidos.

Como se puede ver, los valores que se especifican en el manual del motor están en el mismo orden de magnitud que los proporcionados por el sistema, y para los no especificados no se puede realizar la comparación correspondiente. Un problema que tiene el algoritmo es que se requiere la derivada de la velocidad y la corriente, esto se hizo en forma numérica, por lo cual el resultado es muy ruidoso. El otro problema es que se debe especificar un par de carga que inicialmente no se conoce.

Tabla1. Valores estimados con el Algoritmo

Parámetro	Estimado	Especificado
L	1mH	3mH
R	0.961Ω	2.6Ω
K_e	0.011	-
J/k_m	0.045	-

Para visualizar la respuesta del controlador PI se definió un perfil en el que puede observarse la velocidad deseada, la velocidad del motor y el voltaje aplicado. En la gráfica mostrada (Fig. 10) se puede observar la respuesta del sistema a los valores de velocidad deseada y carga aplicados correspondientes a la Tabla 2. El período de muestreo $T_s = 0.1s$ fue el mínimo tiempo que se consiguió porque se tiene que enviar el comando SCPI a la fuente de voltaje para que configure su salida. La constante de voltaje en la velocidad es de 6.27mV/rpm y la de la corriente del motor es de 2.5 V/A. El perfil de velocidad deseado del sistema se muestra en la Tabla 2.

Para sintonizar el controlador PI fue necesario definir T_s , debido a la velocidad de operación de la fuente de voltaje. La constante $k_p = 0.1$ y $k_i = 0.1$ se obtuvieron experimentalmente. Cuando la velocidad deseada se estableció en 3V el motor se empieza a mover hasta que se vence la fricción de Coulomb. En el sistema, se varió la velocidad deseada y el par de carga aplicado con el generador. Posteriormente se deseó una velocidad de 5V y otra vez 3V (478rpm y 797rpm). La Fig. 10 muestra los resultados experimentales obtenidos. La grafica blanca es la velocidad deseada, la roja la velocidad y la verde el voltaje aplicado.

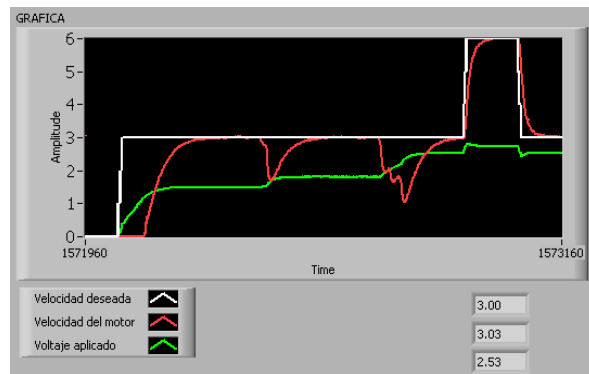


Fig. 10. Prueba con distintas velocidades y cargas.

La respuesta del controlador es un poco lenta, debido al tiempo de integración que se utilizó

(Ts), sin embargo la curva la velocidad del motor siempre es alcanzada, comportamiento que corresponde con el control proporcional integral.

Tabla2. Valores de Velocidad y carga aplicados

Tiempo (seg)	Velocidad deseada (V)	Carga aplicada (V)
0	0	0
2	3	0
7	3	0.5
12	3	1
15	6	1
17	3	1

VI. Conclusiones

En este trabajo se muestra la forma en que se implementó el algoritmo de identificación de parámetros para un motor de corriente directa. La parte de identificación se hizo en Matlab y se manda a llamar desde el LabVIEW, la parte de control se hace desde el mismo LabVIEW. A pesar de ser un sistema con un tiempo de respuesta de 100ms, se tuvo una buena respuesta del controlador PI.

Una característica importante del sistema es que no se requiere de mucha capacidad de cómputo para que el software desarrollado funcione, el único requisito que se debe de cumplir es que se pueda instalar LabVIEW y posea al menos 2 puertos USB o bien un puerto serial y uno USB.

El control Proporcional Integral tuvo una respuesta lenta debido a que para la conexión con la fuente de voltaje, se utilizó el puerto RS-232. Como un trabajo futuro se utilizara el puerto GPIB para obtener una señal de control más rápida y/o usar las salidas analógicas de la tarjeta USB. La ventaja de tener el sistema de esta manera es que la capacidad de corriente de la fuente de voltaje es de hasta 6 A, lo cual permite que este mismo sistema se pueda usar en otra planta de forma más o menos rápida.

En un trabajo a futuro la fuente de voltaje que se utilizó para agregar la señal de perturbación (par de carga) al sistema también deberá ser

controlada por software, también se pretende implementar este sistema como un laboratorio virtual para el entrenamiento en el control de servomecanismos.

VII. Autores

Wilebaldo Martínez Velazco, se graduó en la Universidad Tecnológica de la Mixteca, de la Ingeniería en Electrónica en el año 2004. Actualmente es estudiante de la Maestría en Robótica en la Universidad Tecnológica de la Mixteca y se desempeña como técnico de tiempo completo en el laboratorio de Robótica Inteligente en la misma Universidad.

Felix Emilio Luis Pérez Ingeniero en computación por la Universidad Tecnológica de la Mixteca. Actualmente, estudiante de la Maestría en Robótica y colaborador del Laboratorio de Robótica Inteligente adscrito a la división de estudios de posgrado de la misma universidad.

F. Hugo Ramírez Leyva obtuvo su título de Maestría en Ciencias con especialidad en electrónica por el Instituto Nacional de Astrofísica Óptica y Electrónica. Actualmente se desempeña como profesor investigador, adscrito al Instituto de Electrónica y Mecatrónica de la Universidad Tecnológica de la Mixteca.

VIII. Referencias

- [1] Martínez Nolasco, Juan José. "*Estudio Comparativo de Esquemas de Control Aplicado a un Módulo Didáctico de Nivel de Agua*". Memorias del congreso Senie 2010. ISBN 978-607-477-363-7.
- [2] Lorandi Medina, Alberto Pedro., Herminda Saba, Guillermo., Hernández Silva, José., Ladrón de Guevara Duran, Enrique. "*Los Laboratorios Virtuales y Laboratorios Remotos en la Enseñanza de la Ingeniería*". Revista Internacional de

Educación en Ingeniería. Volumen 4, 2011. ISSN 1940-1115.

- [3] National Instruments Corp. “*La Instrumentación Virtual*”. Reporte técnico. 2003. National Instruments Corporation.
- [4] Burns, Roland S. “*Advanced Control Engineering*”. Butterworth Heinemann. 2001. ISBN 07506 51008.
- [5] National Instruments Corp. “*System Identification Toolkit User Manual*”. 2004–2006. National Instruments Corporation.
- [6] Amira GmbH. “*DR300 – Laboratory Setup Speed Control with Variable Load*”. 2000. Duisburg.
- [7] Agilent. “*Agilent E3631A, Triple Output DC Power Supply, User Guide*”. 2007. Agilent Technologies, inc.
- [8] Measurement Computing “*USB1208FS, Low cost, USB based Module with 8 Channels, 12 bit input, User guide*”. 2010. Measurement Computing.
- [9] Roubal, J., Augusta, P., Havlena, V. “*A brief introduction to control design demonstrated on laboratory model servo DR300-Amira*”. 2005. Acta Electrotechnica et Informatica. No. 4. Vol. 5. Slovak Republic.
- [10] Clar, M., J. Suriñach. “*A Latent Variable Model to Measure Regional Manufacturing Production in Spain*”. 1998. Workshop on Regional Economic Indicators. University of Minho. Braga.

The RFID Tag Performance Applied to the Vehicular Identification

Ricardo Meneses, Laura Montes, Roberto Linares

Instituto Politécnico Nacional Escuela Superior de Ingeniería Mecánica y Eléctrica, Campus Zacatenco.
U.P.A.L.M. Edif. 5, Jefatura de ICE, Col. Lindavista, C. P. 07738, México, D. F. rmeneseg@ipn.mx

Abstract

Understand and analyze the radio signal performance in the RFID link applied to the vehicular identification in the near field as in the far field is important, and at this moment there is not a work about this situation. It will be a big advantage and technologic contribution because of, it will allow to solve actual design problems, as to avoid or to reduce the intersymbolic interference, increase the bit rate, and from the propagation models, to take control on the jamming electromagnetic environment around this application operates. So, the focus of this work is to present some results of the experimental tests and measurements taken in the field

Keywords: tag detection; sensitivity; vehicular identification

I. Introduction

Actually, there are RFID studies, for example [4] and [7], which, they deal the far field or radiation field, but not applied to vehicular identification, where the tag is strapped to a mobile, and this one and the reader are not situated at the same height, living side by side with other communication systems under a complex and violent propagation environment. [6] suggests a propagation model and a probabilistic estimation of the tag position in the UHF band, including antenna characteristics, considering the direct path and multipath in a mobile robot application using a sensor system with six antennas, in the same way, [9] and [11] report the tag characteristics in the near field, and informs the RFID passive UHF performance limitations, focusing to the antenna problems, due to path loss, the radio channel, when the tag is fixed. In fact, there is not near and far field RFID propagation model applied to vehicular identification, so this work represents the proposal model the objective and a real contribution.

The RFID technology in road applications, the passive tag is strapped on the motor vehicle, as a mobile receiver, and it is enquired by the reader, while the vehicle traverses across the electric

field covering area, radiated by the antenna reader. Under a complex electromagnetic environment the reader will conduct several enquires and it is said that successful reading occurred when at least one out of the k enquires was read.

Because of the multiple signals reception, that is, a multipath environment, the mobile antenna, the antenna of the tag, and the antenna of the reader, receive in abundance reflected and scattered waves, producing in the receiver devices signal fading, phase alteration, frequency shifting.

The RFID applied to vehicular identification is a world novel application. The experience is low and there are not much systems working around the world. The actual regulation is due to the work of some organizations and working groups, which, they have participated in the resolution and elaboration of the different aspects of RFID technology rules, nevertheless, about vehicular identification there are not specific rules, producing the existence of open and closed systems.

II. Statement of the Problem

All the disturbances produce a bad reception causing false readings, and because of the intersymbolic interference and/or changes in the

signal structure. This way, the instantaneous power received by the mobile antenna is a random variable, which it depends of the antenna situation and the antenna response to the cancelation effects and/or waves contribution. Moreover, the mobile received signal in any point in the space could be built by a lot of plane waves with phase and amplitude with random distribution.

If the receiver is stationary, as in the case of RFID applied to vehicular identification, the received signal could fades because of the movement of the objects around the radio channel. If the objects are statics, the movement is due the tag, the fading concept is considered as spatial phenomenon, so, the spatial variations of the resultant signal are considered as a time factor meanwhile it moves through the multipath environment.

Because of the constructive and destructive effects of the multipath waves adding in different points in the space, the receiver device moving at high speed could suffer severe fading in a short time interval. In this way, to maintain efficient communication between the reader, the base station, and the mobile station, the tag, is difficult, even more, people or motor vehicles moving around the radio channel disturbs the electric field pattern, then, the received signal probability decreases, producing deep nulls for a long time. The antenna spatial diversity could prevent deep fading nulls.

There are model proposals, starting with the two ray model, which it has been improving according with the field and laboratory tests, including other variables, that, take part in the situation, verifying the value of some parameters, the way to be evaluated, considering a multipath environment, not only two rays, the cross polarization, the road surface roughness in order to establish the covering area, which it has high relation with the power and sensitivity characteristics of the electronic equipment (reader and tag).

And in the same way, applying some techniques as the waveguide model, the tunnel effect, using the modal technique or the IPO (Iterative Physical Optic Method) technique, which allow

to know and analyze the performance system under environment without line sight, that is, there are some obstacles blocking partial or totally the line sight, in this way, putting the system in a big traffic environment and it is necessary the identification.

In the other hand, a deterministic model channel predicts the electric strength field, taking in account the structure, geometry and materials of the covering area. This deterministic model requires a lot of structural information and data, often not available. Meanwhile a pure stochastic model does not take in account the radio propagation details, making easy the analysis and using low computational resources, which, the combination of both of them makes a stronger model. From above mentioned, the ray tracing looks to have a rough idea for specific environment, which, it works as a lot of rays source to all directions bouncing and penetrating with the information of the material and geometry of each object around the environment, and above all the walls and ground. This model is improved with the Uniform Diffraction Theory, because of the corners are considered. But due to a lot information necessary to describe the material, the model turns complex to solve the situation and in the same way the algorithms and computational resources.

Actually basic information (owner name, license, fines, car robbery report, etc.), is resident in the memory tag, which it could be so much, if it is only necessary a key word linked with a data base where this information is contained, or it could be so little, if the text information is not enough, for example the owner photograph, image files, video files and other kind of multimedia information. But this need involves a higher data transmission rate and modulation and codification methods capable to transmit and recover the data efficiently under mobility and environment conditions.

The higher data transmission rate means a challenge, because of the involves the study of some parameters for example the enough power to energize the passive tag, selected for this application, the tag sensitivity, which in this application, at least, it should be detectable 4

meters away from the reader, the necessary backscattered power to recover the data by the reader, also the appropriateness protocol communication and the modulation and codification methods as much the uplink as the downlink.

In the same way, meanwhile the time is essential to acquire the carrier signal and being the RFID system a time sensitive system, then, under the conditions above mention, the reader must be able to receive the backscattered signal from the tag and recover the information, even when the vehicle is running at high speed making possible the error in the data, as consequence of frequency shifting, multi-paths which alter the carrier amplitude and for that reason some wrong bits in the frame synch or subsequent commands. Thus, under these conditions and for this application the RFID technology must be functional when a higher data transmission rate is necessary if it wants to be a new distribution method of digital information when the vehicle is traveling by road, then design parameters must be revised to meet if they satisfy the requirements.

III. Tag Performance Parameters

A. Tag detection

The passive RFID (radio frequency identification) tags answer to a reader query by transponding a modulated signal according to the information already stored in the memory of a tag, operating on the road, where vehicles are passively tagged on the windshield and read at high speeds while passing through a portal is subject to multiple random phenomena.

Additionally, in the same way, a portal holding a mid-lane yagi antenna at 5.4 m height with main beam tilted down 75°. This elevation is a common in some scenarios in order to ensure successful readings in the close proximity of the reader. A minimum tag-reader distance clearance of 4 meters was considered and the readers will conduct several enquires while the mobile traverses the footprint, showed in the figure 1.

During the interrogation, the reader manages three basic operations, the Select Command, used to determine which groups of tags will respond; the Inventory Command used to identify individual tags from a group, and Access Command used once the tag has singulated and ready to be addressed with individual commands, whereas the tag manages the next states: ready, arbitrate, reply, acknowledge, open, secured and kill, in this way, the communication between the reader and the tag must be established during the short interval time when the vehicle crosses the portal where the reader is installed, considering, that, the speed of the vehicle is going from 40 Km/h to 120 Km/h, as showed in graphic of the figure 2.

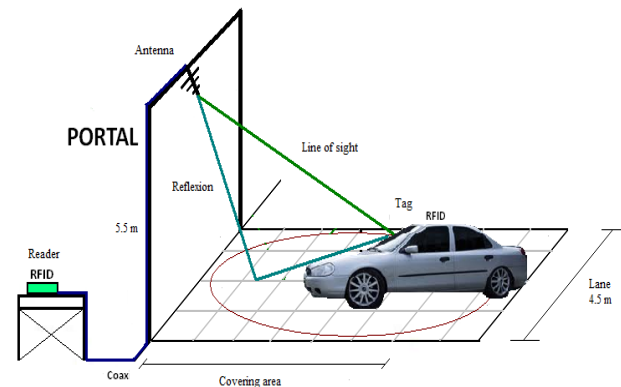


Fig. 1. The vehicle under test crossing the portal

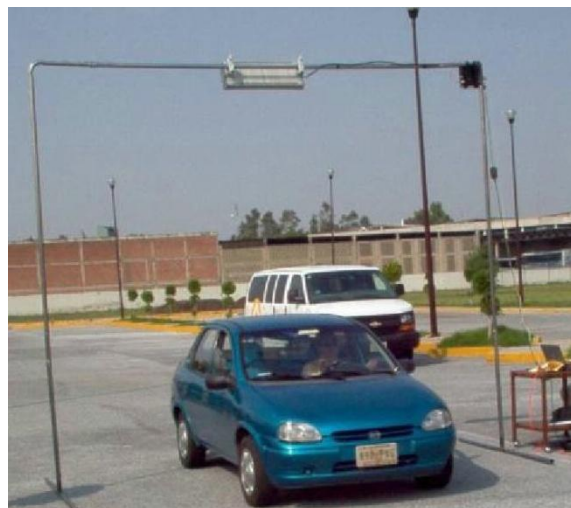


Fig. 2. Tests on road.

The test results in road casts doubts on the tag detection performance, due of the number of successful readings is lower when the vehicle crosses the portal at high speed. The possible reasons could be the Doppler effect, or the modulation and encoding methods are not robust after at certain speed, or because of the BER in the data transmission is very high and the error correction codes fail, or also the multipath should be considered as a cluster and not as single rays.

This tag detection performance at high speed represents a weaknesses of the system applied to vehicular identification, due of, if the vehicle has been stolen, probably it will be running at very high speed to avoid be detected. Actually, at this time the team is working on this problem.

B. Tag sensitivity

The sensitivity of the tag's integrated circuit imposes a physical limitation with nominal sensitivity values typically in the range from -10dBm to -18dBm depending on the manufacturer.

In the uplink direction (Tag to Reader), limitation relates the tag's efficiency of the backscattering signal which is associated with the antenna-chip coupling impedances and with the reader's sensitivity, whose typical levels are in the -80dBm to -95dBm range. In the other hand, the downlink direction (Reader to tag), the tag have no frequency selection but respond to the strongest signal, and in this application, the vehicular identification, the tag confronts the destructive or constructive contribution in the received power due the direct wave and the reflected waves from the reflection surface in accordance of the electric properties of itself. This fluctuation plays a part in the downlink, since the signal is modulated using the ASK technique, affecting the amplitude level, it doesn't matter if the data transmission is low or high.

The results show that the reader sensitivity or tag sensitivity is not a parameter to be modified if the data transmission rate is increased.

C. Transmission power

According to region's 2 international regulations, in the down link, the power transmitted by the reader's antenna to the tag's antenna will be 36dBm EIRP i.e. $P_t G_t = 4$ Watts Effective Isotropic Radiated Power. About the uplink, the power transmitted from the tag's antenna to the reader's antenna is determined by the backscattering efficiency η_{tag} , which in turn, is related to the maximum energy transfer or antenna-tag impedances matching according to equation (1) [10].

$$\eta_{tag} = \frac{P_{re-radiated} G_{tag}}{P_a} = \frac{4R_a}{|Z_a + Z_c|^2} \quad (1)$$

Factor $P_{re-radiated}$ represents the re-radiated power, G_{tag} is the tag's antenna gain, P_a is the incident power in the tag's antenna, $Z_a = R_a + jX_a$ is the tag's antenna impedance and $Z_c = R_c + jX_c$ is the tag's integrated circuit impedance. For the tag, the Effective Isotropic Radiated Power (EIRP) is determined by the product $P_{re-radiates} G_{tag}$. When the impedances are matched the maximum efficiency is ($\eta_{tag} = 1$), the $-3dB$ criteria is obtained for $\eta_{tag} = 0.7$.

The figure 3 shows coverage area reader antenna, it is possible to observe that the maximum power is focused in the middle of the area. During the tests the passive tag strapped in the vehicle was detected even in the corners of the area. In the same way, the transmission power is not a parameter to be modified if the data transmission rate is increased.

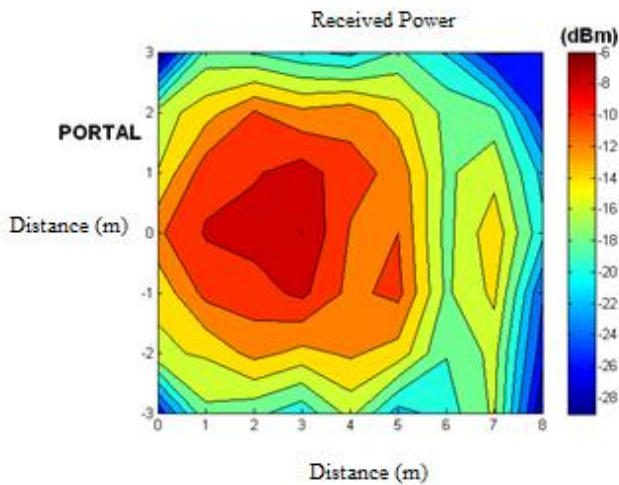


Fig. 3. Coverage area Yagi antenna reader

D. The physical layer

So the fact of the higher data transmission rate applied to vehicular identification is that the responsibility falls on the physical layer, then.

In accordance with the ISO/IEC 18000-6 standard, the reader states communication with one or more tags, modulating a RF carrier using the Double Side Band Amplitude Shift Keying, DSB-ASK, or the Single Side Band, SSB-ASK or the Phase Reversed Amplitude Shift Keying, PR-ASK with Pulse Interval Encoded, PIE, and in the other hand, the uplink, the reader instructs the tag which method of data encoding to use when sending its data back the Miller Subcarrier Encoding or the FM0 Baseband Encoding, but these encoding techniques use low data rates. Actually, UHF Gen 2 establishes for uplink communication, Tag to Reader, ASK or PSK Backscatter modulation, encoding the data using FM0 baseband with data transmission rate from 40 kbps to 640 kbps or Miller subcarrier with data transmission rate from 5 kbps to 320 kbps. In the case of the downlink communication, Reader to Tag, UHF gen 2 establishes ASK modulation encoding the data using Pulse Interval Encoding with data transmission rate from 26.7 Kbps to 128 Kbps. These data transmission rates are not enough for multimedia

applications, but they are suitable for text transmission.

The signaling interface between the reader and the tag is considered as the physical layer of the open architecture, it defines frequencies, modulation, data codification and other required parameters for the radio frequency communications. So, it is necessary bear the considerations about the digital communications system in mind, as RFID, it should carry out with a requirement set as:

- Bit rate
- Maximum acceptable bit error rate
- Maximum Band Width
- Maximum signal transmission power
- Maximum used power by the detector
- Maximum time acquisition by the detector

The bit error rate is defined as the relation between the error bits number and the transferred bits number during a time interval, and it is often expressed as error probability, P_e . From the measurements taken in field, average received power -60.76 dBm in the reader from the tag, the bandwidth 500 KHz, and the expression:

$$P_e = \frac{1}{2} e^{-\frac{E_{prom}}{4\eta}} + \frac{1}{2} Q\left(\sqrt{\frac{E_{prom}}{2\eta}}\right) \quad (2)$$

where:

E_{prom} , average energy

η , background noise

$Q(x)$, Complementary error function

We have found:

$$P_e = 6.559 \times 10^{-4}$$

E. Reader - Tag Communication Procedure in the vehicular identification

The reader - tag signaling is formed by a frame, being the first bits block, the synchrony block, formed by a length fix starting delimiter, 0 data symbol, Reader – Tag calibration symbol and the Tag – Reader Calibration symbol, followed by the query command block, in this way, starting the inventory round. The reader orders the codification option. Then, the tag antenna receives an electromagnetic field, generating a RF voltage, which, it is rectified by a diode and the resultant signal is smoothed by a capacitor creating a constant voltage. This rectified circuit as well as the demodulation circuit must be fast switching to follow the high data transmission rate.

Replying the tag uses backscatter modulation to respond to a reader. It does this by switching the reflection coefficient of its antenna (using a shunt antenna) from a matched load where the maximum reflected signal is created, to a short at the antenna terminals, which reduces the reflection. The tag must make backscatter using a fix modulation format, data codification and suitable data rate during the round inventory, which starts by a query command and ending by a select command. The tag selects the modulation format, not the data encoding after the query command, and the reader must be able to demodulate both kind of modulation, ASK or PSK.

The reader instructs the tag which method of data encoding when sending its data back. The tag must code the data using FMO, bi-phase space, reversing the phase in each symbol as a result the data has a phase reverse in the half of the symbol, or by Miller modulation, modulating a subcarrier to the bit rate, and for the command confirmation, using CRC-16 or CRC-5 in accordance with the international standard 18000-6.

This protocol must be kept for success readings during the interval time when the vehicle crosses the portal, even the high speed of the vehicle.

In the other hand, the TARI length time is equal to “0” symbol, it varies from 6.25 μ s to 25 μ s every 6.25 μ s, $\pm 1\%$ tolerance and the pulse width is aprox. 0.5TARI. The time length for the data “1” is equal to $1.5\text{TARI} \leq \text{Data“1”} \leq 2\text{TARI}$, and if the bit rate is the TARI inverse, the bit rate is equal from 160 Kbps to 40 Kbps.

From the tests in road, diverse kind of reader operation were proved, hybrid (TARI= 12.5 μ s), dense reader M=4 (TARI= 25 μ s), Dense Reader M=8 μ s (TARI=25) and Maximum Throughput (TARI= 7.5 μ s). Even when the latter showed best performance, and the “tari” is the shortest, low successful readings when the speed of the vehicle was above 100 Km/h were obtained.

Meanwhile the pulse width is shorter the bit rate increases, and if the vehicle runs at high speed, the tag detection may be successful but the data transfer fails.

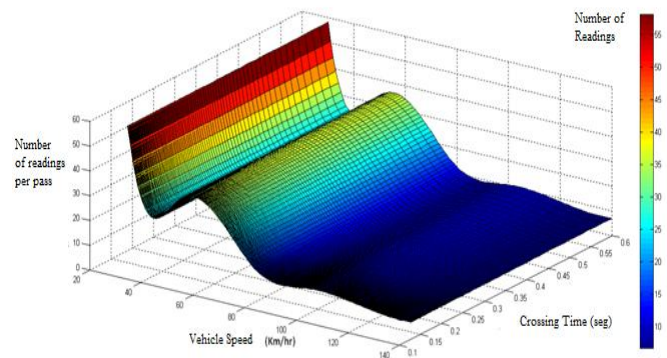


Fig. 4. Number of readings per journey on the footprint.

IV. Conclusions

This work shows some results tests applied to the RFID technology in the vehicular identification, in order to evaluate the performance. From the results this technology shows efficiency and has the qualifications for the application. Nevertheless, for other kind information, not only

text, higher data transmission is necessary, which, involves to reduce the TARI length time, changes in the electronic circuit tag, and/or modulation formats or encoding data method. The team is working on it actually.

V. References

- [1] E. Cooney, RFID+: The Complete Review of Radio Frequency Identification, Cengage Delmar Learning, 2006.
- [2] K. Finkenzeller, RFID Handbook: Radio-Frequency Identification Fundamentals and Applications, 2nd edition. New York: John Wiley & Son, 2003.
- [3] K. Finkenzeller, RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, John Wiley and Sons, West Sussex, England, 2003.
- [4] Glenn A., Lieberman G. “Effect on propagation fading and antenna fluctuations on communication systems in a jamming environment”, IEEE Transactions on Communications, vol. 10, no. 1. March, 1962
- [5] V. D. Hunt, A. Puglia, & M. Puglia, RFID - A Guide to Radio Frequency Identification, New York: John Wiley and Sons.
- [6] Kemp D., Reynolds, M.S., “Probabilistic UHF RFID tag pose estimation with multiple antennas and a multipath RF propagation model. Intelligent Robots and Systems”, IROS 2008, IEEE/RSJ International Conference on, 2008.
- [7] Lazaro A., Girbau D., Salinas D.,” Radio Link Budgets for UHF RFID on Multipath Environments.”, Antennas and Propagation, IEEE Transactions on, Volume 57, Issue 4, April, 2009.
- [8] L. M. Ng, K.. S. Leong, D. M. Hall, & P. H.Cole P. H., “A small passive UHF RFID tag for livestock identification”. IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, vol. 1, pp. 67-70, August 2007.
- [9] Nikitin P. V., Rao K. V. S, and. Lazar S., “An overview of near field UHF RFID”, IEEE RFID Conference, pp.167-174, 2007.
- [10] P. V. Nikitin and K. V. S. Rao, “Theory and Measurement of Backscattering from RFID Tag,” IEEE Antennas and Propagation Magazine, Vol. 48, no. 6, pp 212-218, December 2006.
- [11] Turner L., Mickle M. H. (2007). Overview primer on near field UHF versus near HF RFID tags. International Journal on RFID Technology and Applications, vol. 1, no. 3.

Diseño de Antena para Etiqueta en aplicaciones RFID

Roberto Leyva Hernández, Rubén Ignacio Cruz Baños, Salvador Ricardo Meneses González

Instituto Politécnico Nacional Escuela Superior de Ingeniería Mecánica y Eléctrica, Campus Zacatenco. Edif. 4.
Primer piso, Cubículo 4140. Col. Lindavista, C. P. 07738, México, D. F.
beto90_6@hotmail.com, rnacho_2208@live.com.mx, rmeneseg@ipn.mx

Resumen

Este trabajo fue desarrollado por estudiantes como parte del Programa Institucional de Formación de Investigadores del IPN. El trabajo muestra el diseño de una Antena para la etiqueta de un sistema de Identificación por Radiofrecuencia (RFID) en la banda de UHF. Tomando como base teoría de radiadores electromagnéticos, sistemas de acoplamiento inductivo y técnicas de reducción de tamaño se diseñó una antena pasiva que trabaja a los 915 MHz, y mediante su simulación se obtuvieron sus características y comportamiento en el espacio libre.

Palabras Clave: Diseño de antena RFID, Dipolo serpenteado, Acoplamiento Inductivo, Etiqueta pasiva, UHF.

I. Introducción

Un sistema RFID (Radio Frequency IDentification) es la tecnología inalámbrica que nos permite, básicamente, la comunicación entre un lector y una etiqueta. Esta tecnología es un método automático de identificación, basado en el almacenamiento y recuperación remota de datos [1, 2].

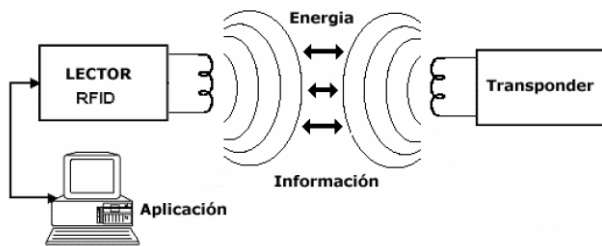


Fig. 1. Esquema de un sistema RFID.

El funcionamiento del sistema (ver Fig. 1) es de la siguiente manera: el lector envía una serie de ondas de radiofrecuencia al tag, que son captadas por la microantena de éste [3]. Dichas ondas activan el microchip, el cual, a través de la microantena y mediante ondas de radiofrecuencia, transmite al lector la información que tengan en su memoria.

Finalmente, el lector recibe la información que tiene el tag y lo envía a una base de datos en la que previamente se han registrado las características del producto o puede procesarlo según convenga a cada aplicación [3].

Un tag, transponder o etiqueta electrónica contiene un microchip y una antena, que puede adherirse a cualquier producto.

Los tags pasivos para UHF de 915 Mhz suelen acoplarse por ondas electromagnéticas con un frente de onda esférico a la antena del lector y utilizan antenas tipo dipolo $\lambda/2$. Sin embargo, para un mejor ajuste a las características del circuito integrado (con alta impedancia de entrada) se utiliza un bucle corto de aproximadamente $\lambda/4$ como estructuras inductivas para una mejor alimentación [4].

La comunicación entre el lector y el tag tiene determinadas características de alcance, velocidad y seguridad según el rango de frecuencia, el tipo de antenas utilizadas, el tipo de etiquetas y demás parámetros que se pueden configurar para una aplicación u otra.

1. Parámetros Importantes para el Diseño de la Antena para Etiqueta RFID

El éxito en el funcionamiento de un sistema RFID recae en la eficiencia de la antena de la etiqueta, debido a que ésta debe tener alta ganancia y una correcta orientación con respecto al lector. Para su diseño es primordial considerar diferentes parámetros como son: rango de lectura, inductancia, ganancia, coeficiente de acoplamiento y resonancia de acuerdo a la aplicación a la que el sistema está orientado.

Rango de lectura

El rango de lectura depende de la potencia de transmisión y, en el caso de los tags pasivos, también los requerimientos de energía de los mismos. El rango efectivo de lectura depende también del factor de absorción del material al cual va unido el tag [3].

El tamaño del tag también juega un papel importante en el rango de lectura.

Cuanto menor es el tag, menor es el área de captura de energía, por lo que menor es el rango de lectura. Un diseño adecuado del sistema, la optimización de la potencia del lector, la orientación de la antena y una colocación óptima del tag ayudan a superar estas limitaciones.

Orientación de la etiqueta

La orientación de la antena de la etiqueta con respecto a la antena del interrogador influye en el rango de lectura. Cuando la onda electromagnética está polarizada linealmente, la antena del tag debe estar orientada en la misma dirección que la del interrogador para permitir la máxima recepción de energía. La situación de peor caso se da cuando la orientación entre ambas antenas forma un ángulo recto. Si la onda electromagnética no está polarizada linealmente no importa la orientación que tenga la antena de la etiqueta. Por ejemplo, si empleamos una onda electromagnética polarizada circularmente podemos emplear cualquier orientación para el tag [3].

Principios físicos de los sistemas RFID

La inmensa mayoría de los sistemas RFID operan de acuerdo con el principio de acoplamiento inductivo, por tanto comprender los procedimientos de transferencia de datos y alimentación requiere un conocimiento detallado de los principios físicos del magnetismo. Los campos electromagnéticos son usados por los sistemas que operan a frecuencias por encima de los 30 MHz [3].

Inductancia L

Cualquier circuito es atravesado por un flujo creado por el mismo y que debe ser proporcional a la intensidad que lo recorre como vemos a partir de la Ec. 1. [8]

$$\Psi = \sum_N \Phi_N = N \cdot \Phi = N \cdot \mu \cdot H \cdot A \quad (1)$$

Definimos como inductancia L, la relación entre el flujo total y la corriente que atraviesa el conductor [3].

$$L = \frac{\Psi}{I} = \frac{N \cdot \Phi}{I} = \frac{N \cdot \mu \cdot H \cdot A}{I} \quad (2)$$

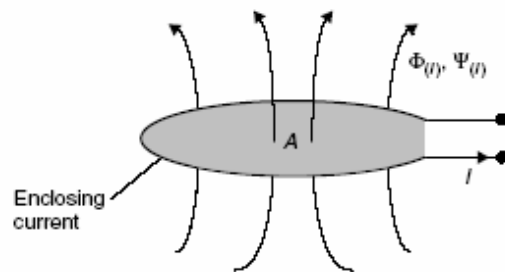


Fig. 2. Definición de Inductancia L.

La inductancia es una de las características variables de este tipo de conductores.

Inductancia Mutua M

La inductancia mutua se produce por la proximidad de dos conductores. La corriente que atraviesa al conductor induce un flujo magnético en el otro y al inverso. La magnitud del flujo inducido depende de las dimensiones geométricas de ambos conductores, la posición de un conductor respecto al otro y las propiedades magnéticas del medio [3]. Para dos conductores de áreas A_1 y A_2 , e I_1 la corriente que circula por la primera espira vemos:

$$M_{21} = \frac{\Psi_{21}(I_1)}{I_1} = \oint_{A_2} \frac{B_2(I_1)}{I_1} \cdot dA_2 \quad (3)$$

Por definición tenemos que la inductancia mutua es igual:

$$M = M_{12} = M_{21} \quad (4)$$

La inductancia mutua siempre está presente entre dos circuitos electrónicos, en este principio físico es en el que se basa el acoplamiento inductivo de los sistemas RFID.

Coefficiente de acoplamiento k

Si la inductancia mutua describía cualitativamente el flujo creado por la corriente que circula por otra espira, el coeficiente de acoplamiento realiza una predicción cualitativa de la inducción creada entre dos espiras independientemente de las dimensiones geométricas de los conductores [3].

$$k = \frac{M}{\sqrt{L_1 \cdot L_2}} \quad (5)$$

Tenemos que $0 \leq k \leq 1$, por lo que en los casos extremos:

$k=0$: No hay acoplamiento debido a la gran distancia no hay acción del campo magnético.

$k=1$: Acoplamiento total. Las dos espiras están sometidas al mismo Φ .

Resonancia

El voltaje inducido en la antena del transponder es usado como alimentación necesaria para el chip en su proceso de almacenamiento de datos en memoria. Para mejorar la eficiencia un capacitador C_2 se conecta en paralelo con el loop del transponder L_2 , de manera que forma un circuito paralelo resonante con una frecuencia resonante que es la frecuencia de operación del sistema de RFID [3]. La frecuencia resonante se puede calcular con la Ec. 6.

$$f = \frac{1}{2\pi\sqrt{L_2 \cdot C_2}} \quad (6)$$

En la práctica existe un capacitador parásito en paralelo C_p por lo que el valor del capacitador sería C'_2 , como vemos en la Ec. 7

$$C'_2 = \frac{1}{(2\pi f)^2 L_2} - C_p \quad (7)$$

Ganancia

La ganancia de una antena es la relación (cociente) de potencias por unidad de superficie, entre una antena dada y una antena isótropa que se toma como referencia, ambas alimentadas con la misma potencia [3,8]:

$$G = \frac{\left(\frac{P}{S}\right)_{ant}}{\left(\frac{P}{S}\right)_{iso}} \quad (8)$$

La potencia por superficie unitaria transportada por una onda electromagnética es:

$$\left(\frac{P}{S}\right)_{ant} = \frac{1}{2} c \epsilon_0 E_{\theta}^2 \simeq \frac{1}{120\pi} E_{\theta}^2 \quad (9)$$

La potencia por superficie unitaria de una antena isotrópica alimentada con la misma potencia es:

$$\left(\frac{P}{S}\right)_{iso} = \frac{\frac{1}{2}R_{serie}I_0^2}{4\pi r^2} \quad (10)$$

En el caso de un dipolo corto, cuando se rempazan los valores, el resultado final es:

$$G = \frac{\pi\left(\frac{L}{\lambda}\right)^2}{\epsilon_0 c \frac{2\pi}{3\epsilon_0 c} \left(\frac{L}{\lambda}\right)^2} \quad (11)$$

Técnica de Serpenteado

En la banda de UHF, especialmente debajo de 1 Ghz, meandro o serpenteo de antenas de línea (MLA, por sus siglas en inglés) son una elección atractiva si se tiene como propósito reducir el tamaño de las etiquetas. El serpenteado consiste [5] en doblar los elementos produciendo una configuración de cables con reactancias tanto capacitivas como inductivas que mutuamente se cancelan. Por lo tanto, la resonancia se produce a una frecuencia mucho menor que las antenas rectas a expensas de tener un ancho de banda reducido y baja ganancia [6].

II. Desarrollo

Se necesita una antena dipolo $\lambda/2$ a una frecuencia de 915 MHz, donde la longitud de la antena se calcula de acuerdo a la siguiente ecuación:

$$\lambda = \frac{c}{f} \quad (12)$$

Donde:

λ = longitud de onda
 c = velocidad de la luz

Quedando la longitud de la antena aproximadamente de 16.4 cm.

Por otra parte, para lograr una máxima transferencia de energía [7] y como se había mencionado en la introducción, de acuerdo a la alta impedancia que presenta el circuito integrado se requiere que el dipolo tenga un acoplamiento

inductivo con un loop de $\lambda/4$ y que sería el que directamente esté conectado con el circuito integrado.

Debido a que la longitud del dipolo es demasiado grande y poco práctica como antena para etiqueta, se necesita emplear técnicas de reducción de tamaño, propiamente para nuestro caso usamos el serpenteado. La antena entonces, quedó aproximadamente de 8cm de largo por 1 cm de altura (ver Fig. 3) para no sobrepasar los estándares en América del Norte que son de 9 cm de largo por 2 cm de ancho.

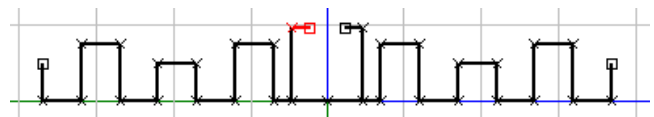


Fig. 3. Antena Dipolo con serpenteado y bucle

El manejo de herramientas como la simulación para el análisis de la antena permite tener una idea clara de su comportamiento y características sin necesidad de construirlo y hacer mediciones físicamente. Además de tener la flexibilidad de hacer pruebas y cambiar sus atributos fácilmente. Así se podrán reducir costos y tiempo para su diseño, y la construcción de la antena definitiva tendrá una mayor confiabilidad.

La simulación se llevo a cabo en el programa MMANA-GAL. Se propusieron los siguientes parámetros para la simulación de la antena:

- Frecuencia de operación: 915 MHz.
- Longitud de dipolo: 16.4 cm
- Perímetro de loop: 3.2 cm

En base a la simulación, se observo que el patrón de radiación deseado se obtuvo con una separación de 0.1 mm entre el loop y el dipolo para generar un buen acoplamiento (ver Fig. 5).

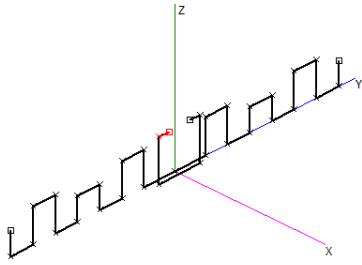


Fig. 5. Vista 3D de Antena Dipolo con serpenteado y bucle.

De acuerdo al patrón de radiación eléctrico (ver Fig. 6) y magnético (ver Fig. 7) de la antena diseñada se observa que sigue el mismo comportamiento de una antena tipo dipolo, como se había esperado, así entonces, también radia en su campo total en forma de “dona” como muestra la figura 8, o desde una vista frontal como ilustra la figura 9.

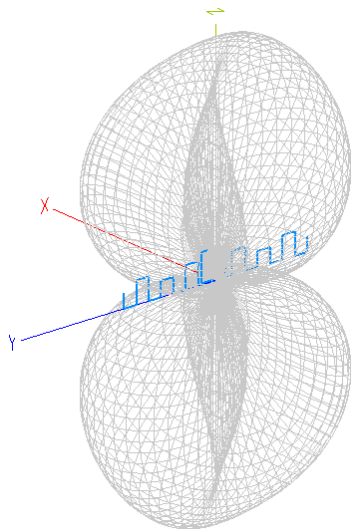


Fig. 6. Diagrama 3D de Campo Lejano V.

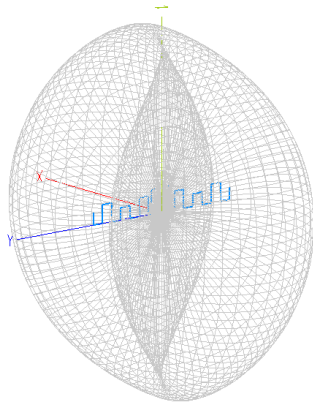


Fig. 7. Diagrama 3D de Campo Lejano H.

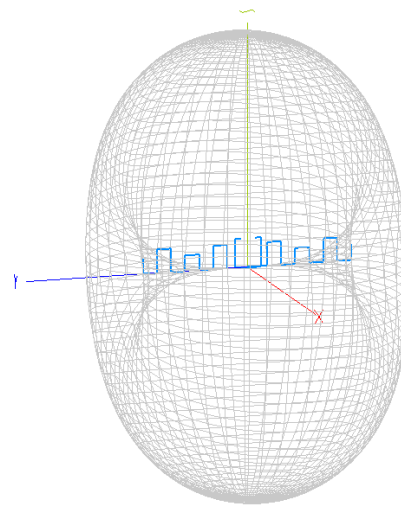


Fig. 8. Diagrama 3D de Campo Lejano Total.

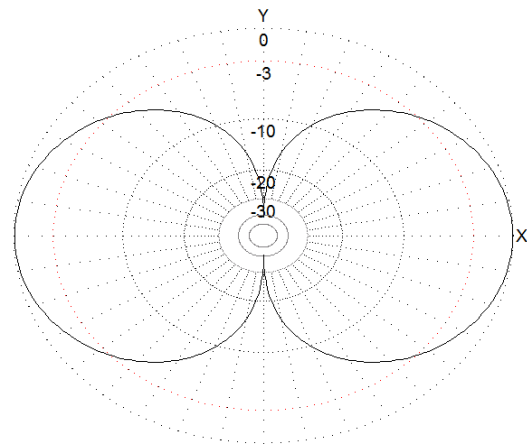


Fig. 9 Diagrama 2D de Campo Lejano Total.

Para su construcción se eligieron materiales de fácil acceso y los más convenientes de acuerdo a la aplicación. Por lo tanto, la antena fue dibujada en una placa fenólica para ser un conductor de cobre de 1 mm de grosor (ver Fig. 9 y 10); y la separación entre el dipolo y el loop es un pedazo de acetato que sirve como dieléctrico de 0.1 mm.

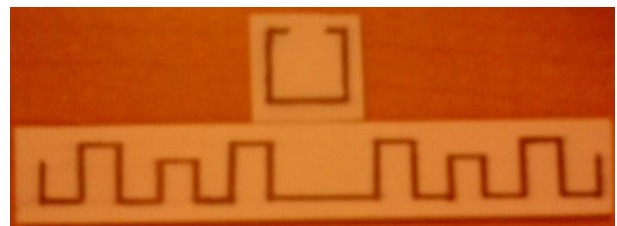


Fig. 10. Antena dipolo y loop en placa fenólica.

III. Resultados

La impedancia de la antena se acopla de una manera favorable a la impedancia del circuito integrado, indicando que la antena tendrá un óptimo desempeño gracias a que cumple con una alta ganancia, un patrón de radiación correcto, y la variación entre el loop y el dipolo mejora aún más el rendimiento.

IV. Discusión

Actualmente, este trabajo todavía está en proceso, aún faltan las mediciones con el chip conectado a la antena y el empleo de la antena diseñada en un sistema real de RFID.

La impedancia de la antena variará de acuerdo al Chip de la etiqueta, por lo que se deberán hacer ajustes para el acoplamiento y lograr una alta ganancia.

V. Conclusiones

La variación óptima de las dimensiones del serpenteo del dipolo lograron una reducción considerable en el tamaño de la antena y conservando las mismas características que el dipolo simple.

El acoplamiento inductivo que se establece entre el dipolo y el loop genera una mayor transferencia de potencia entre la antena y el Chip, logrando una mayor eficiencia en el sistema.

VI. Referencias

- [1] E. Cooney, RFID+: The Complete Review of Radio Frequency Identification, Cengage Delmar Learning, 2006.
- [2] Antenna Design for UHF RFID tags: A review and a practical application. P.VNikitin, K. V. S Rao. IEEE. 12 December 2005.
- [3] María Ciudad Herrera, Estudio Diseño y Simulación de un sistema RFID basado en EPC.
- [4] K. Finkenzeller, RFID Handbook: Radio-Frequency Identification Fundamentals and Applications, 2nd edition. New York: John Wiley & Son, 2003.
- [5] T. J. Warnagiris and T. J. Minardo, "Performance of a meandered line as electrically small transmitting antenna," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 1797–1876, Dec. 1998.
- [6] G. Marrocco, A. Fonte, and F. Bardati, "Evolutionary design of miniaturized meander-line antennas for RFID applications," in *Proc. IEEE AP-S Int. Symp.*, vol. 2, San Antonio, TX, 2002, pp. 362–365.
- [7] Nikitin, P.V. "Performance limitations of passive UHF RFID systems", IEEE Antennas and Propagation Society International Symposium, pp. 1011-1014, 2006.
- [8] C. A. Balanis, Antenna Theory Analysis and Design, John Wiley & Son, 1997.

VII. Autores

Estudiante PIFI Roberto Leyva Hernández, cursa actualmente séptimo semestre de Ingeniería en Comunicaciones y Electrónica en el IPN ESIME Zacatenco.

Estudiante PIFI Rubén Ignacio Cruz Baños, cursa actualmente séptimo semestre de Ingeniería en Comunicaciones y Electrónica en el IPN ESIME Zacatenco.

M. en C. Salvador Ricardo Meneses González, Director de proyecto PIFI y Profesor investigador de la ESIME Zacatenco.

Sistema de localización y seguimiento de unidades móviles en el Distrito Federal

***Hermes Francisco Montes Casiano, **José Jaime López Rabadán, ***Ulises Velez Saldaña**

* Escuela Superior de Cómputo del Instituto Politécnico Nacional, Av. Juan de Dios Bátiz s/n esquina Miguel Othón de Mendizabal. Unidad Profesional Adolfo López Mateos. Col Lindavista C.P 07738. México D.F, (044) 55 34 06 18 46, hermes.montes@gmail.com.

**Escuela Superior de Cómputo del Instituto Politécnico Nacional, Av. Juan de Dios Bátiz s/n esquina Miguel Othón de Mendizabal. Unidad Profesional Adolfo López Mateos. Col Lindavista C.P 07738. México D.F, 57 29 6000 ext. 52004, rabadanlorj@gmail.com.

*** Escuela Superior de Cómputo del Instituto Politécnico Nacional, Av. Juan de Dios Bátiz s/n esquina Miguel Othón de Mendizabal. Unidad Profesional Adolfo López Mateos. Col Lindavista C.P 07738. México D.F, (044) 57 29 6000 ext. 52005, ulises.velez@gmail.com.

Resumen

Existe un gran número de organizaciones gubernamentales y empresariales que se apoyan de unidades móviles para realizar sus actividades de entrega y recolección de objetos, supervisión y vigilancia, por citar las más comunes. Dichas organizaciones necesitan corroborar si sus unidades móviles cumplen con los recorridos asignados, supervisar su uso y trayectoria; con el objetivo de tomar decisiones ante diversas circunstancias. Éste trabajo se enfocó en el diseño e implementación de un *Localizador Automático de Vehículos* (AVL, por sus siglas en inglés) en Web utilizando *Sistemas de Información Geográfica* y dispositivos móviles *iPhone* que permita verificar las trayectorias y regiones de recorridos, mediante la definición de geocercas, a través de la Internet. Además, proveer mecanismos para la generación de historiales de posicionamiento y mapas temáticos de recorridos.

Palabras Clave: AVL, GPS, GIS, Dispositivos Móviles.

I. Introducción

En los últimos años, el crecimiento y la popularidad de la telefonía móvil fue más significativa de lo previsto. El incremento en la capacidad de cómputo de los procesadores, las mejoras en los sistemas operativos modernos para dispositivos móviles y la popularidad de la banda ancha móvil, hacen de los teléfonos celulares el mejor candidato como dispositivo de cómputo móvil; capaces de realizar, incluso, sofisticadas aplicaciones científicas [1].

Se clasifica como “*Sistema de Información Geográfica*” (GIS, por sus siglas en inglés) aquellos sistemas que gestionan información georeferenciada. Un GIS se define como un conjunto de herramientas para: recolectar, almacenar, obtener, transformar y desplegar información espacial de la Tierra para propósitos específicos; son utilizados en ingeniería civil, geología, inteligencia militar, control de tráfico vial, entre otros [2].

Una *Geocerca* permite describir una zona geográfica (es decir, una barrera geográfica) de un punto de interés, por ejemplo, en términos de un círculo o un polígono, y asociar a esa área eventos, acciones e información. Los eventos comunes de localización son entrar y salir de la zona geográfica delimitada por la barrera geográfica o permanecer dentro o fuera una cierta cantidad de tiempo [3].

Por otro lado, el Sistema de Posicionamiento Global (GPS, por sus siglas en inglés) es una tecnología madura para la localización precisa en ambientes al aire libre y de mayor uso en la actualidad [4]. El GPS tiene la capacidad de proveer información de posición y tiempo a los usuarios del sistema. El GPS Asistido (AGPS) es una solución híbrida que utiliza información provista por los satélites GPS y la red celular [5]. El AGPS se auxilia de fuentes externas, como servidores de asistencia o referencias de red, para ayudar al receptor GPS a

disminuir el tiempo de respuesta y el margen de error [6][7].

El seguimiento de vehículos es una de las tecnologías más importantes de los Sistemas Inteligentes de Transporte (ITS). Desempeña funciones de seguimiento en línea de los vehículos en movimiento, describe su estado de movimiento y trayectoria [8].

Éste trabajo propone diseñar e implementar un sistema que permita localizar y seguir, de forma constante, las unidades móviles de una organización dentro de la demarcación del Distrito Federal. Además, proveer los medios necesarios para la generación de historiales de posicionamiento y emisión de alertas.

Trabajos relacionados

A continuación se describen dos implementaciones realizadas en el ámbito de la localización móvil y cuya funcionalidad se enfoca en la localización y seguimiento de vehículos y personas respectivamente:

- *Sistema de seguimiento vehicular en tiempo real* [9]. El diseño e implementación del sistema incluye la adquisición y transmisión de la ubicación del vehículo, junto con información de encendido y estado de las puertas, a la estación de monitoreo o servidor. El sistema también provee una interfaz Web para desplegar en un mapa la información del vehículo transmitida. EL sistema consta de los 2 componentes principales.
 - *Unidad en el vehículo (IVU, por sus siglas en inglés)*: Consta de un receptor GPS para capturar la ubicación y velocidad del vehículo, un CPU para procesar la información provista por el GPS, monitorear el estado de las puertas y controlar el estado de encendido del automóvil. El CPU es el encargado de controlar el intercambio información con el Servidor de Seguimiento a través de un transmisor/receptor GSM.
 - *Servidor de Seguimiento*: esta compuesto de un módem GSM para el intercambio de información con el IVU a través de mensajes SMS. El servidor de seguimiento se instaló en un servidor Apache utilizando

MySQL como sistema gestor de base de datos. La visualización de la posición de vehículos se implementó utilizando el servicio de *Google Maps*.

- *Sistema de seguimiento de personas mayores basado en AGPS* [4]. Este prototipo se desarrolló con la finalidad de mejorar la calidad de vida de las Personas Mayores de la ciudad de Hong Kong, debido a que éstas en su mayoría viven de manera independiente. Provee una precisión en la localización de las personas en un rango de 10 a 20 m cuando están en movimiento y 138 m en reposo. El sistema consta de los siguientes componentes:
 - *Terminal AGPS*: la terminal AGPS, además tiene integrado un chip para comunicación a través de 3G.
 - *Servidor Móvil de Localización*: es un servidor de Localización Segura de Usuarios (SUPL, por sus siglas en inglés), fabricado por Broadcom, que provee un único punto de contacto a todas las terminas AGPS.
 - *Estación de referencia GPS*: es un receptor GPS y un procesador ARM7, tiene la finalidad de asistir a las terminales AGPS para reducir el margen de error.
 - *Servidor de Base de Datos*: la base de datos del prototipo se instaló en MySQL, utilizando su extensión para información geográfica.
 - *Servidor Web*: componente encargado de brindar acceso a los clientes Web, realizar el control de usuarios y despliegue de la ubicación de las terminales AGPS.
 - *Cliente Web*: Provee a los usuarios finales acceso a la aplicación y es por medio de éste que se realiza el seguimiento de las personas mayores.

II. Localización y seguimiento de unidades móviles

Existe un gran número de organizaciones gubernamentales y empresariales que se apoyan de unidades móviles para realizar sus actividades de entrega y recolección de objetos, supervisión y vigilancia. Éstas organizaciones necesitan:

- Corroborar si sus unidades móviles cumplen con las actividades de recorridos asignadas,

debido a que es importante verificar que el personal cumpla con sus labores adecuadamente.

- Supervisar el uso de las unidades móviles, debido a los vicios adoptados por los conductores y al uso de las unidades para actividades personales.
- Conocer la trayectoria que siguen las unidades móviles, debido que les permite supervisar el uso de las unidades y tomar decisiones ante diversas circunstancias.

El trabajo propuso resolver las necesidades de localización y seguimiento proveyendo las siguientes funciones generales:

- *Asignación de regiones de recorridos:* Permite la asignación de geocercas a las unidades móviles.
- *Configuración de políticas:* Facilita la configuración del periodo de localización de las unidades móviles y el tiempo para emitir una alerta de desplazamiento nulo.
- *Localización y seguimiento:* Permite localizar constantemente a las unidades móviles endesplazamiento, visualizando su ubicación y trayectoria en un mapa georeferenciado.
- *Seguimiento:* Se puede visualizar en línea la trayectoria que tiene una unidad móvil o revisar el historial de la trayectoria que realizó la unidad.
- *Emisión de alertas:* El sistema notifica cuando una unidad móvil abandona la geocerca asignada o nose desplace durante un intervalo de tiempo mayor al permitido.
- *Catálogos:* Permite realizar la administración (altas, bajas y modificaciones) de la entidades del sistema, las cuales son: Unidades, Usuarios, Conductores y Geocercas.

III. Diseño arquitectónico

La arquitectura del sistema fue diseñada para que éste operara a través de la Internet (Ver Fig. 1). Además, fue necesario el análisis y diseño los tres componentes principales de ésta solución: Servidor de Aplicación, Cliente Web y Cliente Móvil, ver Fig. 2.

Servidor de Aplicación

Componente de la arquitectura que funge como intermediario entre los clientes Web y Clientes Móviles. Es el encargado de recibir y almacenar la información de posicionamiento de los dispositivos en localización. Por otro lado, es el componente al cuál los Clientes Web le solicitan la posición geográfica de los dispositivos de interés para visualizarla sobre un mapa georeferenciado.

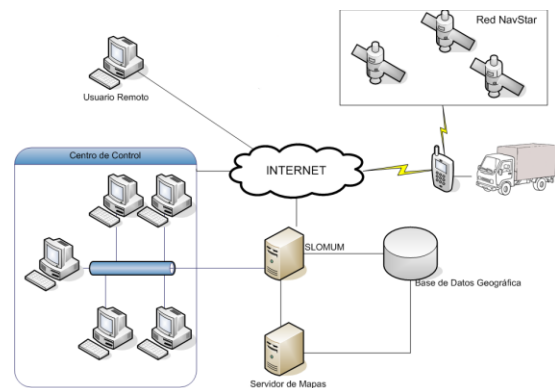


Fig. 1. Arquitectura de comunicación del sistema

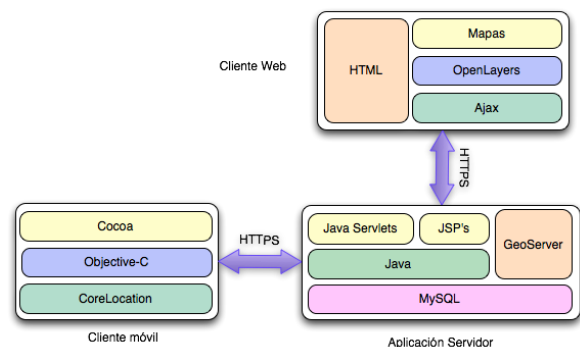


Fig. 2. Diagrama arquitectónico de bloques

La implementación se realizó utilizando una arquitectura en 3 capas, ver la Fig. 3. la generación del contenido web, lógica de negocios y acceso a datos, las cuales se estructuraron de la siguiente manera:

- *Vistas:* en esta capa se encuentra la implementación de las interfaces de usuario que se visualizan en el cliente web y las interfaces de comunicación con los clientes móviles. Los componentes de esta capa se desarrollaron con JSP's y Servlets.
- *Controladores:* esta capa esta formada por las clases que abstraen la lógica de negocios de un

AVL. En esta capa se encuentra la implementación del mecanismo de posicionamiento, el cual consiste en el análisis de la ubicación geográfica de las unidades móviles y emisión de alertas; la automatización del proceso de seguimiento de las unidades y la generación de los mapas temáticos con base en el historial de posicionamiento de las unidades. Además, están implementados los controladores de las altas, bajas y modificaciones de las entidades del sistema.

- *Acceso a datos:* en esta capa se implementan las clases que abstraen las entidades propias de un AVL. Por otro lado, contiene la funcionalidad de acceso a los datos del sistema; esta capa es la encargada de interactuar con MySQL, a través del conector JDBC.

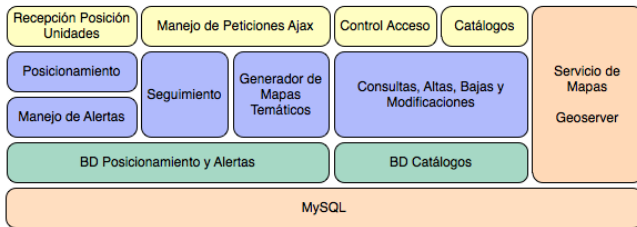


Fig. 3. Arquitectura en capas de la aplicación Web.

En el Servidor de la Aplicación está almacenada la cartografía de calles de la Ciudad de México, la cual se provee a través de un servicio web de mapas. El servicio web se configuró en el servidor de mapas *GeoServer*, el cual es un servidor de mapas de código abierto desarrollado en Java que permite a sus usuarios compartir y editar información geoespacial.

Cliente Web

El Cliente Web es el componente por medio del cual se despliega la posición geográfica de los dispositivos en localización. Después de que el dispositivo móvil captura y envía la posición al servidor, el Servidor de la Aplicación provee esta información a los clientes web para desplegarla. La arquitectura del Cliente Web, ver Fig. 4, está distribuida en tres capas: en la capa superior se encuentran las vistas del sistema en formato

HTML que se gestionan de forma dinámica, en la capa intermedia la gestión y manipulación de mapas implementada utilizando OpenLayers y en la capa inferior las peticiones asíncronas realizadas utilizando Ajax.

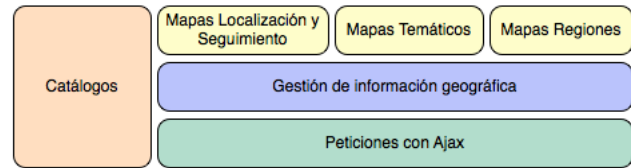


Fig. 4. Arquitectura en capas del Cliente Web

Cliente Móvil

Se denomina Cliente Móvil a los teléfonos iPhone en localización, los cuales se comunican constantemente con el Servidor de la Aplicación, utilizando el mecanismo *polling*, a través de HTTPS para reportar su ubicación geográfica.

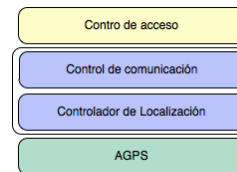


Fig. 5. Arquitectura en capas del Cliente Móvil.

La arquitectura de la aplicación en el iPhone está distribuida, de igual forma, en 3 capas, ver Fig. 5. La capa de las vistas o interfaz de usuario se implementó utilizando el *Framework Cocoa*. La lógica de negocios, es decir, la obtención constante de la posición geográfica y su correspondiente envío al servidor de la aplicación se hizo en *Objective-C*. Para obtener la posición geográfica del móvil se programó la funcionalidad utilizando los servicios provistos por la capa *CoreLocation*.

IV. Implementación de la funcionalidad

En esta sección se describe la implementación de la funcionalidad principal del sistema, la localización y seguimiento de los vehículos, las alertas que se emiten en determinadas circunstancias y el mecanismo de generación de historiales de posicionamiento.

Localización y seguimiento

El proceso de localización y seguimiento de los vehículos se divide en dos bloques:

- *Reportar/Almacenar posición*: inicia en el momento en que el conductor del vehículo, en la aplicación móvil, ingresa al sistema. A partir de ese momento se realiza lo siguiente:
 1. El dispositivo móvil reporta su posición geográfica al servidor de la aplicación invocando un *Servlet*. La actualización de la posición se realiza a través de el mecanismo de actualización *push* del móvil al servidor, hasta que el conductor finalice su sesión en la aplicación móvil.
 2. La posición, una vez recibida en el servidor, se somete al análisis de emisión de alertas.
 3. La posición se almacena con base en los datos del sistema en el servidor.
- *Consulta/Actualización de la ubicación de los vehículos*: los usuarios del sistema, a través de la aplicación web, pueden observar la ubicación y desplazamiento de los vehículos. Una vez que se solicita localizar un vehículo, automáticamente se actualizará en pantalla la ubicación del mismo.
 1. El cliente web solicita al servidor de aplicación la posición más reciente reportada por la unidad a localizar.
 2. El servidor devuelve: la ubicación, fecha y hora, conductor y alerta, en caso de que exista, del vehículo.
 3. El cliente procesa la respuesta y visualiza en la interfaz la ubicación del vehículo solicitado.

Emisión de alertas

El sistema tiene la capacidad de emitir dos tipos de alertas: *desplazamiento nulo* y *región abandonada*. La primera se refiere a emitir una alerta si un vehículo no se ha desplazado durante un determinado intervalo de tiempo. La segunda, se emite si la posición de un vehículo no se encuentra dentro de la geocerca que le fue asociada.

Una alerta de intervalo de tiempo se emite cuando un vehículo no ha tenido desplazamiento

significativo durante un determinado intervalo de tiempo definido. En el algoritmo 1 se muestra la función implementada para emitir o no la alerta. Donde se sugiere que $R \geq meAGPS$ ($meAGPS$, máximo error de tecnología AGPS y R representa el desplazamiento mínimo que debe tener una unidad móvil) debido a que la tecnología de localización AGPS tiene un margen de error de hasta 50 m.

Algoritmo 1: Emisión de alerta de tiempo

Entrada: Arreglo P con las n posiciones reportadas dentro del intervalo de tiempo (Δt) y la distancia R que indica un desplazamiento significativo.

Salida: *true* si se emite la alerta o *false* en caso contrario.

ALERTA_TIEMPO(P,R)

```

alerta ← true;
for  $i \leftarrow 1$  to  $i = n - 1$  OR alerta = false do
    alerta ← alerta AND  $|P[0].distancia(P[i])| < R$ ;
end for
return alerta;
```

Para verificar si un dispositivo móvil ha abandonado la región geográfica que le fue asignada para sus recorridos, cada ocasión que se reporta la posición de un vehículo se verifica si ésta se encuentra dentro o no de la región asociada, ver algoritmo 2.

Algoritmo 2: Emisión de alerta de región abandonada

Entrada: Polígono G de la geocerca y posición actual p .

Salida: *true* si se emite la alerta o *false* en caso contrario.

ALERTA_GEOCERCA(P,R)

```

contiene ← G.contiene(p);
if contiene = true then
    alerta ← false;
else
    alerta ← true;
end if
return alerta;
```

Historial de posicionamiento

Para la generación del historial de posicionamiento de un vehículo, el usuario debe proporcionar el intervalo de tiempo para el que desea generarlo. Las fechas tienen el formato *dd/mm/aaaahh:mm:ss*. En el algoritmo 3 se muestra el proceso que se realiza en el servidor de aplicación para generación de un historial de posicionamiento. Una vez que se genera el

historial de posicionamiento en el servidor es enviado al cliente web, mismo que lo interpreta para visualizarlo en la cartografía de calles de la Ciudad de México y de igual forma en una tabla.

Algoritmo 3: Generación de historial.

Entrada: Identificador del vehículo (v), fecha inicial (f_i) y fecha final (f_f)

Salida: *historial* con la información de los recorridos del vehículo, *null* si no hay recorridos en el intervalo indicado o si existe un error en el mismo.

```

HISTORIAL_POSICIONAMIENTO ( $v, f_i, f_f$ )
  if  $f_i < f_f$  then
     $S \leftarrow$  RecorridosVehiculo( $v, f_i, f_f$ );
    if  $S.length() = 0$  then
      return null;
    else
      for  $i \leftarrow 1$  to  $i = n$  do
        trayectoria[ $i$ ]  $\leftarrow$  obtenerPosiciones( $S[i], f_i, f_f$ );
        historial.agregar(trayectoria[ $i$ ]);
      end for
    end if
  else
    return null;
  end if
return historial;

```

V. Pruebas de funcionamiento

Al finalizar el desarrollo, el sistema se sometió a estrictas pruebas de funcionalidad con el objetivo de verificar el cumplimiento de los requerimientos funcionales y no funcionales establecidos. En la presente sección se describen las pruebas de funcionamiento más significativas evaluadas, para lo cual se tomó como base el seguimiento realizado por el sistema al recorrido de un vehículo.

En las Fig. 6 y 7 se muestra el seguimiento de un vehículo, en la primera se resalta la funcionalidad de localización y seguimiento, mientras que en la segunda la emisión de alertas. El seguimiento de la unidad, que se muestra en la Fig. 6, se realizó un día en que las condiciones climáticas y ambientales de la Ciudad de México fuesen desfavorables al AGPS, con la finalidad de evaluar la precisión de posicionamiento del dispositivo móvil. Sin embargo, a pesar de que el AGPS estaba configurado para proveer el servicio de localización de máxima precisión, las posiciones obtenidas durante el primer minuto de

muestreo tienen un margen de error de 500 m. Posteriormente, una vez superado el primer minuto de muestreo el margen de error se reduce a un margen ≤ 15 m, el cual es aceptable en el dominio de aplicación del sistema. Por otro lado, en la Fig. 6 se muestra la funcionalidad de alertas de región abandona, se resalta las alertas emitidas por el sistema cuando la ubicación reportada por el dispositivo en el vehículo se encuentra fuera de la región de recorrido establecida, es decir, de su geocerca.

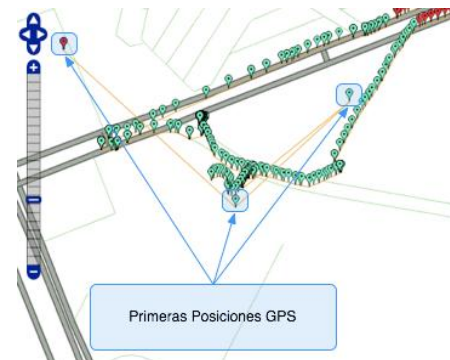


Fig. 6. Error de posicionamiento.



Fig. 7. Visualización en la interfaz del cliente Web alertas de región abandonada.

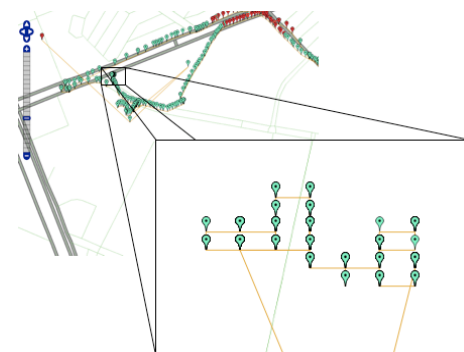


Fig. 8. Margen de error de posicionamiento, $R = 0$.



Fig. 9. Tiempo para el análisis y almacenamiento de una posición.

En la figura 8 se muestran las posiciones reportadas por la unidad mientras se encontraba estacionada, en la prueba el parámetro R , que representa el desplazamiento mínimo que debe tener un vehículo durante un intervalo de tiempo, se estableció a 0 m . Lo anterior se realizó con la finalidad de mostrar que existe, aunque aceptable, un error en la precisión de posicionamiento por parte del AGPS del dispositivo móvil. Por lo anterior, se esperaba que el sistema no emitiera alertas de desplazamiento nulo.

VI. Pruebas de simulación

Debido a la carencia de más de un dispositivo móvil para realizar pruebas de rendimiento al sistema desarrollado, se realizaron pruebas en las cuales se simuló la localización y seguimiento de hasta 100 unidades en movimiento. Para la simulación se implementó un objeto que transmitiera una secuencia de posiciones, obtenidas de un archivo de texto, a la aplicación servidor. Para simular las 100 unidades se ejecutó el objeto en 10 computadoras, con 10 hilos cada una.

En las Fig. 9 y 10 se muestran las gráficas de tiempo requerido para almacenar la posición geográfica de una unidad en función del total de unidades en localización y el tiempo necesario para actualizar la posición y trayectoria de las unidades en la interfaz del Cliente Web, respectivamente. Es importante mencionar que los tiempos que se muestran en la gráfica de

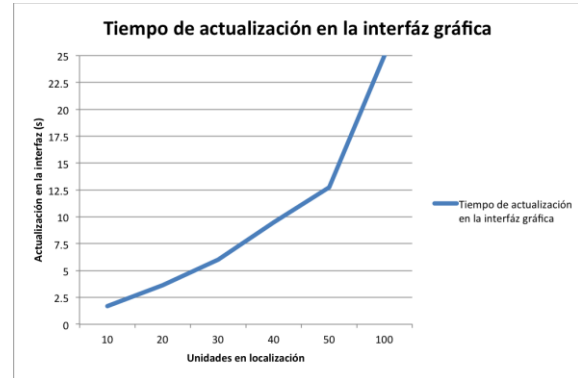


Fig. 10. Tiempo para actualizar la posición y trayectoria de unidades.

la Fig. 6 incluyen tiempo consumido en el proceso de análisis y emisión de alertas.

VII. Conclusiones

Este desarrollo es una opción viable para la localización y seguimiento vehicular, dado que permite a una organización conocer, en línea, la ubicación y trayectoria de sus unidades a través de la Internet con un margen de error $\leq 15\text{ m}$. La emisión de alertas otorga un medio para la rápida toma de decisiones, ya que el tiempo de retardo para la visualización de una alerta es $\leq 25\text{ s}$.

Los mapas temáticos con base en el historial de posicionamiento, permiten analizar la logística de los recorridos realizados por los vehículos y con ello adoptar medidas para optimizar el uso de los recursos de la organización si es necesario.

Sin embargo, la solución desarrollada es susceptible de mejoras debido a:

- La cartografía que se elija es clave para el manejo temático y cobertura de la aplicación. En el caso de este desarrollo se limitó a la Ciudad de México.
- Podría implementarse un mecanismo de comunicación de la aplicación web a los clientes móviles basado en el mecanismo *push*. En el producto final, la comunicación entre los clientes móviles y la aplicación web sigue una arquitectura Cliente – Servidor a través del protocolo HTTP. Por lo cual, la reconfiguración de los parámetros de la aplicación móvil (como el tiempo de

localización) podría tener un retraso considerable ya que se realiza por medio de la técnica *polling*.

- La aplicación móvil puede implementarse en dispositivos para los cuales no se necesite comprar licencias de desarrollo, cómo fue el caso de el iPhone.

VIII. Referencias

- [1] KinYeung, W. "CellPhones as Mobile Computing Devices". *IT Professional* vol.12, no.3, May-June 2010, pp. 40-45.
- [2] Reza, B. *Mobile Computing Principles: Designing and Developing Mobile Applications with UML and XML: Mobility and Location-Based Services*. Cambridge University Press, pages 676 - 722, 2005.
- [3] Bareth, U. et all. "geoXmart - A Marketplace for Geofence-Based Mobile Services," *Computer Software and Applications Conference (COMPSAC), 2010 IEEE 34th Annual*. 19-23 July 2010. pp. 101-106.
- [4] Wong, A.K. et al. "An AGPS-based elderly tracking system". *Ubiquitous and Future Networks, 2009. ICUFN 2009. First International Conference*. 7-9 June 2009. pp.100-105.
- [5] Porcino, D. "Location of third generation mobile devices: a comparison between terrestrial and satellite positioning systems". *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol.4. 2001. pp. 2970-2974.
- [6] Jarvinen, J. et. al. "Assisted GPS: A Low Infrastructure Approach, GPS World". The Business and Technology of Global Navigation and Positioning, 25 de Mayo de 2011. <http://www.gpsworld.com/gps/assisted-gps-a-low-infrastructure-approach-734>.
- [7] Tsalgatidou, A. Veijalainen, J. Markkula, J. Katasonov A. and Hadiefthymiades, S. "Mobile E-Commerce and Location-

Based Services: Technology and Requirements", *Proceedings of ScanGIS'2003 - The 9th Scandinavian Research Conference on Geographical Information Science*, 4-6 June 2003, pp. 1-14.

- [8] Qi, L. "Research on Intelligent Transportation System Technologies and Applications," *Power Electronics and Intelligent Transportation System, 2008. PEITS '08*. 2-3 Aug. 2008. pp.529-531.
- [9] Elahi, M.A. et al. "Design and implementation of real time vehicle tracking system," *Computer, Control and Communication, 2009. IC4 2009. 2nd International Conference*. 17-18 Feb. 2009. pp.1-5.

IX. Autores

Ing. Hermes Francisco Montes Casiano es Ingeniero en Sistemas Computacionales por la Escuela Superior de Cómputo del Instituto Politécnico Nacional.

M. en C. José Jaime López Rabadán obtuvo su título de Maestro en Ciencias en Ingeniería Eléctrica con especialidad en Computación en el Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional.

Lic. Ulises Vélez Saldaña es Licenciado en Computación por la Universidad Autónoma Metropolitana.

Comparación de Mediciones de Atenuación en los Diversos Entornos de Terreno de la Ciudad De México y el Modelo Hata

Jorge R. Sosa Pedroza, Fabiola Martínez Zúñiga, María Elena Acevedo Mosqueda

Instituto Politécnico Nacional, Escuela Superior de Ingeniería Mecánica y Eléctrica, Sección de Estudios de Posgrado e Investigación. Edificio Z-4, 3^{er} piso, Col. Lindavista, C.P.07738, México, D.F
jsosa@ipn.mx, fmartinezzu@ipn.mx, ame1972@gmail.com

Resumen

Este trabajo presenta la comparación de la aplicación del Modelo Hata, con mediciones hechas en la Banda de 400 MHz en la Ciudad de México, se pretende establecer una alternativa de aplicación para las viejas ciudades coloniales españolas de América Latina. El estudio pretende establecer una relación objetiva entre el modelo y mediciones hechas en una ciudad de edificios de paredes gruesas de piedra, con alturas superiores a 15 m y calles angostas, que contrasta estructuralmente con las ciudades japonesas, en donde se hicieron las mediciones del modelo, por parte de Okumura y Hata. Se incluyen también en el análisis, las zonas modernas de paredes ligeras y baja altura de una zona típicamente urbana y sin dejar a un lado las zonas rurales que rodean normalmente a las antiguas ciudades. Los resultados se presentan comparando la atenuación en ambos casos, en función de la distancia. Se establece la comparación con el modelo Hata porque es el más usado en la actualidad, aún y cuando existen algunos modelos de reciente presentación como lo es Erceg sobre el cual los autores han realizado trabajos de investigación. La razón principal sobre el enfoque de nuestro estudio en este modelo es la frecuencia sobre la que operamos.

Palabras Clave: Okumura, Hata, modelos de propagación, sistemas de información geográfica.

I. Introducción

En las características de propagación en un sistema de comunicaciones móviles es un problema ampliamente estudiado en muchas partes del mundo; mediante la medición específica en las ciudades de interés, se han propuesto modelos de atenuación que después se han generalizado para otros sitios, sin embargo, para cada lugar, es importante corroborar si son aplicables o si las mediciones específicas tienen diferencias que deben llevar a un ajuste de los parámetros del modelo. El modelo más ampliamente usado para comunicaciones móviles en el mundo, para frecuencias entre 150 MHz y 1500 MHz, es el desarrollado por Hata [1], que define una formulación empírica a partir de mediciones de propagación hechas en Japón por Okumura [2], posteriormente el modelo fue modificado para incluir frecuencias en las bandas de telefonía celular en el rango de 2 GHz y se propuso para su aplicación por parte de la UIT en

la recomendación COST 231. Este trabajo presenta la comparación de la aplicación del Modelo Hata con mediciones hechas en la Banda de 400 MHz en la Ciudad de México, se pretende establecer una alternativa de aplicación, para las viejas ciudades coloniales españolas de América Latina, con edificios de paredes gruesas de piedra, con alturas superiores a 15 m y calles angostas, incluyendo también las zonas modernas de paredes ligeras, con estructuras metálicas y grandes ventanales de vidrio y sin dejar a un lado las zonas rurales que rodean normalmente a las antiguas ciudades. El estudio es el resultado de un recorrido de aproximadamente 1500 Km por las calles de la Ciudad de México, con un promedio calculado por minuto, sobre mediciones hechas cada 0.3 segundos, lo que aproximadamente representa una medición cada 600 m sobre velocidades que dependían de la cantidad de tráfico y de las vías recorridas.

Hemos escogido para la presentación en este trabajo tres zonas típicas definidas por Hata: una ciudad alta, en el centro histórico de la Ciudad de México, con edificios de altura superior a 15m y calles angostas, una ciudad media en una zona residencial de casas bajas, no mayores a 10 m, y calles relativamente anchas y una zona rural de espacios amplios, con alguna vegetación o sembradíos y casas bajas de 5 m muy espaciadas unas de otras. El procedimiento consiste en comparar las mediciones de atenuación con los resultados empíricos de Hata, para establecer diferencias entre ambos.

II. El modelo Hata

El modelo Hata es el más usado en la planeación y predicción de propagación de los sistemas de comunicaciones móviles, en el rango de 150 MHz a 1500 MHz, partiendo de los estudios de Okumura, modificó el modelo proponiendo el uso como una nueva variable el factor de corrección de altura de la antena del móvil, definida a partir del “tamaño” de la ciudad, en relación con la altura de los edificios, así una ciudad mediana es aquella que tiene edificios de alturas promedio inferiores a 10 m y una ciudad grande la que tiene edificios de altura superior; por otro lado una zona suburbana es la que tiene edificios bajos y espacios amplios y una zona rural la conformada por espacios abiertos con alguna vegetación, por supuesto en una misma ciudad puede haber condiciones diferentes y debe aplicarse aquella que corresponda a la zona en que se desea determinar la atenuación. El modelo de Hata permite determinar las pérdidas de propagación a partir de la ecuación empírica dada por:

$$L_{50}(\text{dB}) = 69.55 + 26.16 \log f_c - 13.82 \log h_{te} - a(h_{re}) + (44.9 - 6.55 \log h_{te}) \log d \quad (1)$$

En donde:

f_c = Frecuencia portadora [150 - 1500 MHz]

h_{te} = Altura de antena transmisora en metros para 30-200 m

h_{re} = Altura de antena receptora en metros para 1-10 m.

$a(h_{re})$ = Factor de corrección para la altura efectiva de la antena móvil, que es función del tipo de área de servicio.

d = Distancia entre transmisor y receptor en km

La corrección para la altura de la antena móvil se establece en la siguiente forma:

Para ciudades pequeñas y medianas:

$$a(h_{re}) = (1.1 \log f_c - 0.7) h_{re} - (1.56 \log f_c - 0.8) \text{ dB} \quad (2)$$

Para ciudades grandes:

$$a(h_{re}) = 8.29 (\log 1.54 h_{re})^2 - 1.1 \text{ dB} \quad (3)$$

donde h_{re} es una función de la frecuencia, dada por:

$$a(h_{re}) = 8.29 (\log 1.54 h_{re})^2 - 1.1 \text{ dB} \quad (4)$$

para $f_c < 300$ MHz

$$a(h_{re}) = 3.2 (\log 11.75 h_{re})^2 - 4.97 \text{ dB} \quad (5)$$

para $f_c > 300$ MHz

Para ambientes suburbanos o rurales, se aplican las siguientes correcciones:

Áreas suburbanas:

$$L(\text{dB}) = L_{50}(\text{urbano}) - 2[\log(f_c/28)]^2 - 5.4 \quad (6)$$

Áreas rurales:

$$L(\text{dB}) = L_{50}(\text{urbano}) - 4.78(\log f_c)^2 + 18.33 \log f_c - 40.94 \quad (7)$$

III. El experimento

El experimento se desarrolló montando la antena receptora sobre el techo de un automóvil,

conectada a un receptor ANRITSU que guarda en memoria 3 datos por segundo, mismos que promedia cada minuto, se hicieron mediciones sobre 22 estaciones base, diseminadas por toda la Ciudad de México, para el caso que nos ocupa las frecuencias son cercanas a los 400 MHz, la clasificación se hizo en función de Hata para el tamaño de las ciudades, zona urbana alta en el centro histórico, una zona urbana media y suburbana en el oriente y una zona rural en el sur de la ciudad. La figura 1 muestra el Sistema de Información Geográfica construido para realizar e análisis de propagación por zona de estudio.

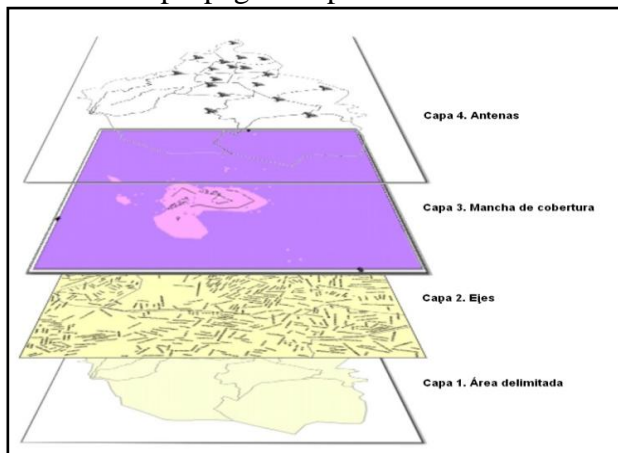


Fig. 1. Integración del Sistema de Información Geográfica.

Todas las mediciones están totalmente georreferenciadas y se cerraron círculos de medición en los recorridos tanto como fue posible, un ejemplo de lo anterior lo muestra el mapa de la figura 2.

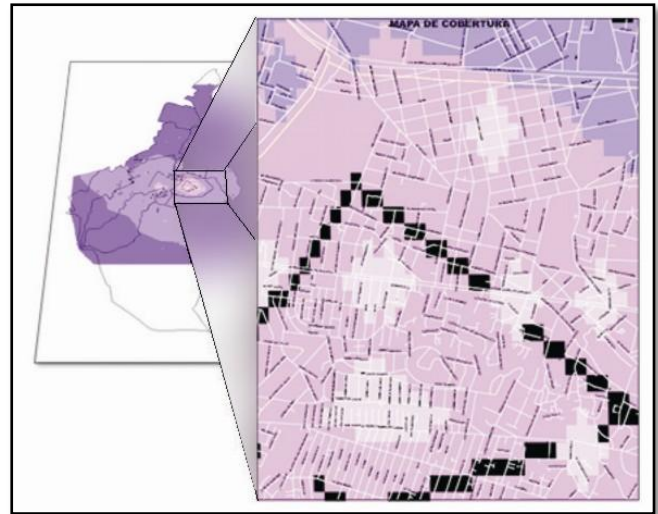


Fig. 2. Mapa de cobertura georreferenciado.

Todas las mediciones están georreferenciadas de tal modo que el cálculo de distancia, entre la estación base correspondiente y los puntos de comparación, se considera totalmente confiable. Partiendo de que las zonas de estudio tienen estructuras de construcción similares presentamos, como un análisis preliminar, un promedio de mediciones en función de la distancia para comparar con los resultados empíricos de Hata, continuamos el estudio de los resultados para definir las características estadísticas en función de la distancia.

En general las características de transmisión para todas las estaciones son:

$$P_t = 44 \text{ dBm}$$

$$\text{Ganancia de Antena transmisora } G_t = 7.5 \text{ dB}$$

$$\text{PIRE} = 49 \text{ dBm}$$

$$\text{Ganancia de Antena receptora } G_r = 2 \text{ dB}$$

La atenuación se refiere a la relación

$$\alpha = \frac{P_R}{\text{PIRE} \times G_t \times G_r} \quad (8)$$

Donde P_R es la potencia promedio medida en el receptor.

En el analizador de espectro la función “*analizador de interferencias*” del BTS, despliega un espectrograma que es una representación en dos dimensiones de frecuencia y tiempo, mientras que los niveles de potencia se muestran en una escala de colores en la parte izquierda como se observa en la figura 3, es posible medir niveles de potencia en el rango de frecuencia 390 - 400 MHz. Dentro de este ancho de banda se encuentran todas las portadoras de las 22 estaciones base, en promedio 8 para cada una. En el recuadro superior del aparato se muestra la posición geográfica, así como la fecha y hora de la medición, por cada archivo de un minuto el BTS almacena solamente una posición GPS. En la parte inferior se muestran las portadoras presentes y su nivel de potencia y en la pantalla principal los trazos de la medición de las portadoras presentes en el ancho de banda seleccionado. Todos los archivos de un minuto generan una base de datos que fue procesada para relacionarla con su posición geográfica, indicada por el GPS, lo que permite georreferenciar los niveles de potencia y mostrarlos como una capa en el software del Sistema de Información Geográfica ArcGIS [4].

Para procesar las mediciones se diseñó un programa en lenguaje C, en primer lugar para discriminar las portadoras que no correspondan a la radiobase de interés y en segundo lugar para generar datos georreferenciados con el promedio de potencia de las mediciones de cada minuto, teniendo en esta forma valores puntuales de potencia promedio con su respectiva localización geográfica. Una vez filtradas las portadoras de la estación de interés se promedian las muestras por minuto obteniendo su potencia promedio, que representa el promedio de la portadora de 180

trazos por segundo. El nivel de potencia de las bases finales para generar los mapas de cobertura es el promedio de los promedios de las portadoras de cada una de las estaciones e indica la medición puntual georeferenciada que entrega la base de datos del programa en C. En total se obtienen 1200 puntos promedio cubriendo el recorrido establecido para las mediciones.

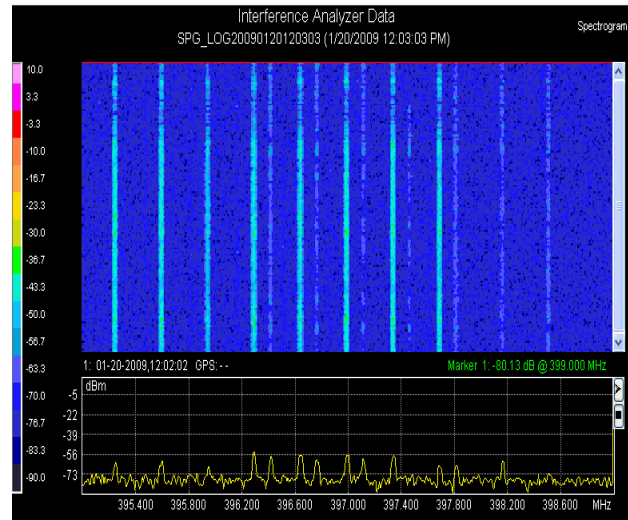


Fig. 3. Espectrograma dentro de la función analizador de interferencias del BTS.

Los datos puntuales se cargan en el software *Easy Krig* donde se realiza la interpolación o cálculo de datos estimados para cubrir áreas geográficas completas o medición continua, además proporciona mediciones de potencia de la misma magnitud (isolínea). De esta manera 2 mapas son generados, uno con mediciones puntuales georreferenciadas y otro con mediciones continuas o mancha de cobertura, ambos serán superpuestos en el ArcGis para generar el mapa de cobertura georreferenciado y obtener niveles de potencia en distancias reales. La base de datos se captura en la capa 1 del SIG. Los valores de potencia, se agruparon en niveles por medio del programa desarrollado en lenguaje C estableciendo un código de colores que permite visualizar desde el nivel más óptimo de

potencia (rosa claro) hasta el menos óptimo (morado), ver figura 4. Los puntos del nivel más óptimo se encuentran cercanos a la radiobase y los demás cambian de color en función de los obstáculos que encuentren en el terreno. Es evidente que conforme la distancia aumenta, la potencia disminuye debido a las pérdidas por trayectoria, incluyendo todos los factores que atenúan una señal.





COLOR	NIVEL DE CAMPO (dBm)
	MAYOR A -63
	ENTRE -63 Y -70
	ENTRE -70 Y -78
	ENTRE -78 Y -86

Fig. 4. Código de colores de niveles de campo.

Se buscó que la diferencia de color fuera muy notoria, para que el límite entre dos colores nos dibujara una isolínea. Conocido el nivel de potencia, se procede a medir la distancia promedio y hacer el análisis de potencia en función de la distancia de acuerdo al tipo de terreno en el que se encuentre la radiobase y así comparar datos con el modelo de propagación.

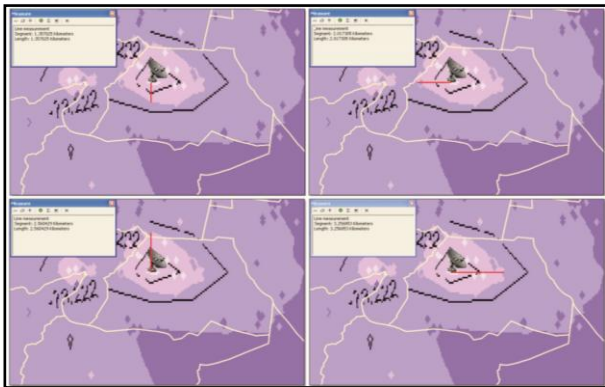


Fig. 5. Medición de distancias promedio.

El software ArcGis contiene una herramienta que permite la medición de distancias georeferenciadas. Una vez ubicada la isolínea y la

estación base, se miden distancias alrededor de la separación transmisor-isolínea y se obtiene un promedio de la distancia. El proceso es el mismo para las 22 estaciones base. De esta manera tenemos las dos variables que nos interesan, valor de la potencia medida y la distancia entre el punto de transmisión y recepción, representado por la isolínea.

La clasificación de los mapas está en función de los tipos de área que propone Hata en su modelo, para el D.F realizamos la siguiente clasificación por zonas:

URBANA ALTA

Estación 8. Ubicación: Delegación Cuauhtémoc.
Estación 17. Ubicación: Delegación Cuauhtémoc.
Estación 20. Ubicación: Delegación Cuauhtémoc.

URBANA MEDIA

Estación 1. Ubicación: Delegación Venustiano Carranza.
Estación 4. Ubicación: Delegación Azcapotzalco.
Estación 5. Ubicación: Delegación Iztapalapa.
Estación 7. Ubicación: Delegación Gustavo A. Madero.
Estación 12. Ubicación: Delegación Iztapalapa.
Estación 13. Ubicación: Delegación Iztacalco.
Estación 15. Ubicación: Delegación Coyoacán.
Estación 16. Ubicación: Delegación Benito Juárez.
Estación 18. Ubicación: Delegación Tlalpan
Estación 19. Ubicación: Delegación Gustavo A. Madero.
Estación 21. Ubicación: Delegación Gustavo A. Madero.

SUBURBANA

Estación 10. Ubicación: Delegación Alvaro Obregon.
Estación 11. Ubicación: Delegación Alvaro Obregon.

ABIERTA O RURAL

Estación 14. Ubicación: Iztapalapa.
Estación 22. Ubicación: Delegación.

Una vez que se hizo la clasificación se generaron los mapas para cada estación, presentaremos únicamente el de la estación 8 ubicada en la Delegación Cuauhtémoc con una altura de 63 m y que por su ubicación geográfica pertenece a la zona urbana alta de acuerdo a la clasificación presentada; se observa la isolínea con promedio de potencia -72.15 dBm de la cual se calcula el promedio de la distancia de 4.94 kilómetros, véase la figura 6.

Se procede de igual forma para las 22 estaciones respetando su clasificación en función del tipo de entorno que se estudia; se encuentran las distancias promedio y su nivel de potencia en función de las isolíneas generadas por el software.

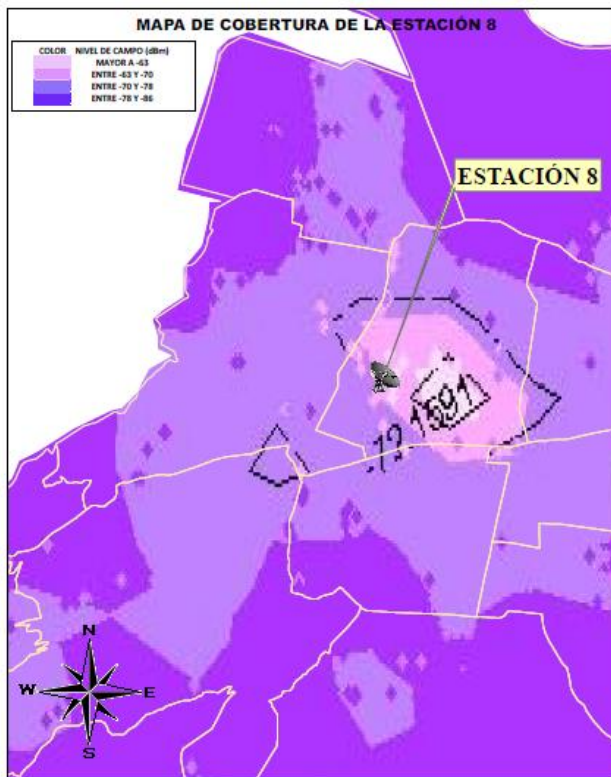


Fig. 6. Mapa de cobertura de la estación 8.

Se genera una base de datos comparativa entre potencias de predicción del modelo y las

mediciones en campo, generando gráficas comparativas en función de las cuatro clasificaciones presentadas, las figura 7 a 10 presentan los resultados de la comparación.

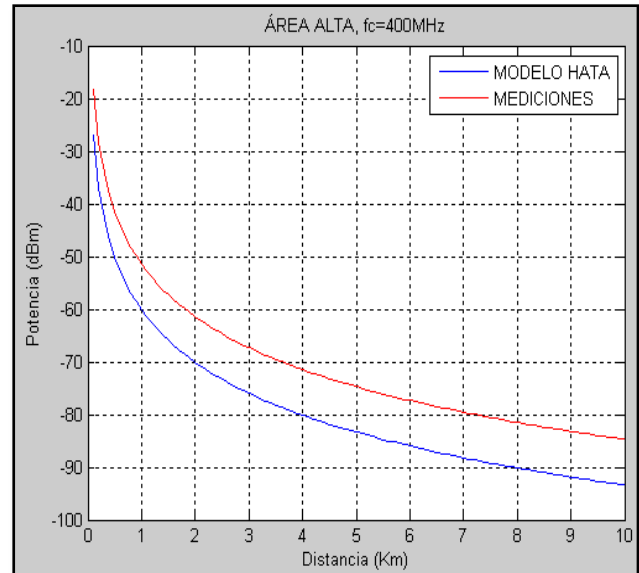


Fig. 7. Comparación entre mediciones de campo y el modelo Hata para áreas urbanas altas.

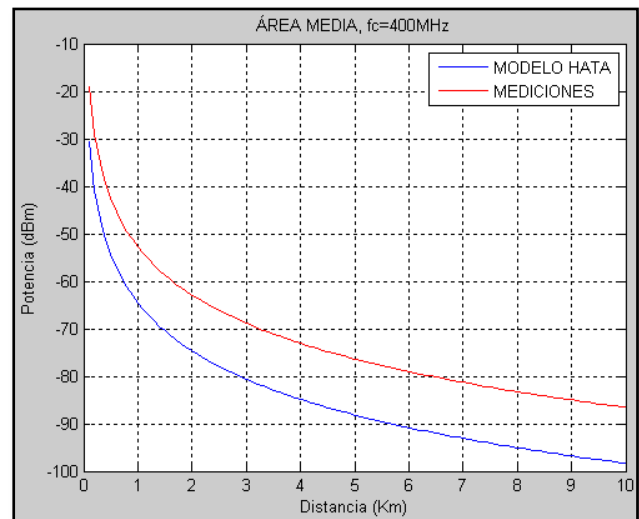


Fig. 8. Comparación entre mediciones de campo y el modelo Hata en áreas urbanas medias.

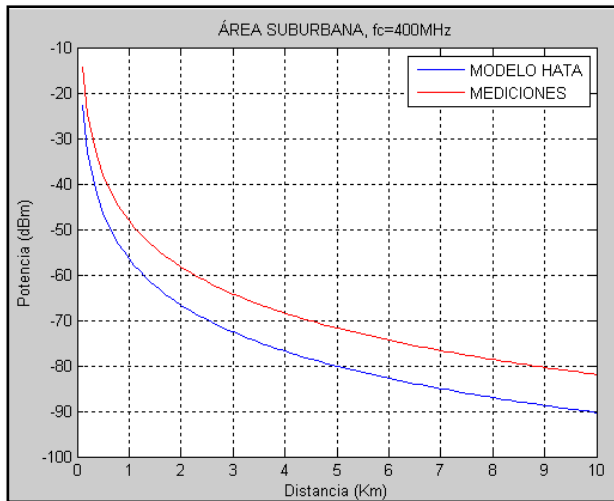


Fig. 9. Comparación entre mediciones de campo y el modelo Hata para áreas suburbanas.

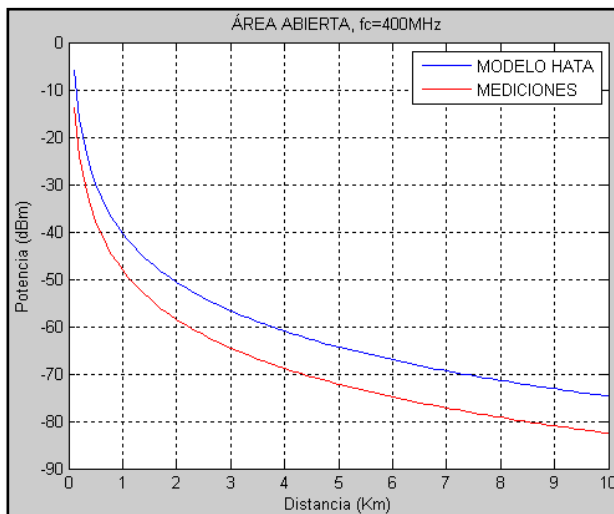


Fig. 10. Comparación entre mediciones de campo y el modelo Hata para áreas abiertas.

IV. Discusión

Una de las aportaciones más importante de este trabajo es la construcción de nuestro Sistema de Información Geográfica ya que nos permiten visualizar de manera muy puntal el comportamiento de la propagación de las ondas electromagnéticas en función del tipo de terreno sobre el que se trabaja. El modelo Hata es desarrollado sobre condiciones muy diferentes a la de nuestro país y ciudad, por eso la relevancia de la metodología presentada; una zona, como en

el estudio original de Hata, es diferente a una ciudad con unidades habitacionales como la nuestra, este entorno geográfico no está propuesto en el modelo original, o bien un estudio sobre las características del centro histórico de la ciudad de México, por ello la relevancia de la metodología aquí propuesta ya se obtienen resultados que muestran que la caracterización de la propagación electromagnética debe estar en función de las condiciones de terreno

Con la información de la distribución territorial puede caracterizarse el propio modelo y por tanto decidir si es apropiado para usarse en las condiciones específicas de esta zona, se entiende sin embargo, que una decisión en este sentido debe basarse en estudios mucho más amplios que los aquí presentados; los autores trabajan actualmente en este sentido.

V. Conclusiones

De acuerdo al experimento realizado podemos concluir que existen diferencias entre el modelo Hata y las mediciones realizadas en la Ciudad de México. Si se deseara ajustar el modelo Hata a las mediciones, tendría que realizarse un ajuste a la fórmula empírica del modelo de predicción sumando o restando las diferencias de acuerdo a las zonas propuestas por Hata: Para la zona alta 8.91 dB, zona urbana media 11.97dB, zona suburbana 8.34dB y para la zona abierta -7.82dB.

VI. Referencias

- [1] Y. Okumura et al. "Field strength and variability in VHF and UHF land-mobile radio service" Rev. Elec. Commun. Lab., vol. 16, 1968
- [2] Masaharu Hata "Empirical Formula for Propagation Loss in Land Mobile Radio Services" IEEE Transactions on Vehicular Technology, Vol. VT-29, No. 3, August 1980.

- [3] David Parsons. The Mobile Radio Propagation Channel, Pentech Press, London, 1994.
- [4] Jorge R. Sosa-Pedroza, Fabiola Martínez-Zúñiga, Miguel Sánchez-Meraz “Comparing Hata model with Mexico city’s measurements in 400 MHz”, Proceedings of LATINCOM workshop, Medellín Colombia, September 2009.

VII. Autores



Dr. Jorge R. Sosa Pedroza, Profesor-Investigador Titular de tiempo completo de la ESIME-IPN de 1973 a la fecha. Ingeniero en Comunicaciones y Electrónica de la ESIME IPN, con estudios de Maestría en

Ciencias con especialidad en Telecomunicaciones en la Sección de Posgrado de la ESIME y doctorado en Telecomunicaciones por la misma institución



M. en C. Fabiola Martínez Zúñiga, Profesora-Investigadora de la ESIME-IPN de 2003 a la fecha. Ingeniera en Comunicaciones y Electrónica, con una Maestría en Ciencias con

especialidad en Telecomunicaciones en la Sección de Posgrado de la ESIME. Profesora Titular de las academias de Electromagnetismo y Comunicaciones.



María Elena Acevedo Mosqueda. Realizó sus estudios de licenciatura en Ingeniería con especialización en Computación en la Escuela

Superior de Ingeniería Mecánica y Eléctrica (ESIME) del Instituto Politécnico Nacional (IPN) en 1996. Se ha desempeñado como profesora en ESIME desde 1994. Recibió su grado de Maestra en Ciencias con especialización en Computación en el Centro de Investigación y de Estudios Avanzados (CINVESTAV) en 2001. Obtuvo su grado de Doctora en Ciencias con especialidad es Computación en el Centro de Investigación en Computación del IPN (CIC-IPN) en 2006. Su principal área de investigación son las Memorias Asociativas Bidireccionales.

Optimización de un Módulo Peltier de Dos Etapas en Régimen Eléctrico Oscilante

* Federico Vázquez Hurtado, **Miguel Ángel Olivares Robles, *** Alfonso Dávila Enriquez

* Facultad de Ciencias, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001, Chamilpa, Cuernavaca 62209, Mor., 01 (777) 3297020, vazquez@uaem.mx .

**Sección de Estudios de Posgrado e Investigación, ESIME-Culhuacan IPN, Av. Santa Ana No 1000, Coyoacan 04430 México D.F., 555 7296000 Ext. 73025, olivares67@mailaps.org.

*** Sección de Estudios de Posgrado e Investigación, ESIME-Culhuacan IPN, Av. Santa Ana No 1000, Coyoacan 04430 México D.F., 555 7296000 Ext. 73025,alfda@hotmail.com.

Resumen

En este trabajo analizamos el desempeño de un módulo termoelectrico semiconductor (MTE) de dos etapas que contiene un número arbitrario diferente de termopares, n_1 y n_2 , en cada etapa. Los MTEs con este tipo de estructura, conocida como piramidal, se describen por medio de un conjunto de ecuaciones de balance de calor que están escritas en términos de propiedades adimensionalizadas del sistema. El dispositivo se opera como refrigerador y por cada etapa del módulo se hace fluir una corriente distinta suponiendo que éstas no son constantes sino que varían armónicamente con el tiempo. En general, las corrientes están fuera de fase y se evalúan los efectos del desfaseamiento entre ellas sobre el coeficiente de desempeño térmico del sistema (COP por sus siglas en inglés). En las condiciones descritas, reportamos regímenes de funcionamiento de carácter resonante que llevan a que los valores del COP sean sustancialmente mayores que los del caso en el que las corrientes son constantes.

Palabras Clave: Optimización, coeficiente de desempeño térmico, termoelectricidad, corriente eléctrica oscilante.

I. Introducción

Los módulos termoelectricos (MTEs) se utilizan ampliamente para control de la temperatura de componentes electrónicos que tienen aplicación en muchas áreas tecnológicas como la refrigeración termoelectrica, aplicaciones espaciales, enfriamiento de dispositivos pequeños, secado, conservación de alimentos, etc. [1]. El desempeño térmico de los MTEs dependen de varios factores entre los que se encuentran las propiedades de los materiales semiconductores, la geometría de los termopares, la temperatura de las juntas, la corriente eléctrica que circula, etc [2-4]. Los métodos para estudiar este tipo de sistemas que se han descrito en la literatura incluyen teoría de control, termodinámica fuera de equilibrio,

termodinámica de tiempo finito, entre otros. El método que se utiliza en este trabajo fue originalmente propuesto por Yamanashi [5], Chen [6] y otros en la segunda mitad de los 90's y fue concebido dentro del marco teórico de la termodinámica de procesos irreversibles. Una aplicación reciente puede verse en [7]. Aquí se presenta un análisis basado en la termodinámica de procesos irreversibles [8] sobre el desempeño de un MTE de dos etapas usando las ecuaciones adimensionales de balance de calor como se pueden ver en [9]. Los dispositivos de más de una etapa han tenido especial relevancia cuando se requiere operar el dispositivo como refrigerador con la máxima diferencia de temperatura entre su lado frío y su lado caliente. Para tal fin es conveniente utilizar la eficiencia

energética en lugar del COP ya que ella permite determinar las condiciones de operación en las que se establece un compromiso entre un COP máximo y una diferencia de temperaturas máxima. En [7] puede encontrarse un estudio detallado sobre el desempeño en esos casos. El criterio de optimización que se utiliza en el presente desarrollo se basa en el COP considerando que las propiedades de los materiales p y n de los termopares son independientes de la temperatura, El desempeño del dispositivo depende de su configuración (número de termopares en cada etapa) y de las características de las corrientes eléctricas que fluyen por ellas. A este respecto, las corrientes son distintas y varían con el tiempo. Se ha supuesto una variación temporal armónica y la oscilación de las corrientes ocurre alrededor de un valor promedio positivo. En la descripción se incluye la contribución de Fourier al flujo de energía y la disipación debida al efecto Joule. No se considera el efecto Thomson. De este modo, las irreversibilidades presentes en el sistema provienen de los dos primeros. En la siguiente sección se plantean las ecuaciones de balance de calor para el MTE y se muestra el cálculo del COP . En la tercera sección se introducen las corrientes armónicas y se realiza el análisis del efecto de la constante de desfase de las corrientes eléctricas sobre el desempeño del dispositivo.

II. Desarrollo

El MTE que se va a analizar consiste de dos etapas caracterizadas por el número de termopares n_1 y n_2 en cada una de ellas. Los termopares están eléctricamente conectados en serie en cada etapa y térmicamente en paralelo. Las dos etapas están conectadas térmicamente en serie. La Figura 1 muestra el dispositivo.

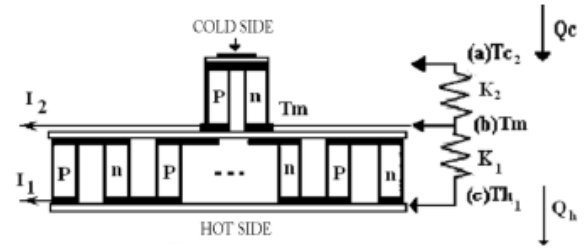


Fig. 1. Diagrama esquemático del refrigerador termoelectrico de dos etapas.

El desempeño de cada etapa está determinado por la resistividad eléctrica ρ , el coeficiente Seebeck α y la conductividad térmica κ de los materiales p y n . Todos estos parámetros se suponen constantes. En cada etapa la resistencia eléctrica R , el coeficiente Seebeck A y la conductancia térmica K están dados por:

$$R = n(\rho_p + \rho_n) \frac{l}{S}, \quad (1)$$

$$A = n(\alpha_p + \alpha_n), \quad (2)$$

$$K = n(\kappa_p + \kappa_n) \frac{S}{l}, \quad (3)$$

donde l es la longitud de los semiconductores en los termopares, S es el área transversal de ellos, n es el número de termopares en la etapa y los subíndices p y n denotan los semiconductores tipo p y n , respectivamente. En principio, R , A y K Tienen valores distintos en cada etapa. Las ecuaciones de balance del calor transferido entre las posiciones (a) a (c) en la Figura 1 se escriben como sigue:

$$Q_{C_2} = (A_2 I_2 T_{C_2} - \frac{1}{2} I_2^2 R_2 - K_2 (T_m - T_{C_2})) n_2, \quad (4)$$

$$Q_{H_2} = (A_2 I_2 T_{H_2} + \frac{1}{2} I_2^2 R_2 - K_2 (T_m - T_{C_2})) n_2, \quad (5)$$

$$Q_{C_1} = (A_1 I_1 T_{C_1} - \frac{1}{2} I_1^2 R_1 - K_1 (T_{H_1} - T_m)) n_1, \quad (6)$$

$$Q_{H_1} = (A_1 I_1 T_{H_1} + \frac{1}{2} I_1^2 R_1 - K_1 (T_{H_1} - T_m)) n_1, \quad (7)$$

donde Q_C es el flujo de calor que ingresa a cada etapa desde el lado frío y Q_H el flujo de calor

que sale por el lado caliente. T_C es la temperatura del lado frío del módulo, T_H la temperatura del lado caliente, T_m es la temperatura de contacto entre las dos etapas del MTE e I la corriente eléctrica que circula por cada una de ellas. Debe notarse que $T_{H_2} = T_{C_1} = T_m$. También se tiene que $n_2 = N_T / (r+1)$ y que $n_1 = rN_T / (r+1)$, siendo N_T el número total de termopares en el MTE, $N_T = n_1 + n_2$, y r el cociente n_1 / n_2 .

Ahora se reescriben las ecuaciones (4) a (7) introduciendo las variables adimensionales que a continuación se definen. Esto se hace para que ellas sean válidas para cualquier sistema MTE.

$$j_1 = \frac{A_1 I_1}{K_1}, \quad j_2 = \frac{A_2 I_2}{K_2}, \quad (8)$$

$$q_{C_1} = \frac{Q_{C_1}}{N_T K_1 T_{C_1}}, \quad q_{C_2} = \frac{Q_{C_1}}{N_T K_2 T_{C_2}}, \quad (9)$$

$$q_{H_1} = \frac{Q_{H_1}}{N_T K_1 T_{H_1}}, \quad q_{H_2} = \frac{Q_{H_2}}{N_T K_2 T_{H_2}}, \quad (10)$$

$$\theta_{H_1} = \frac{T_{H_1}}{T_m}, \quad \theta_{C_2} = \frac{T_{C_2}}{T_m}, \quad \theta_T = \frac{T_{H_1}}{T_{C_2}} = \frac{\theta_{H_1}}{\theta_{C_2}}, \quad (11)$$

$$ZT_{C_2} = \frac{A_2^2 T_{C_2}}{R_2 K_2}, \quad ZT_{H_1} = \frac{A_1^2 T_{H_1}}{R_1 K_1}, \quad (12)$$

$$ZT_{m_1} = \frac{A_1^2 T_m}{R_1 K_1}, \quad ZT_{m_2} = \frac{A_2^2 T_m}{R_2 K_2}. \quad (13)$$

En esta forma, j_1 y j_2 son las corrientes eléctricas adimensionales que fluyen a través de cada etapa del MTE, q_{C_2} y q_{H_1} pueden considerarse como el flujo de entropía normalizado por la conductancia térmica del NTE en las posiciones (a) y (c) en la Figura 1, respectivamente, ZT_C , ZT_H y ZT_m son figuras de mérito termoeléctrico asociadas a las temperaturas T_C , T_H y T_m , respectivamente. Las ecuaciones (4) a (7) pueden ahora reescribirse usando las variables adimensionales. El resultado es:

$$q_{C_2} = (j_2 - \theta_{C_2}^{-1} + 1 - \frac{1}{2} j_2^2 \frac{1}{ZT_{C_2}}) \frac{1}{r+1}, \quad (14)$$

$$q_{H_2} = (j_2 + \theta_{C_2} - 1 + \frac{1}{2} j_2^2 \frac{1}{ZT_{m_2}}) \frac{1}{r+1}, \quad (15)$$

$$q_{C_1} = (j_1 - \theta_{H_1} + 1 - \frac{1}{2} j_1^2 \frac{1}{ZT_{m_1}}) \frac{r}{r+1}, \quad (16)$$

$$q_{H_1} = (j_1 + \theta_{H_1}^{-1} - 1 + \frac{1}{2} j_1^2 \frac{1}{ZT_{H_1}}) \frac{r}{r+1}. \quad (18)$$

Dadas las corrientes que circulan por cada etapa, j_1 y j_2 , por un cálculo directo se puede mostrar que la temperatura de contacto está dada por:

$$T_m = \frac{T_{C_2} + T_{H_1} r + \frac{1}{2} (R_1 K_1 j_1^2 r / A_1^2 + R_2 K_2 j_2^2 r / A_2^2)}{1 - j_2 + r(j_1 + 1)} \quad (19)$$

Ahora se introduce el coeficiente de desempeño (COP) del MTE. El COP se define como:

$$COP = \frac{Q_{C_2}}{Q_{H_1} - Q_{C_2}} = \frac{q_{C_2}}{\theta_T q_{H_1} - q_{C_2}}. \quad (20)$$

Sustituyendo las ecuaciones (9)-(11) en (19) se llega a:

$$COP = \frac{j_2 - \theta_{C_2}^{-1} + 1 - \frac{1}{2} j_2^2 \frac{1}{ZT_{C_2}}}{\Delta}, \quad (21)$$

donde el discriminante se escribe como:

$$\Delta = \theta_T r (j_1 - 1 + \theta_{H_1}^{-1} + \frac{1}{2} j_1^2 \frac{1}{ZT_{H_1}} - j_2 + \theta_{C_2}^{-1} - 1 + \frac{1}{2} j_2^2 \frac{1}{ZT_{C_2}}). \quad (22)$$

Antes de pasar a examinar las implicaciones de las corrientes eléctricas que circulan por cada etapa sobre el desempeño térmico del MTE se

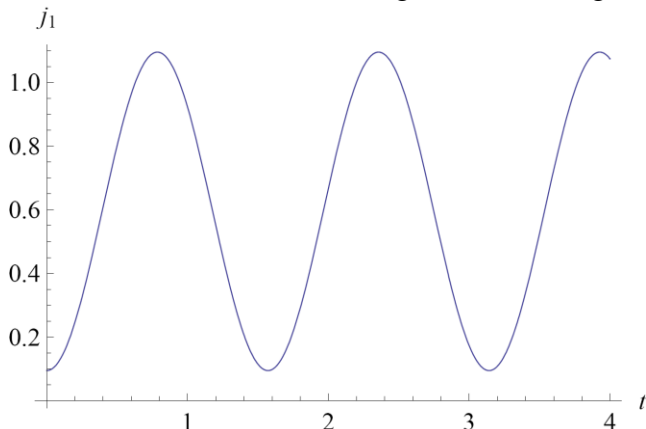
especifican los valores de los parámetros que se utilizan en los cálculos. Los coeficientes Seebeck de los materiales semiconductores tipo p y n tienen los valores $\alpha_p = \alpha_n = 1.8 \times 10^{-4} V/K$, las conductividades térmicas $\kappa_p = \kappa_n = 1.4 \times 10^{-2} W/Kcm$, y las resistividades eléctricas $\rho_p = \rho_n = 7.7 \times 10^{-4} \Omega cm$. La temperatura del lado frío del MTE es $T_{C_2} = 248K$ y la del lado caliente es $T_{H_1} = 308K$. Como los parámetros R y K siempre aparecen como producto en las expresiones no es necesario especificar los valores de la sección transversal de los termopares ni su longitud. Finalmente, la forma genérica que se asigna a las corrientes oscilantes j_1 y j_2 es la siguiente:

$$j = j_{PROM} + C \text{sen}(\omega t + \delta), \quad (23)$$

donde j_{PROM} es el valor promedio de la corriente, ω la frecuencia de oscilación y δ la fase. En la siguiente sección presentamos los resultados obtenidos para el COP cuando se hacen circular corrientes eléctricas oscilantes por cada etapa del MTE.

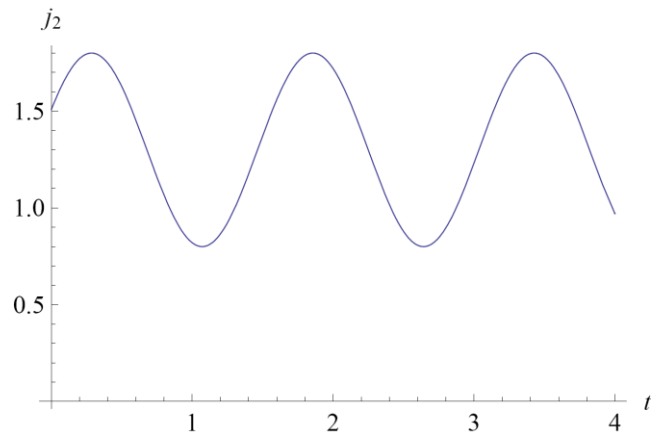
III. Resultados

Iniciamos esta sección con la gráfica en la Figura

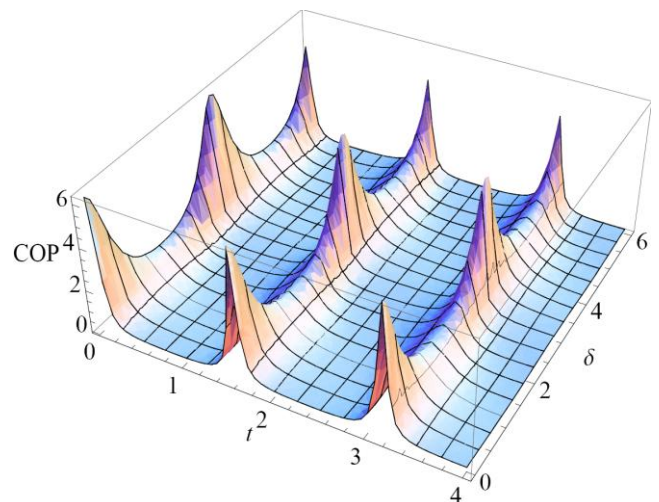


2 en la que se muestra la variación de la corriente eléctrica j_1 en la etapa 1 del MTE con respecto al tiempo. El tiempo está dado en segundos. La

frecuencia ω es de 1 HZ. Se ha tomado como promedio para j_1 el valor de 0.6. La Figura 3 muestra lo respectivo para la corriente j_2 de la etapa 2. Nuevamente, el tiempo está dado en segundos. La frecuencia ω es de 1 HZ. Obsérvese que en este caso el promedio es 1.3. No obstante en la diferencia de la magnitud de las corrientes, ambas tienen la misma amplitud unitaria.



El COP calculado con la ecuación (21) muestra un comportamiento complejo con respecto al tiempo tanto como respecto a la fase δ entre las



corrientes j_1 y j_2 . La Figura 4, donde el tiempo está dado en segundos y la frecuencia ω es de 1

HZ, muestra la gráfica correspondiente donde puede observarse la existencia de máximos tanto con respecto al tiempo como respecto a la fase. Para el propósito de este trabajo los máximos de interés son los que ocurren respecto a la fase entre las corrientes por lo que el comportamiento respecto al tiempo se promedia en el lapso de un ciclo de oscilación de las corrientes. Este promedio se calcula como sigue:

$$COP_{prom}(\delta) = \frac{1}{T} \int_0^T COP(t, \delta) dt, \quad (23)$$

siendo T el lapso entre máximos en la línea temporal del coeficiente de desempeño térmico COP . Para terminar esta sección, en la Figura 5 se presenta el comportamiento del coeficiente de desempeño promediado en el tiempo durante un periodo con respecto a la fase entre las corrientes que circulan en cada etapa del dispositivo. Este es el principal resultado reportado en este trabajo.

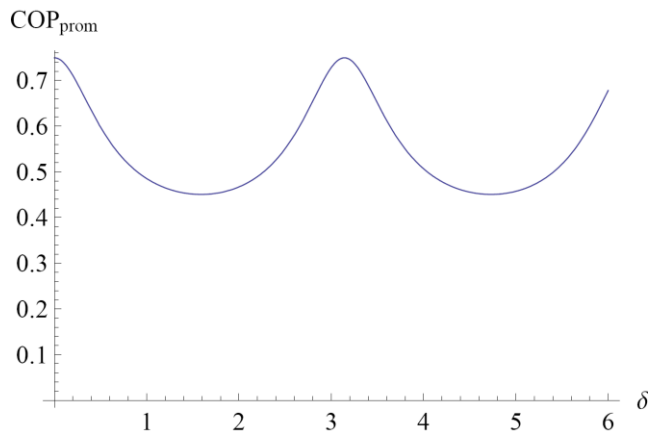


Fig. 5. Coeficiente de desempeño del MTE promedio contra la fase δ .

IV. Discusión

Como es natural, se busca un régimen de funcionamiento del MTE en el que el COP_{prom} sea máximo. Iniciamos esta sección discutiendo las condiciones en las que esto es posible a la luz del comportamiento mostrado por el COP_{prom} en la Figura 5. En dicha figura se puede observar nuevamente la presencia de máximos cuya posición depende de la forma en que aparecen funcionalmente las corrientes de las etapas 1 y 2

del MTE en la ecuación (21) de la sección anterior. El comportamiento observado revela que para valores de la fase 0, 3, 6, etc., en el dispositivo ocurren resonancias térmicas por la forma en que se acoplan los flujos de calor en las diferentes etapas del MTE debido al carácter oscilante de las corrientes j_1 y j_2 . Es claro que la condición para la existencia de estas resonancias es que el módulo termoeléctrico tenga al menos dos etapas y que ellas son el resultado del acoplamiento no lineal de las corrientes como puede observarse en la ecuación (21). Es importante comparar los resultados mostrados en la Figura 5 en las fases de resonancia con los que se obtienen cuando las corrientes j_1 y j_2 son estrictamente constantes. Para ello se calcula a continuación el valor del COP para $j_1 = 0.6$ y $j_2 = 1.3$ que corresponden a los valores promedio de las corrientes oscilantes. La expresión utilizada es nuevamente la ecuación (21). El cálculo reporta que el coeficiente de desempeño tiene el valor 0.11. Comparando este valor con el máximo del COP_{prom} en $\delta = 0, 3, 6, \dots$ que es de 0.75, se encuentra un incremento de 680%.

V. Conclusiones

Se presentan resultados de la operación en régimen de corrientes oscilatorias de un módulo termoeléctrico de dos etapas. El acoplamiento de las corrientes en el coeficiente de desempeño térmico revela la existencia de resonancias térmicas en el sistema que producen un incremento significativo del COP para desfases específicos entre las corrientes. La aparición y descripción detalladas de tales resonancias serán objeto de un estudio posterior.

VI. Referencias

- [1] S.B. Riffat, X.L. Ma, "Thermoelectrics: a review of present and potential applications". Applied Thermal Engineering. Vol 23. Año 2003. Página 913-935.

[2] F. Meng, L. Chen, F. Sun, “Performance optimization for two-stage thermoelectric refrigerator system driven by two-stage thermoelectric generator”. *Cryogenics*. Vol 49. Año 2009. Páginas 57-65.

[3] F. Meng, L. Chen, F. Sun, “Performance analysis for two-stage TEC system driven by two-stage TEG obeying Newton's heat transfer law”. *Mathematical and Computer Modelling*. Vol 52. Año 2010. Páginas 586-595.

[4] F. Meng, L. Chen, F. Sun, “Multivariable optimization of two-stage thermoelectric refrigerator driven by two-stage thermoelectric generator with external heat transfer”. *Indian Journal Pure and Applied Physics*. Vol 48. Año 2010. Páginas 731-742.

[5] M. Yamanashi, “A new approach to optimum design in thermoelectric cooling systems”. *Journal of Applied Physics*. Vol. 80. Año 1996. Páginas 5494-5502.

[6] J. Chen, Y.H. Zhou, H.J. Wang, J.T. Wang, “Comparison of the optimal performance of single- and two-stage thermoelectric refrigeration systems”. *Applied Energy*. Vol 73. Año 2002. Páginas 285-298.

[7] M.A. Olivares-Robles, F. Vázquez, C. Ramírez-López, “Optimization of Two-Stage Peltier Modules: Structure and Exergetic Efficiency”. *Applied Energy*. Enviado.

[8] J. Chen, Z. Yan, L. Wu, “Nonequilibrium thermodynamic analysis of a thermoelectric device”. *Energy*. Vol. 22. Año 1997. Páginas. 979-985.

[9] X.C. Xuan, K.C. Ng, C. Yap, H.T. Chua, “Optimization of two-stage thermoelectric coolers with two design configurations”. *Energy Conversion and Management*. Vol 43. Año 2002. Páginas. 2041-2052.

VII. Autores

Federico Vázquez Hurtado es Doctor en Ciencias con especialidad en Física por la Universidad Nacional Autónoma de México. Pertenece a la Academia Mexicana de Ciencias, a la Academia de Ciencias de Morelos y al Sistema Nacional de Investigadores.

Miguel Ángel Olivares Robles es Doctor en Ciencias en el área de Física por la Universidad Autónoma Metropolitana Unidad Iztapalapa. Pertenece al Sistema Nacional de Investigadores

Alfonso Dávila Enriquez es Ingeniero Mecánico por la Escuela Superior de Ingeniería Mecánica y Eléctrica Unidad Culhuacan.

Diseño de Circuitos Integrados CMOS: Una Aproximación desde la Perspectiva de la Ley de Ohm

Federico Sandoval-Ibarra

CINVESTAV-Unidad Guadalajara, Av. Del Bosque 1145, Col. El Bajío, Zapopan, Jal. (México)
+52 (33) 37773600, sandoval@gdl.cinvestav.mx

Resumen

El objetivo de esta contribución es doble. Por un lado se presenta una metodología de diseño de Circuitos Integrados (CIs) CMOS basada en el uso de la ley de Ohm. Esta propuesta es un intento por rescatar el diseño tradicional, porque es bien sabido, todo diseño no solo inicia con un análisis en corriente directa (cd), sino que la información de ese análisis es la base para determinar la función del circuito en corriente alterna (ca). Por otro lado, y a la par de la omisión en el análisis en cd – comúnmente encontrada en la literatura especializada– se presenta también el nulo impacto que tiene el desarrollo tecnológico en la formación de ingenieros, siendo un ejemplo, el escenario que hoy día presenta la industria de los CIs. Por lo anterior, esta contribución es también un intento por incorporar el avance tecnológico en la instrucción en aula desde la perspectiva del diseño de circuitos simples CMOS, usando modelos simples de diseño y simulándolos según los modelos adoptados por la industria, mas no usando software industrial. Los resultados de esta aproximación de diseño, muestra que la instrucción en aula facilita no solo la comprensión de los conceptos fundamentales del diseño, sino que los modelos analíticos de diseño tienen un significado físico y ello conduce al saber-hacer.

Palabras Clave: Amplificadores, enseñanza, ley de Ohm, tecnología CMOS.

I. Introducción

El avance tecnológico, y el conocimiento racional que de los fenómenos físicos ha alcanzado el hombre, han permitido que a través de la ingeniería el desarrollo de soluciones sea tan amplio como en ninguna otra época sucedió. Los dispositivos inalámbricos de entretenimiento y de comunicación personal, el desarrollo de redes sociales, internet, así como el advenimiento de nuevos productos que integran diversos protocolos de comunicación (PCS, GSM, WLAN, Bluetooth, WCDMA, etc.) son muestra de ese aspecto cultural que hoy identifica a la sociedad postmoderna. La imagen que proyecta el hombre, ser social por naturaleza, hoy día no se concibe sin un dispositivo inalámbrico de comunicación, y esta característica se refleja incluso desde edades tempranas; no es difícil ver a un niño que asiste a la escuela primaria con un teléfono celular. El establecimiento de este escenario obedece a factores de diversa índole,

pero de todos ellos el más importante es la madurez que ha logrado el mayor desarrollo tecnológico de los últimos 50 años: los materiales semiconductores. Uno de esos materiales, por un lado, fue el detonador de una tecnología que dio lugar a los primeros dispositivos portátiles para la comunicación masiva; dispositivos que incluían la etiqueta *solid-state*, y que eran los nuevos productos electrónicos basados en transistores bipolares. Esos productos, que eran radios de transistores, operaban con voltajes de 3, 6, y 9V. En los 70s, por otro lado, el desarrollo de la tecnología MOS permitió integrar en silicio sistemas electrónicos cuya densidad de integración era de varias decenas de miles de transistores. Este nivel de integración, que operaba con voltajes de alimentación no mayores a 5V, impulsó el desarrollo de nuevas tecnologías MOS con la característica de integrar sistemas de señal mixta, es decir circuitos para el procesamiento analógico de señales y digital también. Las soluciones desarrolladas bajo este

esquema tomaban ventaja de la alta densidad de integración de los sistemas digitales, porque es bien sabido, las técnicas digitales de procesamiento presentan un desempeño que no logra su contraparte analógica. En la práctica, la posibilidad de integrar cada vez más sistemas en silicio no debe verse como un macro-sistema que sólo crece –en área– según los sistemas que lo forman, sino como la integración de un mayor número de sistemas en una misma área de silicio. En otras palabras, la integración de muchos sistemas es posible porque la miniaturización es una característica de las tecnologías de fabricación de CIs, de ahí que hoy día se les identifique más por la mínima dimensión que es fabricable (y reproducible también), y no tanto por el sistema que se desarrolla. A manera de comparación, el diámetro promedio de un cabello humano es aproximadamente $50\mu\text{m}$, y la dimensión mínima que podía reproducirse, a nivel CI, al inicio de los 70s era $10\mu\text{m}$; al final de la misma década la dimensión mínima que se logró fue $3\mu\text{m}$. En los 80s las tecnologías fueron de $2\mu\text{m}$ y concluyó la década con tecnologías de $0.7\mu\text{m}$ [1]. De ese tiempo a la fecha las tecnologías han sido desarrolladas de $0.5\mu\text{m}$ a 90nm , y la alimentación de los sistemas ha sido de 5V a 1V , respectivamente [2]. Este desarrollo explicaría la existencia en el mercado de Laptops, USBs, Ipods, teléfonos celulares, etc., pero no debe pasarse por alto que esos productos solo son resultado de la ingeniería aplicada, es decir, de una ingeniería que aprovecha las posibilidades de las tecnologías de fabricación, y el hecho de *aprovechar* implica que el hombre ha generado un conocimiento científico, conocimiento que no es sino un intento racional por conocer y controlar –en la medida de lo posible– los fenómenos físicos.

El propósito de esta contribución, más allá de citar el desarrollo y tendencias de las tecnologías de fabricación de CIs, es presentar ese otro escenario relacionado con la formación de ingenieros de diseño de CIs, y específicamente describir de qué manera el desarrollo de la tecnología impacta en la actualización curricular de los planes de estudio tanto en las instituciones

de educación superior, como en los centros de estudios de posgrado. Sin embargo, si bien es cierto que el autor no posee autoridad para hablar de los planes de estudios universitarios, también es cierto que un alto porcentaje de las instituciones de educación superior no consideran el diseño de CIs en silicio, y aquellas que lo ofrecen son mayoritariamente centros de estudio de posgrado. En uno de ellos, la Unidad Guadalajara de CINVESTAV, el autor ha impartido los cursos Diseño de Circuitos Analógicos I, y II, y en éstos se han incorporado las facilidades tecnológicas de los procesos de fabricación a través de un convenio con MOSIS [3]. Estas facilidades, por un lado, permiten a los ingenieros en formación no solo diseñar circuitos y sistemas a nivel CI, sino someter a fabricación los proyectos académicos, lo que permite –en consecuencia– verificar el desempeño. La posibilidad de fabricar y caracterizar permite cerrar el flujo de diseño, y más importante: se logra la comprensión del fenómeno con la práctica experimental.

¿Cuál es el problema? En ingeniería, como en toda actividad relacionada con la comprensión de los fenómenos físicos, el uso de *software*, información técnica, y modelos de diseño son fundamentales hoy día. En lo que al diseño de CIs CMOS se refiere, la información tecnológica consta fundamentalmente del modelo del transistor para simulación spice. Este modelo, por un lado, hoy día consta de más de 90 parámetros físicos y tecnológicos que hacen imposible su uso como parámetros de diseño en ejercicios a-mano; este modelo se conoce como LEVEL=49 y se recomienda su uso para evaluar el desempeño de circuitos en régimen de pequeña señal. A la par de ese inconveniente, por otro lado, está la ausencia de modelos analíticos de diseño en los que puedan incorporarse esos parámetros, de ahí que el modelo del transistor sea de utilidad solo para propósitos de simulación. Entonces, en este escenario no es fácil determinar cómo llegar a la etapa de simulación de un circuito CMOS, cuando el desempeño de éste es función del valor de la relación W/L , donde W y L son el ancho y largo del canal de inversión de cada transistor. Es

decir, obtener el valor de esas relaciones parece imposible ante la falta de modelos analíticos. O lo que es lo mismo, ¿cómo fueron diseñadas las celdas básicas que ofrece el fabricante? ¿cómo fueron diseñados los circuitos elementales de los productos comerciales de comunicación de venta en el mercado? O en la academia, ¿qué hacer para diseñar un sistema, por ejemplo, de conversión analógico-digital si no hay modelos fundamentales de diseño?

¿Cómo proceder? Los modelos analíticos existentes, adecuados para ejercicios a-mano, y orientado al diseño de transistores MOS son esos que se propusieron al inicio de los 70s, modelos que corresponden al denominado LEVEL=1, y conocido como modelo de simulación Shichman-Hodges [4]. Usar esos modelos pareciera ser un error de método, y más cuando el fabricante no recomienda diseñar incluso con los parámetros del modelo LEVEL=3. Entonces, si LEVEL=49 es la correcta representación del desempeño del transistor en tecnologías submicrométricas, pero carece de modelos analíticos para diseñar circuitos y sistemas a-mano, en esta contribución se propone una metodología para diseñar circuitos CMOS basada en modelos analíticos LEVEL=1, y simulación de circuitos y sistemas con LEVEL=49. La metodología, soportada en una aproximación resistiva, se presenta –en esta contribución– en el diseño de amplificadores, y en el que las especificaciones de la ganancia son tomadas en cuenta no solo para definir el valor de cada relación W/L , sino para establecer el punto de polarización de los amplificadores. Para el análisis en cd el uso de la ley de Ohm es necesario, y también lo es su equivalente en el dominio de la variable de Laplace para analizar el desempeño de los circuitos en el dominio de la frecuencia. Las facilidades en esta contribución corresponden a un proceso de fabricación CMOS estándar, $0.5\mu\text{m}$, $\pm 2.5\text{V}$, pozo N.

II. Modelado del transistor MOS

El modelo analítico pionero es un intento para describir las regiones básicas de operación del transistor MOS en régimen de fuente inversión, a saber: corte, lineal, y saturación [5]:

$$I_{DS} = \begin{cases} 0 & V_{GS} \leq V_{TH} \\ \mu C_{ox} \frac{W}{L} \left(V_{GS} - V_{TH} - \frac{V_{DS}}{2} \right) V_{DS} & V_{DS} < V_{GS} - V_{TH} \\ \frac{\mu C_{ox}}{2} \cdot \frac{W}{L} (V_{GS} - V_{TH})^2 & V_{DS} \geq V_{GS} - V_{TH} \end{cases} \quad (1)$$

Obsérvese, por un lado, que el modelado del transistor corresponde a un transductor, es decir, la corriente I_{DS} es resultado de los voltajes aplicados. Sin embargo, si el voltaje de compuerta V_{GS} no es mayor que el voltaje de encendido V_{TH} , el transistor se considera apagado (o en corte) y no circula corriente alguna. Por otro lado, si $V_{GS} > V_{TH}$ el transistor está en conducción y se dice que opera en la región lineal si $V_{DS} < V_{GS} - V_{TH}$, caso contrario opera en saturación. Es de notar que la diferencia de voltaje $V_{GS} - V_{TH}$ establece el rango de valores para el cual el transistor, en conducción, opera en una u otra región; esa diferencia de potencial es conocido como V_{DSAT} . Sin pérdida de generalidad, suponer que se desea obtener de (1) la característica $I_{DS} - V_{DS}$ de un transistor NMOS con $W/L = 6.0\mu\text{m}/1.8\mu\text{m}$ y $V_{GS} = 2\text{V}$. Entonces, considerando el proceso de fabricación descrito, el producto μC_{ox} suele ser un valor tecnológico proporcionado por el fabricante, y lo mismo es cierto para el valor del voltaje de encendido. Esos parámetros, que pueden consultarse en [2], al ser sustituidas en (1) permiten obtener las curvas que se muestran en la Fig. 1.

Igualando las expresiones para la corriente en la región lineal y saturación, se obtiene una expresión para V_{DSAT} ; éste está representado por la recta vertical, letra **d**, en la Fig. 1. La curva **b**, en la misma figura, es el modelo analítico que aproxima a la *región lineal* de operación del transistor; el rango de valores para la región lineal es $V_{DS} < V_{DSAT}$. Obsérvese que para valores fuera de rango se presenta una reducción en la magnitud de I_{DS} ; tal comportamiento –teórico– es debido a la naturaleza parabólica mostrada en (1). Luego, la curva **a** es el modelo analítico que aproxima la operación en la región de saturación, donde $V_{DS} \geq V_{DSAT}$.

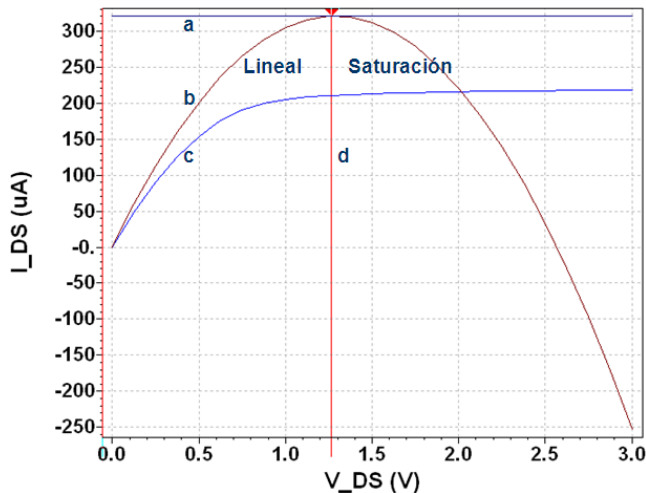


Fig. 1. Comparación de desempeño entre modelo teórico y simulación spice.

Nótese que eliminando las extensiones de los modelos sin significado físico (ver Fig. 1), resulta una aproximación para $I_{DS}-V_{DS}$ que no corresponde a la obtenida de simulación (letra **c** en la Fig. 1). La conclusión inmediata es que los modelos analíticos $LEVEL=1$ no aproximan el desempeño obtenido del simulador (usando $LEVEL=49$).

Por otro lado, y considerando el desempeño en la región de saturación, nótese que la corriente aproximada –por el modelo simple– es independiente del valor de V_{DS} , independencia que no se aprecia en la respuesta obtenida del simulador. Esta diferencia, entre respuestas, suele minimizarse al incluir un factor lineal, $1+\lambda V_{DS}$, en los modelos mostrados en (1). No obstante y aun cuando ese factor solo es un intento por explicar un fenómeno atribuido al *efecto de modulación de canal*, no minimiza el error en la estimación de la corriente –teórica y simulada– que se observa en la Fig. 1; el valor de λ suele obtenerse del tratamiento estadístico de curvas experimentales, o alternativamente de curvas de simulación. Sin embargo, nótese que la inclusión de ese parámetro no evita que el desempeño del transistor, en saturación, se aleje del comportamiento deseado, comportamiento equivalente a una fuente de corriente (independiente del valor de V_{DS}). La Fig. 2, por otro lado, muestra la simulación $I_{DS}-V_{DS}$ para el

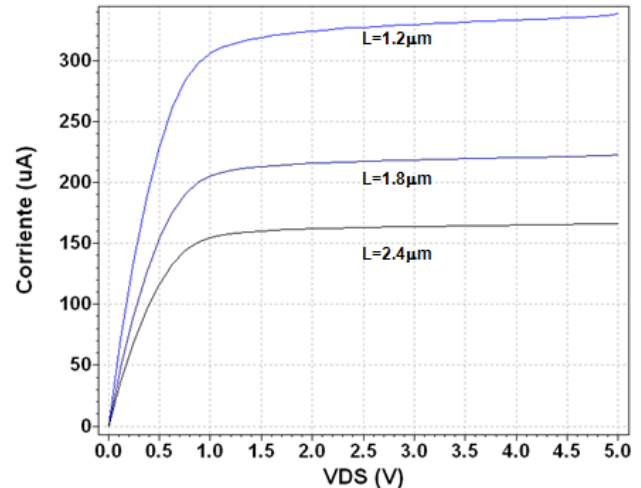


Fig. 2. Impacto del valor de L en la característica $I_{DS}-V_{DS}$.

transistor descrito en la Fig. 1, pero incluyendo diferentes valores de L . Obsérvese que para longitudes de canal menores a $1.8\mu\text{m}$, el transistor no solo conduce más corriente sino que el efecto atribuido a la modulación de canal es más notorio. Por el contrario, aumentos en el valor de L permiten considerar al transistor como una fuente de corriente con poca dependencia de V_{DS} . La conclusión inmediata es que el desempeño del transistor, según (1), se aproxima a la tendencia obtenida de simulación si la longitud de canal es más grande que el mínimo permitido por la tecnología, L_{min} . En otras palabras, (1) puede usarse para diseñar circuitos analógicos –en cualquier tecnología– siempre que $L \gg L_{min}$, lo que es contrario al diseño de circuitos digitales porque en éstos se desea reducir el retardo de la señal; en diseño analógico otros criterios aplican para aumentar el denominado *slew-rate*.

¿Cómo proceder en el diseño analógico? En esta contribución, mientras no se indique lo contrario, el diseño de circuitos –para el procesamiento de señales analógicas– es con transistores que operan en la región de saturación. El modelo de diseño suele llamarse *ley cuadrática* y es una simplificación del modelo mostrado en (1):

$$I_{DS,i} = KP_i \left(\frac{W}{L} \right)_i (|V_{GS}| - |V_{TH}|)^2 \quad (2)$$

donde i denota la conductividad del transistor (tipo **p** o tipo **n**). Pareciera un error de método usar (2) luego de observar que su uso como modelo de diseño a-mano, no reproduce la magnitud de la corriente obtenida del simulador. Sin embargo, ante la falta de modelos de diseño, la aproximación que ofrece (2) se ha usado para modelar el voltaje V_{TH} [6], [7] y la magnitud del ruido $1/f$ [8], entre muchas otras aplicaciones.

Del punto de vista del modelado eléctrico equivalente, un transistor que opera en la región de saturación se le representa como una fuente de corriente controlada por voltaje en paralelo con un resistor de gran valor resistivo. Obsérvese que la alta impedancia de entrada del transistor se modela como un circuito-abierto entre las terminales de *drenaje* (D) y *compuerta* (G). Esa representación se ve en la Fig. 3a, donde la transconductancia, está dada por

$$\frac{\partial I_{DS,i}}{\partial V_{GS}} = g_{m,i} = 2KP_i \left(\frac{W}{L} \right)_i (|V_{GS}| - |V_{TH}|) \quad (3)$$

Es de notar que un transistor que satisface $V_{DG}=0$ opera en saturación, es decir, su circuito eléctrico equivalente de pequeña señal corresponde a un resistor de valor $1/g_{mn}$ porque en esa región de operación $g_{mn} \gg g_{ds}$; la misma condición se satisface para un transistor PMOS.

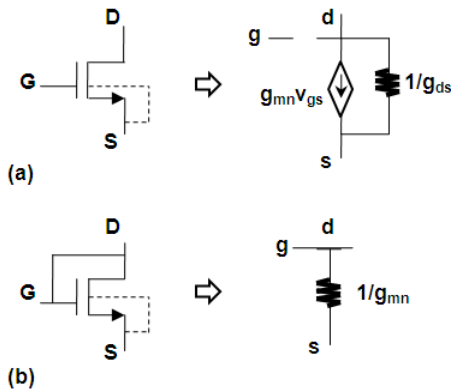


Fig. 3. Modelo eléctrico equivalente para un NMOS en lazo abierto (a), y en configuración diodo (b).

La Fig. 4 muestra la simulación del transistor, citado en la Fig. 1, pero ahora con $V_{DG}=0$. Obsérvese que esta respuesta es la responsable de porqué al transistor, en ese arreglo, se le denomina *configuración-diodo*.

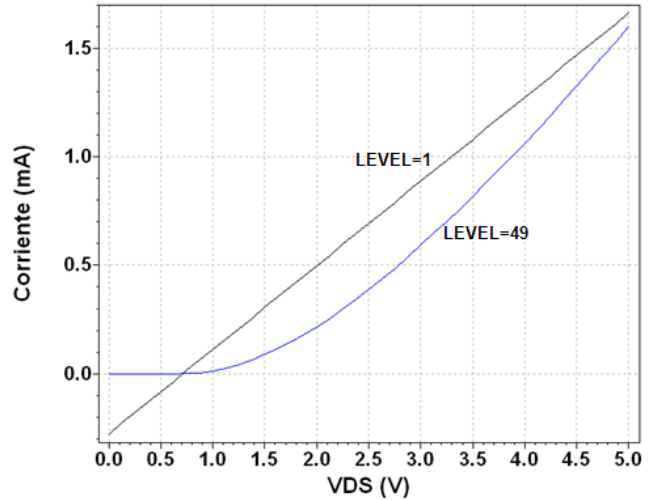


Fig. 4. Característica I_{DS} - V_{DS} para un NMOS en configuración diodo.

De la Fig. 4 se pueden apreciar dos propiedades básicas: la característica I_{DS} - V_{DS} , según LEVEL=49, es la de un dispositivo no-lineal y difiere de la obtenida con (3). Sin embargo, como se mostrará más adelante, es viable usar un transistor en *configuración-diodo* en rededor de un valor específico de V_{DS} de manera que pueda aproximarse a un comportamiento lineal. Es decir, definido el punto de polarización, Q , el desempeño del transistor se aproximará al descrito por la ley de Ohm. Nótese que aunque (3) obedece esa ley, la idea en esta contribución es usarle como aproximación al desempeño obtenido con simulaciones LEVEL=49.

III. Aproximación resistiva

Esta técnica de diseño se basa en el uso de la ley de Ohm en redes activas MOS, donde cada transistor se configura como diodo. Sin pérdida de generalidad la Fig. 5a muestra un circuito resistivo, y donde $V_{REF,1}$ y $V_{REF,2}$ están bajo el control del diseñador. De la ley de Ohm es fácil verificar que la corriente I_{BIAS} está dada por

$$I_{BIAS} = \frac{V_{DD} - V_{REF,1}}{R_1} = \frac{V_{REF,1} - V_{REF,2}}{R_2} = \frac{V_{REF,2} - V_{SS}}{R_3} \quad (4)$$

Y de la cual se obtienen las siguientes relaciones:

$$\frac{R_1}{R_3} = \frac{V_{DD} - V_{REF,1}}{V_{REF,2} - V_{SS}} \quad \wedge \quad \frac{R_2}{R_3} = \frac{V_{REF,1} - V_{REF,2}}{V_{REF,2} - V_{SS}} \quad (5)$$

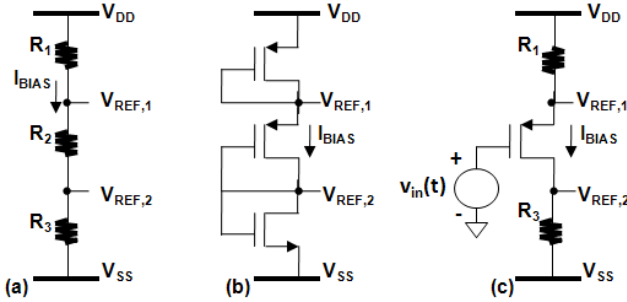


Fig. 5. Transformación de un circuito resistivo a otro activo usando transistores MOS.

Considerando que $V_{SS} = -V_{DD} = 2.5V$ y proponiendo $V_{REF,1} = \frac{1}{2} \cdot (V_{DD} + V_{REF,2})$ y $V_{REF,2} = \frac{1}{2} \cdot (V_{DD} + V_{SS})$, se puede demostrar que las relaciones resistivas que satisfacen esos valores son las siguientes:

$$\frac{R_1}{R_3} = \frac{R_2}{R_3} = \frac{1}{2} \quad (6)$$

Este resultado es importante porque postula que no importa el valor de cada resistencia, sino su relación. Luego, si cada resistor es sustituido por un MOS en configuración diodo (ver Fig. 5b), se obtienen relaciones de transconductancias:

$$\frac{g_{m3}}{g_{m1}} = \frac{g_{m3}}{g_{m2}} = \frac{1}{2} \quad (7)$$

Esta transformación significa que aun cuando la pendiente que muestra cada respuesta, en la Fig. 4, es diferente, lo más importante es que todo circuito diseñado con LEVEL=1 debe generar una relación de transconductancias equivalente al que resulta de simulaciones LEVEL=49. Entonces, sustituyendo (3) en (7) resulta que

$$\frac{g_{mn,3}}{g_{mp1}} = \frac{2KP_n \left(\frac{W}{L}\right)_3 \left(|V_{REF,2} - V_{SS}| - |V_{TH,n}|\right)}{2KP_p \left(\frac{W}{L}\right)_1 \left(|V_{REF,1} - V_{DD}| - |V_{TH,p}|\right)} \quad (8)$$

y al sustituir los valores propuestos, $V_{REF,1} = 1.25V$ y $V_{REF,2} = 0$, se obtiene una relación de razones geométricas en función de parámetros – tecnológicos y de diseño– conocidos:

$$\frac{\left(\frac{W}{L}\right)_3}{\left(\frac{W}{L}\right)_1} = \frac{1}{2} \cdot \frac{KP_p}{KP_n} \cdot \frac{1.25 - |V_{TH,p}|}{2.5 - |V_{TH,n}|} \quad (9)$$

Cabe señalar que el valor del voltaje de encendido, proporcionado por el fabricante, corresponde a $V_{BS} = 0$, es decir, las terminales de *fuelle* (*S*) y *substrato* (*B*) están cortocircuitadas. Del punto de vista del diseño de CIs, y considerando la tecnología de fabricación ya mencionada, la condición $V_{BS} = 0$ es intrínseca al diseño mismo porque M_{n3} se fabrica directamente en el substrato tipo **p**. La misma condición se cumple para M_{p1} porque éste se fabrica en un pozo N. Sin embargo, si M_{p2} –un transistor flotante– se fabricara en el mismo pozo su voltaje de encendido no cumpliría la condición $V_{BS} = 0$, y habría que calcular su voltaje de encendido:

$$V_{TH-p} = V_{TH,p} + \gamma \left(\sqrt{|2\Phi_f + V_{SB}|} - \sqrt{|2\Phi_f|} \right) \quad (10)$$

donde γ es una constante y Φ_f es el potencial superficial [6]. En la práctica, porque el cálculo de esos parámetros físicos no es inmediato, se recomienda integrar un pozo adicional para M_{p2} . Consecuentemente, la relación geométrica entre M_{p2} y M_{n3} está dada también por (9). Este resultado es un modelo simple de diseño, útil para análisis en cd, adecuado para diseñar a mano, y demanda cuatro parámetros tecnológicos; el valor de esos parámetros puede consultarse en [2]. Sustituyendo valores, y proponiendo $L_1 = L_2 = L_3$, se obtiene la siguiente relación de anchos de canal:

$$\frac{W_1}{W_3} = 30.828070 \quad (11)$$

Este resultado muestra que proponiendo un valor para W_3 se obtiene el valor de W_1 . Sin embargo, todo valor que se proponga y/o que resulte debe ser múltiplo de la longitud unitaria de diseño, λ . Para la tecnología descrita $\lambda = 0.3\mu m$; si $W_3 = 6.0\mu m$ entonces

$$W_1 = W_2 = 184.968421 \dots \mu m \approx \begin{cases} 184.8\mu m \\ 185.1\mu m \end{cases} \quad (12)$$

donde los valores ajustados son los valores límites (inferior y superior) del valor original. Es de notar que el hecho de ajustar a múltiplos de λ conlleva un error que se refleja –a nivel simulación– en los valores de los voltajes de referencia propuestos. Entonces, si el ajuste se

interpreta como un *offset*, V_{off} , puede estimarse el ancho que resulta para los transistores si se considera un *offset* de hasta $\pm 100\text{mV}$.

La Tabla 1 muestra los valores de simulación obtenidos para los diferentes valores de W . Nótese que no se cuestiona la verosimilitud del simulador por cuanto éste es una extensión –a nivel de software– del conocimiento alcanzado en el campo de los dispositivos semiconductores, es decir, el simulador no es sino la incorporación de modelos que intentan sustituir diversos fenómenos materiales que se estudian por medio de abstracciones [10]; en esta contribución se usa Tspice versión 5.03, 1997.

W_3 (μm)	W_2 (μm)	W_1 (μm)	$V_{REF,1}$ (V)	$V_{REF,2}$ (mV)
6.0	242.1	242.1	1.19	-110
		242.4	1.19	-110
	184.8	184.8	1.15	-194
		185.1	1.15	-194
	152.4	152.4	1.12	-255
		152.7	1.12	-254

Tabla 1. Resultados de simulación Spice [9] para el diseño de experimentos propuesto.

En la práctica, cuando el circuito activo (ver Fig. 5b) es usado como base para el diseño de una etapa de amplificación, es fácil concluir que $V_{REF,2}$ representa el voltaje en cd del nodo de salida. Idealmente debiera ser un nodo de 0V, pero al existir un corrimiento *offset*, se seleccionaría el diseño que presente menor *offset*, y ese corresponde a $W_2=242.1\mu\text{m}$; en ese diseño $I_{BIAS}\approx 200\mu\text{A}$ según resultados de simulación, donde $L=3.0\mu\text{m}$ para todos los transistores.

IV. Análisis de pequeña señal

Una etapa de amplificación se muestra en la Fig. 5c, donde R_1 y R_3 luego son sustituidos por transistores MOS; el circuito eléctrico equivalente se muestra en la Fig. 6a. Aplicando KCL es fácil verificar que las ecuaciones de nodo son las siguientes:

$$\frac{v_x}{R_1} + (v_x - v_o)g_{dsp2} - g_{mp2}(v_{in} - v_x) = 0 \quad (13a)$$

$$\frac{v_o}{R_3} + (v_o - v_x)g_{dsp2} + g_{mp2}(v_{in} - v_x) = 0 \quad (13b)$$

Despejando v_x de (13a) y sustituyendo en (13b) se tiene la función de transferencia a frecuencia baja:

$$\frac{v_o}{v_{in}} = H_0 = \frac{-g_{mp2}R_3}{g_{mp2}R_1 + g_{dsp2}R_3 + 1} \quad (14)$$

donde v_o y v_{in} son funciones del tiempo; el en Apéndice A muestra el equivalente de (14) en el dominio de la frecuencia. Luego, para propósitos de amplificación se recomienda sustituir R_1 y R_3 por transistores en configuración-diodo o sin él. La Tabla 2 muestra todas las posibles combinaciones, donde un 1 representa un transistor en configuración-diodo, y 0 un transistor en lazo abierto. Considerando el circuito eléctrico equivalente del transistor, es fácil concluir que en un MOS que no recibe señal (como M_{p1} y M_{n3}) su equivalente de pequeña señal corresponde a una conductancia si está en lazo abierto, caso contrario es una transconductancia. Para la configuración de mayor amplificación, obsérvese que el circuito final usa dos circuitos activos equivalentes, uno usado como divisor de voltaje y el otro es propiamente el amplificador. En la práctica, porque los circuitos individuales son réplicas, y del punto de vista del análisis en cd, M_{n3} opera como un transistor con *retroalimentación* porque su potencial V_G es equivalente a V_D ; esta es una propiedad del espejo de corriente formado por los transistores NMOS.

Los resultados de simulación que se muestran comparan los valores obtenidos directamente del simulador y los que se obtienen de cada una de las funciones de transferencia. En éstos se sustituyó cada valor de g_m y g_{ds} obtenido del simulador. En la práctica se esperaría que los valores de g_m cumplieran –con un margen de error– las relaciones mostradas en (7), sin embargo los valores del simulador conducen a relaciones cuyo valor es aproximadamente 0.2, las que justificarían las variaciones en los potenciales obtenidos en los distintos nodos (ver Tabla 1). Un tema de estudio –no analizado en esta contribución– en la metodología de diseño

propuesta es esa que intenta explicar cómo es que las relaciones de transconductancias sí conducen al valor deseado siempre y cuando la longitud de canal del transistor M_{n3} sea el doble del que presenta M_{p2} y M_{p1} , solución que contraviene la condición $L_1=L_2=L_3$ inicialmente propuesta.

M_{p1}	M_{n3}	H_o	$ H_o $	dB	dB_{spice}	ϵ_r (%)
0	0	$H_o^{00} \approx -\frac{g_{mp2}}{g_{dsn3} \left(1 + \frac{g_{mp2}}{g_{dsp1}}\right) + g_{dsp2}}$	2.5	7.99	5.8	37.2
0	1	$H_o^{01} \approx -\frac{g_{mp2}}{g_{mn3} \left(1 + \frac{g_{mp2}}{g_{dsp1}}\right) + g_{dsp2}}$	0.047	-26	-28.94	10.1
1	1	$H_o^{11} \approx -\frac{g_{mp2}}{g_{mn3} \left(1 + \frac{g_{mp2}}{g_{mp1}}\right) + g_{dsp2}}$	4.52	5.6	13.1	57.2
1	0	$H_o^{10} \approx -\frac{g_{mp2}}{g_{dsn3} \left(1 + \frac{g_{mp2}}{g_{mp1}}\right) + g_{dsp2}}$	52.32	34.37	33.89	1.41

Tabla 2. Función de transferencia para diferentes cargas en el amplificador simple.

V. Conclusiones

Cuatro parámetros tecnológicos son necesarios para desarrollar una metodología que conduzca al diseño de circuitos CMOS. Esos parámetros, en conjunción con conceptos fundamentales de la teoría de circuitos, permiten proponer relaciones geométricas que cumplan los voltajes de cd asignados a los distintos nodos. Si bien los resultados de simulación muestran un *offset* de hasta 100mV, también es verdad que el método permite aproximar a los valores deseados y evita, lo que también es un propósito del método, polarizar los circuitos *a prueba-y-error*.

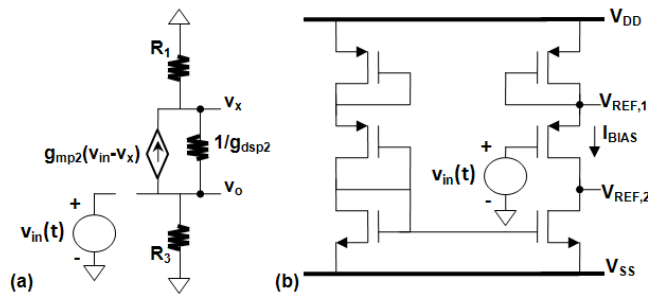


Fig. 6. Circuito eléctrico equivalente para un amplificador inversor con cargas pasivas (a), y diseño CMOS.

La manera en que se usen los circuitos dimensionados, para obtener circuitos con diversas funciones a nivel pequeña-señal, está más en función de la habilidad del diseñador y no de las posibilidades del método propuesto. A la fecha, la metodología de diseño –basada en una aproximación resistiva– se ha usado para instruir a jóvenes universitarios al diseño de circuitos diferenciales [11], circuitos de señal mezclada [12], y a jóvenes graduados para el diseño de sensores de campo magnético [13] y circuitos varios para aplicaciones biomédicas [14].

El método está sujeto a mejoras, fundamentalmente para definir un criterio que permita, por ejemplo, asignar un valor a W_{n3} e incluir especificaciones de diseño, siendo éstas establecidas por la aplicación. En esta etapa del desarrollo el método es útil para introducir a los jóvenes estudiantes al diseño de circuitos integrados en silicio, aplicar los postulados de la teoría de circuitos, y usar el simulador como una herramienta más y no como el medio a través del cuál se *aprende* el diseño de circuitos.

VI. Apéndice A

Una muy útil herramienta en el análisis de circuitos electrónicos es la transformada de Laplace. La transformación por sí misma permite el análisis de circuitos usando las leyes de Kirchhoff –y por ende de la ley de Ohm– y conociendo la impedancia $Z(s)$ de cada componente del circuito bajo análisis. Estas leyes, de uso común en la ingeniería eléctrica y electrónica, facilitan la obtención de modelos analíticos en el dominio de la frecuencia, en contraposición a los modelos que resultan del análisis en el dominio del tiempo vía el planteamiento de ecuaciones diferenciales. Para todo propósito práctico, y considerando que las condiciones iniciales para capacitores e inductores son nulas, las impedancias de esos elementos de red –en el dominio de la variable de Laplace– son las siguientes:

$$Z_C(s) = \frac{1}{sC}, \quad Z_L(s) = sL \quad (15)$$

Este resultado, como todo estudiante de ingeniería eléctrica (y áreas afines) sabe, se obtienen al obtener la transformada de Laplace de las relaciones:

$$i(t) = C \frac{dv(t)}{dt}, \quad v(t) = L \frac{di(t)}{dt} \quad (16)$$

VII. Referencias

- [1] Ch. Hu, "MOSFET Scaling in the Next Decade and Beyond", Semiconductor International, June 1994, 105-114
- [2] www.mosis.org
- [3] Los lineamientos del acuerdo con las instituciones de educación superior pueden consultarse en el *site* de Mosis.
- [4] D. Foty, Mosfet Modeling with Spice, 1^a edición, 1997, Prentice Hall, New Jersey, p. 90
- [5] HSPICE, User's manual, Meta-Software, Inc., 1990, Chapter 7, p. 70-72
- [6] Yan-Yu Huang *et al*, "Highly linear RF CMOS variable attenuators with adaptive body biasing", IEEE J. of Solid-State Circuits, vol. 46, no. 5, May 2011, 1023-1033
- [7] Yao-Huang Kao *et al*, "Study of front end of CMOS RFID tag with inductively-coupled broadband antenna", Proc. of IEEE Asia-Pacific Microwave Conference, 2007
- [8] Y. Nemirovsky *et al*, "1/f noise in advanced CMOS transistors", IEEE Instrumentation and Measurement Magazine, January 2011, 1-10
- [9] Tspice circuit simulator, 1997 Tanner Research Inc.
- [10] A. Rosenblueth, *Mente y cerebro*, Siglo veintiuno, México, p. 14
- [11] R. Peralta-López, F. Sandoval-Ibarra, "Diseño de Circuitos Integrados Analógicos CMOS: Una propuesta", Memoria técnica XII Reunión de Otoño de Potencia, Electrónica y Computación, 10-12 Noviembre 2010, Manzanillo, México
- [12] F. Sandoval-Ibarra, *et al*, "Prácticas Profesionales y Diseño de Circuitos Integrados CMOS: Un Caso de Estudio", Memoria Técnica CONCyE, 1^{er} Congreso Nacional de Computación y Electrónica, 14-15 Abril 2011, Zacatecas, México
- [13] M. A. Dávalos-Santana and F. Sandoval-Ibarra, "A 540 μT^{-1} Silicon-based MAGFET", Rev. Mex. Fís., vol. 53, No. 3, pp. 218-221, June 2007
- [14] José H. Martínez-Aldrete *et al*, "Consideraciones de Diseño en un Amplificador Miller para Aplicaciones de Frecuencia Baja y Ruido 1/f Bajo", Memoria Técnica XVII Workshop Iberchip, Febrero 23-25, 2011, Bogotá, Colombia

VIII. Autor

Federico Sandoval nació en 1964 en la ciudad de San Luis Potosí, México. Recibió el grado Lic. en Electrónica-Física por la Universidad Autónoma de San Luis Potosí en 1988, y el grado de Doctor en Ciencias, con especialidad en Electrónica, por el Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Puebla, México, en 1998. En 1999 se incorporó como investigador-profesor a la Unidad Guadalajara de CINVESTAV, donde inició su participación en el diseño de circuitos integrados de señal mixta en tecnologías CMOS. De 2002 a 2004 fue responsable de los programas especiales de maestría para Siemens y Solectron (Guadalajara). De 2002 a 2006, fue coordinador del Grupo de Diseño Electrónico de la Unidad Guadalajara de CINVESTAV. Sus áreas de interés son el diseño de circuitos analógicos en tecnología CMOS, y el desarrollo de metodologías de enseñanza. El Dr. Sandoval es miembro del SNI.

FPGA BASED PARAMETRIC PLATFORM FOR EVOLVABLE HARDWARE

Juan M. Xicoténcatl-Pérez, Arturo Lezama, Rubén O. Hernandez, Miguel Liceaga

Universidad Politécnica de Pachuca. Km. 20 Carretera Pachuca Cd. Sahagún, Zempoala, Hidalgo,
México. jmxico@upp.edu.mx

Abstract

Evolvable hardware is a special hardware that can reconfigure its structure autonomously. Conventional designs for evolvable hardware include embedded processors or external computer resources to achieve the evolvable engine. In this paper, a platform for evolvable hardware applications is described. The proposed architecture implements an intrinsic evolution algorithm without dependences of the application nature or the number of population members and it does not need any external resources to compute the artificial evolution process. The platform is implemented in an FPGA device and validated with a function optimization problem. Performance and parallel scaling are discussed.

Keywords: FPGA, Evolvable hardware, Genetic Algorithms

I. Introduction

In conventional hardware, the structure is irreversibly fixed in the design process and it cannot be altered in response to environmental changes. Reconfigurable devices, like FPGAs (Field Programmable Gate Arrays) allow hardware reconfiguration on the fly as it executes a task. However, users must specify the details of the reconfiguration, so a limited amount of adaptation is gained. On the other hand, Evolvable Hardware (EHW) can reconfigure its hardware structure autonomously, i.e., the EHW is designed to adapt itself. These capacities for adaptation are achieved by employing efficient search algorithms known as genetic algorithms (GA). A GA handles a population of possible solutions where each solution is represented as a chromosome. The chromosome is a binary string representation of a solution and it could represent the architecture configuration of a programmable logic device. Depending on the architecture bit string, the logic device will behave differently. The performance of all these logic structures on a given tasks evaluated in parallel and the resulting performance is the fitness measurement of the possible solutions to the task. Commonly the evolvable hardware (EHW) is based on primitive gates such as AND-gate and OR-gate [1]. However, this

technique has the common limitation that the evolved problem is small. To overcome this problem, an evolvable hardware at function-level is proposed in this work.

This paper proposes a platform for generic genetic applications. The proposed platform includes all the necessary circuitry to implement the genetic engine on chip. The paper is organized as follows. Section 2 describes the basic concepts of EHW. In sections 3 and 4 detail the architecture implementation using VHDL as design and simulation tool. Section 5 outlines an application of the chip as function optimizer. Finally, the performance of the proposed architecture and future works are discussed.

II. Background and related works

Evolvable hardware may be considered as a technology where its architecture has the ability to adapt itself to a dynamically changing environment. In this kind of systems, the elements connection of a reconfigurable platform are autonomously synthesized by a genetic algorithms. These interconnections are coded into chromosomes defining possible solutions to an application. Traditionally programmable logic, like Field Programmable Gate Arrays (FPGA), has been used to carry out

the experiments of the hardware evolution at different granularity levels. The hardware reconfiguration ability of FPGA devices is used to implement and evaluate different hardware solutions to a given task. In conventional EHW, the adaptation is executed by software on a host computer and only the best configuration is downloaded to hardware in each generation (extrinsic evolution). However, this method is not always practical due to slow evolution speed and large system sizes. On the other hand, intrinsic EHW simulates evolution directly in its hardware. The system currently presented in this paper uses intrinsic evolution and every chromosome is used to reconfigure the hardware in a parallel way.

There are a variety of applications of EHW reported in the literature. In [2] it is proposed an evolvable reasoning hardware system where the circuits are synthesized from true tables converted from the case database, and the rules behind the case data are embedded directly in circuits through genetic algorithms. The previous hardware is classified as extrinsic EHW because it carries out the evolutionary computation. The EHW proposed in [3] implements the GA operations in hardware, and integrates the GA hardware together with the reconfigurable device within a single LSI circuit. The chip is used as an artificial hand controller to recognize appropriate hand actions using myoelectric signals. Another applications of the chip is in adaptive navigation tasks for a real world mobile robot that must track a moving colored ball while avoiding obstacles.

The architecture described in [5] uses a genetic algorithm to evolve minimum-cost circuits from a population of initially random circuits. Evolution is first driven by the relative correctness circuits that have been generated, evolution is driven by gate number of the circuits.

In most of the previous works there is a common factor: the GA is implemented in hardware by employing a steady-state population with random parent selection where fitter children replace lesser-fit parents in the population [6]. This is in contrast to Holland's original GA, which uses a generational population scheme with parent selection

according to fitness to promote evolution. The hardware adaptation mentioned earlier serializes a parallel algorithm. It is possible to note that this method always requires a personal computer or workstation, which increases the size of the

system. In others EHW, the GA is implemented using an embedded processor. But, such applications only translate the external sequential genetic software to an internal processor.

In our work, the proposed architecture implements all possible solutions in a two dimensional array with all genetic operators integrated in an inherent way. Furthermore, the platform solves the communication problems between members in a population that a serial or sequential genetic algorithm avoids, because of the necessity to establish interconnection lines among arbitrary elements under evolution. The proposed architecture segments a high density FPGA in similar modules with different implemented architectures, increasing the performance due to the parallel implementation.

III. The EHW platform architecture

The proposed architecture is divided in some structural levels. The low level includes a set of modules that supports some stages of the evolution process: the chromosome module, the fitness module, the selection module and the hardware application. This set is joined with the application hardware to form the evolvable cell. The chromosome module stores the chromosome data that can be accessed for the application. The fitness module computes the accumulative fitness, which is necessary for the selection stage. The selection module contains the state of the cell; such state indicates if the cell can be selected or no selected for future processing. The last part of the cell, the hardware application implements the test hardware for the chromosome and the selection criteria at the same time. A cell set with a priority decoder and a bus arbiter is called an evolvable row. An evolvable row implements a local control over the cells to avoid time consume and lines from the global control. A set of evolvable rows plus additional logic like an arbiter bus, encoders and decoders is called an evolvable array. Here, the evolvable array

contains in a single module a designer-defined cell number. Finally, the evolvable array plus the global control and the PRNG (Pseudo Random Number Generator) define the evolvable architecture, which is reported in this paper. The architecture modules operate concurrently with each other and together form a course-grained pipeline. All modules are written in VHDL and are independent of the implementation technology (e.g. Xilinx FPGAs or fabricated chips). Figure 1 shows the block diagram of the system and the structural levels discussed before.

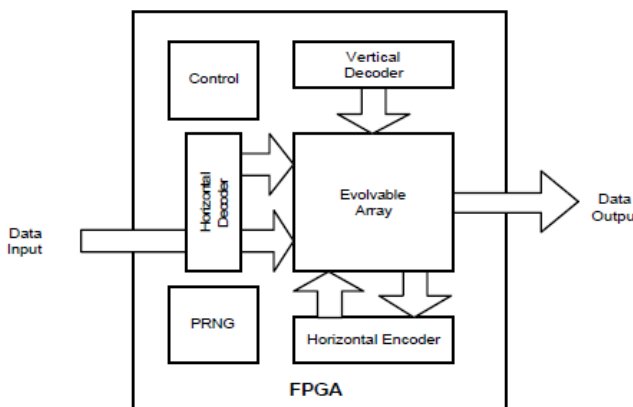


Fig. 1. Block diagram of the proposed EHW platform. The evolvable array is a bi-dimensional cell array

In fig. 1, the evolvable module integrates a memory, which can be updated using a global fast bus, and the necessary hardware to store and to manage the chromosome. The application hardware evaluates the chromosome fitness. The fitness module computes the selection probability for the cell in the selection stage. Furthermore, this module has special interfaces with cells in its neighborhood to obtain the accumulative fitness, which is used in the selection probability. Finally, the selection module sets internal flags that contain the cell state. The state indicates to the global control if the cell has been selected to be a parent or not. The evolvable row adds a bus arbiter to a one-dimensional cell array and a special control to generate a local control sequence to update the cells. The bus arbiter gives a cell access to the row output bus using a priority table. At this level, the evolvable row can evolve a design in a one-dimensional scheme. The next structural level in the architecture corresponds to the

evolvable array. In the same way as the evolvable row, the evolvable array includes a designer-defined number of rows. Additional circuitry is necessary to facilitate communication among arbitrary cells and to allow genetic material interchange.

The global control indicates to the evolvable array the actual processing stage in GA execution. This control does not give a complete sequence control for each cell in the array, because cells can work in different ways at the same time. Finally, the pseudo random number generator based upon neighborhood of four cellular automata (CA) with asymmetrical, non-local connections is selected for the architecture given its characteristics in GA applications. A neighborhood size of four allows a single CA cell to be implemented with a four-input lookup table and a one-bit register. With these characteristics, the FPGA implementation is compact and the implemented PRNG can produce a 64-bit random number for each clock cycle.

IV. Workflow

The proposed design presents in essence the classical GA structure but it executes each phase in a different form and in parallel way (see figure 2). Here, the evaluation phase is divided in local and global fitness computation and the selection stage is divided in roulette selection and data dispersion.

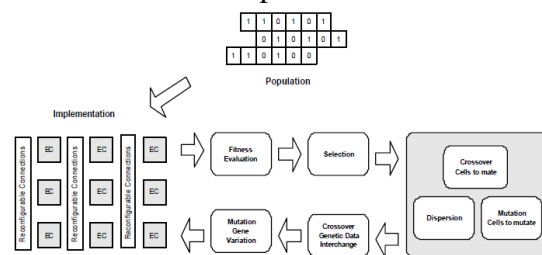


Fig. 2. The flow diagram indicates the processing steps in the proposed architecture

In the fig. 2 the GA is represented the GA in an abstract way. However inside the architecture, genes from the different chromosomes are stored in an internal memory associated to each cell and the fitness value for the proposed chromosome is computed in the evaluation phase using the hardware application. The initial population generation takes $2MN$ clock

cycles, where M and N are the array of cell dimension and the $M \cdot N$ product is equal to the population members. The evaluation time is independent of the number of members, so the evaluation time is constant in the entire process. There are two ways to compute the global or accumulative fitness: the classical way used in PCs or embedded microprocessors or, a parallel implementation, which is based in the partial sums to obtain the accumulative fitness. Using this alternative, the required processing time is

$$T_{acc_fitness} = (M+N-2)/f \quad (1)$$

In (1) f is the system frequency. The roulette technique selects the fitter members in the population in first place. For the second phase, it disperses the genetic content from the selected members to create the offspring. The first part in the modified roulette implementation takes a computing time of

$$T_{roulette} = 2M \cdot N / f \quad (2)$$

In the dispersion stage, an unmodified offspring is created. The time required to compute the dispersion is

$$T_{selection} = 2(N_{genes} + N_{selected_members}) / f \quad (3)$$

Where N_{genes} and $N_{selected_members}$ are the number of genes for the chromosome and the number of times that a member was selected by the roulette to generate an offspring. The selection stage second phase has four steps to propagate the genetic information from a selected cell in the array to non-selected cells.

Step 1. A cell in the array is selected with the modified roulette

Step 2. The selected cell propagates their chromosome to others non-selected cells, in order to create unmodified copies of itself

Step 3. An additional value is concatenated to the selected cells. It contains the number of times that a cell was selected. The global control takes this number for step four.

Step 4. The selected cell and copied cells are inhibited to avoid further reprocessing.

The offspring generation has a variable processing time, since it depends on the number of selected cells. For multiple selected cells, the previous four steps are executed in sequence for each selected cell. When there are not selected cells the selection stage ends. The crossover stage interchanges genetic data between two random selected chromosomes. The architecture compensates the random selection time implementing this phase in parallel with the selection stage, because the selection stage does not use the PRNG, so it is free for other process. One of the parameters of a genetic system is probability of crossover p_c . This probability gives the expected number $p_c \cdot M \cdot N$ of chromosomes which undergo the crossover operation.

Architecture proceeds in the following way, for each chromosome in the array: the PRNG generates a random number r from the range $[0,1]$ and If $r < p_c$, the architecture selects given chromosome for crossover. Finally, for each pair of coupled chromosomes we generate a random integer from the range $[1, m-1]$, (m is the total length of genes in a chromosome). The number indicates the position of the crossing point. All this data are stored in a crossover memory, which is external to the evolvable array. The second crossover phase obtains the genetic material from selected chromosomes to mate them. For each chromosome pair, this process takes place. The crossover stage is implemented in

$$T_{crossover} = p_c (2 \cdot N_{genes} + 1) M \cdot N / f \quad (4)$$

In this stage, an interconnection is generated between two arbitrary cells to interchange genetic contents. Because of the random connection, structured connections like used in systolic arrays do not work, because structured connections take more time to pass the genetic data. At the same time that the first phase of the crossover stage is been processed the mutation is performed on a gene-by-gene basis. Another parameter of the genetic system, probability of mutation p_m , gives to the architecture the expected number of mutated genes $p_m \cdot M \cdot N \cdot N_{genes}$. For each chromosome in the current array and for each gene within the

chromosome: the PRNG generates a random number r from $[0,1]$ and if $r < pm$, mutate the gene. Like selection and crossover stages, the mutation stages is divided in two phases, the second phase is similar to the crossover second phase and it takes

$$T_{mutation} = 4 pm \cdot M \cdot N / f \quad (5)$$

V. RESULTS

The platform was implemented in a modular way and in a structured code. A hierarchy tree from the design is shown in fig. 3. In the evolvable array, only a cell is specified and the synthesizer replicates it $M \cdot N$ times. The bus size among cells is parametric and it depends from the application. In the implementation reported here, the application takes numbers in the range $[-5.25, 5.25]$; so, the implementation uses a 16 bits bus in fixed-point notation with sign. The architecture only implements and distributes data among cells in an evolvable scheme, but it is independent about the data application.

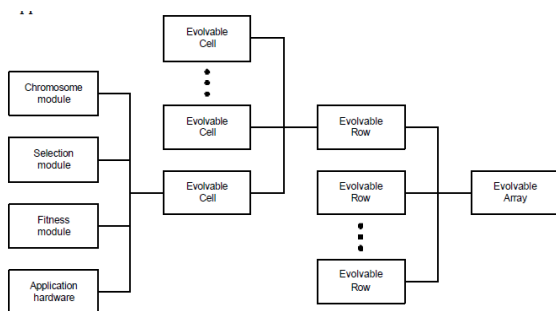


Fig. 3. There are four hierarchy levels in the architecture; each level contains arbiter buses to gain access to further levels.

Cells are distributed in an $M \cdot N$ array. Arbiter buses contain priority tables with random data for the selection stage. Actually, there are two control versions in the architecture: a wired control for permanent applications and a microcoded control for test. The proposed architecture is completely implemented in a FPGA, so, the outside world interface is a single output 16 bits bus and three control lines. START, ACK (ACKnowledge) and OBF (Output Buffer Full). Parameters like the PRNG seed pc and pm are designer-defined, since the

design targets an autonomous on-chip evolution. In order to validate the architecture a non trivial application was implemented in the evolvable platform. This application requires finding the maximum of a five variable function, which is represented in equation 6. The equation is a particularly interesting since it has multiple local minimums, it is not differentiable in some points and where gradient based methods fails. This function belongs to five functions used commonly as test bench. These problems have been popular in the field of evolutionary computation because they represent different kinds of complexity in function optimization. The function expressed in 6 is illustrated in fig. 4a in a specific range with three variables.

$$f(x_1, x_2, x_3, x_4, x_5) = \sum_{i=1}^5 \text{int}(x_i) \quad (6)$$

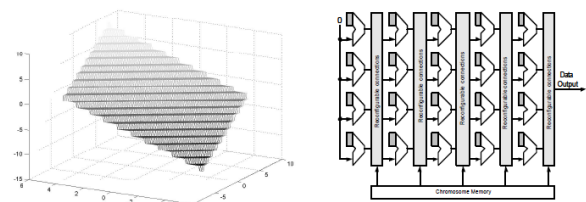


Fig.4. a) Three dimensional graph of the used function in the design and b) The function level evolvable array

An implementation block diagram of the evolvable hardware for this application is shown in the fig. 4b. The hardware exists in each cell of the bi-dimensional array and consists of a bi-dimensional array of accumulators with a predefined content. Connections in this array are in evolution. Here, a function level evolvable hardware was designed. Implementation was done using 64 cells in VHDL and the XST synthesis tool from the XILINX ISE configured for speed using a Xilinx FPGA from the Virtex family: the XCV2000BG560-6. As no temporal restriction of LOW level were imposed on the place and route process, further improvements in area and speed could be obtained if timing and area restrictions are imposed. The virtex device is included in a board from AlphaData attached to

a PC computer using the PCI bus. In order to compare the hardware and software performance for the specified application, the same genetic algorithm was implemented in a 2.2GHz Dual Core from INTEL, using Visual C++ from Microsoft. Results from both implementations are compared in figure 5a. Software implementation was implemented with a three decimal precision to approximate the software implementation performance to the hardware implementation performance. Increasing the number of implemented cells and running the same application a curve is obtained to indicate the work frequency versus cell number (See fig. 5b).

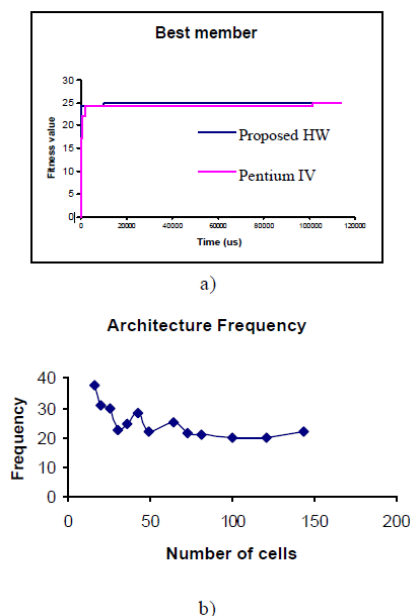


Fig. 5. a) Performance comparison between the architecture and a standard PC; b) Frequency performance in the hardware implementation

There is a slight increment in the work frequency when the cells number fits the FPGA. This characteristic can be used to decide between two numbers of implemented cells. Finally, the technical specification of the platform for the implemented function is given in table 1.

Feature Specification
Power supply 1.8V
FPGA percentage per Evolvable Cells 106 slices, less than 1%
FPGA percentage, address and control logic 201 slices, less than 2%
Estimated Power Consumption 1.52W

Table 1. Technical specifications of the EHW architecture

The proposed hardware running time for the problem was 3.5us, which was a 10.42x speed-up over a software version of the algorithm running on a 2.2 GHz personal computer. Here, the high parallelism included in the original GA is implemented, which approximates the proposed architecture to a bio-inspired design. Additionally, the hardware hierarchy and modular design allows multiple dimensions arrays, i.e., the architecture can include cells in other FPGAs boards in a parallel configuration, for example, using a Revolarc-100 from NALLATEC with Virtex II 2V6000 is possible to implement 2028 evolvable cells in a parallel structure and to accelerate the hardware implementation over 100x. The control in this new architecture is the same control used in this paper. The main problem in the massive EHW parallel implementations is the communication among cells. However, in the proposed design the communication is not a problem.

VI. Conclusions and future work

A platform for generic operations using a two-dimensional array has been presented. The genetic operators are implemented as hardware structures to achieve intrinsic evolution. Additional advantages reached with the proposed design are the total independence from the hardware application and the all on-chip concept without an embedded processor in the design. Conclusions about the design are the linear dependence between the number of cells and the FPGA resources and the compromise between FPGA resources (slices) and the work frequency of the design.

Other important characteristics are the cell arrangement suitable for bio-inspired designs (not discussed here) and the unrestricted and parameterized structure added to the design.

Furthermore, the design presents a flexible architecture, i.e., for implementation, the application only needs to have a data bus and a pair of control signals to read the memory chromosome in the evolvable module. Future works includes evolving systolic designs for computer visions and bio-inspired designs.

VII. References

- [1] A. J. Hirst, Notes on the evolution of adaptive hardware, in Proc. Of the 2nd Int. Conf. On Adaptive Computing in Engineering and Design (ACEDC96) (I. Parmee, ed., (1996).
- [2] Yasunaga, Moritoshi. Hwan Kim, Jung. Yoshihara, Ikuo. Evolvable Reasoning Hardware: Its Prototyping and Performance Evaluation. Genetic Programming and Evolvable Machines, 2, 211-230. Kluwer Academic Publishers. (2001).
- [3] M. Salami, M. Murakawa and T. Higuchi. Lossy Image Compression by Evolvable Hardware. *Proc. of IJCAI-97 Workshop on Evolvable Systems*, pp.53-59, IJCAI, 1997.
- [4] M. Salami & G. Cain. Implementation of Genetic algorithms on reprogrammable architectures, in Proc. Of the Eighth Australian Joint Conference on Artificial Intelligence (AI'95) (X. Yao, ed.), World Scientific Publ. Co., Singapore, p.581, 1995.
- [5] G. Howells, K. D. McDonald-Maier, T. Binzegger, and M. P. Young, MAVIS: A Secure Formal Computational Paradigm Based on the Mammalian Visual System, The 2007 ECSIS Symposium on Bio-inspired, Learning, and Intelligent Systems for Security, August 9-10, 2007, Playfair library Hall, Edinburgh, UK BLISS-2007
- [6] Scrubbing away transients and jiggling around the permanent: Long survival of FPGA systems through evolutionary self-repair. 10th IEEE Int. On-Line Testing Symposium 2004.

Characteristics Study of an Electronic Systems Design Strategy that Includes the Traditional Strategy and the Modified Traditional Strategy.

*Miguel Ángel San Pablo Juárez, **Alexander Zemliak, *** Eduardo Ríos Silva.

*FCE. BUAP. Puebla, Pue. sanpablojuarez@yahoo.com.mx.

**FCFM. BUAP, Puebla, Pue. azemliak@yahoo.com.

***FCE. BUAP. Puebla, Pue. erios@ece.buap.mx.

Abstract

This work presents the theoretical study considerations and the characteristics of a general design methodology in optimal time for electronic systems using numerical methods and optimal control theory. Through this the design problem of a system is formulated in terms of optimal control in minimal time. This general design methodology includes the traditional design strategy (TDS), and the modified traditional design strategy (MTDS), where the model of the system is part of optimization procedure but an objective function of optimization process is constructed such includes the traditional objective function and some penalty functions that feign the model of the system. A whole of special control functions are introduced artificially to generalize the methodology and producing several design trajectories for the same optimization process, the first and final trajectories correspond to TDS and MTDS respective. The combination of a whole these trajectories produce an infinite number of design strategies, some of these are quasi-optimal in time and only one is optimal in time. These ideas are proposed using the method of gradient optimization and acceleration effect. Qualitative and numeric results of this iterative process are generated in a personal computer in a C++ language elaborated with a visual C++ graphic user interface.

Keywords: General Design Strategy, Numerical Methods, Optimal Control, Optimal in Time Design Strategy, Traditional Design Strategy.

I. Introduction

Actually one of the most important goals in large systems design is to reduce time spent to get the optimal point of the objective function of the design process. This problem is more difficult when the system has non linear elements. Solving this problem is very important, because it is presented in systems of different knowledge areas, not only in a type physic systems, such as systems in electronics, mechanics, pneumatics, thermodynamics, optics, and optoelectronics, but in other areas as biologic systems, chemistry, economy, administrative and statistics. Specifically this problem is observed too much in the very large scale integration circuits VLSI. This problem is a part of another more general problem about optimal systems design [1].

The design process itself includes optimization of the structure of the future system, but since

this stage is related to an artificial intelligence problem it is still unsolved, in general it is performed “by hand”, and thus is absent in the CAD systems. In other words, the traditional approach to computer-aided design consists of two main parts: a model of the system set up in the form of a network described by some algebraic or integro-differential equations, and the parametric optimization procedure. There are some powerful methods that reduce the necessary time for the circuit analysis. As the matrix of the large-scale circuit is very sparse, special sparse matrix techniques are used successfully for this purpose [2-5]. Other approach to reduce the amount of computation required for the linear and nonlinear equations is based on the decomposition techniques. The partitioning of a circuit matrix into diagonal bordered block form can be done by branches tearing as in [6], or by nodes tearing as in [7] and jointly with direct solution algorithms gives the

solution of the problem [8]. Other approach for achieving decomposition at nonlinear level consists in special iteration techniques for example in [9, 10] showing the iterated timing analysis and circuit simulation. Optimization technique that is used for the circuit optimization and design exerts a very strong influence on the total necessary computer time. The numerical methods are developed for both, the unconstrained and the constrained optimization [11, 12]. The practical aspects using these methods are developed for VLSI circuit design, yield, timing and area optimization [13-15]. It is possible to suppose that the circuit analysis methods and the optimization procedures will be improved later on. Meanwhile, it is possible to reformulate the total design problem and generalize it to obtain a set of extended possibilities for the selection of one or several design strategies for optimal time or quasi-optimal-time. The time required for optimization grows rapidly as the system complexity increases. The known measures of reduction of time for advanced system analysis (in the traditional approach) turned out to be insufficiently. By convention, the generally accepted ideas of network design will be called the traditional design strategy, meaning that the method of analysis is based on *Kirchhoff's* laws. A new formulation of the network optimization problem without strict adherence to Kirchhoff's laws was suggested in [18, 19]. This process was called the generalized optimization and used the idea of ignoring *Kirchhoff's* laws for the whole network or some part of it. In this case, apart from minimization of the previously defined objective function, we also had to minimize the residual of the equation system describing the network model. In the extreme case, when the residual function included all equations of the network mathematical model, this idea was practically implemented in two CAD systems [20, 21]. The authors of these works asserted overall that design time was reduced considerably. This latter idea may be termed: the modified traditional design strategy. As distinct from the traditional approach proper, including network model analysis at every step of the optimization procedure, the modified traditional strategy of design may be defined as a strategy

which does not include at all the model analysis in the process of optimization. Another formulation of the network design problem based on the idea proposed in [22-30] can be introduced by generalization and formulation of this idea to obtain a set of different design strategies. Here we may pass to the problem of selecting, among this set, an optimal strategy in a certain way – for instance, from the running-time viewpoint. Then the optimal strategy of design may be defined as a strategy that lets us to reach the optimal point of the objective function in minimal time. The main issue in this definition is which conditions have to be fulfilled to build the algorithm provided for the optimal time. The answer to this question will make it possible to reduce substantially the computing time necessary for the design.

II. Problem formulation

Traditional design strategy means, the problem of design of an analogue network with a given topology based on the process of unconditional minimization of an objective function $C(X)$ in a space \mathbb{R}^K , where K is the number of independent variables. Simultaneously, we are seeking the solution to a system of M dependent on some components of the vector X . It is assumed that the physical model can be described by a system of nonlinear algebraic equation:

$$g_j(X) = 0, \quad j = 1, 2, \dots, M \quad (1)$$

The vector $X \in \mathbb{R}^N$ is divided in two parts: $X = (X', X'')$, where the vector $X' \in \mathbb{R}^K$ is the vector of independent variables, the vector $X'' \in \mathbb{R}^M$, is the vector of dependent variables and $N = K+M$. This partition of independent and dependent variables is a matter of convention, because some parameters of the design process may be considered independent or dependent. Due to such definition, frequency, temperature, etc., are beyond our consideration. We can easily include them in the general design procedure, but here we presume them to be constant and include them in the coefficients of system (1).

In the general case, the process of minimization of the objective function $C(X)$ in the space \mathbb{R}^K of independent variables for the two-step procedure

can be described by the following vector equation:

$$X^{s+1} = X^s + t_s H^s \quad (2)$$

where s is the iterations number, t_s is an iteration parameter, $t_s \in \mathbb{R}^1$, and H is a function establishing the direction of lowering the objective function $C(X)$. The constraints for independent variables can be bypassed easily, which is shown in the examples given in the second part of this work.

A particular feature of the design process, at least for electronic network applications, is that we are not obliged to fulfill conditions (1) at every step of the optimization procedure. It is sufficient to satisfy these conditions at the final point of the design process. In this event the vector function H depends on the objective function $C(X)$ and on some additional penalty function $\varphi(X)$, whose structure includes all the equations of system (1) and can be defined, for instance, as:

$$\varphi(X^s) = \frac{1}{\varepsilon} \sum_{j=1}^M g_j^2(X^s) \quad (3)$$

In this case we define the design process as an unconditional optimization problem:

$$X^{s+1} = X^s + t_s H^s \quad (4)$$

in the space \mathbb{R}^N without any additional system of constraints, but for a new objective function $F(X)$, which can be defined, for instance, as an additive function:

$$F(X) = C(X) + \varphi(X) \quad (5)$$

Then at the minimum point of the objective function $F(X)$ we also have the minimum of the objective function $C(X)$, and system (1) is satisfied at the final point of the optimization process. This method can be called the modified traditional method of design, it reproduces a different strategy of design and a different trajectory in the space \mathbb{R}^N . On the other hand, we can generalize the idea of using of an additional penalty function, if the penalty function is formed only from a part of system (1)

while the remaining part is regarded as a system of constraints. In this event the penalty function includes, for example, only Z first terms of:

$$\varphi(X^s) = \frac{1}{\varepsilon} \sum_{i=1}^Z g_i^2(X^s)$$

where $Z \in [0, M]$ and other $M-Z$ equations form, instead of (1), a modified system makes up one modification of the system (1):

$$g_j(X) = 0, \quad j = Z + 1, Z + 2, \dots, Z + M \quad (6)$$

Obviously, every new value of the parameter Z generates a new design strategy and a new trajectory. This notion can be easily extended to a situation when the penalty function $\varphi(X)$ includes Z arbitrary system equations (1). The overall number of different design strategies in this case equals 2^M . All these strategies exist within the same optimization procedure. The optimization procedure is realized in the space \mathbb{R}^{K+Z} . The number of dependent parameters M grows together with complexity of the system while the number of different design strategies grows by exponential law. These strategies are characterized by different operations number and overall running time. Accordingly, we may formulate the problem of searching for the design strategy optimal in time, i.e., having a minimum running time of the processor.

The most general approach to the problem of construction of the optimal design algorithm can be worked out based on the optimal control theory. We can define the design strategy with the aid of equations (4) and (6) with the variable parameter Z during the whole process of optimization. It means that we may change the number of independent variables and the number of terms in the penalty function formula at every point of optimization procedure. Also, it is worth adding into our consideration a vector function control $U = (u_1, u_2, \dots, u_m)$, where $u_j \in \Omega$ and $\Omega = \{0; 1\}$. In other words, every control function u_j can take the value 0 or 1. These functions have the meaning of control functions of the design process and generalize this process. Particularly, the meaning of the control function u_j is as follows: the equation with the ordinal

number j belongs to system (6), while the term $g_j^2(X)$ is eliminated from the right-hand part of (3) if $u_j = 0$, and vice versa the j^{th} equation is excluded from system (6) and the respective term appears in the right-hand part of (3) if $u_j = 1$. Then the system model equations and the type of the penalty function can be rewritten in the form:

$$(1 - u_j)g_j(X) = 0; \quad j = 1, 2, \dots, M \quad (7)$$

$$\varphi(X) = \frac{1}{\varepsilon} \sum_{j=0}^M u_j \cdot g_j^2(X) \quad (8)$$

All the control functions u_j are functions of a current point of the design process. In this case the directed motion vector $H=F(X,U)$ is the function of the vectors X and U . The number of various design strategies generated within a single optimization procedure is virtually unlimited. Among all these strategies, there are one or several optimal strategies, that accomplish all design goals in a minimum possible time. Hence, the problem of search for the optimal strategy is now formulated as a typical problem of minimization of some functions in optimal control theory. The functional value represents the actual running time of the processor. The main difficulty of such definition consists in unknown optimal dependencies of the control functions u_j . However, if we have an optimal vector of control functions, the optimal design strategy will be realized with the aid of this vector. The idea of formulating optimal system design from the viewpoint of time as a function minimization problem invoked from the optimal control theory, does not depend on some specific optimization algorithm realization, and can be embedded into an arbitrary optimization procedure. All this has been shown in [23-25], with approbation of three different algorithms, which are typical representatives of three major optimization methods groups: the gradient method, the *Newton's* method and the *Davidon-Fletcher-Powell* method (DFP). Now the process of network optimization is formulated as a controllable dynamic process. We have the possibility to control the design process through the control vector U variation.

In the above formulation, every possible design strategy, defined by the vector U , has its own trajectory in the space of variables. Obviously, the comparison of different trajectories by the time of moving over them, or by some other parameters, is consistent only in situation when these trajectories have identical initial and final points. On the other hand, the objective function $C(X)$ has a number of local minima, since design problems are nonlinear in principle, even if the design concerns a physically linear system. In this case, for consistent comparison of different strategies and their trajectories, it is desirable to impose additional conditions of single-value for attainment of one and the same point in the space parameters. At the same time, the problem of ambiguity is not a peculiar feature of the new formulation of the design methodology. We face this problem every time when starting the design process from different initial points. In future, both in theoretical reasoning and in practical examples, we shall assume that the problem of single-value of the final point is overcome by imposing some additional conditions on the variables. It should be specially stressed that this problem is essential only in comparison of different strategies and their trajectories, while in actual design we do not need any additional conditions except for feasibility requirements. The process of system design, formulated in terms of control theory, can be described either in discrete or continuous form. The continuous form is traditional for the optimal control theory. To represent the problem in the continuous form, we assume that numerical equation (2) corresponding to the optimization process can be replaced by a differential equation:

$$\frac{dX}{dt} = f(X, U) \quad (9)$$

Where the right-hand part $f(X, U)$ is the vector of directed motion of H and depends on the generalized objective function $F(X, U)$.

It means that the design process is formulated as the problem of integration of system (9) with additional conditions (7). The function's structure can be defined as follows:

$$f(F(X, U)) = -F'(X, U) \quad (10)$$

for the gradient method.

In this case the optimal design problem, in terms of running time, design algorithm is formulated as a typical problem of functional minimization of the control theory for differential equation system (9). The right-hand part of (9) depends on the particular method of optimization, for instance, (10), and has an objective function defined by formulas (5), (8) and constraints (7). An additional difficulty is that the righthand parts of system (9) are piecewise continuous functions rather than strictly continuous. Such a problem for system (9) with piecewise continuous control functions can be solved most effectively based on the known maximum principle [31], but straightforward application of this principle for nonlinear problems of large dimensionality is highly problematical. This problem can be solved based on provisions developed in the process of approximate solution of control theory problems [32].

The fundamental equations system design process for the three algorithms of optimization is given below. System (9), which can be rewritten in component-wise form as

$$\frac{dx_i}{dt} = f_i(X, U), \quad i = 1, 2, \dots, N \quad (11)$$

In combination with (7) defines the design process.

In the case of gradient method, the right-hand part of (11), i.e.,

$$f_i(X, U) = \frac{\delta}{\delta x_i} F(X, U), \quad i = 1, 2, \dots, K \quad (12)$$

$$f_i(X, U) = -u_{i-K} \frac{\delta}{\delta x_i} F(X, U) + \frac{(1-u_{i-K})}{t_s} \{-x_i^s + \eta_i(X)\}, \quad i = K + 1, K + 2, \dots, N \quad (13)$$

where

$$F(X, U) = C(X) + \sum_{j=1}^M \frac{1}{\varepsilon} u_j g_j^2(X) \quad (14)$$

The function $\eta_i(X)$ written in an implicit form, defines the current value of the variable x_i^{s+1} ($x_i^{s+1} = \eta_i(X)$) obtained after resolving the system (7), and the control variables u_j are the functions of “current time”.

III. Problem solution

The topology of the single node nonlinear passive circuit that is studied in this example is shown in figure 1. In this circuit we have an independent variable, $K=1$, given by resistor R , and a dependent variable, $M=1$, represented by nodal voltage V_1 [17].

The total number of variables is $N=K+M=2$. It is assumed that R_n has a nonlinear behavior governed by law:

$$R_n = a + b(V_1 - V_0)^2$$

where a is a constant and b is a nonlinear parameter nonzero. We want to find the R value for make V_1 get a desired value, i.e., we search $V_1 = wV_0$, with $w=0.1$ for example; $w \in [0,1]$.

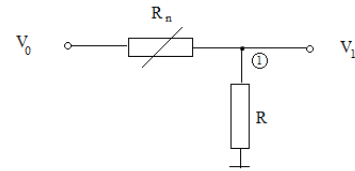


Fig. 1. A nonlinear single node circuit.

The mathematic model that describes the behavior of the circuit for analysis, according to *Kirchhoff's* law, is a voltage divider:

$$V_1 = V_0 \frac{R}{R + R_n} = V_0 \frac{R}{R + a + b(V_1 - V_0)^2}$$

In this example the phase coordinates vector only has two elements:

$$X = (x_1, x_2)$$

A physical characteristic of the elements is that they have only positive values. In order to consider this property we represent R in a squared form:

$$\begin{aligned} x_1^2 &= R \\ x_2 &= V_1 \end{aligned}$$

The nonlinear element takes the form:

$$R_n = a + b(x_2 - V_0)^2$$

and the voltage divider is:

$$x_2 = V_0 \frac{x_1^2}{x_1^2 + a + b(x_2 - V_0)^2} \quad (15)$$

The physical model is:

$$g_j(X) = g_1(x_1, x_2) = 0$$

Hence, (15) becomes in a quadratic equation that defines the circuit model for analysis:

$$g(X) = (x_1^2 + a + b(x_2 - V_0)^2)x_2 - V_0x_1^2 = 0 \quad (16)$$

This example includes the two strategies design: Traditional and modified traditional. Each design strategy has its own trajectory in the phase plane. The objective functions $F(X)$ of the trajectories always include the terms of function $C(X)$. The $C(X)$ function was determined as:

$$C(X) = (x_2 - wV_0)^2 = 0$$

The objective function $F(X)$ can be seen as a value of acceptable criterion and it should to spread zero with an error (ϵ). Define design process as this minimization. For the examples consider:

$$F(X) < \epsilon \approx 10^{-9}$$

Different algorithms were developed and their codification in the C++ program language to find numeric values generated by the design process and to graphic the general methodology trajectories and the cuasi-optimal strategy. This program calculates also the iterations number and the computer time. For this example is introduced only a control variable: u . Therefore the general methodology shows two cases: $u=0$ (traditional strategy) and $u=0$ (modified traditional strategy).

For these examples the gradient method is applied as an optimization procedure. For this

method the calculus of phase coordinates was obtained in each iteration step integrating the gradient vector generated in each case. Is used the *Euler's* integration numeric method adapted to each strategy. For this example (11) acquires the form [17]:

$$\begin{aligned} \frac{dx_1}{dt} &= -b_d \frac{dF}{dx_1} \\ \frac{dx_2}{dt} &= -b_d \frac{dF}{dx_2} u + (1 - u) \frac{(\eta_2 - x_2(t - dt))}{dt} \end{aligned}$$

where (14) is:

$$F(x_1, x_2) = C(x_1, x_2) + \frac{1}{\epsilon} u \cdot g^2(x_1, x_2)$$

and (7) takes the form:

$$(1 - u)g(x_1, x_2) = 0$$

Consequently, the objective function takes the form:

$$\text{Case 1: } u = 0, \quad F(X) = C(X) = (x_2 - wV_0)^2$$

$$\text{Case 2: } u = 1, \quad F(X) = C(X) + \frac{1}{\epsilon} u g^2(X)$$

$$= (x_2 - wV_0)^2 + [(x_1^2 + a + b(x_2 - V_0)^2)x_2 - V_0x_1^2]^2$$

with $\epsilon = 1$.

Derivatives of F respect to x_i are obtained with calculus rules, then the variables in this two different cases are obtained with *Euler's* method.

Figure 2 shows the trajectories programmed for TDS and for MTDS in a console application from Visual C++, with initial conditions $x_1=1$, $x_2=1$, and $a=1$, $b=1$, $h=0.01$ (integration step), and $V_0=1V$. Numeric results are shown in table 1.

Design n:	Control Functions U=(u _i)	Design Time (ms)	Iterations number
1	0	11.7	1570
2	1	6.9	2600

Table 1. General design strategy for the gradient method in a single node passive circuit ($K=1$, $M=1$).

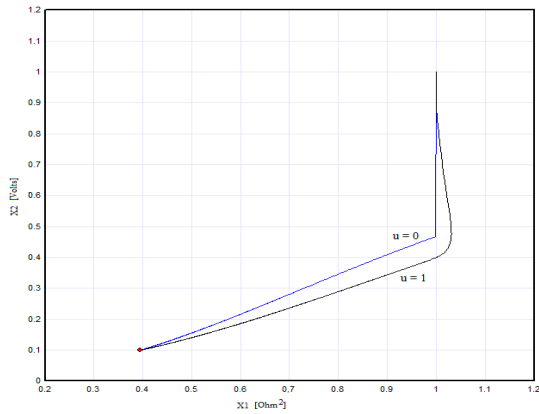


Fig. 2. Phase map for the traditional strategy and modified traditional strategy trajectories with $K=1$ and $M=1$.

The acceleration effect of the design process has been discovered through analysis of different design strategies with many different initial points. Acceleration effect reduces the computational necessary time for the design process in a considerable way and gives a perspective for accelerating this process, however this effect not ever is present, since it depends of initial design point and of the selected strategies to build the quasi-optimal trajectory, as soon as the point where is made the change of it [28].

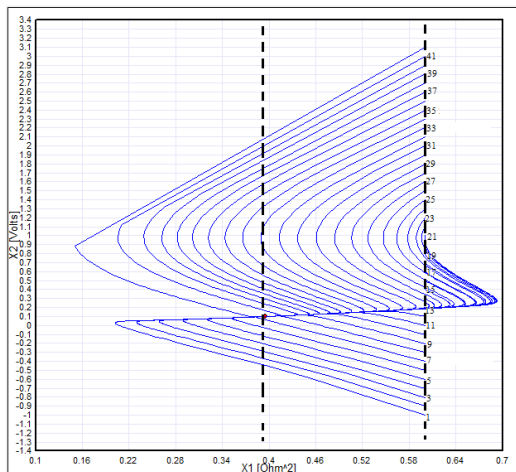


Fig. 3. Trajectories for modified traditional strategy in different initial conditions.

The acceleration effect of the design process is described and analyzed in [27, 28]. This effect appears in this kind of circuits, how the one analyzed previously. For observing the acceleration effect in the single node circuit we can vary the initial conditions of the system and

to analyze graphic results. Figure 3 shows many trajectories varying from negative initial conditions to positive initial conditions for MTDS in x_2 variable, note that the TDS ever will draw only a trajectory if we change initial conditions, therefore we only graphic trajectories for modified strategy. If we measure the compute time that need each strategy for solving the solution, we have results as shown in figure 4; the strategy number 7 (From down to up) spend the minimum time of all, it shows that exist a initial condition that spend the minimum time, and is a negative initial condition for x_2 .

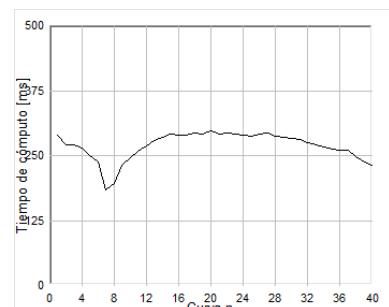


Fig. 4. Compute time in function of n curve.

A sufficient condition so that exist acceleration effect will be the selection of a initial condition of negative x_2 variable and sufficient big. This condition is considered a sufficient condition to exist the effect.

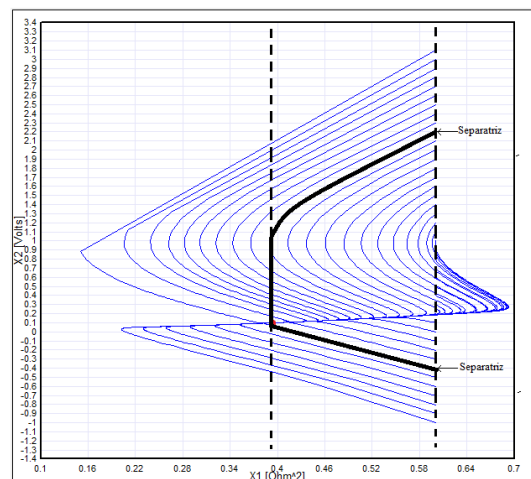


Fig. 5. Separatrix in the phase map of the trajectories in the ETMD.

Figure 5 shows two dotted lines that delimit the solutions vector in trajectories between the solution point and constant initial conditions of x_1 . Suppose that the X vector began to solve with

the down trajectory of MTDS in figure 5, and then, just in the left dotted line we change to TDS for solving X , or any different negative initial condition for x_2 variable of MTDS. Then the solution vector has been solved without spend more compute time.

Now suppose that we began in the initial conditions of the trajectories that spend the minimum time in figure 5, the solution doesn't need to make a change of strategy, but until this trajectory, the acceleration effect can be applied, if we try with other trajectory up and we change from one to other, the system will spend more time, and so on if we try with other trajectory until on or upper other trajectory, trajectory that touch tangentially (in bold) to left dotted line, that in union with the minimum consumption time trajectory we call *separatrix*.

If the initial point is inside of *separatrix*, we can't make the acceleration effect, it is a sufficient and necessary condition. If the initial point is outside of *separatrix*, is possible to make the acceleration effect.

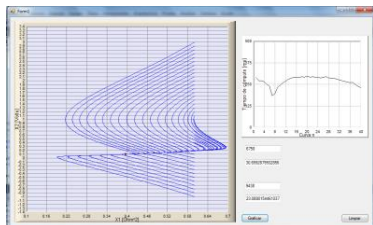


Fig. 6. Graphic user interface generated by C++ code programmed in Windows 7.

In figure 6 is presented the graphic user interface in a IBM PC with Windows 7 operating system.

Finally we can add other optimization into design process, since the problem solution of optimization consist in applying an acceleration effect switching between control variables u , by example change $u=1$ for $u=0$ or vice versa.

Now we can consider the three nodes non linear circuit shown in figure 7. In this example the X vector is:

$$X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)$$

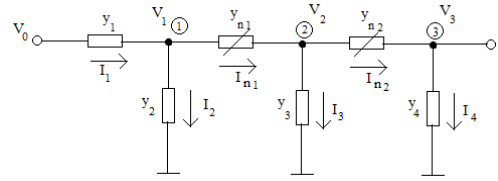


Fig. 7. Diagram of a nonlinear 3 nodes passive electronic circuit (with $K=4, M=3$).

We have 4 independent variables, admittances, and 3 dependent variables, nodal voltages. For obtaining the model we make:

$$x_1^2 = y_1; x_2^2 = y_2; x_3^2 = y_3; x_4^2 = y_4; x_5 = V_1; x_6 = V_2; x_7 = V_3$$

And non linear elements take the form:

$$y_{ni} = a + b(x_{k+i} - x_{k+2i})^2$$

Analyzing a three nodes passive electronic circuit of figure 7, we can see that the trajectories generated are separated in two subsets of trajectories, figure 8; one of them contain trajectories similar to TDS, and the another contain trajectories similar to MTDS, printed trajectories are of output voltage x_7 and value of admittance in the output node x_4 . Then we can see clearly that acceleration effect can be applied first taking one trajectory of MTDS subset and then switching the control variables to take a trajectory subset TDS. We can see acceleration effect give a gain in time when we change trajectories, but we need first take one of MTDS subset and then other one of TDS subset, but we can take other changes too.

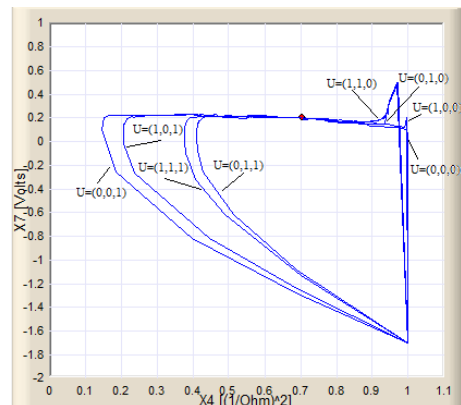


Fig. 8. Trajectories of the x_4 versus x_7 for a nonlinear 3 nodes passive electronic circuit (with $K=4, M=3$)

For applying acceleration effect, first we need take a MTDS trajectory subset, and after to make a commutation to a TDS trajectory subset. How is shown in figure 9, commutation point could be taken in dotted line, making a trajectory change to another TDS subset strategy. We need a commutation point that spends minimum time, but we can make more than one change.

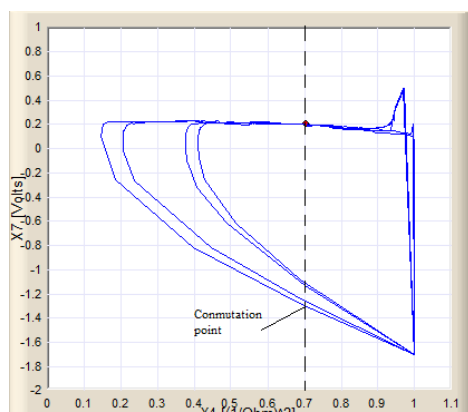


Fig. 9. A possible commutation point in three nodes example.

IV. Conclusions

The methods and the programs for the design of passive electronic systems was developed using control theory and was made the analysis of the characteristics in each case of the traditional design strategy and the modified strategy, the code was programmed in a C++ language compiler of a visual program suite in computers with compatibility with IBM PC, is possible to see qualitative behavior of the solution system evolution in time for each one of the strategies that it generates. Phase map with many different initial conditions in the modified traditional strategy was obtained, we used to graph a *separatrix* for delimiting the design trajectories conjunct that are utile for applying an *acceleration effect* of the system solution and those to have a criterion to can using this effect in the design quasi-optimal algorithm. Different projections of the passive two nodes electronic system was programmed in a conjunct of code functions for analyzing projections of the solutions in the phase space, where the optimization of the temporal numeric integration step was added to system across a code function that finds the optimal temporal step. Was made the study of the mathematic methods and

algorithms that allow to make an correct analysis of different electronics circuits as linear as non linear, by example the *Euler's* numeric integration methods, the *Newton-Raphson* method for solving algebraic equations of nth order, the gradient optimization method, the temporal step optimization using interpolating parabola and the compiler graphic libraries to print trajectories in the screen, and the time measuring in each strategy too.

V. References

- [1] P. Y. Papalambros, D. J. Wilde, Principles of Optimal Design, Cambridge University Press, 2000.
- [2] J.R. Bunch, D.J. Rose, Eds., Sparse Matrix Computations, Acad. Press, N.Y., 1976.
- [3] I.S.Duff, J.K. Reid, Some Design Features of a Sparse Matrix Code, ACM Trans. on Mathematical Software, vol. 5, no. 1, 1979, 18-35.
- [4] O. Osterby, Z. Zlatev, Direct Methods for Sparse Matrices, Springer-Verlag, N.Y., 1983.
- [5] A. George, On Block Elimination for Sparse Linear Systems, SIAM J. Numer. Anal. vol. 11, no.3, 1984, 585-603.
- [6] F.F. Wu, Solution of Large-Scale Networks by Tearing, IEEE Trans. Circuits Syst., vol. CAS-23, no. 12, 1976, 706-713.
- [7] A. Sangiovanni-Vincentelli, L.K. Chen, L.O. Chua, An Efficient Cluster Algorithm for Tearing Large-Scale Networks, IEEE Trans. Circuits Syst., vol. CAS-24, no. 12, 1977, 709-717.
- [8] N. Rabat, A. E. Ruehli, G. W. Mahoney, J. J. Coleman, A Survey of Macromodeling, IEEE Int. Symp. Circuits Systems, April 1985, 139-143.
- [9] A. Ruehli, A. Sangiovanni-Vincentelli, G. Rabbat, Time Analysis of Large-Scale Circuits Containing One-Way Macromodels, IEEE Trans. Circuits Syst., vol. CAS-29, no. 3, 1982, 185-191.

- [10] A. E. Ruehli, G. Ditlow, Circuit Analysis, Logic Simulation and Design Verification for VLSI, Proc. IEEE, vol. 71, no. 1, 1983.
- [11] R. Fletcher, Practical Methods of Optimization, John Wiley & Sons, N.Y., vol. 1, 1980, vol. 2, 1981.
- [12] P.E. Gill, W. Murray, M.H. Wright, Practical optimization, Academic Press, London, 1981.
- [13] R.K. Brayton, G.D. Hachtel, A.L. Sangiovanni-Vincentelli, A survey of optimization techniques for integrated-circuit design. IEEE Proc., vol. 69, 1981.
- [14] A.E. Ruehli, Ed., Circuit Analysis, Simulation and Design, vol. 3, part 2, Elsevier Science Publishers, Amsterdam, 1987.
- [15] R.E. Massara, Optimization Methods in Electronic Circuit Design, Longman Scientific & Technical, Harlow, 1991.5
- [16] A. Zemliak, Frontiers in Robotics, Automation and Control. In-Tech. Croatia 2008.
- [17] E. Ríos, Aspectos teóricos de diseño en tiempo óptimo de sistemas electrónicos. Tesis de Doctorado. Optoelectrónica, FCFM, BUAP. Puebla 2003.
- [18] I.S. Kashirskiy, General Optimization Methods, Izvestiya VUZ USSR - Radioelectronica, vol. 19, no. 6, 1976.
- [19] I.S. Kashirskiy, I.K. Trokhimenko, General Optimization for Electronic Circuits, Tekhnika, Kiev, 1979.
- [20] V. Rizzoli, A. Costanzo, C. Cecchetti, Numerical optimization of broadband nonlinear microwave circuits, IEEE MTT-S Int. Symp., vol. 1, 1990, 335-338.
- [21] E.S. Ochotta, R.A. Rutenbar, L.R. Carley, Synthesis of High-Performance Analog Circuits in ASTRX/OBLX, IEEE Trans. CAD, vol.15, no. 3, 1996, 273-294.
- [22] A. Zemliak, General Methodology for System Design, Chapter of the book “Modern Applied Mathematics Techniques in Circuits, Systems and Control”, Ed. N. Mastorakis, WSES Press, 1999, 150-155.
- [23] A. Zemliak, One Approach to Analog System Design Problem Formulation, IEEE Int. Symp. Quality Electronic Design – ISQED2001, San Jose, CA, USA, March 2001, 273-278.
- [24] A. Zemliak, System Design Problem Formulation by Control Theory, IEEE Int. Symp. Circuits and Syst. – ISCAS2001, Sydney, Australia, 2001, vol. 5, 5-8.
- [25] A.M. Zemliak, Analog System Design Problem Formulation by Optimum Control Theory, IEICE Trans. Fundam., vol. E84-A, no. 8, 2001, 2029-2041.
- [26] A. Zemliak, Novel Approach to the Time-Optimal System Design Methodology, WSEAS Trans. Syst., vol. 1, no. 2, 2002.
- [27] A. Zemliak, On Start Point Selection for the Time-Optimal System Design Algorithm, IEEE Int. Symp. Circuits and Syst. – ISCAS2002, Scottsdale, USA, May, 2002, vol. IV, 465-468.
- [28] A.M. Zemliak, Acceleration Effect of System Design Process, IEICE Trans. Fundam., vol. E85-A, no. 7, 2002.
- [29] A. Zemliak, Main Properties Study of the Time-Optimal System Design Algorithm, WSEAS Trans. Circuits Syst., vol. 3, no. 4, 2004, 759-764.
- [30] A. Zemliak, Analog Circuits Design by the Methods of Control Theory, Part 1, Theory, Izv. VUZ Radioelectronica, vol. 47, no. 5, 2004, 18-28.
- [31] L. S Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, “The mathematical Theory of Optimal Process”. Interscience Publishers, Inc., New York. 1962.
- [32] R. Pytlak, “Numerical Methods for Optimal Control Problems with State Constraints”, Springer-Verlag, Berlin. 1999.

“Formación del educador universitario desde la visión del educando” Diagnóstico Aplicado en el Área de Electrónica y Computación

***Graciela Lara López, **Martha Patricia Macías Pérez, ***Norma Ramírez Hernández,
****José Luis Ramos Quirarte**

*Universidad de Guadalajara (CUCEI) Depto. de Cs. Computacionales, Bv. Marcelino García Barragán No. 1451. Guadalajara, Jalisco, México 44430, (33)13785900 ex 7735, lara_graciela@hotmail.com.

** Universidad de Guadalajara (CUCEI) Depto. de Química, Bv. Marcelino García Barragán No. 1451. Guadalajara, Jalisco, México 44430, (33)13785900 ex 7735, mpatricia11@hotmail.com.

*** Universidad de Guadalajara (CUCEI) Depto. de Cs. Computacionales, Bv. Marcelino García Barragán No. 1451. Guadalajara, Jalisco, México 44430, (33)13785900 ex 7735, mirandapartida@terra.com.mx.

**** Universidad de Guadalajara (CUCEI) Depto. de Electrónica, Bv. Marcelino García Barragán No. 1451. Guadalajara, Jalisco, México 44430, (33)13785900 ex 7735, jlquirarte@yahoo.com.mx.

Resumen

La formación del educador se percibe como un proceso de preparación inicial y permanente que vincula las prácticas de enseñanza y de aprendizaje orientadas a los docentes-discentes, que se manifiesta directamente dentro del ambiente áulico, en la práctica con los alumnos y en el desarrollo docente en general; con el propósito de coadyuvar a la construcción del conocimiento de los estudiantes [1]. Por lo expuesto, se presenta la siguiente investigación, que muestra los datos recabados del diagnóstico realizado con los alumnos de las carreras de Licenciatura en Informática, la Licenciatura en Ingeniería en Computación y la Licenciatura en Ingeniería en Comunicaciones y Electrónica que oferta el Centro Universitario de Ciencias Exactas e Ingenierías, relacionado con la formación y la práctica docente de sus profesores. La información obtenida ha sido analizada y tiene como intención utilizarla para la mejora del proceso educativo.

Palabras Clave: Aprendizaje, Diagnóstico, Formación y práctica docente, Proceso educativo.

I. Introducción

En el informe de la UNESCO de 1996 titulado “La educación encierra un tesoro, entiende que la educación ha de articularse entorno a cuatro ejes denominados los cuatro pilares:” aprender a conocer, aprender a hacer, aprender a vivir juntos y aprender a ser”. [2]

Podemos entender por educación un proceso sistemático e intencional para conseguir el desarrollo integral de los educandos junto a su máxima integración social en el seno de una sociedad, a través de una sociedad personalizada, para que puedan llevar a cabo un proyecto de vida en el marco de un aprendizaje a lo largo de su vida.

El papel de los maestros implica la tarea de formar, además de enseñar a los dicentes para su desarrollo cognitivo y así aprenda a ser autogestivo, aprovechando no sólo la información del maestro, que indague por su cuenta en conocimientos extras, disponibles con relativamente poco esfuerzo, en donde el problema no es obtener la información, sino que tiene que aprender a discriminarla y a extraerla del cúmulo de datos disponibles en la Internet.

Es en este contexto en donde el docente debe ser reflexivo con una mentalidad analítica y crítica hacia el trabajo, que le permita hacer modificaciones en período corto de clase, debido a los dicentes cambiantes que tendrá a su cargo.

Su actualización profesional deberá ser continua, renovada y con la disposición real de utilizar los métodos y técnicas con los alumnos con los cuales desarrolla su actividad como docente. Para que lo anterior suceda es necesario analizar el binomio educando-educador.

El diagnóstico de la formación docente nos permite examinar la situación actual de los profesores en su práctica docente tanto en Universidades como en otras Instituciones Educativas, de modo que los educadores pueden ser evaluados por ellos mismos, por sus pares o sus propios alumnos con el fin de marcar nuevos cambios en los procesos y actividades formativas de los docentes, con la intención de mejorar sus métodos de enseñanza [3]. Es por ello que los docentes en su labor académica coadyuvan, para que los estudiantes logren su misión que tienen de aprender nuevas formas de estudio, desarrollar sus habilidades y competencias científicas y tecnológicas.

Resulta de suma importancia que los docentes muestren sentido de responsabilidad y compromiso con la sociedad y la institución que los forman como personas, con un nivel de estudios superior y una perspectiva clara del mundo de hoy y del futuro.

Dado lo anterior, es importante reconocer que vivimos en una época en la que los estudiantes están marcando un cambio significativo en el ambiente áulico, ya que se encuentran inmersos en un mundo rodeado de nuevas tecnologías que van innovando día con día. Es común ver como un estudiante aprende de forma autodidacta un curso o programa de software, con la ayuda de un Sistema Multimedia Interactivo, quizás montado en línea o desde su PC, sin el apoyo de un profesor, logrando así, que para él, el proceso de aprendizaje sea más atractivo y deseado. A los profesores cada vez les cuesta más trabajo atrapar la atención de los alumnos, como a los alumnos también les cuesta entender y aceptar las formas tradicionales de enseñanza que muchos profesores aplican. Si a esto le sumamos que un profesor universitario sólo cuenta con la

formación de su área profesional, que rara vez ha podido tomar un curso de pedagogía y más aún no se actualiza en su área de formación, resulta un problema serio, para las instituciones universitarias ya que esto desmotiva gravemente al alumnado y es una de las posibles causas de un alto índice de deserción. Es por eso, que resultó de interés el desarrollo de esta investigación e identificar los factores que influyen en el bajo rendimiento de ellos, ocasionado por el profesor bajo la perspectiva de los alumnos y con ello generar una propuesta de mejora, para así lograr una combinación de formación docente con conocimiento teórico-práctico tanto en su área de formación como en la pedagógica y así poner en práctica todos los conocimientos adquiridos a través de cursos, entrenamientos y todo lo que forme parte de la formación docente que facilite en los estudiantes el aprendizaje y por ende un mejor desempeño en la universidad.

II. Desarrollo

El proceso para el desarrollo del diagnóstico fue el siguiente:

1. Se elaboró una encuesta estructurada con 11 preguntas de formato de escalas nominales, escala de intervalos y abiertas.
2. De un total de 129 alumnos de cinco grupos distintos de las carreras de Licenciatura en Informática, la Licenciatura en Ingeniería en Computación y la Licenciatura en Ingeniería en Comunicaciones y Electrónica, se realizó un sistema de muestreo aleatorio simple [4], dando como tamaño de la muestra 84 alumnos.

La fórmula para un sistema de muestreo aleatorio simple es:

$$n = \frac{Z_{\alpha/2}^2 S^2}{\varepsilon^2} \quad (1)$$

Donde:

n : = tamaño necesario de la muestra.

$Z_{\alpha/2}$ = margen de confiabilidad o número de unidades de desviación estándar en la distribución normal que producirá el nivel

deseado de confianza (para una confianza de 95% o un $\alpha = 0.05$, $Z = 1.96$; para una confianza de 99% o un $\alpha = 0.01$, $Z = 2.58$).

S = desviación estándar de la población.

ε = error o diferencia máxima entre la media muestral y la media de la población que se está dispuesto a aceptar con el nivel de confianza que se ha definido.

Se decidió tomar un margen de confianza de 95% ($1 - \alpha$) (que corresponde a $Z = 1.96$), con desviación estándar $S = 0.4$, un error de estimación $E = 5\%$, se supone que no conoce el tamaño de la población ($N = \text{infinito}$), para el uso de esta función.

Sin embargo, para el desarrollo de este caso de estudio si se conocía el tamaño de la población (N), la fórmula que se utilizó para estimar el tamaño de la muestra fue:

$$n = \frac{S^2}{\frac{\varepsilon^2}{Z^2} + \frac{S^2}{N}} \quad (2)$$

$$n = \frac{(0.4)^2}{\frac{(0.05)^2}{(1.96)^2} + \frac{(0.4)^2}{129}} = 84 \quad (3)$$

3. Se procedió a aplicar la encuesta en varios días y a levantar la información.

4. Posteriormente las respuestas de las preguntas con escalas nominales, así como las de intervalo se capturaron en Excel. Se contabilizaron y se calcularon porcentajes, para posteriormente representar en gráficas de barras, permitiendo ver así la visión de los alumnos con respecto a la formación docente de sus profesores.

5. Por último se tomaron las respuestas de las preguntas abiertas para ser analizadas y conocer métodos de mejora de los profesores con respecto a su formación y práctica docente.

III. Resultados

En este apartado se presentan los resultados de las encuestas realizadas, las cuales revelan que de los alumnos encuestados el 88.10% (74) de los estudiantes de las Licenciatura en Informática, la Licenciatura en Ingeniería en Computación y la Licenciatura en Ingeniería en Comunicaciones y Electrónica, creen que sus profesores si dominan la materia que imparten y el 11.90% (10) creen que no. (ver Fig. 1).

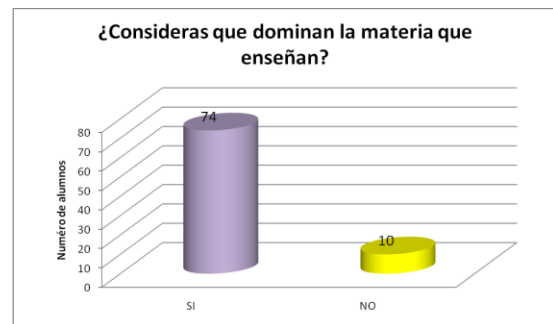


Fig. 1. Porcentaje de estudiantes que consideran que sus profesores dominan las materias que enseñan.

El 59.52 % (50) de los estudiantes consideran que sus profesores se encuentran al día con los conocimientos de la materia que imparte y el 39.29% (33) cree que no. (ver Fig.2).

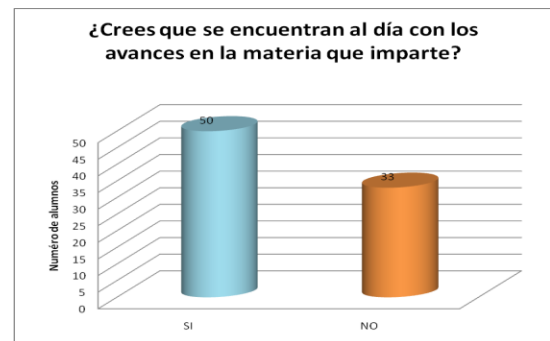


Fig. 2. Porcentaje de estudiantes que creen que sus docentes se encuentran al día con los avances de la materia.

El 69.05% (58) de los estudiantes evalúan que la habilidad del profesor para organizar el contenido y presentarlo en forma clara, lógica e imaginativa es regular, mientras que un 21.43% (18) creen que en mucho presentan esa habilidad, 8.33% (7)

opinan que poco y el 1.19% (1) muy poco. (ver Fig. 3).



Fig. 3. Porcentaje de estudiantes que consideran que sus profesores organizan sus contenidos

El 45.24% (38) de los alumnos opinan que algunas veces los docentes relacionan su materia con otras esferas de conocimiento, en tanto el 32% (27) señala que casi siempre y el 21.43% (18) cree que siempre. (ver Fig. 4).

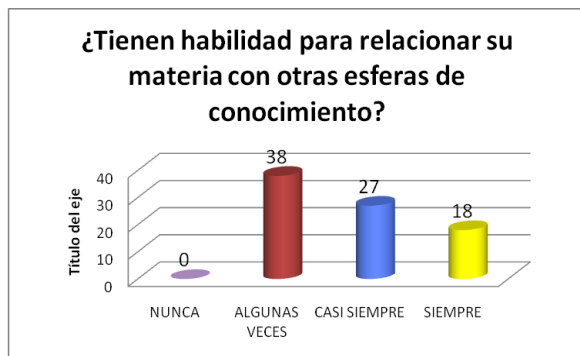


Fig. 4. Porcentaje de estudiantes que opinan que sus profesores relacionan su materia con otras esferas de conocimiento.

El 57.14% (48) de los estudiantes percibe que el docente algunas veces emplean métodos y estrategias adecuadas, incluyendo el uso de la tecnología educativa para una enseñanza efectiva, 34.52% (29) consideran que casi siempre, mientras que el 8.33% (7) cree que siempre. (ver Fig. 5).

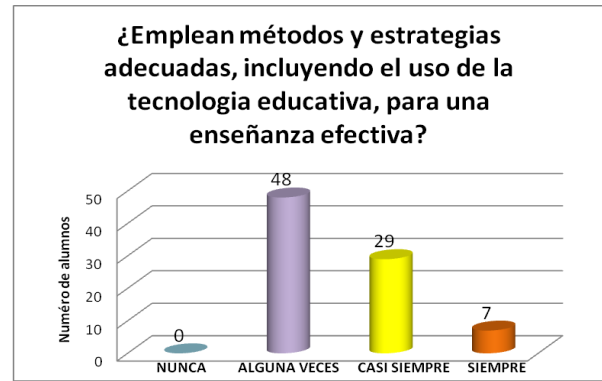


Fig. 5. Porcentaje de alumnos que opinan que sus profesores emplean métodos y estrategias adecuadas.

El 60.71% (51) señala que los profesores se comunican casi siempre bien, mientras que el 21.43% (18) opina que algunas veces y el 17.86% (15) considera que siempre (ver Fig. 6).

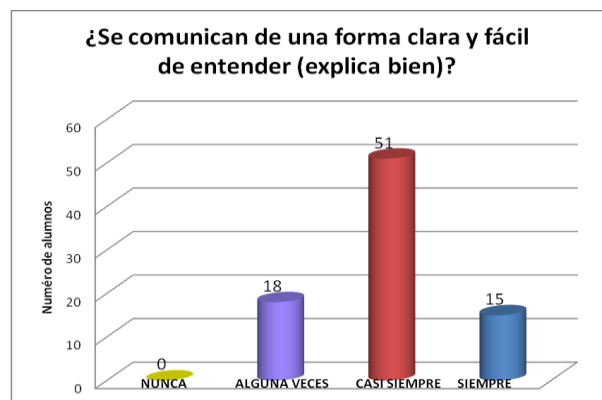


Fig. 6. Porcentaje de estudiantes que opinan que sus profesores se comunican de una forma clara.

IV. Discusión

Actualmente existen algunas instituciones educativas que empiezan a tomar conciencia sobre la necesidad de crear nuevas políticas y estrategias orientadas a reducir la brecha entre el currículum, la investigación y la práctica pedagógica [5], lo cual ha llegado a incluir una revisión crítica de la educación tal y como es impartida actualmente por parte de los maestros de acuerdo a Carr y Kemmis. [6]

Lo anterior y de acuerdo a la información obtenida en las encuestas, nos permite observar que dentro de las instituciones educativas existen carencias en diversos aspectos fundamentales para su misión, y uno de ellos es la desvinculación entre la investigación, la práctica y el desarrollo docente con la práctica pedagógica.

En este momento, dentro de las universidades se está impulsando fuertemente el trabajo de investigación y es muy importante dado que se promueve el desarrollo de la sociedad, que aporta ciencia y tecnología nueva. Sin embargo, no debemos de olvidar el propósito principal de toda institución educativa que es ayudar a formar futuras personas profesionales en un área de conocimiento específica, pero sobre todo ayudarlos a construir sus competencias y a descubrir sus habilidades propias.

El papel de los formadores no es tanto "enseñar" (explicar-examinar) unos conocimientos que tendrán una vigencia limitada y estarán siempre accesibles, como ayudar a los estudiantes a "aprender a aprender" de manera autónoma en esta cultura del cambio y promover su desarrollo cognitivo y personal mediante actividades críticas y aplicativas que, aprovechando la inmensa información disponible y las potentes herramientas TIC, tengan en cuenta sus características (formación centrada en el alumno) y les exijan un procesamiento activo e interdisciplinario de la información para que construyan su propio conocimiento y no se limiten a realizar una simple recepción pasiva-memorización de la información.

La necesidad de profesionalizar las habilidades didácticas en cualquier ámbito educativo es innegable, pero, en todo caso, la habilidad didáctica se consolida al eliminar las actividades inútiles y lograr la armonía entre la teoría y la práctica del conocimiento de las tareas y competencias, así como la capacidad para lograr la armonía entre las disciplinas de la cual es especialista, (Ingeniería, Electrónica, Química, etc.) y la didáctica: aprender la interdisciplinariedad entre el *qué* se enseña y se

aprende y el *cómo* se enseña y se aprende; armonía que también está presente en saber dirigirse a las tres áreas de desarrollo: adquisición de conocimientos por medio de las operaciones mentales o estrategias cognitivas para el logro de aprendizajes significativos; generación de actitudes positivas, y desarrollo de habilidades manuales e intelectuales[8].

En el caso de México, se ha dado gran importancia a la formación pedagógica de los docentes universitarios desde los años setenta hasta la fecha. Organismos como la Secretaría de Educación Pública (SEP), la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES) y el Consejo Nacional de Formación de Profesores de acuerdo con Ezcurra se han dedicado a impulsar programas para capacitar profesores [3]. Debido a esto, se ha incrementado el número de maestros a nivel licenciatura, sin embargo, el reto no es sólo tener un mayor número de profesores, sino que tengan una formación tanto de la disciplina que imparten, como de la didáctica-pedagógica [7].

Es por esto que, los profesores deben adquirir una formación docente inicial y procurar mantener una capacitación constante, que desde luego se proyecta en su carrera profesional, pero sobre todo se reconoce y valora por sus estudiantes.

De acuerdo a las opiniones compartidas de los estudiantes, la formación del educador debe permitir ver atractiva, dinámica y divertida su clase entre sus estudiantes, debe atrapar la atención de ellos, pero sobre todo el docente tiene que demostrar que posee los conocimientos suficientes para enseñar y transmitir información fresca y de utilidad, llevándola a la práctica, con la resolución de caso de estudios o el desarrollo de proyectos reales.

Es importante mencionar lo que Chehaybar y Kur comentan sobre la formación: "Es un proceso permanente, dinámico, integrado, multidimensional, en el que convergen, entre otros elementos, la disciplina y sus aspectos teóricos, metodológicos, epistemológicos, didácticos, psicológicos, sociales, filosóficos e

históricos, para lograr la profesionalización de la docencia.”[7]

La formación del educador es todo un proceso formal bien estructurado, dirigido al desarrollo del docente en sus diversos contextos de su entorno laboral, nos referimos en sus actividades de docencia, investigación, gestión y tutorías; es decir a tener un perfil deseable de profesor en todas sus áreas donde se desenvuelve.

V. Conclusiones

Al obtener esta información y analizarla encontramos que a los alumnos les gustaría que los profesores despertarán en ellos más su interés, que usaran métodos más interactivos durante sus cursos, que se perfeccionaran en los temas, no sólo teóricamente sino también en la práctica para que pudieran aclarar en su momento todas las dudas que surgieran durante el desarrollo del tema.

Coinciden en gran medida los educandos, en que la formación docente debe ser constante, así como la capacitación dentro de la industria, para relacionar lo enseñado con el campo laboral. La mayoría de los estudiantes concuerdan en que los docentes tienen que ampliar sus estrategias de enseñanza, para que ellos puedan comprender clara y sencillamente los contenidos temáticos expuestos en los cursos.

Resultó interesante encontrar dentro de esta investigación como a los estudiantes les llama la atención ser comprendidos por los profesores. A los docentes en el nivel universitario se les olvida algunas veces que trabajan con seres humanos que necesitan respeto, tolerancia, atención y reconocimiento en algún momento.

Al analizar todo lo anterior fue sugestivo formularnos la siguiente pregunta: ¿Cómo lograr que un profesor alcance ese perfil deseable para los alumnos? Resulta difícil responder esta pregunta, sin embargo, se pueden plantear estrategias y planes de acción a seguir, para llegar al objetivo deseado. Es importante mencionar que

para lograr la formación docente, la conexión entre el compromiso personal del educador y la responsabilidad de la institución educativa son elementos claves para el éxito de esta labor. La institución educativa debe brindar todo el apoyo al docente para buscar u otorgar cursos de capacitación internos o externos bajo un Plan de Formación Constante (PFC), plenamente establecido en tiempos y espacios definidos. Además de reconocer su interés de participación con incentivos cualitativos o cuantitativos.

Dentro del PFC sería deseable, que los profesores en períodos de verano o en invierno realizaran estancias en la industria. Esto elevaría ampliamente su perspectiva de cómo se encuentra fuera de la universidad el campo laboral, en su área práctica. Se recomienda que dentro del Plan de Trabajo de cada docente se establezca la participación de dos cursos de formación y actualización en su área de conocimiento, así como un curso de proceso en la práctica pedagógica, para no desligar ambas líneas.

Debemos mencionar lo importante que es asumir un compromiso profesional por parte del docente, para asistir y participar pero sobre todo poner en práctica los conocimientos y las técnicas adquiridas.

Este trabajo representa una contribución parcial, en la comprensión del fenómeno de la formación docente, requerida en el proceso de enseñanza-aprendizaje, ya que siendo este campo tan complejo es indispensable el profundizar con nuevos estudios y propuestas de mejora, por ejemplo establecer la correlación existente entre la formación académica afín a su área del educador y sus habilidades pedagógicas.

VI. Referencias

- [1] I. C. Gorodokin. La formación docente y su relación con la epistemología. <http://www.rioei.org/deloslectores/1164Gorodokin.pdf>. [10/05/2011].

- [2] J. Delors, “La educación encierra un tesoro”, Informe a la UNESCO de la Comisión Internacional sobre educación para el siglo XXI. 1997. Ediciones UNESCO, Editorial Santillana. Madrid, p 34.
- [3] V. Álvarez Rojo, y otros, Diagnóstico pedagógico. Alfar, 1984, España, Tema 1.
- [4] C.A. Bernal T., Metodología de la Investigación para Administración y Economía. Primera. 2000. Prentice Hall. Colombia, pp 161-169.
- [5] J. Quintero–Corzo; R. Ancízar; I.F. Munévar–Molina. "Semilleros de investigación: una estrategia para la formación de investigadores", 2008. Revista Educación y Educadores, Vol. 11, No 1. Colombia.
- [6] M.V. Arenas Castellanos, T. Fernández de Juan. Formación pedagógica docente y desempeño académico de alumnos en la facultad de Ciencias Administrativas de la UABC. Rev. educ. sup [online]. 2009, vol.38, n.150 [citado 2011-03-04], pp. 7-18. Disponible en: <http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-27602009000200001&lng=es&nrm=iso>. ISSN 0185-2760.
- [7] C. Zarza Charur, “La Definición de Objetivos de Aprendizaje. Una Habilidad Básica para la Docencia”. Perfiles Educativos UNAM. Enero-Marzo. Núm. 63. 1994, p 3.
- [8] E.M. Villalobos Pérez-Cortés, Didáctica Integrativa y el Proceso de Aprendizaje. Primera Edición. 2002. Trillas. México, p 74.

VII. Autores

M.S.I. Graciela Lara López obtuvo su título de Licenciada en Informática por la Universidad de Guadalajara y el grado de Maestro en Sistemas de Información.

M en C. Martha Patricia Macías Pérez es Químico Farmacobiólogo y obtuvo la Maestría en Ciencias por la Universidad de Guadalajara.

M en C. Norma Ramírez Hernández es Ingeniera en Comunicaciones y Electrónica y obtuvo la Maestría en Ciencias en Ingeniería Electrónica por la Universidad de Guadalajara.

Dr. José Luis Ramos Quirarte es Ingeniero en Comunicaciones y Electrónica por la Universidad de Guadalajara, obtuvo la Maestría en Ciencias en Ingeniería Eléctrica (Telecomunicaciones) en el CINVESTAV (México) y su grado de Doctor en Tecnología de Alta Frecuencia por la Universidad Chalmers en Gotemburgo, Suecia.

Cómo Capturar Alumnos Usando Videojuegos.

* Jorge Elpidio Cruz Gálvez, **Luis Antonio Álvarez Oval, ***Arón de la Cruz Vázquez, ****
Christian Mauricio Castillo Estrada

* Facultad de Contaduría Pública del Campus IV de la Universidad Autónoma de Chiapas. 7a Oriente número 14, Colonia Centro, Tapachula, Chiapas, México. C.P. 30700. +52 01 962-112-4243. jorge.cruz@unach.mx

** Facultad de Contaduría Pública del Campus IV de la Universidad Autónoma de Chiapas. Carretera a Puerto Madero Km 1.5, Tapachula, Chiapas, México. +52 01 962-626-8372 y 625-1723, loval@unach.mx

*** Facultad de Contaduría Pública del Campus IV de la Universidad Autónoma de Chiapas. Calle Guillermo el Conquistador Manzana 44 lote 8 Colonia: Colinas del rey. Tapachula, Chiapas, México. +52 01 962-112-7792, delacruzvazquez.aron61@gmail.com

**** Facultad de Contaduría Pública del Campus IV de la Universidad Autónoma de Chiapas. Fraccionamiento Solidaridad Las Vegas Manzana 1 Casa 7, Tapachula, Chiapas, México. +52 01 962-107-1753, cmce@unach.mx

Resumen

De acuerdo con la PROFECO, en México hay más de 13 millones de consolas y las ventas anuales de videojuegos rondan los 900 millones de dólares [9], a demás nuestro país es considerado el mayor consumidor de este tipo de entretenimiento en América Latina, lo cual nos da una idea de las jugosas ganancias en éste sector. Para un desarrollador de software esto representa una importante oportunidad que se puede aprovechar sólo si se está preparado para asumir semejante reto. Los videojuegos representan una nueva e interesante área de investigación que combina la experiencia de un amplio rango de habilidades computacionales, desde programación e ingeniería de software hasta algoritmos, resolución de problemas y redes, así como habilidades de diseño. Por lo anterior, enseñar a estudiantes de la computación a crear juegos efectivos con un propósito serio en sólo un semestre es todo un reto [1].

Palabras Clave: educación, e-learning, videojuegos.

I. Introducción

Durante varias décadas millones de niños y jóvenes alrededor del mundo han hecho de los videojuegos parte de su vida cotidiana, sin embargo, la sociedad muestra cierto rechazo a este tipo de programas porque algunos incluyen contenido que puede desarrollar actitudes negativas. Afortunadamente después de numerosos estudios se ha demostrado que si el interés de niños y jóvenes por los videojuegos se encausa de manera correcta, provee una excelente herramienta de aprendizaje interactivo.

Para muchos desarrolladores de software y en particular para los recién iniciados, construir un videojuego es más interesante que desarrollar otro tipo de aplicación como por ejemplo un

sistema administrativo, pues los videojuegos generan una gratificación instantánea, ya que desarrollar una aplicación gráfica es mucho más satisfactorio que calcular el salario promedio de una nómina[17].

II. Desarrollo

Connotaciones negativas hacia los videojuegos.

Las opiniones emitidas por los jugadores acerca de los videojuegos son extremadamente variadas, desde curiosidad hasta indiferencia, e inclusive fascinación; aunque la opinión general es de aceptación. Sin embargo, las actitudes de los jugadores en base a sus experiencias como usuarios finales y la forma en que los videojuegos forman parte de su vida diaria,

contrasta con la de los políticos, líderes educacionales y muchos otros profesionales y críticos. Aunque algunos aceptan los videojuegos, la mayoría expresa una profunda preocupación, y algunos muestran un rechazo total, atribuyéndoles el crecimiento de una cultura de violencia [15].

Cuando ocurren desagradables incidentes que involucran a jóvenes entusiastas de los videojuegos, los líderes de la opinión pública frecuentemente los ponen a juicio, exhibiendo a los jugadores y sus videojuegos de una forma que causa alarma general. En términos similares a los usados para criticar la televisión, generalmente se denigra a los videojuegos. Ellos culpan a los jugadores atribuyendo a los juegos violencia, contenido sexual o características adictivas.

Bajo esta atmósfera, no es difícil entender las razones por las cuales los legisladores han tomado acciones para regular la industria de los videojuegos e incluso han sido debatidos en el senado de los Estados Unidos. Curiosamente, durante este debate, muchos expertos testificaron que hace falta evidencia científica que demuestre una relación directa entre el uso de los videojuegos y cualquier efecto negativo imputable a ellos, excepto en casos muy raros bajo circunstancias inusuales [15]. Por lo anterior, el uso de los videojuegos no puede asociarse de manera directa con ningún efecto negativo, pues es tan sólo un medio a través del cual el jugador se involucra en determinadas actividades y causa los mismos efectos que el cine, la televisión o la Internet, por lo que cuidar el contenido que se exponga al jugador y la forma en que se usen es suficiente para garantizar la ausencia de efectos negativos.

Videojuegos en la educación

Los videojuegos han sido usados en las escuelas por muchos años para ayudar a niños a adquirir habilidades en matemáticas, lenguaje, geografía, ciencias, y en muchas otras áreas de una forma motivante e interesante. Algunas investigaciones muestran que el uso de videojuegos dentro de un

salón de clases puede ser benéfico para el alcance académico, la motivación y las dinámicas de clase [4].

Muchos autores han probado que el uso de videojuegos ayuda a los alumnos a percibir sus deficiencias y ayuda a corregirlas. De acuerdo a estos investigadores, la adaptabilidad y versatilidad de los videojuegos, así como de los jugadores que los controlan, motivan y estimulan el aprendizaje, y hacen que el juego sea útil en casos en donde los niños tienen problemas de concentración. La retroalimentación inmediata que proveen los videojuegos y la necesidad de respuesta continua durante el juego, retan y estimulan a los niños y adolescentes, y despierta la curiosidad, lo cual puede ser extremadamente útil en el aprendizaje [15].

Al aprender jugando se obtienen varios beneficios adicionales, pues de manera intrínseca se desarrollan habilidades motrices, se fomentan el trabajo en equipo, la colaboración y el manejo de conflictos; además, se incrementa la atención, se acelera el aprendizaje, se mejora la memoria y se estimula al alumno para estudiar por largos periodos de tiempo, ya que la interacción con los videojuegos permiten mantener al alumno entretenido y emocionado, eliminando la sensación de aburrimiento durante el estudio.

Para alcanzar metas educativas hemos agregado a los videojuegos desarrollados en nuestra Facultad trivias, videos explicativos, crucigramas, memoramas y algunas otras características que aparte de ser didácticas son divertidas.

Beneficios didácticos para los desarrolladores de videojuegos

Para los alumnos involucrados en el desarrollo de videojuegos las ventajas también son numerosas, pues ponen en práctica el aprendizaje obtenido en un gran número de disciplinas de estudio, entre otras las siguientes:

1. **Algoritmos:** Los videojuegos se encuentran entre el software que consume mayor capacidad computacional, ya que para su funcionamiento es necesario leer

las entradas de parte del usuario, evaluarlas y reaccionar en consecuencia, construyendo complejas gráficas que deben presentarse varias veces en cada segundo, al tiempo que se procesa audio y se analizan señales de red, por lo anterior, los algoritmos que se implementen deben de ser prácticos y eficientes para obtener el máximo desempeño del hardware.

2. **Programación:** Un desarrollador de videojuegos necesita tener profundos conocimientos de programación para hacer uso eficiente del hardware, además, se necesita aprender librerías adicionales para la manipulación de imágenes, audio, señales de red, etcétera.
3. **Graficación:** Las gráficas computacionales forman parte primordial de los videojuegos y su diseño guarda relación directa con la aceptación de los videojuegos. Los juegos más sencillos pueden hacer uso de gráficas 2d y se pueden procesar con hardware básico, a diferencia de las gráficas 3d que necesitan hardware mucho más potente y herramientas más complejas para construirlos.
4. **Programación de dispositivos móviles:** Los juegos para las videoconsolas son demasiado complejos para abordarse en un proyecto escolar, pero los dispositivos móviles como teléfonos celulares y computadoras de mano modernos cuentan con capacidades de procesamiento y almacenamiento lo suficientemente poderosas como para ejecutar videojuegos complejos, además, este tipo de tecnología está presente en prácticamente todos los niveles socioeconómicos y supera ampliamente al número de computadoras personales en el país.
5. **Redes de computadoras:** La interacción entre jugadores hace que los videojuegos se vuelvan más interesantes y atractivos, por lo que dotarlos con características de red es muy importante si se desea mejorar las posibilidades de éxito de un videojuego. En

nuestra Facultad hemos encontrado que la interacción a través de BlueTooth, es bastante práctica y eficiente cuando el número de jugadores es reducido y la distancia entre ellos es corta, además, esta tecnología de comunicaciones es mucho más barata y común en los dispositivos móviles.

6. **Inteligencia artificial:** La inteligencia artificial de cada juego varía considerablemente en función de su tipo, así, un juego de sudoku necesitará una inteligencia artificial reducida, pero uno de fútbol requerirá de complejos algoritmos que lo hagan interesante y entretenido. Un videojuego debe ser inteligente para retar al jugador, pero es importante encontrar el punto correcto de complejidad, ya que si el juego es demasiado sencillo el jugador lo abandona muy pronto por aburrimiento, pero si es demasiado complejo, terminará siendo frustrante; por lo anterior, la dificultad adaptativa está siendo usada con grandes beneficios en las versiones más recientes de videojuegos.
7. **Ingeniería de software:** Los videojuegos se obsoletan con gran rapidez y es frustrante que el trabajo que se está realizando se quede en desuso porque han aparecido en el mercado versiones muy similares antes del lanzamiento del nuestro, por lo que la metodología de desarrollo de un videojuego debe ser rápida, práctica, flexible y eficiente, lo que convierte al diseño de videojuegos en una excelente manera de poner en práctica esta disciplina.
8. **Otras disciplinas:** además de las áreas ya mencionadas, el desarrollo de videojuegos requiere del entendimiento y manejo de varias disciplinas más como: la resolución de problemas, el manejo de audio, capacidades de diseño gráfico, etcétera.

Funcionamiento de un videojuego.

Cada vez que juegas a tu videojuego preferido, dentro del ordenador están ocurriendo muchas

cosas. Se hace patente que un factor importante en un juego es que se mueva con soltura y a buena velocidad. Todos los elementos del juego parecen funcionar de forma independiente, como con vida propia y a la vez. Sin embargo esto es sólo una apariencia, ya que dentro del programa se van sucediendo las diferentes fases de ejecución de forma secuencial y ordenada [2]. Sin embargo el funcionamiento de un videojuego se basa en la ejecución repetitiva de varias tareas:

1. **Entrada:** A fin de comunicarse con el jugador es necesario estar censando constantemente algún dispositivo de entrada como un ratón, un teclado o un joystick, lo cual puede parecer sencillo, pero hacerlo de manera incorrecta puede hacer que el juego sea frustrante.
2. **Visualización:** para crear las gráficas del videojuego constantemente se recurre a la representación repetitiva de imágenes que varían un poco en cada ciclo y provocan la ilusión de movimiento.
3. **Sonido:** Otra forma vital de comunicación con el usuario es la emisión de sonidos, los cuales también guardan relación con la calidad del juego, pues elevan la experiencia y sensación del jugador.
4. **Comunicaciones:** Una de las formas más comunes de comunicar videojuegos es a través de redes TCP/IP, sin embargo en juegos pequeños las piconets son suficientes para brindar ésta característica.
5. **El Game Loop (o ciclo de juego):** Lo primero que hacemos es leer los dispositivos de entrada para ver si el jugador ha realizado alguna acción. Si hubo alguna acción por parte del jugador, el siguiente paso es procesarla, esto es, actualizar su posición, disparar, etcétera, dependiendo de qué acción sea. En el siguiente paso realizamos la lógica de juego, es decir, todo aquello que forma parte de la acción del juego y que no queda bajo control del jugador, por ejemplo, el movimiento de los enemigos,

cálculo de trayectoria de sus disparos, comprobamos colisiones entre la nave enemiga y la del jugador. Fuera de la lógica del juego quedan otras tareas que realizamos en la siguiente fase, como son actualizar el fondo (si lo hubiera), activar sonidos (si fuera necesario), realizar trabajos de sincronización, etcétera. Ya por último nos resta volcar todo a la pantalla y mostrar el siguiente cuadro del juego. Normalmente, el game loop tendrá un aspecto similar a lo siguiente:

```
int done = 0;
while (!done) {
    // Leer entrada
    // Procesar entrada
    // Lógica de juego
    // Otras tareas
    // Mostrar cuadro
} [2].
```

Videojuegos desarrollados en la Facultad de Contaduría Pública del Campus IV de la UNACH

Al percibir las enormes ventajas inherentes al desarrollo de videojuegos, algunos maestros de la Facultad los hemos utilizado como proyectos conjuntos de materia, agrupando a los alumnos en parejas para que trabajen de manera colaborativa. Al inicio del proyecto los alumnos eligen el videojuego que quieren desarrollar, se discuten las partes que debe contener, y se diseña un calendario de actividades que se debe cumplir estrictamente para poder completar el juego en tan solo 3 meses, a continuación exponemos algunos de los más destacados:

Serpientes y escaleras: una implementación electrónica del clásico juego de mesa “Serpientes y escaleras” al que agregamos una serie de preguntas y respuestas de distintos niveles educativos.



Fig. 1. Videojuego desarrollado por: Jesús Malo Escobar y Fredy Alexander Orellana Flores

Contillo Aprendí Astronomía: Un grupo de 6 juegos didácticos para aprender acerca del universo, cuyo nombre alude a la mascota de la Facultad, el periquito “Contillo”. En este juego se decidió situar al jugador en un escenario varias décadas después de la actual, donde la falta de cuidado y amor por la tierra ha ocasionado una catástrofe ecológica que lo ha destrozado (con lo cual se pretende formar conciencia ecológica), por lo que los protagonistas de esta historia deben abandonar la tierra en busca de otro planeta habitable para mudarse.

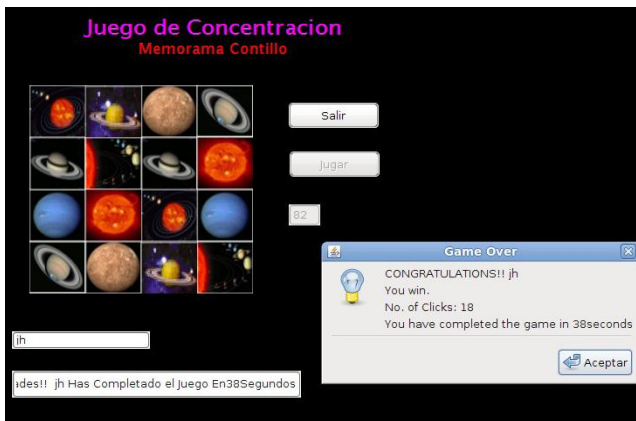


Fig. 2. Módulo desarrollado por Beatriz González Velázquez.

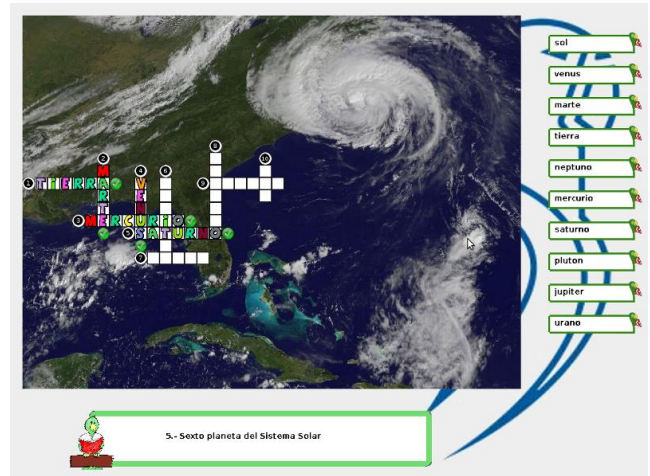


Fig. 3. Módulo desarrollado por Fernando Franco Pérez

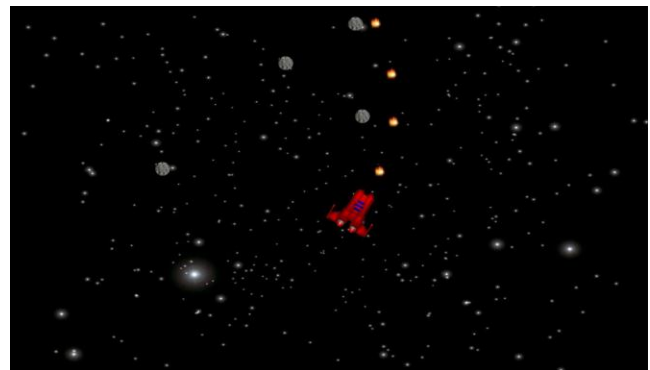


Fig. 4. Módulo desarrollado por Gabriela Alonso Sánchez

Arkanoid: el legendario juego de destrucción de bloques en donde el jugador controla la plataforma para impedir que la pelota salga de la zona de juego:



Fig. 5. Videojuego desarrollado por: Luis Guillermo López Díaz y Diego de Jesús López de León

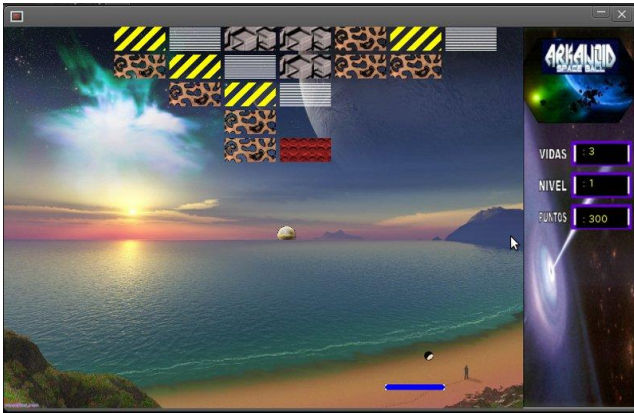


Fig. 6. Videojuego desarrollado por: Luis Alberto Guillén Hernández y Pablo Didier Juárez Alfaro



Fig. 9. Videojuego desarrollado por: Hugo Armando Hernández Marroquín y Gilberto Bartolón López

PuzleBuble: Videojuego cuya intención principal es juntar canicas de colores lanzadas por un cañón situado en la parte baja del juego, antes de que el techo móvil aplaste al jugador.

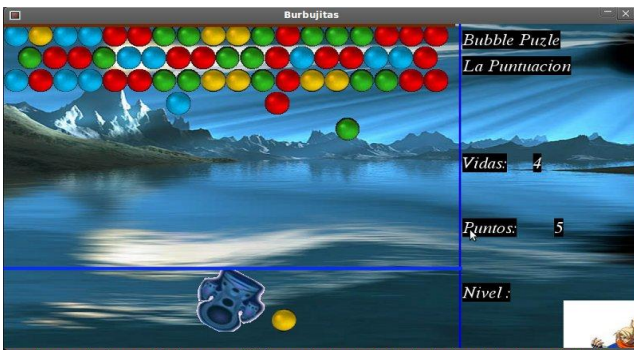


Fig. 7. Videojuego desarrollado por: José Efraín Ramos Gómez y Aremy Aguilar Pacheco



Fig. 10. Videojuego desarrollado por: Alexis Galindo Tovar y José Ramón Moreno Sánchez



Fig. 8. Videojuego desarrollado por: Floridalma López Gómez y Norma Pérez Vicente

Bob & Arrow: En este juego un arquero dispara flechas para destruir globos que si no son destruidos causan pérdida de puntos y vidas.

III. Resultados

Gracias a la implementación de los videojuegos como proyecto de materia en nuestra Facultad, hemos alcanzado diversos resultados, ya que algunos han sido expuestos en foros de investigación como el SENIE 10, el Canal 4, el Planetario de Bachilleres de Chiapas y el III congreso de Investigación UNACH; se han escrito varios artículos al respecto y se tienen dos alumnos titulados mediante el proyecto de tesis denominado “Diseño de un sistema didáctico para niños y adolescentes basado en juego electrónico con tecnología Bluetooth”. Actualmente seis alumnos se encuentran

escribiendo su tesis de licenciatura con juegos educativos como tema.

IV. Conclusiones

El uso de videojuegos en la educación tiene grandes ventajas en la construcción del conocimiento, pues además de todos los beneficios pedagógicos ya señalados para quienes los juegan, también representa una importante oportunidad de aprendizaje para los alumnos de nuestra facultad, pues la programación de videojuegos exige del dominio de múltiples disciplinas del conocimiento que son parte fundamental de nuestro programa de estudios, por tanto, el desarrollo de videojuegos es una excelente manera de evaluar aprendizaje obtenido por nuestro alumnado.

V. Referencias

- [1] A. Chaffin, T. Barnes. "Lessons from a Course on Serious Games Research and Prototyping". 2010 ACM 978-1-60558-937-4/10/06. 8 pages.
- [2] A. García Serrano. Programación de Videojuegos con SDL. Primera edición. Creative Commons. 155 páginas.
- [3] A. Inge Wang. "Extensive Evaluation of Using a Game Project in a Software Architecture Course". 2011 ACM 1946-6626/2011/02-ART5. 28 pages.
- [4] A. Inge Wang, B. Wu. "An Application of a Game Development Framework in Higher Education". International Journal of Computer Games Technology. Accepted 24 February 2009. Volume 2009. Article ID 693267. 12 pages.
- [5] B. COWLEY, D. CHARLES, M. BLACK, R. HICKEY. "Toward an Understanding of Flow in Video Games". ACM Comput. Entertain. 6, 2, Article 20 (July 2008). 27 Pages.
- [6] C. Linehan, B. Kirman, S. Lawson, G. G Chan. "Practical, Appropriate, Empirically-Validated Guidelines for Designing Educational Games". 2011 ACM 978-1-4503-0267-8/11/05. 10 pages.
- [7] D. CLARKE, P. R. DUIMERING. "How Computer Gamers Experience the Game Situation: A Behavioral Study". ACM 1544-3574/06/0700-ART6. 23 pages.
- [8] D. Grammenos. "Game Over: Learning by Dying". 2008 ACM 978-1-60558-011-1/08/04. 10 pages.
- [9] Estudia Profeco a videojuegos en México. <http://www.eluniversal.com.mx/articulos/58456.html>. Consultado el día 09 de Junio de 2011.
- [10] F. Bellotti, R. Berta, A. De Gloria, L. Primavera. "Enhancing the Educational Value of Video Games". 2009 ACM 1544-3574/2009/06-ART23. 18 pages.
- [11] J. Puerto, A. Tamir, F. Perea. "Some new cooperative coverage facility location games". 2010 ACM 978-1-60558-919-0. 3 pages.
- [12] K. McGee, A. T. Abraham. "Real-time team-mate AI in Games". 2010 ACM 978-1-60558-937-4/10/06. 8 pages.
- [13] K. Sung. "Computer Games and Traditional CS Courses". Communications of the ACM. December 2009. Volume 52. No. 12. 5 pages.
- [14] L. Doucet, V. Srinivasan. "Designing Entertaining Educational Games Using Procedural Rhetoric: A Case Study". Association for Computing Machinery, Inc. Los Angeles, California, July 28 – 29, 2010. 6 pages.
- [15] M. DE AGUILERA, A. MÉNDIZ. "Video Games and Education (Education in the Face of a "Parallel School")". ACM Computers in Entertainment, Vol. 1, No. 1, October 2003, Article 01. 14 pages.
- [16] S Björk. "Games, Gamers, and Gaming". 2008 ACM 978-1-60558-197-2/08/10. 5 pages.

- [17] S. Kurkovsky. “Engaging Students through Mobile Game Development”. ACM Special Interest Group on Computer Science Education. 2009 ACM 978-1-60558-183-5/09/03. 5 pages.

VI. Autores

M.G.T.I. Jorge Elpidio Cruz Gálvez obtuvo su título de Maestro en Gestión de Tecnologías de la Información por la Universidad TecMilenio, y es profesor de asignatura de la Licenciatura en Sistemas Computacionales en la Facultad de Contaduría Pública del Campus IV en la Universidad Autónoma de Chiapas.

M. en C. Luis Antonio Álvarez Oval obtuvo su título Maestro en Ciencias de la Computación por la Universidad de Houston – Clear Lake y es profesor de Tiempo Completo de la Licenciatura en Sistemas Computacionales en la Facultad de Contaduría Pública del Campus IV en la Universidad Autónoma de Chiapas.

M.C. Arón de la Cruz Vázquez obtuvo su título de Maestro en Ciencias de la Computación con especialidad en Redes, en el Centro Universitario Cultural del Soconusco y es profesor de Tiempo Completo de la Licenciatura en Sistemas Computacionales en la Facultad de Contaduría Pública del Campus IV en la Universidad Autónoma de Chiapas.

L.S.C. Christian Mauricio Castillo Estrada es Licenciado en Sistemas Computacionales por la Universidad Autónoma de Chiapas y se encuentra cursando el último trimestre de la Maestría en Comercio Electrónico con área terminal en Tecnologías de Información en la Universidad Virtual del ITESM y es profesor de asignatura de la Licenciatura en Sistemas Computacionales en la Facultad de Contaduría Pública del Campus IV en la Universidad Autónoma de Chiapas.

IMS-Learning Desing: Retos de Adaptabilidad Evaluada y Metamodelos de Personalización en las UoL

***Carolina Franco Espinosa, *Leonardo Yamasaki Maza**

*Universidad Autónoma de Chiapas, Facultad de Contaduría Pública, Campus IV. Carretera Puerto Madero Km. 1.5. Tapachula, Chiapas. Tel.(962)625-17-23. {cfranco, lyama}@unach.mx

Resumen

IMS Learning Design (IMS-LD), es una especificación de e-Learning que describe y codifica las metodologías educativas implícitas en un proceso de enseñanza, de forma que sean procesables por un LMS. La adaptabilidad y la personalización, son funcionalidades que están consideradas en la especificación de las “unidades de aprendizaje” (Unit of Learning, UoL). Uno de los objetivos de brindar adaptabilidad en los entornos de aprendizaje, es presentar a los estudiantes los contenidos de acuerdo a su perfil inicial y a los avances que vaya obteniendo durante el desarrollo de un curso. La personalización complementa la adaptabilidad, ya que provee a cada estudiante la capacidad de elegir ciertas características del contenido, de la presentación, de la evaluación y del itinerario, dentro de un marco prefijado de aprendizaje. Sin embargo, considerando que la esencia de una UoL es una secuencia de actividades organizadas para conseguir un aprendizaje por parte de un alumno o grupo de ellos, actualmente ¿cómo se evalúa el éxito de estas secuencias?, es decir, ¿cómo se comprueba que esta secuencia sugerida por un profesor es eficiente para todos los alumnos o un grupo de ellos? Y ¿cómo se pueden obtener los metamodelos eficientes de la personalización de los estudiantes exitosos? En este trabajo se presentan una propuesta para incluir nuevas funcionalidades en el área de aprendizaje adaptativo y personalización.

Palabras Clave: adaptabilidad, metamodelos, diseño instruccional, unidades de aprendizaje, evaluación, personalización.

I. Introducción a IMS-LD

IMS Learning Design, IMS-LD [1], es una especificación de e-Learning, es el resultado de la integración dentro de IMS de la especificación Educational Modeling Language (EML) [2], desarrollada inicialmente en la *Open Universiteit in the Netherlands* (OUNL).

IMS-LD describe y codifica el diseño pedagógico, es decir las metodologías educativas implícitas en un proceso de enseñanza, de forma que sean procesables por un LMS. Su filosofía esencial es: *“independientemente de la aproximación pedagógica, una persona toma un rol en un proceso de aprendizaje-enseñanza, normalmente el rol de alumno o docente. En este rol la persona tiene que trabajar para conseguir ciertos resultados mediante la realización de*

actividades de aprendizaje o soporte más o menos estructuradas en un entorno”. El escenario de aprendizaje específico normalmente se define por los roles que una persona puede tomar, las actividades a realizar y las características particulares del entorno.

En IMS-LD se introduce a la Unidad de Aprendizaje como concepto nuevo y unidad mínima de intercambio entre sistemas (UoL, Units of Learning), el elemento clave de una Unidad de Aprendizaje es la actividad o tarea, que se concibe como uno o más actores (alumnos, profesores) que trabajan para lograr un cierto objetivo educativo en un determinado entorno. El entorno contiene los recursos y los servicios necesarios para realizar la actividad

propuesta. El principio subyacente es que los alumnos aprenden realizando actividades en un entorno, en el cual los objetos de aprendizaje son recursos que permiten o facilitan la tarea.

De acuerdo a lo anterior se puede establecer que la visión es más amplia que la de los objetos de aprendizaje, ya que se contempla el uso de herramientas o de procesos, como la comunicación entre alumnos o entre alumnos y profesores. El rol o papel de un alumno podría cambiar en un determinado momento, por ejemplo, para supervisar el trabajo realizado por otros alumnos.

Una de las funcionalidades que también permite IMS-LD es crear itinerarios de aprendizaje personalizados y adaptativos (permite que el estudiante tenga una experiencia personalizada en función de su rendimiento o de su propia elección, adaptando los contenidos, los servicios proporcionados y el itinerario formativo), asimismo IMS-LD puede representar una gran variedad de modelos pedagógicos y permite que el profesor adapte sus recursos y sus programaciones de aula a clases virtuales de una manera completamente flexible.

II. Qué es una Unidad de Aprendizaje (UoL)

La definición formal de Unidad de Aprendizaje (Unit of Learning, UoL), de acuerdo a [1] es “una unidad de educación o formación completa y autónoma, curso, módulo o lección. La creación de una unidad de aprendizaje implica la creación de un diseño de aprendizaje y también la compilación de sus recursos asociados, bien como archivos contenidos en la unidad o como referencias Web, incluyendo evaluaciones, materiales de aprendizaje e información para configurar el servicio de aprendizaje.”

Se pueden observar los elementos que conforman una UoL, (ver Fig. 1) la cual es un archivo ZIP con:

- Un manifiesto XML
- Archivos opcionales HTML con contenido,

- Archivos (opcionales) XML con referencias al manifiesto y
- Un conjunto de recursos (archivos) de cualquier tipo, referenciados desde el *manifiesto*.

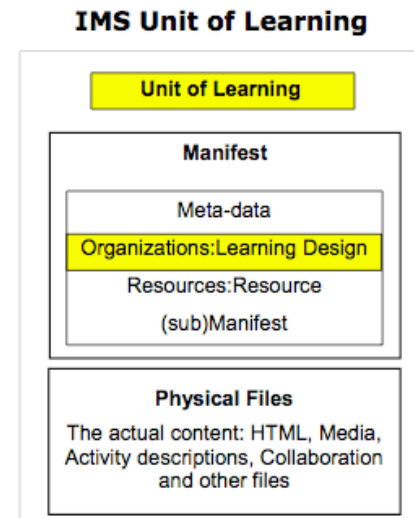


Fig. 1 IMS Unit of Learning

III. Aplicaciones de IMS-LD en *e-learning* y *blended learning*

De acuerdo a [1], [3] las aplicaciones más frecuentes de IMS-LD, desde un punto de vista didáctico son:

- Aprendizaje activo,
- Aprendizaje colaborativo,
- Aprendizaje adaptativo,
- Personalización,
- Retroalimentación dinámica,
- Seguimiento en tiempo de ejecución,
- Nuevas alternativas de evaluación y
- e- Portafolios.

El interés de este trabajo es centrarnos en las funcionalidades de *aprendizaje adaptativo* y en la *personalización*.

IV. Adaptabilidad

Las personas tienen estrategias distintas para aprender aunque compartan una serie de

circunstancias como: la edad, el nivel académico, un tema de estudio o incluso una misma motivación.

Estas estrategias personales dan como resultado aprendizajes que pueden ser más o menos eficaces y a distintas velocidades. Esta es una de las razones por las cuales se presentan resultados tan diversos entre los estudiantes que toman una misma asignatura, con los mismos materiales didácticos, con un mismo profesor y en una misma plataforma educativa.

De acuerdo a [4][5], cuando se brindan funcionalidades adaptativas se busca, elegir los contenidos y presentarlos al estudiante, en función de su perfil inicial y de los resultados progresivos que vaya proporcionando durante el desarrollo del curso.

Asimismo se espera que las plataformas educativas puedan ofrecer a los alumnos una interactividad diferenciada (tomando en cuenta sus particularidades), durante su proceso de aprendizaje. Es decir, que tengan en cuenta sus características personales, preferencias y necesidades a la hora de presentarles información y actividades, asimismo tomen en cuenta sus intereses y objetivos en cada etapa.

Problemática

Si bien es cierto que la esencia de una UoL es una secuencia de actividades organizadas para conseguir un aprendizaje por parte de un alumno o grupo de ellos, uno de los problemas que actualmente se enfrentan es: ¿cómo se evalúa el éxito de estas secuencias?, es decir, ¿cómo se comprueba que esta secuencia diseñada por un profesor, es eficiente para todos los alumnos o un grupo de ellos?

La solución que se propone integrar a las funcionalidades que ya ofrece IMS-LD en materia de *adaptabilidad* son:

Evaluar cada secuencia propuesta por un profesor/diseñador

Secuencia 1

Secuencia 2

...

Secuencia N

Esta evaluación debe estar asociada a la información del alumno, específicamente a los datos guardados sobre sus estilos de aprendizaje, diferenciado por componentes

Alumno X (estilo A 25%, estilo B 30%, estilo C 35%, estilo D 10%)

Los resultados de esta evaluación se verifican contra los inicialmente etiquetados por los profesores/diseñadores de UoL.

De este proceso de verificación se etiquetan con un valor de éxito ponderado

Se realizan los ajustes necesarios

Se integran nuevamente a las UoL.

Para explicar la solución propuesta (ver Fig. 2).

Estas nuevas funcionalidades se están implementando sobre los módulos presentados en [6].

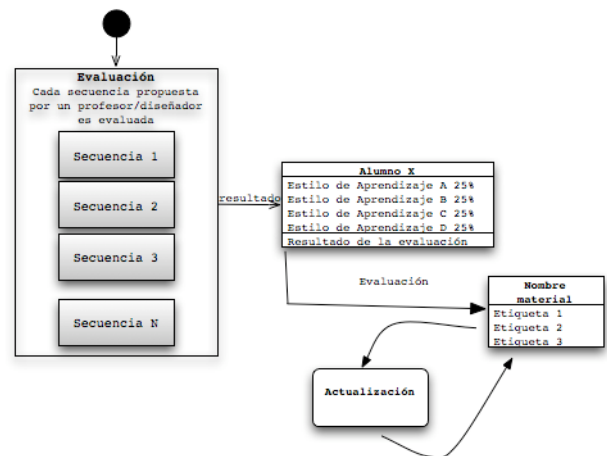


Fig. 2 Evaluación de secuencias

V. Personalización

La personalización complementa el aprendizaje adaptativo ya que permite que cada estudiante pueda elegir ciertas características del contenido, de la presentación, de la evaluación y del itinerario, dentro de un marco prefijado de aprendizaje [5].

La relación entre personalización, *patrones* y *metamodelos*

Los estudiantes necesitan aprender habilidades y métodos de pensamiento, mucho más que absorber una cantidad de conocimientos sobre un tema específico. Debido a que hay algunos estudiantes que logran desarrollar estas habilidades a lo largo de su carrera, pero otros no, uno de los problemas conocidos en el ámbito educativo (y muchas veces ignorado) es cómo enseñar a aprender. Más específicamente, en el tema que nos ocupa: ¿cómo pasar la experiencia de los estudiantes experimentados a los nuevos estudiantes?

En el trabajo realizado por [7][8] dio como resultado un manual de 40 *patrones*, que los estudiantes de nuevo ingreso reciben y que categoriza las diferentes situaciones académicas que todo alumno vive a lo largo de su paso por la Universidad, y en cada una de ellas, describe una serie de actividades ordenadas, con las cuales se puede hacer frente a una determinado reto académica. Por ejemplo: Cómo seleccionar un tema de investigación. Este grupo de trabajo [7][8], tiene una aplicación Web [9] sobre los patrones de aprendizaje, pero no han incluido estos patrones en ningún LMS.

Este manual de patrones, es un claro ejemplo de lo que son los metamodelos aplicados a la vida académica de los estudiantes, ya que son formas exitosas en las que se resuelven retos académicas comunes.

Metamodelos

Los metamodelos son formas con las que se pueden detectar las estrategias exitosas para realizar una determinada actividad. La idea del metamodelo ha sido introducida por la

Programación Neurolingüística [10][11][12][13], y básicamente plantea cómo *modelar-copiar* esa organización de secuencias y obtener los mismos resultados.

Problemática

¿Cómo transmitir las experiencias exitosas en el uso de UoL a otros estudiantes?

Una vez que varios alumnos han utilizado las UoL, actualmente no se guarda la meta información o los patrones exitosos de su utilización. Como patrones nos referimos al resultado de la personalización, que permite actualmente IMS-LD a los alumnos.

Las modificaciones que se propone integrar a las funcionalidades que ya ofrece IMS-LD en materia de *personalización* son:

Medir los resultados obtenidos por cada estudiante en una UoL.
Elegir los mejores resultados
Guardar información de esta personalización (metamodelo)
Hacer un repositorio de metamodelos exitosos en uso de una determinada UoL
Proponer esos metamodelos a otros estudiantes como "tips" de casos de éxito anteriores.

La idea esencial es, (ver Fig. 3) obtener los metamodelos de los estudiantes exitosos en el uso de UoL, guardar esos metamodelos y proponerlos como *patrones* de aprendizajes.

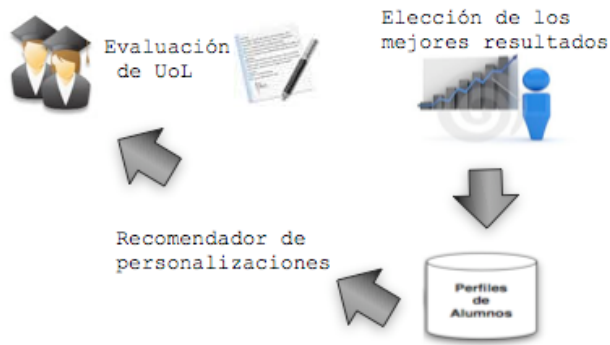


Fig. 3 Obtención de patrones (personalizaciones exitosas)

VI. Metodología

Es necesario analizar y criticar la especificación IMS-LD desde un punto de vista pedagógico y técnico, basándose en la investigación teórica y en la realización concreta de Unidades de Aprendizaje que implementen o intenten implementar situaciones y necesidades educativas con el objetivo de forzar, modificar, extender y complementar la especificación para mejorar el aprendizaje a distancia o mixto asociado.

Actualmente IMS-LD no proporciona elementos para guardar patrones de aprendizaje. En este trabajo se plantearon soluciones específicas a las oportunidades detectadas en las áreas de: *Adaptabilidad y Personalización*.

Esta solución implica el desarrollo de nuevas estructuras y elementos de modelado que permitan guardar metamodelos de aprendizaje (patrones). Estos elementos y estructuras serán utilizados como parte de otros desarrollados de manera específica para sustentar procesos de aprendizaje adaptativo. Constituyen, pues, un conjunto básico para ser utilizado directamente o como parte de un enfoque de personalización.

Desarrollo de nuevas estructuras y elementos de modelado que permita incorporar los patrones desarrollados por [7]. Estos elementos y

estructuras serán utilizados para sugerir estrategias sobre cómo aprender en situaciones académicas concretas.

Con todo ello se pretende que IMS LD sea extendida para lograr una respuesta suficiente y optimizada que proporcione un mayor grado de identificación entre objetivos educativos, recursos tecnológicos y resultados de aprendizaje en la práctica educativa real. En definitiva, una mejora de sus capacidades pedagógicas.

VII. Conclusiones

Debido a que actualmente IMS-LD no proporciona elementos para guardar patrones de aprendizaje. En este trabajo se plantearon soluciones específicas a las oportunidades detectadas en las áreas de: *Adaptabilidad y Personalización*.

El objetivo fundamental se centra en complementar el modelo pedagógico de IMS-LD, añadiendo el concepto de metamodelos, la evaluación de las UoL, así como permitiendo más granularidad en el aprendizaje adaptativo.

En el escenario ideal, es necesaria la evaluación de las soluciones planteadas, en un ambiente real de aprendizaje en educación superior, mediante experimentos controlados, y como una continuidad del trabajo iniciado en [6].

VIII. Referencias

- [1] IMS “IMS Learning Design. Information Model, Best Practice and Implementation Guide”
<http://www.imsglobal.org/content/learningdesign>. 2003.
- [2] R. Koper. *Modeling units of study from a pedagogical perspective: the pedagogical meta-model behind EML*. Educational Technology Expertise Center (OTEC) Open University of the Netherlands. 2001.
- [3] Burgos, D. and Tattersall, C. and Koper, R. *Utilización de estándares en el aprendizaje virtual. Funcionalidades didácticas de la especificación IMS Learning Design*. Educational Technology Expertise Center (OTEC) Open University of the Netherlands. 2005.
- [4] P. Brusilovsky and C. Peylo. *Adaptive and intelligent web-based educational systems*. International Journal of Artificial Intelligence in Education, 13(2- 4):159–172, 2003.
- [5] D. Burgos. Extension of the IMS Learning Design specification based on adaptation and integration of units of learning (Extensión de la especificación IMS Learning Design desde la adaptación y la integración de Unidades de Aprendizaje). Leganés, Madrid, Spain: Doctoral Thesis. Carlos III University of Madrid, 2008.
- [6] C. Franco, L. Yamasaki, A. Domínguez. *Implementación de B-learning en EducAD-UNACH utilizando MOODLE*. Congreso Internacional en Innovación y Desarrollo Tecnológico, CIINDET 2009. Cuernavaca, Morelos, México. 7-9 Octubre 2009.
- [7] T. Iba and T. Miyake. *Learning patterns: A pattern language for creative learners II*. In 1st Asian Conference of Pattern Language of Programs, Mar. 2010.
- [8] T. Iba, T. Miyake, M. Naruse, and N. Yotsumoto. *Learning patterns: A pattern language for active learners*. In 16th Conference on Pattern Languages of Programs, Chicago, Illinois, USA, 2009.
- [9] Learning Patterns.
<http://learningpatterns.sfc.keio.ac.jp/>
- [10] J. M. Paul Tosey, “Neurolinguistic programming: its potential for learning and teaching in formal education,” European Conference on Educational Research, p. 16, September 2003.
- [11] J. S. Joseph O’Connor, Introducción a la PNL, T. A. Press, Ed. Ediciones Urano, 1995, no. ISBN: 978-84-7953-096-9
- [12] R. Dilts, Modeling with NLP. Capitola, Calif.: Meta Publications, 1998.
- [13] R. Bandler and J. Grinder, The structure of magic: a book about language and therapy. Palo Alto, Calif.: Science and Behavior Books, 1975.

IX. Autores

Carolina Franco Espinosa, Ingeniera en Sistemas Computacionales, egresada del Instituto Tecnológico de Tuxtla Gutiérrez (ITTG) en 1996. Obtuvo el grado de Maestría en Administración de Tecnologías de Información en 2000 en el Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM). Profesora de tiempo completo de la Universidad Autónoma de Chiapas. Actualmente candidata a Doctora en Ingeniería Telemática, en la Universidad Carlos III de Madrid, España. Becaria CONACyT.

Leonardo Yamasaki Maza, Diseñador Industrial egresado de la Universidad Autónoma Metropolitana (UAM), Unidad Xochimilco, México, D.F. 1985. Obtuvo el grado de Maestría en Comercio Electrónico en 2006 en el Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM). Profesor de tiempo completo de la Universidad Autónoma de Chiapas. Actualmente es estudiante del Doctorado en Informática y Matemática Computacional de la Universidad de Valencia. Becario CONACyT.

Simulación en SIMNON del Control Directo de Par de un Motor de Imanes Permanentes

^{1,2}F.H. Ramírez-Leyva, ¹E. Peralta-Sánchez, ²J. J. Vásquez-Sanjuan

¹Departamento de Ingenierías, Universidad Popular Autónoma del Estado de Puebla (UPAEP)

²Instituto de Electrónica y Mecatrónica, Universidad Tecnológica de la Mixteca (UTM)

Carretera a Acatlima Km. 2.5 Huajuapán de León, Oaxaca, C.P. 69000, Tel. (01953) 5320399 ext. (555)

Correo electrónico: hugo@mixteco.utm.mx

Resumen

Con sus características de alta densidad de potencia y eficiencia, los motores de imanes permanentes dominan la mayoría de aplicaciones donde se tienen requerimientos de control de alto desempeño de posición, velocidad y par. Existen varias técnicas de control para estos motores, el control directo de par (DTC) o también conocido como control vectorial es el que se presenta en este trabajo. Matlab/Simulink se ha colocado como uno de los simuladores más usados en la actualidad sin embargo SIMNON es un simulador que tiene una interfaz muy simple y con el cual se logran tiempos de simulación mucho mejores que los que se consiguen con Matlab®. En este trabajo se muestra la forma en que se implementó el DTC en Simnon® y una comparación entre las respuestas obtenidas con ambos simuladores.

Palabras clave: Control Directo de Par, Matlab/Simulink, Simnon, Motor de imanes permanentes.

I. Introducción

Recientemente el uso de los motores de imanes permanentes de corriente directa (CD) o alterna (CA) en diversas aplicaciones industriales especialmente para las de baja potencia, se ha incrementado significativamente. Sus principales características son: alta densidad de potencia y eficiencia. El motor de imanes permanentes de CD sin escobillas se utiliza en aplicaciones en el que su alto contenido de rizado de par no impacte, se utiliza principalmente por tener un control simple. El motor de CA sin escobillas, conocido como Motor Síncrono de Imanes Permanentes (PMSM por sus siglas en inglés) se usa principalmente en aplicaciones que requieren un alto desempeño de control. Actualmente el precio de este tipo de motores en unión con los procesadores digitales de señales (DSP) de alta velocidad es competitivo, lo cual ha incentivado su uso, especialmente cuando se requiere en una aplicación de alta eficiencia y desempeño [1].

El método de control más popular de control vectorial es el de campo orientado (FOC por sus

siglas en inglés), fue desarrollado en la década de los 70's. En este método las ecuaciones del motor son transformadas a un sistema de coordenadas que rota en sincronía con el flujo de los imanes permanentes. Esto permite separar las ecuaciones del motor y con ello controlar indirectamente el flujo y torque mediante un controlador PI de corriente, de forma similar al muy conocido control de una máquina de CD [2].

El control DTC inicialmente fue usado para el control de motores de inducción, pero ha mostrado muy buenos resultados en el control de motores PMSM. Su principio de operación se basa en seleccionar el vector de voltaje que se le aplica al estator en función de las diferencias entre el par de referencia y el actual, así como el flujo magnetizante [3]. Cuando se requiere de un control de velocidad se incluye un controlador PI cuya entrada es el error de velocidad y la salida es el par deseado. En este control no se requiere conocer los parámetros del motor, con excepción de la resistencia del estator, además de que no necesita de un modulador PWM para el inversor.

En los últimos años Matlab/Simulink se ha posicionado como una herramienta para la simulación de sistemas de control aplicados a diferentes áreas y en particular en la de electrónica de potencia. El que sea gráfico lo hace una herramienta muy versátil que permite entender en forma visual la forma en que trabaja el sistema [4]. Sin embargo existen simuladores matemáticos como SIMNON, con los cuales es importante experimentar para saber en qué condiciones es mejor usar uno u otro simulador.

SIMNON es un programa que resuelve ecuaciones diferenciales y en diferencias en forma numérica. Por tanto es capaz de simular el comportamiento de sistemas continuos y sistemas discretos representados mediante ecuaciones en diferencias. Ha sido utilizado en muchas universidades del mundo para apoyo en educación e investigación en disciplinas como el Control Automático, Biología, Ingeniería Química, Economía, Ingeniería Eléctrica, etc. También SIMNON ha sido utilizado en la industria para la Simulación de Sistemas de Control [5].

En este trabajo se muestra el procedimiento de las simulaciones en SIMNON y Matlab/Simulink del control de DTC de un motor PMSM, se realiza una comparación de lo que se obtiene con ambos simuladores con el fin de determinar la pertinencia y las condiciones para las cuales es preferible usar uno u otro.

El trabajo se divide en cuatro secciones. En la sección II se describe el modelo del motor de imanes permanentes en el plano dq y la estructura del control de campo orientado. En la sección III la forma en que se implementó el sistema con el controlador en el simulador Simnon. En la sección IV se muestran los resultados y en la V las conclusiones

II. Modelo del Motor

Los motores de imanes permanentes está compuesto por el rotor y el estator al igual que el de inducción o el de DC, la diferencia con estos últimos motores es que usa imanes permanentes para producir el campo magnético en del rotor. En la figura 1(a) se muestra el rotor y como se colocan los imanes en forma superficial para 6 pares de polos. En la figura 1(b) la fotografía de del estator y el rotor. Para un motor PMSM de AC, el embobinado es similar al de inducción en la parte del estator [2].

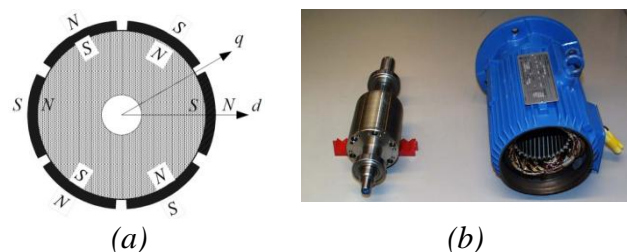


Fig. 1. (a) Diagrama del rotor de un motor de imanes permanentes. (b) Fotografía del motor de imanes mostrando su estator y rotor

De acuerdo a la forma en cómo se colocan las barras de magnetos en el rotor, se clasifica en dos tipos: Motores Síncronos con Imanes Permanentes en el Interior (IPMSM) y Motores Síncronos con Imanes Permanentes Superficiales (SPMSM). Los primeros tienen una mayor robustez mecánica, por lo cual se pueden usar en aplicaciones de alta velocidad.

En el estator se montan 3 bobinas que tienen una separación espacial de 120° , las cuales se alimentan con voltaje trifásico con también 120° eléctricos de desfase entre las fases en secuencia. La corriente que consumen es función de la inductancia propia y la resistencia de cada bobina, y los flujos acoplados entre cada bobina y su interacción con el flujo que genera el imán permanente.

El modelo del motor de imanes permanentes es muy complicado de analizar en un sistema de referencia trifásico estático, sin embargo éste se simplifica mucho cuando se realiza la transformación de Clarke a un nuevo sistema de

referencia de dos ejes (Ecuación 1). Los dos tipos de transformaciones usadas son $\alpha\beta$, que es un sistema estacionario con α y β ortogonales. La transformación de Park asociada al rotor se le llama dq , y es un sistema de referencia de dos ejes rotando a la velocidad del rotor (Ecuación 2). En la figura 2 se muestra la relación de ejes en dq .

$$\begin{bmatrix} V_\alpha \\ V_\beta \end{bmatrix} = \frac{2}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} V_d \\ V_q \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} V_\alpha \\ V_\beta \end{bmatrix} \quad (2)$$

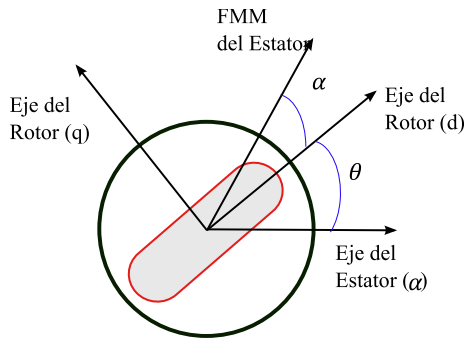


Fig. 2. Representación del motor en los ejes dq

El motor PMSM en el sistema de referencia dq se modela con dos ecuaciones diferenciales para la parte eléctrica (Ecuación 3), donde V_d , V_q , i_d e i_q son los voltajes y corrientes en el eje de referencia dq , L_d y L_q son las inductancias, R la resistencia del estator, Ψ_{PM} es la magnitud del flujo de los imanes permanentes y ω_r la velocidad eléctrica del rotor.

$$\begin{aligned} V_d &= L_d \frac{di_d}{dt} - \omega_r L_q i_q + R i_d \\ V_q &= L_q \frac{di_q}{dt} + \omega_r (L_d i_d + \Psi_{PM}) + R i_q \end{aligned} \quad (3)$$

La ecuación diferencial 4 corresponde a la parte mecánica del motor, donde τ_e es el par eléctrico del motor, T_L el par de la carga, T_f el par de la fricción, J el momento de inercia del rotor y P el número de par de polos. τ_e es función de las corrientes en dq , como se muestra en la misma ecuación.

$$\begin{aligned} \tau_e - T_L - T_f &= \frac{J}{P} \frac{d\omega_r}{dt} \\ \tau_e &= \frac{3}{2} P [i_d i_q (L_d - L_q) + i_q \Psi_{PM}] \end{aligned} \quad (4)$$

La velocidad en revoluciones por minuto (ω_{rpm}) está relacionada con la velocidad eléctrica y el par de polos. La ecuación (5) es la que se utiliza para hacer esta conversión.

$$\omega_{rpm} = \frac{30 \omega_r}{P\pi} \quad (5)$$

El control DTC aplica el vector de voltaje que requiere el motor en función de donde se encuentra localizado el rotor. La posición del rotor es función del flujo generado por los campos del estator y la posición inicial. Una vez que se tiene esta información se le aplica el vector de voltaje adecuado para incrementar o decrementar el par o el flujo magnético. Todos estos cálculos se realizan en el eje de referencia $\alpha\beta$. El flujo magnético en α y β (φ_α y φ_β) es función del voltaje v_α y v_β , así como la corriente i_α e i_β como se muestra en la ecuación (6).

$$\begin{aligned} \varphi_\alpha(t) &= \int_{t_0}^t [v_\alpha - R i_\alpha] dt + \varphi_\alpha(t_0) \\ \varphi_\beta(t) &= \int_{t_0}^t [v_\beta - R i_\beta] dt + \varphi_\beta(t_0) \\ \varphi_S &= |\varphi_S| e^{j\theta} = \sqrt{\varphi_\alpha^2 + \varphi_\beta^2} e^{j \tan^{-1} \left(\frac{\varphi_\beta}{\varphi_\alpha} \right)} \end{aligned} \quad (6)$$

El único parámetro que se requiere conocer es la resistencia del embobinado así como el estado inicial del flujo ($\varphi_\alpha(t_0)$ y $\varphi_\beta(t_0)$). Ambos flujos como son ortogonales se pueden representar como un vector de magnitud $|\varphi_S|$ y ángulo de fase θ , la cual en condiciones de estado estable corresponde a la posición del rotor.

El par electromagnético generado por el rotor es función de la corriente y flujo en el sistema $\alpha\beta$ como se muestra en la ecuación (7).

$$\tau_e = \frac{3}{2} P [\varphi_\alpha i_\beta - \varphi_\beta i_\alpha] \quad (7)$$

Cuando se aplica la transformación Clarke a los voltajes trifásicos, la grafica v_α vs v_β , describe un círculo en el cual cada punto desde el origen hasta el extremo del círculo es el vector V_s . Esto se muestra en la figura 3. Como se puede ver está dividido en 6 sectores, los cuales abarcan un ángulo de 60° . En esta figura también se muestran los 8 vectores que se pueden generar con el inversor (de U_0 a U_7).

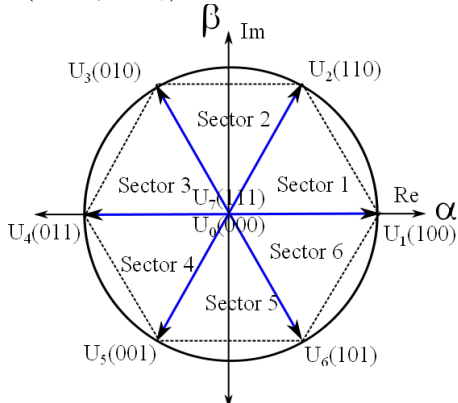


Fig. 3. Gráfica de V_s en el eje de referencia $\alpha\beta$

En la figura 4 se muestra el diagrama a bloques del control DTC. Intrínsecamente sirve para el control de par o del flujo (para este último el flujo deseado es el del imán), sin embargo si se quiere controlar la velocidad se incluye un control PI para convertir la velocidad deseada en el torque deseado. El error de torque y flujo es la diferencia entre lo medido y lo deseado. Éste error entra a un comparador de histéresis cuya salida es 0 o 1. Existen tablas de conmutación que especifican el vector a aplicar (de U_1 a U_6) en función de la salida de los comparadores y del sector en que se encuentra el rotor. La tabla 1 muestra ésta información.

La forma en cómo funciona la tabla se muestra en la figura 3, en ella están marcados cada uno de los 6 vectores en el plano $\alpha\beta$. Si la salida del comparador es $\phi=1$ y $\tau=1$ y el rotor se encuentra en el sector 1, el vector de voltaje que se aplica es U_2 , pero si el torque hubiera sido $\tau=0$, el vector a aplicar es U_6 . De la misma manera se explica para los otros casos.

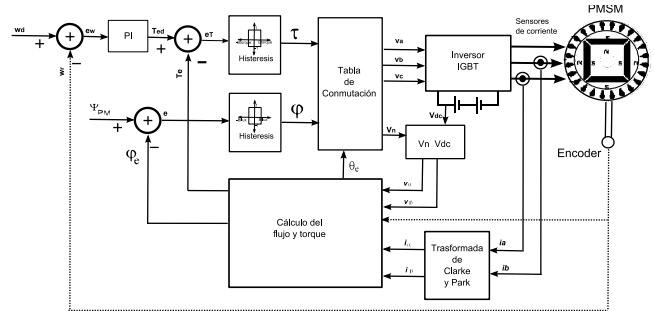


Fig. 4. Diagrama a bloques del control directo de par.

ϕ	τ	θ (Sector)					
		$\theta(1)$	$\theta(2)$	$\theta(3)$	$\theta(4)$	$\theta(5)$	$\theta(6)$
1	1	U_2	U_3	U_4	U_5	U_6	U_7
	0	U_6	U_1	U_2	U_3	U_4	U_5
0	1	U_3	U_4	U_5	U_6	U_1	U_2
	0	U_5	U_6	U_1	U_2	U_3	U_4

Tabla 1. Tabla de conmutación para el inversor

Los vectores de voltaje son generados por un inversor, el cual está hecho por un puente con seis IGBT's, el cual es alimentado por un bus de corriente directa (CD) de 180V o 360V. La salida se toma en la unión de las ramas como se muestra en la figura 5. De acuerdo a las combinaciones de encendido y apagado de cada transistor, solo se tienen 8, a cada una de ellas se les llama vector. En la tabla 2 se muestra el voltaje de cada fase por cada vector.

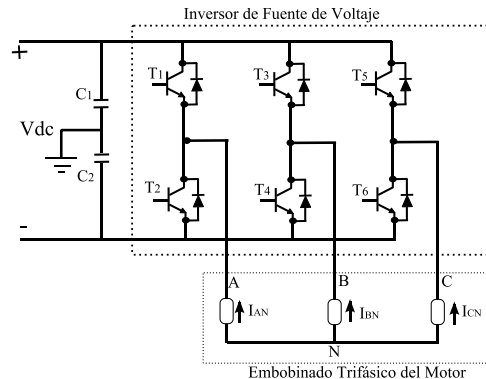


Fig. 5. Puente de IGBT's del inversor

La corriente que demanda el motor es monitoreada por sensores de corriente, por ser un sistema balanceado solo se necesitan medir dos y la tercera se estima. El voltaje del bus de CD también es medido y con estas dos señales se calcular el flujo y el torque de la máquina.

Vector	V_a	V_b	V_c
U_0	0	0	0
U_1	$\frac{-V_{DC}}{3}$	$\frac{-V_{DC}}{3}$	$\frac{2V_{DC}}{3}$
U_2	$\frac{-V_{DC}}{3}$	$\frac{2V_{DC}}{3}$	$\frac{-V_{DC}}{3}$
U_3	$\frac{-2V_{DC}}{3}$	$\frac{V_{DC}}{3}$	$\frac{V_{DC}}{3}$
U_4	$\frac{2V_{DC}}{3}$	$\frac{-V_{DC}}{3}$	$\frac{-V_{DC}}{3}$
U_5	$\frac{V_{DC}}{3}$	$\frac{-2V_{DC}}{3}$	$\frac{V_{DC}}{3}$
U_6	$\frac{V_{DC}}{3}$	$\frac{V_{DC}}{3}$	$\frac{-2V_{DC}}{3}$
U_7	0	0	0

Tabla 2. Voltajes de salida del inversor para cada vector

III. Implementación en Matlab y Simnon

Matlab/Symulink cuenta con el toolbox SimPowerSystems que permite simular sistemas eléctricos de potencia, a la vez que se puede combinar con las demás herramientas que posee. Se usó la versión 7.1 de Matlab. En la figura 6 se muestra el diagrama en Symulink con la implementación del modelo del motor así como del control DTC.

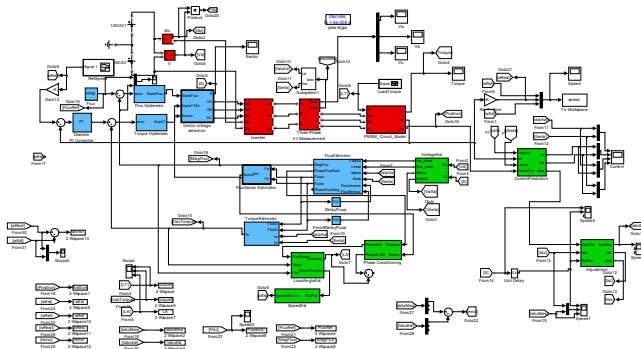


Fig. 6. Modelo del motor y control DTC en Matlab/Symulink

Aunque tiene muchos bloques el modelo del DTC, permite tener una representación gráfica de cómo fluyen las señales y su interrelación. Los parámetros del motor se muestran en la tabla 3. El modelo del motor se implementó usando fuentes de voltaje controladas por corriente, el voltaje del inversor se hace pasar por la transformada de Clarke y Park. Matlab tiene modelos de IGBT's a los cuales se les puede medir la corriente, por lo cual en un bloque se

implementó las tablas de conmutación de los transistores.

Voltaje del bus V_{DC}	300V
Par de polos P	3
Magnetización de los Imanes Ψ_{PM}	0.2547 Wb
Momento de Inercia J	0.000141 kg/m ²
Resistencia R	0.86 Ω
Inductancia en d L_d	0.00665 H
Inductancia en q L_q	0.00665 H
Fricción Viscosa B	0.00038

Tabla 3. Parámetros del motor

Simnon usa una representación en variables de estado, por lo cual el modelo del motor se tuvo que poner en esta forma, de igual manera el controlador PI. Se usaron 4 archivos para simular el DTC, los cuales se muestran en el apéndice 1, estos son:

- MotPMSM.T. Este archivo implementa el modelo del motor. Recibe como entrada el voltaje de cada fase y da como salida las corrientes de cada fase y la velocidad. En esta parte también se define el torque de carga.
- ContDTCk.T. En este archivo se definen los bloques que conforman el control DTC. Recibe como entrada las corrientes de fase y la velocidad y envía la el voltaje de cada fase. También en esta parte se define la velocidad deseada.
- Conxion.T. En este archivo se define la conexiones de las señales que van a cada uno de los archivos que conforman el modelo más el control.
- Macro1.T. Para facilitar la simulación y la gráfica de las señales se hizo esta macro, el cual es el que se ejecuta de la línea de comandos en Simnon.

En la figura 7 se muestra un diagrama a bloques conceptual de la forma en que fue hecho el control DTC en Simnon. Entre líneas punteadas

se pone en que archivo se encuentra definido y las líneas de colores representan el flujo de señales de y hacia donde van. En cada línea se le pone el nombre de la señal como se manejan en cada archivo de la simulación.

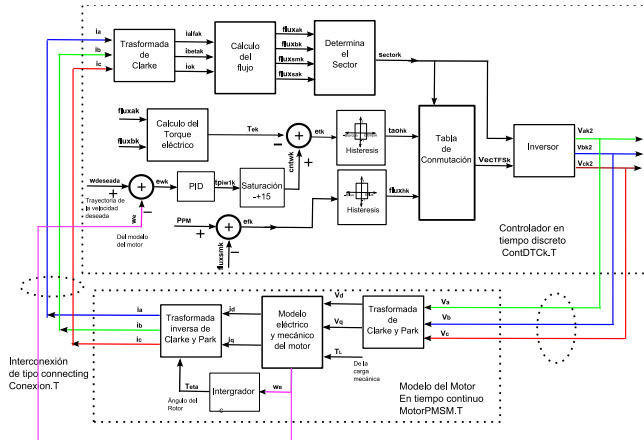


Fig. 7. Diagrama a bloques conceptual de la implementación DTC en Simnon.

El principal problema que se tuvo para implementar el control DTC fue que Simnon no tiene forma de definir tablas, solo la función IF ELSE, así como funciones booleanas. Por tal motivo se tuvo que emplear ambas en conjunto para que se pudiera realizar la tabla de los vectores. Lo mismo sucedió para definir los voltajes del inversor. Para la histéresis se usaron estado actual y anterior con el fin de que se pudiera tener la memoria y realizar la histéresis.

El controlador PI se hizo con una variable de estado para la parte del integrador. Las constantes se sintonizaron por el método de respuesta en frecuencia, para lo cual se obtuvo un modelo linealizado del motor, despreciando la componente de secuencia cero. Se definió un margen de fase de 60° y una frecuencia de corte de un orden menor a la frecuencia de conmutación del inversor. Con estas consideraciones las ganancias proporcional e integral usadas fueron $k_p=7.9$ y $k_i=34.247$ [8]. También se les implementó la función saturación con límites de ± 15 .

IV. Pruebas y Resultados

Para verificar el correcto funcionamiento del modelo hecho en Matlab y Simnon, se procedió a probarlo ante una variación de la velocidad deseada y el torque de la carga. En la figura 8 se muestran los perfiles de cada una de estas dos señales. La velocidad deseada es cero y en $t=0.2s$ cambia a 1000rpm en $0.4s$ cambia a -1000rpm y a partir de $0.6s$ se mantiene en 1000rpm. Para el torque de carga en $t=0.1s$ cambia de 0Nm a 9Nm, en $t=0.8s$ cambia a -9Nm y en $1s$ vuelve a regresar a 9Nm, finalmente en $t=1.2s$ éste queda en 0Nm.

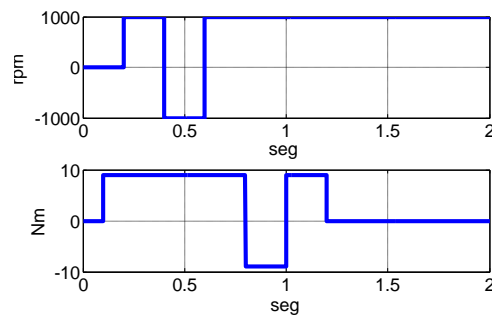


Fig. 8. Perfil de velocidad y par aplicado al modelo del sistema

En la figura 9 se muestran los resultados de la simulación en Simnon de la velocidad del motor (wrpm) el error de velocidad (ewk) y el par generado por el motor (thk). En la figura 10 se muestran las mismas simulaciones obtenidas con Matlab. Como se puede ver, cada que ocurre un cambio de par de carga o de la velocidad deseada ocurre un cambio de la velocidad del motor con respecto a la deseada. Una vez que se estabiliza el motor sigue a la deseada.

En la gráfica del error de velocidad se observa este comportamiento, ya que cuando se estabiliza el error es casi cero, excepto cuando ocurre un cambio en el par de carga o en la velocidad deseada, en condiciones estables el error es menor a 10rpm. Con respecto al par el perfil del par del motor sigue muy de cerca al par que demanda la carga, excepto cuando hay cambios, ya que en este momento se necesita aplicar más o menos par para compensar las mismas.

En la figura 11 se muestra la grafica del flujo en los ejes α β , así como los vectores de voltaje en el mismo plano. Como se puede ver se genera un circulo continuo ya que el rotor gira continuamente, cuando $\varphi_{\beta}=0$ la magnitud $\varphi_{\beta}=0.2547$ que corresponde al flujo del imán permanente. El voltaje en los ejes $\alpha\beta$ generan lo mismo que se puso en forma conceptual en la figura 3, las líneas continuas se generan ya que salta de un vector a otro.

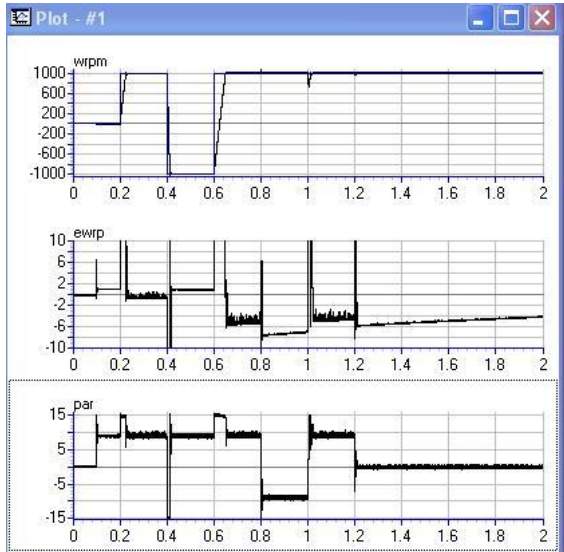


Fig. 9. Velocidad, error de velocidad y par del motor en Simnon

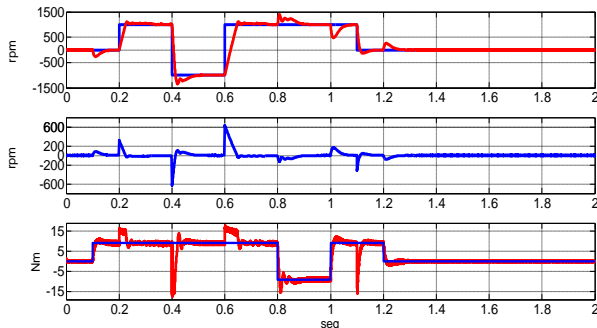


Fig. 10. Velocidad, error de velocidad y par del motor en Matlab

Se midieron los tiempos que tardaban ambos simuladores en ejecuta el sistema. El tiempo total de simulación fue de 2s, con un tiempo de muestreo $T_s=50\mu s$ en Matlab y $10\mu s$ en Simnon, se usó una Laptop Marca Dell Modelo Inspiron con un procesador Pentium IV. El tiempo que tardó en hacer la simulación Matlab fue de 157s, en Simnon se llevó 20s. Como se puede ver es mucho más rápida la simulación en este último.

Sin embargo no se tienen disponibles todas las señales de la simulación, a lo más se pueden definir 10 variables (en el macro), si se requiere otra se necesita redefinir el macro con las variables deseadas.

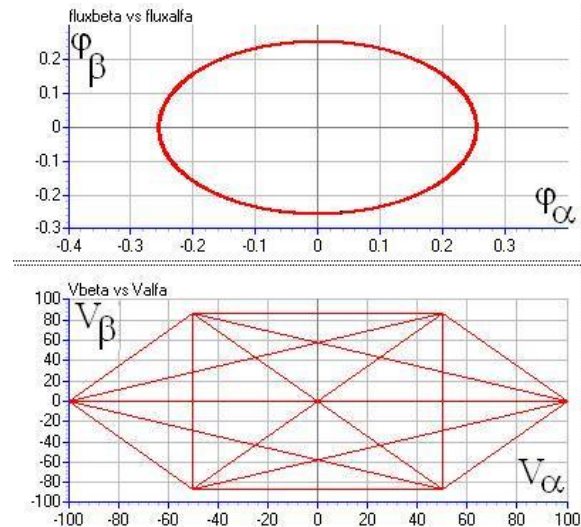


Fig. 11. Gráfica del flujo en α vs β y el voltaje en ángulo α vs β

En la figura 12 se muestra la grafica de las corrientes de la fase a, b y c, como se puede ver mientras el par de carga es cero, la corriente por fase es cero. Cuando hay par de carga existe corriente cuya magnitud es menor a 10 A.

En la figura 13 se muestran las gráficas de los voltajes generados por el inversor en la fase a, b y c. Cuando se muestra en todo el tiempo de simulación no se observan los detalles, por lo cual solo se muestra la simulación de 0.95s a 1.05s. Como se puede ver los voltajes son discretos y solo se tienen 4 valores diferentes tal como se especifica en la tabla 2.

En la figura 14 se presentan las gráficas de los voltajes en d y q y las corrientes en este mismo marco de referencia.

VI. Conclusiones

En el presente trabajo se mostró la forma en que fue implementado el control directo de par a nivel de simulación en los simuladores matemáticos Simnon y Matlab/Simulink. Se describieron todas las ecuaciones necesarias para realizar la simulación, así como las constantes del

controlador PI. Las simulaciones obtenidas entre uno y otro ambiente fueron muy similares.

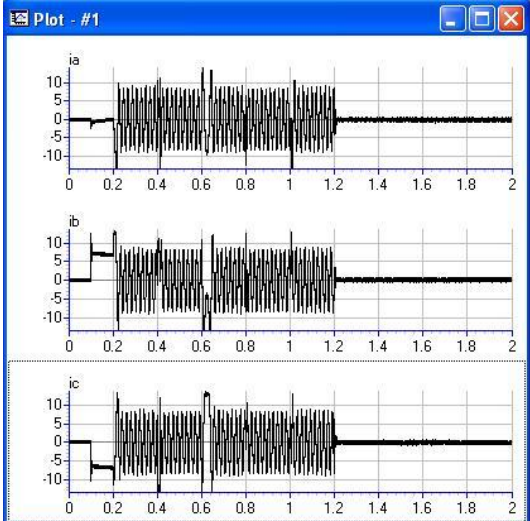


Fig. 12. Gráficas de las corrientes de fase a, b y c

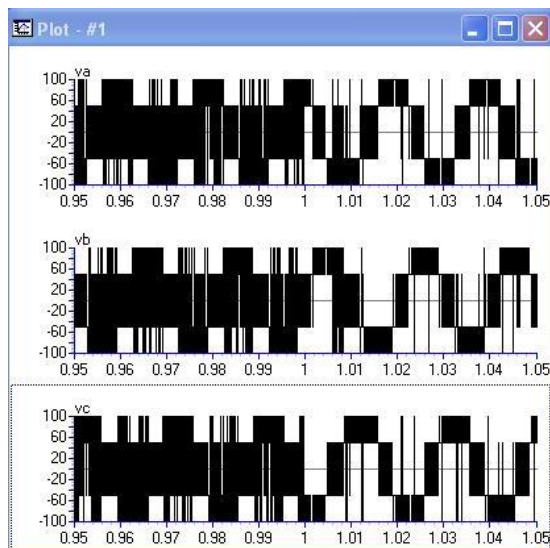


Fig. 13. Gráficas de los voltajes de fase a, b y c

Los problemas que se tuvieron al implementar el DTC en Simnon, fueron que hay que realizar tablas y estas no se pueden definir de forma directa por lo cual se tuvo que usar funciones lógicas con sumas algebraicas para poder generarlas. Otra ventaja de haberlo realizado en Simnon es que para implementación práctica en un DSP que sea programado en C, es más directo recodificar el controlador hecho en Simnon que el que se hizo en Matlab/Simulink.

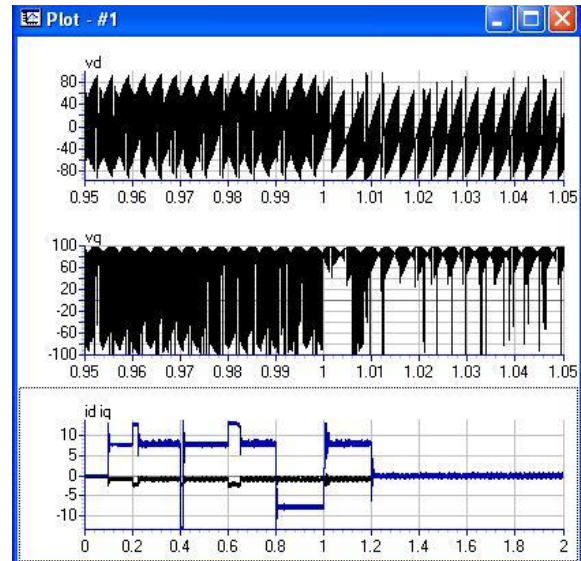


Fig. 14. Gráficas de los voltajes V_d , V_q y corrientes i_d e i_q

Se recomienda usar Simnon cuando se quiere obtener resultados rápidos del modelo ya que éste se ejecuta más rápido que el realizado en Matlab, sin embargo Matlab cuenta con una representación visual más clara de la forma en que interactúan los diferentes módulos del controlador, además de que se disponen de todas las señales en forma inmediata.

Como posibles trabajos a futuro se tiene el realizar un modelo más real de inversor. También el experimentar con otros tipos de controladores como son el de Control Directo de Par en su versión sin sensores.

VII. Autores

F. Hugo Ramírez Leyva obtuvo su título de Maestría en Ciencias con especialidad en electrónica por el Instituto Nacional de Astrofísica Óptica y Electrónica. Actualmente se desempeña como profesor investigador, adscrito al Instituto de Electrónica y Mecatrónica de la Universidad Tecnológica de la Mixteca. Actualmente es estudiante de doctorado en Ingeniería Mecatrónica en la Universidad Popular Autónoma del Estado de Puebla

Edgar Peralta Sánchez. Licenciado en Electrónica por la BUAP, Puebla, México

(1995). Maestro en ciencias con especialidad en Ingeniería Eléctrica por LA UDLAP, Puebla, (1999). PhD en Ingeniería Eléctrica por la Universidad de Manchester, Reino Unido. (2006). Actualmente es Catedrático en el Departamento de Ingenierías de la Universidad Popular Autónoma del Estado de Puebla (UPAEP).

VIII. Referencias

- [1] Zeroug. H and Hocine. L, “Sensorless Control of a Surface Mounted Permanent Magnet Motor by Signal Injection Associated with Space Vector Modulation Technique”, pag 1573-1580, 2008
- [2] Dariusz Świerczyński, “Direct Torque Control with Space Vector Modulation (DTC-SVM) of Inverter-Fed Permanent Magnet Synchronous Motor Drive”, Doctoral Thesis, Warszawa, 2005, Warsaw University Of Technology.
- [3] L. Zhong, M. F. Rahman, W. Y. Hu & K. W. Lim. “Analysis of Direct Torque Control in Permanent Magnet Synchronous Motro Drives”. IEEE Transactions on Power Electrónica, Vol. 12, No. 3, May 1997.
- [4] Mats Leksell, Lennart Harnefors and Magnus Jansson “Direct Sensorless Speed Control of PM-Motors - a Simple and Effective Sensorless Method” 2001
- [5] Enrique L. Carrillo Arroyo, "Modeling And Simulation Of Permanent Magnet Synchronous Motor Drive System", University Of Puerto Rico Mayagüez Campus, 2006 .
- [6] Simnon, “<http://dctrl.fi-b.unam.mx/tutoriales/digital/simnon21.html>. Fecha de último acceso 22/Ocubre/2010
- [7] Muhammed Fazlur Rahman, Md. Enamul Haque, Lixin Tang, and Limin Zhong. “Problems Associated With the Direct Torque Control of an Interior Permanent-Magnet Synchronous Motor Drive and Their Remedies”. IEEE Transactions on

Industrial Electronics, Vol 51, No. 4, August 2004.

- [8] Pedro C. Castellanos Morales, Edgar Peralta Sánchez, Jaime J. Rodríguez Rivas. “Control del Motor de Imanes Permanentes sin Sensor de Posición”. Reunión de Verano RVP-AI’2010, Acapulco Gro., del 11 al 17 de Julio de 2010.

Apéndice 1. Modelo en Simnon

Modelo de la planta ModPMSM.T

```
DISCRETE SYSTEM ContDTCk
INPUT iakk ibkk ickk wrpmk
OUTPUT Vakk Vbkk Vckk

" States and time variables:
STATE fluxaka fluxbka secka fluxhka intwka taohka
zvalfaka zvbetaka zsecak
NEW fluxak fluxbk seck fluxhk intwk taohk
zvalfak zvbetak zseck
TIME t
TSAMP ts
ts=t+h
h=1e-5

"Envia y recibé Variables
Vakk=vak2
Vbkk=vbk2
Vckk=vck2
iak=iakk
ibk=ibkk
ick=ickk
wrpmk=wrpmk
"ialfak=(iak)
"ibetak=(ibk-ick)/sqrt(3)
ialfak=(2/3)*(iak-ibk/2-ick/2)
ibetak=(2/3)*(sqrt(3)/2)*(ibk-ick)
iok=iak+ibk+ick
"Calculo del flujo en alfa y beta
fluxak=fluxaka+(valfak-R*ialfak)*h
fluxbk=fluxbka+(vbetak-R*ibetak)*h
fluxaka:0.2547
fluxsmk=sqrt(fluxak*fluxak+fluxbk*fluxbk)
"Magnitud del flujo
fluxsak=atan2(fluxbk,fluxak)
"Ángulo de fase del flujo
fluxsakg=180*fluxsak/3.1416
" Initializations:
"Determina el Sector
tp1sek=sign(fluxak)
tp2sek=fluxak*tan(pi/6)
tp3sek=sign(fluxbk)
tp4sek=abs(fluxbk)>abs(tp2sek) "Compara si
la magitud de b es mayor que a

"Con flujo en a positivo
tp2sek= if ((not tp1sek) and (not tp3sek) and
(not tp4sek) ) then 4 else 0
tp3sek= if ((not tp1sek) and (not tp3sek) and
tp4sek ) then 5 else 0
tp4sek= if ((not tp1sek) and tp3sek and (not
tp4sek) ) then 4 else 0
tp5sek= if ((not tp1sek) and tp3sek and
tp4sek ) then 3 else 0
tp6sek= if (tp1sek and (not tp3sek) and (not
```

```

tplabsek) ) then 1 else 0
  tp7sek= if (tplsek and (not tplbsek) and
tplabsek) then 6 else 0
  tp8sek= if (tplsek and tplbsek and (not
tplabsek) ) then 1 else 0
  tp9sek= if (tplsek and tplbsek and tplabsek )
then 2 else 0
    "Con flujo en a negativo

tp10sek=tp2sek+tp3sek+tp4sek+tp5sek+tp6sek+tp7sek+
tp8sek+tp9sek
  sectork=tp10sek
  seck=sectork
  secka:1
"CONTROLADOR DE VELOCIDAD Y FLUJO
  ulk=if ts<0.2 then 0 else 1000
  u2k=if ts<0.4 then 0 else -2000
  u3k=if ts<0.6 then 0 else 2000
  u4k=if ts<1.1 then 0 else 0
  Wdesk=ulk+u2k+u3k+u4k "Velocidad deseada en
RPM
  ewk=(Wdesk-Wrpmk)*P*pi/30
  intwk=intwka + h*ewk
  tpiw1=Kpw*ewk + Kiw*intwk
  tpiw2= if tpiw1 > 15 then 15 else tpiw1
  tpiw3= if tpiw2 < -15 then -15 else tpiw2
  cntwk=tpiw3
  Tek=(3*P/2)*(fluxak*ibetak-fluxbk*ialfak)
  etk=cntwk-Tek
  efk=Ppm-fluxsmk

"IMPLEMENTA LA HISTERESIS, TABLA DE CONMUTACION Y
VECTOR
  "Implementación de la histeresis para el flujo y
torque

  temp1fk= if efk > Bflux then 1 else 0
  temp2fk= if efk < -Bflux then 1 else 0
  temp3fk= if ( (efk > -Bflux) and (efk < Bflux))
then fluxhka else 0
  fluxhk=temp1fk + ((not temp2fk) and temp3fk)
  stork=etk
  temp1tk= if stork > Btorque then 1 else 0
  temp2tk= if stork < -Btorque then 1 else 0
  temp3tk= if ((stork > -Btorque) and (stork <
Btorque)) then taohka else 0
  taohk=temp1tk + ((not temp2tk) and temp3tk)
  "Identificación del sector
  temp1s=sectork>0 and sectork<1.5
  temp2s=sectork>1.5 and sectork<2.5
  temp3s=sectork>2.5 and sectork<3.5
  temp4s=sectork>3.5 and sectork<4.5
  temp5s=sectork>4.5 and sectork<5.5
  temp6s=sectork>5.5 and sectork<6.5
  "Vector para flujo 1 torque 1 y sector 1-6
  v11k=if (temp1s) then 2 else 0
  v12k=if (temp2s) then 3 else 0
  v13k=if (temp3s) then 4 else 0
  v14k=if (temp4s) then 5 else 0
  v15k=if (temp5s) then 6 else 0
  v16k=if (temp6s) then 1 else 0

vft11k=(v11k+v12k+v13k+v14k+v15k+v16k)*(taoh
k and fluxhk)
  "Vector para flujo 1 torque 0 y sector x
  v101k=if (temp1s) then 6 else 0
  v102k=if (temp2s) then 1 else 0
  v103k=if (temp3s) then 2 else 0
  v104k=if (temp4s) then 3 else 0
  v105k=if (temp5s) then 4 else 0
  v106k=if (temp6s) then 5 else 0

vft10k=(v101k+v102k+v103k+v104k+v105k+v106k)*((not
taohk) and fluxhk)
  "Vector para flujo 0 torque 1 y sector x
  v011k=if (temp1s) then 3 else 0
  v012k=if (temp2s) then 4 else 0
  v013k=if (temp3s) then 5 else 0
  v014k=if (temp4s) then 6 else 0
  v015k=if (temp5s) then 1 else 0
  v016k=if (temp6s) then 2 else 0

vft01k=(v011k+v012k+v013k+v014k+v015k+v016k)*(taoh
k and (not fluxhk))
  "Vector para flujo 0 torque 0 y sector x
  v001k=if (temp1s) then 5 else 0
  v002k=if (temp2s) then 6 else 0
  v003k=if (temp3s) then 1 else 0
  v004k=if (temp4s) then 2 else 0
  v005k=if (temp5s) then 3 else 0
  v006k=if (temp6s) then 4 else 0

vft00k=(v001k+v002k+v003k+v004k+v005k+v006k)*((not
taohk) and (not fluxhk))
  "VECTOR A USAR
  zseck=vft11k + vft10k + vft01k + vft00k
  vecTFSk=zsecak
"GENERA EL INVERSOR Y LOS VOLTAJES
  VDCk=150*2/3
  VDC2k=150
  tetainvk=(vecTFSk-1)*60*3.1416/180
  valfk=(VDCk)*cos(tetainvk)
  vbetk=(VDCk)*sin(tetainvk)
  zvalfak=valfk
  zvbetak=vbetk
  zvalfaka:0
  valfak=zvalfaka
  vbetak=zvbetaka
  "Genera los voltajes de linea discretos
  vask1=if (vecTFSk>0.5 and vecTFSk<1.5) then
2*VDC2k/3 else 0
  vask2=if (vecTFSk>1.5 and vecTFSk<2.5) then
VDC2k/3 else 0
  vask3=if (vecTFSk>2.5 and vecTFSk<3.5) then -
VDC2k/3 else 0
  vask4=if (vecTFSk>3.5 and vecTFSk<4.5) then -
2*VDC2k/3 else 0
  vask5=if (vecTFSk>4.5 and vecTFSk<5.5) then -
VDC2k/3 else 0
  vask6=if (vecTFSk>5.5 and vecTFSk<6.5) then
VDC2k/3 else 0
  vask2=vask1+vask2+vask3+vask4+vask5+vask6
  vbsk1=if (vecTFSk>0.5 and vecTFSk<1.5) then -
VDC2k/3 else 0
  vbsk2=if (vecTFSk>1.5 and vecTFSk<2.5) then
VDC2k/3 else 0
  vbsk3=if (vecTFSk>2.5 and vecTFSk<3.5) then
2*VDC2k/3 else 0
  vbsk4=if (vecTFSk>3.5 and vecTFSk<4.5) then
VDC2k/3 else 0
  vbsk5=if (vecTFSk>4.5 and vecTFSk<5.5) then -
VDC2k/3 else 0
  vbsk6=if (vecTFSk>5.5 and vecTFSk<6.5) then -
2*VDC2k/3 else 0
  vbsk2=vbsk1+vbsk2+vbsk3+vbsk4+vbsk5+vbsk6
  vcsk1=if (vecTFSk>0.5 and vecTFSk<1.5) then -
VDC2k/3 else 0
  vcsk2=if (vecTFSk>1.5 and vecTFSk<2.5) then -
2*VDC2k/3 else 0
  vcsk3=if (vecTFSk>2.5 and vecTFSk<3.5) then -
VDC2k/3 else 0
  vcsk4=if (vecTFSk>3.5 and vecTFSk<4.5) then
VDC2k/3 else 0
  vcsk5=if (vecTFSk>4.5 and vecTFSk<5.5) then
2*VDC2k/3 else 0
  vcsk6=if (vecTFSk>5.5 and vecTFSk<6.5) then

```



```

VDC2k/3 else 0
  vck2=vcsk1+vcsk2+vcsk3+vcsk4+vcsk5+vcsk6
  " Parameter values:
Bflux=0.001
Btorque=0.001
pi=3.1416
Vdc=300
P=3
Ppm=0.2547
J=14.1e-4
R=0.86
Lq=0.00665
Ld=0.00665
TL=0
Tf=0
B=0 "0.00038
Kt=0.272
Kpw=10*0.79"0.1*0.79
Kiw=342.47*0.01"342.47*0.01
END

```

Modelo del controlador ContDTCK.T

```

DISCRETE SYSTEM ContDTCK
INPUT iakk ibkk ickk wrpmk
OUTPUT Vakk Vbkk Vcck
" States and time variables:
STATE fluxaka fluxbka secka fluxhka intwka taohka
zvalfaka zvbetaka zsecak
NEW fluxak fluxbk seck fluxhk intwk taohk
zvalfak zvbetak zseck
TIME t
TSAMP ts
ts=t+h
h=1e-5
"Envia y recibí Variables
Vakk=vak2
Vbkk=vbk2
Vcck=vck2
iak=iakk
ibk=ibkk
ick=ickk
wrpmk=wrpmk
"ialfak=(iak)
"ibetak=(ibk-ick)/sqrt(3)
ialfak=(2/3)*(iak-ibk/2-ick/2)
ibetak=(2/3)*(sqrt(3)/2)*(ibk-ick)
iok=iak+ibk+ick
"Calculo del flujo en alfa y beta
fluxak=fluxaka+(valfak-R*ialfak)*h
fluxbk=fluxbka+(vbetak-R*ibetak)*h
fluxaka:0.2547
fluxsmk=sqrt(fluxak*fluxak+fluxbk*fluxbk)
"Magnitud del flujo
fluxsak=atan2(fluxbk,fluxak)
"Ángulo de fase del flujo
fluxsakg=180*fluxsak/3.1416
" Initializations:
"Determina el Sector
tplsek=sign(fluxak)
tplasek=fluxak*tan(pi/6)
tplbsek=sign(fluxbk)
tplabsek=abs(fluxbk)>abs(tplasek) "Compara si
la magnitud de b es mayor que a
"Con flujo en a positivo
tp2sek= if ((not tplsek) and (not tplbsek) and
(not tplabsek) ) then 4 else 0
tp3sek= if ((not tplsek) and (not tplbsek) and
tplabsek ) then 5 else 0
tp4sek= if ((not tplsek) and tplbsek and (not
tplabsek) ) then 4 else 0
tp5sek= if ((not tplsek) and tplbsek and

```

```

tplabsek ) then 3 else 0
  tp6sek= if (tplsek and (not tplbsek) and (not
tplabsek) ) then 1 else 0
  tp7sek= if (tplsek and (not tplbsek) and
tplabsek ) then 6 else 0
  tp8sek= if (tplsek and tplbsek and (not
tplabsek) ) then 1 else 0
  tp9sek= if (tplsek and tplbsek and tplabsek )
then 2 else 0
  "Con flujo en a negativo
tp10sek=tp2sek+tp3sek+tp4sek+tp5sek+tp6sek+tp7sek+
tp8sek+tp9sek
  sectork=tp10sek
  seck=sectork
  secka:1
"CONTROLADOR DE VELOCIDAD Y FLUJO
ulk=if ts<0.2 then 0 else 1000
u2k=if ts<0.4 then 0 else -2000
u3k=if ts<0.6 then 0 else 2000
u4k=if ts<1.1 then 0 else 0

```

```

Wdesk=ulk+u2k+u3k+u4k "Velocidad deseada en
RPM

```

```

ewk=(Wdesk-Wrpmk)*P*pi/30
intwk=intwka + h*ewk
tpiwl=Kpw*ewk + Kiw*intwk
tpiw2= if tpiwl > 15 then 15 else tpiwl
tpiw3= if tpiw2 < -15 then -15 else tpiw2
cntwk=tpiw3
Tek=(3*P/2)*(fluxak*ibetak-fluxbk*ialfak)

```

```

etk=cntwk-Tek
efk=Ppm-fluxsmk

```

```

"IMPLEMENTA LA HISTERESIS, TABLA DE CONMUTACION Y
VECTOR

```

```

"Implementación de la histeresis para el flujo y
torque

```

```

temp1fk= if efk > Bflux then 1 else 0
temp2fk= if efk < -Bflux then 1 else 0
temp3fk= if ( (efk> -Bflux) and (efk < Bflux) )
then fluxhka else 0
fluxhk=temp1fk + ((not temp2fk) and temp3fk)
stork=etk
temp1tk= if stork > Btorque then 1 else 0
temp2tk= if stork < -Btorque then 1 else 0
temp3tk= if ((stork > -Btorque) and (stork <
Btorque)) then taohka else 0
taohk=temp1tk + ((not temp2tk) and temp3tk)

```

```

"Identificación del sector
temp1s=sectork>0 and sectork<1.5
temp2s=sectork>1.5 and sectork<2.5
temp3s=sectork>2.5 and sectork<3.5
temp4s=sectork>3.5 and sectork<4.5
temp5s=sectork>4.5 and sectork<5.5
temp6s=sectork>5.5 and sectork<6.5

```

```

"Vector para flujo 1 torque 1 y sector 1-6
v11k=if (temp1s) then 2 else 0
v112k=if (temp2s) then 3 else 0
v113k=if (temp3s) then 4 else 0
v114k=if (temp4s) then 5 else 0
v115k=if (temp5s) then 6 else 0
v116k=if (temp6s) then 1 else 0

```

```

vft11k=(v11k+v112k+v113k+v114k+v115k+v116k)*(taoh
k and fluxhk)

```

```

"Vector para flujo 1 torque 0 y sector x
v101k=if (temp1s) then 6 else 0
v102k=if (temp2s) then 1 else 0
v103k=if (temp3s) then 2 else 0

```

```

v104k=if (temp4s) then 3 else 0
v105k=if (temp5s) then 4 else 0
v106k=if (temp6s) then 5 else 0

vft10k=(v101k+v102k+v103k+v104k+v105k+v106k)*((not
taohk) and fluxhk)
"Vector para flujo 0 torque 1 y sector x
v011k=if (temp1s) then 3 else 0
v012k=if (temp2s) then 4 else 0
v013k=if (temp3s) then 5 else 0
v014k=if (temp4s) then 6 else 0
v015k=if (temp5s) then 1 else 0
v016k=if (temp6s) then 2 else 0

vft01k=(v011k+v012k+v013k+v014k+v015k+v016k)*(taoh
k and (not fluxhk))

"Vector para flujo 0 torque 0 y sector x
v001k=if (temp1s) then 5 else 0
v002k=if (temp2s) then 6 else 0
v003k=if (temp3s) then 1 else 0
v004k=if (temp4s) then 2 else 0
v005k=if (temp5s) then 3 else 0
v006k=if (temp6s) then 4 else 0

vft00k=(v001k+v002k+v003k+v004k+v005k+v006k)*((not
taohk) and (not fluxhk))
"VECTOR A USAR
zseck=vft11k + vft10k + vft01k + vft00k
vecTFSk=zseck
"GENERA EL INVERSOR Y LOS VOLTAJES
VDCk=150*2/3
VDC2k=150
tetainvk=(vecTFSk-1)*60*3.1416/180
valfk=(VDCk)*cos(tetainvk)
vbetk=(VDCk)*sin(tetainvk)
zvalfak=valfk
zvbetak=vbetk
zvalfaka:0
valfak=zvalfaka
vbetak=zvbetaka

"Genera los voltajes de linea discretos
vask1=if (vecTFSk>0.5 and vecTFSk<1.5) then
2*VDC2k/3 else 0
vask2=if (vecTFSk>1.5 and vecTFSk<2.5) then
VDC2k/3 else 0
vask3=if (vecTFSk>2.5 and vecTFSk<3.5) then -
VDC2k/3 else 0
vask4=if (vecTFSk>3.5 and vecTFSk<4.5) then -
2*VDC2k/3 else 0
vask5=if (vecTFSk>4.5 and vecTFSk<5.5) then -
VDC2k/3 else 0
vask6=if (vecTFSk>5.5 and vecTFSk<6.5) then
VDC2k/3 else 0
vak2=vask1+vask2+vask3+vask4+vask5+vask6
vbsk1=if (vecTFSk>0.5 and vecTFSk<1.5) then -
VDC2k/3 else 0
vbsk2=if (vecTFSk>1.5 and vecTFSk<2.5) then
VDC2k/3 else 0
vbsk3=if (vecTFSk>2.5 and vecTFSk<3.5) then
2*VDC2k/3 else 0
vbsk4=if (vecTFSk>3.5 and vecTFSk<4.5) then
VDC2k/3 else 0
vbsk5=if (vecTFSk>4.5 and vecTFSk<5.5) then -
VDC2k/3 else 0
vbsk6=if (vecTFSk>5.5 and vecTFSk<6.5) then -
2*VDC2k/3 else 0
vbk2=vbsk1+vbsk2+vbsk3+vbsk4+vbsk5+vbsk6
vcsk1=if (vecTFSk>0.5 and vecTFSk<1.5) then -
VDC2k/3 else 0
vcsk2=if (vecTFSk>1.5 and vecTFSk<2.5) then -

```

```

2*VDC2k/3 else 0
vcsk3=if (vecTFSk>2.5 and vecTFSk<3.5) then -
VDC2k/3 else 0
vcsk4=if (vecTFSk>3.5 and vecTFSk<4.5) then
VDC2k/3 else 0
vcsk5=if (vecTFSk>4.5 and vecTFSk<5.5) then
2*VDC2k/3 else 0
vcsk6=if (vecTFSk>5.5 and vecTFSk<6.5) then
VDC2k/3 else 0
vck2=vcsk1+vcsk2+vcsk3+vcsk4+vcsk5+vcsk6

" Parameter values:
Bflux=0.001
Btorque=0.001
pi=3.1416
Vdc=300
P=3
Ppm=0.2547
J=14.1e-4
R=0.86
Lq=0.00665
Ld=0.00665
TL=0
Tf=0
B=0 "0.00038
Kt=0.272
Kpw=10*0.79"0.1*0.79
Kiw=342.47*0.01"342.47*0.01
END

```

Archivo de Interconexion Conexion.T

```

CONNECTING SYSTEM Conexion
"Archivo de interconexion
"Created: 13/03/2011
" Time, if needed:
TIME t
Vakk[ModPMSM]=Vakk[ContDTCK]
Vbkk[ModPMSM]=Vbkk[ContDTCK]
Vcck[ModPMSM]=Vcck[ContDTCK]
iakk[ContDTCK]=iakk[ModPMSM]
ibkk[ContDTCK]=ibkk[ModPMSM]
ickk[ContDTCK]=ickk[ModPMSM]
wrpmkk[ContDTCK]=wrpmkk[ModPMSM]
" Connections:
END

```

Macro Macro1.T

```

MACRO Macro1
"Macro para la simulacion del contron DTC con el
PMSM
"Created: 13/03/2011
" Enter commands here:
syst ModPMSM ContDTCK Conexion
store Tek Te seck fluxbk wrpm vecTFSk fluxhk
taohk valfak vbetak ewk Wdesk
simu 0 1 1e-5
split 3 1
ashow wrpm wdesk
text 'wrpm'
ashow ewk
text 'ewk'
ashow Tek
text 'thk'
export store < store /0
END

```

Electronic Ballast Based on a Series Active Input Current Shaper

*Claudia V. Hernández Gutiérrez, *Eugenio Camargo Trigueros, *Nimrod Vázquez Nava,
*Elías Rodríguez Segura, **René Osorio Sánchez

*Instituto Tecnológico de Celaya, Av. Tecnológico y A. G. Cubas s/n, Tel: (461) 611 7575, n.vazquez@ieee.org

**UAM Azcapotzalco, Av. San Pablo 180, Col. Reynosa Tamaulipas, Del. Azcapotzalco, México D. F., Tel:
(55) 5318 9000, reneosorios@yahoo.com.mx.

Abstract

Great part of the energy generated in the industrialized world is consumed by lighting, and to reduce the energy consumption fluorescent lamp are used; to operate a fluorescent lamp properly an electronic ballast is required. Different techniques to improve the characteristics of the topologies of electronic ballast are reported in literature; to fill the standard specifications at low cost, one option is to use the active input current technique. In this paper an electronic ballast based on a series active input current shaper is discussed. Design and experimental results are presented.

Palabras Clave: AICS, electronic ballast, low cost.

I. Introduction

There are several electronic ballasts reported in literature for fluorescent lamps, the traditional approach consists of two stages additional to the rectifier; the first one is used to demand a current in phase with the input voltage to assure a high power factor, the second one is dedicated to operate the lamp properly [1]; this option offers good characteristics, however also some disadvantages: low efficiency and high cost because two converters and two controllers are used.

In [2] is presented a topology with high power factor, two stages are integrated; the input inductor is operated in discontinuous conduction mode (DCM) therefore a high power factor is achieved, but a high semiconductor stress and low efficiency are obtained. It is found in literature [2]-[8] single stage topologies; however these are actually integrated topologies. For example a flyback converter integrated with a push pull converter is presented in reference [2]. The flyback converter is operated in discontinuous conduction mode in order to assure a good power factor, and then efficiency is penalized significantly.

In [9] a single stage topology based on charge pump power factor correction, certainly just one

controlled semiconductor is used; however a lot of passive components and it have a complex operation and design. In [10] a two stage topology, but one controller is proposed; a modified valley fill converter plus an inverter is analyzed, this permit to reduce the cost because one controller is eliminated.

An alternative to reduce the cost is the use of the active input current shaper (AICS) technique; this is because traditionally one controller is eliminated from the traditional method; however the unity power factor cannot be achieved, but only the standard can be fulfilled.

An electronic ballast topology based on AICS is presented in [11]. In Fig. 1 is shown the topology, it uses a separate switch with its respective controller, the flyback converter is operated in DCM.

In this paper a different approach for electronic ballast based on the AICS technique is presented, the converter is operated partially in DCM, and some angle is operated in continuous conduction mode (Fig. 2); at the end the operation could be similar to the converters operated at the boundary conduction mode, however no additional controller is required and it has constant switching frequency. A high power factor is achieved. The operation of the converter is presented, also experimental results.

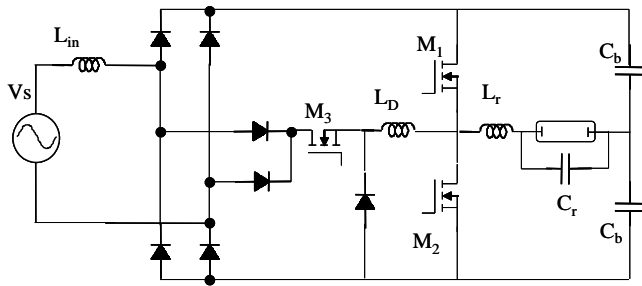


Fig. 1. Electronic ballast based on AICS

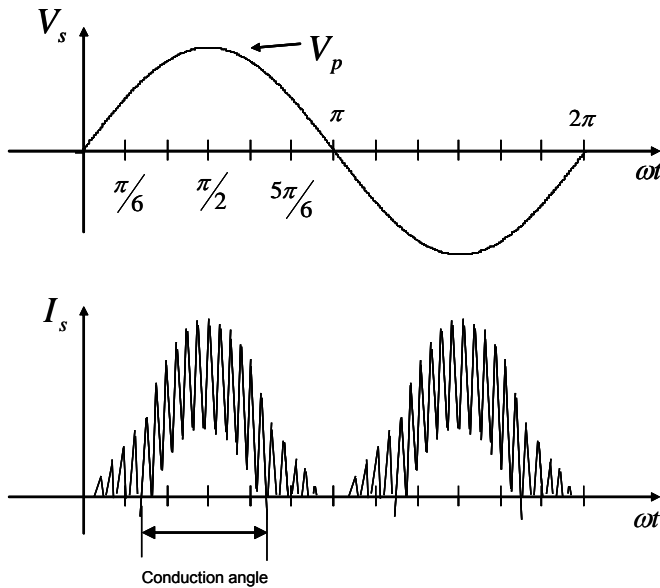


Fig. 2. Waveforms of the converter.

II. Proposed Converter

The proposed converter is shown in Fig. 3. It is composed of a shaping cell and the half bridge resonant inverter: The inverter controls the start up and the output power of the lamp, the shaping cell permits to fulfill the standard specifications about harmonic content at the input current.

The inductor L_{in} is operating in continuous conduction mode during some time (conduction angle) and the rest in discontinuous operation mode, the inductor L_D permits to shape the input current and determines the change between the two operating modes. The waveform of the converter input current is shown in Fig. 2, as it can be observed the inductor current has an evolution near to a sinusoidal current, therefore the standard is satisfied. The control signals M_1 and M_2 are 180° out of phase to produce a square

output voltage; the control signal M_3 is equal to M_2 , then no additional controller is required for the shaping cell.

In Figs. 4 and 5 are shown the waveforms for each operating region of the converter, near to the zero crossing the converter operated completely in discontinuous conduction mode, but after the converter operates in continuous conduction mode (conduction angle). In Fig. 4 is illustrated the operation in DCM and in the Fig. 5 is shown the operation in CCM.

Independently of the operation mode of the input inductor the lamp always is feed with a square output voltage at high frequency, the semiconductors M_1 and M_2 are complementary.

To illustrate the behavior of the system, only the positive semicycle of the input voltage is analyzed, since the operation for the other semicycle is similar. When the converter is working in DCM operates as follows:

- During t_0-t_1 : M_2 and M_3 are on and M_1 is off, then the inductors L_{in} and L_D are being charged in a series configuration. The equivalent subcircuit is shown in Fig. 6.
- During t_1-t_2 : M_2 and M_3 are turned off and M_1 is on, then the inductors L_{in} and L_D are being discharged, but their slope are different, due to the conduction paths are different. One is with bulky capacitors (L_D) and the other with the bulky capacitor minus the input voltage (L_{in}). The subcircuit is illustrated in Fig. 7.
- During t_2-t_3 : M_2 and M_3 are turned off and M_1 is on, the inductor L_{in} is still being discharged, but L_D is completely discharged. The equivalent subcircuit is shown in Fig. 8.

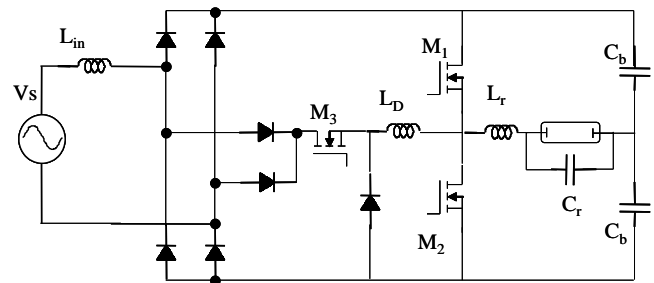


Fig. 3. Proposed electronic ballast.

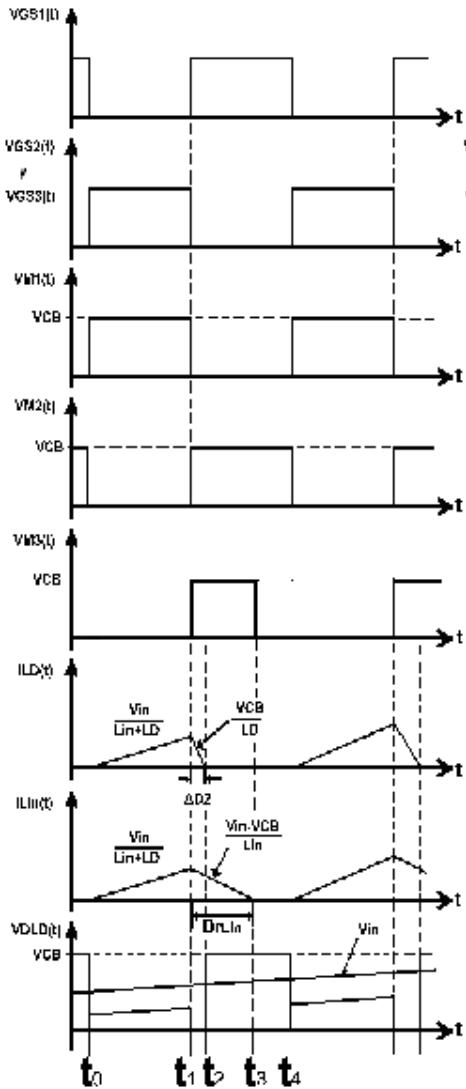


Fig. 4. Waveforms of the converter in DCM.

- During t_3-t_4 : M2 and M3 are turned off and M1 is on, the inductors L_{in} and L_D are completely discharged. The equivalent subcircuit is shown in Fig. 9.

When the converter is working in CCM operates as follows:

- During t_0-t_1 : M2 and M3 are on and M1 is off; then the inductors L_{in} and L_D are being charged, but with different slope, L_D was completely discharged. This interval ends when the current inductor L_D reaches the current inductor L_{in} . The equivalent subcircuit is shown in Fig. 10. This stage permits the operation of the shaping technique.

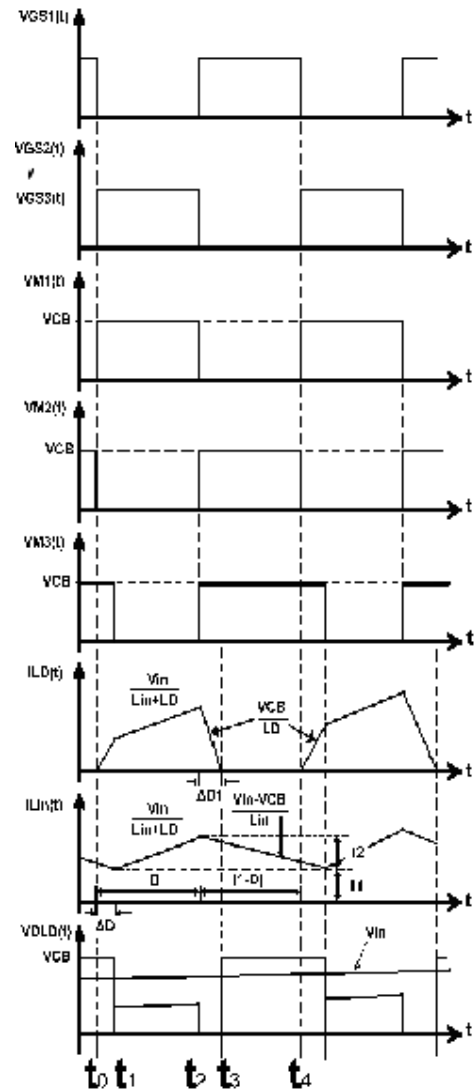


Fig. 5. Waveforms of the converter in CCM

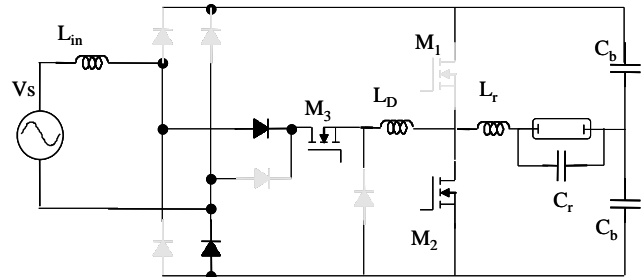


Fig. 6. Subcircuit when M2 and M3 are on and M1 is off, but L_{in} and L_D are in serie.

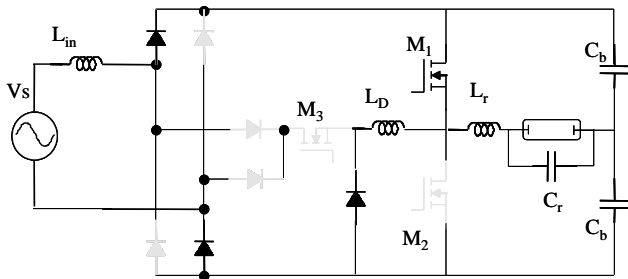


Fig. 7. Subcircuit when M2 and M3 are off and M1 is on, but Lin and LD are both charged.

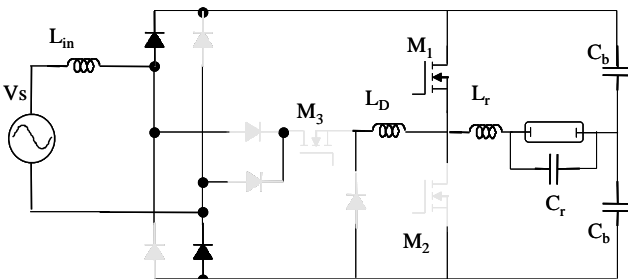


Fig. 8. Subcircuit when M2 and M3 are off and M1 is on, but Lin is charged and LD discharged.

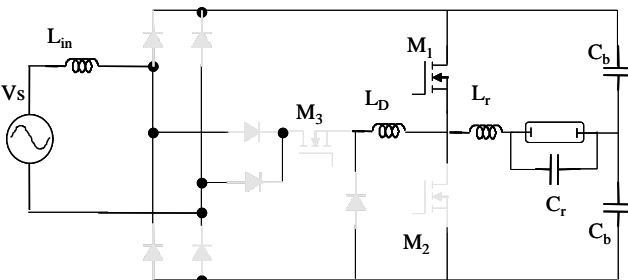


Fig. 9. Subcircuit when M2 and M3 are off and M1 is on, but Lin and LD are both discharged.

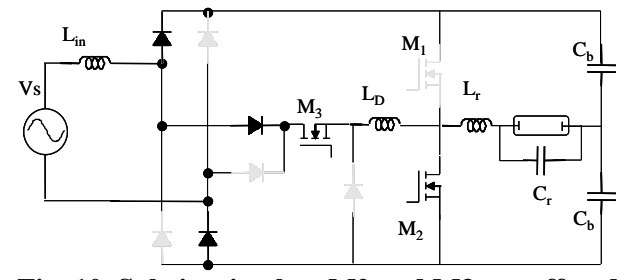


Fig. 10. Subcircuit when M2 and M3 are off and M1 is on, but Lin and LD are both charged, they have different values of current

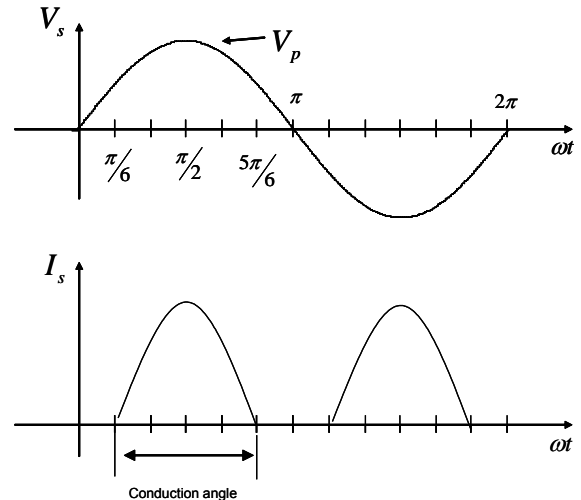


Fig. 11. Waveforms of the traditional AICS.

- During t1-t2: M2 and M3 are still on and M1 is still off, then the inductors Lin and LD are being charged in a series configuration. The subcircuit is illustrated in Fig. 6.
- During t2-t3: M2 and M3 are turned off and M1 is on, then the inductors Lin and LD are being discharged, but their slope are different, due to the conduction paths are different. This stage ends when the inductor LD is completely discharged. The subcircuit is illustrated in Fig. 7.
- During t3-t4: M2 and M3 are turned off and M1 is on, the inductor Lin is still being discharged, but LD has no energy. This stage ends when the semiconductors changes its operating conditions that are M2 and M3 are on and M1 is off. The equivalent subcircuit is shown in Fig. 8.

A. Converter design

The AICS are traditionally designed to operate with conduction angles higher than 120°, this is because the inductor is designed to operate during almost all the time in continuous conduction mode. The waveforms of this type of operation is shown in Fig. 11, as it can be observed the input inductor current ripple is neglected, and taking this into account only with conduction angles higher than 120° the standard specifications can be fulfilled [5-7].

When the conduction angle is reduced under 120° certainly the converter does not fill the standard, however if is reduced even more than

50° the standard is satisfied again; and this operation form has not been studied before. The reason why the converter fills again the standard is due to the approximation of the waveform to the discontinuous operation mode that in a natural way has a high power factor, in this form the waveform is more like the shown in Fig. 2, when the ripple of the input inductor current is relatively high.

The operation of the converter with conduction angles lower than 50° can be compared with the boundary control method (BCM); where the power factor correction is naturally achieved, with the advantage of lower semiconductor stress than the DCM operation, however in BCM the switching frequency is not constant and a specific controller is required. With the AICS technique the switching frequency is constant, it is not required a specific controller, and the semiconductor stresses are similar to the BCM.

The converter can not be operated with conduction angles higher than 120°, because a restriction was considered: the control signal of M_3 is exactly to M_2 , this is made to simplify the operation and implementation of the converter.

The conduction angle of the converter is given by:

$$\theta_c = 2 \cos^{-1} \left[\left(\frac{V_{CB}}{V_{inp}} \right) \cdot \left(\frac{\frac{L_{in}}{L_D} + 1}{2 \frac{L_{in}}{L_D} + 1} \right) \right] \quad (1)$$

Where: V_{CB} is the voltage of the dc side.
 V_{inp} is the peak of the input voltage
 L_{in} is the input inductor
 L_D is the delayed inductor

To assure a proper operation the conduction angle must be selected lower than 50°; it was selected of 45°; the input voltage selected is 120Vac, then V_{inp} is 170V approx.; the capacitor voltage is selected to 240V, the input inductor can be designed by conventional methods (2.8mH), then using (1) the delayed inductor is 1.5mH approximately.

If L_D is chosen higher then a lower conduction angle is obtained and the operation is approached

to the DCM, then a higher stress at the semiconductors is obtained. But if it is selected a lower L_D then a higher conduction angle is obtained and therefore the standard is not fulfilled.

III. Experimental Results

The converter was designed, built, and tested. The operation of the converter with a conduction angle lower than 50° is shown in Figs. 12 and 13. The input current and voltage, also the current of the inductor L_D are shown in Fig. 12; in Fig. 13 is shown the voltage and current of the lamp in steady state. The inductor are $L_{in} = 2.8\text{mH}$ and $L_D = 1.5\text{mH}$.

The operation of the converter with an angle between 50° and 120° is shown in Fig 14, as it can be observed the angle is increased, and then voltage and current are more similar to the Fig. 11. In this case L_{in} was increased for simplicity to 6mH and L_D was maintained to 1.5mH

In 15 is shown a comparison of the harmonic content of both cases with the standard, as it can be observed when the angle is lower than 50° the standard is fulfilled and the other case does not.

IV. Conclusions

In this paper is analyzed an electronic ballast based on the active input current shaper, the analyzed converter can fulfill the standard specifications if the conduction angle is lower than 50°. The converter was operated in a different way as the traditional AICS. The system offers a simple operation due to a single controller for the lamp is considered, the system fills the standard specifications by just controlling the conduction angle. The system was tested experimentally.

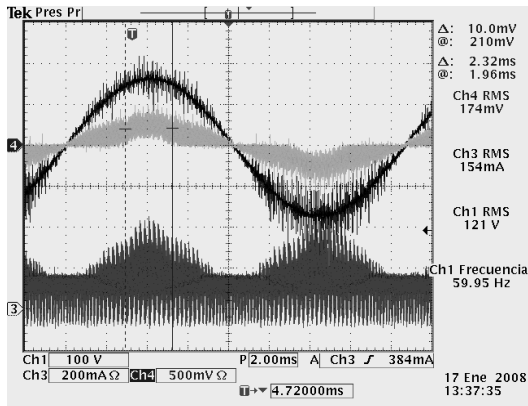


Fig. 12. Converter operation with conduction angle lower than 50o Top to down: Input voltage, input current and auxiliary inductor current

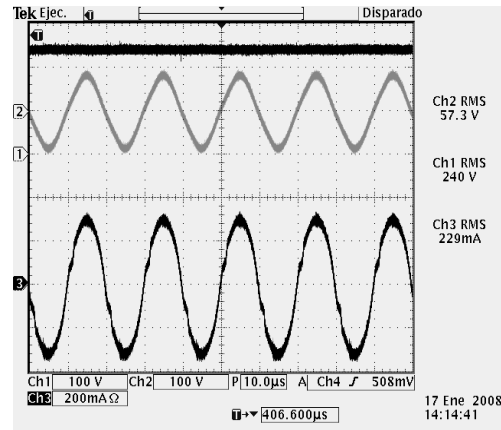


Fig. 13. Output characteristics. Top to down: Bulky capacitor voltage, lamp voltage, lamp current

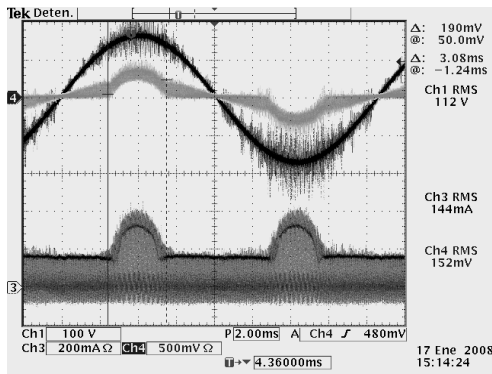


Fig. 14. Converter operation with conduction angle higher than 50o, but lower than 120o. Top to down: Input voltage, input current and auxiliary inductor current

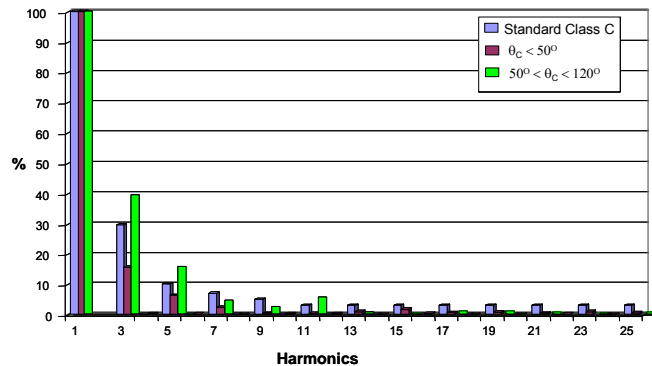


Fig. 15. Harmonic content under different conduction angles versus the standard.

V. References

- [1] S. Y. Ron Hui, L. M. Lee, H. S-H Chung, and Y. K. Ho. "An Electronic Ballast with Wide Dimming Range, High PF, and Low EMI". IEEE Transactions on Power Electronics, Vol. 16, No. 4, July 2001. pp 465-472
- [2] R. Nederson, S. Azzolin. "A High-Power-Factor Electronic Ballast Using a Flyback Push-Pull Integrated Converter". IEEE Transactions on Industrial Electronics, Vol. 46, No. 4, August 1999, pp 796-892.
- [3] M. Brumatti, M. Almeida, D.S. L. Simonetti, J. L. F. Vieira. "Single Stage Self-Oscillating HPF Electronic Ballast". IEEE Transactions on Industry Applications, Vol. 41, No. 3, May/June 2005, pp 735-741.
- [4] C.M. Wang. "A Novel Single-Stage High-Power-Factor Electronic Ballast With Symmetrical Half-Bridge Topology". IEEE Transactions on Industrial Electronics, Vol. 55, No. 2, February 2008, pp 969-972
- [5] J.A. Vilela, A. Rodrigues, V.J. Farias, L.C. de Freitas, E.A. Alves, J. Batista. "An Electronic Ballast With High Power Factor and Low Voltage Stress". IEEE

Transactions on Industry Applications, Vol. 41, No. 4, July/August 2005, pp 917-926

- [6] C.M. Wang. “A Novel Single-Switch Single-Stage Electronic Ballast With High Input Power Factor”. IEEE Transactions on Power Electronics, Vol. 22, No. 3, May 2007, pp 797-803
- [7] Y.C. Chuang and H.L. Cheng. “Single-Stage Single-Switch High-Power-Factor Electronic Ballast for Fluorescent Lamps”. IEEE Transactions on Industry Applications, Vol. 43, No. 6, November/December 2007, pp 1434-1440.
- [8] C.S. Moo, K.H. Lee, H.L. Cheng, W.M. Chen. “A Single-Stage High-Power-Factor Electronic Ballast With ZVS Buck–Boost Conversion”. IEEE Transactions on Industrial Electronics, Vol. 56, No. 4, April 2009, pp 1136-1146
- [9] Chien-Ming Wang , “A Novel Single-Stage High-Power-Factor Electronic Ballast With Symmetrical Half-Bridge Topology” IEEE Transactions on industrial electronics, Vol. 55, No. 2, February 2008, pp 969-972
- [10] C.S. Lin, C.L. Chen. “Single-Switch Electronic Ballast with Continuous Input Current Charge Pump Power-Factor Correction”. IEEE Transactions on industrial electronics, Vol. 47, No. 6, December 2000, pp 1263-1270.
- [11] J.C. Wo Lam, P. K. Jain. “A Modified Valley Fill Electronic Ballast Having a Current Source Resonant Inverter With Improved Line-Current Total Harmonic Distortion (THD), High Power Factor, and Low Lamp Crest Factor”. IEEE Transactions on industrial electronics, Vol. 55, No. 3, March 2008, pp 1147-1149
- [12] J.C. Wo Lam, P. K. Jain. “A Modified Valley Fill Electronic Ballast Having a Current Source Resonant Inverter With Improved Line-Current Total Harmonic Distortion (THD), High Power Factor, and Low Lamp Crest Factor”. IEEE

Transactions on industrial electronics, Vol. 55, No. 3, March 2008, pp 1147-1149

VI. Authors

M.C. Claudia V Hernández Gutiérrez obtuvo el título de Ingeniero en Electrónica por el Instituto Tecnológico de Celaya en 1995 y el grado de Maestra en Ciencias en Ingeniería Electrónica por el cenidet en 2000, actualmente es docente y Coordinadora de Posgrado en el Instituto Tecnológico de Celaya.

M. C. Eugenio Camargo Trigueros obtuvo el título de Ingeniero Electrónico y el grado de Maestro en Ciencias en Ingeniería Electrónica por el Instituto Tecnológico de Celaya, actualmente trabaja en la PROFECO.

Dr. Nimrod Vazquez Nava obtuvo el título de Ingeniero en Electrónica por el Instituto Tecnológico de Celaya en 1994, los grados de Maestro en Ciencias en Ingeniería Electrónica y Doctor en Ciencias en Ingeniería Electrónica por el cenidet en 1997 y 2003 respectivamente, actualmente es profesor-investigador en el Instituto Tecnológico de Celaya e investigador nivel 1 en el SNI.

Dr. Elías Rodríguez Segura obtuvo el título de Ingeniero en Electrónica por la Universidad Autónoma Metropolitana en 1994, los grados de Maestro en Ciencias en Ingeniería Electrónica y Doctor en Ciencias en Ingeniería Electrónica por el cenidet en 1997 y 2003 respectivamente, actualmente es profesor-investigador en el Instituto Tecnológico de Celaya e investigador nivel 1 en el SNI.

Dr. René Osorio Sánchez obtuvo el título de Ingeniero en Electrónica por la Universidad Autónoma de Campeche en 1999, los grados de Maestro en Ciencias en Ingeniería Electrónica y Doctor en Ciencias en Ingeniería Electrónica por el cenidet en 2001 y 2007 respectivamente, actualmente es profesor-investigador en la Universidad Autónoma Metropolitana y candidato a investigador en el SNI.

Convertidor Push-Pull Alimentado en Corriente para Baja Tensión de Salida

Luis Alejandro Flores Oropeza, Alejandro Román Loera

Universidad Autónoma de Aguascalientes, Av. Universidad # 940 Cd. Universitaria, Centro de Ciencias Básicas,
Departamento de Sistemas Electrónicos, +52 449 9 10 84 22 {lflores, aroman}@correo.uaa.mx

Resumen

Este artículo analiza la aplicabilidad del convertidor Push-Pull alimentado en corriente con rectificación síncrona para alimentar cargas de baja tensión y alta corriente. La ventaja principal de ésta topología es que la rectificación síncrona puede ser fácilmente implementada incluso para las tensiones de salida muy bajas (3.3V, 1V, etc.). Se construye y se prueba un prototipo experimental de laboratorio de 66W, 3.3V. Se presenta el principio de operación y los resultados experimentales.

Palabras Clave: acoplamiento, baja tensión, rectificación síncrona.

I. Introducción

En los últimos años, la fuente de alimentación de los sistemas de telecomunicación ha cambiado. Ahora se utilizan sistemas de alimentación distribuidos en vez de sistemas centralizados. Para reducir al mínimo tamaño los sistemas de telecomunicación, las fuentes de alimentación se colocan directamente en las tarjetas electrónicas (Convertidores de tarjeta). Esta situación impone especificaciones más exigentes a los convertidores de energía: alta densidad de energía, alto rendimiento y tamaño muy reducido. Al mismo tiempo, las cargas electrónicas requieren voltajes de alimentación cada vez más bajos (3.3V, 1.5V, 1V...) y corrientes más elevadas (20A, 30A, 50A...) haciendo más complejo el diseño de los Convertidores de tarjeta.

Diversas soluciones se han desarrollado para satisfacer los requisitos anteriores [1-5], y todas ellas tienen implementadas técnicas de Rectificación Síncrona para alcanzar altos rendimientos, ya que del 40% al 50% de las pérdidas en los Convertidores de tarjetas son causadas por la etapa de rectificación [2].

Las topologías alimentadas en tensión que hacen conducir simétricamente al transformador (medio puente, Push-Pull, etc.) son ampliamente

utilizadas para estas aplicaciones. Aplicar técnicas de Rectificación Síncrona Autoexcitada para estas topologías es complejo puesto que los rectificadores síncronos deben conducir durante los tiempos muertos del transformador cuando los voltajes son cero en el transformador. Varias técnicas se han propuesto para solucionar este problema.

La ventaja principal de la topología propuesta en éste artículo, es que se puede implementar fácilmente Rectificación Síncrona Autoexcitada únicamente con la adición de dos devanados auxiliares acoplados a los devanados del transformador. En ésta topología, los rectificadores no conducen durante los tiempos muertos del transformador, haciendo más fácil controlarlos.

Las ventajas y desventajas de esta topología se estudian en éste artículo para analizar su aplicabilidad en la alimentación de cargas electrónicas de baja tensión.

II. Convertidor Push-Pull Alimentado en Corriente

La Fig. 1 muestra el convertidor Push-Pull Alimentado en Corriente con Rectificación Síncrona que se propone. El principio de operación de esta topología es exactamente igual

que con diodos de rectificación puesto que la influencia de los rectificadores no cambia el diseño. Esta topología necesita dos devanados auxiliares (*Aux1*, *Aux2*) para controlar los rectificadores de la salida (*SR1*, *SR2*) puesto que no es posible obtener rectificación síncrona con la señal de voltaje en el lado secundario del transformador.

La Fig. 2 muestra las formas de onda principales de este convertidor. Hay dos señales de puerta para los MOSFETs primarios (V_{GS1} y V_{GS2}) para controlar *S1* y *S2*, estas señales de puerta tienen un ciclo de trabajo mayor o igual al 50% y están desfasadas 180°.

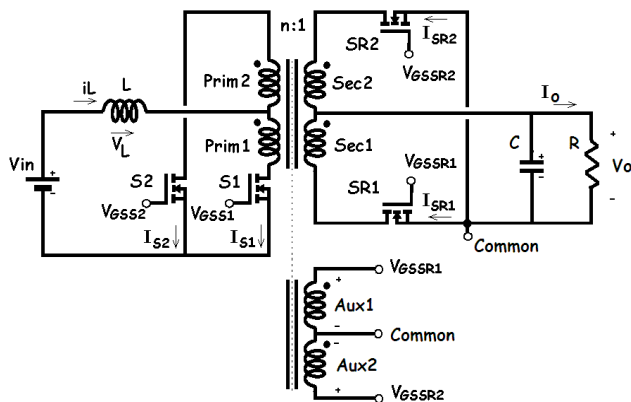


Fig. 1. Convertidor Push-Pull alimentado en corriente con devanados auxiliares (*Aux1*, *Aux2*) y rectificación síncrona

Modo de Operación

Se pueden identificar básicamente cuatro modos de operación en un período de conmutación, la Fig. 2 muestra los estados del convertidor, y los modos de operación se definen a continuación:

$t_0 - t_1$.- En éste intervalo *S1* y *S2* se encienden y la corriente en la bobina aumenta linealmente, ambos devanados en el primario del transformador comparten la corriente de la bobina e imponen cero voltios en el primario del transformador. En éste intervalo, la carga esta alimentada por la energía previamente almacenada en el condensador *C*. Los rectificadores síncronos están apagados (*SR1*,

SR2), puesto que el voltaje cero aparece en sus correspondientes señales de puerta.

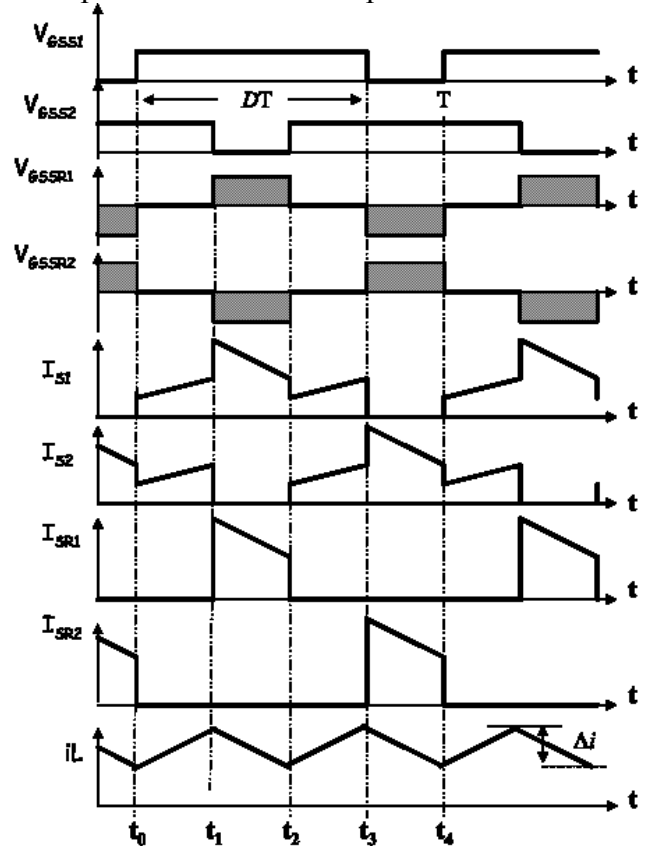


Fig. 2. Formas de onda principales del convertidor propuesto

t_1-t_2 .- *S2* está apagado, causando que la corriente en la bobina disminuya linealmente y transfiera la energía a la carga a través del transformador con *Prim1*, *Sec1* y *SR1*. En éste intervalo, la corriente total fluye a través de *S1* y el condensador de salida toma energía del primario del convertidor. Al mismo tiempo un voltaje positivo aparece en la señal de puerta de *SR1* para evitar que la corriente de salida fluya a través del diodo parásito del MOSFET rectificador. Con esto se consigue un mejor funcionamiento de operación puesto que las pérdidas de conducción en el rectificador son disminuidas.

$t_2 - t_3$.- En éste intervalo, **S2** se enciende otra vez y el convertidor funciona como el intervalo ($t_0 - t_1$), incrementando linealmente la corriente en la bobina.

$t_3 - t_4$.- En éste intervalo **S1** se apaga, causando que la energía sea transferida hacia la carga a través de **Prim2**, de **Sec2** y del rectificador síncrono **SR2**. En éste intervalo, V_{GSSR2} es positivo para permitir que la corriente no fluya a través del diodo parásito en **SR2**. En éste intervalo, la corriente total de la bobina fluye a través de **S2**. La corriente magnetizante no es considerada en este análisis, debido a que es insignificante comparada con la corriente de carga. Éstas son todas las etapas en un período de conmutación y se repiten cada periodo de conmutación.

Múltiples voltajes de salida

La Fig. 3 muestra la estructura de la topología para implementar múltiples voltajes de salida. Puesto que la topología utiliza devanados auxiliares para implementar la rectificación síncrona, resulta fácil construir éste convertidor para proveer múltiples voltajes salida. La adición de nuevos devanados secundarios es suficiente para proveer distintos voltajes de salida. La utilización de devanados auxiliares hace de este convertidor un buen candidato para proporcionar voltajes bajos de salida como 3.3V, 1.5V, 1V o menores, puesto que el control del voltaje de los rectificadores síncronos es independiente del voltaje de la salida.

III. Diseño del convertidor

Para verificar el principio de operación del convertidor propuesto, un prototipo de laboratorio de 66W fue construido y probado. Las especificaciones del prototipo se muestran en la Tabla 1, el análisis y consideraciones de diseño se muestran en ésta sección.

Voltage de entrada (V_{in})	38V ~ 58V
Voltage de salida(V_o)	3.3V
Potencia de salida (P_o)	66W
Frecuencia de conmutación (f_s)	100kHz

Tabla 1. Especificaciones del prototipo

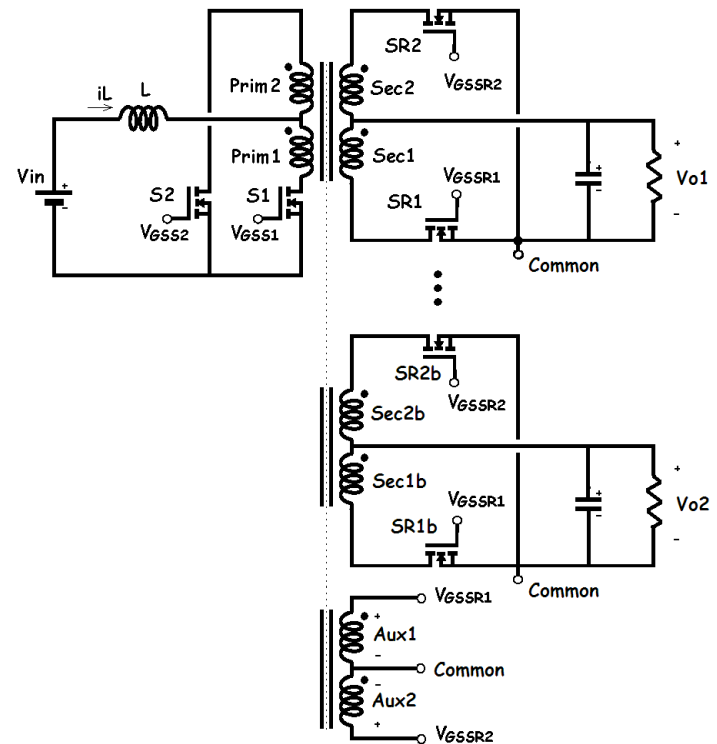


Fig. 3. Estructura para implementar múltiples voltajes de salida

Consideraciones de diseño

Para el diseño adecuado del convertidor, es necesario conocer voltajes y corrientes a través de los MOSFETs, rectificadores y de la bobina de entrada. El diseño de éste convertidor comienza con la corriente promedio a través de la bobina de entrada. La ecuación (1) define ésta corriente.

$$\hat{i}_L = \frac{P_o}{V_{in}} \quad (1)$$

Al mismo tiempo, el rizado de corriente (Δi) está definido por el voltaje aplicado a la bobina de entrada.

$$V_L = L \frac{di}{dt}$$

La ecuación (2) define el rizado de corriente en la bobina:

$$\Delta i = \frac{V_{in}}{L} (D - 0.5)T \quad (2)$$

La ecuación (3) define el ciclo de trabajo D del convertidor propuesto en modo de conducción continuo (MCC). Esta ecuación se obtiene como una función de la tensión de entrada (V_{in}), de la tensión de salida (V_o) y de la relación de transformación n aplicando balance voltios-segundos en la bobina L del convertidor en un período de conmutación.

$$D = 1 - \frac{V_{in}}{2nV_o} \quad (3)$$

La ecuación (4) define el valor de la bobina de entrada, donde su valor depende de (V_{in}), ciclo de trabajo D , frecuencia de conmutación (fs) y el rizado de corriente (Δi). El valor de la inductancia en la bobina es de la mitad comparado con el valor de las bobinas de otras topologías con la misma frecuencia de conmutación.

$$L = \frac{V_{in} (D - 0.5)}{\Delta i fs} \quad (4)$$

En ésta topología, el condensador de salida C es uno de los elementos principales que almacenan y entregan energía. Este dispositivo como el condensador de un convertidor Flyback, entrega toda la energía a la salida mientras que ambos interruptores principales ($S1$ y $S2$) están encendidos. La ecuación (5) se obtiene aplicando el balance de energía en el condensador en un período de conmutación; esta ecuación define el condensador de salida como función del ciclo de trabajo (D), voltaje de la salida (V_o), potencia de

salida (P_o), frecuencia de conmutación (fs) y el rizado de voltaje (ΔV).

$$C = \frac{P_o}{\Delta V} \frac{(D - 0.5)}{V_o fs} \quad (5)$$

Para completar el diseño del convertidor es conveniente conocer otras ecuaciones que ayuden a seleccionar adecuadamente los MOSFETs principales y los rectificadores síncronos. Estas ecuaciones son para todos los componentes del convertidor, voltajes de estrés y corrientes promedio y eficaces. La relación de transformación n se selecciona de acuerdo al voltaje de entrada, voltaje de salida y ciclo de trabajo para trabajar con todo el rango de la tensión de entrada.

Cálculo de estreses de voltaje y corrientes

MOSFETs de baja resistencia en conducción se deben seleccionar para obtener las pérdidas más bajas de conducción en los rectificadores ($SR1$, $SR2$). La selección de estos rectificadores depende solamente del voltaje que aparece en los terminales drenador fuente del MOSFET (V_{DSSR1} , V_{DSSR2}). Los MOSFETs con menor tensión drenador fuente tienen menor resistencia en conducción. De acuerdo con la topología, el voltaje que aparece en estos rectificadores síncronos es el doble del voltaje de salida. La ecuación (6) define la tensión drenador fuente que soportan a estos rectificadores.

$$V_{DSSR1} = 2V_o \quad (6)$$

La ecuación (7) define la tensión que soportan los MOSFETs del primario. Al igual que los rectificadores síncronos, consideraciones similares se deben hacer para seleccionar los MOSFETs principales ($S1$, $S2$). La tensión drenador fuente depende directamente de la relación de transformación n , y dos veces la tensión de salida.

$$V_{DSS1} = 2nV_o \quad (7)$$

La Fig. 4 muestra las formas de onda de la corriente a través de los MOSFETs del primario y la bobina de entrada. Obteniendo estas dos corrientes, resulta fácil calcular el resto de corrientes en el convertidor. Cualquier hoja matemática de cálculo se puede utilizar para calcular las corrientes en todo el rango de voltajes de entrada y de potencia de salida del convertidor.

De la Fig. 4, la corriente de pico en la bobina de entrada:

$$i_{pk} = \hat{I}_L + \frac{\Delta i}{2} \quad (8)$$

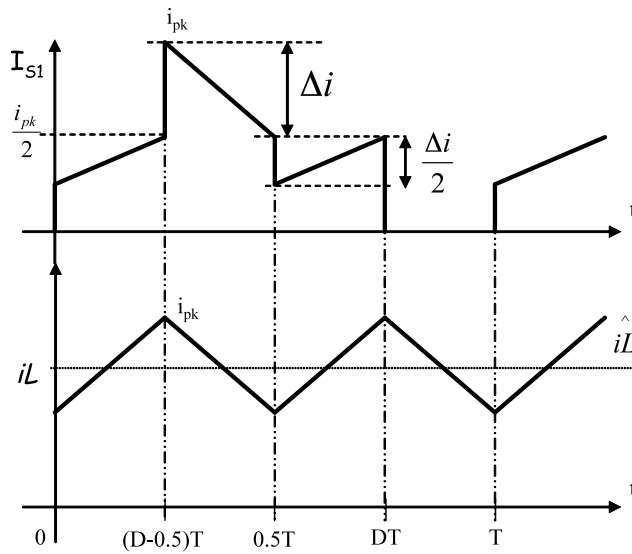


Fig. 4. Formas de onda principales para calcular corrientes promedio y eficaces en el convertidor.

La corriente promedio en los MOSFETs principales es:

$$\hat{I}_S = \frac{Po}{2Vin} \quad (9)$$

La corriente instantánea a través de los MOSFETs principales esta definida por:

$$i_s(t) = \begin{cases} 0 \leq t < (D-0.5)T & ; \frac{\Delta i}{2(D-0.5)T} t + \left[\frac{i_{pk}}{2} - \frac{\Delta i}{2} \right] \\ (D-0.5)T \leq t < 0.5T & ; \frac{-\Delta i}{(1-D)T} [t - (D-0.5)T] + i_{pk} \\ 0.5T \leq t < DT & ; \frac{\Delta i}{2(D-0.5)T} (t - DT) + \frac{i_{pk}}{2} \\ DT \leq t < T & ; 0 \end{cases}$$

La ecuación (10) define la corriente eficaz en los MOSFETs principales.

$$I_{Srms} = \sqrt{\frac{1}{T} \int_0^T i_s(t)^2 dt} \quad (10)$$

Al mismo tiempo, la corriente promedio en los rectificadores síncronos está definida como:

$$\hat{I}_{SR} = \frac{Po}{2Vo} \quad (11)$$

Y la corriente instantánea a través de los rectificadores síncronos está definida como:

$$i_{SR}(t) = \begin{cases} 0 \leq t < (D-0.5)T & ; 0 \\ (D-0.5)T \leq t < 0.5T & ; \frac{n\Delta i}{(1-D)T} [t - (D-0.5)T] + ni_{pk} \\ 0.5T \leq t < T & ; 0 \end{cases}$$

La ecuación (12) define la corriente eficaz en los rectificadores síncronos.

$$I_{SRrms} = \sqrt{\frac{1}{T} \int_0^T i_{SR}(t)^2 dt} \quad (12)$$

Particularización del diseño del convertidor

Seleccionar adecuada relación de transformación en el transformador es un trabajo importante para producir mínimas pérdidas en el convertidor. Tienen que ser considerado todos los elementos en los que afecte la selección de éste parámetro. Con las ecuaciones de la sección anterior se muestran una serie de curvas que ayudan a seleccionar el mejor diseño del convertidor dependiendo solo de la relación de transformación.

Variando la relación de transformación de 15 a 30, y sustituyendo los valores de la Tabla I en la ecuación (3), resulta la Fig. 5. Esta figura muestra que el valor mínimo de n para satisfacer todo el rango de tensiones de entrada es 18. Con ésta relación de transformación y con el máximo voltaje de entrada (58V), el mínimo ciclo de trabajo es de 51.4%. En esta topología el mínimo ciclo de trabajo para operar adecuadamente es el 50%.

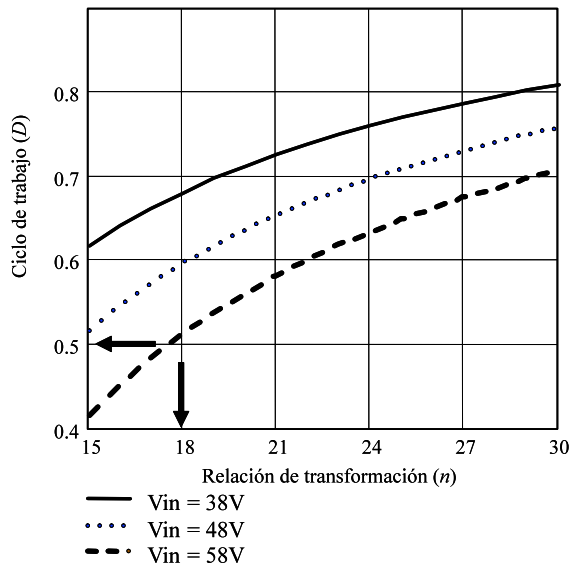


Fig.5. Mínima relación de transformación para todo el rango de Vin

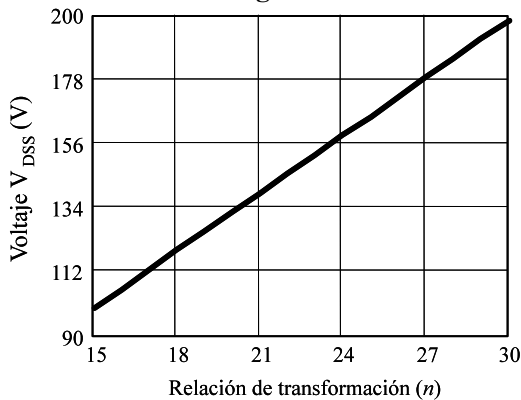


Fig.6. Tensión drenador fuente en el MOSFET principal

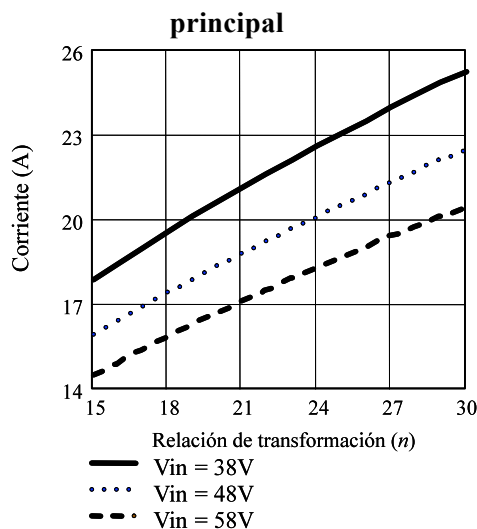


Fig.7. Corriente a través de los rectificadores síncronos

La Fig. 6 muestra la tensión drenador fuente a través de los MOSFETs principales. Se puede ver que si aumenta la relación de transformación n , también aumenta el voltaje en estos MOSFETs. La Fig. 7 contiene información muy importante, en ella se muestra la corriente eficaz de los rectificadores síncronos de la salida. Es importante que ésta corriente sea lo más bajo posible ya que es la que causa las principales pérdidas por conducción en el convertidor.

Como se puede ver, seleccionar el menor valor de la relación de transformación beneficia el diseño del convertidor debido a que los voltajes en MOSFETs principales y las corrientes en los rectificadores síncronos son bajas comparadas con otros valores más altos. Esto significa que el mejor valor para la relación de transformación en este caso particular, es el mínimo como sea posible. En este caso el valor corresponde a 18.

Diseño del transformador

El transformador fue hecho con un núcleo RM10 de perfil bajo, se utilizaron capas de PCB para los devanados para alcanzar un alto acoplamiento entre los devanados. La relación de transformación fue elegida de 18:1 (**Prim1:Sec1**) y la relación de vueltas del devanado auxiliar fue elegida de 18:2 para (**Prim1:Aux1**). Además, para reducir las pérdidas en conducción, varias capas fueron conectadas en paralelo.

Señales de control

El diseño de las señales de control fue realizado con el microcontrolador comercial UC3825. Este controlador proporciona dos salidas desfasadas 180° y su ciclo de trabajo máximo es del 50% para cada señal. Es necesario invertir éstas dos salidas para obtener las señales de control para éste convertidor. La Fig. 8 muestra las formas de onda que se obtienen de UC3825 y las señales que se obtienen al invertir las.

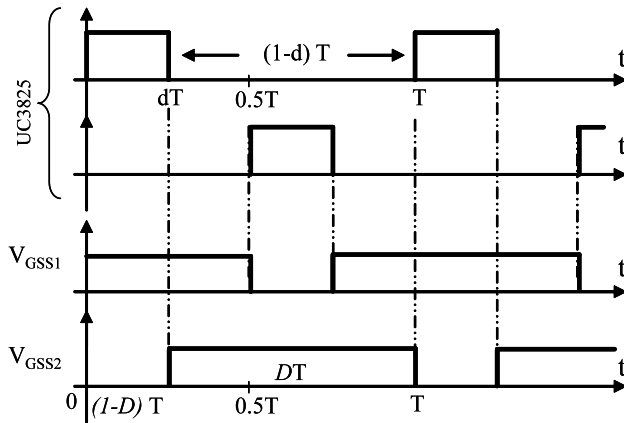


Fig.8. Salida del UC3825 y sus correspondientes señales invertidas

IV. Resultados

Para validar el convertidor propuesto, un prototipo experimental fue construido y probado con las especificaciones de la Tabla 1. Para obtener baja resistencia en conducción, tres MOSFETs SI4410DY fueron conectados en paralelo usándolos como rectificadores síncronos en cada rama de la salida. Por otra parte, dos MOSFETs BUZ31L fueron utilizados en el primario, y la bobina fue construida para tener un rizado de corriente del 10% pico a pico con la máxima carga. Un núcleo RM8/I mecanizado de material 3F3 ($A_L = 250\text{nH}$) con 39 vueltas y 3 conductores de cobre en paralelo y con 29mm del diámetro. Se obtuvo una inductancia de $380\mu\text{H}$. El rizado de voltaje (ΔV) en el condensador de la salida fue seleccionado del 6%, este valor corresponde a un condensador de $178\mu\text{F}$. En el prototipo, dos condensadores OSCON 6.6V-150 μF con baja resistencia fueron colocados. El control fue hecho con el controlador comercial UC3825 que invertía los pulsos para controlar los MOSFETs del primario.

La Fig. 9 muestra la corriente en la bobina de entrada y en los MOSFETs principales. Durante los tiempos en que las señales de control (V_{GSS1} y V_{GSS2}) se traslapan, la corriente instantánea en el secundario del transformador y en los rectificadores síncronos es cero. Lo anterior ocasiona que las corrientes eficaces en el secundario del transformador incrementen si éste

tiempo es muy grande. Por lo tanto, se aumentan las pérdidas de conducción y conmutación cuando este tiempo es muy grande debido a que la corriente que circula es más alta en todo el convertidor. Según lo muestra la Fig. 10, el rendimiento máximo se alcanza con el voltaje de entrada máximo ya que el ciclo de trabajo debe ser más pequeño y los tiempos de traslape son más pequeños, reduciendo la corriente eficaz que circula en el primario y secundario del transformador. La Fig. 10 también muestra cómo el convertidor tiene rendimiento de (90%) con carga media (12A) sin embargo disminuye hasta el 86% con la carga máxima (20A). Es porque el efecto de la corriente circulante aumenta cuando la corriente de salida aumenta.

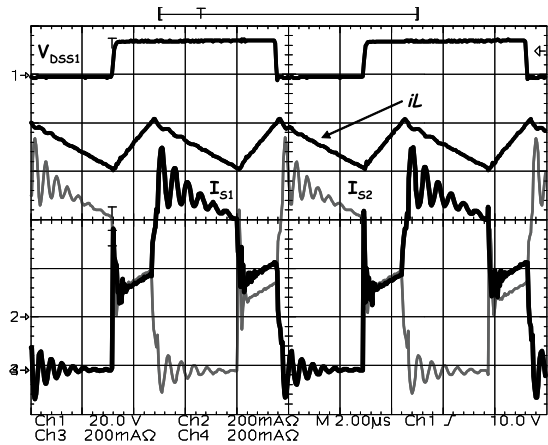


Fig.9. Ch1:20V/div V_{DSS1} ; Ch2:200mA/div i_L ; Ch3:200mA/div I_{S1} and Ch3:200mA/div I_{S2}

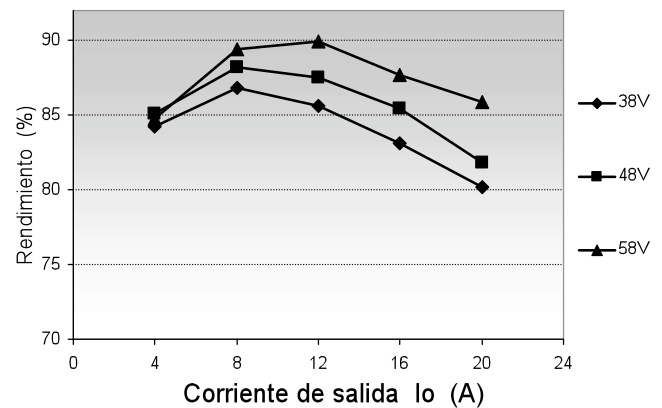


Fig.10. Rendimiento del convertidor

V. Conclusiones

La topología del convertidor Push-Pull Alimentado en Corriente se analiza para ser utilizada en aplicaciones de baja tensión y alta corriente. Se puede implementar fácilmente Rectificación Síncrona Autoexcitada agregando dos devanados auxiliares al transformador. Un convertidor de 3.3V y 20A se ha diseñado y se ha probado para evaluar el funcionamiento de ésta topología.

Durante los tiempos muertos, el flujo de la entrada a la salida se interrumpe, aumentando la corriente que circula a través del convertidor. Cuanto más grande es el tiempo muerto o cuanto más alta es la corriente de salida, más altas son las pérdidas. Por lo tanto, ésta topología se puede aplicar en tensiones muy bajas (0.8V - 3.3V) pero sin corrientes de salida muy elevadas (menos de 20A). Las pruebas del convertidor muestran un rendimiento de 90% en media carga (10A) pero va disminuyendo hasta llegar al 86% con carga máxima (20A).

Se pueden tener distintas tensiones de salida en el convertidor únicamente agregando mas salidas al convertidor, ya que los devanados auxiliares pueden manejar a los rectificadores síncronos para más salidas.

VI. Referencias

- [1] W.A. Tabisz, F.C. Lee, D.Y. Chen, "A MOSFET resonant synchronous rectifier for high-frequency DC/DC converters", IEEE PESC 1990.
- [2] J.A. Cobos, J. Uceda, "Low output voltage DC/DC conversion", Industrial Electronics, Control and Instrumentation, IECON '94.

- [3] J.A. Cobos, O. Garcia; J. Sebastian; J. Uceda, "Active Clamp Forward Converter whit Self Driven Synchronous Rectification", IEEE International Telecommunication Conference (INTELEC) 1993.
- [4] J.A. Cobos, O. Garcia, J. Sebastian, J. Uceda, "RCD clamp PWM forward converter with self driven synchronous rectification", Industrial Electronics, Control, and Instrumentation, IECON '93.
- [5] H. Visairo,; E. Rodriguez, P. Alou, J.A. Cobos, "Multi-output half-bridge converter with single-winding self-driven synchronous rectification", Applied Power Electronics Conference and Exposition (APEC) 2003

VII. Autores

Dr. Luis Alejandro Flores Oropeza nació en Aguascalientes, México en 1976. Recibió el título de Ingeniero Eléctrico en el Instituto Tecnológico de Aguascalientes en 1999. Realizó estudios de Doctorado en la Universidad Politécnica de Madrid en España en 2005. Actualmente es profesor de tiempo completo en la Universidad Autónoma de Aguascalientes.

M. en C. Alejandro Román Loera obtuvo título de Ingeniero en Electrónica y Maestría en Ciencias en Ingeniería Eléctrica por el Instituto Tecnológico de Aguascalientes en 1999 y 2002 respectivamente. Actualmente es profesor de tiempo completo en la Universidad Autónoma de Aguascalientes

A Three Port Converter for Renewable Energy Applications

*Nimrod Vázquez Nava, *Carlos M. Sánchez González, *Claudia V. Hernández Gutiérrez, *Raúl Lesso Arroyo

*Instituto Tecnológico de Celaya, Av. Tecnológico y A. G. Cubas s/n, Tel: (461) 611 7575, n.vazquez@ieee.org

Abstract

A converter suitable for renewable energy applications is proposed in this paper, a three port configuration is chosen, one port for the renewable source, the second one for the battery set and the last one for the output. The converter can deliver energy from each input source in an independently way or in a mixed form, but also the battery can be charged from the renewable source if it is required. The operation and analysis is presented, but also simulation and experimental results.

Palabras Clave: Battery set, renewable energy, three ports.

I. Introduction

Applications like photovoltaic/wind systems use converters capable to accept two inputs [1-11]. In order to increase the efficiency of the system is preferred to have a low voltage at the solar cell array, and some wind systems produce a relatively low voltage. The converters in these applications are then preferred of the boosting type.

The energy provided from these systems is variable and dependent of the climatic conditions, this make that the energy that can be delivered to the load is also variable. Then the converters used in these applications must permit to demand power from both input voltages simultaneously or from each one independently, but depending on the availability of the voltage sources.

This is made traditionally with two independent converters, with a common DC bus voltage (Fig. 1), but these increase the cost of the system. Converters with two inputs have been reported in literature (Fig. 2), which permits to demand current from both inputs [2,5]. Complex controller and circuits are considered in the schemes reported in literature.

In autonomous systems one input is traditionally a battery set, this is due the output power must be maintained even if the energy available in the renewable source is not enough. Some three port converters have been proposed in the literature for this type of applications [6-13]. In Fig. 3 is shown the converter proposed in [7], the main disadvantage of this proposal is that the battery set is always in use during all the different modes of operation, then reducing the battery set useful life. In Fig. 4 is shown another converter that it is based on full-bridge inverter and a transformer, the main disadvantage of this topology is the semiconductors count.

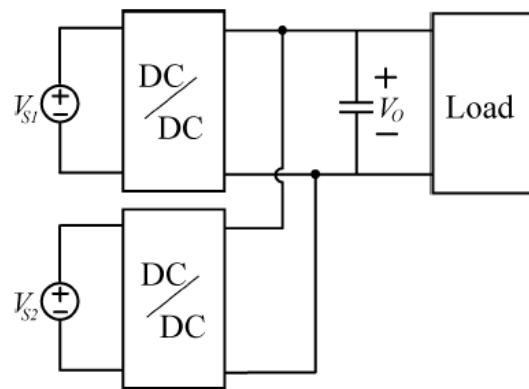


Fig. 1. System with two independent converters, with a common dc bus voltage.

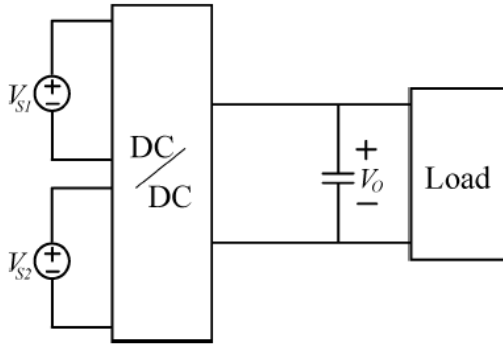


Fig. 2. System with two inputs converter.

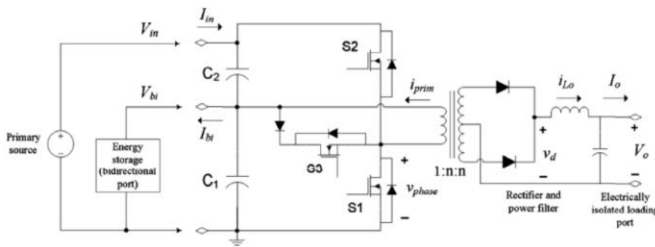


Fig. 3. Three port converter proposed in [7].

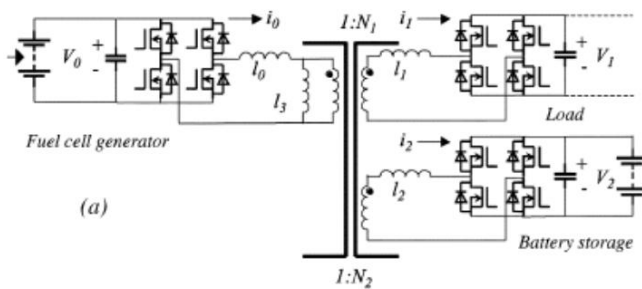


Fig. 4. Three port converter proposed in [8].

There are other three port converters reported in literature, but they are similar to the two previous mentioned.

In this paper is presented a different three port converter capable to accept two input voltages, the power can be demanded from both converters simultaneously or independently depending on the availability of the voltage source. One of these sources is the battery set, which can be charged from the other input if it is necessary, in this case a renewable source is considered. In the proposal the semiconductors count is low and also the battery set can be out of operation when it is not required, increasing its useful life.

The operation of the converter is simple, and the control of the system is easy, because each switch is used for just one purpose.

The paper is organized as follows: in section two the proposed converter is addressed, the operation is discussed; in section three simulations and experimental results of the proposed converter is presented; and finally the conclusions are given.

II. Proposed Converter

The proposed three port converter is based on a step-up converter and a buck converter. As it can be observed in Fig. 5, the system is composed by one step-up converter, one buck converter, but they are integrated in a single stage. An additional switch is used, that permits to integrate the operation of the converter.

The converter is operated in four modes, the first one and second occurs when the energy is demanded from one of the inputs, the third one occurs when the power becomes from both sources, and the last one when the battery set requires to be charged.

1) Power delivered from one voltage source

The equivalent circuit of the converter in this form is shown in Fig. 6. In this mode the switch S1 is turned off, however due to the antiparallel diode of the mosfet, the conduction of the semiconductor is possible in just one way.

The main switch S₂ is switching operating as a dc/dc boost converter to regulate the output voltage. If the switch S₃ is turned on, the energy is taken only from the renewable source (V_{rs}), this is illustrated in Fig. 6(a), the antiparallel diode of S₁ is not conducting due to the renewable source has a higher voltage than the battery set, the inductor that is operating is only L₁. If the switch S₃ is turned off, the power is delivered to the load from the battery set (V_{bat}), this is illustrated in Fig. 6(b), the antiparallel diode of S₃ is conducting when the Switch S₂ is turned off due to the natural conduction of the stored energy in the inductors, in this stage the inductor L₁ and L₂ are in series connection, the diode D_{rs} is not in conduction.

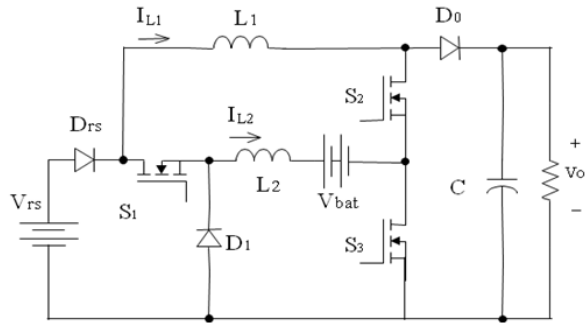
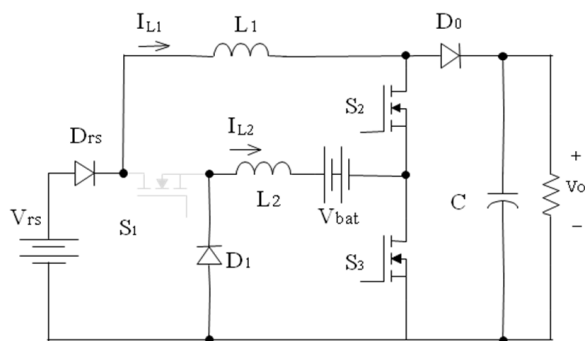
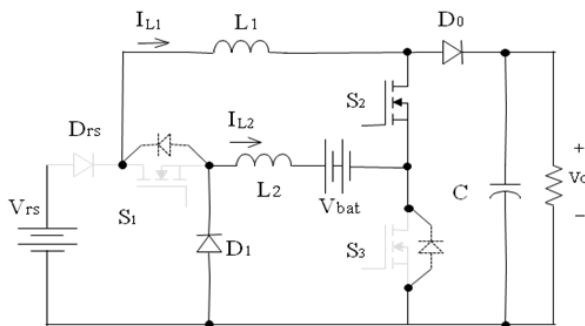


Fig. 5. Proposed converter.



(a) Energy taken from V_{rs}



(b) Energy taken from V_{bat}

Fig. 6. Equivalent circuits of the first operation mode

The main switch S_2 is switching operating as a dc/dc boost converter to regulate the output voltage. If the switch S_3 is turned on, the energy is taken only from the renewable source (V_{rs}), this is illustrated in Fig. 6(a), the antiparallel diode of S_1 is not conducting due to the renewable source has a higher voltage than the battery set, the inductor that is operating is only L_1 . If the switch S_3 is turned off, the power is delivered to the load from the battery set (V_{bat}),

this is illustrated in Fig. 6(b), the antiparallel diode of S_3 is conducting when the Switch S_2 is turned off due to the natural conduction of the stored energy in the inductors, in this stage the inductor L_1 and L_2 are in series connection, the diode D_{rs} is not in conduction.

2) Power delivered from both voltage sources

In this mode the switch S_1 is maintained off, S_2 is commutating to regulate the output voltage; and to deliver energy from both sources (the renewable source and the battery set) S_3 must be switching in synchronization with S_2 , but with different duty cycle (Fig 7). This permit to use an maximum power point tracker for the renewable source in case of necessity, then if it is desired to deliver less energy from a specific input, then the auxiliary switch S_3 must be commutated properly.

This operation mode occurs because the energy available from the renewable source is lower than the power required by the load, and then the battery set must deliver the complement, the duty cycle of S_3 must be adjusted to assure that the renewable source delivers the maximum power available.

3) Battery charging

Once the energy of the battery set is null, this must be charged from the renewable source, if the climatic condition permits this operation form. The auxiliary switch S_3 must be turned on all the time, but now the switch S_1 must be commutating to assure the charge of the battery set (Fig. 8), the battery set is charged with a buck converter topology.

If the energy available from the renewable energy is enough to satisfy the requirements of the load and the battery, the switch S_2 can be used to control the output voltage, however if this is not possible, the system output must be disabled in the output stage (could be an inverter).

Maybe in this mode is better to not deliver energy to the load, because it is possible that the renewable source could not satisfy the energy demanded by the load and the battery set at the same time, and a malfunction may occur.

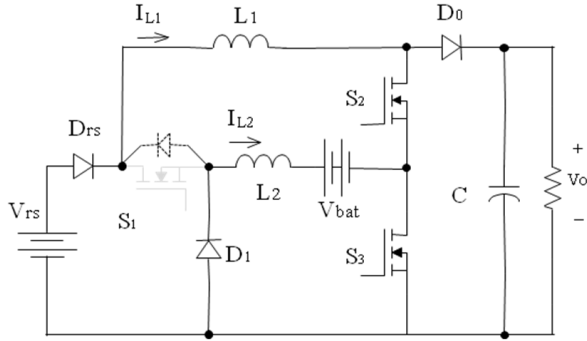


Fig. 7. Equivalent circuits when both sources deliver energy

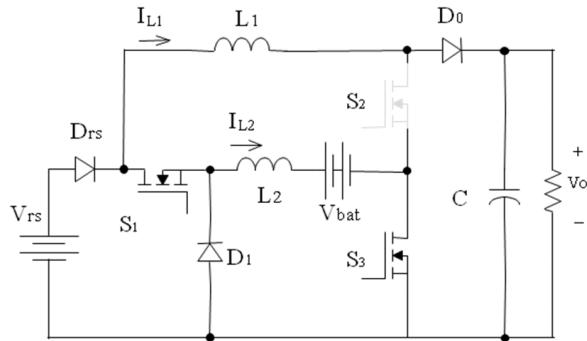


Fig. 8. Equivalent circuits when the battery set is being charged

A. Waveforms of the converter

In order to illustrate more clearly the operation of the converter in Fig. 9 are shown the waveforms under the different modes of operation.

In the first section is shown when the energy is only delivered from the renewable source (Fig. 9), the converter operated as a traditional boost converter with the switch S_2 as the main switch, the battery set is disabled.

In the second section is illustrated when the energy is only obtained from the battery set, as

the switch S_3 is off the renewable source is disabled, but the switch S_2 is also the main switch of the boost converter.

In the third section the energy is demanded from both sources, the switching operation of S_3 determined the amount of energy demanded to each source, and again the switch S_2 operated as the main switch of a dc/dc boost converter.

And finally the fourth section illustrates when the battery set is being charged, but also is considered that the output is feed. In the figure also are shown the currents and control signals of the different semiconductors involved.

B. Controlling the converter

The output voltage of the converter can be controlled using a traditional voltage mode controller, the switch used for this purpose is S_2 ; this is independently from which input voltage the energy is taken.

The switch S_3 determines the amount of energy taken from each voltage source, that is from the renewable source or the battery set. A maximum power point tracker (MPPT) algorithm can be used to assure to deliver the maximum power from the renewable source.

The main controller used is a sliding mode control in order to introduce a good dynamic response to the system [14-17]. This method is used due to their robustness to input voltage variation, which is the case of renewable sources, and also of the batteries.

1) The sliding surface

The controller used in the main stage considers the following sliding surface [14]:

$$\sigma = s_1 x_1 + s_2 \left(\int_0^t (V_o - V_{ref}) dt \right) \quad (1)$$

Where: $\dot{x}_1 = di_{L1}/dt - k_1 x_1$

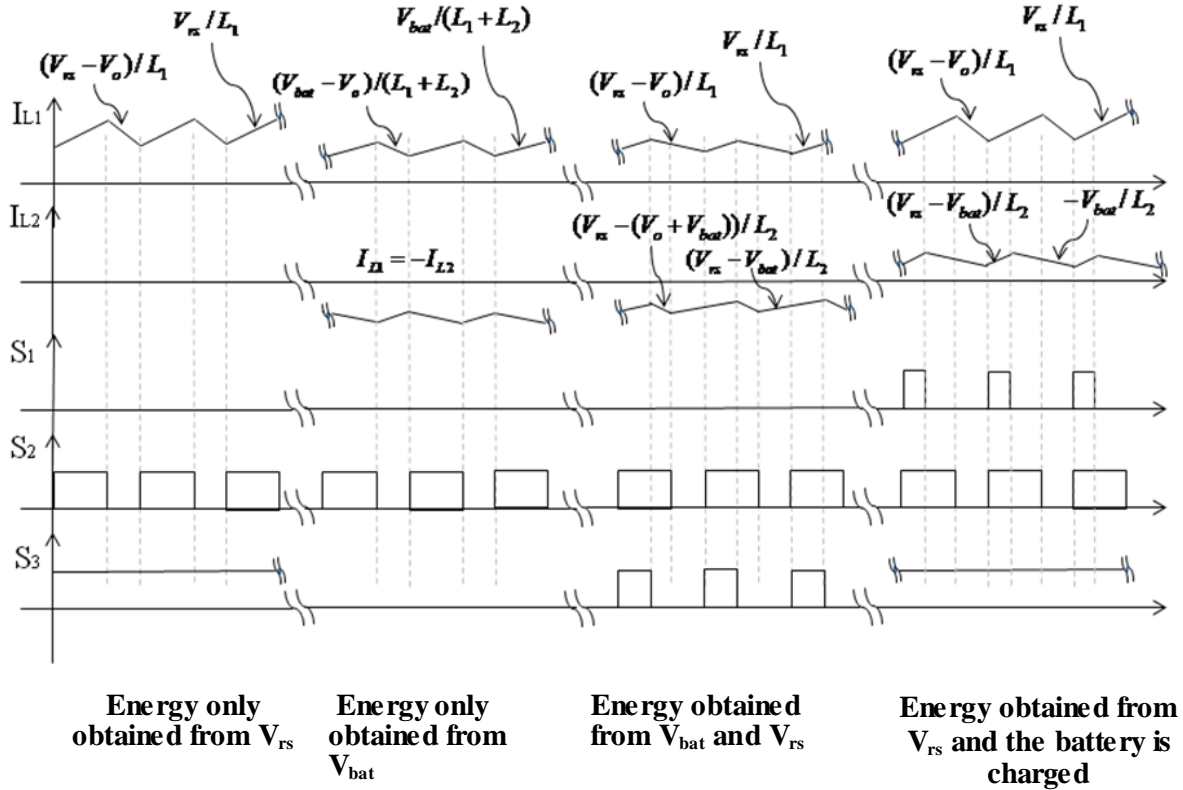


Fig. 9. Waveforms of the converter.

s_1, s_2 are the controller parameters
 ki is a integration constant
 xl is a error current emulator

The control law is:

$$u = \begin{cases} 1 & \sigma < 0 \\ 0 & \sigma > 0 \end{cases} \quad (2)$$

The existence of the sliding mode is determined by [15]:

$$\sigma \dot{\sigma} < 0 \quad (3)$$

III. Simulation and Experimental Results

The proposed converter was designed and numerically simulated, the battery set is of 24V, the renewable source was emulated with a photovoltaic panel, its power characteristics was changed to illustrated the behavior of the system. The output voltage was selected of 60V, the output power is 100W. In the Fig. 8 are shown the simulation results.

In Fig 10(a) is shown the operation when one input voltage is available, in this case the renewable source considered as a photovoltaic panel. The energy available from by the photovoltaic panel is higher that the load power, then the battery no needs to be used.

In Fig 10(b) is shown the operation when the energy available from the renewable source is null, then all the energy is obtained from the battery set.

In Fig. 10(c) is shown the operation then the energy available from the photovoltaic panel is less that the required by the output, then the battery set is used as a compliment, as it can be observed both sources deliver energy to the load.

In Fig 10(d) is shown the operation of the system, when the battery set is being charged form the renewable source. Also in this case is considered that the energy demanded to the renewable source can satisfy the requirements of the load.

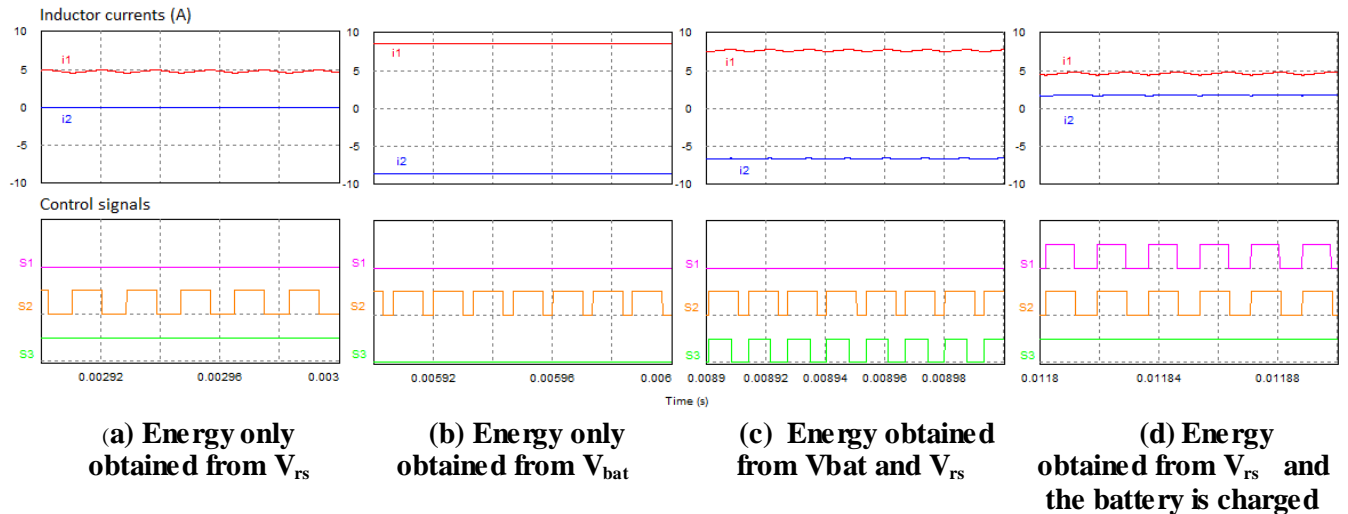


Fig. 10. Simulations of the converter.

In Figs. 11 through 14 are shown the experimental results, the different modes can be observed in the tests and they were obtained with a mixed signals oscilloscope. In order to illustrate the operation and to permit the comparison with the theoretical waveforms, they were not graphed at the designed output power; this is because as it was observed in the simulations the ripple values cannot be appreciated well at the full output power.

In Fig 11 is shown the operation when one input voltage is available, the renewable source considered is a photovoltaic panel emulator. In this case the energy available from the photovoltaic panel is higher that the load power, then the battery no needs to be used. As it can be observed only the inductor L_1 provides the energy to the load, since the inductor L_2 has no current the battery is not in use.

In Fig. 12 is shown the operation when the energy available from the renewable source is null, and also its voltage; then all the energy is obtained from the battery set. As it can be observed the inductor L_1 and L_2 has the same current in magnitude, but opposite sign, in this case the battery is being discharged through both inductors, operating the dc/dc boost converter.

In Fig. 13 is shown the operation when the energy available from the photovoltaic panel emulator is less that the required by the output, then the battery set is used as a compliment, as it

can be observed both sources deliver energy to the load. As it can be observed the current in both inductors are different, but they are delivering energy to the load.

In Fig. 14 is shown the operation of the system, when the battery set is being charged form the renewable source. In this case is considered that the energy demanded to the renewable source can satisfy the requirements of the load, and then as it can be observed both inductors have current, L_1 provides energy to the load and L_2 to the battery set.

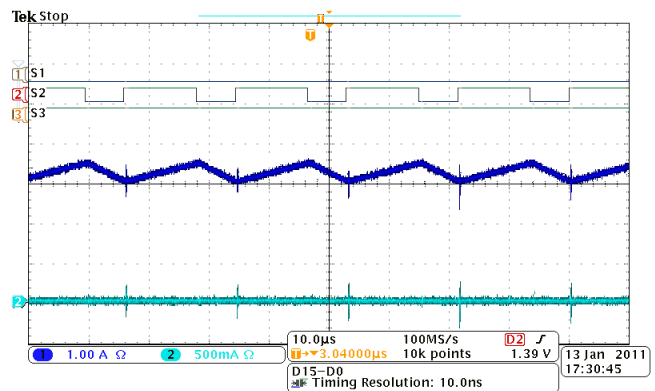


Fig. 11. Operation when the energy is obtained from the renewable source. Top to bottom: S_1 , S_2 , S_3 , i_{L1} , i_{L2} . The current voltage reference of both inductors is the same

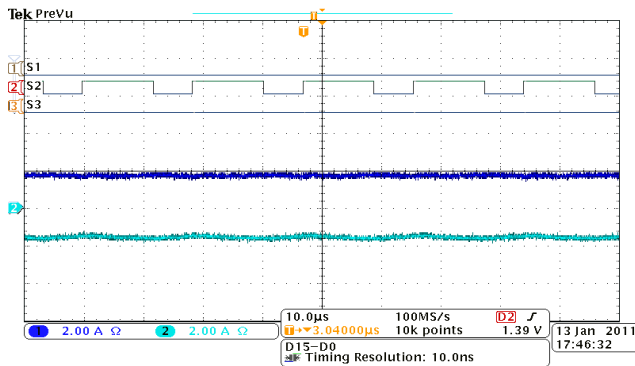


Fig. 12. Operation when the energy is obtained from the battery set. Top to bottom: S_1 , S_2 , S_3 , I_{L1} , I_{L2} . The current voltage reference of both inductors is the same

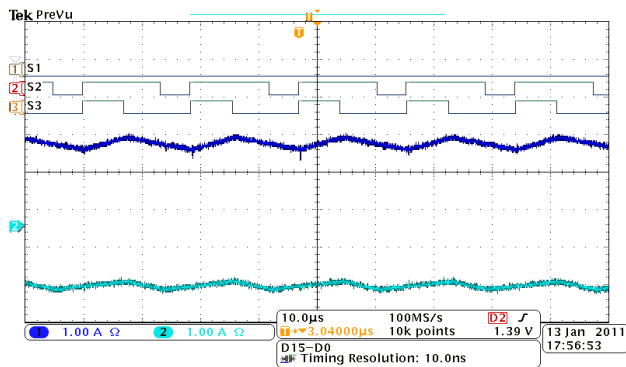


Fig. 13. Operation when the energy is obtained from the battery set and the renewable source. Top to bottom S_1 , S_2 , S_3 , I_{L1} , I_{L2} . The current voltage reference of both inductors is the same

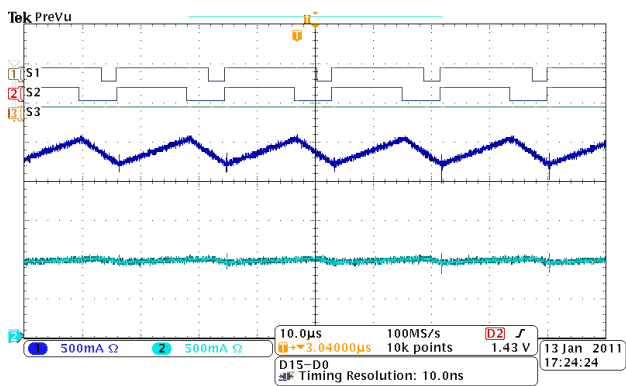


Fig. 14. Operation when the battery set is being charged. Top to bottom: S_1 , S_2 , S_3 , I_{L1} , I_{L2} . The current voltage reference of both inductors is the same

IV. Conclusions

In standalone applications where photovoltaic/wind systems are considered, dc/dc converters that are able to handle the renewable source, and the battery set in a smart way are required; this is because the energy can be taken from this sources depends on their availability due to the climatic conditions.

A different three port DC/DC converter is proposed in this paper, one port for the renewable source, the second for the battery set and finally for the output power. The power can be demanded to both input voltages simultaneously or to each one independently. The operation of the converter is presented in detail. Simulation and experimental results are also presented.

V. References

- [1] Y-M Chen, Y-C Liu, and F Y Wu. "Multi-Input DC/DC Converter Based on the Multiwinding Transformer for Renewable Energy Applications", IEEE transactions on industry applications, vol. 38, no. 4, July/August 2002.
- [2] Y-M Chen, Y-C Liu, and S-H Lin. "Double-Input PWM DC/DC Converter for High-/Low-Voltage Sources", IEEE Transactions on Industrial Electronics, vol. 53, no. 5, October 2006.
- [3] B G. Dobbs and P L. Chapman. "A Multiple-Input DC-DC Converter Topology", IEEE Power Electronics Letters, vol. 1, no. 1, March 2003.
- [4] Y-M Chen, Y-C Liu, S-C Hung, C-S Cheng. "Multi-Input Inverter for Grid Connected Hybrid PV/Wind Power System", IEEE transactions on power electronics, vol. 22, no. 3, May 2007.
- [5] N. Vázquez, A. Hernández, C Hernández, E. Rodríguez, R. Orozco, J. Arau. "A Double Input DC/DC Converter for

- Photovoltaic/Wind Systems". IEEE Power Electronics Specialists Conference 2008, PESC'08. Rhodes, Greece; pp. 2460-2464
- [6] V. M. Pacheco, L. C. Freitas, J.B. Vieira Jr., E. A. A. Coelho and V.J. Farias. "A DC-DC Converter Adequate for Alternative Supply System Applications". IEEE Applied Power Electronics Conference 2002, Dallas, TX, USA, pp 1074-1080.
- [7] H. Al-Atrash, F. Tian, I. Batarseh. "Tri-Modal Half-Bridge Converter Topology for Three-Port Interface". IEEE Transactions on Power Electronics, Vol. 22, No. 1, January 2007. pp. 341-345.
- [8] J. L. Duarte, M. Hendrix, M. Godoy. "Three-Port Bidirectional Converter for Hybrid Fuel Cell Systems". IEEE Transactions on Power Electronics, Vol. 22, No. 2, March 2007, pp. 480-487.
- [9] H. Tao, J. L. Duarte, M. A. M. Hendrix. "Three-Port Triple-Half-Bridge Bidirectional Converter With Zero-Voltage Switching". IEEE Transactions on Power Electronics, Vol. 23, No. 2, March 2008. pp. 782-792.
- [10] H. Krishnaswami, N. Mohan. "Three-Port Series-Resonant DC-DC Converter to Interface Renewable Energy Sources With Bidirectional Load and Energy Storage Ports". IEEE Transaction on Power Electronics, Vol. 24, No. 10, October 2009. pp 2289-2297
- [11] Z. Qian, O. Abdel-Rahman, H. Al-Atrash, I. Batarseh. "Modeling and Control of Three-Port DC/DC Converter Interface for Satellite Applications". IEEE Transactions on Power Electronics, Vol. 25, No. 3, March 2010, pp. 637-649.
- [12] Lu, D.D-C.; Agelidis, V.G., "Photovoltaic-Battery-Powered DC Bus System for Common Portable Electronic Devices," Power Electronics, IEEE Transactions on , vol.24, no.3, pp.849-855, March 2009
- [13] Yang Du; Lu, D.D-C, "Analysis of a battery-integrated boost converter for module-based series connected photovoltaic system," Power Electronics Conference (IPEC), 2010 International , vol., no., pp.694-698, 21-24 June 2010
- [14] N. Vazquez, C Hernandez, J Alvarez, J Arau, "Sliding Mode Control for DC/DC Converters: A new Sliding Surface", IEEE International Symposium on Industrial Electronics 2003, CD proceeding
- [15] R. A. DeCarlo, S. Zak, G. P. Matthews, "Variable Structure Control of Nonlinear Multivariable Systems: A Tutorial", Proceedings of the IEEE, vol. 76 No. 3, March 1988, pp. 212 - 232.
- [16] J. Y. Hung, W. Gao, J. C. Hung, "Variable Structure Control: A Survey", IEEE Transactions on Industrial Electronics, vol. 40, No. 1, Feb. 1993, pp. 2 - 18.
- [17] V.I Utkin, Sliding Modes And Their Application In Variable Structure Systems, MIR Publishers, Moscow, 1974.

VI. Authors

Dr. Nimrod Vazquez Nava obtuvo el título de Ingeniero en Electrónica por el Instituto Tecnológico de Celaya en 1994, los grados de Maestro en Ciencias en Ingeniería Electrónica y Doctor en Ciencias en Ingeniería Electrónica por el cenidet en 1997 y 2003 respectivamente, actualmente es profesor-investigador en el Instituto Tecnológico de Celaya e investigador nivel 1 en el SNI.

Ing. Carlos M. Sánchez González obtuvo el título de Ingeniero Eléctrico por la Universidad Michoacana de San Nicolás de Hidalgo y actualmente es estudiante de la Maestría en Ciencias en Ingeniería Electrónica en el mismo Instituto.

M.C. Claudia V Hernández Gutiérrez obtuvo el título de Ingeniero en Electrónica por el

Instituto Tecnológico de Celaya en 1995 y el grado de Maestra en Ciencias en Ingeniería Electrónica por el cenidet en 2000, actualmente es docente y Coordinadora de Posgrado en el Instituto Tecnológico de Celaya.

M.C. Raúl Lesso Arroyo obtuvo el título de Ingeniero en Electrónica y el grado de Maestro en Ingeniería por la Facultad de Ingeniería Mecánica, Eléctrica y Electrónica de la Universidad de Guanajuato en 1991 y 1993 respectivamente, actualmente es docente en el Instituto Tecnológico de Celaya.

A Six-Switch Three-level Current Source Inverter

*Nimrod Vázquez Nava, *Héctor López Tapia, *Luz García Rodríguez, *Claudia V. Hernández Gutiérrez, **Juan Carlos Iturria Echalar, **Miguel Nadal Satorre

*Instituto Tecnológico de Celaya, Av. Tecnológico y A. G. Cubas s/n, Tel: (461) 611 7575, n.vazquez@ieee.org

**Carrocerías Diversas Halcón SA de CV, Carr. Panamericana Km 284, Celaya, Tel: (461) 611 0195, ingenieria@carroceriashalcon.com

Abstract

In several applications DC/AC converters are widely used; traditionally they can be classified in voltage source inverters (VSI) and current source inverters (CSI). Their use depends on the application. Other possibility for the DC/AC conversion is the multilevel configuration, and the most analyzed is the VSI. In this paper a multilevel current source inverter (MCSI) is proposed. The operation, analysis and implementation are presented; also simulation and experimental results are shown.

Palabras Clave: CSI, DC/AC converter, Multilevel inverter.

I. Introduction

The dc/ac converters can be classified in two types: the voltage source inverter (VSI) and the current source inverter (CSI) [1]–[15].

The VSI has been studied extensively [1]–[8], this type of converter can be classified in half bridge inverter (HBI), full bridge inverter (FBI) and multilevel inverter (MI); and the last one can be subdivided in neutral point clamped inverter (NPC), flying capacitor inverter, cascade full bridge inverter [1] and recently the hybrid and hexagram inverter [3]–[5]. Also many works have been studied the MI [8]. In Fig. 1 the topologies of different type of multilevel VSI are presented.

Nowadays the CSI has been studied less than the VSI, however it is used extensively in motor drives and it is recommended in applications when boosting capabilities are required; this type

also two multilevel topologies of CSI [17],[18]; the last type can be subdivided in the inverters that we referred as the “embedded” configuration [13], and the “two stage” configuration [14]. In

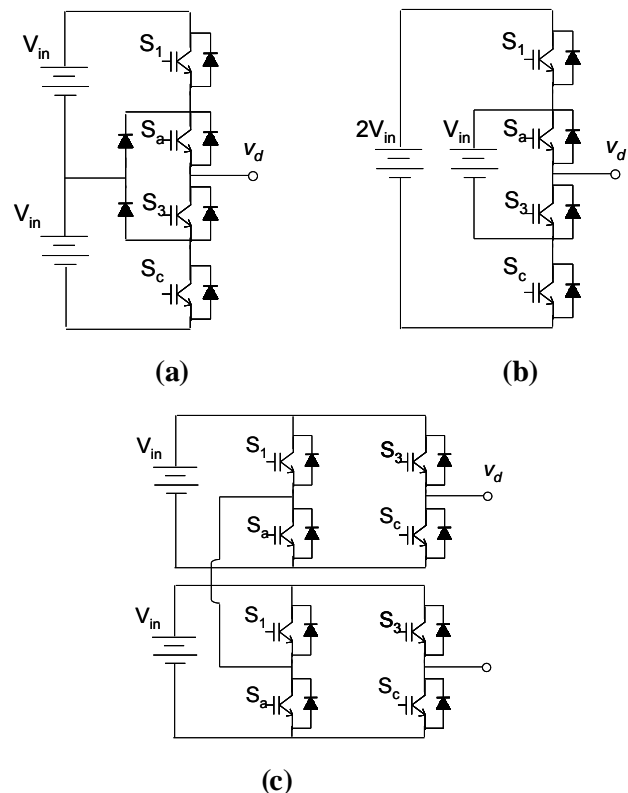


Fig. 1. Different topologies of multilevel VSI. (a) NPC converter. (b) Flying capacitor converter. (c) Cascaded converter.

This work was sponsored by CONA CyT under project ECO-2010-C01-00000000147044

of converter can be classified in pulse with modulated (PWM) CSI [13]–[15], load-commutated inverters (LCIs) [12] and there are

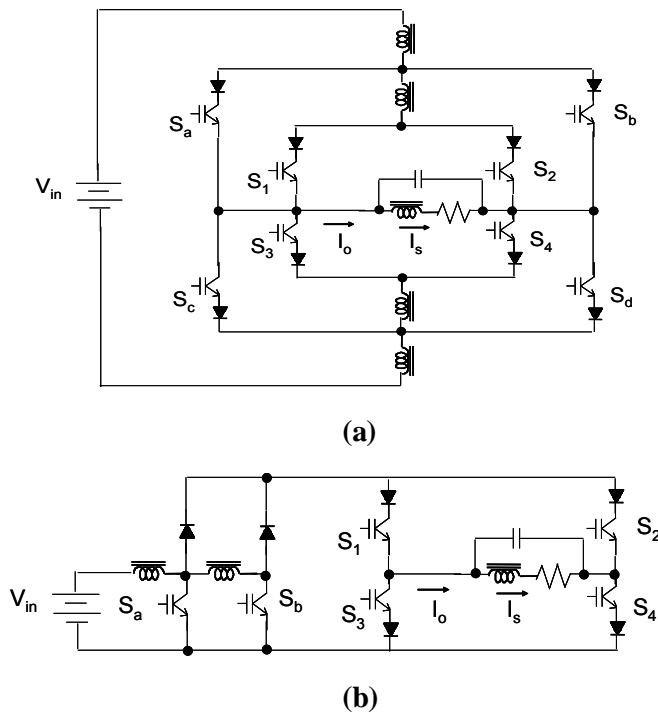


Fig. 2. Different topologies of multilevel CSI (a) Embedded converter, (b) Two stage converter.

Fig. 2 the topologies of multilevel CSI are presented.

Basic concepts are different for each inverter. VSI requires considering a blanking time, but an overlapping time is prohibited in one inverter leg. CSI needs an overlapping time, but blanking time is strictly prohibited in the upper/lower switches of the inverter legs. Finally, the ZSI allows both in one inverter leg.

Current source converters feature a simple structure, low switching dv/dt , and reliable over-current/short-circuit protection. On the other side, multilevel converters offer the advantage that a low harmonic content can be obtained with relatively low switching frequency [1].

In this paper a type of CSI is discussed; in fact a multilevel type, that combines the characteristics of CSI and multilevel converter. This converter is quite simple, it uses fewer or equal amount of semiconductors than the multilevel CSI reported

in literature; and also the current is equally distributed in the inductors associated to the topology

II. Proposed Converter

The proposed idea consists of a CSI, but a multilevel type that permit to produce an output current with a lower harmonic content compared with the semiconductors switching frequency, the proposed converter injects current to the output in a parallel connection.

In Fig. 3 the proposed three level topology is shown; it consists of two CSI connected in parallel.

Operation of the converter

In order to explain the operation of the proposed converter, the input current is considered constant (large inductors); and when the switches S_1 , S_2 , and S_3 are turned on the other switches are turned off, in fact the control signal of S_1 is exactly the opposite of S_a and so on for the other switches. In practice, there is a small overlapping in the control signals S_1 and S_a , this is because always a current path must be provided to the inductors in order to avoid the semiconductor damage.

Fig. 4 shows the current generated by the proposed MCSI. These waveforms are at low frequency and only the positive semicycle is shown in order to illustrate the behavior. Control signals for the switches S_1 , S_2 and S_3 are also included. According to Fig. 4 the converter operates as follows:

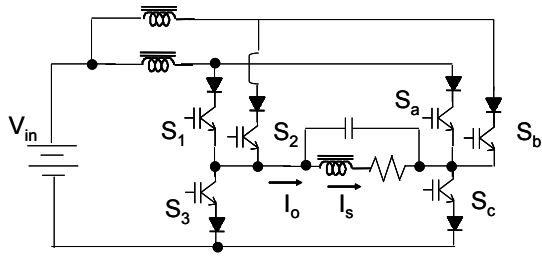


Fig. 3 Proposed three level current source inverter

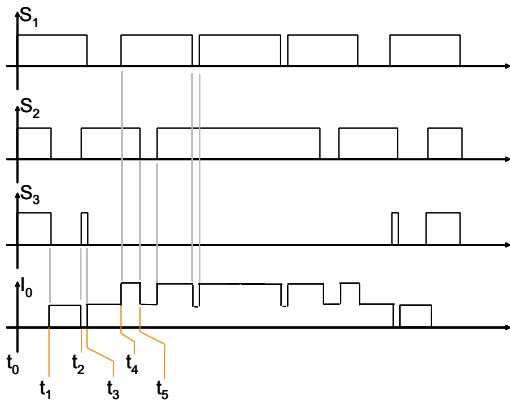


Fig. 4. Control signals and output current. Top to down: Control signal of S₁, control signal of S₂, control signal of S₃ and output current

During $t_0 - t_1$. S_1 , S_2 and S_3 are turned on, the equivalent subcircuit is shown in Fig. 5(a), then zero current is injected (also a zero current can be produced when S_1 , S_2 and S_3 are turned off).

During $t_1 - t_2$. S_1 is maintained on, S_2 and S_3 are turned off, the equivalent subcircuit is shown in Fig. 5(b), and then a positive current is delivered.

During $t_2 - t_3$. The same switches combination that during t_0-t_1 is used. Then a zero current output is obtained.

During $t_3 - t_4$. A positive current is injected to the load, but in this case the other combination is used. S_2 is turned on, S_1 and S_3 are turned off.

During $t_4 - t_5$. S_1 and S_2 are turned on, and S_3 is turned off, the equivalent subcircuit is shown in Fig. 5(c), and then a double magnitude of positive current is obtained.

After t_5 . During the rest of time the converter operation is similar to the before mentioned.

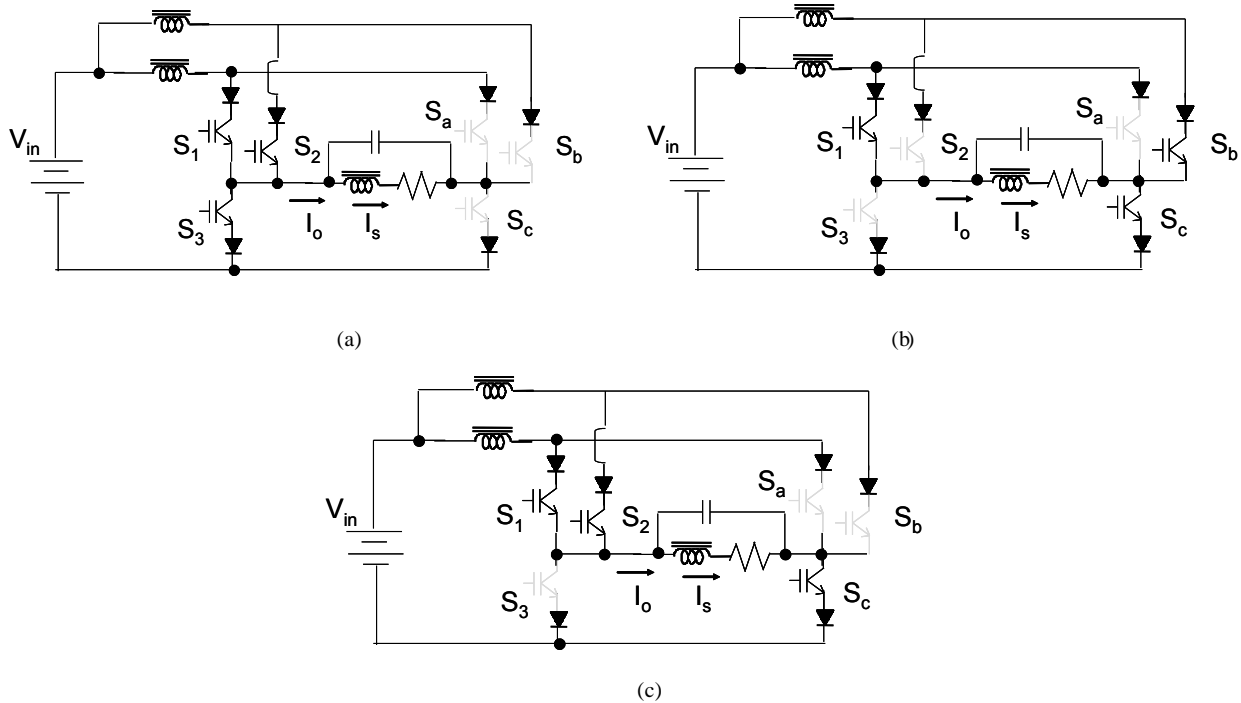


Fig. 5. Subcircuits of the proposed multilevel CSL. (a) Subcircuit for zero current. (b) Subcircuit for positive current. (c) Subcircuit for double positive current.

Balancing the inductors current

A natural commutation of the switches was selected; this permit to equalize the inductor currents in ideal conditions and in experimental situation only a slight unbalance will appear.

Converter redundancy is taken into account to generate the same output level with different switching states. There are two combinations of the switches to deliver a positive current; alternating both of them gives balance at the inductor currents.

The implementation was made alternating the two combinations in steady state; however a simple controller is employed to regulate the inductor currents in a much better way, just in case of current unbalance due to parasitic elements or disturbances in the system.

III. Experimental and Simulation Results

Numerical simulations are shown in Figs. 6 and 7. They were made at 600Hz, with an inductor of 40mH due to the response time of the simulator. Fig. 6 shows the output current with a modulation index of 0.99 and also the inductors current, and

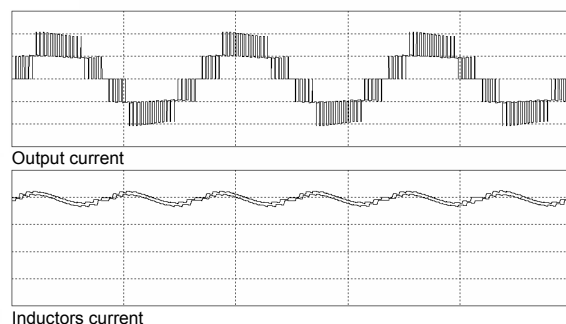


Fig. 6. Output current and inductors current. Top to down: Output current (500mA/div), inductors current (125mA/div). Time: 1ms/div

as it can be observed they are well balanced.

Fig. 7 illustrates the performance of the system with a lower modulation index, particularly for this case the output current and the control signals of the switches S_1 , S_2 and S_3 are shown. As it can be observed the output has a higher switching frequency than the semiconductors, this permits to reduce the losses significantly and also the requirements of the output filter are lower compared to the switching frequency.

In Figs. 8 and 9 some experimental results of the system are shown. The input inductors are of 400mH; the switching devices are IRG4PC40 without anti-parallel diode, as a consequence the extra topology diodes are not required; the use of this device reduces the cost of the system and also the conduction losses are reduced.

In Fig. 8 the output current of the proposed converter and the inductor currents are shown; it can be observed that they are well balanced; the converter operates as it was expected.

In Fig. 9, the output current and its respective harmonic content are shown; it can be observed that the first high frequency harmonic appears

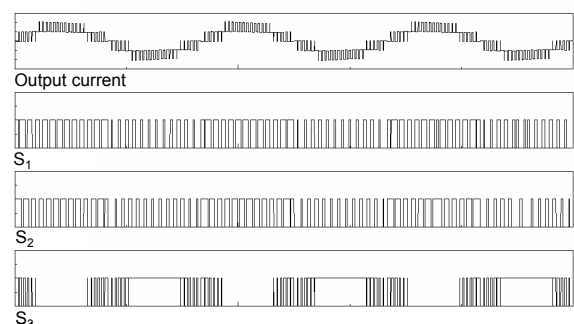


Fig. 7 Output current and control signals of the switches. Top to down: Output current (500mA/div), Control signal S1, S2 and S3 respectively. Time: 1ms/div.

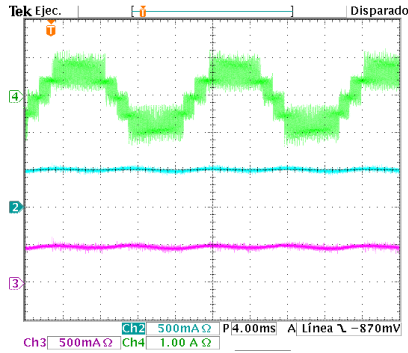


Fig. 8. Output current and inductor currents.

around 50 KHz and the carrier frequency is 25 KHz. The THD obtained in this test is lower than 5%.

IV. Conclusions

This paper presents a different method to build a multilevel current source inverter, it is suggested a parallel connection. The converter is different to the previous published schemes. For a single-phase three level CSI, six semiconductors and two inductors are used. The proposed idea can be extrapolated to three phase systems, but also more levels can be added.

V. References

- [1] Muhammad H. Rashid. "Power Electronics: Circuits, Devices, and Applications" Third Edition, Ed Prentice Hall
- [2] M. A. Pérez, P. Cortés, J. Rodríguez. "Predictive Control Algorithm Technique for Multilevel Asymmetric Cascaded H-Bridge Inverters" IEEE Transactions on Industrial Electronics. Vol 55, No 12, December 2008. pp 4354-4361.
- [3] A. Chen, X. He. "Research on Hybrid-Clamped Multilevel-Inverter Topologies". IEEE Transactions on Industrial Electronics. Vol 53, No 6, December 2006. pp 1898-1907.
- [4] C. Rech, J. R. Pinheiro. "Hybrid Multilevel Converters: Unified Analysis and Design

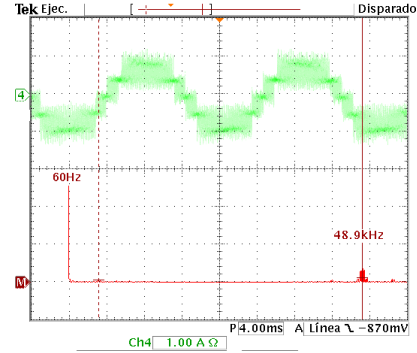


Fig. 9. Output current and harmonic content.

- [5] Considerations". IEEE Transactions on Industrial Electronics. Vol 54, No 2, April 2007. pp 1092-1104.
- [6] J. Wen, K. M. Smedley. "Hexagram Inverter for Medium-Voltage Six-Phase Variable-Speed Drives" IEEE Transactions on Industrial Electronics. Vol 55, No 6, April 2008. pp 2473-2481.
- [7] H. du T. Mouton. "Natural Balancing of Three-Level Neutral-Point-Clamped PWM Inverters". IEEE Transactions on Industrial Electronics. Vol 49, No 5, October 2002. pp 1017-1025.
- [8] G.P. Adam, S.J. Finney, A.M. Massoud, B.W. Williams. "Capacitor Balance Issues of the Diode-Clamped Multilevel Inverter Operated in a Quasi Two-State Mode" IEEE Transactions on Industrial Electronics. Vol 55, No 8, August 2008. pp 3088-3099
- [9] J. Rodríguez, J. Pontt, P. Correa, P. Cortés, C. Silva. "A New Modulation Method to Reduce Common-Mode Voltages in Multilevel Inverters". IEEE Transactions on Industrial Electronics. Vol 51, No 4, August 2004. pp 834-839.
- [10] G. Joos, G. Moschopoulos, P. D. Ziogas, "A High Performance Current Source Inverter", IEEE Transactions on Power Electronics, Vol 8, No 4, Oct. 1993, pp 571-579
- [11] J. R. Espinoza, G. Joos. "Current-Source Converter On-Line Pattern Generator Switching Frequency Minimization". IEEE

Transactions on Industrial Electronics, Vol 44, No 2, April 1997. pp 198-206.

VI. Autores

Dr. Nimrod Vazquez Nava obtuvo el título de Ingeniero en Electrónica por el Instituto Tecnológico de Celaya en 1994, los grados de Maestro en Ciencias en Ingeniería Electrónica y Doctor en Ciencias en Ingeniería Electrónica por el cenidet en 1997 y 2003 respectivamente, actualmente es profesor-investigador en el Instituto Tecnológico de Celaya e investigador nivel I en el SNI.

M. C. Héctor López Tapia obtuvo el título de Ingeniero Electrónico y el grado de Maestro en Ciencias en Ingeniería Electrónica por el Instituto Tecnológico de Celaya, actualmente es docente en el Instituto Tecnológico de Celaya.

Ing. Luz García Rodríguez obtuvo el título de Ingeniero Electrónico por el Instituto Tecnológico de Celaya y actualmente es estudiante de la Maestría en Ciencias en Ingeniería Electrónica en el mismo Instituto.

M.C. Claudia V Hernández Gutiérrez obtuvo el título de Ingeniero en Electrónica por el Instituto Tecnológico de Celaya en 1995 y el grado de Maestra en Ciencias en Ingeniería Electrónica por el cenidet en 2000, actualmente es docente y Coordinadora de Posgrado en el Instituto Tecnológico de Celaya.

Juan Carlos Iturria Echalar es licenciado en Administración de Negocios por la Universidad de Celaya en el año 1994. Se ha desempeñado en diferentes puestos y funciones administrativas, actualmente es el Gerente General de la empresa Carrocerías Diversas Halcón SA de CV

Miguel Nadal Satorre obtuvo el título de Ingeniero Mecánico en la Escuela de Ingeniería Industrial de Alcoy (España) en el año 1971 . Se ha desempeñado en diferentes puestos en empresas españolas y mexicanas y actualmente es ingeniero de procesos de la empresa Carrocerías Diversas Halcon S.A. de C.V

A Single Stage Electronic Ballast Based on the ‘Z’ Source Inverter

*Nimrod Vázquez Nava, *Alejandro Perea Franco, *Claudia V. Hernández Gutiérrez, **Jaime Arau Roffiel

*Instituto Tecnológico de Celaya, Av. Tecnológico y A. G. Cubas s/n, Tel: (461) 611 7575, n.vazquez@ieee.org
 ** *cenidet*, Int. Internado Palmira S/N, Tel: (777) 318 77 41, jarau@cenidet.edu.mx

Abstract

Electronic ballast for fluorescent lamp with high power factor is traditionally implemented in two stages: first one to obtain a unity power factor and second one to feed properly the lamp. Different topologies have been studied in order to able reduce the number of stages and employed controllers. In this paper a different electronic ballast, which is based on the ‘Z’ source inverter is analyzed, as a result of this the system becomes a single stage and high power factor electronic ballast. Operation, analysis, and experimental results are presented.

Palabras Clave: Electronic ballast, two stages, Z source inverter.

I. Introduction

The use of electronic ballasts for fluorescent lamps has increased because they offer a good performance: high efficiency, high power factor (PF), and light weight. In the literature have been reported different topologies [1]-[19]; the two stage topology is the traditional approach [1] (Fig. 1). For electronic ballast there are two important requirements: the power quality standard and the current crest factor at the lamp. In the two stage approach, each stage is used for each requirement; however the disadvantage of this approach is the amount of semiconductors, the reduced efficiency, and also the use of two controllers. Different topologies have been proposed in order to reduce the drawbacks of the two stage approach.

Literature [2-3] suggest a two stage topology, where only a single controller is used; these converters are based on the known “Valley fill” circuit, and it is used to demand a low harmonic content, not a unity power factor; the topology reported in [2] is shown in Fig. 2. Even when two stages are employed, a reasonable efficiency is obtained. It is important to notice the

semiconductors count and the amount of passive elements used in the topology.

There exist another converter [4], which demands a low harmonic content; it is based on the active input current technique. In spite of being based on techniques that considers the use of just one stage, the final implementation was made in two stages, but certainly one controller is used.

It is found in literature [5]-[11] single stage topologies; however these are actually integrated topologies. For example a flyback converter integrated with a push pull converter (Fig. 3) is presented in reference [5]. The flyback converter is operated in discontinuous conduction mode in order to assure a good power factor, and then efficiency is penalized significantly.

There are also found in literature [12-14] topologies based on the charge pump technique. This technique offers a good power factor, however operation is complex and efficiency is reduced. A topology for this configuration is shown in Fig. 4.

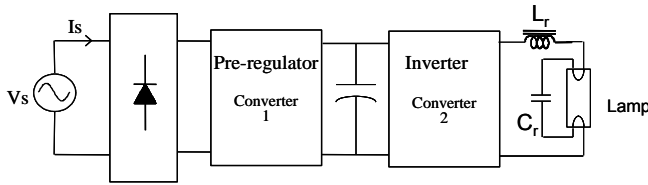


Fig. 1. Traditional two stage approach.

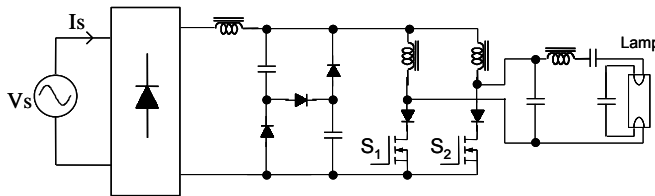


Fig. 2. Converter based on Valley Fill circuit.

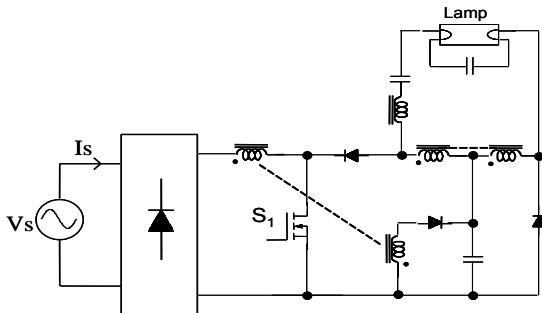


Fig. 3. Converter based on the integration of a Flyback and push-pull converters.

It should be noticed from Figs. 2 to 4 that it is required an EMI filter for the ac side, which it has not being included in the figs.

Reducing the amount of stages is traditionally the main target in electronic ballast; however efforts are focused on dimming control; being more specific some recent works [15-17] are made magnetically. Some other works [18-19] deals with self-oscillating converter that does not require a control circuit.

This paper proposes a single stage electronic ballast based on a different inverter topology: a 'Z' source inverter (Fig. 5(a)). It is not considered the integration between stages or complex operation. The implemented system

offers in a single stage good power factor feature and properly lamp operation. The demanded current at the input is similar to the active input current shaper converter (a parallel version [21]), however topology is completely different. Only the inverter stage with its inherent passive elements of the topology is used (Fig. 5(b)), it should be noticed that the EMI filter is also required.

Paper is organized as follows: proposed converter is addressed in section two, not only operation and waveforms but also analysis in details about this proposal are included; experimental results are discussed in section three, and some final conclusions are given.

II. Proposed Converter

Inverters traditionally used in electronic ballast are voltage and current source inverters (VSI and CSI, respectively) as it is found reported in literature [1-20]. However it is also published [22] a different inverter: the 'Z' source inverter (ZSI), but not used as electronic ballast.

Basic concepts are different for each inverter. VSI requires considering a blanking time, where an overlapping time is prohibited in one inverter leg. CSI needs an overlapping time, however, blanking time is strictly prohibited in the upper/lower switches of the inverter legs. And finally, the ZSI allows both in one inverter leg; this feature let this inverter may be operated at more switching states than the VSI and CSI, therefore it is possible to have more possibilities in its functions.

'Z' source inverter is used in this paper as electronic ballast, and the two ballast requirements (power quality standard and current crest factor) are satisfied in a single stage. Some switching states are used in order to control the lamp output voltage, and other states are used for satisfying input current specifications, where properly lamp operation is kept.

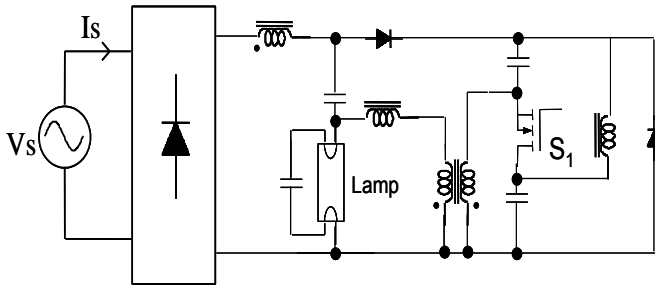


Fig. 4. Converter based on charge pump technique.

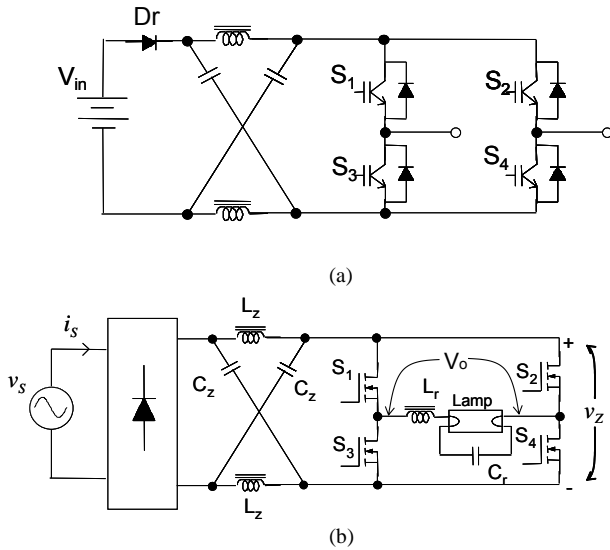


Fig. 5. The 'Z' source inverter. (a) The traditional topology, (b) The proposed electronic ballast

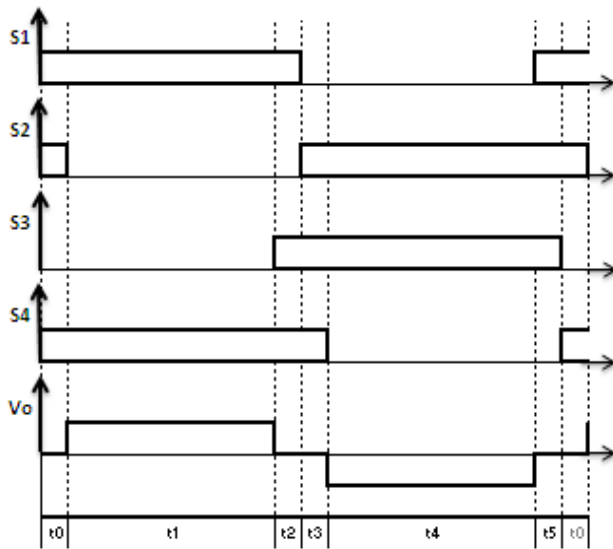


Fig. 6. Output waveform of the 'Z' source inverter at one switching period.

A. 'Z' source invertir

'Z' source inverter topology is shown in Fig. 5(a). Diode Dr and passive elements are important in the topology. Proposed electronic ballast is shown in Fig. 5(b), where the diode Dr is a full bridge rectifier.

According to Fig. 6 the converter operates as follows:

During t0. S1, S2 and S4 are turned on, the equivalent subcircuit is shown in Fig. 7(a), and then a zero voltage is applied to the resonant tank, where the inverter input voltage is zero (vz); it should be noticed that the full bridge diode is not conducting.

During t1. S1 and S4 are maintained on, S2 is turned off, the equivalent subcircuit is shown in Fig. 7(b), and then a positive output voltage is delivered to the resonant tank; the full bridge diode is conducting.

During t2. S1 and S4 are maintained on, S3 is turned on, the equivalent subcircuit is shown in Fig. 7(c), then a zero voltage is applied to the resonant tank, but also the inverter input voltage is zero (vz).

During t_3 . S_2 and S_4 are maintained on, S_1 is turned off, but S_2 is turned on, the equivalent subcircuit is shown in Fig. 7(d), then a zero voltage is still applied to the resonant tank.

During t_4 . S_2 and S_3 are maintained on, and S_4 is turned off, the equivalent subcircuit is shown in Fig. 7(e), and then a negative voltage is applied to the resonant tank.

During t_5 . S_2 and S_3 are maintained on, and S_1 is turned on, the equivalent subcircuit is shown in Fig. 7(f), then a zero voltage is applied to the resonant tank.

After t_5 . The switching sequences are being repeated since this time.

Table I summarize different switching states for producing the required output voltage. Reader should be aware that in order to generate zero output voltage there are more switching states than the traditional inverter; this redundancy is used to correct the power factor as it will be shown later.

Converter output voltage is actually slightly different from shown in Fig. 6; this is because inductors L_Z are operated in discontinuous conduction mode (DCM), and then time t_1 and t_4 are splitted in two stages as it will be explained later.

B. Output characteristics

Output for the proposed inverter stage is intended to be a quasi square output voltage (Fig. 6); traditionally this is not considered for electronic ballast.

Converter is operating at the resonant frequency; then the filtered output voltage is only composed by the fundamental component. The only difference between using the square output and the quasi square output voltage is the magnitude of the fundamental component, and then the resonant tank must be designed according to this consideration.

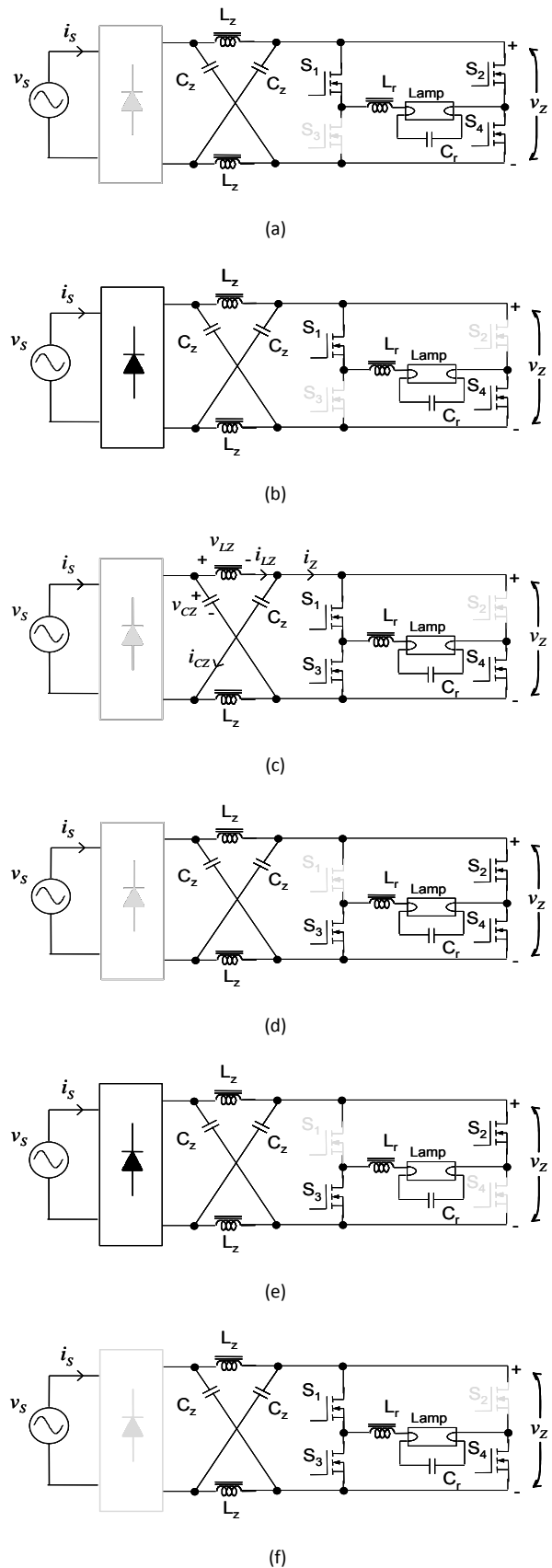


Fig. 7. Subcircuits of the proposed converter. In (a) is indicated the nomenclature of the equations

Resonant tank is made with L_r and C_r in a parallel configuration; its transfer function is calculated by [23]:

$$\frac{V_{OLF}}{V_{INV}}(s) = \frac{1}{1 + s L_r / R + L_r C_r s^2} \quad (1)$$

Where:

L_r, C_r are the inductance and capacitance of the resonant tank

R is the lamp resistance in steady state.

V_{OLF} is the lamp voltage with the filter

V_{INV} is the voltage of the inverter without the filter.

Resonant frequency and gain at this frequency are established according to [22]:

$$\omega_r = \frac{1}{\sqrt{L_r C_r}} \quad (2)$$

$$Q = \frac{R}{\omega_r L_r} \quad (3)$$

As converter is operated at resonant frequency and a quasi square voltage is used.

C. Input characteristics

The inverter input is the ‘Z’ array and also the full bridge rectifier. In spite of having same elements for the mentioned inverter, a high power factor is obtained. The inductors LZ are operated in discontinuous conduction mode.

Converter operates in a similar way as the parallel active input current shaper proposed in

Output voltage	Switches			
	S ₁	S ₂	S ₃	S ₄
Zero	1	1	X	X
	X	X	1	1
Positive	1	0	0	1
Negative	0	1	1	0

‘1’ means ‘on’
‘0’ means ‘off’
‘X’ means ‘don’t care’

Table 1. Switching states of the converter

literature [21] does; this is by regarding the input current waveform. However, topology is completely different.

Equations concerning to the input are:

1) When the converter is working at the switching states during $t1$ and $t4$.

Voltage applied to the inverter (v_z) when the full bridge diode is conducting (stage 3, Fig. 8) is determined by:

$$v_z = 2v_{CZ} - |v_s| \quad (4)$$

Where:

v_{CZ} is the capacitor voltage of CZ, and it is almost constant

$|v_s|$ is the rectified input voltage

This voltage is actually applied to the lamp during $t1$ and $t4$ according to Fig. 6; certainly with different polarity in each case. During these times, the capacitors CZ are charged and the inductors LZ are discharged.

The inductor LZ voltage is:

$$v_{LZ} = |v_s| - v_{CZ} \quad (5)$$

The rectified input current is:

$$|i_s| = 2i_{LZ} - i_z \quad (6)$$

Where: i_{LZ} is the inductor current of LZ

i_z is the current demanded by the inverter

The current of the capacitor is:

$$i_{CZ} = \frac{|i_s| - i_z}{2} \quad (7)$$

It should be noticed that at the zero crossing of input voltage, due to expression (6) inductors LZ are discharged more than any other time; this is a natural behavior due to the input voltage evolution. And consequently due to (9), also input current in a natural way has an evolution

which allows demanding a current with a high power factor, certainly not a unity.

Applied voltage to the inverter (v_z) when the full bridge diode is not conducting is determined by:

$$v_z = v_{CZ} \quad (8)$$

This voltage is applied during most of the time to the load in a quasi square output voltage (Fig 8, stage 4).

2) When the converter is working at the rest of switching states, which is during t_0 , t_2 , t_3 and t_5 .

Voltage applied to the inverter (v_z) is zero (Fig. 8, stage 1 and 2) and the rectifier is not conducting ($|i_s|=0$).

At this time, inductors LZ are charged with the capacitor CZ voltage. Then the inductor voltage is:

$$v_{LZ} = v_{CZ} \quad (9)$$

The capacitor current is:

$$i_{CZ} = -i_{LZ} \quad (10)$$

These four stages (Fig. 8), mentioned previously, occur at the switching frequency but also a low frequency component is presented, which appears due to the input voltage.

III. Experimental Results

System functionality was experimentally tested by a built prototype, so that, the proposed idea was validated.

Experimental results are shown in Fig. 8 to 12, these were obtained for a 75W lamp (OSRAM S068 SL75W), the switching frequency is 47

KHz, the resonant tank is composed by L_r is 1.3mH and C_r is 9.15nF, the ‘Z’ inductor LZ is 290 μ F and ‘Z’ capacitor CZ is 10 μ F. For generating the control signals a microcontroller (PIC16HV616 of Microchip) was used.

Input voltage and current are shown in Fig. 8, it is easily seen a good agreement with the theoretical and simulated results, the obtained power factor (PF) is 0.98 and the total harmonic distortion (THD) of the input current is 14%. Fig. 9 shows the input current with its respective harmonic content; due to the good PF and THD the converter satisfy the standard specifications.

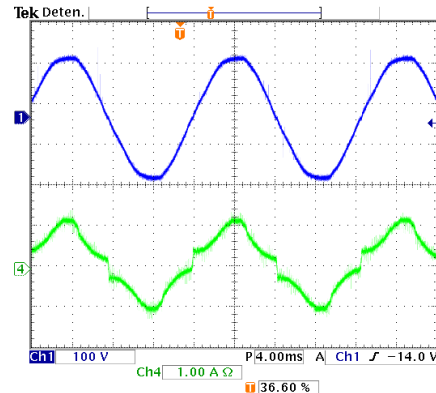


Fig. 8. Input voltage and current (from top to bottom respectively).

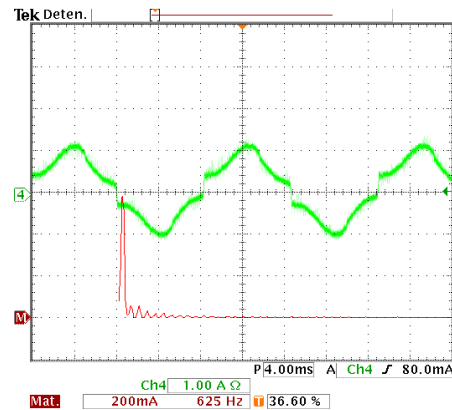


Fig. 9. Harmonic content of the input current. From top to bottom: Input current and harmonic content.

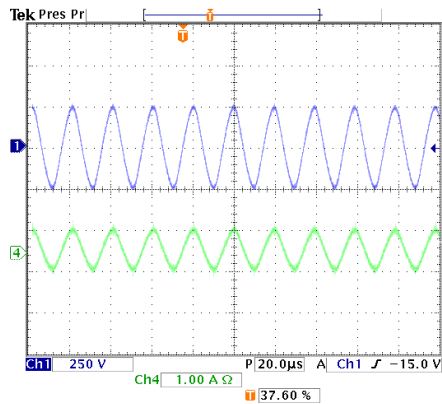


Fig. 10. Voltage and current lamp.

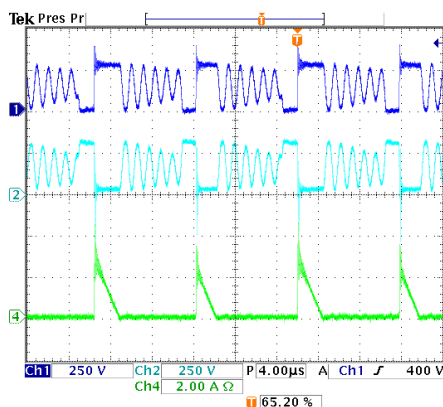


Fig. 11. Inverter input voltage (V_z), rectified ac mains voltage, and rectified ac mains current

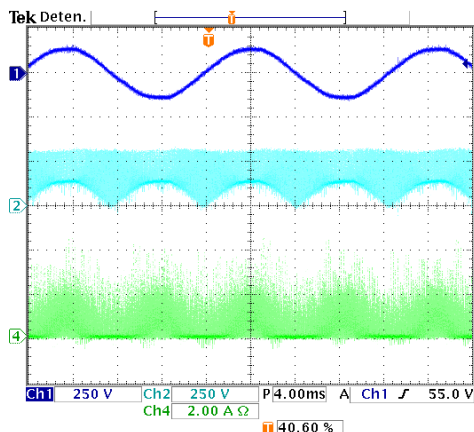


Fig. 12. Input voltage, rectified ac mains voltage, and inductor current (IL_z)

Waveforms regarding to the output are shown in Fig. 10, where current and voltage lamp are

illustrated. It illustrates that the output voltage is sinusoidal in spite of a quasi square output voltage being applied; then the lamp is properly fed.

Converter waveforms at the switching frequency are shown in Fig. 11. Inverter input voltage (V_z), rectified ac mains voltage, and rectified ac mains current are graphed. It is important to notice that an oscillation waveform is observed in (V_z) when the full bridge rectifier is not conducting, at this time the capacitor CZ voltage is applied through the inductor L_z to the load.

Converter waveforms at the input voltage are shown in Fig. 12, but the low frequency component it can be observed. Input voltage, rectified ac mains voltage, and inductor current (IL_z) are illustrated in good agreement with the theoretical analysis.

IV. Conclusions

This paper shows different electronic ballast for fluorescent lamps. Proposal is based on the 'Z' source inverter. Converter offers good input and output characteristics in spite of a single stage being used; it is not an integrated one. A high power factor is obtained and a sinusoidal output voltage is applied to the lamp.

Since a single stage is considered, just one controller is used; having disposed in open loop, it makes for this proposal a simple solution for electronic ballast. The demanded current to the ac mains is similar to that converters based on the active input current shaper techniques, but the topology is completely different.

Operation, analysis, and finally experimental results were presented.

V. References

- [1] J. Ribarich, and J.J. Ribarich. "A New Procedure for High-Frequency Electronic

- Ballast Design". IEEE Transactions on Industry Applications, Vol. 37, No. 1, January-February 2001, pp 262-267.
- [2] J.C. Wo, P.K. Jain. "A Modified Valley Fill Electronic Ballast Having a Current Source Resonant Inverter With Improved Line-Current Total Harmonic Distortion (THD), High Power Factor, and Low Lamp Crest Factor". IEEE Transactions on Industrial Electronics, Vol. 55, No. 3, March 2008, pp 1147-1159.
- [3] J Song, J.H. Song, I. Choy, J.Y. Choi. "Improving Crest Factor of Electronic Ballast-Fed Fluorescent Lamp Current Using Pulse Frequency Modulation". IEEE Transactions on Industrial Electronics, Vol. 48, No. 5, October 2001, pp 1015-1024.
- [4] A.J. Calleja, J. M. Alonso, J. Ribas, E. López, M. Rico, J. Sebastián. "Design and Experimental Results of an Input-Current-Shaper Based Electronic Ballast". IEEE Transactions on Power Electronics, Vol. 18, No. 2, March 2003, pp 547-557.
- [5] R. Nederson, S. Azzolin. "A High-Power-Factor Electronic Ballast Using a Flyback Push-Pull Integrated Converter". IEEE Transactions on Industrial Electronics, Vol. 46, No. 4, August 1999, pp 796-892.
- [6] M. Brumatti, M. Almeida, D.S. L. Simonetti, J. L. F. Vieira. "Single Stage Self-Oscillating HPF Electronic Ballast". IEEE Transactions on Industry Applications, Vol. 41, No. 3, May/June 2005, pp 735-741.
- [7] C.M. Wang. "A Novel Single-Stage High-Power-Factor Electronic Ballast With Symmetrical Half-Bridge Topology". IEEE Transactions on Industrial Electronics, Vol. 55, No. 2, February 2008, pp 969-972
- [8] J.A. Vilela, A. Rodrigues, V.J. Farias, L.C. de Freitas, E.A. Alves, J. Batista. "An Electronic Ballast With High Power Factor and Low Voltage Stress". IEEE Transactions on Industry Applications, Vol. 41, No. 4, July/August 2005, pp 917-926
- [9] C.M. Wang. "A Novel Single-Switch Single-Stage Electronic Ballast With High Input Power Factor". IEEE Transactions on Power Electronics, Vol. 22, No. 3, May 2007, pp 797-803
- [10] Y.C. Chuang and H.L. Cheng. "Single-Stage Single-Switch High-Power-Factor Electronic Ballast for Fluorescent Lamps". IEEE Transactions on Industry Applications, Vol. 43, No. 6, November/December 2007, pp 1434-1440.
- [11] C.S. Moo, K.H. Lee, H.L. Cheng, W.M. Chen. "A Single-Stage High-Power-Factor Electronic Ballast With ZVS Buck-Boost Conversion". IEEE Transactions on Industrial Electronics, Vol. 56, No. 4, April 2009, pp 1136-1146
- [12] C.S. Lin, C.L. Chen. "Single-Switch Electronic Ballast with Continuous Input Current Charge Pump Power-Factor Correction". IEEE Transactions on Industrial Electronics, Vol. 47, No. 6, December 2000, pp 1263-1270
- [13] H.J. Chiu, S.J. Cheng. "Single-Stage Voltage Source Charge-Pump Electronic Ballast With Switched-Capacitor Dimmer for Multiple Fluorescent Lamps". IEEE Transactions on Industrial Electronics, Vol. 54, No. 5, October 2007, pp 2915-2918
- [14] C. Bitencourt, A.J. Perin. "High Power Factor Electronic Ballast for Fluorescent Lamps With Reduced Input Filter and Low Cost of Implementation". IEEE Transactions on Industrial Electronics, Vol. 55, No. 2, February 2008, pp 711-721
- [15] J.M. Alonso, M.A. Dalla, M. Rico-Secades, J. Cardesín, J. García. "Investigation of a New Control Strategy for Electronic Ballasts Based on Variable Inductor". IEEE Transactions on Industrial Electronics, Vol. 55, No. 1, January 2008, pp 3-10
- [16] M.S. Perdigão, J.M. Alonso, M.A. Dalla, E. Sousa. "Comparative Analysis and Experiments of Resonant Tanks for Magnetically Controlled Electronic

- Ballasts". IEEE Transactions on Industrial Electronics, Vol. 55, No. 9, September 2008, pp 3201-3211
- [17] S. Borekci. "Dimming Electronic Ballasts Without Striations". IEEE Transactions on Industrial Electronics, Vol. 56, No. 7, July 2009, pp 2464-2468
- [18] Á. Raniere, F. Ecke, R. Nederson. "A Design Methodology for a Self-Oscillating Electronic Ballast". IEEE Transactions on Industry Application, Vol. 43, No. 6, November/December 2007, pp 1254-1533.
- [19] H.-L. Do, K.-W. Seok and B.-H. Kwon. "Single-stage electronic ballast with unity power factor". IEE Proceeding on Electrical Power Applications, Vol. 448, No. 2, March 2001, pp 171-176.
- [20] C. Ekkaravarodome, Nathakaranakule, A. Boonyaroonate, I. "Single-Stage Electronic Ballast using Class-DE Low-dv/dt Current-Source-Driven Rectifier for Power-Factor Correction", IEEE Trans. Ind. Electron., Vol. 57, No. 10, OCT 2010, pp 3405 – 3414.
- [21] N. Vázquez, J. López, J. Arau, C. Hernández and E. Rodríguez "A Different Approach to Implement an Active Input Current Shaper", IEEE Transactions on Industrial Electronics, Vol. 52, No. 1, January 2005, pp 132-138.
- [22] F. Z. Peng. "Z-Source Inverter". IEEE Transactions on Industry Applications, Vol. 39, No 2, March/April 2003. pp 504-510.
- [23] M. Rashid. "Power Electronics: Circuits, devices, and applications". Third Edition. Ed. Prentice Hall

VI. Authors

Dr. Nimrod Vazquez Nava obtuvo el título de Ingeniero en Electrónica por el Instituto Tecnológico de Celaya en 1994, los grados de Maestro en Ciencias en Ingeniería Electrónica y Doctor en Ciencias en Ingeniería

Electrónica por el cenidet en 1997 y 2003 respectivamente, actualmente es profesor-investigador en el Instituto Tecnológico de Celaya e investigador nivel 1 en el SNI.

Ing. Alejandro Perea Franco obtuvo el título de Ingeniero Electrónico por el Instituto Tecnológico de Celaya y actualmente es estudiante de la Maestría en Ciencias en Ingeniería Electrónica en el mismo Instituto.

M.C. Claudia V Hernández Gutiérrez obtuvo el título de Ingeniero en Electrónica por el Instituto Tecnológico de Celaya en 1995 y el grado de Maestra en Ciencias en Ingeniería Electrónica por el cenidet en 2000, actualmente es docente y Coordinadora de Posgrado en el Instituto Tecnológico de Celaya.

Dr. Jaime Arau Roffiel obtuvo el título de Ingeniero en Electrónica por el Instituto Tecnológico de Minatitlán y el grado de Doctor en Ciencias por la Universidad Politécnica de Madrid, en Madrid, España. Actualmente es profesor-investigador en el Centro Nacional de Investigación y Desarrollo Tecnológico (*cenidet*) e investigador nivel 2 en el SNI.

Balastro Electrónico para Lámparas de HID MH con Control de Intensidad Luminosa

Gabino González Herrera, Elías Rodríguez Segura, René Osorio Sánchez, José Luis Hurtado Chávez

Instituto Tecnológico de Celaya. Departamento de Ingeniería Electrónica. Av. Tecnológico y A. G. Cubas S/N, Col. Alfredo V. Bonfil, Celaya, Gto. (461)6117575 ext. 384, gagohe2@hotmail.com

Resumen

En este artículo se presenta un balastro electrónico para una lámpara de HID de halogenuros metálicos. Se implementó un prototipo para una lámpara de 35 W. Se consiguió un factor de potencia de 0.99, se cumple con la norma IEC 1000-3-2. No existe resonancia acústica y es posible extrapolarlo a mayor potencia. Se muestran resultados experimentales y aspectos críticos de diseño.

Palabras Clave: Balastro electrónico, convertidor cd/cd reductor, corrector del factor de potencia.

I. Introducción

Las lámparas de alta intensidad de descarga HID (por sus siglas en inglés, high intensity discharge) forman una familia que incluye los siguientes tipos de lámparas: mercurio de alta presión, halogenuros metálicos y sodio de alta presión [1]. En lámparas de halogenuros metálicos, la luz se produce mediante la descarga eléctrica en vapor de mercurio de alta presión, lo cual produce una banda continua de radiación, cuenta con un alto índice de rendimiento cromático y muy buena eficiencia. Los balastos electrónicos [2], principalmente formados por un circuito corrector del factor de potencia, un convertidor reductor y un circuito inversor, proporcionan las características necesarias para el encendido y estabilización de la lámpara una vez encendida.

El control de la atenuación de intensidad luminosa en balastos electromagnéticos emplea técnicas de *control de intensidad luminosa continuo* y *operación bi-nivel* en baja frecuencia. La primera consiste en variar el voltaje primario de todo el sistema completo mientras que la segunda lo realiza cambiando la impedancia del balastro mediante la conmutación de capacitores [3]. Actualmente el control de la intensidad luminosa en lámparas de HID se realiza con balastos electrónicos operados en alta frecuencia

[4, 5, 6] y en baja frecuencia [7, 8]. Para el control de intensidad luminosa en alta frecuencia se tienen tres formas: controlando el voltaje de alimentación al inversor, variando el ciclo de trabajo y controlando la frecuencia de conmutación del inversor. Estas técnicas por lo general operan en frecuencias mayores a los 400 KHz, con la finalidad de evitar el fenómeno de las resonancias acústicas. Para el control de intensidad luminosa en baja frecuencia, se controla la corriente o potencia de la lámpara. La frecuencia de operación varía entre 200 – 400 Hz.

En la figura 1 se observa una curva típica característica de voltaje contra corriente de una lámpara de HID bajo condiciones de atenuación luminosa [9].

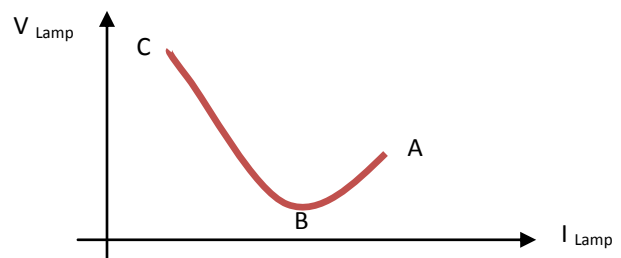


Fig. 1. Voltaje contra corriente de una lámpara de HID bajo condiciones de atenuación luminosa.

En lámparas de HID no se recomienda hacer atenuación de la intensidad luminosa por debajo del 50%, ya que la temperatura en el interior de la

lámpara es un parámetro crítico para el funcionamiento de los ciclos de disociación o recombinación de los halogenuros, lo cual puede reducir la vida útil de la lámpara y además el arco de descarga de la misma puede extinguirse si se excede del 50%.

En este artículo se presenta un balastro electrónico de tres etapas para una lámpara de HID de 35 W de halogenuros metálicos, con la capacidad de controlar la intensidad luminosa en un rango del 100% al 50% (Lm), mediante la variación de la referencia de corriente. Se presentan resultados con respecto a lúmenes, temperatura de color y el índice de rendimiento cromático en condiciones de atenuación luminosa.

II. Desarrollo

El balastro electrónico está conformado por un circuito corrector del factor de potencia, un convertidor cd/cd reductor como limitador de corriente, un circuito inversor de potencia y por ultimo un circuito ignitor externo (ver figura 2). Con el fin de eliminar el problema de resonancia acústica, la lámpara es operada en baja frecuencia. El control de intensidad luminosa se realiza en el convertidor cd/cd reductor.

En el balastro electrónico se implementó un convertidor Flyback corrector del factor de potencia, el cual puede ser alimentado con un rango de tensión universal (90–240 VCA), la frecuencia de operación a la que trabaja es de 50 KHz. El voltaje de salida es de 200 V, suficiente para alimentar el convertidor reductor y la etapa del inversor, así como el voltaje de ruptura del interruptor utilizado en el circuito ignitor para la generación del pulso de alto voltaje necesario para el encendido de la lámpara.

En la figura 3 se puede observar el convertidor Flyback corrector del factor de potencia operando en modo de conducción frontera, empleando el circuito de control L6561.

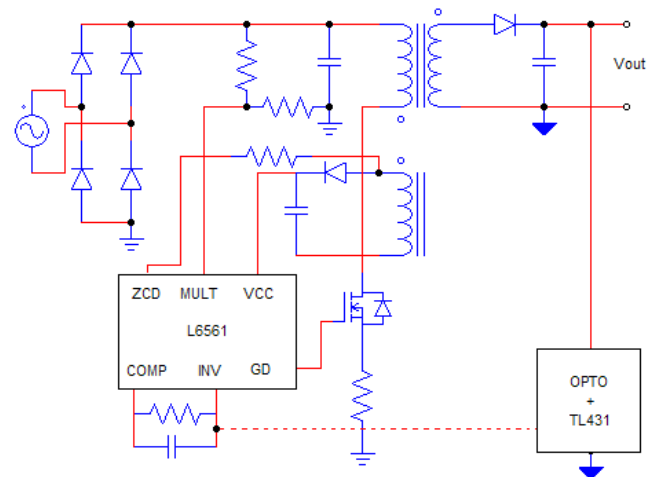


Fig. 3. Convertidor Flyback CFP

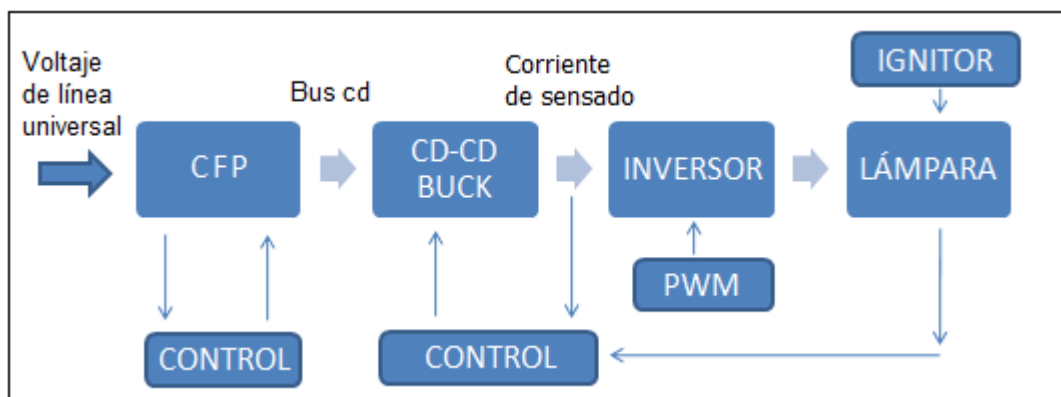


Fig. 2. Diagrama a bloques del balastro de tres etapas.

Los balastos electrónicos requieren de un circuito limitador de corriente para el óptimo funcionamiento de la lámpara. Cuando la descarga de gases ha sido iniciada y el arco de descarga no se ha estabilizado, es necesario un tiempo de calentamiento. Durante el tiempo de calentamiento la temperatura del gas aumenta, lo que implica que la potencia de la luz incrementa. El voltaje de la lámpara comienza aproximadamente desde un cuarto del voltaje de la lámpara. Durante esta fase la corriente máxima debe ser limitada y mantenida un 30% más alta del valor nominal. Las lámparas de HID tienen un tiempo de calentamiento que puede variar entre los 2 y 5 minutos. Después de la fase de calentamiento, cuando la lámpara alcanza su potencia nominal, la corriente de la lámpara debe ser regulada tomando en cuenta el voltaje de ésta para asegurar la potencia de la lámpara.

Para garantizar las condiciones previamente comentadas, se hace uso de un circuito convertidor reductor (buck). La mayoría de los circuitos limitadores de corriente, son alimentados con un voltaje superior al voltaje nominal de la lámpara.

Para el diseño del convertidor cd-cd reductor, se tomaron en cuenta las especificaciones eléctricas de la lámpara MASTER Colour CDM-T 35W/830 G12 1CT de Philips. En la figura 4, se observa el circuito convertidor cd/cd reductor implementado en el balastro.

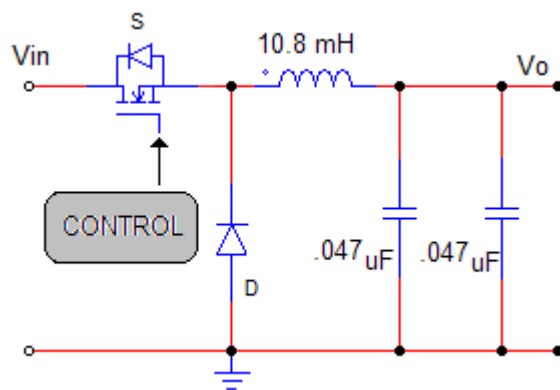


Fig. 4. Convertidor cd/cd reductor

En la tabla 1, se pueden observar las características eléctricas de la lámpara, mientras que en la tabla 2 se encuentran las características más relevantes de la fuente de luz. En la tabla 3, se observan las especificaciones del convertidor cd-cd reductor diseñado.

Características Eléctricas	Valor
Potencia de la lámpara estimada	35W
Potencia de lámpara con balastro electrónico.	38 W
Voltaje de la lámpara	88 V
Corriente lámpara con balastro electrónico	0.47 A

Tabla 1. Características eléctricas de la lámpara.

Características de Fuente de Luz	
Código de color	[CCT of 3000K]
Índice Reproducción Cromática	81 Ra8
Designación de color	Blanco Cálido
Temperatura de color	3095 K
Flujo luminoso con balastro electrónico	3300 Lm
Eficacia luminosa con balastro electrónico	83 Lm/W

Tabla 2. Características de fuente de luz de la lámpara.

Voltaje entrada	V_i	200 V
Voltaje salida	V_o	88 V
Potencia salida	P_o	40 W
Rizo de corriente	Δi_L	10 % de I_o
Rizo de voltaje	ΔV_c	0.1 % de V_o
Frecuencia de conmutación	F_s	50 KHz

Tabla 3. Especificaciones del convertidor cd-cd reductor.

c) Inversor

La principal función de un inversor es cambiar un voltaje de entrada de corriente continua a un voltaje simétrico de salida de corriente alterna, con la magnitud y frecuencia deseada.

Para la implementación de este balastro se utilizó un inversor puente completo, el control del inversor se realizó de manera digital, por lo que fue necesario realizar un algoritmo que fuera capaz de generar dos señales complementarias a una frecuencia de 400 Hz. En la figura 5 se observa la configuración del inversor puente completo.

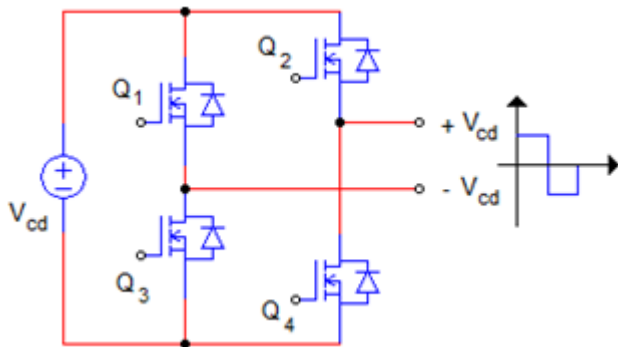


Fig. 5. Inversor puente completo.

d) Ignitor

El ignitor es un dispositivo utilizado para encender lámparas de alta intensidad de descarga. El ignitor es el encargado de proporcionar los pulsos de alta tensión necesarios para iniciar el proceso de encendido de la lámpara. Debe poseer características especiales, como es la de proporcionar una tensión mayor a los 3 KV. La figura 6 muestra el circuito ignitor implementado en el balastro.

Es importante mencionar que un aspecto crítico en la implementación del ignitor es el transformador, ya que no toda la energía almacenada en el capacitor es transferida a los electrodos de la lámpara, debido a que existen pérdidas de energía significativas entre las capacitancias parásitas de los inter-devanados.

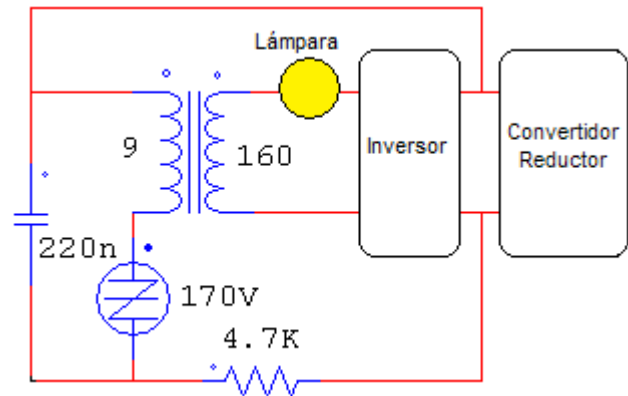


Fig. 6. Circuito ignitor.

e) Control de la intensidad luminosa

El control de la intensidad luminosa se realiza de manera directa dentro del circuito convertidor cd-cd reductor mediante la variación de la corriente de sentido con un control analógico, el cual consiste en tomar una muestra de la corriente de salida que sirve como referencia para el lazo de control. La atenuación de la intensidad luminosa se realiza con la ayuda de un control de 0 a 10 V comercial (ver figura 7).

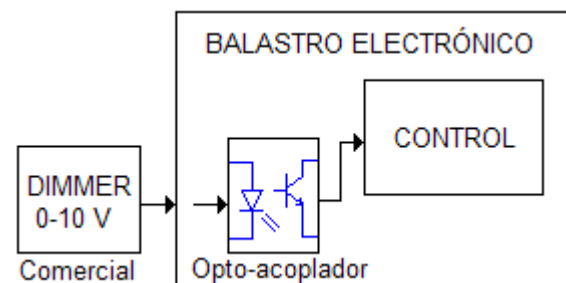


Fig. 7. Interfaz del control de 0-10V comercial, con el balastro electrónico.

III. Resultados

La figura 8, presenta las formas de onda de corriente y voltaje de entrada del circuito corrector del factor de potencia conectado directo a la línea de 120 V de C.A. Se aprecia una corriente con forma de onda senoidal en fase con el voltaje de entrada, se obtuvo un factor de potencia de 0.99 y cumple con la norma IEC 1000-3-2 clase A.

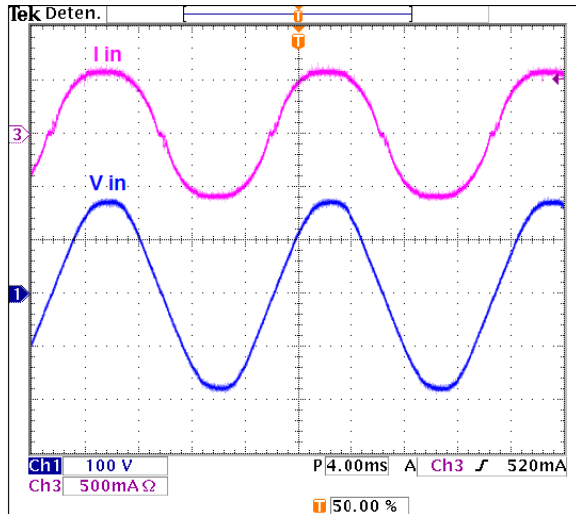


Fig. 8. Forma de onda del voltaje y corriente del CFP.

En la figura 9, se pueden observar pulsos de voltaje de 3.4 KV provenientes del circuito ignitor, necesarios para el encendido de la lámpara. La figura 10 presenta el proceso de ignición del sistema, en ella se puede observar que una vez generado el arco de la lámpara, la corriente tiende a crecer de manera significativa, de ahí la importancia de contar con un buen circuito de control que limite y establezca la corriente de la lámpara una vez encendida. De lo contrario la corriente continúa creciendo hasta extinguir el arco de la lámpara y en el peor de los casos dañar la lámpara.

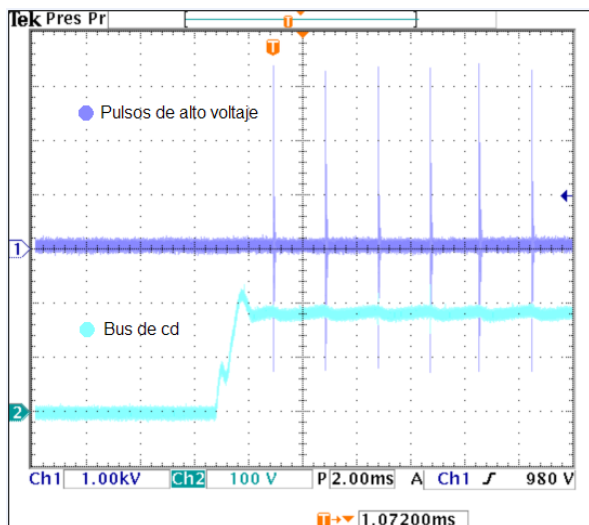


Fig. 9. Pulsos del ignitor y bus de cd

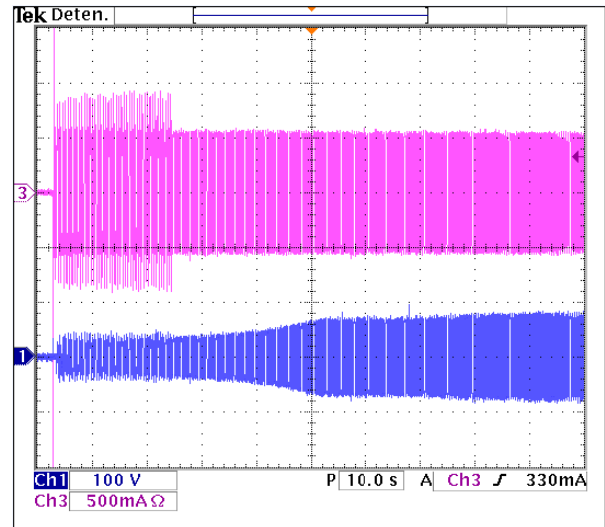
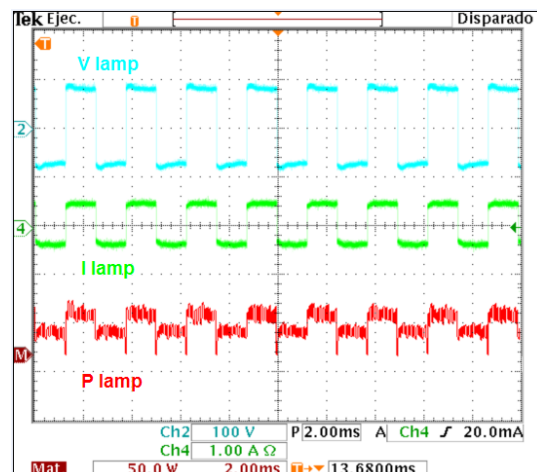


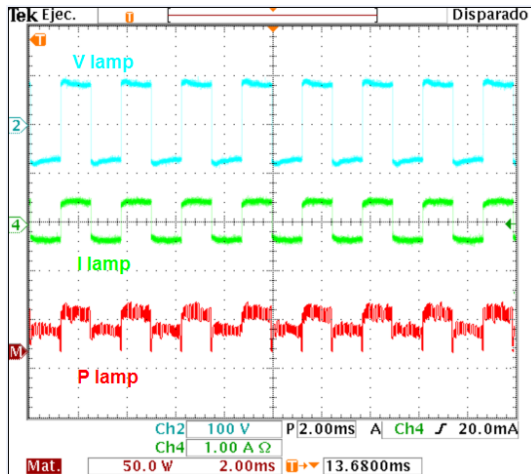
Fig. 10. Proceso de encendido de la lámpara.

En las gráficas de la figura 11, se pueden observar las formas de onda de la corriente, voltaje y potencia de la lámpara bajo el proceso de atenuación de la intensidad luminosa.

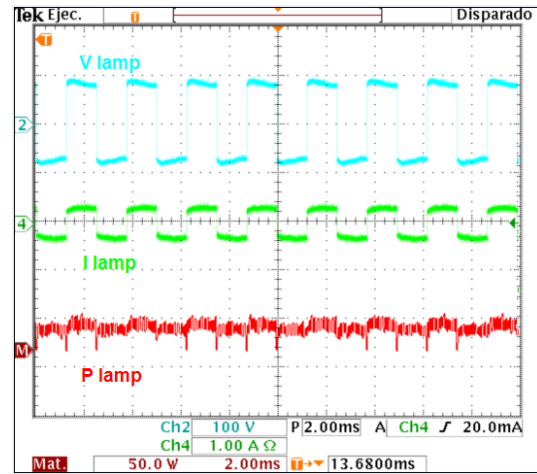
Estas gráficas se tomaron cada 15 minutos con cambios de referencia de corriente de 30 mA aproximadamente. La medición de los luxes se realizó con la ayuda de un luxómetro dentro de una caja oscura con volumen de un metro cúbico.



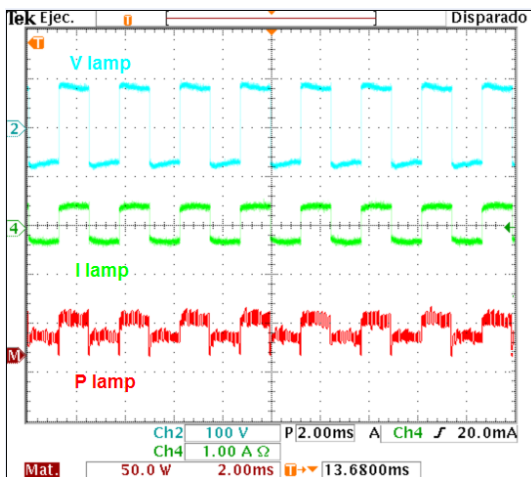
a)



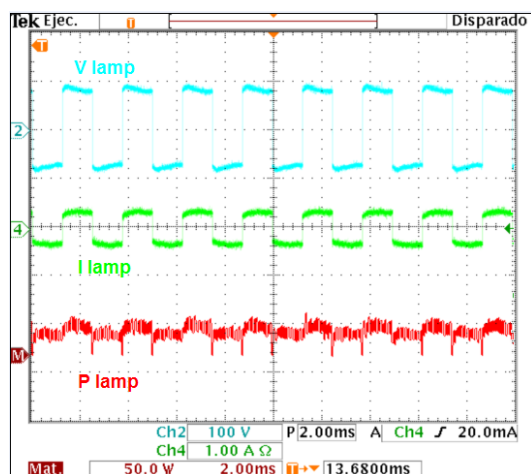
b)



e)



c)



d)

Fig. 11. Voltaje, corriente y potencia de la lámpara bajo condiciones de atenuación de intensidad luminosa. Luxes correspondientes a cada cambio de referencia. A) 79 V, 425 mA, 33.11 W con 458 Luxes b) 79.4 V, 397 mA, 31.06 W con 418 Luxes c) 78.7 V, 369 mA, 28.34 W con 374 Luxes d) 80.1 V, 326 mA, 25.69 W con 336 Luxes e) 79.9 V, 303 mA, 24.09 W con 302 Luxes

Se realizaron mediciones de lúmenes, temperatura de color y rendimiento cromático empleando una esfera integradora (Labsphere's Light Measurement Integrating Spheres).

En la figura 12, se observa como los lúmenes descienden al disminuir la corriente de la lámpara. En esta gráfica se puede apreciar que con una corriente de 300 mA, se alcanza el 50% del flujo luminoso.

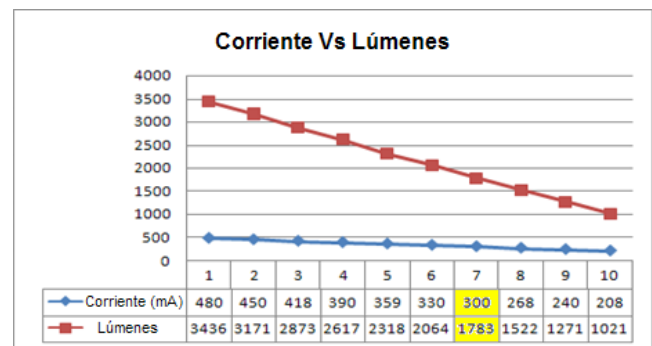


Fig. 12. Corriente vs Lúmenes.

En la gráfica de la figura 13, se puede observar como la temperatura de color aumenta cuando descende la corriente de la lámpara. Teniendo el 50% del flujo luminoso la temperatura de color es de 4061 °K

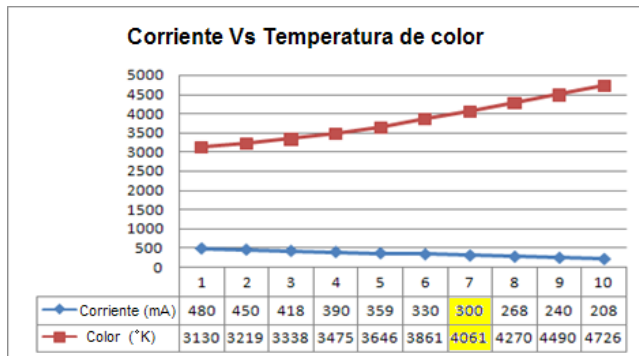


Fig. 13. Corriente Vs Temperatura de color.

El índice de rendimiento de color (IRC) se define como la capacidad que tiene una fuente de luz para reproducir correctamente los colores de los cuerpos iluminados. En la figura 14, se puede observar que durante el proceso de atenuación de la intensidad luminosa, se registraron pérdidas de 17 unidades de IRC para quedar en 64.12, lo cual es aceptable si se hace una comparación con las lámparas de vapor de sodio la cuales no alcanzan los 60 IRC en condiciones normales.

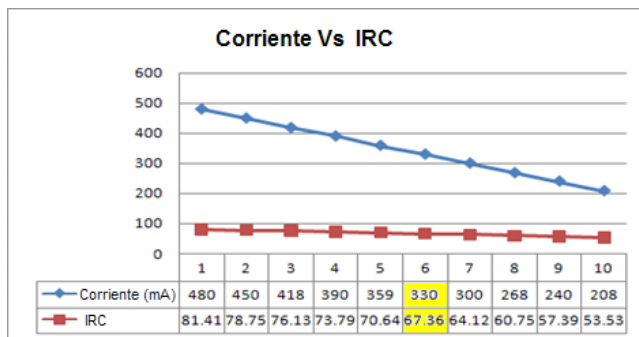


Fig. 14. Corriente Vs IRC.

En la tabla 4, se observan diferentes parámetros bajo condiciones de atenuación de la intensidad luminosa.

PF	I lamp	V lamp	P lamp	Luxes
.996	490 mA	80.2 V	38.4 W	570
.996	460 mA	79.9 V	35.2 W	504
.996	426 mA	78.9 V	33.1 W	458
.996	398 mA	79.3 V	31.2 W	418
.996	369 mA	79.7 V	28.41 W	374
.994	325 mA	80 V	25.7 W	336
.995	305 mA	81.3 V	24 W	302
.993	267 mA	82.6 V	21.74 W	253

Tabla 4. Balastro de tres etapas en funcionamiento, bajo condiciones de AIL.

En la figura 15 se muestran dos fotografías de una manzana, tomadas con diferentes niveles de intensidad luminosa. La primera fotografía corresponde al 100%, mientras que la segunda representa el 50% de la intensidad luminosa.



a) b)

Fig. 15. Fotografías con diferentes niveles de intensidad luminosa. a) 100%, b) 50%

IV. Conclusiones

Se implementó un balastro electrónico para lámparas de HID MH de 35 W con control de intensidad luminosa. El prototipo implementado presenta un alto factor de potencia, buena estabilidad y un excelente control de intensidad luminosa. No hay resonancia acústica y se logró

controlar la atenuación de intensidad luminosa en un rango de 100 al 50%.

Se realizaron mediciones de algunos parámetros importantes tales como: lúmenes, temperatura del color y el índice de recombinación cromático, bajo condiciones de atenuación de intensidad luminosa. Los lúmenes alcanzados por el balastro fueron de 3436 lm operando al 100%, una vez iniciado el proceso de atenuación este parámetro fue disminuyendo. El 50% del flujo luminoso (1783 lm) se alcanzó con una corriente de 300 mA. Por otro lado, la temperatura del color inició en 3130 °K y a medida que la corriente de lámpara fue disminuyendo, la temperatura de color fue aumentando. Con flujo luminoso del 50%, este parámetro alcanzó los 4061 °K. El índice de rendimiento cromático alcanzado por la lámpara en condiciones nominales fue de 81.41 con una excelente tonalidad como caracteriza a las lámparas de halogenuros metálicos.

Mediante la incorporación del control luminoso en lámparas de alta intensidad de descarga de halogenuros metálicos y de acuerdo a los resultados obtenidos se puede alcanzar un ahorro de energía alrededor del 32%, tomando en cuenta que el 50% de flujo luminoso, según los parámetros obtenidos, equivalen al 68% de la potencia nominal de la lámpara.

V. Referencias

- [1] J. Correa. “Estrategias de Control en Lámparas de Alta Presión para la Eliminación de Resonancias Acústicas”. Tesis Doctoral, CENIDET, Junio 2003.
- [2] Dalla Costa, M.A. Alonso, J.M. Marchesan, T.B. Cervi, M. Prado, R.N., “ Electronic Ballast for HID Lamps”, IEEE Industry Applications Magazine, vol. 17, issue 2, p 54. 2007
- [3] Chung, H.S.-H. Ngai-Man Ho Wei Yan Pok Wai Tam Hui, S.Y., “Comparison of Dimmable Electromagnetic and Electronic Ballast System – An Assessment on Energy Efficiency and Lifetime”, IEEE Transactions on Industrial Electronics, vol. 54, issue 6, p. 3145. 2007.
- [4] X. Cao, W. Yan, S. Y. Hui, and H. S. H. Chung, “Dimming Control Characteristics of High-Frequency Operated Metal Halide Lamps”, IEEE Trans. Power Electronic, vol. 19, n0. 3, pp. 854-861, May 2004.
- [5] Huang-Jen, Hsiu-Ming Huang, Yu-Kang Lo, and Chia-Hsing Li. “High-Frequency Dimmable electronic Ballast for HID Lamps” IEEE Trans. Power Electronic, vol. 21, n0. 5, September 2006.
- [6] Wei Yan. Hui, S.Y, “An analysis Into the Dimming Control and Characteristic of Discharge Lamps”, Power Electronics, IEEE Transactions on, Nov 2005, p 1432-1440.
- [7] Y.-T. Huang, S.T. Chen, C.-R. Lee, and L.-L- Lee. “Designs and implementation of the Dimmable Electronic Ballast for Metal Halide Lamps”, The 33rd Annual Conference of the IEEE Industrial Electronics Society (IECON), Nov. 5-8, 2007. Taipei Taiwan.
- [8] Stankovic, A. Nerone, L. Kulkarni, P. “Modified Synchronous Buck Converter for Dimmable HID Electronic Ballast” Industrial Electronics, IEEE Transactions on, vol. pp, issue 99, 2011.
- [9] Wei Yan. Hui, S.Y, “An analysis Into the Dimming Control and Characteristic of Discharge Lamps”, Power Electronics, IEEE Transactions on, Nov 2005, p 1432-1440.

Análisis y reducción de emisión de ruido en una máquina cortadora de plasma X Y

Hilda Olvera Hernández, Francisco Javier López Jáquez

Instituto de Ingeniería y Tecnología, departamento de Ingeniería Industrial y Manufactura
Av. Del Charro 450 Nte. Colonia Partido Romero C.P. 32310, Ciudad Juárez, Chih., México
olvera.hilda@gmail.com, frlopez@uacj.mx.
Tel: 656 6884843, frlopez@uacj.mx

Resumen

Este trabajo trata sobre la reducción del ruido de una máquina de corte XY. El objetivo fue que la diferencia entre el nivel de ruido y el ruido de fondo sea lo más cercana posible a 0dB. Para el diseño del concepto se analizaron las fuentes de ruido y la dirección del ruido. Se enumeraron las posibles soluciones y por medio de una matriz de Pugh se seleccionó la opción de un cerramiento acrílico, esta opción se implementó y se recolectaron datos para hacer un análisis el cual resultó en una reducción de decibeles. Se instaló corcho para sellar los orificios, ya que es comúnmente utilizado para el aislamiento acústico en general. El compresor fue una de las fuentes de ruido que también fue necesario atacar. Por lo tanto se decidió crear un cerramiento acústico, se trabajó con lámina y material fibroso aislante de ruido. La reducción obtenida fue de aproximadamente el 19%. Se presenta la implementación de todos los diseños que se propusieron así como las pruebas para verificar la eficacia de las propuestas, las cuales arrojaron resultados favorables.

Palabras Clave: Estadística, Decibeles, Ruido.

I. Introducción

En este caso el objeto de estudio es una máquina cortadora de plasma XY. El ruido producido por la máquina XY sobrepasa el nivel de ruido máximo permisible en un salón de clases y en consecuencia resulta molesto para las personas que se encuentran contiguas a esta.

Por tal motivo, fue necesaria una alternativa de reducción de ruido para que la máquina cumpla con los niveles admisibles y pueda ser utilizada sin ninguna reserva.

Cuando las ondas sonoras chocan con una superficie una parte de la energía se refleja, otra se transmite y otra es absorbida por la superficie [1,2]. Si la energía absorbida y transmitida son bajas se trata de una superficie reflectante, acústicamente dura. Esto indica que para problemas de control de ruido se debe tener presente que este tipo de superficies incrementa el nivel de presión sonora en el interior [3].

Los niveles usualmente recomendados para un aula son de 40 a 45 decibeles, según las normas

básicas de edificación NBE [4]. Sin embargo, se midió un salón con 25 alumnos en un estado de calma y el nivel de decibeles osciló entre 50 y 55. Es normal que los valores sean diferentes ya que este tipo de normas muestran valores óptimos a alcanzar, obtenidos bajo condiciones controladas y con un mínimo o casi nulo ruido de fondo. Por lo tanto, en la vida real estos valores pueden variar. Evidentemente el nivel de ruido es mucho mayor que el valor óptimo sugerido.

Para poder analizar el ruido emitido por la máquina se utilizó un decibelímetro Simpson 898-2. Se determinó el número mínimo de posiciones del micrófono que fueron 3 y su localización a 1m, 2m, 4m, 6m, y así sucesivamente. Se calibró el instrumento, se posicionó a una altura de 1.2 metros, con una distancia para tomar la lectura de 0.5 metros y una orientación del micrófono de 70° hacia la fuente de ruido [5,6]

Derivada de la hoja de recolección de datos se tiene el diagrama direccional que muestra gráficamente la distribución de la presión sonora.

Las superficies de contacto tales como las paredes, muebles y puertas son altamente reflejantes, dado que cuando las mediciones se hacen cerca de este tipo de superficies el nivel de presión sonora es mayor. El nivel promedio de ondas sonoras fue 72dB, en este caso.

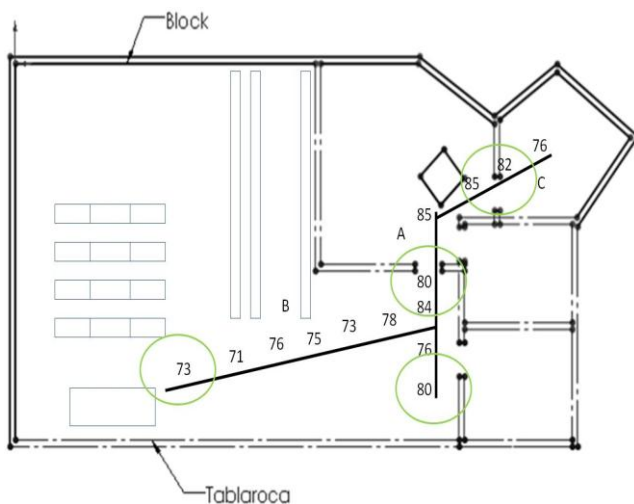


Fig. 1. Diagrama direccional del ruido para la máquina de corte X Y.

II. Revisión de Literatura

Las máquinas convencionales y las no convencionales producen cierto nivel de ruido, este puede ser un factor no deseado que se pretende eliminar, pero en otros casos puede ser un factor de análisis para el monitoreo o control del proceso de maquinado y elementos como la vibración. Desarrollos actuales e investigaciones en este campo proponen métodos para el monitoreo del proceso de maquinado mediante la emisión acústica, determinación de la estabilidad del maquinado mediante el análisis de las frecuencias y amplitudes del sonido, etc. En contraste con otras investigaciones que lo que buscan principalmente no es el análisis si no la reducción del ruido.

Iulian Marinescu [7] propone un nuevo método de supervisión de corte para procesos con cortadores de múltiples dientes simultáneos, tomando en cuenta patrones de señales. La nueva técnica de monitoreo está basada en la

combinación del análisis de tiempo-frecuencia de la emisión de ondas sonoras usando técnicas avanzadas de procesamiento como el análisis Formant, entre otros.

Primero muestra que es posible identificar las condiciones de corte (corte con uno dos o tres dientes) usando solo la emisión acústica, también muestra que es posible la detección de anomalías en la superficie. Aplicando los diferentes tiempo-frecuencia se encontró que un patrón específico de frecuencia puede ser identificado para ayudar a la detección de una fase específica de corte (entrada o salida del cortador).

Por su parte Gullem Quintana [8] tuvo como objetivo determinar el diagrama de lóbulo de estabilidad de un proceso de fresado aplicando la metodología de mapeo de sonido, que es una representación gráfica de los niveles de sonido, los diagramas de lóbulo de estabilidad muestran la frontera de estabilidad como una combinación de la velocidad de rotación de usillo y la profundidad del corte. Una plataforma de adquisición de datos fue desarrollada para recolectar el sonido del procesó de maquinado. En el sonido de fresado se analizan frecuencias y amplitudes esto produce un buen resultado y permite obtener una aproximación precisa para entender la incidencia de la vibración en el proceso de fresado.

Un mapa del sonido 3D fue hecho mediante la grafica de cada punto correspondiente en una malla descrita sobre la amplitud del sonido y la frecuencia donde se muestran las diferencias entre estable e inestable. Mostrando el mapa de fresado 3D es posible detectar regiones donde la señal del sonido sube, esto permite distinguir regiones de condiciones de cortes estables e inestables e identificarlas en un diagrama de lóbulo de estabilidad. Es posible seleccionar los parámetros apropiados de corte para asegurar operaciones libres de fracturas o hendiduras.

En contraste J. Hesselbach [9] Desarrollo un sistemas de sujeción por vacío para la reducción de ruido y vibración, los sistemas tradicionales usados para el maquinado de tableros son deficientes. Para encontrar si la fuente acústica

dominante está localizada en el área de contacto de la herramienta o en el área de la superficie de la pieza y la remoción de viruta, varias medidas y análisis fueron hechos con una cámara acústica que es un sistema de medición que usa la formación de haz como una técnica de procesamiento de señal para visualmente localizar la emisión acústica y se concluye que el ruido es generado mayormente por la pieza de trabajo y no por el contacto de la herramienta.

Se desarrolló un sistema donde se adaptó al método de sujeción por vacío un actuador con resorte, en total cuatro sujetadores y además cuatro sensores para monitorear la vibración. Este experimento arrojó que mediante el uso de este nuevo método de sujeción se reduce el ruido y la vibración.

III. Desarrollo

Se seleccionó de todas las posibles fuentes, la que emitía el mayor nivel de decibeles que era la antorcha, ya que al momento del corte, la boquilla de la antorcha está orientada hacia el suelo, el aire y el plasma salen por la boquilla y cortan la lamina, todas estas ondas sonoras se impactan contra el piso que es tipo mármol y contra las paredes de la máquina que son de aluminio lo que causaba un efecto de reflexión (ver Fig. 2).

En base al análisis de la situación se generaron una serie de posibles soluciones, tales como: 1) cama de sacrificio para corte; 2) mesa de agua; 3) cerramiento parcial; 4) instalar material absorbente de sonido en habitación y techo y 5) aislamiento de antorcha y área de corte. Por medio de una matriz de Pugh (ver Tab.1) se le dio peso a cada una de ellas. La alternativa más factible fue aislar la antorcha y el área de corte para, evitar la propagación de las ondas sonoras.

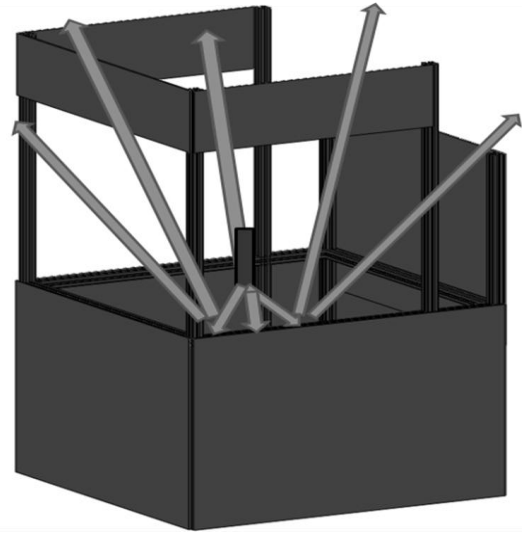


Fig. 2. Dispersión del ruido en máquina XY.

		evaluación de posible solución				
Atributos		1	2	3	4	5
1	Fácil construcción	0	-1	1	-1	0
2	Efectividad	1	0	1	1	1
3	Tempo de implementación	0	-1	-1	-1	1
4	Costos	1	1	0	-1	1
total		2	-1	1	-2	3

Tabla 1. Matriz de decisión de Pugh.

Se construyó el cerramiento acrílico y se hizo un análisis en software de computadora minitab (ver Tabla 2) para comprobar si la media global del nivel de decibeles de la máquina en condiciones iniciales era la misma media global de la máquina con el cerramiento de acrílico, en esta prueba se observó que con el cerramiento del acrílico se disminuyeron las ondas sonoras emitidas por el área de corte, dejando ahora expuesta otra fuente de ruido importante que es el compresor, fue evidente que el solo accionamiento del compresor generaba un nivel de ondas sonoras elevado.

Para analizar si efectivamente el nivel de ruido del compresor solo es mayor que el nivel de ruido normal existente en aéreas como esta (50dB), se hizo una prueba de hipótesis (ver Tabla 3), y se comprobó que efectivamente este nivel de ruido es mayor que el nivel deseado.

Por lo tanto se decidió reubicarlo en el cubículo de enseguida para observar que efecto tenía esto en la disminución de las ondas sonoras.

Paired T-Test and CI: inicial, acrilico				
Paired T for inicial - acrilico				
	N	Mean	StDev	SE Mean
inicial	112	78.093	4.460	0.421
acrilico	112	74.739	4.996	0.472
Difference	112	3.354	2.828	0.267
95% lower bound for mean difference: 2.911				
T-Test of mean difference = 0 (vs > 0): T-Value = 12.55 P-Value = 0.000				

Tabla 2. Prueba t pareada para diferencia de medias.

One-Sample T: compresor						
Test of mu = 50 vs > 50						
			95% Lower			
Variable	N	Mean	StDev	SE Mean	Bound	T P
compresor	14	67.314	3.176	0.849	65.810	20.40 0.000

Tabla 3. Prueba t para el nivel de ruido del compresor.

Adicionalmente se instaló corcho en los orificios y área de alimentación de material. Comúnmente el corcho es utilizado para el aislamiento acústico en general, el índice de reducción sonora preestablecido para este tipo de material es de 43 dB, además es resistente al fuego lo cual es una característica muy favorecedora debido a que en el área de corte generalmente se forman chispas que podrían ocasionar fuego espontaneo.

Se hizo un análisis estadístico para probar si la media de el nivel de ruido de la máquina es mayor con el solo cerramiento del acrílico y es menor añadiendo la instalación del corcho y cambiando el compresor de lugar; se determinó que es posible decir que el nivel sonoro con estos cambios es menor que el anterior, dicha diferencia es de 5.8dB.

Suplementariamente se hizo una prueba para analizar si el nivel de ruido disminuía al cerrar la

primera puerta que se encuentra antes de la máquina XY y se observó que si existe diferencia entre estas dos condiciones aparentemente al tener la puerta cerrada el nivel de decibeles promedio es 1.2 dB menor que al tener la puerta abierta (ver Tabla 4), esto con un nivel de significancia de 0.05.

Paired T-Test and CI: corcho, puerta cerrada				
Paired T for corcho - puerta cerrada				
	N	Mean	StDev	SE Mean
corcho	72	67.410	3.497	0.412
puerta cerrada	72	66.167	3.509	0.413
Difference	72	1.243	1.986	0.234
95% lower bound for mean difference: 0.853				
T-Test of mean difference = 0 (vs > 0): T-Value = 5.31 P-Value = 0.000				

Tabla 4. Prueba t pareada para diferencia de medias.

Uno de los primero ajustes que fue necesario hacer fue facilitar el acceso al área de corte por lo tanto una de las paredes del cerramiento acrílico fue rediseñada para agregar una puerta de acceso (ver Fig. 3a). Otro ajuste tuvo que ver con el compresor, ya que en si es una fuente de ruido. Se decidió aislar el ruido por medio de la interposición de una barrera física. En esta ocasión se decidió trabajar con lamina que es un material reflejante y para complementar se agrego un material fibroso que dé un efecto de absorción al interior del cerramiento (ver Fig. 3b). Para el material fibroso se utilizó un aislante de fibra de vidrio, este es un buen material de absorción con un índice de reducción acústica mínimo de 36dB en la clasificación STC y máximo de 63dB. Esta fibra de vidrio inorgánica, tiene la capacidad de absorber la vibración, es muy ligera y no absorbe humedad.

Cuando el ensamble terminó se metió al compresor dentro del cerramiento acústico y se reubicó en el lugar que se encontraba originalmente (ver Fig. 4).

Finalmente se hicieron pruebas para medir el nivel de ruido, y poder comparar los resultados anteriores contra los resultados actuales, de esta

forma se puede saber si el construir un cerramiento para el compresor ayudó a disminuir el nivel de ruido.



Fig. 3. a) Puerta de fácil acceso. b) Cerramiento metálico.

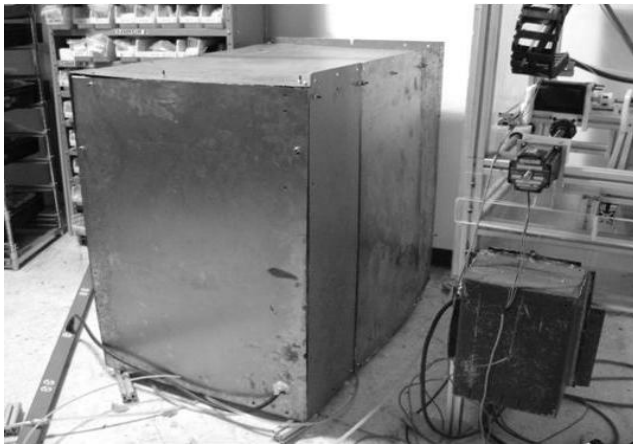


Fig. 4. Cerramiento metálico

Se analizaron los datos por medio de Minitab y fue posible concluir que en nivel de decibeles emitidos por el cerramiento mientras la máquina trabaja es menor que el anterior aproximadamente en 2.056dB.

IV. Resultados

Gradualmente se fue reduciendo el nivel sonoro en todas las corridas de prueba (ver Fig. 5), es importante señalar que durante la prueba final hubo lecturas altas en algunos puntos debido a los momentos en que el corte iniciaba y/o paraba.

En el cerramiento para el compresor el resultado fue primero que el peso del cerramiento era bastante considerable debido a su tamaño, y la fibra de vidrio cambio el sonido que las láminas de metal hacían al momento de ser golpeadas.

El ruido verdadero de la máquina, según cálculos es de 58.2dB y comparado con los 72dB iniciales, se redujo el ruido en 13.8 unidades, es decir aproximadamente una reducción del 19%.

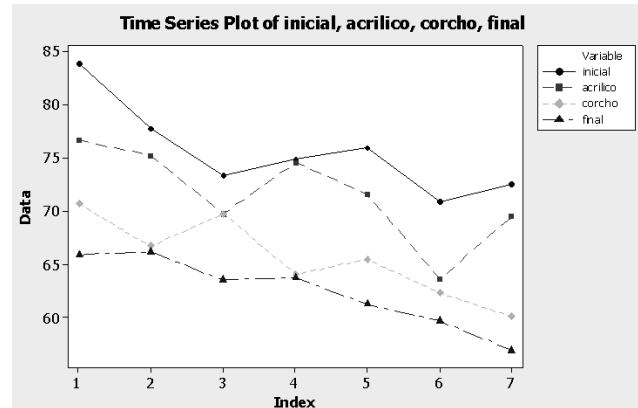


Fig. 5. Gráfica comparativa.

V. Discusión

El nivel de ruido propuesto como objetivo fue el nivel de ruido más bajo detectado en un salón de clases que es 50dB, este mismo nivel fue usado como nivel de ruido de fondo.

Se tienen dos factores muy importantes que ayudan a la propagación del ruido, el primero es que las paredes no son completas por lo cual las ondas sonoras sobrepasan esas barreras existentes y se propagan por todo el recinto, el segundo factor es que las los elementos de construcción cuentan con un bajo coeficiente de absorción, eso significa que el ruido no es absorbido si no que es reflejado, Las superficies de contacto tales como las paredes, muebles y puertas son altamente reflejantes. Los problemas de ruidos excesivos se deben tratar combinando absorción con aislamiento.

Con el objetivo planteado fue necesario hacer el análisis de las fuentes de ruido, la máquina XY tiene varios componentes que emiten un nivel importante de ruido como lo son, el compresor, el generador de plasma y la antorcha.

El propósito de un cerramiento completo es contener y absorber la energía acústica irradiada por la fuente. Esto reduce los niveles de presión sonora para cualquier distancia de la fuente [6].

Para el cerramiento se utilizó acrílico, este material tiene ciertas ventajas tales como ser liviano, lo cual es muy importante ya que la máquina puede no soportar mucho el peso de todo el cerramiento, claridad y transmisión de la luz, teóricamente las ondas sonoras serán aisladas por el acrílico.

Las hojas de acrílico son unas buenas reductoras de sonido, para una hoja con un espesor de 1/2", el índice de reducción acústica esperado es de 33dB [3].

La caja aislante (ver Fig. 6), tiene una altura adecuada que no afecta la función de la oruga por la cual pasa el cable de la antorcha y contiene toda el área de corte al igual que un motor que se encuentra en la parte inferior de la máquina, se instalaron soportes para la base de la caja que ayudan a fijar la firmemente a la máquina.

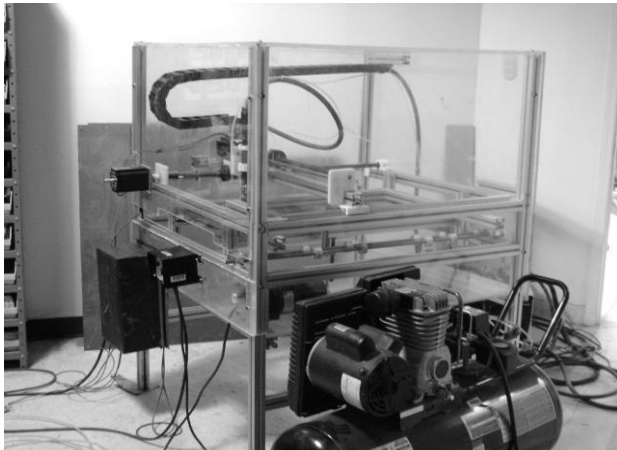


Fig. 6. Máquina XY con cerramiento acrílico.

La prueba t pareada fue utilizada a lo largo de esta investigación ya que es la que más se ajustaba a las comparaciones requeridas de análisis de la diferencia ó el efecto que tuvo cada cambio en cada zona de medición que se seleccionó, no una diferencia promedio que se pudiese obtener con una prueba z.

El diagrama direccional también fue una herramienta muy útil ya que este ayuda a ver de una manera gráfica la distribución de las ondas sonoras [3].

VI. Conclusiones

Comparando la media contra el valor meta de 50dB por medio de una prueba de hipótesis t, se tiene que la media es diferente y el intervalo de confianza muestra que el valor real promedio del sonido debe estar alrededor de 58-59db (ver Tabla 5) por lo tanto es posible concluir que el nivel de ondas sonora, aunque no es el esperado, disminuyó notoriamente.

Con los datos se hizo un diagrama direccional para conocer la dispersión final de las ondas sonoras, también se analizaron con el objetivo de saber el nivel de ruido verdadero de la máquina, el cual según el resultado es de 58.2dB, 8.2dB por encima de lo esperado, y comparado con los 72dB iniciales, se redujo el ruido en 13.8 unidades, es decir aproximadamente una reducción del 19%.

One-Sample T: ruido

Test of $\mu = 50$ vs not = 50

Variable	N	Mean	StDev	SE Mean	95% CI	T	P
ruido	8	58.625	0.744	0.263	(58.003, 59.247)	32.79	0.000

Tabla 5. Cerramiento de compresor prueba t.

Este proyecto sentó antecedentes para continuar con la reducción de ruido emitido por la máquina XY.

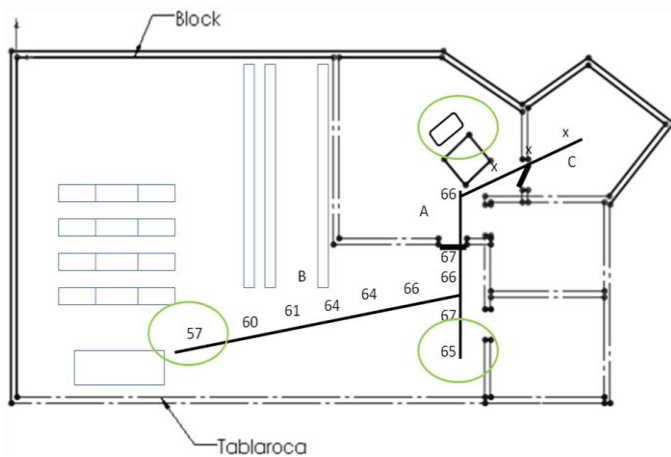


Fig. 7. Diagrama direccional del ruido para el cerramiento del compresor.

Se implementaron todos los cambios diseñados con buenos resultados, y se generaron propuestas para mejoras en ambos diseños.

Todo esto con el mismo objetivo inicial, la reducción de sonido de la máquina XY, para que esta pueda ser usada sin restricciones.

Referencias

- [1] Garcia, Armando. La contaminación acústica. Valencia. 1988, Cultura universitaria popular.
- [2] Perez, Juan M. Ochoa. Medida y control del ruido. Barcelona. 1990, Marcombo boixareu editors.
- [3] Harris, Cyril M. Handbook of acoustical measurement and noise control. New York. 1995, Mc Graw Hill.
- [4] J.R. hassall, M.Sc. Acoustic Noise Measurements. 1988, Bruel & Kjaer.
- [5] South, Tim. Managing noise and vibration at work. London. 2004, Elsevier.
- [6] Behar, Alberto. El ruido y su control. México, D.F. 1994 Trillas.
- [7] Iulian Marinescu. A time-frequency acoustic emission-based monitoring technique to identify workpiece surface malfunctions in milling whit multiple teeth cutting simultaneously. International Journal of Machine Tools & Manufacture, 49, 2009, 53–65

[8] Guillem Quintana. Sound mapping for identification of stability lobe diagrams in milling processes. International Journal of Machine Tools & Manufacture, 51, 2011, 363–376.

[9] J. Hesselbach. Development of an active clamping system for noise and vibration reduction. CIRP Annals - Manufacturing Technology, 59, 2010, 395–398 .

VII. Autores

Ing. Hilda Olvera Hernández es egresada de la carrera de Ingeniería en Manufactura por la Universidad Autónoma de Ciudad Juárez y actualmente se encuentra en el proceso de titulación para obtener el grado de M.I. en el área de manufactura opción diseño del producto, en la misma universidad.

Dr. Francisco Javier López Jáquez obtuvo su título de Doctorado en filosofía en el área de Manufactura por la Universidad de Liverpool en el Reino Unido, Inglaterra. Actualmente está adscrito como profesor investigador al departamento de Ingeniería Industrial y Manufactura de la Universidad Autónoma de Ciudad Juárez.

Comentarios finales:

Es necesario analizar la emisión de ruido en la antorcha ya que, al inicio y al final del corte el ruido es mayor, probablemente el mayor ruido producido por la máquina será el que se tiene que tomar en cuenta como base para la reducción de la emisión de ruido.

Separar las paredes de acrílico de los soportes de la máquina utilizando corcho ayudaría a seguir reduciendo el ruido; también la instalación de soportes transversales o longitudinales en la parte superior del cerramiento es necesaria ya que por el propio peso la lámina de acrílico, se tiende a pandear.

Poner una capa de fibra de vidrio a cada uno de los cerramientos acrílicos, es también una opción para disminuir el ruido transmitido a través de las paredes del acrílico.

Por último la utilización de un decibelímetro que cuente con ponderación A es importante para que las mediciones sean más apegadas al nivel de ondas sonoras que capta el oído humano.

Sistema de Adquisición de Datos Automatizado para Pruebas Mecánicas Uniaxial y Triaxial

***Fernando Rivas Cruz, Enrique Contreras López**

*Instituto de Investigaciones Eléctricas, Reforma 113 Cuernavaca, Morelos, 777 3623811,
fernando.rivas@iie.org.mx.

Resumen

Se presenta el desarrollo de un sistema de monitoreo para realizar las pruebas mecánicas uniaxial y triaxial aplicados a muestras de roca, por ejemplo petroleras o mineras. El desarrollo permite conocer algunas de las características de la roca para evaluar la explotación de los recursos energéticos y minerales del subsuelo. El sistema muestra el monitoreo de las variables que intervienen en el proceso de pruebas mediante una interfaz gráfica y amigable para el usuario. Las pruebas se desarrollan mediante equipos especializados para simular las condiciones del subsuelo en el que se encuentra inmersa la muestra de roca. Los resultados se presentan en línea y con reportes detallados de las deformaciones y otros parámetros.

Palabras Clave: Automatización, adquisición de datos, control de fuerza, esfuerzo mecánico, monitoreo,

I. Introducción

El conocimiento de las propiedades termofísicas, mecánicas, eléctricas y de transporte de las rocas es un aspecto importante para muchas aplicaciones relacionadas con actividades de localización, evaluación y explotación de los recursos energéticos y minerales del subsuelo, así como para otras áreas de interés científico y tecnológico que involucran el estudio de eventos en los que las rocas de la corteza terrestre actúan como elementos estructurales recipientes de contención y medios de transporte de fluidos y de energía mecánica y calorífica [1].

El conocimiento de las propiedades de las rocas tienen varias áreas de aplicación como: en la industria petrolera, la geotermia, la minería, la construcción de túneles y presas, la selección y adecuación de sitios para depositar desechos tóxicos y radiactivos, por mencionar algunos ejemplos.

La localización, la evaluación y el aprovechamiento de los recursos energéticos y minerales del subsuelo dependen en gran medida del conocimiento de las propiedades de las rocas que lo conforman.

En una diversidad de sistemas subterráneos profundos como los yacimientos geotérmicos y petroleros, las rocas se encuentran sometidas a condiciones ambientales peculiares, alta temperatura, esfuerzos litostáticos y tectónicos, presión de poro e interacciones entre los fluidos de poro y roca. Estos parámetros ambientales pueden afectar en forma considerable las propiedades de las rocas [2]. La obtención de estos datos en sistemas subterráneos profundos sólo es posible a partir de mediciones en muestras de roca específica del sitio de interés.

En el Departamento de Geotermia del Instituto de Investigaciones Eléctricas (IIE) se cuenta con un laboratorio especializado, donde se obtiene la información del comportamiento de las rocas en sistemas subterráneos profundos relacionados con las áreas de aplicación antes mencionadas. Además, se realiza la medición de propiedades mecánicas, térmicas, eléctricas y de transporte de fluidos en muestras de roca sometidos a condiciones de alta presión y temperatura [1].

Algunas otras pruebas y mediciones que se realizan en el laboratorio de yacimientos del IIE son: compresión hidrostática, compactación,

consolidación, medición de la permeabilidad absoluta, medición de la conductividad, medición del calor específico, medición de la resistividad eléctrica, deformación uniaxial, deformación triaxial, por mencionar algunas.

Los trabajos desarrollados con respecto a pruebas del tipo mecánico (uniaxial y triaxial) se mencionan a continuación:

El procedimiento de pruebas triaxiales son aplicados de una forma convencional en suelos agrícolas no saturadas para determinar la influencia de la succión del suelo en la carga y el comportamiento del cambio de volumen [3]. En [4] se propone un nuevo método para medir la relación esfuerzo-deformación en masas de roca. Las notas técnicas de aplicación se obtienen a partir de diferentes composiciones de roca. Se aplican pruebas mecánicas.

Con respecto a sistemas de monitoreo en [5] se presenta el desarrollo de un programa de pruebas de tipo virtual, en donde se realizan experimentos para complementar estudios y cursos de las características de rocas. En [6] se rediseña un sistema de cómputo para pruebas del tipo triaxial que permite analizar el comportamiento de la relación esfuerzo-deformación para evitar los problemas encontrados durante y después de la construcción de estructuras.

La originalidad de este trabajo, consiste en el desarrollo de un sistema capaz de monitorear y almacenar las variables implicadas en el proceso de pruebas mecánicas del tipo uniaxial y triaxial, según sea las características de la roca, así como la generación de reportes detallados durante y al final de la prueba. La aplicación se desarrolló con la herramienta de software comercial LabVIEW®.

II. Desarrollo

2. Propiedades mecánicas de las rocas

La determinación de las propiedades mecánicas de la roca al punto de falla es parte de la especialidad de la “mecánica de rocas”, incluye el estudio de la deformación de las rocas, por lo

que es necesario conocer los conceptos fundamentales como: esfuerzo y deformación para evaluar las propiedades de la roca.

2.1 Esfuerzo

Comúnmente denotado por σ , se refiere a la fuerza aplicada a la roca que tiende a cambiar de dimensión. La fuerza aplicada normalmente se refiere como carga:

$$\sigma = \frac{F}{A} \quad (1)$$

Donde σ es el esfuerzo ($\text{Pa} = 1 \text{ N/m}^2$), F es la fuerza en N y A es el área de la muestra en m^2 . Las tres condiciones básicas de esfuerzo son: a) tensión, b) compresión y c) esfuerzo cortante, ver Fig. 1.

2.2 Deformación

El efecto del esfuerzo aplicado a la roca es parte del estudio para medir la deformación producida por la aplicación de una carga. La deformación, comúnmente denotada por ϵ , es el cambio relativo en la forma o el tamaño de la roca debido a la fuerza aplicada. En otras palabras se mide la deformación del material cuando se le aplica una carga [7], ver Fig. 2.

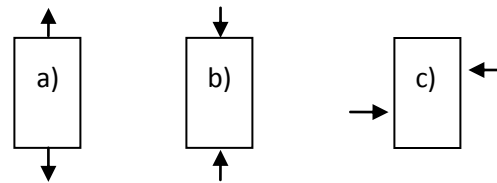


Fig. 1. Representación esquemática de esfuerzos

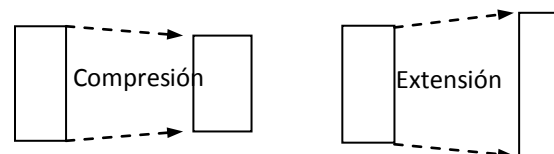


Fig. 2. Deformación

Por ejemplo, considerar un núcleo de una longitud original L_0 que está sujeta a esfuerzo de

tensión. Después de aplicar la carga, la longitud original se incrementa a L , por lo que la deformación axial está definida como:

$$\varepsilon = \frac{L - L_0}{L} \quad (2)$$

Un núcleo es una muestra continua de roca, que se obtiene de la formación mediante perforaciones con una barrena especial hueca, la cual es capaz de cortar tramos de roca de hasta 20 m de longitud con un diámetro máximo de 20 cm [2].

2.3 Fallas mecánicas (relación esfuerzo-deformación)

Algunos materiales al aplicarle un esfuerzo en un tiempo prolongado permite que la deformación se incremente, esta deformación es proporcional al esfuerzo aplicado. Si se retira la carga, la deformación del material regresa a su estado inicial cero. A este fenómeno se le llama “deformación elástica”. Si el esfuerzo continúa incrementando, alcanza el punto de saturación o de falla; la fuerza de la roca se supera y la deformación es permanente. A este fenómeno se le llama “deformación plástica”, ver Fig. 3.

En la Fig. 4 se muestra la aplicación de un esfuerzo de compresión (par de pistones) y un esfuerzo de diferente magnitud aplicado a la circunferencia del cilindro (presión de confinamiento). Si la presión radial (confinamiento) es cero, y solo se aplica la compresión se denomina “Prueba de esfuerzo uniaxial”. Cuando a la muestra se le realiza una prueba con presión de confinamiento diferente a cero y se aplica compresión, esta prueba se le denomina “Prueba de esfuerzo triaxial”. Con este ensayo es posible obtener una gran variedad de estados reales de carga.

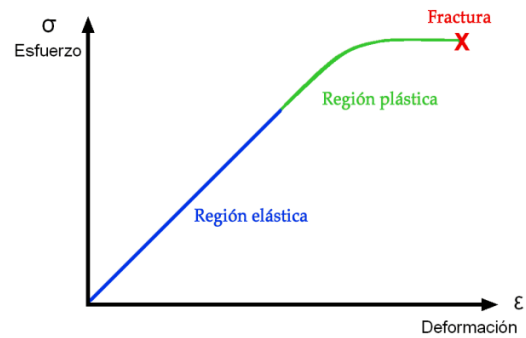


Fig. 3. Relación esfuerzo-deformación.

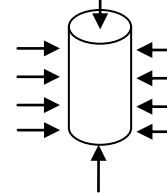


Fig. 4. Pruebas uniaxial y triaxial de muestras de roca de 38 mm de diámetro.

Una prueba triaxial usualmente se realiza al incrementar la carga axial (compresión) y la presión de confinamiento se mantiene constante, hasta que el nivel de esfuerzo hidrostático se alcance. La carga axial normalmente se aplica tal, que se llegue a la condición de ruptura o deformación de la muestra. En la Fig. 3 se muestra la relación esfuerzo-deformación en una prueba uniaxial.

2.4 Descripción del proceso de pruebas

El Proceso de Pruebas Mecánicas (PPM) está diseñado con la capacidad de generar la carga aplicada a la roca, la presión de confinamiento, el monitoreo, almacenamiento y el procesamiento de los datos en los ensayos de compresión uniaxial y triaxial ejercidas sobre muestras de núcleos de rocas.

Ensayo de compresión Uniaxial

Para el ensayo del tipo uniaxial la roca se coloca en el equipo, del cual aplica una carga de compresión, hasta llevarla a la fractura. Se adquieren los datos de carga, deformación axial y

transversal mediante sensores *LVDTs* (*Linear variable differential transformer*) conectados en un arreglo mostrado en la Fig. 5. Un *LVDT* es un dispositivo sensor que mide el desplazamiento lineal para transformarlo en una señal eléctrica en milivolts. Cada uno de los sensores se conecta a la etapa de acondicionamiento de señal, tarjeta de adquisición de datos y a la computadora de adquisición de datos.

Ensayo de compresión Triaxial

Para el caso de este tipo de pruebas, la muestra de roca se deposita en una camisa, donde se le aplica presión de confinamiento, se mantiene a una presión determinada y se le aplica la carga de compresión. Esta carga se aplica hasta que la roca llega a su fractura. El arreglo de esta prueba se muestra en la Fig. 6.

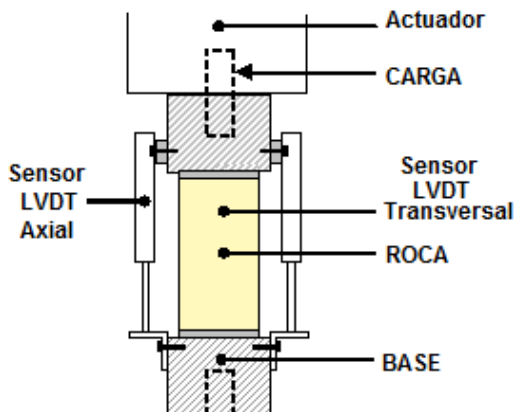


Fig. 5. Arreglo de ensayo Uniaxial.

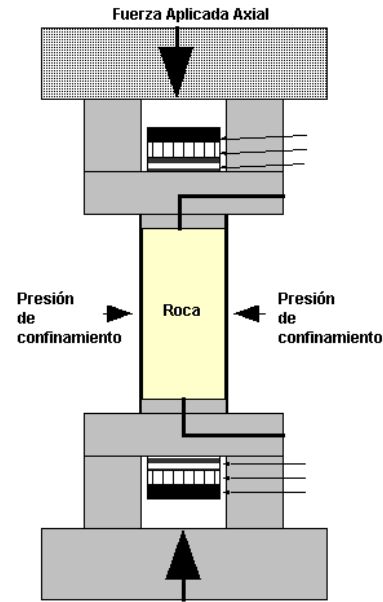


Fig. 6. Arreglo de ensayo Triaxial.

3. Sistema de adquisición de datos

Se utilizó un equipo experimental configurado alrededor de una máquina de compresión para pruebas mecánicas, equipada con un sistema de control manual (no servocontrolado) para regular la rapidez de aplicación de la carga, con capacidad para aplicar niveles de carga axial compresiva de hasta 150 Ton (aproximadamente 150,000 kilogramos), ver Fig. 7.

La muestra se instrumenta con los sensores adecuados, para medir las deformaciones axiales en función de la carga aplicada. La deformación axial se mide con un dispositivo que utiliza como elementos sensores, dos transformadores diferenciales linealmente variables (LVDT), con rango de medición nominal de $\pm 0.100''$, resolución infinita y respuesta altamente lineal dentro de su rango de medición nominal.

La deformación transversal se mide en la parte media de la muestra en una dirección diametral, empleando un dispositivo sensor tipo cantiléver instrumentado con extensómetros eléctricos. La deformación de la muestra en la dirección radial propicia que en el sensor se genere un voltaje de

salida que es linealmente proporcional a la magnitud de dicha deformación.



Fig. 7. Sistema de control manual no servocontrolado.

Todos los sensores primarios (LVDTs) que se utilizaron para medir las deformaciones axial y transversal fueron calibrados mediante la aplicación de valores conocidos de desplazamiento o de deformación que se generaron y midieron con micrómetros de alta exactitud, los que a su vez están calibrados por un laboratorio acreditado contra un instrumento patrón. En la Fig. 7 se muestra un anillo dinamométrico integrado con un indicador de caratula con el que se calibra la celda de carga, la cual aplica la fuerza necesaria para deformar la muestra de roca instrumentada de aproximadamente 1.5 pulgadas de diámetro con una longitud de 4 pulgadas de longitud.

La calibración de la celda de carga se realiza con base a un sistema patrón, el cual indica la relación de desviaciones micrométricas-carga aplicada. La relación de desviaciones se correlaciona con la carga aplicada y el voltaje del sensor de presión como se puede ver en la Fig. 8. La calibración indica que a 0 volts de corriente continua no hay carga aplicada y aproximadamente 10 volts equivale a 40000 kg fuerza.

Los sensores de desplazamiento diferencial (LVDTs) se calibran mediante el uso de micrómetros; la relación del desplazamiento en milímetros (mm) genera una señal de voltaje de corriente directa de 0-10 Volt al aplicar los acondicionadores de señal, ver Fig 9.

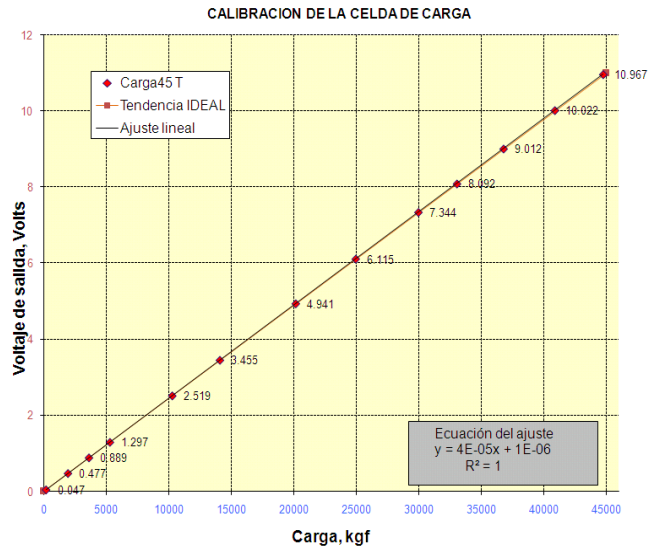


Fig. 8. Calibración de la celda de carga.

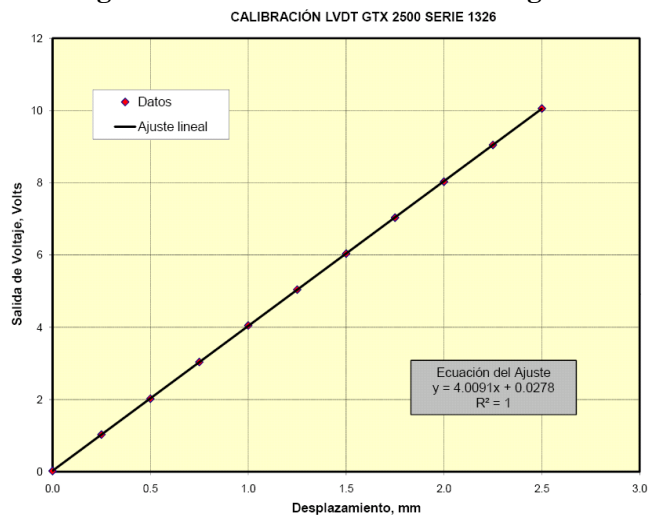


Fig. 9. Calibración de los sensores LVDT.

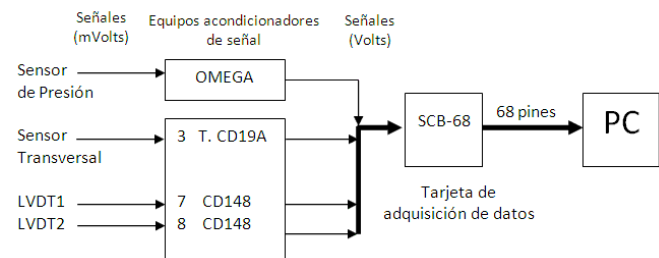


Fig. 10. Diagrama a bloques del sistema de adquisición de datos.

Los sensores previamente calibrados se conectan a la electrónica necesaria para generar las señales de desplazamiento y de carga (señales de voltaje de corriente directa, 0- 10 volts). En la Fig. 10 se muestra el diagrama a bloques de la conexión entre sensores, acondicionadores de señal y tarjeta de adquisición de datos conectados a la computadora.

4. Desarrollo del sistema de adquisición de datos

Para el control de la SPM (Sistema de pruebas mecánicas), se requiere de una programación que proporcione una interfaz gráfica donde el usuario puede seguir la evolución de las variables sensadas en línea. LabVIEW® es el entorno de programación gráfica capaz de cumplir con estas necesidades además de realizar las tareas de adquisición, análisis y presentación de los datos.

En el diseño del sistema de monitoreo se tienen cuatro principales operaciones a realizar:

a) Aplicación de control de botones y barra de menú; b) Aplicación leer datos de la instrumentación, c) Aplicación almacenar los datos de todas las variables y d) Aplicación de ejecución de acciones.

A partir de estas consideraciones se genera un diagrama de flujo principal encargado de monitorear los dispositivos encargados del proceso de pruebas mecánicas [8]. En la Fig. 11 se muestra el diagrama de flujo principal, y las operaciones que se ejecutan simultáneamente utilizando ciclos *while*. A continuación se explican los subdiagramas principales en el desarrollo del sistema de pruebas mecánicas aplicadas a muestras de núcleos de roca.

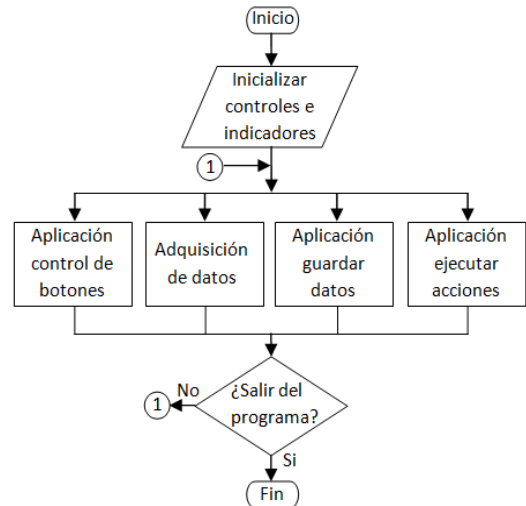


Fig. 11. Programa principal SPM.

Inicializar controles e indicadores.

Este subprograma se encarga de asignar un valor igual a cero a los controles e indicadores de la pantalla principal del sistema de monitoreo antes de iniciar la prueba. Se incluye la rutina para indicar los parámetros de: tipo de prueba, muestra, diámetro, longitud, T muestreo.

Aplicación control de botones.

Estructura de casos que permite llamar a la acción indicada en la barra de menú o botones.

Adquisición de datos.

Consiste en la ejecución de un ciclo que adquiere los datos de cada uno de los elementos que constituyen al SPM en tiempo real.

Aplicación ejecutar acciones.

En este subprograma se ejecuta las aplicaciones dadas desde la barra de menú como son: graficación en tiempo real, tabla de datos, etc. Estos procesos se realizan en paralelo.

Aplicación guardar datos.

Almacena los datos para generar el reporte en formato Excel. El tiempo de muestreo es variable mediante el selector del tiempo de adquisición de datos. Se almacenan tres tipos de archivos:

1. Archivo de especificaciones del tipo de prueba y las características de la roca (diámetro, longitud, etc.).
2. Archivo de adquisición de datos de las variables de prueba (en unidades de ingeniería), el formato es en hoja de cálculo. El nombre del archivo se especifica el año, mes, día y nombre la muestra.
3. Archivo de respaldo de la adquisición de datos en voltajes.

III. Resultados

Con el objetivo de verificar la funcionalidad de cada aplicación del sistema de pruebas mecánicas y con base a las normas establecidas para las pruebas uniaxiales “*Standard test method for elastic modul of intact rock core specimens in uniaxial compression*” y para las pruebas triaxiales “*Standard test method for triaxial compressive strength of undrained rock core specimens without pore pressure measurement*”, se realizan pruebas de la puesta en marcha del SPM a muestras de rocas.

La Fig. 12 muestra la interfaz del sistema de monitoreo SPM. La pantalla principal contiene los elementos como son almacenamiento de datos, diámetro y longitud de la muestra, cálculo el área, nombre de la muestra (dadas por el cliente), tipo de muestra, tiempo de muestreo (variable).

Los resultados son: la etapa de graficación que permite al usuario conocer el comportamiento de la deformación axial y transversal. En la Fig. 13 muestra el efecto de la presión sobre la resistencia a la compresión y sobre el comportamiento general del esfuerzo-deformación con compresión uniaxial. Se aprecia que al incrementarse la presión la deformación axial y transversal se incrementan.

En la Fig. 14 muestra la ventana de adquisición de datos en tiempo real, se muestra en una tabla los datos de tiempo de adquisición, voltajes, carga, esfuerzo, por mencionar algunos.

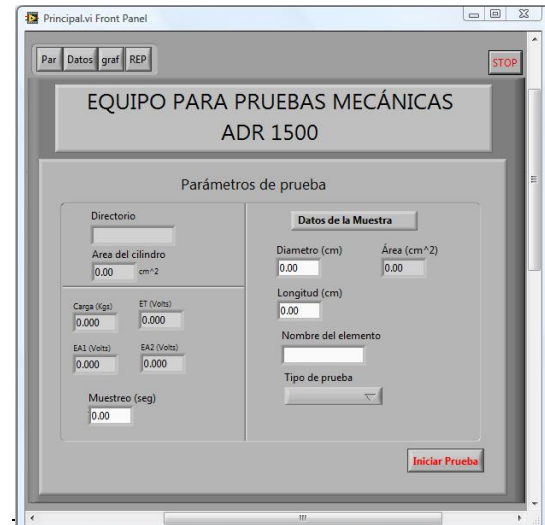


Fig. 12. Programa principal SPM.

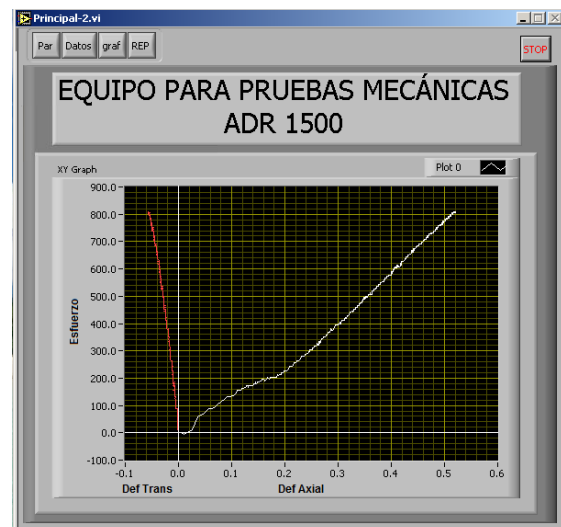


Fig. 13. Graficación en línea de las deformaciones axial y transversal.

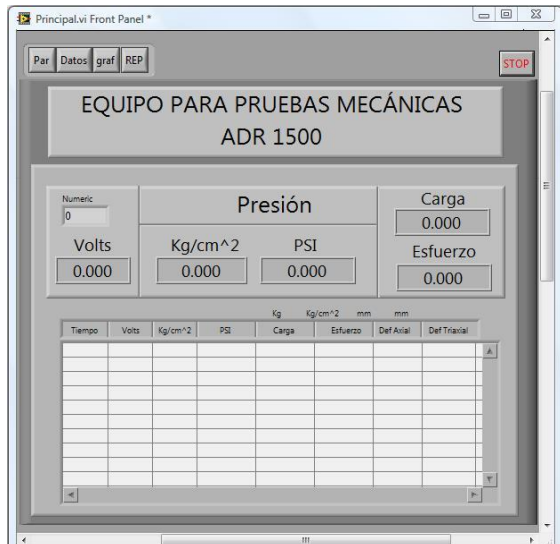


Fig. 14. Datos en línea.

5. Generación de Reportes

Para las pruebas de compresión uniaxial, se genera en el Sistema de Pruebas Mecánicas la gráfica en tiempo real del comportamiento de la muestra al aplicarle presión a la misma. Se genera adicionalmente el reporte a detalle, ver la Fig. 15, con la finalidad de entregar al cliente detalles de las pruebas realizadas. Para este caso algunos datos adicionales en el reporte son: longitud de la muestra, diámetro, área, esfuerzo hasta fractura de la roca donde termina la prueba y el nombre de la muestra.

Otra característica del reporte final es la generación de las gráficas de los Círculos de Mohr. Es una técnica usada en ingeniería y geofísica para representar gráficamente un tensor simétrico y calcular con ella momentos de inercia, deformaciones y tensiones, adaptando los mismos a las características de una circunferencia (radio, centro, etc.). También es posible el cálculo del esfuerzo cortante máximo absoluto y la deformación máxima absoluta. En la Fig. 16 se muestra el resultado de graficar los Círculos de Mohr a diferentes muestras de roca del mismo grupo litológico. Entiéndase por grupo litológico al total de muestras obtenidas de un núcleo de perforación, en la cual el núcleo tiene las mismas características como son: tamaño de grano, partículas, características físicas y químicas.

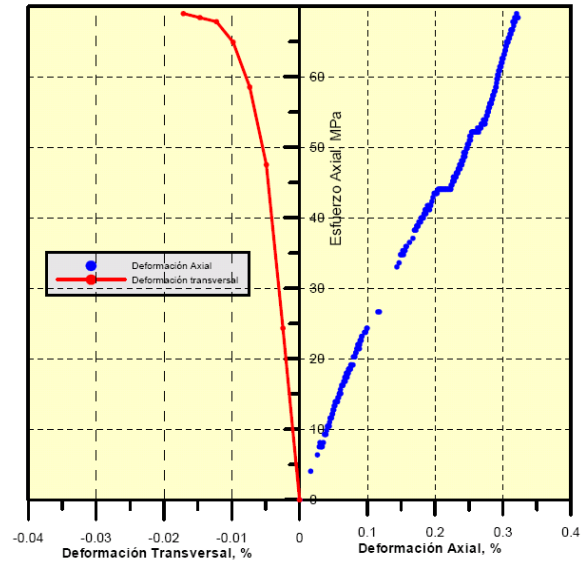


Fig. 15. Datos en línea de las deformaciones.

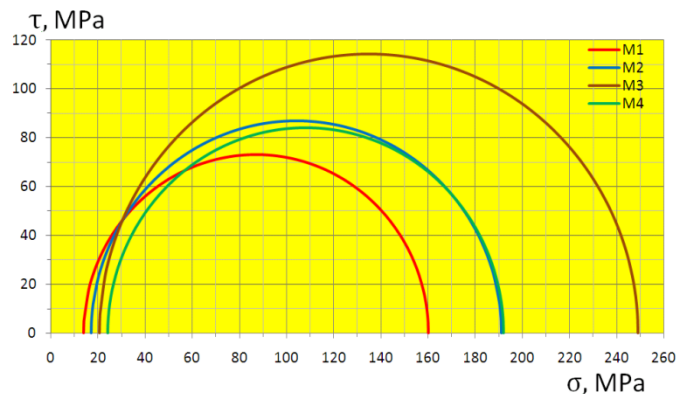


Fig. 16. Resultados y graficación de Círculos de Mohr.

IV. Conclusiones

El sistema desarrollado permite monitorear las diversas variables físicas que intervienen en el proceso de pruebas mecánicas. Esta aplicación en tiempo real, tiene la gran ventaja de analizar el comportamiento del proceso, para la toma de decisiones durante la ejecución del mismo, con la finalidad de mejorar el desempeño del proceso y finalmente incrementar el nivel de confiabilidad en las pruebas y los niveles de seguridad, ya que la fuerza aplicada a una roca puede llegar hasta las 45 toneladas.

La importancia del SPM dispone de un sistema de adquisición de datos a bajo costo de

instrumentación y permite de una manera efectiva y flexible, un control óptimo y de gran calidad.

Las pruebas del sistema se verificaron al colocar muestras de roca aplicando diferentes pruebas uniaxiales y triaxiales. El sistema genera reportes detallados de la prueba en formato Excel.

V. Referencias

- [1] E. Contreras López, “Determinación de propiedades mecánicas y térmicas en muestras de roca sometidas a condiciones de alta presión y temperatura”, Sociedad Mexicana de Mecánica de Rocas, Taxco Guerrero, 1993.
- [2] E. Contreras López, “Importancia de las actividades de planificación, corte, manejo y análisis de los núcleos de perforación de pozos petroleros”, Boletín interno del Instituto de Investigaciones Eléctricas IIE. Año 31. Julio-Septiembre 2007. ISSN 0185-0059, 2007.
- [3] D. Wulfsohn, A. Bankole and D. Fredlund, “*Triaxial Testing of Unsaturated Agricultural Soils*”, Journal of Agricultural Engineering Research. Vol. 69., pp 317-330, 1998.
- [4] T. Okada, K. Tani, H. Ootsu, Y. Toyooka, T. Hosono and S. Kaneko, “*Development of in-situ triaxial test for rock masses. International Journal of the JCRM*”. Japanese Committee for Rock Mechanics. Vol. 2, November 2006.
- [5] D. Penumadu, A. Prashant and D. Frost, “*A virtual triaxial compression test simulator*”. 17th ASCE Engineering Mechanics Conference. University of Delaware, Newark, DE, June 2004.
- [6] M. Bilir and Y. Muftuoglu, “*A computer-controlled triaxial swelling test apparatus*”, Turkish Journal Of Engineering & Enviromental Sciencies. ISSN: 1303-6157. Vol. 28, pp. 269-280, 2004.
- [7] A. Dandekar, “*Petroleum reservoir rock and fluid properties*”. Edit Taylor & Francis Group. ISBN: 978-0-8493-3043-8, 2004.
- [8] E. Fjaer et al., “*Petroleum related rock mechanics*”, Edit. Elsevier Science Publishers B.V. Netherlands. ISBN: 0-444-88913-2, 1992.
- [9] F. Rivas Cruz, C. Astorga, R. Longoria, G. Valencia y F. López, “*Desarrollo de una estación de monitoreo y control en línea. Aplicación a una columna de destilación*”. Asociación Mexicana de Control Automático, UNAM, México, D.F, Octubre 2006
- [10] “*Standard test method for triaxial compressive strength of undrained rock core specimens without pore pressure measurement*”, American Society for Testing and Materials, Commitee D18 on Soil and rock. Designation: D 2664-95a. June 2002.
- [11] “*Standard test method for elastic moduli of intact rock core specimens in uniaxial compression*”. American Society for Testing and Materials, Commitee D18 on Soil and rock. Designation: D 3148-02. April 2002.

VI. Autores

M. en C. Fernando Rivas Cruz. Ingeniero en electrónica egresado del Instituto Tecnológico de Orizaba, Veracruz, México, en 2003. Obtuvo el grado de maestría como Ingeniero en Mecatrónica por el Centro Nacional de Investigación y Desarrollo Tecnológico en 2006.

En el 2009 ingresa al Instituto de Investigaciones Eléctricas como investigador en el área de yacimientos de la Gerencia de Geotermia, participando en proyectos de caracterización petrofísica y geomecánica en muestras de núcleos.

Los intereses de su investigación incluyen adquisición de datos, supervisión de sistemas y control, mecatrónica e implementación de los mismos.

M. en I Enrique Contreras López. Ingeniero mecánico egresado de la Facultad de Ingeniería de la Universidad de Guadalajara (1974). Obtuvo la maestría en ingeniería mecánica con especialidad en termociencias en la División de Estudios de Postgrado de la Facultad de Ingeniería de la UNAM (1978).

A su ingreso como investigador al Departamento de Geotermia (1980); se desarrolló en las disciplinas de Petrofísica y Mecánica de Rocas dentro del Área de Ingeniería de Yacimientos. Hasta el 2010 generó extensas bases de datos petrofísicos de los principales campos geotérmicos de México y realizó estudios de caracterización petrofísica de algunos de los más importantes yacimientos petroleros del país.

Realizó también numerosas estancias de capacitación en las disciplinas de petrofísica y mecánica de rocas en diversas compañías e instituciones del extranjero. Impartió cursos de capacitación en El Salvador y Costa Rica.

Diseño y Calibración de un Fotómetro con Interfaz Virtual

Marco Antonio Olvera López, Juan Ricardo Damián Zamacona, José Rufino Díaz Uribe

CCADET/UNAM, circuito exterior UNAM, s/n, 56228602/1184, olveram11@gmail.com
CCADET/UNAM, circuito exterior UNAM, s/n, 56228602/1184, ricardo.damian@ccadet.unam.mx
CCADET/UNAM, circuito exterior UNAM, s/n, 56228651/1117, rufino.diaz.@ccadet.unam.mx

Resumen

Este artículo describe el proceso de diseño y calibración de un fotómetro para el equipamiento de un laboratorio de óptica, con interfaz virtual construido en el Laboratorio de Electrónica del CCADET-UNAM; así como su funcionamiento, tanto del fotómetro como de la interfaz virtual, y las pruebas realizadas al mismo.

Las mediciones realizadas por el fotómetro pueden ser guardadas en una PC para un análisis posterior. La calibración se realiza utilizando un medidor comercial Newport, con éste se toman lecturas que servirán de referencia para ajustar el fotómetro desarrollado.

Palabras Clave: calibración, fotómetro, interfaz virtual, laboratorio, láser.

I. Introducción

El fotómetro consiste de un sensor óptico, un amplificador con tres escalas de selección de ganancia; después de esto se tiene una etapa de procesamiento que realiza los cálculos necesarios de tal forma que calcula la intensidad de la luz incidente en el sensor para finalmente mostrar el resultado en un display LCD y en una PC gracias a la interfaz virtual creada en LabVIEW.

Es capaz de medir la intensidad luminosa en Watts de dos diferentes longitudes de onda: 670 nm y 780 nm. Cuenta con tres escalas que aumentan la precisión de las mediciones. Estas características son suficientes [6] para realizar lecturas que se requieren en las prácticas de laboratorio de Óptica de la carrera de Física en la Facultad de Ciencias de la UNAM.

Se pueden tener lecturas a otras longitudes de onda, pero en este caso, sólo serán lecturas relativas dado que el fotómetro no está calibrado para ellas.

Para la calibración se utiliza un fotómetro comercial de Newport de doble canal modelo 2832-C con un detector modelo 818-UV que cuenta con un atenuador el cual disminuye tres décadas la intensidad de la luz incidente. Se

utiliza un láser de 670 nm de Thorlabs modelo CPS198 y otro de 780 nm modelo CPS192. Para la calibración se tomaron lecturas con el equipo comercial, las cuales son tomadas como referencia para ajustar el fotómetro desarrollado.

Por otro lado, el instrumento virtual permite al usuario visualizar las mediciones en un display y en una gráfica, además de que estas mediciones se pueden guardar en un archivo con extensión .lvm que pueden ser vistos con procesadores de texto u hojas de cálculo para realizar un análisis posterior de los resultados obtenidos en el proceso de medición. Esto es muy útil dada la gran cantidad de datos que se toman al realizar las prácticas de óptica.

II. Desarrollo

DISEÑO

En el fotómetro desarrollado se utiliza un sensor de silicio de Thorlabs modelo SM1PD1A (Figura 1) cuya curva de responsividad (Figura 2) permite realizar el cálculo necesario para obtener una aproximación de la intensidad luminosa en Watts que incide en el sensor, y que se requiere medir. Para el cálculo de la corriente se emplea la

fórmula (1) la cual nos muestra su relación con la responsividad y la potencia P_o [3].

$$I = R_\lambda P_o \quad (1)$$

Donde:

I: Corriente producida en el sensor

R_λ : Responsividad

P_o : Intensidad de luz incidente en el sensor



Fig. 1. Sensor de silicio con conector BNC

Spectral Responsivity:

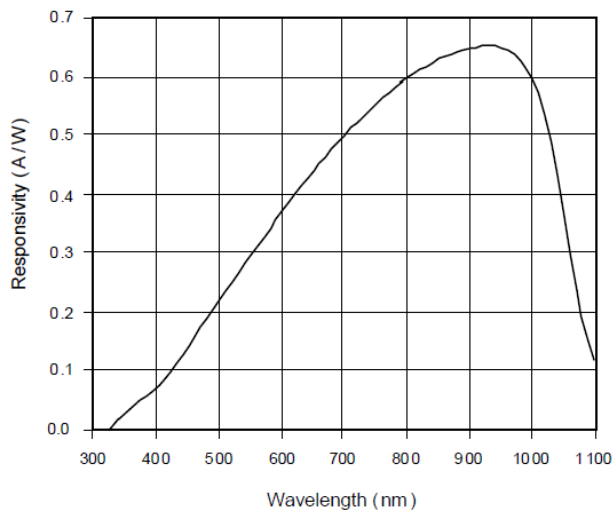


Fig. 2. Responsividad espectral del SM1PD1A

Se puede ver de la ecuación anterior que para obtener el valor de la intensidad luminosa (P_o) es necesario dividir la corriente generada entre la responsividad del sensor. Es importante mencionar que este valor sólo es aproximado ya que de la gráfica de responsividad no se puede obtener un valor de R_λ preciso para 670 nm. Para acondicionar la señal del fotodiodo se cuenta con un amplificador de transimpedancia. Éste convierte la corriente directamente en un voltaje.

Mediante tres resistencias de realimentación seleccionables (una a la vez por cada escala) conectadas a la salida del amplificador (Figura 3) se puede seleccionar la ganancia del amplificador, por lo que el cálculo de la intensidad luminosa ahora involucra a la resistencia R_f [1,2]. Esto se puede ver en la ecuación (2):

$$V_{OUT} = R_f R_\lambda P_o \quad (2)$$

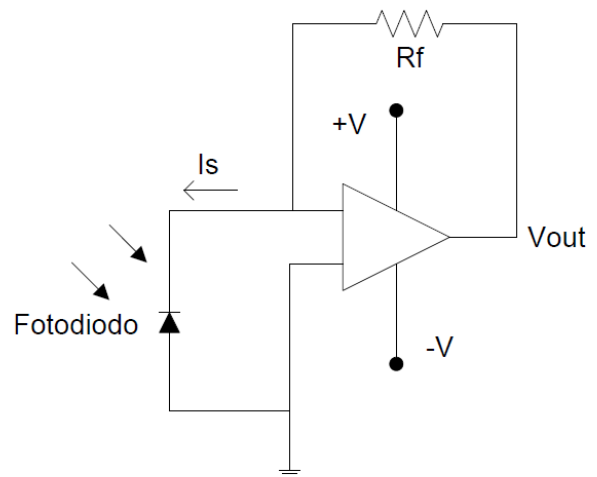


Fig. 3. Amplificador de transimpedancia

Estas resistencias introducen un error más en la lectura, producen una menor exactitud al hacer el cálculo de la intensidad. Por lo que hay que utilizar resistencias de precisión para reducir este problema.

El voltaje de salida del amplificador de transimpedancia es leído por un microcontrolador

pic16F688 que es el que realiza las operaciones para calcular P_o de la ecuación (2). Antes de realizar los cálculos, el microcontrolador realiza un promedio 16 muestras, lo que sirve como un filtro digital. Después de los cálculos, despliega

las mediciones en un display LCD y en una PC mediante el pin TX. El circuito completo se muestra en la (figura 4).

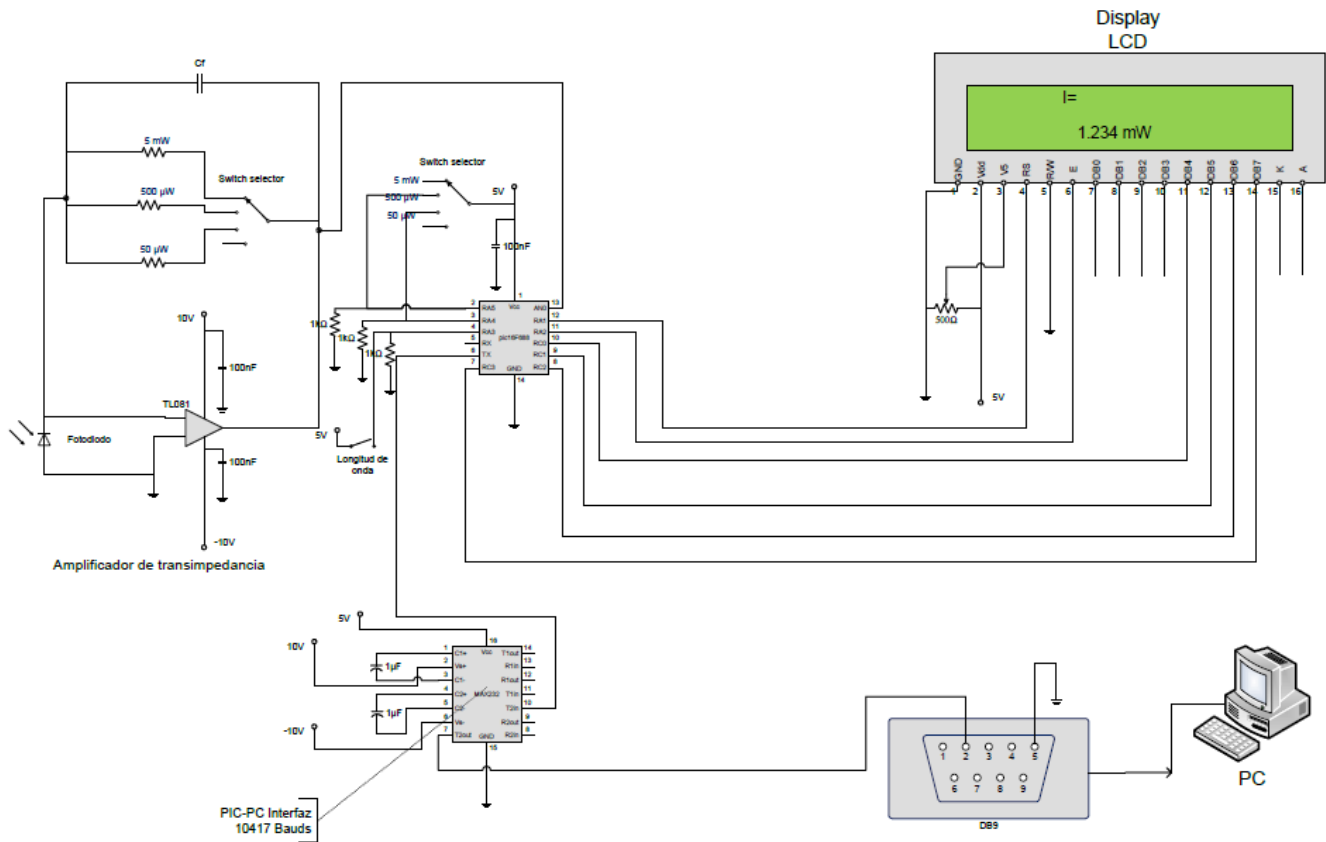


Fig. 4. Circuito completo del fotómetro desarrollado.

CALIBRACIÓN

El fotómetro que sirve como referencia para hacer la calibración es, como ya se dijo anteriormente, un fotómetro de Newport de doble canal modelo 2832-C (figura 5) en cual tiene un sensor del tipo 818-UV (figura 6).



Fig. 5. Fotómetro Newport



Fig. 6. Sensor de silicio 818-UV

Este fotómetro ya tiene grabado en un módulo de calibración la curva de responsividad del sensor a utilizar. Se utiliza en modo DC continuo, en rango automático, en el canal A y se le indica al medidor que el sensor tiene atenuador. Además se le aplican los filtros paso bajas que el fotómetro Newport tiene, esto es, un analógico más un digital para evitar ruidos de alta frecuencia.

Para realizar las lecturas que servirán como referencia, es necesario indicarle al instrumento que se tiene un offset. Para simular un laboratorio a oscuras, se obstruye la entrada de luz al sensor y se presiona la tecla ZERO dos veces para hacer que en esas condiciones (laboratorio a oscuras) el nivel de intensidad en el sensor sea 0 W. Una vez realizado este paso, se procede a realizar las mediciones de referencia.

Uno de los láseres que se utiliza es un CPS198 de 670 nm (figura 7). Por lo que el fotómetro Newport se ajusta a esta longitud de onda



Fig. 7. Láser CPS198, 670 nm.

Mediante un arreglo óptico en el cual se mantiene el láser apuntando a un mismo punto y el sensor en posición fija para obtener mediciones estables. Se toman diferentes mediciones para poder tener un promedio. Con los ajustes del fotómetro Newport arriba mencionados, se obtuvieron las mediciones de la tabla 1.

No.	I [mW]
1	2.31
2	2.29
3	2.31
4	2.28
5	2.27

Tabla 1. Mediciones de referencia.

Se puede ver una pequeña variación en las lecturas posiblemente debidas a vibraciones de la mesa de trabajo.

Para la etapa de calibración, se realizan mediciones con el fotómetro desarrollado reproduciendo las mismas condiciones (manteniendo láser y sensor fijos). Se procedió a ajustar el valor de I para que checara con la referencia, como ya se mencionó, se tomó el promedio de las lecturas de la tabla 1. La calibración se consigue multiplicando por un factor de ajuste en el cálculo de P_o , es decir, la relación entre el valor calculado (que es el mismo al que despliega el fotómetro desarrollado sin calibrar) y el valor de referencia medido. Y se consigue así un error de sólo 1 %.

$$e = \frac{2.292}{2.270} \quad (3)$$

Donde 2.292 es el promedio de la tabla anterior y 2.270 el valor leído en el fotómetro.

De esta manera se calibran las tres escalas del fotómetro, ya que cada una tiene asociada una resistencia diferente. Una vez calibrado el instrumento ya se pueden tener lecturas confiables.

El factor es introducido en el programa del microcontrolador que realiza el cálculo.

En la (figura 8) se muestra un segmento del programa en ensamblador grabado en el microcontrolador y muestra la rutina con la cual se ajusta el factor mencionado.

```

escala_5m1:
movlw   high .4192
movwf   multiplicador + 1
movlw   low .4192
movwf   multiplicador
clrf    multiplicador + 2      ; multiplicador <-- k1
goto    factor
escala_500u1:
movlw   high .4078
movwf   multiplicador + 1
movlw   low .4078
movwf   multiplicador
clrf    multiplicador + 2      ; multiplicador <-- k2
goto    factor
escala_50u1:
movlw   high .4076
movwf   multiplicador + 1
movlw   low .4076
movwf   multiplicador
clrf    multiplicador + 2      ; multiplicador <-- k3
goto    factor
escala_5m2:
movlw   high .3456
movwf   multiplicador + 1
movlw   low .3456
movwf   multiplicador
clrf    multiplicador + 2      ; multiplicador <-- k4
goto    factor
escala_500u2:
movlw   high .3361
movwf   multiplicador + 1
movlw   low .3361
movwf   multiplicador
clrf    multiplicador + 2      ; multiplicador <-- k5
goto    factor
escala_50u2:
movlw   high .3360
movwf   multiplicador + 1
movlw   low .3360
movwf   multiplicador
clrf    multiplicador + 2      ; multiplicador <-- k6
goto    factor

```

Fig. 8. Segmento del programa en el microcontrolador

INTERFAZ

El fotómetro cuenta además con un software que permite visualizar la intensidad medida en una PC. La conexión se realiza mediante un cable serial, o bien, un adaptador serial para computadoras que no cuentan con puerto serial. Este software se desarrolló en LabVIEW y puede además almacenar la información de las intensidades que se midan en un lapso de tiempo determinado. Ver (figura 9).

Para utilizar este software es necesario instalar un programa RunTime (LabVIEW Run Time Engine) y una librería para la comunicación serial (VISA) de National Instruments que funciona como *framework* para que el software pueda operar en la PC sin tener LabVIEW instalado.

Este software, que es un archivo ejecutable (.exe) se utiliza de la siguiente manera:

Una vez instalado el runtime y la librería, se ejecuta el archivo (.exe); se selecciona el puerto COM al cuál se comunica la PC con el fotómetro; se le da clic al botón RUN (o Ctrl+R); se elige nombre y ubicación del archivo .lvm que guarda las mediciones. Finalmente, el fotómetro virtual muestra el valor de la intensidad incidente en el sensor y lo despliega también en una gráfica como se puede ver en la (figura 9).

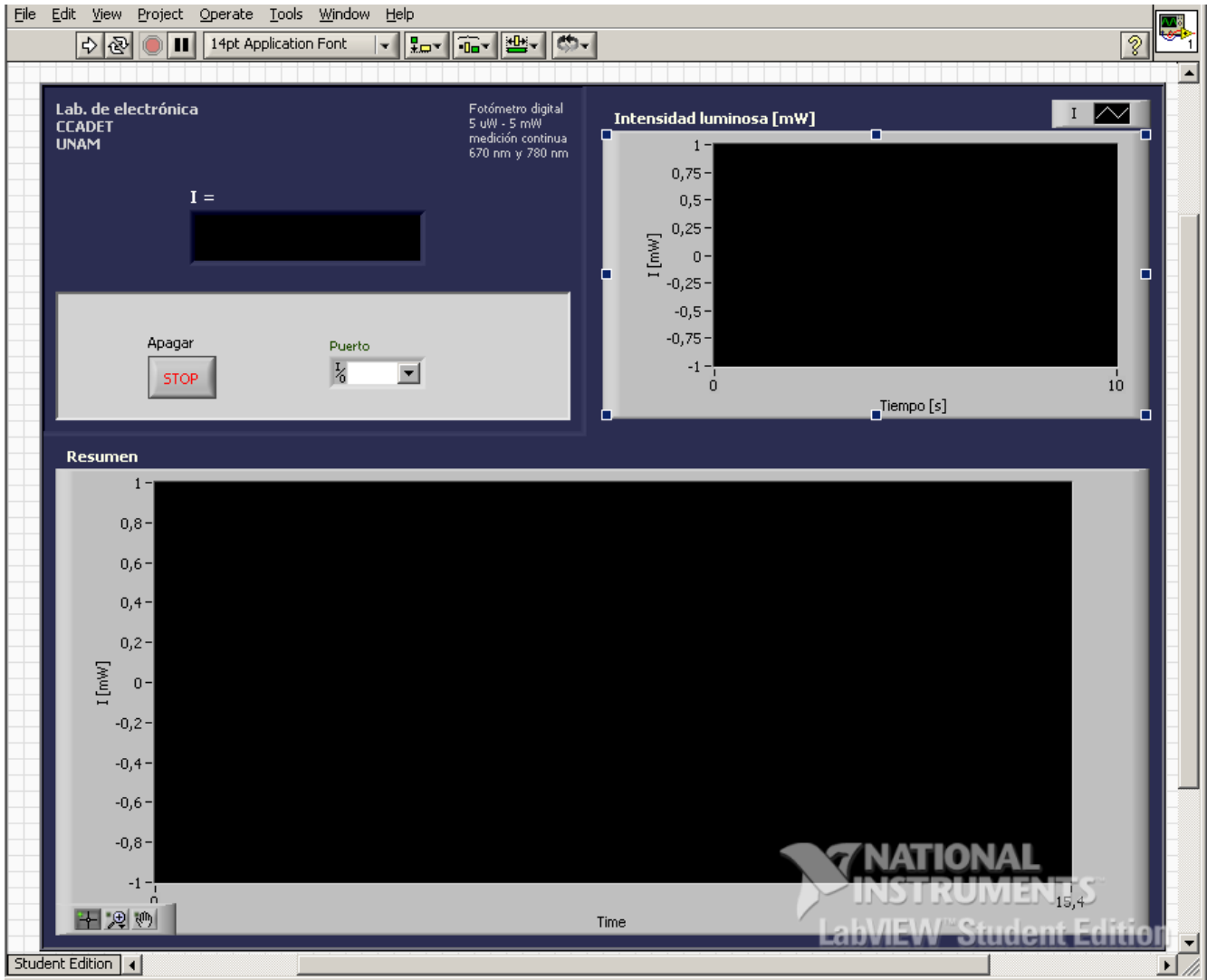


Fig. 9. Instrumento virtual

Al terminar de realizar las mediciones y dar clic en el botón Apagar, el fotómetro despliega una gráfica con todas las mediciones realizadas. Además se puede abrir el archivo .lvm con un procesador de textos u hojas de cálculo para un análisis posterior más profundo.

El diagrama de bloques de este instrumento virtual se muestra en la (figura 10).

Con el bloque “Visa serial” se configura el puerto serial con el que se establecerá la

comunicación, esto es, la velocidad de transmisión, número de bits, bits de paridad, etc.

El bloque “Visa Read” es el que lee lo que se transmite del pin TX del pic16F688 y lo despliega después en el indicador “I”.

El resto de los bloques se encargan de extraer los números para desplegarlos en una gráfica “Intensidad luminosa” mientras se guardan en la ubicación especificada con el bloque “Write to measurement file” [5].

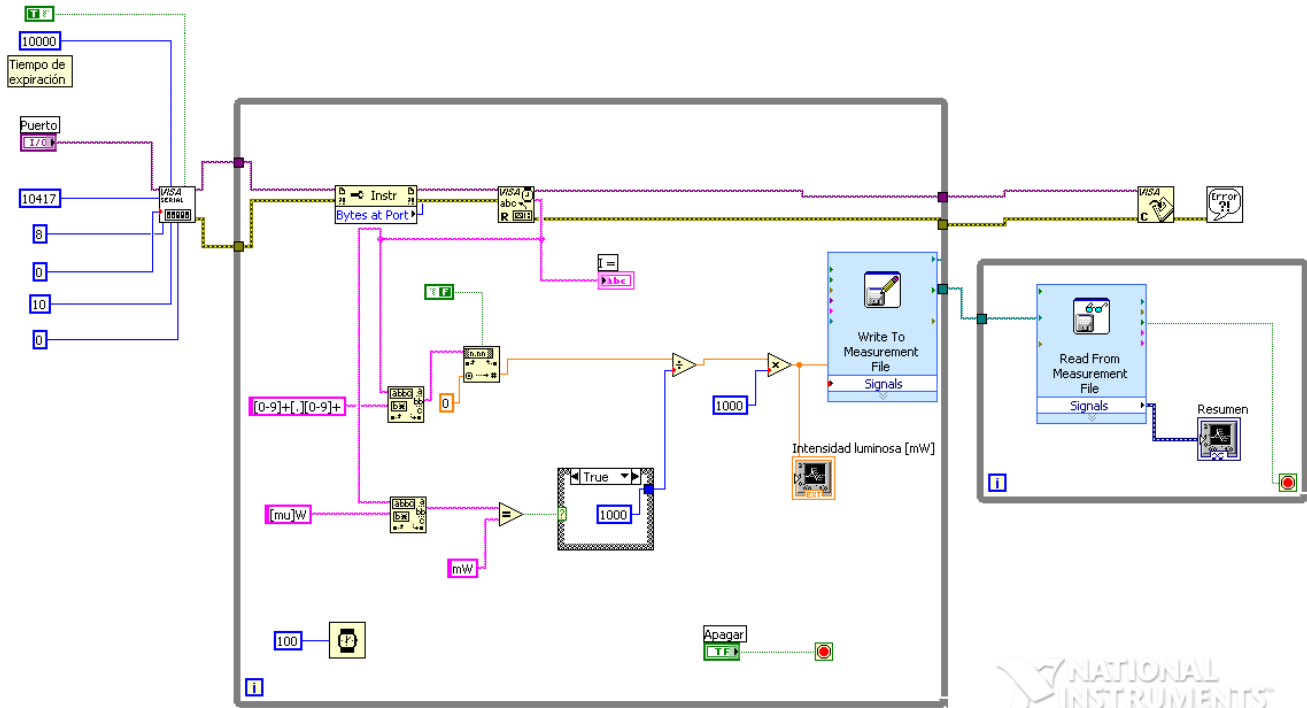


Fig. 10. Diagrama de bloques



Los últimos bloques se encargan de leer la información cuando ya se terminaron de hacer las mediciones “Read from measurement file” para desplegarlo en una gráfica “Resumen”.

La (figura 11) muestra un ejemplo de cómo se puede leer desde un procesador de textos el archivo generado por el programa.

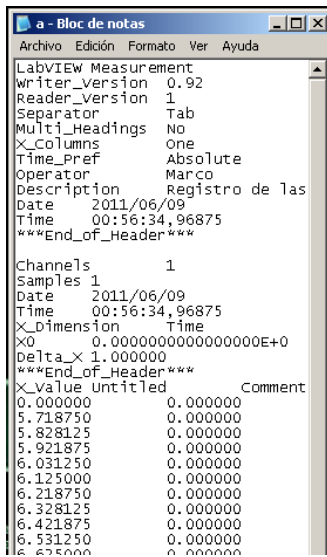


Fig. 11. Ejemplo lectura desde procesador de texto

La columna de la izquierda muestra el tiempo transcurrido en segundos y la columna de la derecha muestra los valores de intensidad leídos en mW.

III. Resultados

El fotómetro calibrado cumple con las características de precisión requeridas para cumplir con su objetivo principal de servir en las prácticas de laboratorio. Puede medir desde $1 \mu\text{W}$ hasta 4 mW . Las escalas son como se muestran en la tabla 2.

Escala	Rango
1	00,00 – 40,00 μW
2	000,0 – 400,0 μW
3	0,000 – 4,000 mW

Tabla 2. Rango de las escalas

Además de tener la opción de medir entre dos diferentes longitudes de onda (670 nm y 780 nm), se pueden realizar lecturas de cualquier otra longitud de onda dentro del rango $350\text{--}1100 \text{ nm}$ como ya se mencionó anteriormente, la intensidad mostrada en este caso por el fotómetro estará afectada por una constante que involucra la relación entre la responsividad del sensor a 670 o 780 nm y la responsividad a una longitud de onda diferente. De esta manera sólo se pueden realizar mediciones relativas.

Este fotómetro cuenta con un led de *backlight* para poder leer las intensidades sin problemas en un laboratorio a oscuras.

El fotómetro virtual es de gran utilidad ya que si se realizan mediciones durante un largo período de tiempo, se pueden guardar los datos de las mediciones y visualizarse en el mismo instrumento virtual al terminar el proceso de medición para un análisis posterior.

El fotómetro se colocó en un gabinete relativamente pequeño que le otorga al instrumento una buena maniobrabilidad dentro del laboratorio, además de ser ligero. Ver (Figura 12)



Fig. 12. Fotómetro

IV. Discusión

El instrumento está hecho para prácticas de laboratorio, su funcionamiento es adecuado para este propósito. Se tomaron varias precauciones para la disminución del ruido eléctrico que pudiera afectar las lecturas tales como: elegir un sensor de bajo ruido, un amplificador operacional igualmente de bajo ruido y con buen desempeño, resistencias de precisión, así como la implementación de un filtro analógico y uno digital [4].

El instrumento virtual muestra exactamente lo mismo que el fotómetro físico por lo que la intensidad de la luz se puede leer desde cualquiera de los dos. Por otro lado, el instrumento el virtual tiene además la opción de guardar lecturas de la intensidad medida y de visualizarlas al mismo tiempo en una gráfica.

Una de las ventajas de este instrumento, comparado con fotómetros comerciales, es su bajo costo, además todos los componentes, con

excepción del sensor, se pueden conseguir fácilmente y son de bajo costo.

Se podría mejorar al instrumento aumentando el número de longitudes de onda disponibles. También utilizando un microcontrolador con un convertidor analógico-digital con mayor número de bits que aumentaría la resolución en las lecturas.

V. Conclusiones

El fotómetro está pensado para mejorar la calidad de la enseñanza en las prácticas del laboratorio de Óptica, siendo un instrumento que se puede mover fácilmente de un lugar a otro del laboratorio. Como ya se mencionó, tiene la resolución y precisión necesarias, que permiten demostrar algunas de las características de la luz como son la reflectancia, transmitancia y la Ley de Malus.

Además de su interfaz con la PC hace más amigable el análisis de los resultados de las prácticas.

Se está desarrollando una nueva versión del fotómetro el cual además de incluir todas las características antes descritas, permite el control de la selección de la longitud de onda y de la escala desde la PC, esto es, se puede controlar completamente el comportamiento del fotómetro desde una PC con el instrumento virtual.

VI. Referencias

- [1] Jerald G. Graeme, *Photodiode amplifiers: op amp solutions*. 1995. MacGraw-Hill. 22-45.
- [2] *Design Considerations for a Transimpedance Amplifier*, Application Note AN-1803. National Semiconductor. EU, 2008
- [3] <http://www.edmundoptics.com/technical-support/electro-optics/basic-principles-of-silicon-detectors/>, 2010, Edmund Optics Inc., USA
- [4] *Low Noise Amplifiers for Small and Large Area Photodiodes*, Design Note 399. Linear Technology Corporation. 2006.
- [5] *LabVIEW Help*, Agosto 2009, National Instruments.
- [6] Wilmar Hernandez, *Robust photometer circuit*, IEEE, 2008

VII. Autores

M. en I. Juan Ricardo Damián Zamacona: Obtuvo su título de Maestría en Ingeniería eléctrica con especialidad en sistemas electrónicos en la UNAM. Su área de especialidad es la Instrumentación, actualmente labora en el Laboratorio de Electrónica del Departamento de Instrumentación y Medición del CCADET.

Dr. José Rufino Díaz Uribe: El Dr. Rufino Díaz Uribe estudió la Licenciatura, la Maestría y el Doctorado en Física en la Facultad de Ciencias de la UNAM. Su área de especialidad es la

Óptica. Actualmente es Investigador Titular C en el Departamento de Óptica y Microondas del Centro de Ciencias Aplicadas y Desarrollo Tecnológico de la UNAM.

Marco Antonio Olvera López: Estudiante, actualmente está desarrollando su tesis en Ingeniería Eléctrica-Electrónica en el Laboratorio de Electrónica del Departamento de Instrumentación y Medición del CCADET.

Un agradecimiento especial al programa PAPIME por su apoyo para la realización de este trabajo.

Este trabajo forma parte del proyecto PAPIME PE104410 “Manual de prácticas para el laboratorio de Óptica de la carrera de Física en la Facultad de Ciencias de la UNAM”.

Desarrollo de un Paquete Tecnológico para Mantenimiento Predictivo Industrial

***Julio Laria Menchaca, ** Juan Carlos Sepulveda, *José Antonio Castán Rocha, *Salvador Ibarra Martínez, *Emilio Castán Rocha**

*Facultad de Ingeniería, Universidad Autónoma de Tamaulipas, Centro Universitario Tampico-Madero, Tamaulipas, 52-833-1403455, jlaria.uat.edu.mx

**Facultad de Ingeniería Informática, Instituto Politécnico José Antonio Echeverría, calle 114 #11901, e/n 119 y 129, Marianao, La Habana, Cuba, 537-2600641, jcarlos@ceis.cujae.edu.cu

Resumen

En el presente trabajo presentamos el desarrollo de un paquete tecnológico para el monitoreo de la condición de la máquina en base a la técnica de medición y análisis de vibraciones mecánicas. El paquete tecnológico consta de un sistema electrónico y un software. El sistema electrónico que se presenta permite el trabajo hasta con 3 sensores pero puede ser escalado hasta un total de 16 sensores lo que le permite su uso para el monitoreo de la condición de la máquina en línea o fuera de línea. Para el desarrollo del software se optó por el lenguaje “object pascal” utilizando como ambiente de desarrollo al binomio Lazarus/Free pascal por ser multiplataforma, código abierto y libre. Esto hace que a diferencia del software similar existente, este pueda ser adaptado fácilmente a las condiciones concretas de cada empresa. Permite la medición de las diferentes unidades de vibración así como el análisis por frecuencias de las señales y por medio de un sistema de expertos brinda un diagnóstico sobre las posibles fallas que presenta la máquina.

Palabras Clave: acelerómetros, estado de la máquina, mantenimiento predictivo, vibraciones mecánicas

I. Introducción

Hoy en día, el mantenimiento predictivo basado en el monitoreo de la condición de la máquina se está convirtiendo en una poderosa herramienta de productividad, clave para enfrentar la creciente competitividad de mercados. Este cambio es parte de una transformación global que abarca toda la industria [1].

El mantenimiento predictivo basado en el monitoreo de la condición de la máquina se centra en la determinación del estado de la máquina en operación. La mayoría de las partes de la máquina darán un tipo de aviso antes de que fallen. Para percibir los síntomas con que las máquinas nos están advirtiendo se requieren varias pruebas no destructivas tales como el análisis de lubricantes, el análisis de corrientes de Eddy, el monitoreo y análisis de vibraciones mecánicas y la termografía infrarroja. De todas estas pruebas no destructivas, la más importante en el mantenimiento predictivo lo constituye el

monitoreo y análisis de vibraciones mecánicas ya que esta proporciona la cantidad de información más relevante acerca de su funcionamiento interno [2].

Los sistemas para la medición y análisis de las vibraciones mecánicas pueden ser manuales o automáticos [1]. Los manuales son aquellos donde la colección de datos se hace de forma manual. Estos han tenido un crecimiento rápido dentro de la industria, convirtiéndose en la instrumentación más común utilizada. Los automáticos envían los datos colectados de forma automática a una computadora donde se puede almacenar una gran cantidad de información, generando tendencias de series de mediciones y presentando diversos tipos de gráficos para su evaluación por analistas. Este tipo de sistema tiene la ventaja sobre el manual que debido a la gran cantidad de datos que son manejados, permiten observar cambios en el comportamiento de las máquinas, eventualmente automatizar la detección de estos cambios y en casos más

avanzados, son incorporados a sistemas de expertos.

En el campo del monitoreo de vibraciones mecánicas los sistemas expertos se utilizan fundamentalmente como herramientas de diagnóstico. Se trata de que el programa pueda determinar en cada momento el estado de funcionamiento de sistemas complejos, anticipándose a los posibles incidentes que pudieran acontecer. Así, usando un modelo computacional del razonamiento de un experto humano, proporciona los mismos resultados que alcanzaría dicho experto [3].

Los sistemas expertos no sólo deben poder indicarle probables fallas y facilitarle la labor de mantenimiento, deben también ser parte de todo su proceso productivo, adaptarse a su infraestructura de telecomunicaciones, ser cómodos para que sus ingenieros y técnicos los adopten y tener la capacidad de crecer junto con sus necesidades.

Existen varias tecnologías para el monitoreo de la condición de la máquina, entre las que se encuentran [4, 5, 6, 7,8] CSI Emerson Process Management RBMware 4.70, Machinery Condition Management (MCM2000), SKF Condition Monitoring, Microlog CMXA 50 Portable Data Collector/FFT Analyzer que son de amplio uso en la industria pero estas no incluyen sistemas de expertos. Sin embargo, existen algunas opciones de software comercial para el monitoreo de la condición de la máquina, como el de la marca Bently–Nevada y el Expert Alert de DLI Engineering Corporation que incluyen sistemas de expertos pero debido a sus costos puede dejar de ser convenientes.

Los requerimientos de la industria son muy variables y resulta a veces complicado encontrar un sistema experto que cuente con las características requeridas. En estos casos, una opción muy recomendable es utilizar un sistema con la posibilidad de que los mismos desarrolladores le hagan las adecuaciones necesarias al software según las necesidades de la industria. Es por ello que en el presente trabajo mostramos el desarrollo de un paquete tecnológico que incluye una electrónica de fácil configuración y escalable desde 1 hasta 16

sensores y un software que incluye un sistema de expertos y desarrollado en código abierto y libre para su libre adecuación a las necesidades de los clientes.

II. Desarrollo

Una vez que un problema específico de una máquina ha sido identificado por su firma de vibración, la pregunta siguiente consiste en como estimar la gravedad de esas vibraciones.

Una buena base de partida para estos niveles de alarma, es utilizar los niveles de severidad contenidas en publicaciones de organizaciones como la American National Standards Institute (ANSI), la Asociación Alemana de Ingenieros (VDI) y la International Standards Organization (ISO), en los estándares internacionales ISO 10816-1 1995 que reemplazó la norma ISO2372 e ISO 3945, VDI 2056 y la BS 4675 [2, 4, 5, 9, 10, 11]. En estas normas se estipulan las posiciones de medición así como los límites máximos permisibles por clases de equipos. Estos indicadores contemplan la medición del nivel total de velocidad rms dentro de un rango de frecuencias específico. A partir de estas normas y de otras tablas de severidades como las de Azima DLI y las de IRD Mechanalysis, en las tablas 1 y 2 se presenta una relación de los niveles de vibración en función de la frecuencia y que se puede aplicar para un gran número de máquinas rotativas con una confianza aceptable.

Por otra partes, la norma ISO 13374-1: 2003 [12], establece las guías generales para las especificaciones del software relacionado al procesamiento de los datos, comunicación y presentación de la información del monitoreo de la condición de la máquina y diagnóstico. Esta sirve de guía para el analista de mantenimiento predictivo donde el procesamiento de los datos relevantes y los procedimientos de análisis son requeridos para interpretar los datos recibidos de las actividades de monitoreo.

Para realizar el diagnóstico los sistemas de expertos se basan en el análisis de tendencias. El análisis de tendencias consiste en almacenar las firmas de las mediciones grabadas a tiempos específicos. La manera más sencilla de utilizar

las tendencias en las vibraciones es la de establecer un espectro de vibraciones representativo de la máquina operando normalmente como punto de referencia. Después, comparar esta referencia con espectros que se vayan grabando posteriormente [2]. Cuando se hace la comparación, es necesario tener en cuenta que las condiciones en que se operan las máquinas cuando se graba el espectro nuevo, deben ser lo más similares posibles a las que se operaban cuando se grabó el espectro de referencia. Para ello, los sensores deben colocarse en los mismos lugares de medición en ambos casos y se deberá utilizar en lo posible los mismos sensores.

Por lo general en los sistemas de expertos, en lugar de representar los espectros de las vibraciones en unidades de frecuencias, muchas veces es deseable utilizar órdenes o múltiplos de las rpm de la máquina. La primera orden se llama 1x, la segunda 2x, etc. Esto es especialmente valorable, si se quiere comparar varias mediciones de una misma máquina, tomados en momentos diferentes y que la velocidad de la vibración haya cambiado un poco entre las mediciones [13].

En el desarrollo del paquete tecnológico que se presenta nos basamos en los requerimientos de la norma ISO 13374-1: 2003 y para el análisis de las señales y diagnóstico de posibles fallas, utilizamos los indicadores que se definen en las publicaciones y normas expuestas arriba.

Para el desarrollo del software del paquete tecnológico, se utilizó la metodología ágil de análisis y diseño XP (eXtreme Programming) que se basa en la retroalimentación continua entre el cliente y el equipo de desarrollo. Se optó además, por el lenguaje “object pascal” utilizando como ambiente de desarrollo al binomio Lazarus/Free pascal por ser multiplataforma, código abierto y libre, y por las buenas prestaciones, velocidad y eficiencia del código generado.

III. Resultados

En la Fig. 1 se muestra el diagrama en bloques del paquete tecnológico desarrollado. El mismo consta de un sistema electrónico y un software.

El sistema electrónico está constituido por los sensores y el módulo de la electrónica encargado del acondicionamiento de las señales de los sensores y la digitalización de estas para su posterior envío a la computadora. Este módulo permite el trabajo entre 1 y 16 sensores. Cada sensor se conecta a un circuito de acondicionamiento formado por un amplificador y un filtro pasa bajo del tipo Bessel de segundo orden con una frecuencia de corte de 20KHz. La salida del circuito acondicionador de señal está conectada a la entrada del digitalizador. Como circuito de digitalización se utilizó una tarjeta de la National Instrument modelo NIUSB-6211 la cual permite digitalizar hasta 16 señales con una resolución de 16 bits, con una frecuencia de adquisición de hasta 250 KHz. Las señales digitalizadas se envían a la computadora a través de un puerto USB.

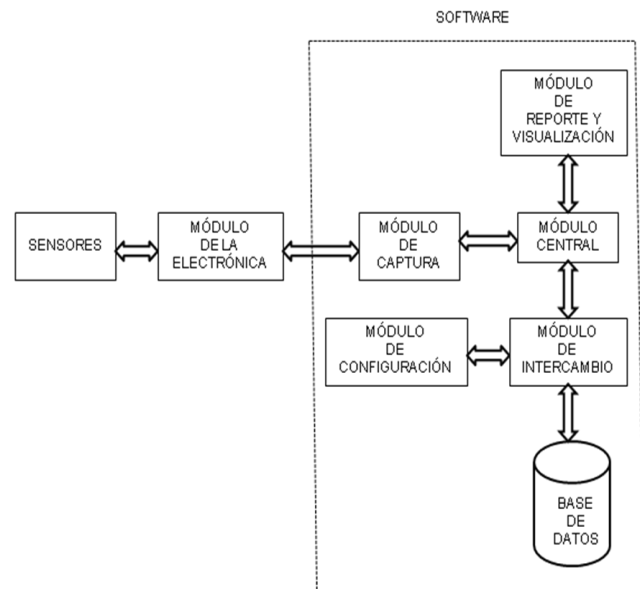


Fig. 1. Arquitectura del paquete tecnológico.

Tanto el circuito de acondicionamiento de las señales como el software del sistema se concibió de forma modular con el objetivo de hacerlo más flexible y adaptable. Por eso para el desarrollo del software del sistema se definieron las distintas interfaces a seguir en la comunicación entre cada uno de los módulos, permitiendo esto cambiar o modificar un módulo sin tener que alterar al resto de los módulos ni la arquitectura general del software del sistema. La estructura en

bloques del software del sistema se muestra en la Fig.1 en el cuadro con líneas discontinuas.

IV. Discusión

El software del sistema se concibió de la siguiente manera: el *módulo de captura* es el encargado de comunicarse con el circuito de digitalización del módulo de la electrónica y en la Fig. 2 se muestra la estructura interna de dicho módulo.



Fig. 2. Estructura interna del módulo de captura.

El mismo está compuesto por 3 bloques, el primero es el bloque de interface con el firmware de la electrónica de captura, en este caso con la tarjeta de digitalización NIUSB-6211 y es completamente dependiente de esta. El módulo de capturo está diseñado de tal manera que si se desea utilizar otro circuito de digitalización solo se deberá cambiar el software de este primer bloque. El segundo bloque es para realizar determinado ajuste o control de los datos de ser necesario. Y por último se tiene al bloque de interface con el módulo central el cual garantiza la independencia entre el módulo de la electrónica y el módulo central. Este módulo de captura también tiene la posibilidad de leer datos directamente desde archivos en disco. Esto permite modelar y simular señales que pueden ser analizadas por el sistema y lo que es más importante aún, las señales puedes ser capturada en una zona lejana y ser enviadas por la red para ser procesadas por el especialista sin necesidad de que el especialista se traslade físicamente al lugar donde se realiza la medición.

En el *módulo central* (ver Fig.1) se encuentra la “inteligencia” del sistema. Este está constituido (ver Fig.3) por los módulos del:

-Monitor en línea: que está diseñado para estar funcionando ininterrumpidamente, realizando mediciones constantemente y chequeando si las

mediciones se encuentran dentro de los parámetros adecuados. De existir alguna falla, se encarga de activar una señal eléctrica de alarma que puede ser utilizada para detener la maquinaria o realizar cualquier otra acción.

-Análisis de vibraciones y sistema de expertos: que se encarga de realizar un análisis más detallado del sistema, y en base a un sistema de expertos, brindar un diagnóstico de las posibles fallas que se presentan en la máquina.

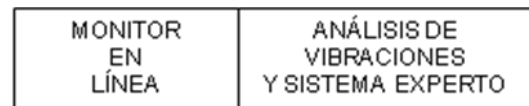


Fig. 3. Estructura interna del módulo central.

Para realizar el diagnóstico el sistema de expertos utiliza los análisis de tendencias entre los datos de referencias anteriormente almacenados y el comportamiento actual de las vibraciones de la máquina. Además, los requerimientos de la norma ISO 13374-1: 2003 y demás indicadores que se definen en las publicaciones y normas expuestas arriba, en el punto de desarrollo.

El *módulo de configuración* (ver Fig.1) se utiliza para la configuración del sistema y realizar la persistencia en la Base de Datos de dicha configuración. Para ello en la computadora se despliega una pantalla con pestañas que permite configurar desde las características del tipo de sensor que se usará en las mediciones hasta los derechos de acceso de los usuarios o la configuración para el acceso a la Base de Datos. Durante el proceso de configuración varias opciones del menú principal son ocultadas para evitar conflictos en el sistema. Este módulo permite la configuración para cualquier tipo de sensor ya sea acelerómetro o velocímetro.

El *módulo de intercambio* (ver Fig.1) es el gestor de base de datos y en su diseño se utilizaron los componentes de ZeosLib ya que estos son de código abierto y software libre, están disponible para una gran variedad de de compiladores y ambientes de desarrollos como Delphi, C++, Lazarus (FreePascal), Eide+MSEgui (FreePascal), y Kylix. Además, tienen conexión directa a sistemas de gestión de base de datos

como MySQL, PostgreSQL, Firebird, Interbase, Microsoft SQL Server, Sybase ASE, Oracle 9i, y SQLite disponibles en distintos sistemas operativos como Windows, Linux y Mac.

El *módulo de reporte y visualización* tiene la responsabilidad de facilitar la realización de los reportes así como de la visualización de los datos procesados respectivamente.

La *Base de Datos* del sistema está formada por 12 tablas y ha sido utilizado como SGBD a "Firebird 2-5". En la base de datos se almacena la configuración del sistema así como los datos capturados.

El software del paquete tecnológico se desarrollo de tal forma que intervienen dos actores diferentes, el analista de vibraciones y el operador.

El *analista de vibraciones* tiene como responsabilidad configurar el sistema, generar los reportes así como posicionar de forma correcta los sensores. Para esto tiene que interactuar con los módulos de configuración y de visualización y reportes.

El *operador* tiene como responsabilidad energizar e inicializar al módulo de captura cada vez que se vaya a realizar las mediciones, igualmente es el encargado de estar chequeando las mediciones para tomar las medidas pertinentes en caso de existir alguna anomalía.

V. Conclusiones

Como resultado del presente trabajo se desarrolló de un paquete tecnológico para el mantenimiento predictivo industrial. El mismo puede ser utilizado de forma manual o automática para el monitoreo de la condición de la maquina en base a la medición y análisis de las vibraciones mecánicas. Está compuesto por un sistema electrónico y un software. El sistema electrónico permite el trabajo desde 1 hasta 16 sensores ya sean acelerómetro o velocímetros. El software del paquete tecnológico incluye un sistema de expertos para el diagnóstico más aceptado de las posibles fallas que presente la máquina. Se desarrollo en lenguaje "object pascal" utilizando como ambiente de desarrollo

al binomio Lazarus/Free pascal por ser multiplataforma, código abierto y libre. Esto hace que a diferencia del software similar existente, este pueda ser adaptado fácilmente a las condiciones concretas de cada empresa.

VI. Referencias

- [1] R. K. Mobley, An Introduction to Predictive Maintenance, Keith Mobley; Hardcover, (2002).
- [2] G. White, Introducción al Análisis de Vibraciones, AZIMA DLI, www.azimadli.com, (1990).
- [3] R. S. Holek, Aspectos claves para un exitoso programa de monitoreo de vibraciones y la norma ISO 13374-1, Reliability world Latin America, www.noria.com/sp/rw2005.asp, (2005).
- [4] International Standards Organization, ISO Standard 2372: Mechanical Vibration of Machines with Operating Speeds from 10 to 200 rev/sec, (1974).
- [5] International Standards Organization, ISO Standard 10816-3: Mechanical Vibration-Evaluation of Machine Vibration by Measurements on Non-Rotating Parts, (1998).
- [6] APMA, Australian Pump Technical Handbook, (1987).
- [7] R. S. Beebe, Machine condition monitoring, MCM Consultants reprint (1995).
- [8] W. William, R. Cato, K. Mobley, Computer-Managed Maintenance Systems in Process Plants: A Step-By-Step Guide to Effective Management of Maintenance, Labor, and Inventory in Your operation, (1998).
- [9] M. P. Srivastava, IRD Mechanalysis, Vibration Monitoring for Predictive Maintenance, Purchase, (1993).
- [10] J. S. Rao, Vibratory Condition Monitoring of Machines, Narosa Publishing House, pp. 354-356, (2000).
- [11] B. R. Satyan and H. N. Nagarajan, Predictive Maintenance through Vibration Monitoring, Technical article, Noise and Vibration Laboratory, CMTI, (1988).

- [12] International Standards Organization, ISO Standard 13374-1: Condition monitoring and diagnostic of machines- data processing, communication and presentation, (2003).
- [13] B. C. Nakra, G. S. Yadava and L. Thuestad, Vibration Measurement and Analysis, National Productivity Council, New Delhi, pp. 54-60, (1989).

Agradecimientos

Los autores del presente trabajo le agradecen al Fondo Mixto CONACYT-Gobierno del Estado de Tamaulipas, México, su apoyo en recursos para el desarrollo de los trabajos así mismo le agradece a la Universidad Autónoma de Tamaulipas los apoyos recibidos.

VII. Autores

Dr. Julio Laria Menchaca obtuvo su título de Doctor en Tecnología de Avanzada con especialidad en física, en el Centro de Investigación en Ciencia Aplicada y Tecnología de Avanzada del IPN. Tiene más de 20 años de experiencia en el diseño y desarrollo de instrumentación electrónica.

Dr. Juan Carlos Sepúlveda obtuvo el título de Doctor en Ciencias con especialidad en informática en el Instituto Superior Politécnico José Antonio Echeverría en La Habana, Cuba. Tiene más de 15 años de experiencia en el desarrollo de software para sistemas y equipos electrónicos y robóticos.

M. en C. José Antonio Castán Rocha obtuvo el título de Maestría en Ciencias con especialidad en Tecnología Educativa en la Universidad Autónoma de Tamaulipas. Tiene más de 5 años de experiencia en el desarrollo de software.

Dr. Salvador Ibarra Martínez obtuvo el título de Doctor en Ciencias con especialidad en robótica en la Universidad de Girona, España. Tiene más de 5 años de experiencia en el desarrollo de equipos electrónicos y robóticos.

M. en C. Emilio Castán Rocha obtuvo el título de Maestría en Ciencias con especialidad en Electrónica en el Instituto Tecnológico de Madero. Tiene más de 5 años de experiencia en el desarrollo de equipos electrónicos.

Propuesta de diseño de un sensor para determinar la concentración de LiBr mediante las técnicas ópticas de Transmitancia (T) e Interferencia Multimodal (MMI)

*Edgar Eduardo Antúnez Cerón, **Miguel Ángel Basurto Pensado, ***José Javier Sánchez Mondragón

*** Centro de Investigaciones en Ingeniería y Ciencias Aplicadas (CIICAp)
Universidad Autónoma del Estado de Morelos,
Av. Universidad 1001, Col. Chamilpa, Cuernavaca, Mor., 62209, México
(777)3297084, eantunez@uaem.mx

*** Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE)
Apdo. Postal 51 y 216, Tonantzintla, Puebla 72000, México

Resumen

Las bombas de calor trabajan con una mezcla refrigerante de bromuro de litio-agua (LiBr-H₂O), y en función de la concentración de esta solución se puede ajustar la eficiencia de una bomba de calor. Razón por la que este trabajo presenta un análisis de la concentración de este par líquido mediante dos técnicas ópticas: Transmitancia (T) e Interferencia Multimodal (MMI) en una fibra óptica con estructura SMS (Singlemode-Multimode-Singlemode), con la finalidad de establecer el diseño para un sensor óptico colocado “*en línea*” con el sistema capaz de determinar la concentración de LiBr durante el proceso. De lo anterior, surge la necesidad de implementar un circuito electrónico de instrumentación para cada uno de los arreglos experimentales propuestos para ambas técnicas, y de este modo generar una respuesta en voltaje equivalente a la intensidad (potencia) transmitida por la solución (a una determinada concentración) para su posterior análisis y la formulación de una ecuación que estime la concentración de LiBr a partir de la respuesta en voltaje obtenida.

Palabras Clave: Circuito de Instrumentación, Interferencia Multimodal, Fibra Óptica, Fotodiodo, Transmitancia, Sensor.

I. Introducción

Las actuales tecnologías de fibra óptica y Optoelectrónica han alcanzado niveles de madurez técnica, calidad y rentabilidad que están más allá de los disponibles hace algunas décadas [1]. En la actualidad, los sensores de fibra óptica (SFO) se han establecido como una alternativa real, desplazando en ciertas aplicaciones a los tradicionales, además también han crecido en técnica y sofisticación y en el rango de posibles aplicaciones. Este tipo de sensores entre otros parámetros pueden medir: tensión, temperatura, humedad, posición lineal y rotacional, tensión, vibración, campos eléctricos y magnéticos; así como realizar mediciones en ambientes

biológicos y químicos [2-3]. Los sensores basados en fibra en la mayoría de los casos son empleados como instrumental de medición, caracterizadores y como elementos de control.

Por otro lado, las técnicas basadas en la interacción de la luz (radiación electromagnética) con la materia (espectroscópicas) revelan información tanto cualitativa como cuantitativa de un compuesto en una solución bajo estudio, esto representa una herramienta indispensable para la estimación de la concentración de dicho compuesto en la solución. El presente trabajo tiene como objetivo implementar y caracterizar un sistema óptico para determinar la concentración de LiBr en soluciones de LiBr-H₂O utilizando las técnicas ópticas de

Transmitancia (T) e Interferencia Multimodal (MMI) para el diseño de un sensor óptico”

El funcionamiento de las bombas depende en gran medida de la mezcla refrigerante o “*par de trabajo*” que se usa en el sistema de absorción. La mezcla refrigerante más conocida es la solución LiBr-H₂O, en la que el agua funciona como fluido de trabajo y el LiBr como absorbente. Actualmente, la concentración de LiBr no es medida en línea, para ello se utilizan dos métodos: 1) *Diagramas de equilibrio de Düring* que sólo aporta datos teóricos y alejados de los resultados reales, consiste básicamente en determinar la concentración de LiBr a partir de diagramas de estado cuando se conocen variables como la presión y temperatura de operación de la bomba y 2) *Refractometría* que si bien sus resultados son más confiables sacrifica coste en tiempo y la medición no se hace “*en línea*” con el sistema, hay que llevar la muestra de la solución para su análisis en un refractómetro (instrumento que determina el índice de refracción, IR, de una solución) para posteriormente mediante una ecuación establecida determinar la concentración de LiBr.

Hoy en día empresas extranjeras ofrecen a la venta sensores ópticos para estimar la concentración en determinadas soluciones (como por ejemplo: azúcar-agua, acetona-agua, etanol-agua, etc.), pero no para la solución refrigerante LiBr-H₂O. Estos sensores comerciales operan bajo el principio de absorción o transmitancia óptica o bien mediante un análisis ultrasónico de la solución; sin embargo, el adquirir uno de estos equipos representa una fuerte inversión económica (entre de \$12,000 y \$15,000 USA) sin contemplar la compra de la unidad de procesamiento y despliegue de la información y gastos de envío.

Un método alternativo para determinar la concentración de LiBr es utilizar sensores que funcionen en base a algunas de las propiedades ópticas de la solución, como por ejemplo: la absorción y transmisión de la radiación electromagnética (luz) y el índice de refracción

(IR). Haciendo uso de una metodología en función de las propiedades ópticas de la solución de trabajo, se proporciona información relacionada con la identidad de la especie atómica o molecular que está en la muestra (análisis cualitativo) o bien se proporciona información numérica como la cantidad de uno o más de los componentes (análisis cuantitativo) [4,5]. Tomando en cuenta el tipo de información que proporciona esta metodología el presente trabajo propone dos técnicas ópticas: **Transmitancia (T)** e **Interferencia Multimodal (Multi-Mode Interference, MMI)** para la determinación de la concentración de LiBr.

Transmitancia (T)

En la Fig. 1 se ilustra un haz de radiación paralela antes y después de que ha pasado a través de un medio con una solución absorbente. Como consecuencia de las interacciones entre los fotones y los átomos o las moléculas absorbentes de la solución, la potencia del rayo se atenúa de P_0 a P . La *transmitancia* del medio es entonces la fracción de la radiación incidente transmitida por el medio. A menudo, la transmitancia se expresa como porcentaje y se suele interpretar mediante perfiles en donde la respuesta de máxima transmitancia se aproxima al 100%:

$$\%T = \frac{P}{P_0} \times 100 \quad (1)$$

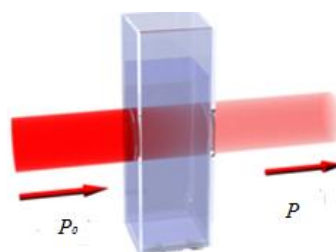


Fig. 1 Transmitancia óptica en una solución absorbente

Cuando hablamos de Transmitancia también nos referimos al comportamiento óptico de una solución, es decir, como consecuencia de las propiedades ópticas de la solución, ésta puede permitir o no el paso de una fracción de la radiación (luz) incidente y dicha fracción de radiación transmitida se utiliza como una señal de

respuesta. Para el caso particular de tener diferentes concentraciones de la misma solución se obtendrá distintas respuestas, respectivamente. Esta es una de las técnicas se utilizará posteriormente para caracterizar el diseño de un sensor.

Interferencia Multimodal (Multi-Modal Interference, MMI)

La técnica MMI es una herramienta base para la implementación de diferentes dispositivos de guía de onda ópticos. MMI fue inicialmente investigada y propuesta para guías de ondas planas [6]. Dispositivos basados en efectos MMI implementados en guías de ondas planas han sido desarrollados para procesamiento de señales ópticas [7, 8], y para aplicaciones de sensado [9].

Recientemente, los efectos de MMI en una fibra óptica con estructura SMS fueron investigados y utilizados para el desarrollo de aplicaciones de sensado y de procesamiento de señales [10, 11, 12]. Esos dispositivos ópticos ofrecen una solución basada totalmente en fibra con las ventajas de fácil fabricación, empaquetado e interconexión con otros tipos de fibra óptica mediante un sistema basado en mediciones de intensidad.

Una base útil para visualizar y obtener un mejor entendimiento de la técnica MMI en una guía de onda multimodal es el fenómeno de *autoimagen* (*self-imaging*) que es una propiedad de guías de onda multimodales por la cual un perfil de campo de entrada es reproducido en una única o múltiples imágenes en intervalos periódicos a lo largo del vector de propagación de la guía [9]. La Fig. 2 muestra un ejemplo de una guía de onda multimodal colocada entre dos secciones de guía de onda monomodales, lo que conformaría una estructura SMS. Así como el principio de autoimagen en el interior de la sección multimodal a los 2708, 5415 y 8122 μm ., distancias en las que se reproducen múltiples imágenes del perfil de entrada. . Múltiples imágenes (*multi-fold images*) de la distribución de campo de entrada también se pueden observar, por ejemplo, la generación

de dos imágenes se puede encontrar a los 1354, 4062, 6770 y 9478 μm ., éstas se pueden utilizar para desarrollar multiplexores ópticos. La generación de auto-imágenes ocurre a distancias específicas de la sección multimodal y solo para ciertas longitudes de onda [13].

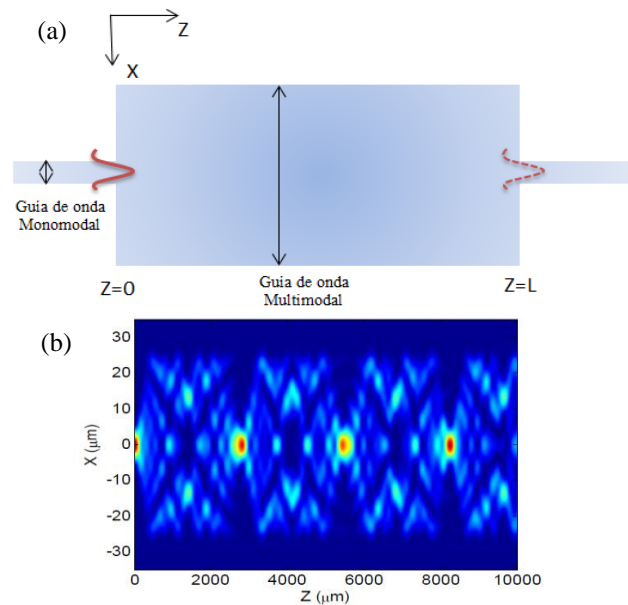


Fig. 2 (a) Esquema de una guía de onda con estructura SMS y (b) Principio de auto-imagen

II. Desarrollo

En este apartado se explican los arreglos experimentales diseñados para determinar la concentración de LiBr con ambas técnicas ópticas. Inicialmente se prepararon 13 soluciones de LiBr-H₂O con distintas concentraciones para su análisis de la concentración.

La Fig. 3 muestra el arreglo experimental utilizado con la Transmitancia para obtener una respuesta en voltaje de cada una de las soluciones de LiBr-H₂O. El arreglo experimental se constituye de: 1) Controlador de corriente del diodo láser (500mA) y diodo láser tipo mariposa (1555nm), 2) Medio de transmisión (fibra óptica monomodal SMF28), 3) Celda de cuarzo, lentes colimadores y sujetador de celda, 4) Fotodetector de InGaAs SM05PD4A (respuesta espectral de 800-1800nm), 5) Circuito de instrumentación convertidor de intensidad a voltaje del fotodetector y 6) Multímetro digital.

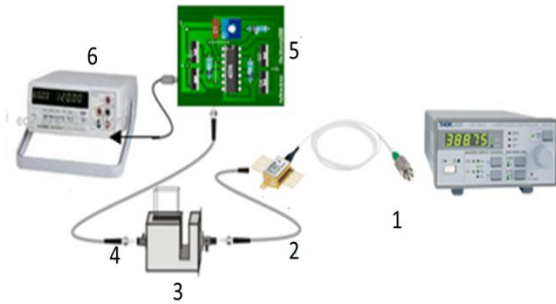


Fig. 3. Arreglo experimental para Transmitancia.

Con la ayuda de este arreglo se analiza la respuesta en intensidad transmitida (o absorbida) por cada una de las soluciones al hacer incidir sobre éstas un haz de luz a una intensidad determinada procedente del diodo láser. Cada una de las soluciones de LiBr-H₂O con distinta concentración (de 44.30 a 60.69%) genera una respuesta en intensidad transmitida diferente, la cual puede interpretarse como una respuesta en voltaje con la ayuda un circuito de instrumentación, y finalmente, mediante una caracterización con todas las soluciones, se establece una ecuación que relacione la respuesta en voltaje generada y la concentración de LiBr presente.

El circuito de instrumentación utilizado para este arreglo consiste básicamente de un fotodetector PIN configurado en modo fotovoltaico (sin polarización) conectado a un amplificador de instrumentación de alta precisión (INA101) con una ganancia ajustable por el usuario para obtener el equivalente en respuesta en voltaje.

Por otro lado, para la técnica MMI se utiliza una fibra óptica con estructura monomodo-multimodo-monomodo (Singlemode-Multimode-Singlemode, SMS). En este tipo de estructura la sección multimodal (Multi-Mode Fiber, MMF) es el elemento de sensado, ésta tiene un perfil de índice escalonado, un diámetro de revestimiento de 125µm, un índice de refracción de núcleo de 1.444 y una longitud de 14.55 mm y se empalma entre dos secciones de fibra monomodal (Single-Mode Fiber, SMF) de índice escalonado, 9µm y 125µm de diámetro del núcleo y revestimiento,

respectivamente. La fibra SMS utilizada se presenta en la Fig. 4.

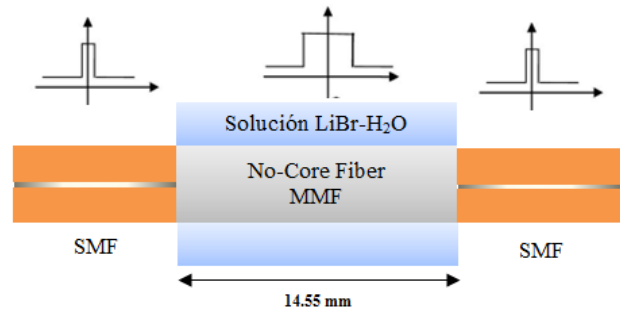


Fig. 4 Fibra SMS utilizada en MMI

La sección MMF no tiene revestimiento de manera que su núcleo expuesto interactúa con diferentes concentraciones de LiBr-H₂O con índices de refracción entre 1.423 y 1.472. Los diferentes índices de refracción de las soluciones que recubren al núcleo expuesto generan variaciones en el acoplamiento de los modos de la sección MMF de la fibra SMS, y por tanto, generan una respuesta diferente para cada caso.

La fibra SMS utilizada fue diseñada para generar una primera auto-imagen del perfil de entrada justo en el extremo de salida de la sección MMF (Fig. 5) a una longitud de onda de 1555nm.



Fig. 5 Reproducción de la primera auto-imagen en la sección MMF de la fibra SMS a 1555nm

El perfil de entrada reproducido sufrirá un *desfasamiento* al momento de que la fibra SMS entre en contacto con cada una de las soluciones (moviéndose de la distancia de reproducción original). Esto se debe al hecho de que cada solución presenta un índice de refracción, IR, diferente, y por lo tanto, generará una respuesta en intensidad transmitida distinta para cada solución.

La Fig. 6 presenta el arreglo experimental para la respuesta en voltaje con MMI y se constituye principalmente de: 1) Diodo láser tipo mariposa (1555nm) y un controlador de corriente para el

diodo láser (500mA), 2) Fibra SMS (elemento de sensado) y contenedor, 3) Fotodetector de InGaAs FGA04 (con respuesta espectral de 800-1800nm) y circuito de instrumentación para convertir la intensidad a una respuesta en voltaje equivalente y 4).Multímetro digital.

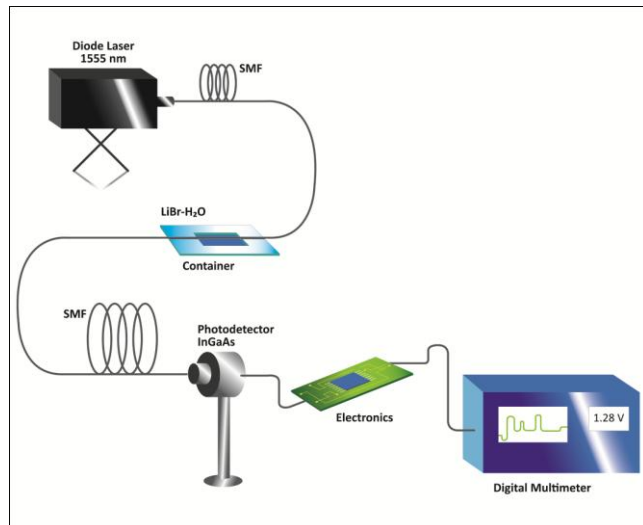


Fig. 6. Arreglo experimental para MMI.

El circuito de instrumentación de la Fig. 6 consiste básicamente del fotodetector con su circuito de respuesta típico (modo de operación con polarización) conectado a un amplificador de instrumentación de precisión (INA122P) con ganancia unitaria (para evitar amplificar la señal de ruido) y a la salida de éste se implementó un filtro RLC para dar una mejor estabilidad a la respuesta.

III. Resultados

Los resultados aquí reportados se realizaron con 13 soluciones de LiBr-H₂O con concentraciones e índices de refracción como los que se muestran en la Tabla 1.

Es importante mencionar que para concentraciones mayores del 60% aproximadamente la solución de LiBr-H₂O presenta cristalización y resulta desfavorable llegar a esta condición en el equipo.

Índice de refracción (IR)	Concentración (%) de LiBr
1.476	60.69
1.47	59.43
1.466	58.53
1.465	58.29
1.46	57.05
1.457	56.27
1.453	55.17
1.451	54.83
1.447	53.41
1.444	52.49
1.439	50.87
1.426	46.24
1.425	45.86
1.421	44.30

Tabla 1 Índices de refracción y concentraciones de las soluciones de LiBr-H₂O

Los resultados se clasifican para cada una de las técnicas ópticas, respectivamente. Para los resultados en T se sigue la siguiente metodología: 1) Se obtiene la respuesta óptica de las diferentes concentraciones de la solución para observar su comportamiento a la longitud de onda de 1555nm (diodo láser) con una fuente de luz blanca y un espectrómetro de infrarrojo cercano, NIR (ya que en el UV y VIS no generan información de interés), 2) Se obtienen los perfiles de transmitancia de cada solución en los que se toma como referencia para su normalización la mayor concentración, recordemos que la T se presenta de forma porcentual con respecto a alguna referencia. En este caso la mayor concentración presenta la mayor transmisión de intensidad y 3) Se obtiene la respuesta de intensidad transmitida en su equivalente en voltaje con el circuito de instrumentación para transmitancia.

Las respuestas ópticas para las concentraciones de LiBr-H₂O en el NIR se presentan en la Fig. 7. Las repuestas generadas muestran que a la longitud de onda de 1555nm los perfiles de las concentraciones se encuentran muy cercanos, lo que nos dice que la solución presenta una fuerte absorción a esta longitud de onda.

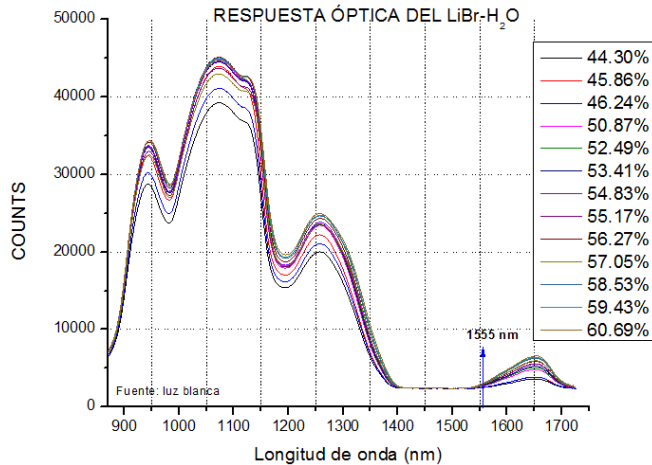


Fig. 7 Respuesta óptica en el NIR para las diferentes concentraciones de LiBr-H₂O.

La Fig. 8 presenta los perfiles de transmitancia con respecto a la mayor concentración (60.69%). Se observa que los perfiles están muy traslapados, razón por la cual se espera una respuesta en intensidad transmitida muy baja.

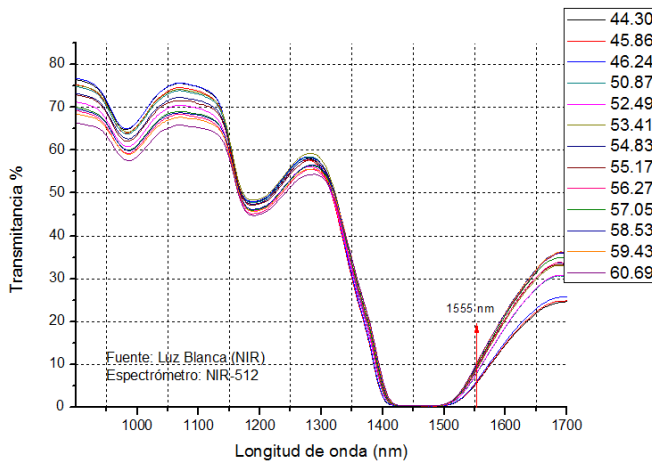


Fig. 8 Perfiles de transmitancia normalizados respecto a la concentración mayor.

Con el antecedente de los resultados obtenidos en T, la señal de intensidad transmitida se alimenta al circuito de instrumentación diseñado particularmente para ampliar el rango de respuesta en voltaje de cada una de las concentraciones. Pruebas preliminares generaban respuestas en voltajes en un intervalo de 3.283 y 3.330 V para la menor y mayor concentración de LiBr, respectivamente. Con el circuito de

instrumentación diseñado se generan las respuestas en voltaje de la Fig. 9.

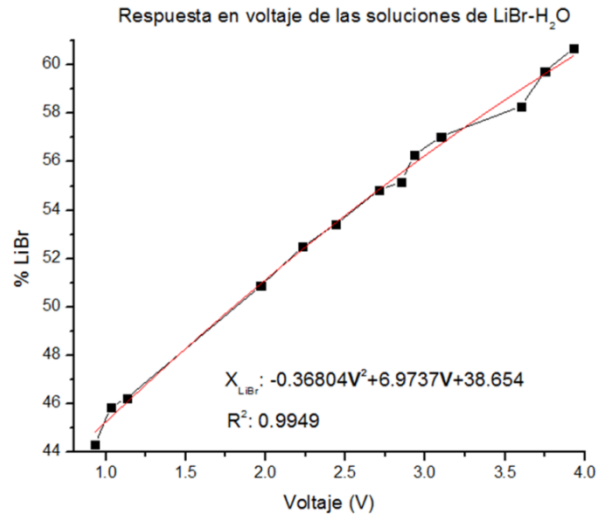


Fig. 9 Respuesta en voltaje de las concentraciones de LiBr-H₂O para transmitancia

De esta forma el nuevo rango para las respuestas en voltaje es de 0.93 a 3.75 V para la menor y mayor concentración, respectivamente. El comportamiento de los datos permite correlacionarlos mediante una ecuación de segundo orden:

$$X_{LiBr} = -0.36804 (V^2) + 6.9737 (V) + 38.654 \quad (2)$$

La Ecuación 2 tiene un factor de ajuste de 0.994. Lo que se traduce en una respuesta precisa y confiable. Una comparativa de los resultados reales y los generados con la ecuación propuesta se presentan en la Tabla 2.

Los resultados de la Tabla 2 que están sombreados corresponden a pruebas realizadas con soluciones de LiBr-H₂O distintas a las utilizadas en la caracterización de la respuesta y que se tenían reservadas con el objetivo de comprobar tanto el arreglo experimental y la ecuación establecida.

Voltaje (V)	%LiBr Experimental	%LiBr Ecuación	Error
0.93	44.3	44.82	-0.52
1.03	45.86	45.44	0.41
1.13	46.24	46.06	0.17
1.97	50.87	50.96	-0.09
2.23	52.49	52.37	0.11
2.44	53.41	53.47	-0.06
2.71	54.83	54.85	-0.02
2.85	55.17	55.53	-0.36
2.93	56.27	55.92	0.34
3.1	57.05	56.73	0.31
3.6	58.29	58.99	-0.70
3.75	59.73	59.63	0.099
3.93	60.69	60.37	0.31
2.04	51.85	51.34	0.50
1.42	47.73	47.81	-0.08
1.87	50.2	50.40	-0.20
2.46	53.11	53.58	-0.47
3.59	58.53	58.94	-0.41

Tabla 2 Resultados obtenidos para transmitancia.

Por otra parte, para los resultados con MMI se utiliza el arreglo experimental mostrado en la Fig. 6. La respuesta en voltaje para las distintas concentraciones en MMI se presentan en la Fig. 10.

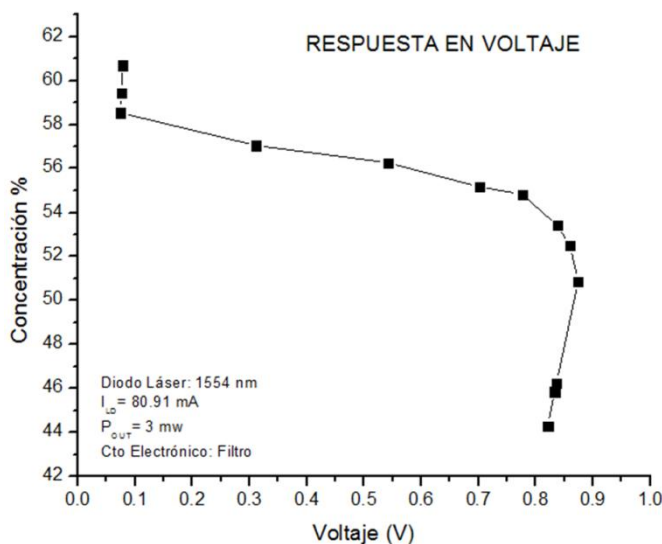


Fig. 10 Respuesta en voltaje para concentraciones de LiBr-H₂O con MMI.

Los resultados de la Fig. 10 presentan tres comportamientos distintos. El primero viene

dato por las concentraciones de LiBr-H₂O que se encuentran entre el 44.3 y 50.87 % como se puede apreciar también en la Tabla 3 (sombreado azul). Para estas concentraciones la respuesta en voltaje presenta un comportamiento creciente, aunque en pequeños incrementos de 0.821 a 0.874 V. Para estas concentraciones el índice de refracción oscila entre 1.421 y 1.439.

Índice de refracción, IR	Voltaje (V)	Concentración % LiBr
1.421	0.821	44.3
1.425	0.833	45.86
1.426	0.836	46.24
1.439	0.874	50.87
1.444	0.86	52.49
1.447	0.838	53.41
1.4518	0.776	54.83
1.453	0.701	55.17
1.457	0.542	56.27
1.46	0.312	57.05
1.466	0.0746	58.53
1.47	0.0769	59.43
1.476	0.0781	60.69

Tabla 3 Índice de refracción, respuesta en voltaje y concentración de las soluciones de LiBr-H₂O para MMI.

Un segundo comportamiento de la caracterización de la respuesta en voltaje viene dado por las concentraciones entre 52.49 y 57.05%. En este intervalo la respuesta empieza a decrecer considerablemente desde 0.86 a 0.312 V como se aprecia en los datos de la Fig. 10 como en la Tabla 3. En teoría se supondría un comportamiento creciente para la respuesta en voltaje de las concentraciones, pero evidentemente un factor muy importante nos genera este comportamiento: el índice de refracción del núcleo de la sección de MMF. Este parámetro intrínseco de la fibra óptica reporta un índice de refracción de aproximadamente 1.444, y como se puede ver en la Tabla 3 este valor es superado por el índice de refracción de las concentraciones de LiBr mayores de 50.87%.

Lo expuesto anteriormente explica la razón por la cual se genera una respuesta con tendencia decreciente en intensidad transmitida. Además, es

importante mencionar que el principio básico para confinar y transmitir radiación por el interior del núcleo de la fibra no se está cumpliendo: *la reflexión total interna*, ya que para que ésta suceda es necesario que índice de refracción del núcleo de la fibra sea ligeramente mayor al índice de refracción del revestimiento, situación que para concentraciones superiores al 50.87% no sucede. Esto resulta en una pérdida de la luz (dispersión) y por consecuencia en una atenuación de la intensidad transmitida por la fibra SMS.

No obstante lo anterior, podemos observar un tercer comportamiento en las tres últimas concentraciones analizadas (58.53 a 60.69%) es que aunque ciertamente se encuentran dentro una respuesta en voltaje muy baja, lo interesante es que conforme la concentración de la solución aumenta también lo hace la respuesta en voltaje, aunque una vez más en muy pequeños incrementos, pero que son detectables por los instrumentos utilizados en el arreglo experimental. La respuesta en voltaje experimenta un pequeño crecimiento de 0.0746 a 0.0781 V. Esto se puede explicar desde un punto de vista químico. Se sabe por las propiedades de la solución de LiBr-H₂O que a altas concentraciones la solución experimenta cristalización. Suponemos que estas concentraciones son lo suficientemente altas para indicar que estamos en el umbral de un nuevo estado físico de la solución: el estado sólido (cristales).

IV. Discusión

De la sección III podemos considerar dos vertientes claramente definidas: los resultados generados por transmitancia (T) y los generados por la interferencia multimodal (MMI) en una fibra óptica con estructura SMS. De lo anterior se establecen dos criterios para llevar a cabo la determinación de la concentración de LiBr en soluciones de LiBr-H₂O:

1) En transmitancia el parámetro que determina la estimación de la concentración de LiBr viene representado la intensidad transmitida

en cada una de las soluciones con diferente concentración. A partir de ésta se generan sus respuestas equivalentes en voltaje, y mediante la ecuación de segundo orden establecida mediante la caracterización de los datos, el usuario puede calcular el valor de dicha concentración. El paso determinante para esta técnica es colocar la muestra bajo las mismas condiciones de medición, ya que el resultado dependerá de que la fuente de luz (diodo láser) a una potencia (intensidad) establecida interactúe con la muestra de la solución bajo estudio para posteriormente recolectar la información mediante un fotodiodo para su posterior interpretación en voltaje con el circuito de instrumentación. De este modo, se tiene un arreglo experimental que puede representar un prototipo de diseño de un sensor en línea, en el que el usuario con la respuesta en voltaje generada puede determinar la concentración de la solución. Los resultados obtenidos muestran precisión y confiabilidad para generar datos aproximados a los reales.

2) En la técnica MMI, la respuesta viene al igual que en la técnica de transmitancia, definida por la cantidad de información (intensidad) que el arreglo transmita a su salida. No obstante, el óptimo funcionamiento con MMI dependerá fundamentalmente de un parámetro intrínseco a la naturaleza de la fibra óptica y del principio bajo el cual se rige: el índice de refracción (IR). El arreglo propuesto para esta técnica depende en gran medida de que la reflexión total interna se genere para obtener una transmisión de intensidad en el interior de la fibra, situación que se ve amenazada cuando la sección de sensado (MMF) de la fibra SMS interactúa con concentraciones que tienen un IR superior a dicha sección con el núcleo expuesto (de aprox. 1.444). Cuando esta condición se presenta el resultado es una dispersión de la intensidad del interior de la fibra hacia el medio que juega el papel del revestimiento (solución) y como consecuencia se obtienen resultados muy

pequeños en intensidad y por lo mismo en su equivalente a señal de voltaje. Hay que recordar que para que una fibra óptica transmita información el IR del núcleo debe ser ligeramente mayor que el IR del revestimiento. Situación que no sucede para concentraciones de LiBr mayores a 50.87%. No obstante, los resultados obtenidos presentan información para una posible caracterización y estudio de respuesta para concentraciones mayores. Inclusive, para concentraciones altas en donde la solución tiende a cristalizarse.

V. Conclusiones

Las propiedades ópticas de la solución de LiBr-H₂O como por ejemplo: transmisión y/o absorción de la radiación electromagnética (luz infrarroja) y el índice de refracción (IR) pueden utilizarse para obtener resultados cuantitativos de la concentración de LiBr presente en la solución.

La técnica de transmitancia genera resultados confiables y muy cercanos a los datos reales, de este hecho se considera que los sensores ópticos pueden tener mayor sensibilidad y precisión en sus respuestas. Con esta técnica se es capaz de estimar el porcentaje de LiBr presente en la solución con solo incluir el valor de la respuesta en voltaje generado por el arreglo en una ecuación establecida y de esta forma se obtiene el dato de interés. No obstante, el arreglo experimental propuesto para esta técnica representa un prototipo de diseño para un sensor en línea con el equipo. Asegurando la obtención de resultados en tiempo real.

Sensores ópticos de esta naturaleza no existen actualmente para soluciones de trabajo refrigerantes como lo es la solución de LiBr-H₂O, por lo que esta propuesta representa un prototipo de sensor óptico con un amplio campo de aplicación, si consideramos que los sistemas de refrigeración (sistemas de aire acondicionado) utilizan este par de trabajo en muchas industrias.

La técnica MMI permite generar una estimación de la concentración de LiBr en solo un intervalo de concentraciones específico, pero si esta

técnica se aplicara para determinar concentraciones en soluciones que no superen el índice de refracción del núcleo de la fibra SMS utilizada en el arreglo experimental, generaría resultados altamente confiables. Cabe mencionar que el implementar un sensor con esta técnica y una fibra SMS es ideal para lugares de trabajo con condiciones extremas de temperatura, presión, interferencias electromagnéticas y sobre todo en ambientes altamente corrosivos como lo es la solución de LiBr-H₂O.

Como trabajo a futuro de estos diseños se propone la implementación física del sensor para poder ser utilizado “en línea” con la bomba de calor y una interfaz mediante un sistema adquirente de datos (en LabView, ya está implementada) para el procesamiento de los datos, esto para la técnica de Transmitancia. Para el caso de MMI se está trabajando en el diseño y construcción (un proceso laborioso) de una fibra SMS que tenga un IR en la sección de sensado (MMF) ligeramente mayor al que presenta la máxima concentración de la solución de LiBr-H₂O, con el propósito de poder generar una respuesta en voltaje para cada una de las concentraciones de trabajo.

VI. Referencias

- [1] S. Sudho and K. Okamoto, *New Photonics Technologies for the Information Age: The Dream of Ubiquitous Services*, Artech House Optoelectronics Library, Londres, 2006, pp. 1-5.
- [2] F. T. S. Yu .S. Yin. *Fiber Optic Sensors*. Marcel Dekker, 2002, USA. Chapter 1.
- [3] E. Udd. *Fiber Optic Sensors-An introduction for Engineers and Scientists*, Jonh Wiley & Sons, 1991, USA. Chapter 1.
- [4] A. Skoog, F. Holler y S. Crouch. *Principios de Análisis Instrumental*, 6ta Edición, CENGAGE Learning, México, 2008, pp. 147-159.
- [5] A. Skoog, D. West, J. Holler y S. Crouch. *Fundamentos de Química Analítica*, 8va

Edición, Editorial Thompson, México, 2005, pp. 723-729.

- [6] L. Soldano y E. Pennings, "Optical multimode interference devices based on self-imaging: principles and applications", *Journal of Lightwave Technol.*, vol.13, 1995, pp. 615-627.
- [7] S. Nagai, G. Morishima, H. Inayoshi, y K. Utaka, "Multimode Interference Photonic Switches (MIPS)," *Journal of Lightwave Technol.*, vol. 20, 2002, pp. 675-681.
- [8] M. R. Paiam y R. I. MacDonald, "A 12-channel phased-array wavelength multiplexer with multimode interference couplers," *Photon. Technol. Lett.*, vol. 10, 1998, pp. 241-243.
- [9] K. R. Kribich, R. Copperwhite, H. Barry, B. Kolodziejczyk, J.-M. Sabatié, K. O'Dwyer, y B.D. MacCraith, "Novel chemical sensor/biosensor platform based on optical multimode interference (MMI) couplers," *Sensors and Actuators B: Chemical*, vol. 107, 2005, pp. 188-192.
- [10] Q. Wang, G. Farrell y W. Yan, "Investigation on Single-Mode-Multimode-Single-Mode Fiber Structure," *Journal of Lightwave Tech.*, vol. 26(5), 2008, pp. 512-519.
- [11] W. S. Mohammed, P. W. E. Smith, y X. Gu, "All-fibre multimode interference bandpass filter," *Opt. Lett.*, vol. 31, 2006, pp. 2547-2549.
- [12] S. M. Tripathi, A. Kumar, R. K. Varshney, Y. B. P. Kumar, E. Marin, y J. P. Meunier, "Strain and temperature sensing characteristics of single-mode-multimode-single-mode structures," *J. Lightwave Technol.*, vol. 27, 2009, pp. 2348-2356.
- [13] Muhamad A., Tesis: "Singlemode-Multimode-Singlemode Optical Fiber Structure for Optical Sensing". Dublin Institute of Technology, 2009, pp. 1-10.

VII. Autores

Ing. Edgar Eduardo Antúnez Cerón obtuvo su título de Ingeniería Tecnológica en Electrónica en el Instituto Nacional de Astrofísica, Óptica y Electrónica. Actualmente es estudiante de la Maestría en Ciencias Aplicadas con opción terminal en Eléctrica con Especialidad en Óptica en el Centro de Investigaciones en Ingeniería y Ciencias Aplicadas (CIICAp) de la Universidad Autónoma del Estado de Morelos.

Aplicación de la Domótica en el monitoreo de sensores electrónicos a través de una interfaz de comunicación con Dispositivos móviles vía Web Services (SOAP-XML)

***Christian Mauricio Castillo Estrada, **Leonardo Yamasaki Maza, *** Jorge Elpidio Cruz Gálvez**

Facultad de Contaduría Pública del Campus IV de la Universidad Autónoma de Chiapas. Carretera a Puerto Madero Km 1.5, Tapachula, Chiapas, México
(962) 6251723

*cmce@unach.mx, **lyama@unach.mx, ***jorge.cruz@unach.mx

Resumen

En este trabajo se presentan los resultados derivados de la implementación de la tecnología móvil y servicios web aplicados, para el monitoreo de sensores electrónicos y la automatización de lámparas en una casa habitación aplicando la terminología de Domótica. Para ello, se requiere de un concentrado de información de control realizado a través de una computadora central y la utilización de líneas de alimentación de baja tensión, propias de la casa como medio de comunicación. Así también, es necesario el desarrollo e implementación de aplicaciones de bajo nivel que permiten el monitoreo de sensores a través de un microcontrolador utilizando el puerto Serial y el uso de servicios web con la tecnología JAX-WS de la plataforma JavaEE para la comunicación con dispositivos móviles. Las pruebas experimentales se realizaron con una maqueta a escala de una casa habitación que posee componentes electrónicos conectados a un equipo de cómputo con acceso a Internet. Por lo anterior, resulta interesante estudiar la aplicación de la Domótica para proporcionar automatización dentro de una vivienda a través de la integración de las áreas electricidad, electrónica, robótica, informática y telecomunicaciones ofreciendo a las personas un aumento de confort, seguridad y ahorro energético en sus hogares.

Palabras Clave: Sensores Electrónicos, Domótica, Web Services, JAX-WS, JavaME.

I. Introducción

En la actualidad, los dispositivos móviles poseen múltiples áreas de aplicación, son usados como una herramienta para guiar a personas en museos, centros comerciales, excursiones por la ciudad; así mismo, permiten la interacción remota con electrodomésticos y también ayudan a personas con capacidades diferentes, entre otras aplicaciones más. [1]

Por otra parte, cada día resulta ser más evidente la necesidad de realizar la automatización de aparatos electrónicos en los hogares, permitiendo a las personas obtener un mayor confort en su estancia y un ahorro de energía; en ese sentido, el uso de sistemas y dispositivos que proporcionan cierto nivel de automatización dentro de una casa

habitación, nos conduce a la implementación de los principios básicos de la Domótica [2].

Motivados por el interés práctico de implementar las tecnologías móviles y los principios de domótica para el monitoreo de sensores y automatización de dispositivos electrónicos en casas habitación; se realizó una simulación a escala de una casa aplicando los conceptos antes mencionados, con la finalidad de elaborar una propuesta que sirva de punto de partida para los estudiantes que cursan las unidades académicas de Interacción Humano-Computadora e Inteligencia Artificial correspondientes a la Licenciatura en Sistemas Computacionales, para elaboración de sus proyectos de fin de curso.

En el presente trabajo se presenta el desarrollo de los módulos necesarios para crear una interfaz de comunicación a bajo nivel con el puerto serial y sensores electrónicos que emiten señales que son enviadas a una interfaz de alto nivel representada por un servicio web, mismo que se comunica con una aplicación cliente desarrollada con la tecnología JavaME, la cual ha sido instalada en un dispositivo móvil. Así mismo, se realizan recomendaciones para la implementación real de la propuesta en una casa habitación, haciendo mención de los requerimientos electrónicos, tecnológicos y aspectos de seguridad.

II. Desarrollo

Este documento está organizado por módulos de la siguiente manera: en el Módulo I se presenta el desarrollo de la interfaz de bajo nivel que permite la comunicación entre el puerto serial y los sensores electrónicos; en el Módulo II se muestra la implementación de una aplicación que se encarga de leer e interpretar las señales emitidas por los sensores y funciona como puente con la interfaz de alto nivel; el Módulo III está dedicado al desarrollo de los servicios web, los cuales permiten la transformación de datos de control hacia el estándar XML y la lectura de parámetros de control enviados desde la aplicación móvil, lo anterior representa la interfaz de alto nivel; finalmente en el Módulo IV se presenta el desarrollo de la aplicación cliente usando la tecnología JavaME, misma que contiene un componente de comunicación con los servicios web.

Módulo Interfaz de Bajo Nivel

Para la construcción de la interfaz de bajo nivel, se utilizó el microcontrolador PIC 16F88 integrado dentro de un módulo electrónico PICAXE que posee 5 entradas y 8 salidas a rele (ver Fig. 1); el cual es un sistema de microcontrolador que permite su programación de una forma fácil, y que resulta ser de bajo costo, por tal motivo, está logrando un impacto destacable en el ámbito educativo [3].

Con la utilización de este módulo, es posible controlar dispositivos a través de los relevadores de salida, para el caso del proyecto, permitió el control del encendido y apagado de LEDs. Así también, se diseñó un circuito electrónico para la implementación del sensor de temperatura; para ello, se usó un termistor conectado a un amplificador operacional lógico modelo LM741 (ver Fig. 2), con salida de voltaje controlado por un potenciómetro; este circuito permite simular un detector de incendio dentro de la casa habitación, mismo que emite una señal hacia el modulo del microcontrolador que se encuentra conectado al puerto Serial de la Computadora (Servidor).

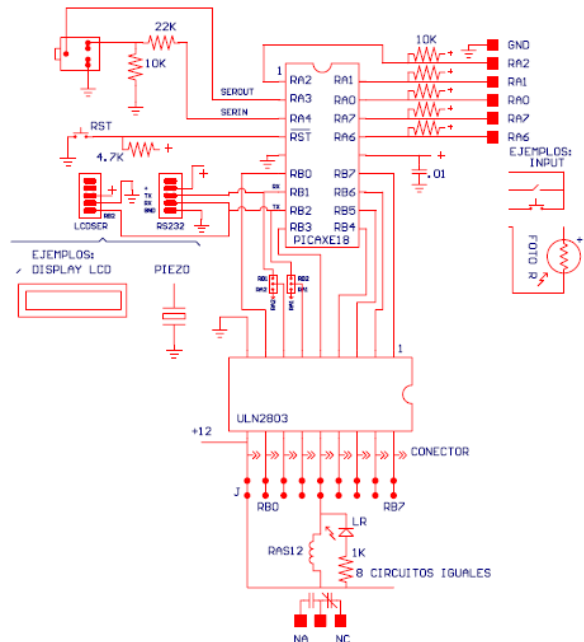


Fig. 1. Diagrama esquemático del Módulo 2202 que posee el Microcontrolador 16F88

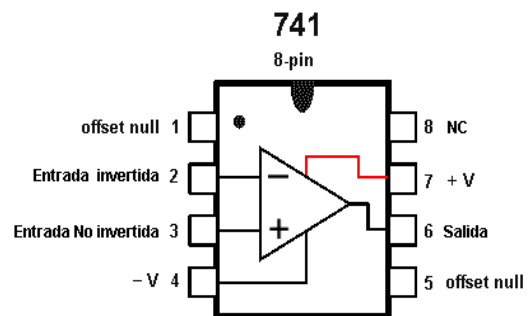


Fig. 2. Diagrama interno del Amplificador Operacional Lógico LM741

Módulo Controlador (Middleware)

La aplicación middleware es la responsable de la interacción entre la interfaz de bajo nivel para la lectura de las señales emitidas por los sensores a través del puerto serial y la interfaz de alto nivel representada por el módulo de servicios web.

Las señales electrónicas emitidas por el circuito electrónico del sensor de temperatura son enviadas a una de las entradas del módulo del Microcontrolador, mismo que transfiere la señal al puerto Serial del Servidor, y es leída por la aplicación middleware desarrollada en el lenguaje de programación Ruby haciendo uso de la librería (ruby-serialports). Las señales son interpretadas como códigos binarios que posteriormente son transformados en datos de control para ser transmitidos a la interfaz de alto nivel, misma que se encarga de interpretarlos y enviarlos al dispositivo móvil a través del protocolo de comunicación SOAP/XML, o bien, enviar un correo electrónico por medio del servicio de correo SMTP. De forma similar, cuando la aplicación middleware es llamada por la interfaz de alto nivel, esta envía un código al módulo PICAXE, que se encarga de activar o desactivar las salidas con la finalidad de controlar el funcionamiento de los diferentes Leds, que representan la automatización de lámparas dentro de la casa habitación. Esta acción, puede ser implementada en un caso real a través del uso de un actuador, que sea capaz de ejecutar y/o recibir una orden desde un controlador y realizar una acción sobre una lámpara (encendido/apagado, apertura/cierre), considerando como Bus a la instalación eléctrica de la propia casa, la cual representa el medio de transmisión que transporta la información entre los distintos dispositivos electrónicos y el Microcontrolador [4].

Las principales operaciones soportadas por la aplicación middleware incluyen: lectura y escritura del puerto serial, interpretación y transformación de los datos binarios, envío de datos de control a la interfaz de alto nivel. Dichas operaciones, se puede visualizar en la figura 3.

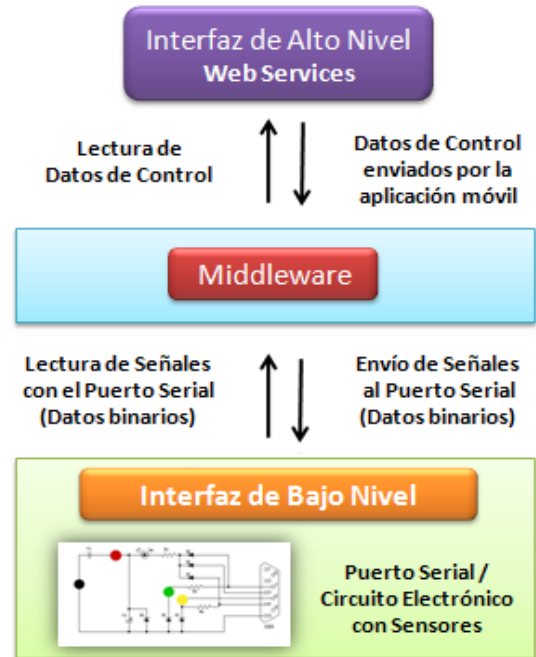


Fig. 3. Comunicación entre las Interfaces de bajo y alto nivel usando como puente la aplicación middleware.

Módulo Interfaz de Alto Nivel (Web Services)

La interfaz de alto nivel es representada por los servicios web, que son los encargados de establecer una comunicación con el dispositivo móvil a través del Protocolo SOAP (Simple Object Access Protocol), el cual define un formato común para el envío de mensajes XML entre aplicaciones usando como medio de transporte la Internet. Así también, se genera el archivo WSDL (Web Services Description Language), mismo que provee el mecanismo mediante el cual las operaciones o funciones de los web services son expuestas al mundo y que las aplicaciones clientes necesitan para poder implementarlas. EL WSDL permite describir los tipos y estructuras de datos usados para el intercambio de mensajes e implementación de las operaciones. [5]

Es importante puntualizar, que se implementan dos operaciones básicas para el Servicio Web "WSCasa", las cuales se denominan Monitorear() y enviarDatos(). Estas operaciones están desarrolladas con la tecnología Java Edición

Empresarial usando el API JAX-WS y son las responsables de recibir los parámetros enviados por el dispositivo móvil para ser interpretados y transferidos hacia la aplicación middleware, quien se encarga de enviar los códigos adecuados al módulo del microcontrolador. La funcionalidad de los web services es representada en la figura 4.

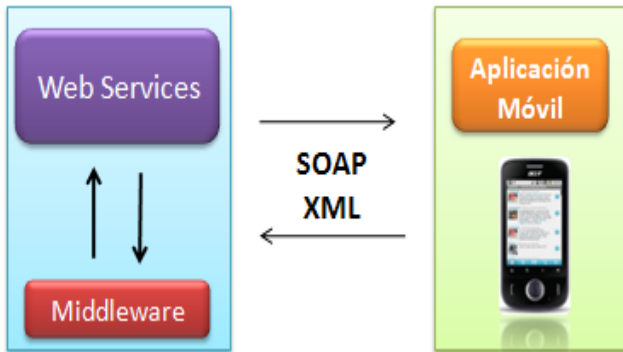


Fig. 4. Comunicación entre la Interfaz de alto nivel y la aplicación móvil, usando el protocolo SOAP.

En el momento que la aplicación móvil se comunica con el WebService para llamar a la operación `Monitorear()`; la aplicación middleware ejecuta la rutina de monitoreo almacenada en el microcontrolador del módulo PICAXE, la cual se encarga de verificar el estado de cada salida y regresa como resultado una serie de códigos que son interpretados por el middleware con la finalidad que la interfaz de alto nivel pueda enviar una respuesta a la aplicación móvil.

Para generar la respuesta emitida por el Servicio Web, se utilizó la especificación SOAP, la cual define la estructura envolvente de codificación de las normas y convenciones para la representación de las invocaciones y respuestas de servicios web. Estas respuestas se transmiten como mensajes SOAP (archivos XML) a través del protocolo HTTP. A continuación, se muestra un ejemplo de la donde se observa que el led número 2 tiene el estado de encendido y todos los demás leds se encuentran apagados.

SOAP Response

```
<?xml version="1.0" encoding="UTF-8"?>
<S:Envelope
xmlns:S="http://schemas.xmlsoap.org/soap/envelope/">
  <S:Body>
    <ns2:MonitorearResponse
xmlns:ns2="http://paq1/">
      <return>
        <led1>0</led1>
        <led2>1</led2>
        <led3>0</led3>
        <led4>0</led4>
        <led5>0</led5>
        <led6>0</led6>
      </return>
    </ns2:MonitorearResponse>
  </S:Body>
</S:Envelope>
```

En el ejemplo, se aprecia una codificación de tipo SOAP que define un conjunto de reglas para el mapeo de datos compuesto de estructuras, tipos de matriz, y tipos de referencia. Para el desarrollo de este proyecto, se utilizaron estructuras de datos a través del enfoque serializar los datos como elementos, y el nombre de cualquier elemento dado coincide con el nombre del campo de datos establecido a nivel de programación [5]. Por lo anterior, se desarrolló una clase Java denominada `Monitoreo`, la cual contiene una serie de campos que hacen referencia a los leds.

Módulo Aplicación Móvil

La aplicación móvil o programa cliente está desarrollado en el lenguaje de programación Java en su versión Micro Edition, la cual se conecta de manera remota al Servicio Web a través de la dirección de dominio correspondiente al Servidor (<http://casaws.serverftp.com:8080/Casa/WSCasaService?WSDL>), dicha URL es configurada dentro del componente denominado `WebService Client`, mismo que se encarga de llamar a las operaciones o métodos del servicio web y les envía parámetros en formato XML usando el protocolo SOAP. En la figura 4 se muestra la interfaz gráfica de la aplicación, en la cual se

aprecia una lista de opciones que son activadas o desactivas y que corresponden a cada una de las áreas que forman parte de la casa inteligente.

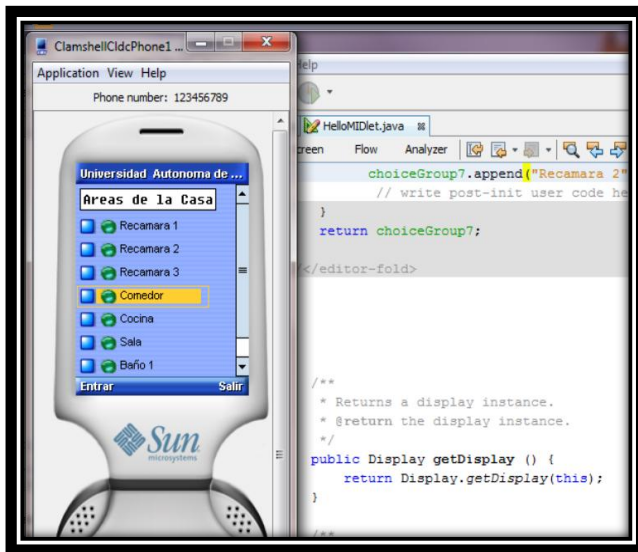


Fig. 5. Aplicación móvil desarrollada en Java Micro Edition usando el componente WebService Client.

Seguridad

Los Servicios Web están basados en arquitectura insegura, la cual usa como medio de comunicación el protocolo HTTP. Este protocolo, es susceptible a una infinidad de amenazas y vulnerabilidades que existen en las redes, específicamente en la Internet. Para brindar seguridad a los Servicios Web, es necesario tomar en consideración los siguientes conceptos: Identificación y Autenticación, Autorización, Integridad, No repudio, Confidencialidad y Privacidad; para poder establecer las políticas de seguridad adecuadas [6].

Así también, es importante señalar que la especificación de invocaciones y respuestas SOAP, no se diseñó pensando en el tema de seguridad; en ese sentido, los mensajes SOAP son vulnerables, a tal grado de ser modificados por usuarios atacantes durante su trayectoria hacia su destino. Sin embargo, hoy en día existen varias opciones disponibles para la seguridad de los mensajes generados por los Servicios Web, destacando principalmente el uso del protocolo

HTTP sobre SSL / TLS (HTTPS), para evitar que los mensajes SOAP que son transmitidos a través de HTTP puedan ser modificados. Para ello, es necesario configurar el servidor de aplicaciones donde se encuentran almacenados de los Servicios Web; para el caso del Servidor GlassFish se requiere ingresar al sistema administrador y activar el servicio SSL en la sección de los puertos de lectura (listeners) del protocolo HTTP (ver Fig. 6).

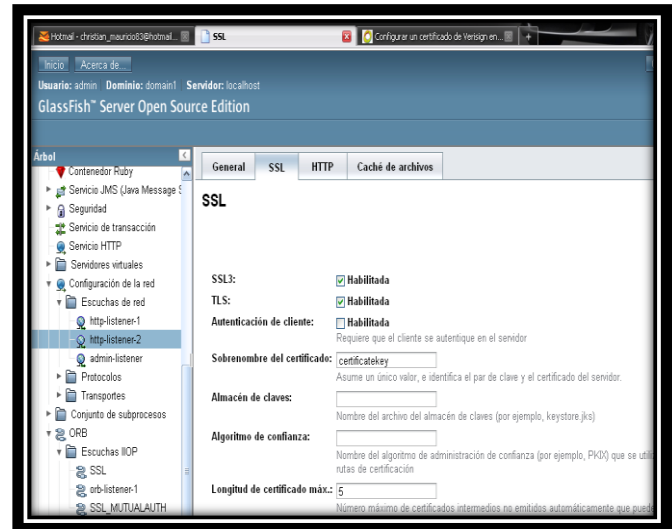


Fig. 6. Consola de Administración del Servidor de Aplicaciones GlassFish.

Para la habilitación del SSL/TSL, se requiere del uso de los certificados y la generación de claves para proveer seguridad en el envío de mensajes; en ese sentido, se recomienda la implementación de la tecnología PKI que permita el cifrado de los mensajes a través de la creación de llaves y certificados para el Servidor y la aplicación cliente utilizando la herramienta keytools, que se incluye en el Kit de Desarrollo del plataforma Java [7].

Así también, es importante considerar las normas de cifrado y firmas XML desarrolladas por el Consorcio World Wide Web (W3C); las cuales permiten que el contenido del mensaje sea firmado y cifrado; lo cual aplica a los mensajes de respuesta de la especificación SOAP que están basados en el estándar XML.

III. Resultados

Para comprobar la funcionalidad del proyecto de la casa inteligente, se realizaron diferentes acciones que consistieron en: configurar el Servidor de Aplicaciones para recibir y enviar información a través el puerto 80 del modem/router, configurar un Dominio de Internet con la finalidad que el Servicio Web sea localizado, y realizar la instalación de la aplicación móvil en un teléfono marca Ericsson modelo T715a. Una vez realizada las acciones, se procedió a iniciar con las pruebas; para ello, se activaron cada una de las opciones que corresponden a las áreas de la casa (recamaras, sala, cocina, etc.) que se muestra en la interfaz gráfica de la aplicación móvil; cuando el usuario seleccionaba una opción, el dispositivo móvil mandaba a una solicitud al Web Service enviándole como datos de control el número de LED debería encenderse, tal como se puede apreciar en la figura 5.



Fig. 5. Aplicación móvil desarrollada en JavaME, siendo ejecutada en un teléfono móvil marca Sony Ericsson modelo T715a.

Así mismo, en la figura 6 se puede apreciar la simulación de un incendio generado en el interior de la casa, resultado de encender un fósforo a una distancia cercana al sensor de temperatura (termistor), el cual al detectar un incremento de temperatura envía una señal a la tarjeta controladora y posteriormente es reenviada hacia

la aplicación middleware, encargada de ejecutar el script que permite enviar al usuario, una alerta de incendio haciendo uso de mensajes de correo electrónico a través del servicio de correo SMTP.

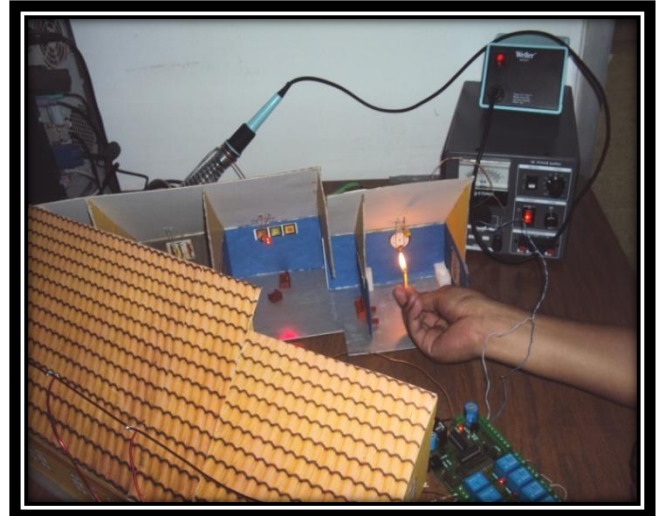


Fig. 6. Simulación de un incendio para activar el sensor de temperatura.

IV. Conclusiones

En la actualidad, la aplicación de la domótica en los hogares resulta ser un tema de gran interés para la sociedad, debido a que permite brindar mayor confort y a la larga un ahorro de energía. Así mismo, un tema relevante es el uso de los dispositivos móviles en nuestra vida cotidiana con un enfoque de herramienta de apoyo que genera grandes beneficios personales. En este proyecto se da a conocer las ventajas de implementar los principios básicos de domótica en combinación con las aplicaciones móviles a través de los Servicios Web, permitiendo monitorear ciertas áreas o elementos de nuestro hogar a través de sensores electrónicos y la automatización de algunos elementos como son las lámparas. Así también, representa un punto de partida para nuestros estudiantes que pertenecen a la Licenciatura en Sistemas Computacionales, al brindarles una propuesta que los motive al desarrollo de proyectos relacionados con la domótica en la unidad académica de Interacción Humano-Computadora; y al mismo tiempo, cumplir con uno de los objetivos planteados en

los planes de estudios; siendo uno de ellos, el desarrollo de proyectos multidisciplinarios. En este sentido, podemos mencionar que el proyecto presentado involucra a distintas disciplinas como son: algoritmos, programación orientada a objetos, electrónica básica, sistemas digitales, programación de dispositivos móviles, entre otras, por lo que consideramos que el desarrollo de este proyecto, permitió a los alumnos colaboradores poner en práctica el aprendizaje obtenido durante el transcurso de su carrera.

Por lo anterior, y con la finalidad de dar continuidad al proyecto que se presenta, se tiene considerado que los alumnos participantes elaboren su tesis documentando este proyecto permitiendo dejar las bases para futuras mejoras, que pueden ser tema de tesis o proyectos de investigación para fin de curso. Para ello, es necesario puntualizar que la etapa en que se encuentra actualmente el proyecto es prueba y adaptación; requiere de la implementación de relevadores de salida que trabajen con voltajes de 110, para poder controlar lámparas reales. Así también, es necesario la adaptación del módulo del microcontrolador (PIC 18F4550) para que pueda conectarse al puerto USB de la Computadora. Así también, una mejora relevante que podría implementarse al proyecto, es la incorporación de una API Java desarrollada por la empresa Descom, que ofrece ciertas clases que permiten los envíos de SMS integrados desde cualquier aplicación Java [10]; con esta integración, se podrá mejorar el desempeño de la aplicación middleware, debido que podrá enviar el aviso de alerta de incendio directamente al dispositivo móvil de la persona a través del envío de un SMS.

Finalmente, es primordial mencionar que este proyecto, es un prototipo que refleja una solución práctica de la aplicación de la domótica; y estamos conscientes, que en un futuro cercano ciertas empresas ofrecerán paquetes tecnológicos para la automatización de dispositivos electrónicos y monitoreo de la casas habitación, incluyendo el uso de dispositivos móviles; lo cual requiere de altos costos en relación a la instalación de cableados especiales y de todos los

requerimientos tecnológicos involucrados, como son: actuadores, microcontroladores, routers, firewalls, entre otros; pero que a la larga brindarán un ahorro de energía, mayor confort y seguridad a las personas.

V. Referencias

- [1] André Santos, Luís Tarrataca, João Cardoso, "*The Feasibility of Navigation Algorithms on Smartphones using J2ME*". Mobile Networks and Applications archive, Volume 15 Issue 6, December 2010, Pág 12. Consultado en ACM Digital Library.
- [2] Domínguez Hugo. "*Domótica: Un enfoque sociotécnico*". CeDInt. 2006. Consultado en http://www.gsi.dit.upm.es/~fsaez/intl/libro_domotica.pdf
- [3] PICAXE. "*What is PICAXE?*". Consultado en <http://www.rev-ed.co.uk/picaxe/>
- [4] Grupo Tecma Red S.L. "*Canales de Domótica/Inmótica - Introducción*". Consultado en <http://www.casadomo.com/noticiasDetalle.aspx?c=14>
- [5] "*DataSheet LM741 Single Operational*". Consultado en el sitio web de componentes electrónicos y semiconductores <http://www.datasheetcatalog.org/datasheet/airchild/LM741.pdf>.
- [6] Anoop Singhal. "*Guide to Secure Web Services*". NIST 2007. Consultado en <http://csrc.nist.gov/publications/nistpubs/800-95/SP800-95.pdf>.
- [7] Cowan Tony. "*Configure Web Services Security with WebSphere: Part 1, HTTPS, .NET, and UsernameToken*". IBM Technical Library. Consultado en http://www.ibm.com/developerworks/websphere/library/techarticles/0504_cowan/0504_cowan.html
- [8] Eric Newcomer. "*Understanding Web services: XML, WSDL, SOAP, and UDDI*". Addison-Wesley, 2002. Consultado en el mes de Mayo de 2011 en <http://books.google.com>.

- [9] Martin Gudgin."SOAP Encodings, WSDL, and XML Schema Types". Consultado en <http://www.xml.com/pub/a/2002/02/20/endpoints.html>.
- [10] SDKDescomSMS."API Java Enviar SMS". Consultado en <http://www.descomsms.com/elegir-api-sms/java/enviar-sms-java.html>.
- [11] Soláns Campo Domingo."Las Nuevas Tecnologías al servicio de los mayores - Domótica". Artículo consultado en <http://mayores.uji.es/proyectos/proyectos2005/domotica.pdf>

Contaduría Pública del Campus IV en la Universidad Autónoma de Chiapas.

Ervin Sánchez López es egresado de la carrera Licenciatura en Sistemas Computacionales en la Facultad de Contaduría Pública del Campus IV en la Universidad Autónoma de Chiapas, actualmente se encuentra en su proceso de titulación.

VI. Autores

L.S.C. Christian Mauricio Castillo Estrada es Licenciado en Sistemas Computacionales por la Universidad Autónoma de Chiapas y se encuentra cursando el último trimestre de la Maestría en Comercio Electrónico con área terminal en Tecnologías de Información en la Universidad Virtual del ITESM. Es profesor de asignatura de la Licenciatura en Sistemas Computacionales en la Facultad de Contaduría Pública del Campus IV en la Universidad Autónoma de Chiapas.

M.C.E. Leonardo Yamasaki Maza obtuvo su título de Maestro en Comercio Electrónico por la Universidad Virtual Instituto Tecnológico y de Estudios Superiores de Monterrey, candidato a Doctor en Informática y Matemática Computacional por Universidad de Valencia, España, y es profesor de tiempo completo en la Universidad Autónoma de Chiapas.

MGTI. Jorge Elpidio Cruz Gálvez obtuvo su título de Maestro en Gestión de Tecnologías de la Información por la Universidad TecMilenio, y es profesor de asignatura de la Licenciatura en Sistemas Computacionales en la Facultad de Contaduría Pública del Campus IV en la Universidad Autónoma de Chiapas.

José Luis Mendoza Villalobos es estudiante de noveno semestre de la carrera Licenciatura en Sistemas Computacionales en la Facultad de

Cálculo de Bioimpedancia Mediante Detección Síncrona de Fase Basado en Hardware Reconfigurable.

Coral Martínez Nolasco, Juan Prado Olivarez, Javier Díaz Carmona, Agustín Ramírez Agundis

Instituto Tecnológico de Celaya, Departamento de Ingeniería Electrónica
Av. Tecnológico y A. García Cubas S/N, Col. Alfredo Bonfil, Celaya, Guanajuato, C.P. 38010
Tel. (461) 611 75 75 ext. 215, correo electrónico: juan.prado@itcelaya.edu.mx

Resumen

En el presente trabajo se describe una propuesta de diseño e implementación del cálculo de bioimpedancia mediante el método de detección síncrona de fase. La solución propuesta es implementada de manera digital en hardware reconfigurable, en donde las señales de referencia y de entrada son generadas utilizando la técnica de Síntesis Directa Digital. El lenguaje de descripción de hardware utilizado en la propuesta de diseño es VHDL y la herramienta de diseño utilizada en la implementación fue System Generator. El sistema de desarrollo utilizado para dicha implementación fue la tarjeta de desarrollo Spartan 3E–Starter Kit de Digilent. Dentro de las pruebas realizadas el retraso de la señal de referencia es generado digitalmente y de acuerdo a los resultados obtenidos es posible concluir que el sistema propuesto es completamente funcional para la estimación de bioimpedancia basada en detección síncrona de fase.

Palabras Clave: Bioimpedancia, Detección Síncrona, FPGA, System Generator.

I. Introducción

Desde el siglo pasado se han realizado investigaciones sobre las propiedades eléctricas de los tejidos biológicos. Existe hoy en día una gran cantidad de metodologías propuestas en donde la mayoría de ellas describen las propiedades en base a la obtención de la impedancia eléctrica del tejido en estudio.

Uno de los métodos comúnmente utilizados para la obtención de la impedancia del tejido es a través de la detección síncrona de la fase entre voltaje y corriente obtenidos de una célula de medición [1].

En la actualidad existen varias técnicas para la medición de impedancia eléctrica compleja, como son los siguientes métodos: del puente, resonante, I-V, RF I-V, de análisis de redes, de puente auto-equilibrio [2].

El detector síncrono de fase analógico presenta inconvenientes de sincronización de señales y para tener un buen funcionamiento se necesitan componentes de alto rendimiento (offset,

linealidad y estabilidad). Mientras que en una aplicación digital, el principio fundamental de la detección síncrona de fase analógica se conserva. La principal diferencia es que la señal de entrada es digital.

La señal de referencia, bien pueden ser sintetizada a través del uso de una tabla de valores en la memoria de la computadora o por un generador de señales digitalizadas. La primera solución es más eficiente porque la señal de referencia es carente de ruido (si se codifica en un número suficiente de bytes) y su amplitud es conocida y estable.

La aportación de éste trabajo es el diseño e implementación de un sistema que permita el cálculo de la bioimpedancia eléctrica compleja, haciendo uso de la técnica de detección síncrona de fase entre voltaje y corriente obtenidos de una célula de medición.

La siguiente sección describe brevemente el concepto de bioimpedancia y el método de

medición considerado en el proyecto. En la cuarta sección se presentan los bloques que componen el esquema de medición utilizado. La propuesta desarrollada se presenta en la quinta sección. La sexta y quinta secciones se presentan los resultados y las conclusiones obtenidas, respectivamente.

II. Bioimpedancia

La bioimpedancia eléctrica se define como la medición de la impedancia eléctrica de una muestra biológica. La impedancia eléctrica es la oposición que presenta un cuerpo, en nuestro caso un tejido biológico (por ejemplo piel humana), al paso de una corriente a través de él. Este parámetro puede reflejar algunas condiciones fisiológicas interesantes.

La impedancia (Z) es un número complejo definido como la relación entre la tensión medida (V) y el flujo de corriente total (I) [3].

Un elemento indispensable para la medición de bioimpedancia es el electrodo, el cual actúa como transductor entre la corriente electrónica en el sistema de medida y la corriente iónica del tejido biológico. Esto crea una discontinuidad que se resuelve mediante un mecanismo de conducción, en la corriente de desplazamiento en los electrodos y reacción de oxidación-reducción de los electrodos no polarizables [3].

La impedancia del electrodo (Z_e) puede ser modelada por una resistencia (R) y una reactancia (X) en serie cuyo valor decrece cuando la frecuencia se incrementa.

$$Z_e = R - jX \quad (1)$$

A pesar de que se puedan identificar los parámetros de la impedancia del electrodo, no se puede determinar un modelo preciso, ya que la magnitud depende del metal del electrodo, de su área, del electrolito utilizado, de la densidad de corriente y de la frecuencia de la corriente utilizada en las medidas.

Los electrodos que se utilizan varían de acuerdo a la aplicación. En nuestro caso, las medidas contempladas a futuro son sobre el tejido de la piel humana, por lo cual el uso de electrodos superficiales serían los adecuados, como se muestra en Fig.1. El método de medición a considerar es a través de un arreglo de 4 electrodos, el cual se describe a continuación.

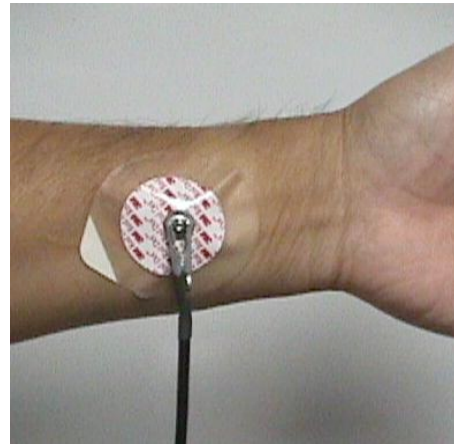


Fig. 1. Electrodo superficial.

Método de medición de 4 electrodos.

Este método consiste en aplicar una corriente (I_o) a través de dos electrodos (Z_{e1} y Z_{e4}) y detectar la tensión en otros dos electrodos distintos (Z_{e2} y Z_{e3}), este procedimiento se puede apreciar en la Fig. 2.

Con este método se consigue eliminar la impedancia de contacto de los electrodos, siempre y cuando los electrodos empleados tengan una impedancia pequeña en comparación con la impedancia de entrada del circuito utilizado para detectar la tensión V_o y no provoquen una saturación de los circuitos que inyectan la corriente I_o . En el caso de realizar medidas superficiales, también se reduce la impedancia de la piel ya que queda en serie con cada electrodo [4].

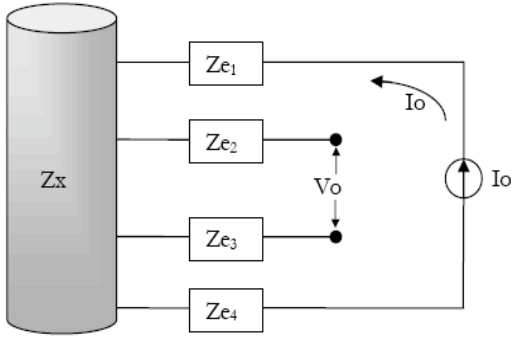


Fig. 2. Configuración del método a 4 electrodos.

III. Detección síncrona

Un detector de fase es un circuito mezclador de frecuencias o multiplicador analógico que genera una señal de voltaje que representa la diferencia en fase entre dos señales de entrada [5].

El sistema detecta la diferencia en fase de dos señales periódicas. Una señal periódica está caracterizada por tener amplitud, factor de forma, frecuencia y fase, siendo esta última la que es detectada por el sistema. El sistema proporciona la información en forma de corriente que sale o entra del dispositivo dependiendo del resultado de la diferencia entre las fases de las entradas, es decir la corriente puede ser positiva o negativa, lo cual es una indicación de si la señal está adelantada o atrasada respecto de una referencia de tiempo dada.

En la detección síncrona de fase dos señales, un voltaje de referencia S_{REF} y un voltaje de entrada V_{IN} , estas son originalmente señales senoidales desfasadas una de la otra, obtenidas de la célula de medición. El objetivo es determinar dicho desfaseamiento, para ello ambas señales se multiplican obteniendo como el resultado la componente DC proporcional a dicho desfaseamiento. Para ello se aplica un filtro pasabajas (promediador) a la salida del multiplicador como se muestra en la Fig. 3.

A la salida del filtro se obtiene la parte real e imaginaria de voltaje y corriente, con las cuales se determina la impedancia compleja [6] [8] [11].

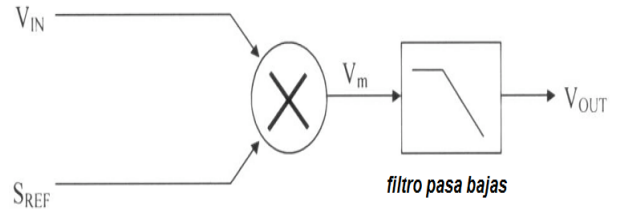


Fig. 3. Detector síncrono de fase.

Por lo tanto, con $V_{IN} = A_1 \sin(\omega_0 t + \phi)$ y $S_{REF} = A_2 \sin(\omega_0 t)$, después del filtraje se tiene:

$$V_{OUT} = -\frac{A^* A_1}{2} \cos(\phi) \quad (2)$$

Que es la parte real de la señal.

Tomando $V_{IN} = A_1 \sin(\omega_0 t + \phi)$ y $S_{REF} = A_2 \cos(\omega_0 t)$, entonces después del filtraje se tiene:

$$V_{OUT} = \frac{A^* A_1}{2} \sin(\phi) \quad (3)$$

siendo igual a la parte imaginaria de la señal.

IV. Esquema de medición

Se diseñó un detector síncrono de fase digital con la finalidad de evitar los inconvenientes que se presentan cuando se realiza de manera analógica. Ver Fig. 4. El principal inconveniente es la atenuación de las señales a altas frecuencias (1-10MHz).

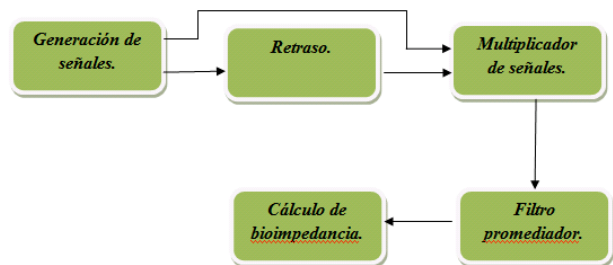


Fig. 4. Diagrama a bloques del esquema de medición.

Las etapas principales de la detección síncrona de fase digital son:

- a) **Generación de señales:** Este bloque se encarga de generar dos señales (senoidal y cosenoidal) mediante la técnica DDS.
- b) **Retraso.** Este bloque representa el retraso que sufre la señal de referencia (señal generada con el método DDS) debido a la bioimpedancia del tejido en medición. Siendo su salida la señal obtenida como voltaje de los electrodos de medición.
- c) **Multiplicador de señales.** Este bloque se encarga de realizar el producto necesario dentro de la detección síncrona de fase de acuerdo a la Fig. 3.
- d) **Filtro promediador.** Se trata de un filtro pasabajas promediador, el cual obtiene el promedio de la señal obtenida a la salida del bloque multiplicador.
- e) **Cálculo de Bioimpedancia.** Este bloque se encarga de dividir V/I para obtener la bioimpedancia Z .

V. Sistema Propuesto

El sistema propuesto consiste en la implementación de cada una de las etapas que conforman la detección síncrona de fase (descritas en la sección anterior) utilizando la tarjeta de desarrollo Spartan 3E–Starter Kit de Digilent, mostrada en la Fig. 5

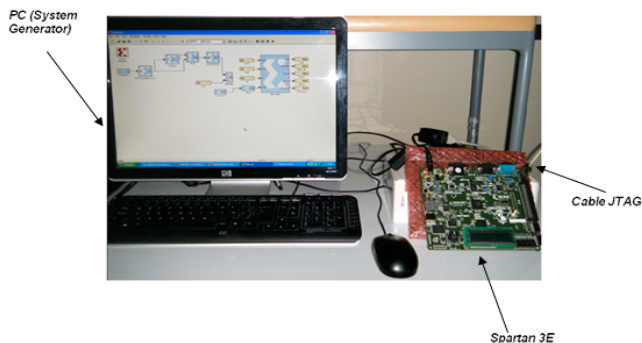


Fig. 5. Tajeta de desarrollo.

Generación de señales.

La generación de señales se realizó mediante el método (DDS) [7], para la implementación se utilizó el bloque DDS Compiler 4.0 de System Generator configurando los siguientes parámetros: $F_{Clock} = 50 \text{ MHz}$, $F_{salida} = 1 \text{ MHz}$

Retraso

Para el proyecto se simulará ese desfase, ya que las etapas anteriores a la etapa de detección síncrona de fase aún no se tienen implementadas. El desfase se implementó con el bloque Delay de System Generator. El parámetro con el cual se indica el desfase es *Latencia*. Dependiendo de la latencia es el desfase en grados que se obtiene en la señal.

Multiplicador de señales

Se desarrolló el multiplicador de señales el cual es una de las etapas más importantes del detector síncrono de fase. En esta etapa se multiplican dos señales: la señal de referencia y la señal de entrada, siendo ambas señales senoidales, sólo que desfasada una con respecto a la otra. Para la simulación e implementación de esta etapa se utilizó el bloque **Mult** de System Generator. Los parámetros utilizados son: Precisión: Full y Latencia: 3.

Filtro Promediador.

Los filtros digitales tienen muchas ventajas sobre los filtros analógicos, por lo cual ha sido muy amplia la utilización y aplicación de éstos. Sin embargo, la implementación de los filtros digitales en hardware reconfigurable no resulta tan simple, o en ocasiones el diseño se enfoca a alguna aplicación específica, dando como resultado la dificultad de rediseñarlos. La frecuencia de corte es el valor de la frecuencia a partir del cual se considera que ya está filtrando las señales. El orden de un filtro describe el grado de aceptación o rechazo de frecuencias por arriba o por debajo de la respectiva frecuencia de corte.

Se diseñó un filtro FIR tipo ventana con los siguientes parámetros: Frecuencia de corte: 0.0025π , Orden del filtro: 50, Ventana: Hamming.

$$Z = \frac{V}{I} = \frac{a + bj}{c + dj} \tag{4}$$

En la Fig. 6 se muestra el diagrama a bloques del sistema propuesto en System Generator para el cálculo de la bioimpedancia.

Cálculo de Impedancia.

Como etapa final para obtener la impedancia compleja se empleó el método de división de dos números complejos.

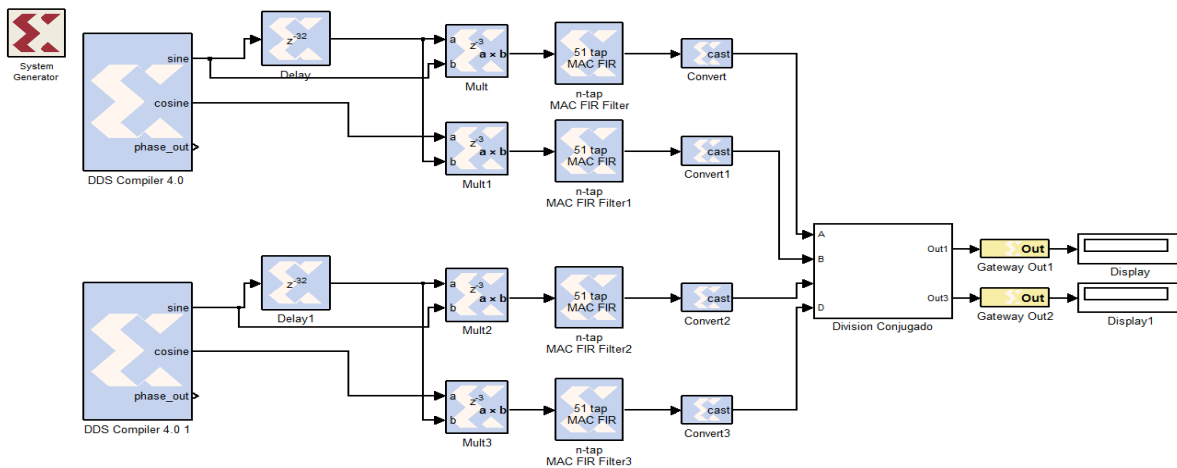


Fig. 6. Diagrama a bloques para el cálculo de la bioimpedancia.

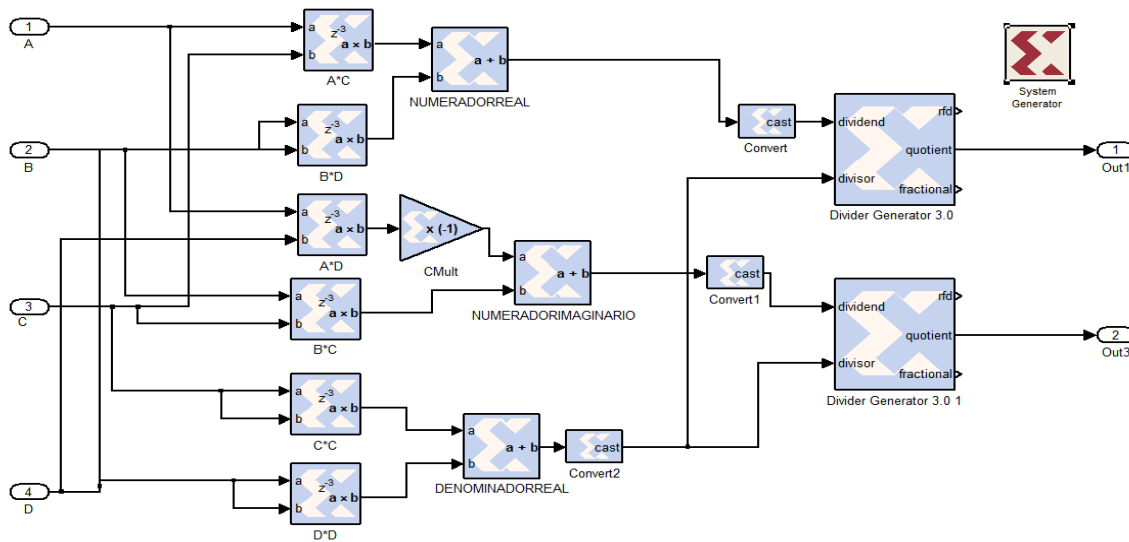


Fig. 7. Diagrama a bloques de la división compleja.

En la Fig. 7. Se observa el diagrama a bloques de la división compleja el cual está contenido en el subsistema del cálculo de bioimpedancia.

VI. Resultados

Se generaron las señales de referencia y de entrada con la técnica DDS, estas señales se muestran en las Fig. 8 y Fig. 9.

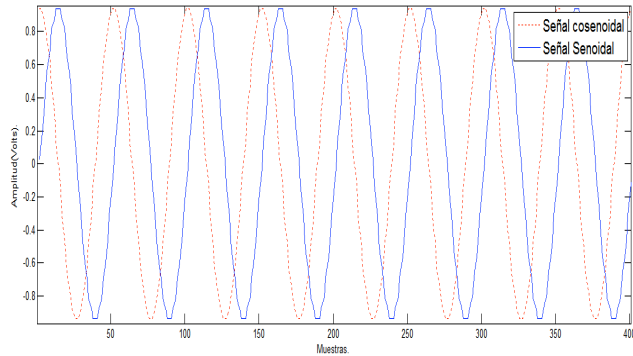


Fig. 8. Señales generadas para parte imaginaria.

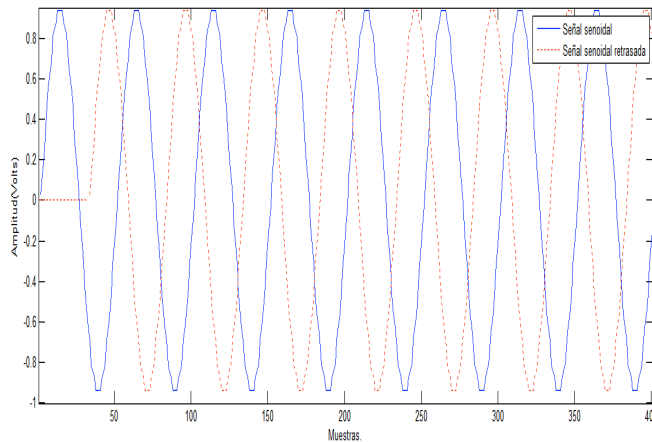


Fig. 9. Señales generadas para parte real.

En Fig. 10. Se muestra que la amplitud de la señal de referencia y de la señal de entrada es de 0.469 V por lo tanto la amplitud de la señal del producto de las señales es igual a $0.469 \text{ V} * 0.469 \text{ V} = 0.2199 \text{ V}$.

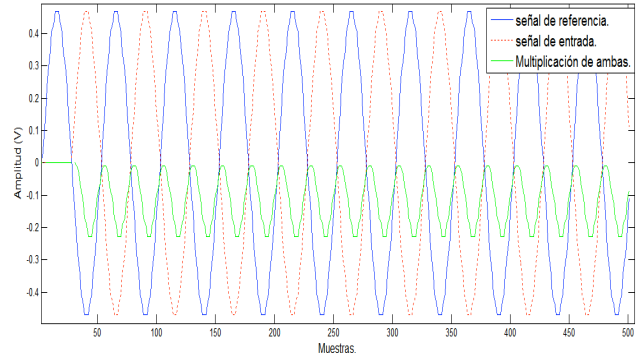


Fig. 10. Multiplicación de señales.

En la Fig. 11 se muestra la gráfica de error de la parte real del voltaje con porcentaje de error máximo de 5.5%.

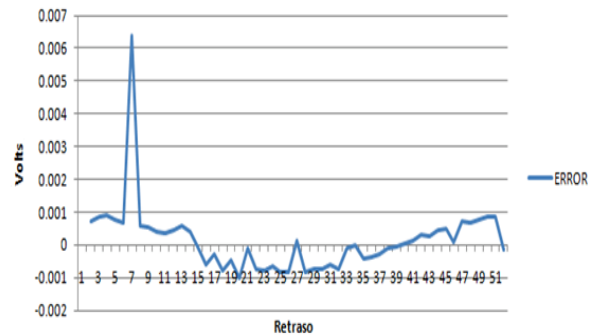


Fig. 11. Error voltaje real.

En la Fig. 12 se muestra la gráfica de error de la parte imaginaria del voltaje con porcentaje de error máximo de 3.84%.

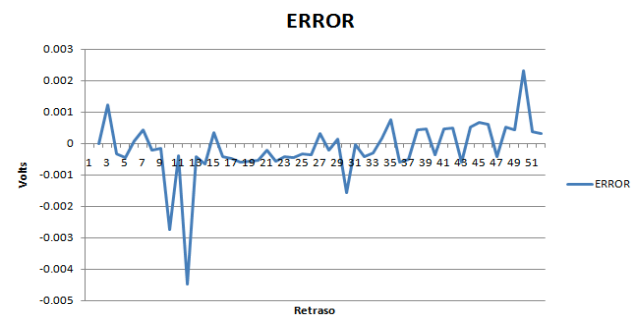


Fig. 12. Error voltaje Imaginario.

Se obtuvieron los resultados de impedancia calculados y medidos, los errores para varios retrasos se ven en la tabla 1.

$$\text{Error} = \frac{\text{Valor calculado} - \text{Valor medido}}{\text{valor calculado}} * 100, \quad (5)$$

Delay	Retraso	Error
0	0 °	0.64%
5	36 °	6.81%
8	57.6 °	0.68%
43	309.6 °	0.71%
48	345.6 °	0.81 %

Tabla 1. Error de Bioimpedancia para diferentes retrasos.

Peak Memory Usage	2 de 20	10 %
Total REAL time to MAP completion	7 secs	
Total CPU time to MAP completion	7 secs	

Tabla 2. Utilización de recursos del FPGA XC3S500E.

Logic Utilization		
Number of Slice Flip Flops	2606 de 9312	27 %
Number of 4 input LUTs	2460 de 9312	26 %
Logic Distribution		
Number of occupied Slices	1568 de 4656	33 %
Number of Slices containing only related logic	1568 de 1568	100 %
Number of Slices containing unrelated logic	0 de 1568	0 %
Total Number of 4 input LUTs	2596 de 9312	27 %
Number used as logic	1754	
Number used as a route-thru	136	
Number of bonded IOBs	12 de 232	5 %
Number of BUFGMUXs	1 de 24	4 %
Peak Memory Usage	2 de 20	10 %

VII. Conclusiones

En este trabajo se propuso un sistema para calcular la bioimpedancia eléctrica de un tejido biológico utilizando el método de detección síncrona de fase utilizando hardware reconfigurable, se comprobó su correcto funcionamiento implementándolo de manera digital.

Se obtuvieron los errores en base a los resultados medidos y calculados de bioimpedancia dando el mayor error

El presente proyecto está enfocado al desarrollo de la parte digital en el cálculo de la bioimpedancia, dejando como trabajo a futuro el desarrollo de la etapa analógica, por ello el retraso de la señal de referencia es generado digitalmente.

VIII. Referencias

- [1] Martínez N. C. "Diseño e implementación de medidor de Bioimpedancia: etapas de detección sincrónica de fase y control digital"; Tesis, Instituto Tecnológico de Celaya, 2011.
- [2] Y Salazar Muñoz, "Introducción a la medida de impedancia eléctrica de los tejidos biológicos", tutorial.
- [3] M. Parve, T. Ronk, A. Annus, P. Paavle Min, "Synchronous Sampling and Demodulation in an Instrument for Multifrequency Bioimpedance Measurement," *Instrumentation and Measurement, IEEE*, no. 9606767, pp. 1365-1372, Aug. 2007.
- [4] J.F.,(2000), "Conception et réalisation d'une cellule de caractérisation des tissus biologiques par spectroscopie de bioimpédance dans la gamme fréquentielle [100 Hz-1MHz], application aux tissus osseux, prise en compte de l'anisotropie", Thèse, Université Henri Poincaré, Nancy.
- [5] Laureano A. Bulus Rossini, José M. Juárez, and Hugo E. Lorente, "Sintetizador de frecuencia PLL controlado por DDS," in *XII Reunion de trabajo en Procesamiento de la información y control*, Universidad Nacional de la Plata, Argentina., 2007, pp. 1-6.
- [6] XILINX, *Spartan-3E Starter Kit Board User Guide.*, 2006.
- [7] J. Stoltenberg, D. Pengra, R. VanDyck, and O. Vilches, "The phase sensitive (Lock-in) Detector".
- [8] Sigfredo Pagel, "Síntesis Digital Directa de frecuencias, DDFS," in *Revista Española de Electrónica*, La Rioja, España., 2007, pp. 66-76.

IX. Autores

M. en C. Coral Martínez obtuvo su grado de Maestría en Ciencias en Ingeniería Electrónica con especialidad Sistemas Digitales en Junio del 2011.

Dr. Juan Prado Olivarez obtuvo el grado de doctor en 2006 por la Université Henri Poincaré, Nancy I en Francia. Actualmente es profesor investigador de tiempo completo en el departamento de Ing. Electrónica del Instituto Tecnológico de Celaya. Realiza trabajos en el área de la instrumentación especialmente biomédica y en particular sobre espectroscopia de impedancia eléctrica en tejidos biológicos (Bioimpedancia).

Dr. José Javier Díaz Carmona obtuvo el grado de doctor en 2003 por el Instituto Nacional de Astrofísica, Óptica y Electrónica. Actualmente es profesor investigador de tiempo completo en el departamento de Ing. Electrónica del Instituto Tecnológico de Celaya. Sus áreas de interés abarca el procesamiento digital de señales, diseño de filtros digitales, así como desarrollo de soluciones basadas en sistemas digitales empleando FPGAs, microcontroladores y procesadores DSP.

Dr. Agustín Ramírez Agundis obtuvo el grado de doctor en 2008 por la Universidad Politécnica de Valencia. Actualmente es profesor investigador de tiempo completo en el departamento de Ing. Electrónica del Instituto Tecnológico de Celaya. Sus áreas de interés son el desarrollo de soluciones basadas en sistemas digitales empleando FPGAs y microcontroladores, así como la integración de sistemas automáticos para aplicaciones industriales.

Método Geométrico para la Orientación Ortogonal en Luces Robóticas

Sandoval Gutiérrez Jacobo, López Ortega Jorge, Cosme Aceves José Francisco y Ferreyra Ramírez Andrés.

Departamento de Electrónica, División de Ciencias Básicas e Ingeniería, Universidad Autónoma Metropolitana
Unidad Azcapotzalco, México D.F., México, jsg@correo.azc.uam.mx

Resumen

El presente trabajo describe un método de análisis geométrico de iluminación dinámica sobre la programación de luces robóticas de un escenario y objeto conocido (sistema de referencia). La problemática de la iluminación se enfoca en obtener una orientación ortogonal entre los robots y objetivo(s); de manera común el programador requiere de tener los equipos en funcionamiento, así como las condiciones de luz ambiental suficientes para visualizar las orientaciones respectivas, esto se logra por una técnica cualitativa; sin embargo, con el método propuesto se optimiza la forma de programación en comparación al método de aproximación por prueba y error.

Palabras Clave: Programación de robots, Sistemas de iluminación, Tecnologías de entretenimiento.

I. Introducción

La iluminación dinámica [2], [4], [5] es un sistema que evoluciona en el tiempo, formando escenas en un espacio tridimensional, refiriéndose a orientar los haces de luces a un objetivo en específico a través de una combinación de múltiples entradas digitales en eventos discretos [3], tal como se muestra en la Fig. 1

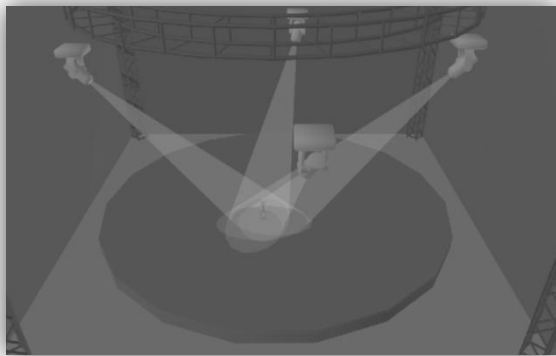


Fig. 1. Representación de la iluminación dinámica.

De la Fig.1 se tiene un objeto que se ha iluminado, orientando las luces robóticas en una forma independiente. Para lograr dicha configuración, la programación comúnmente se

ejecuta mediante ensayos de prueba y error; así mismo, las condiciones reales de iluminación deben ser las adecuadas para distinguir de manera visual los cambios, por mencionar un ejemplo: la restricción de luz solar evita trabajar en la programación debido a que los cambios no se notan. El hecho de trabajar en condiciones de alta luminosidad o restar vida útil a las luces robóticas, crean la necesidad de aplicar métodos alternos a los tradicionales para ser aplicados en forma real.

En este sentido, el presente método es una propuesta de programación virtual que es fácil de implementar para este tipo de aplicaciones.

II. Conceptos de luces robóticas

II.1. Definiciones básicas

Las luces robóticas se clasifican por el tipo de iluminación que proporcionan según las siguientes características [1]:

- *Consumo:* estas se conocen por medio de la potencia utilizada en relación a su luminosidad: 1000 watts, 2500 watts, etc...
- *Por el tipo de fuente:* Lámpara y LED.

- *Textura que proyecta:* Puede ser con globos (figuras) o sin ellos.
- También por la cantidad de opciones que puede dar, como son: colores, texturas, frecuencia de parpadeo (todas referidas en canales por cada propiedad)

II.2. Protocolo/Interfaces de comunicación.

La comunicación estándar utilizado para las luces robóticas es conocida como Digital MultipleX basado en el protocolo RS-485 [6]. La comunicación se logra mediante envío de una trama de 512 paquetes (canales) de datos con una longitud de ocho bits cada uno. Cada 512 canales se refieren a un universo, por lo tanto, para mayor cantidad de canales podemos utilizar múltiples universos.

Lograr una correcta comunicación, requiere de conectar en serie todas las luces robóticas y asignar a cada una de ellas una dirección diferente, en otras palabras ningún canal debe estar asignado a más de un robot. Por lo tanto, una comunicación mínima se logra mediante un sistema de envío de paquetes de datos por medio de una consola, computadora y/o sistema en conjunto de luces robóticas.

II.3. Parámetros de entrada y salida

El algoritmo utilizado para controlar las luces robóticas depende de las condiciones del hardware, del escenario a iluminar y de la evolución de las escenas. Es decir, una programación secuencial de un sistema de múltiples entradas múltiples salidas MIMO.

La combinación de parámetros a la entrada que permite lograr una escena en una tonalidad #A6448B, es necesaria la combinación de los canales para los colores primarios Magenta, Cyan y Amarillo (CYM), señalando que cada canal es un byte (HxFF). Por lo tanto, la combinación resultante se expresa como #000000 (Negro) y #FFFFFF (Blanco).

De igual forma, para lograr una orientación específica debemos combinar los parámetros de x (pan) y y (tilt), tal como se muestra en la Fig. 2.

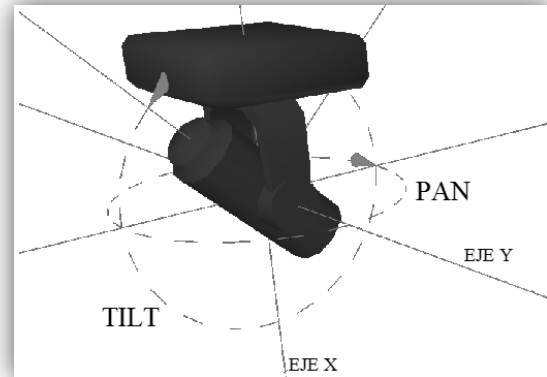


Fig. 2. Ejes de rotación de pan y tilt (x & y).

Los parámetros para combinar dependen del número de canales de cada robot, por citar un ejemplo para dos luces robóticas de cinco canales, en primer lugar se asigna una dirección de las 512 disponibles y se suma la cantidad de canales de la primera para operarlas en forma independiente, como se muestra en la Tabla.1.

Canal DMX	No. robot	Descripción
1	1	Pan x
2	1	Tilty
3	1	Shutter
4	1	Color
5	1	Gobo
6	2	Pan x
7	2	Tilt y
8	2	Shutter
9	2	Color
10	2	Gobo

Tabla 1. Direccionamiento de parámetros de luces robóticas

Una escena es el resultado de una combinación de una cadena de 512 paquetes de ocho bits, es

decir, $(2^8)^{512}$ combinaciones si utilizamos todo un universo.

La escena en negro puede ser obtenida al cerrar los canales de los obturadores, con ello todas las demás combinaciones no serán visibles, sin embargo las demás posiciones se mantienen almacenadas en memoria, la idea es utilizar estos lapsos para acercarse a la nueva posición logrando una transición suave. En forma análoga todas las escenas son logradas mediante combinaciones diferentes en los parámetros de cada robot.

III. Descripción del método de orientación

Las trayectorias son compuestas por las rotaciones sobre los ejes x y y respecto del objetivo, de tal forma, podemos expresar la geometría como una matriz de rotación [9] para las luces robóticas

$$R_i = \begin{pmatrix} c\alpha c\beta & -s\alpha & -c\alpha s\beta & x \\ s\alpha c\beta & c\alpha & s\alpha s\beta & y \\ -s\alpha & 0 & c\alpha & z \\ 0 & 0 & 0 & 1 \end{pmatrix}_i \quad (1)$$

Donde:

- c : coseno
- s : seno
- α : ángulo formado por la base y el eje y
- β : ángulo formado por la base y el eje x
- x, y, z : posición del robot
- i : número de luz robótica

y la matriz para cada objeto j como:

$$\begin{pmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{pmatrix}_j \quad (2)$$

Donde:

- j : número del objeto

x_m, y_m, z_m : posición del objeto

La primera etapa consiste en definir las posiciones de los objetos, direccionamiento de las luces robóticas y definir una posición inicial, tal como se muestra en la Fig. 3.

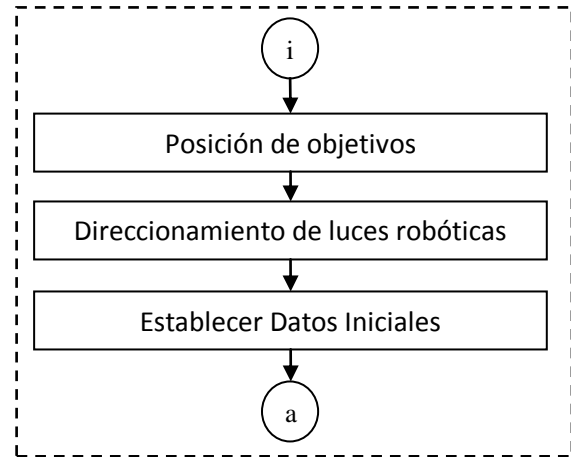


Fig. 3 Diagrama de flujo de la etapa de referencia inicial del sistema.

De la Fig. 3, se establecen las condiciones iniciales que permitirán aplicar la programación de manera virtual.

En la segunda etapa se definen las razones de cambio para las variables de orientación acopladas $\overrightarrow{(\theta_i, \varphi_i)}$ en donde θ_i es el valor angular para *tilt* y φ_i para *pan*, mostrado en la Fig. 2. De forma inicial se tiene q_{i-1} como la posición inicial, por lo tanto al variar las orientaciones se describen las posiciones relativas como:

$$q_i = q_{i-1} \cdot \overrightarrow{(\theta_i, \varphi_i)} \quad (3)$$

Por lo tanto, con la unión de cada i se obtiene la escena deseada como el conjunto de S_n

Al cumplir:

$$\bigcup_{i=1}^n q_i = S_n \quad (4)$$

Por otra parte la orientación ortogonal de las luces robóticas respecto de un objeto se cumple cuando el conjunto O_j se posicionan respecto del objeto:

$$\bigcap_{i=1}^m q_i = O_j \quad (5)$$

Donde:

n : total de luces robóticas

m : total de luces robóticas que se posicionan ortogonalmente al objeto.

La hipótesis es planteada de acuerdo a las siguientes consideraciones: se conocen la posición de un objeto m ubicado en el espacio de operación, en forma reducida de (2):

$$O_m = \begin{pmatrix} x \\ I & y \\ z \\ \vec{0} & 1 \end{pmatrix}_m$$

Así mismo, una luz robótica con una posición inicial q_{i-1} , en forma reducida de (1).

$$q_{i-1} = \begin{pmatrix} q_x \\ R_{i-1} & q_y \\ q_z \\ 0 & 1 \end{pmatrix}_{i-1}$$

Donde

R_{i-1} : Orientación inicial

q_x, q_y, q_z : Posición inicial en coordenadas absolutas.

Calculando los parámetros de (3) como la distancia ortogonal de O_m respecto de q_{i-1} en el espacio:

$$k_i = \|q_{i-1} - O_m\|$$

Donde:

k_i : Tiene las componentes, de $k_{i_x}, k_{i_y}, k_{i_z}$ correspondientes a las los ejes de coordenadas

De este modo, se calculan los valores correspondientes de θ_i y φ_i

$$\theta_i = \arctan2\left(\sqrt{k_{i_x}^2 + k_{i_y}^2}, k_{i_z}\right)$$

$$\varphi_i = \arctan2\left(k_{i_y}, k_{i_z}\right)$$

De esta forma, se obtienen las rotaciones de *pan* φ_i y *tilt* θ_i expresadas en grados, por lo que debe realizar la conversión a valores digitales de ocho bits. Por lo tanto, $\varphi_{i_{max}} = \text{pan}_{max}$ y $\theta_{i_{max}} = \text{tilt}_{max}$ se mantiene un $\varphi_i = \text{pan}_i$ y $\theta_i = \text{tilt}_i$.

Simplificando, la segunda etapa consiste en calcular las razones de cambio y definir la relación entre objetos y luces robóticas, tal como se muestra en la Fig. 4.

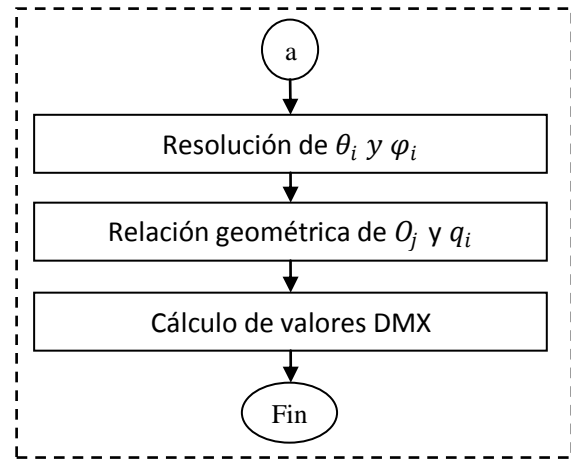


Fig. 4 Diagrama de flujo de la etapa de asignación de luces robóticas a objetos

Unificando la Fig.3 y 4, se establece el método para orientar y programar el sistema respecto de objetos particulares.

IV. Pruebas.

Las pruebas han sido ejecutadas en el simulador Daslight Virtual Controller [8], con una luz

robótica VL2500 Spot [7] y dos objetivos en un solo universo DMX512.

Se han considerado los siguientes datos para una luz móvil con $\text{pan}_{\max} = 540^\circ$ y $\text{tilt}_{\max} = 270^\circ$ y la posición y orientación inicial [2].

$$q_{i-1} = \begin{pmatrix} 1 & 0 & \frac{\sqrt{2}}{2} & 10 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Así mismo un objeto en una posición:

$$o_i = \begin{pmatrix} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & 12 \\ 0 & 0 & 1 & 1.48 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Se obtiene

$$k_i = 4.39$$

$$\theta_i = 124.95$$

$$\varphi_i = 33.69$$

Con los valores encontrados de orientación, los canales respectivos son expresados en valor decimal como:

$$\text{pan}_i = 26.81 \approx 27$$

$$\text{tilt}_i = 84.26 \approx 84$$

El valor de aproximación se debe a la restricción de manejar enteros en incrementos de unidad.

V. Resultados

Las simulaciones en Daslight Virtual Controller y un escenario de prueba, se muestra en la Fig. 5.



Fig. 5 Escenario propuesto para las pruebas.

Como se muestra en la Fig. 5 el robot se encuentra detrás del objetivo, por lo tanto la rotación con los valores calculados de $\text{pan}_i \approx 27$ y $\text{tilt}_i \approx 84$, dan por resultado la orientación de la Fig. 6, en los canales de la Tabla 2.

Canal DMX	Valor	Descripción
1	255	Dimmer
2	27	Pan x
4	84	Tilty
17	255	Iris

Tabla 2. Valores DMX y descripción del canal para el objetivo 1

De la Tabla 2, se describen cuatro de los 21 canales útiles, 2 y 4 corresponden a la orientación, mientras 1 y 17 para iluminar.

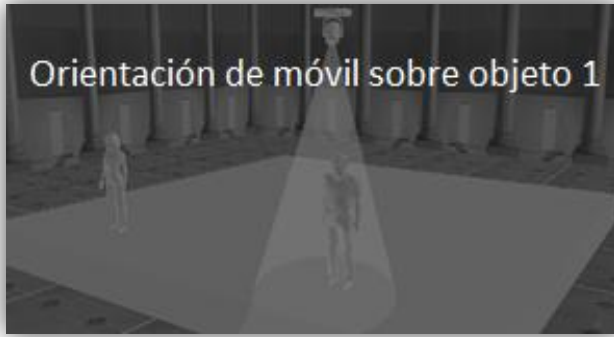


Fig. 6 Iluminación sobre el objeto 1.

Cómo se puede observar en la Fig. 6 la orientación resultante más la apertura de iris y atenuación del móvil nos da un resultado correcto que debe seguir la iluminación.

Para el objeto 2 se ha propuesto una posición corrida en ± 1 bit para conocer el resultado en forma visual de una incertidumbre producida por cada bit, dicha variación ayuda a estimar la precisión de la luminaria en forma cualitativa, tal como se muestra en la Fig. 7 con los datos de la Tabla 3. El objeto 2 tiene la siguiente posición dentro del ambiente propuesto.

$$O_i = \begin{pmatrix} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 1.44 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Canal DMX	Valor	Descripción
1	255	Dimmer
2	60	Pan x
4	84	Tilty
17	255	Iris

Tabla 3. Valores DMX y descripción del canal para objeto 2

De la Tabla 3, se tiene los valores correspondientes para un segundo objeto.



Fig. 7 Iluminación sobre el objeto 2.

Como se puede observar en la Fig. 7, la orientación que sigue la iluminación se encuentra corrida y su error puede notarse en esta perspectiva.

VI. Discusión

El método de análisis propuesto únicamente requiere las posiciones de los objetivos por iluminar, ya que las orientaciones de todos los móviles son calculadas automáticamente. En este sentido la programación puede ser realizada conociendo solamente los planos del escenario.

En términos prácticos, aplicar el método requiere de un conocimiento previo para utilizarlo de forma adecuada, en cambio la programación con parámetros cualitativos es más intuitiva.

VII. Conclusiones

El método proporciona una lógica de programación con parámetros cuantificables en contraste con la programación por ensayo de prueba y error.

La programación puede darse fuera de línea, razón por la que se optimizan los tiempos, así mismo, se ahorra en el consumo de energía ya que no se requiere del encendido de las mismas para programar todo el universo de escenas.

El problema de interpretar las rotaciones compuestas x, y es minimizado al dar el valor adecuado para alcanzar un mismo objetivo con varios móviles u otras combinaciones deseadas.

VIII. Referencias

- [1] Lighting Control Protocols, IES TM-23-11, Illuminating Engineer Society, 2011. ISBN # 978-0-87995-252-5.
- [2] Wei Jiang ,Yujian Jiang, HuiRen, Analysis and Prospect of Control System for Stage Lighting, 2010 3rd International Congress on Image and Signal Processing (CISP2010).
- [3] Wang Yizhu, Malimin, Wandonghua. History of Stage lighting control technology. Entertainment Technology, 2005, (6):53-56.
- [4] Youjin Kim, Insu Kim, Tae-gyu Kang, Seong-hee Park. Analysis of IP-based Control Networksfor LED Lighting Fixture Communication. New Trends in Information Science and Service Science (NISS), 2010 4th International Conference 2010
- [5] Y. F. Wang, a. Mitiche, and j. K. Aggarwal. Computation of Surface Orientation and Structure of Objects Using Grid Coding. IEEE Transactions on Pattern Analysis and Machine intelligence, vol. 9, No. 1, January 1987.
- [6] A DMX 512 Handbook, Elation Professional, 2008.
- [7] Specification of VL2500 Spot, Philips Varilite 2011.
- [8] Dashlight Virtual Controller Manual, 2010.

- [9] Bruno Siciliano, Oussama Khatib, Handbook of robotics, Springer VL2500 Spot, 2008.

IX. Autores

Dr. Jacobo Sandoval Gutiérrez, actualmente es profesor en el departamento de Electrónica en la UAM-A, obtuvo su Doctorado y Maestría en Tecnología Avanzada, Ingeniería en Robótica Industrial todos ellos en el Instituto Politécnico Nacional en la Ciudad de México, México.

Ing. Jorge López, actualmente es profesor en el departamento de Electrónica en la UAM-A, obtuvo su grado en Ingeniero en Electrónica en el área de Control e Instrumentación en la UAM-A, Cd. de México, México.

Ing. José Francisco Cosme Aceves, actualmente es profesor en el departamento de Electrónica en la UAM-A, obtuvo su grado en Ingeniero en Electrónica en el área de Control e Instrumentación en la UAM-A, Cd. de México, México.

Dr. Andrés Ferreyra Ramírez, actualmente es profesor en el departamento de Electrónica en la UAM-A. Obtuvo su Doctorado en Control Automático en el CINVESTAV, Maestría en Ciencias en Ingeniería Biomédica, UAM-I y la Ingeniería en Ingeniería Mecánica Eléctrica (Electrónica y Comunicaciones) en la ENEP Iztacala. México.

Modelado y Análisis de los Circuitos Serie RC y LC Descritos por una Ecuación Diferencial de Orden Fraccionario

Manuel Guía Calderón, **José Francisco Gómez Aguilar, ***J. Juan Rosales García, *J. Jesús Bernal Alvarado.**

* Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato, Carretera Salamanca-Valle de Santiago km 3.5+1.8 km. Comunidad de Palo Blanco. Tel. (464) 647 9940, ext. 2464. guia@ugto.mx

** Departamento de Ingeniería Física, División de Ciencias e Ingenierías del Campus León, Universidad de Guanajuato, 37150, León Gto. México. Tel. +52 (477) 788-5100, ext. 8449. jfga@fisica.ugto.mx

*** Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato, Carretera Salamanca-Valle de Santiago km 3.5+1.8 km. Comunidad de Palo Blanco. Tel. (464) 647 9940, ext. 2369. rosales@ugto.mx

**** Departamento de Ingeniería Física, División de Ciencias e Ingenierías del Campus León, Universidad de Guanajuato, 37150, León Gto., México. Tel. +52 (477) 788-5100, ext. 8455. bernal@fisica.ugto.mx

Resumen

En el presente trabajo se usa la derivada temporal de tipo Caputo para analizar el comportamiento de los circuitos serie RC y LC modelados mediante una ecuación diferencial fraccionaria. El orden de la derivada considerada es $0 < \gamma \leq 1$. Para ser consistentes con la ecuación física se introduce el nuevo parámetro σ el cual caracteriza la existencia de componentes temporales fraccionarias en el sistema. Se encuentra una relación entre el orden de la derivada fraccionaria γ y el nuevo parámetro σ . Las soluciones están dadas en términos de la función de Mittag-Leffler.

Palabras Clave: Cálculo Fraccionario, Funciones de Mittag-Leffler, Circuito Serie RC, Circuito Serie LC, Ecuación Diferencial Fraccionaria.

I. Introducción

El cálculo fraccionario (CF) es tan antiguo como el cálculo entero convencional, más de tres siglos, pero la popularidad dentro de las comunidades científicas y de ingeniería solo tiene unas cuantas décadas. Lo maravilloso de esto es que, tanto la derivada como la integral fraccionaria, no tienen propiedades locales. Así, éste considera la historia y los efectos no locales distribuidos de cualquier sistema físico. Otra peculiaridad del cálculo fraccionario es la inclusión de nuevos grados de libertad al sistema, lo cual incrementa la información que puede ser obtenida de la naturaleza. Las aplicaciones del CF a problemas de las ciencias e ingeniería han crecido considerablemente en los últimos años [1]-[2].

El reciente interés por el CF y, en particular, por las ecuaciones diferenciales de orden fraccionario se ha estimulado por la aplicación en las áreas de la física, química, biología e ingeniería [3-11], entre otras. Sin embargo, la derivación de tales ecuaciones partiendo de algunas leyes fundamentales no es un problema sencillo. El operador fraccionario refleja procesos disipativos intrínsecos suficientemente complicados en la naturaleza. Su relación teórica con el CF no está todavía completamente comprobada. Además, es interesante analizar sistemas físicos simples y tratar de entender su funcionamiento a partir del modelo diferencial fraccionario.

En este trabajo se analiza, en el dominio del tiempo, el comportamiento de dos sistemas físicos simples, modelados, ambos, mediante una

ecuación diferencial de orden fraccionario: el circuito serie RC y el circuito serie LC, respectivamente. Se ha reemplazado el operador diferencial ordinario (entero) por su correspondiente operador diferencial fraccionario.

En este trabajo, se identificara al operador diferencial fraccionario con la definición de la diferencial fraccionaria de Caputo (DFC) [12] esto es

$${}_0^C D_t^\gamma f(t) = \frac{1}{\Gamma(n-\gamma)} \int_0^t \frac{f^{(n)}(\tau)}{(t-\tau)^{\gamma+1-n}} d\tau, \quad (1)$$

donde $n-1 < \gamma \leq n$, $n \in \mathbb{N}$

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad \Re(x) > 0 \quad (2)$$

es la función gama de Euler y

$$f^{(n)}(\tau) = \frac{d^n(\tau)}{d\tau^n}. \quad (3)$$

Considerando el caso cuando $n=1$, esto es, en el integrando está solamente la primera derivada. Una de las principales ventajas de la DFC es que las condiciones iniciales para la ecuación diferencial fraccionaria son dadas en la misma forma que para el cálculo diferencial ordinario; con su conocida interpretación física.

En virtud de que la idea es encontrar la solución de las ecuaciones diferenciales fraccionarias utilizando el método de la transformada de Laplace, es conveniente definir la transformada de Laplace de la DFC, así como una función útil para resolver la transformada inversa de Laplace. La transformada de Laplace de la DFC está dada por

$$\begin{aligned} L\left[\frac{d^\alpha f(t)}{dt^\alpha}\right] &= \int_0^\infty \left[\frac{d^\alpha f(t)}{dt^\alpha}\right] e^{-st} dt \\ &= s^\alpha F(s) - \sum_{k=1}^n s^{\alpha-k} f^{(k-1)}(0^+). \end{aligned} \quad (4)$$

$n-1 < \alpha < n, \quad n \in \mathbb{N},$

donde $F(s)=L[f(t)]$.

La transformada inversa de Laplace requiere de la introducción de una función especial. La función de uniparamétrica de Mittag-Leffler que será utilizada en este trabajo, se define mediante la siguiente fórmula

$$E_\gamma(t) = \sum_{m=0}^\infty \frac{t^{\gamma m}}{\Gamma(\gamma m + 1)}, \quad (\Re \gamma > 0). \quad (5)$$

II. Desarrollo

Haciendo uso de la ley de voltajes de Kirchoff y del circuito de la Fig. 1, se toman las siguientes consideraciones:

$$V_e(t) = \frac{e}{\sigma_e^{(1-\gamma)\beta}} \frac{d^{\beta\gamma} q(t)}{dt^{\beta\gamma}}, \quad (6)$$

donde $\beta = \{1, 2\}$ es un parámetro que determina si el elemento e es un resistor o un inductor $0 < \gamma \leq 1$, el producto $\beta\gamma$ determina el orden fraccionario de la ecuación diferencial, $e = \{R, L\}$ define el elemento de que se trata o las características de e , σ es un parámetro que determina las componentes fractales del elemento y sus dimensiones son [seg].

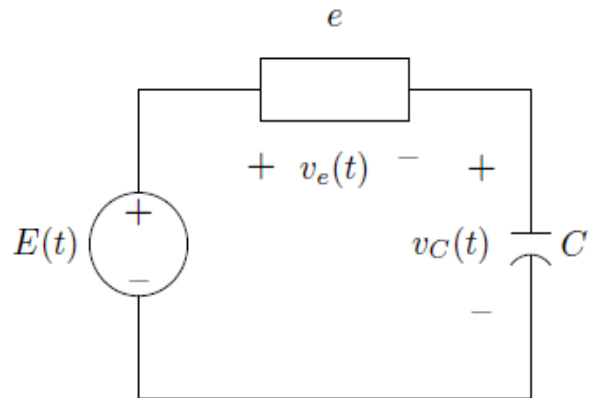


Fig. 1. Circuito general serie de un elemento fractal y un capacitor.

Circuito serie RC fraccionario

Haciendo en (6) $e=R$, $\beta=1$, $\sigma_e = \sigma_R$ se obtiene la ecuación para un circuito serie resistor fractal capacitor como sigue

$$\frac{R}{\sigma_R^{1-\gamma}} \frac{d^\gamma q(t)}{dt^\gamma} + \frac{q(t)}{C} = E(t), \quad (0 < \gamma \leq 1) \quad (7)$$

La introducción del parámetro σ_R permite una consistencia en la dimensionalidad cuando $[\sigma_R]=\text{seg.}$, esto es

$$\left[\frac{1}{\sigma_R^{1-\gamma}} \frac{d^\gamma}{dt^\gamma} \right] = \frac{1}{\text{seg}}, \quad (0 < \gamma \leq 1) \quad (8)$$

de esta forma, con $\gamma=1$ la ecuación (8) se convierte en una diferencial ordinaria de primer orden respecto al tiempo, t . Esto solo se cumple, como se dijo anteriormente, si σ_R tiene dimensiones de segundos.

Usando la expresión (7) la ecuación diferencial fraccionaria para un circuito serie de un capacitor y un resistor fractal tiene la forma

$$\frac{d^\gamma q(t)}{dt^\gamma} + \frac{q(t)}{\tau_\gamma} = \frac{C}{\tau_\gamma} E(t), \quad (9)$$

donde

$$\tau_\gamma = \tau^\gamma \left[\frac{\tau}{\sigma_R} \right]^{1-\gamma} \text{ y } \tau = RC, \quad (0 < \gamma \leq 1), \quad (10)$$

A la constante τ_γ también puede llamársele constante de tiempo fractal, debido a su dimensionalidad $[\text{seg}]^\gamma$. Cuando $\gamma = 1$, de (10) se recupera la constante de tiempo ordinaria, esto es $\tau_1 = \tau = RC$.

Suponiendo condiciones iniciales cero, esto es $q(0) = 0$ y $E(t) = V_0 \mathbf{u}(t)$, donde V_0 es una constante, t , y $\mathbf{u}(t)$ es la función escalón unitario, la transformada de Laplace aplicada a (9) da como resultado,

$$Q(s) = \frac{CV_0}{\tau_\gamma s \left(s^\gamma + \frac{1}{\tau_\gamma} \right)}. \quad (11)$$

La transformada inversa de Laplace (11) está dada por

$$q(t) = CV_0 \left\{ 1 - E_\gamma \left[- \left(\frac{\sigma_R}{\tau} \right)^{1-\gamma} \left(\frac{t}{\tau} \right)^\gamma \right] \right\}, \quad (12)$$

El parámetro γ , que representa el orden de la derivada temporal fraccionaria en (12), puede relacionarse con el parámetro σ_R , el cual

caracteriza la presencia de componentes fraccionarias (fluctuaciones) en el sistema. En este caso la relación está dada por la expresión

$$\gamma = \frac{\sigma_R}{\tau}, \quad (13)$$

entonces, la magnitud

$$\delta = 1 - \gamma. \quad (14)$$

caracteriza la presencia de estructuras fractales en el sistema. Se puede apreciar esto de la siguiente manera: Si $\gamma=1$, entonces, de (13) se tiene $\sigma_R=\tau$ y por consiguiente $\delta=0$ en (14), lo cual quiere decir que en el sistema no existen estructuras fractales, es decir, es un circuito RC ordinario. Sin embargo, en el rango $0 < \gamma < 1$ o lo que es lo mismo $0 < \sigma < \tau$ la magnitud δ crece y tiende a la unidad, ya que en el sistema aparecen cada vez más estructuras fractales. Sustituyendo (13) en (12) se tiene

$$\hat{q}(t) = 1 - E_\gamma \left[-\gamma^{1-\gamma} \hat{t}^\gamma \right], \quad (15)$$

donde $\hat{q} = \frac{q}{CV_0}$ y $\hat{t} = \frac{t}{RC}$, son variables adimensionales. De (15), se tiene el voltaje en el capacitor

$$\hat{v}_C(t) = 1 - E_\gamma \left[-\gamma^{1-\gamma} \hat{t}^\gamma \right]. \quad (16)$$

donde $\hat{v} = \frac{v}{V_0}$.

La corriente se puede obtener derivando (15) con respecto al tiempo, esto es

$$\hat{i}(t) = \frac{d}{d\hat{t}} \left\{ 1 - E_\gamma \left[-\gamma^{1-\gamma} \hat{t}^\gamma \right] \right\}. \quad (17)$$

Las ecuaciones (15) y (16) son idénticas, aunque la primera tiene unidades de Coulomb/Coulomb y la segunda tiene unidades de Volts/Volts, de forma que la gráfica es idéntica para ambas, la ecuación (17) tiene dimensiones de Ampere/Ampere.

Circuito serie LC fraccionario

Haciendo en (6) $e = L$, $\beta = 2$, $\sigma_e = \sigma_L$ se obtiene la ecuación para un circuito serie inductor fractal capacitor como sigue

$$\frac{L}{\sigma_L^{2(1-\gamma)}} \frac{d^{2\gamma} q(t)}{dt^{2\gamma}} + \frac{q(t)}{C} = E(t), \quad (0 < \gamma \leq 1) \quad (18)$$

La introducción del parámetro σ_L permite una consistencia en la dimensionalidad cuando $[\sigma_L] = \text{seg.}$, esto es

$$\left[\frac{1}{\sigma_L^{1-\gamma}} \frac{d^\gamma}{dt^\gamma} \right] = \frac{1}{\text{seg}}, \quad (0 < \gamma \leq 1) \quad (19)$$

de esta forma con $\gamma = 1$ la ecuación (19) se convierte en una diferencial ordinaria de primer orden respecto al tiempo, t . Esto solo se cumple, como se dijo anteriormente, si σ_L tiene dimensiones de segundos.

Usando la expresión (18), la ecuación diferencial fraccionaria para un circuito serie de un capacitor y un inductor fractal tiene la forma

$$\frac{d^{2\gamma} q(t)}{dt^{2\gamma}} + \frac{q(t)}{\tau_{L\gamma}} = \frac{C}{\tau_{L\gamma}} E(t), \quad (20)$$

donde

$$\tau_{L\gamma} = \tau_L^\gamma \left[\frac{\tau_L}{\sigma_L} \right]^{1-\gamma} \text{ y } \tau_L = \sqrt{LC}, \quad (0 < \gamma \leq 1), \quad (21)$$

A la constante $\tau_{L\gamma}$ también puede llamársele constante de tiempo fractal, debido a su dimensionalidad $[\text{seg}]^\gamma$. Cuando $\gamma = 1$, de (21) se recupera la constante de tiempo ordinaria, esto es $\tau_{L1} = \tau = \sqrt{LC}$.

Suponiendo las condiciones iniciales $q(0) = q_0$, $\dot{q}(0) = 0$ y una fuente de alimentación $E(t) = 0$, la transformada de Laplace aplicada a (20) da como resultado

$$Q_h(s) = \frac{q_0}{\tau_{L\gamma} \left(s^{2\gamma} + \frac{1}{\tau_{L\gamma}} \right)} \quad (22)$$

La transformada inversa de Laplace de (22) está dada por

$$q_h(t) = q_0 E_{2\gamma} \left[- \left(\frac{\sigma_L}{\tau_L} \right)^{2(1-\gamma)} \left(\frac{t}{\tau_L} \right)^{2\gamma} \right], \quad (23)$$

donde $E_{2\gamma}$ es la función de Mittag-Leffler de parámetro 2γ .

El parámetro γ , como antes, que representa el orden de la derivada temporal fraccionaria en (20), puede relacionarse con el parámetro σ_L , el cual caracteriza la presencia de componentes fraccionarias (fluctuaciones) en el sistema. En este caso la relación está dada por la expresión

$$\gamma = \frac{\sigma_L}{\tau_L}, \quad (24)$$

entonces, la magnitud

$$\delta = 1 - \gamma. \quad (25)$$

al igual que para el circuito RC, caracteriza la presencia de estructuras fractales en el sistema. Se puede apreciar esto de la siguiente manera: si $\gamma = 1$, entonces, de (24) se tiene $\sigma_L^2 = LC$ y, por consiguiente, $\delta = 0$ en (25), lo cual quiere decir que en el sistema no existen estructuras fractales, es decir, es un circuito LC ordinario. Sin embargo, en el rango $0 < \gamma < 1$, o lo que es lo mismo, $0 < \sigma < \sqrt{LC}$ (frecuencia natural no amortiguada), la magnitud δ crece y tiende a la unidad, ya que en el sistema aparecen cada vez más estructuras fractales. Sustituyendo (24) en (23) se tiene

$$\hat{q}(t) = E_{2\gamma} \left[-\gamma^{2(1-\gamma)} \hat{t}^{2\gamma} \right], \quad (26)$$

donde $\hat{q} = \frac{q}{q_0}$ y $\hat{t} = \frac{t}{\sqrt{LC}}$, son variables adimensionales. De (26), se tiene el voltaje en el capacitor

$$\hat{v}_C(t) = E_{2\gamma} \left[-\gamma^{2(1-\gamma)} \hat{t}^{2\gamma} \right], \quad (27)$$

donde $\hat{v} = \frac{v}{Cq_0}$.

La corriente se puede obtener derivando (26) con respecto al tiempo, esto es

$$\hat{i}(t) = \frac{d}{dt} \left\{ 1 - E_{2\gamma} \left[-\gamma^{2(1-\gamma)} \hat{t}^{2\gamma} \right] \right\}, \quad (28)$$

Las ecuaciones (26) y (27) son idénticas, aunque la primera tiene unidades de Coulomb/Coulomb y la segunda tiene unidades de Volts/Volts, de forma que la gráfica es idéntica para ambas, la ecuación (28) tiene dimensiones de Ampere/Ampere.

La solución general de (20) puede encontrarse sumando a la solución encontrada (23), solución homogénea, la solución particular, esto es

$$q(t) = q_h(t) + \int_0^t E(\tau) E_{2\gamma} \left[-\gamma^{2(1-\gamma)} \left(\frac{t-\tau}{\sqrt{LC}} \right)^{2\gamma} \right] d\tau \quad (29)$$

III. Resultados

Además de los resultados obtenidos hasta el momento, dados por las expresiones para la carga en el capacitor para ambos circuitos, la tensión en el capacitor, también para los circuitos analizados y la corriente en ambos casos, es conveniente dar algunos resultados del comportamiento del voltaje en el capacitor, cuando se varía el orden fraccionario de la ecuación diferencial γ .

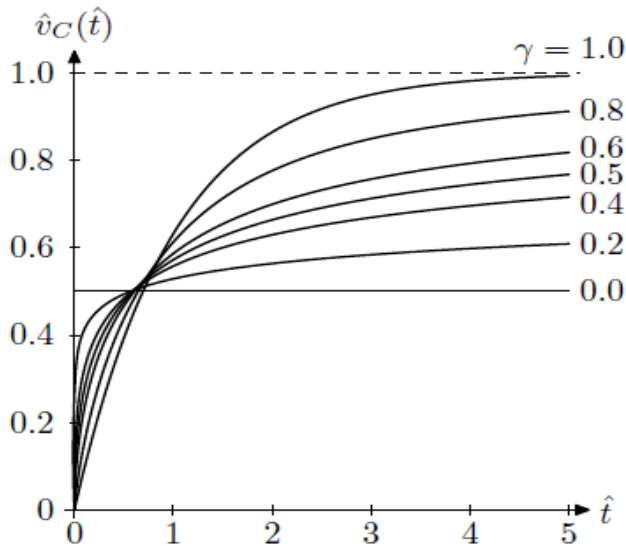


Fig. 2. Tensión en el capacitor en el circuito RC fraccionario para diferentes valores de γ .

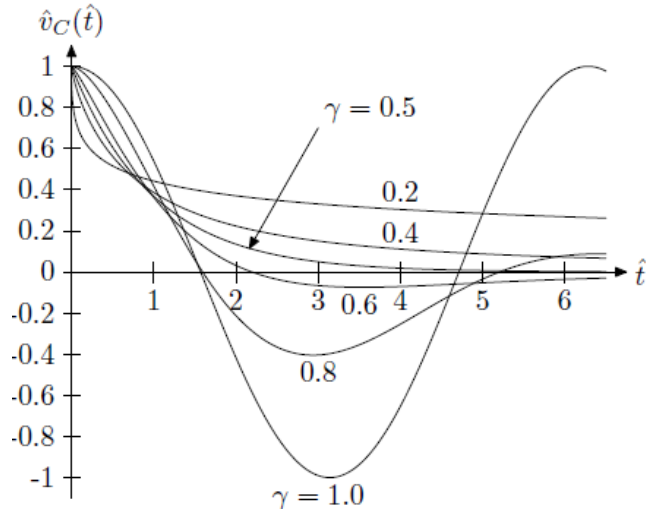


Fig. 3. Tensión en el capacitor en el circuito LC fraccionario para diferentes valores de γ .

De la Fig. 2 se puede observar que a medida que γ crece desde cero hasta 1, el elemento fractal se comporta como un capacitor de capacitancia C , para $\gamma = 0$ hasta un resistor de resistencia R para $\gamma = 1$; valores intermedios de γ determinan un comportamiento del elemento entre un capacitor y un resistor. En la Fig. 3 se observa, por otro lado, un comportamiento del elemento entre un inductor puro, $\gamma = 1$ y un capacitor puro $\gamma = 0$, con un comportamiento entre resistor e inductor para valores intermedios de γ .

IV. Discusión

Se ha encontrado que, para todo $\gamma \in (0, 1]$, la carga almacenada en el circuito RC es directamente proporcional al voltaje en el capacitor, de esto se sigue que el comportamiento de la carga en el capacitor dará curvas idénticas a las mostradas en la Fig. 2.

De la Fig. 2 se puede observar que, a medida que el orden de la ecuación diferencial fraccionaria aumenta de 0 a 1.0, el tiempo de retardo, \hat{t}_d , aumenta, lo mismo que el tiempo de elevación, \hat{t}_r , también aumenta el tiempo de asentamiento, \hat{t}_s . En cuanto al valor de asentamiento, considerado como el 95% del valor final aparentemente disminuye, pero eso solo es obvio para $\gamma = 0$, ya que se requiere conocer el valor final del voltaje en el capacitor para los diferentes valores de γ .

El comportamiento de la corriente se puede inducir del comportamiento del voltaje en el capacitor, para $\gamma = 0$, se deberá tener un impulso de corriente, ya que se trata de dos capacitores ideales, en tanto que para $\gamma = 1$ se tendrá un valor de la corriente igual al cociente entre el voltaje de alimentación y la resistencia, para $t = 0$, e irá decreciendo con el tiempo hasta llegar a cero; un circuito RC de orden entero.

En lo que se refiere al circuito fraccionario LC, se observa, de la Fig. 3 una respuesta sinusoidal pura, para $\gamma = 1$ y una respuesta sinusoidal amortiguada conforme γ decrece, además, se nota que existen pérdidas en el circuito serie LC, esto se puede ver debido a que existe un amortiguamiento en las curvas de voltaje; típico de un circuito RLC con pérdidas. En esta Fig. también se grafica el voltaje en el capacitor.

En general, se ha estudiado el comportamiento de los circuitos serie RC y LC de orden fraccionario. Se ha introducido, en ambos casos, un parámetro, σ , que permite tener una relación física consistente en cuanto a la dimensionalidad y que representa la fractalidad del tiempo. Este parámetro tiene una estrecha relación con el orden de la derivada fraccionaria. Además, se encontró una relación entre el parámetro γ , que representa el orden de la derivada fraccionaria, con el parámetro σ , el cual caracteriza la presencia de componentes fractales temporales (fluctuaciones temporales) en el sistema.

V. Conclusiones

El estudio de circuitos eléctricos de orden fraccionario y sus implicaciones tecnológicas es, definitivamente, un área emergente de la investigación interdisciplinaria. Cabe hacer notar que el conocimiento del cálculo fraccionario y sus diversas técnicas está muy difundido entre los bioquímicos, la teoría de control y el procesamiento de señales, mientras que en la teoría de circuitos y aplicaciones de circuitos los avances son pocos. El modelado de elementos de circuito eléctrico, mediante diferenciales de orden fraccionario, y su inclusión en el análisis

de circuitos eléctricos permitirá comprender mejor el comportamiento de estos elementos los cuales, hasta ahora, se han venido modelando mediante ecuaciones diferenciales de orden entero. Es cierto que los resultados son aceptables pero existen varios fenómenos que no es posible comprender del todo bien. Una vez comprendido el funcionamiento de los elementos de orden fraccionario, lo que sigue es el diseño, empleando nuevos materiales y geometrías, su simulación y la fabricación de tales elementos con un orden fraccionario específico. Todo esto permitirá diseñar circuitos eléctricos con características dinámicas más precisas y más adecuadas a la aplicación particular de que se trate.

Por otro lado, es posible tener un mejor estudio sobre el comportamiento transitorio y abre las puertas a la creación de mecanismos que permitan la predicción de fallas en equipo eléctrico y al estudio de la forma en que evolucionan las probabilidades de falla en equipo eléctrico.

Esperamos que esta forma de tratar con ecuaciones diferenciales de orden fraccionario ayude a una mejor comprensión del comportamiento de los sistemas de orden fraccionario.

VI. Referencias

- [1] D. Cafagna, Past and present—fractional calculus: a mathematical tool from the past for present engineers, IEEE Industrial Electronics Magazine, vol. 1, no. 2, pp. 35–40, 2007.
- [2] J. A. T. Machado, A probabilistic interpretation of the fractional-order differentiation. Fractional Calculus and applied Analysis, vol. 6, no. 1, pp. 73–80, 2003.
- [3] Bisquert J. (2003). *Phys. Rev. Lett.* 91, 010602.

- [4] Boldyrev S., Gwinn C.R., (2003). *Phys. Rev. Lett.* 91(131101).
- [5] Chechkin A.V., Gonchar V.Y., Szydlowsky M. (2002). *Phys. Plasma* 9(78).
- [6] Lacasta A.M., Sancho J.M., Romero A.H., Sokolov I.M., Lindenberg K., (2004). *Phys. Rev. E* 70(051104).
- [7] Mainardi F. (1997), in *Fractals and Fractional Calculus in Continuum Mechanics*, Edited by A. Carpinteri, F. Mainardi (Springer-Verlag, New York), 291-348.
- [8] Serebnyanska M., Hanyga A., (2000). *J. Math. Phys.* 41(2135).
- [9] Stanislavsky A.A., (2004). *JETP* 98(705).
- [10] Tarasov V.E., Zaslavsky G.M., (2005). *Physica A* 354(249).
- [11] Podlubny I., (1994). The Laplace transform method for linear differential equations of the fractional order. Tech. Rep., Slovak Academy of Sciences, Institute of Experimental Physics.
- [12] I. Podlubny, *Fractional Differential equations*. New York: Academic, 1999.

VII. Autores

M. I. Manuel Guía Calderón obtuvo su título de Maestría en Ingeniería Eléctrica en la Facultad de Ingeniería Mecánica Eléctrica y Electrónica de la Universidad de Guanajuato con especialidad en Instrumentación y Sistemas Digitales. Actualmente trabaja en el Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato. Sus intereses son efectos electromagnéticos en sistemas biológicos.

M. I. José Francisco Gómez Aguilar obtuvo su título de Maestría en Ingeniería Eléctrica en la Facultad de Ingeniería Mecánica Eléctrica y Electrónica de la Universidad de Guanajuato con

especialidad en Instrumentación y Sistemas Digitales. Actualmente estudia su Doctorado en el Departamento de Ingeniería Física, División de Ciencias e Ingenierías del Campus León, Universidad de Guanajuato. Sus intereses son el bioelectromagnetismo, el modelado matemático de sistemas biológicos y las aplicaciones del cálculo fraccionario.

Dr. J. Juan Rosales García obtuvo su título de Doctor en Ciencias en la UAM Iztapalapa. Actualmente trabaja en el Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato. Sus intereses son efectos electromagnéticos en sistemas biológicos y aplicaciones del cálculo fraccionario.

Dr. J. Jesús Bernal Alvarado obtuvo su título de Doctor en Ciencias en la Universidad de Guanajuato. Actualmente trabaja en el Departamento de Ingeniería Física, División de Ciencias e Ingenierías del Campus León, Universidad de Guanajuato. Sus intereses son la espectroscopia de impedancia eléctrica, el modelado matemático de sistemas biológicos y bioelectromagnetismo.

Automatización de una Prueba Triaxial Rápida “LG-TRIAxIAL”

***Rubén Herrera Galicia, **Jonathan Rasgado Jiménez, Boris Solís Zavaleta.**

Instrumentación y Automatización, Ingeniería Eléctrica y Electrónica, Instituto Tecnológico de Tuxtla Gutiérrez, Carretera Panamericana Km. 1080, Col. Terán, Tuxtla Gutiérrez, Chiapas, 29070, México. TEL. 6150380, *h_galicia24@hotmail.com, **jonny-873112@hotmail.com.

Resumen

El trabajo aquí descrito está enfocado a diseñar un sistema capaz de adquirir, amplificar y filtrar señales provenientes de un sensor de desplazamiento y un sensor de presión, adaptados a una cámara triaxial, para obtener señales con la menor interferencia y ruido posibles, estas garantizan un registro confiable de los esfuerzos a los que son sometidas las muestras de suelo, son transportadas por un medio de comunicación, para finalmente procesarlas en una computadora y visualizarlas en un monitor.

Palabras Clave: Cámara Triaxial, Celda de Carga, LVDT, Micro Controladores.

I. Introducción

Geotecnia es la ciencia que se encarga del estudio de las características físicas, químicas y dinámicas de los suelos. Objeto de estudio de esta ciencia son: la porosidad, color, textura, cohesión, compresión y el ángulo de fricción. Algunas características son reconocidas mediante inspección visual. Sin embargo, para conocer las características dinámicas es necesario usar herramientas especializadas: acelerómetro, centrifugadora y cámara triaxial [1-7].

Una cámara triaxial permite conocer el comportamiento de un suelo bajo la acción de fuerzas axiales. La prueba triaxial rápida permite estimar la cohesión del suelo “ c ” y el ángulo de fricción interna “ ϕ ”, siendo éstos los parámetros de resistencia del suelo ensayado. La finalidad de la prueba rápida es definir dichos parámetros (c , ϕ) a través de la medición de la relación esfuerzo-deformación. La información que entrega la prueba triaxial es la más representativa del esfuerzo cortante que sufre una masa de suelo al ser cargada [2, 3, 5].

La Fig. 1 presenta una imagen de una cámara triaxial típica. El equipo de ensayo triaxial consta de una cámara triaxial, una prensa de

compresión, un sistema de presión de confinamiento y elementos de medición.

La Cámara triaxial está formada por un cilindro, una base y una tapa. El cilindro está construido con un material transparente que soporta altas presiones internas, la base es metálica y posee tres conexiones al exterior con sus correspondientes válvulas. La tapa, también metálica, posee una válvula de purga de aire.

El sistema de presión de confinamiento consiste en un compresor de aire con escalonamiento de presiones, accionado eléctricamente. El sistema posee un tanque de reserva de aire comprimido y un tanque intermedio. En este último es donde la presión de aire es transmitida al agua que se usa para llenar la cámara triaxial. La presión del agua, llamada presión de confinamiento, proporciona dos de los tres esfuerzos aplicados a la muestra de suelo.

La prensa de compresión es un pistón neumático, ubicado en la parte inferior, que empuja a la cámara hacia arriba proporcionando un tercer esfuerzo axial a la muestra de suelo, llamado esfuerzo desviador. La prensa posee un

dispositivo que permite variar la velocidad con que se aplica la carga vertical sobre la probeta.

Los elementos de medición que se usan son mecánicos. Un manómetro mide la variación de la altura de la muestra. Esta deformación es igual al desplazamiento de la cámara de prueba provocado por el movimiento del pistón neumático y corresponde también a la reducción que sufre la muestra durante la ejecución del ensayo. Otro manómetro, acoplado a un anillo de acero, mide la fuerza vertical aplicada a la muestra de suelo. El anillo sufre una deformación comprendida entre 0 y 0.5 mm que es registrada por el manómetro. Esta deformación es proporcional al esfuerzo desviador.



Fig. 1 Cámara triaxial típica (Lab. Mecánica de Suelos, Facultad de Ingeniería Civil, Universidad Autónoma de Chiapas).

Antes de ejecutar la prueba triaxial se deben hacer mediciones de la muestra a la que se pretende someter al proceso esfuerzo deformación. Se miden la altura y los diámetros (inferior, medio y superior). El análisis comienza colocando una muestra cilíndrica de suelo dentro de una membrana de caucho, de altura aproximada

$H=8\text{cm}$ y diámetro aproximado $D=4\text{cm}$. La muestra, así impermeabilizada con el caucho, se introduce en el interior de la cámara triaxial y se le aplica una presión igual en todo sentido y dirección. Alcanzado ese estado de equilibrio se aumenta la presión axial (σ_1) sin modificar la presión de confinamiento (σ_3).

Las lecturas de los manómetros se toman en forma equidistante. La primera lectura se toma cuando la muestra sufre una deformación de 0.1mm, luego cada 0.25mm. El ensayo termina cuando la muestra de suelo sufre una falla, lo cual se identifica con dos lecturas similares en el manómetro del esfuerzo desviador. Concluida la prueba se llena una tabla de registro desechando las lecturas poco significativas. Una vez concentrada la información se determina el esfuerzo desviador de falla y se grafica la relación esfuerzo vs deformación.

La compresora hidráulica suministra agua a la cámara de prueba a una presión constante, igual a 0.5, 1.0, 1.5 ó 2.0 [kg/cm^2]. La prueba rápida se repite para cada una de las presiones de agua y se concentran los resultados de los esfuerzos desviadores de falla a diferentes presiones. Con los resultados se trazan los círculos de Mohr [1-7] y gráficamente se determina la envolvente que mejor se ajuste a ellos. A partir de la envolvente se determina el ángulo de fricción " ϕ " y el coeficiente de cohesión " c ". Estos parámetros sirven para calcular la capacidad de carga del suelo analizado. Conociendo estos valores un ingeniero civil determina las características que debe cumplir la cimentación, sobre la cual se construirá una edificación, para evitar hundimientos o colapsos.

Los problemas de la operación de una cámara triaxial mecánica son: el tiempo que ocupan las mediciones, el número de operarios requeridos, y el factor del error humano en las mediciones y los cálculos. Una prueba dura entre 15 y 30 min. Se hacen 3 ó 4 pruebas para la estimación de una envolvente. Este proceso lo hacen dos personas.

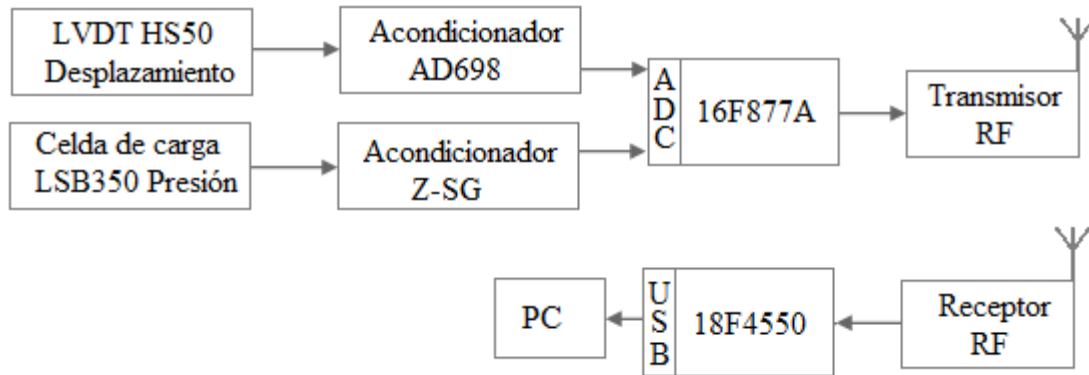


Fig. 2 Diagrama a bloques del sistema de automatización de una prueba triaxial rápida.

El uso frecuente de las cámaras triaxiales, provoca que existan diferentes modelos de estos equipos. Cada uno de ellos ofrece diferentes ventajas. Algunas cámaras son completamente automatizadas y poseen comunicación con la computadora [9]. Sin embargo, el modelo mecánico es el más popular debido a su bajo costo. Algunos modelos utilizan tres pistones neumáticos para proporcionar los esfuerzos axiales que se requieren durante la prueba. Estos equipos son ampliamente utilizados para el estudio de la porosidad de los suelos [6,7].

En México existen deficiencias relacionadas con cámaras triaxiales digitalizadas que den una respuesta eficaz a un precio competitivo, las existentes son de origen italiano [9] y tienen precios elevados. En el estado de Chiapas la industria de la construcción ocupa un lugar importante en la economía regional y es una de las principales fuentes de empleo en el estado. Un requisito para la autorización de una obra civil es un estudio previo de mecánica de suelos para determinar las características geotécnicas del terreno destinado a la construcción de la obra. Es aquí donde entra la aplicación de la cámara triaxial y de ahí resulta su importancia.

El prototipo que se reporta en este artículo es un sistema automatizado de medición y registro de una prueba rápida de compresión de una muestra de suelo en una cámara triaxial. La presión

vertical se mide mediante una celda de carga y la reducción de la probeta mediante un LVDT. El sistema facilita las tareas de medición - registro y disminuye el factor de error humano.

El problema que aquí se resuelve es implementar un sistema automatizado para medir en tiempo real las deformaciones de la muestra y las presiones a las que se somete. Y que dichos valores sean filtrados y transportados de manera confiable a través de un medio de comunicación hasta una computadora para procesarlos, graficarlos e imprimir un reporte de ellos.

II. Desarrollo

El sistema aquí propuesto se compone de las siguientes etapas; transducción, acoplamiento del sensor al módulo de conversión, muestreo y conversión, comunicación por Radio Frecuencia, comunicación a través del puerto USB, captura de datos y procesamiento en una PC. Un diagrama a bloques del sistema se presenta en la Fig. 2.

Acondicionamiento.- El sensor LVDT HS50 de VISHAY genera una señal de corriente alterna proporcional a la deformación de la muestra, provocada por la compresión. El voltaje de corriente alterna del LVDT es amplificado y acondicionado por el circuito AD698 de Analog

Device. La celda de carga LSB350 de Logicbus genera un voltaje proporcional a la carga ejercida por la compresora hidráulica sobre la muestra. El voltaje de la celda es entregado al acondicionador de celda de carga Z-SG de Logicbus. Ambos acondicionadores entregan como respuesta un voltaje en corriente continua.

Muestreo y conversión.- los voltajes que entregan los acondicionadores AD698 y Z-SG son muestreados por el módulo de conversión de un micro controlador, cuya función es convertirlos a señales digitales de 8 bits. Una vez que la señal ha sido convertida el micro controlador entrega el resultado al codificador CIP-8E y a un LCD, Fig.3. Además pone a nivel bajo al bit de habilitación de transmisión, para que el codificador CIP-8E active la transmisión.

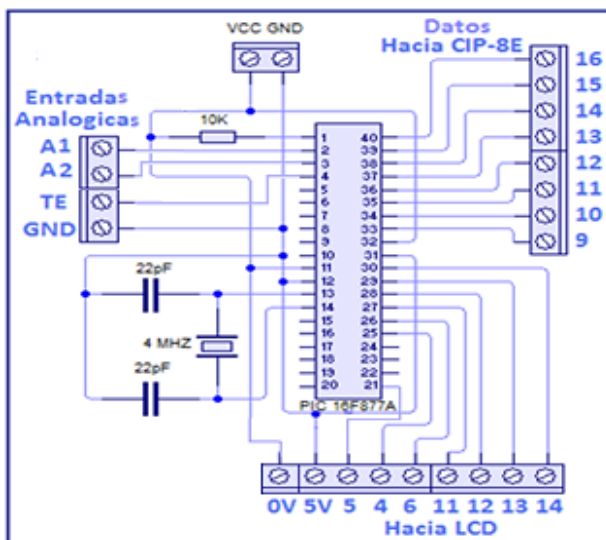


Fig. 3 Micro controlador principal.

Comunicación inalámbrica por RF.- las señales de control son generadas por el micro controlador. Si el pin TE habilitación de transmisión del codificador CIP-8E está a nivel bajo, se inicia la transmisión, Fig.4. Se termina la transmisión actual, incluso si TE retorna a nivel alto antes de concluir el envío del paquete. Mientras TE esté a nivel alto el codificador se encuentra en estado de bajo consumo. En el proceso de una transmisión el codificador lee el

dato y la dirección, y con ellos forma paquetes de siete bytes. Un paquete se compone de un byte de preámbulo, un byte de sincronización, dos veces el dato, dos veces la dirección, y un byte de checksum. Los bits de cada paquete son entregados en forma serial a través del pin de salida de datos al TWS-434, el cual se encarga de su transmisión. Los valores de la dirección y del dato se actualizan durante la formación de cada nuevo paquete. El CIP-8E inserta un retardo de 10 mseg entre cada transmisión. Esto con la finalidad de otorgar un espacio de tiempo al decodificador para que alcance a recibir, decodificar, y comprobar la integridad del paquete.

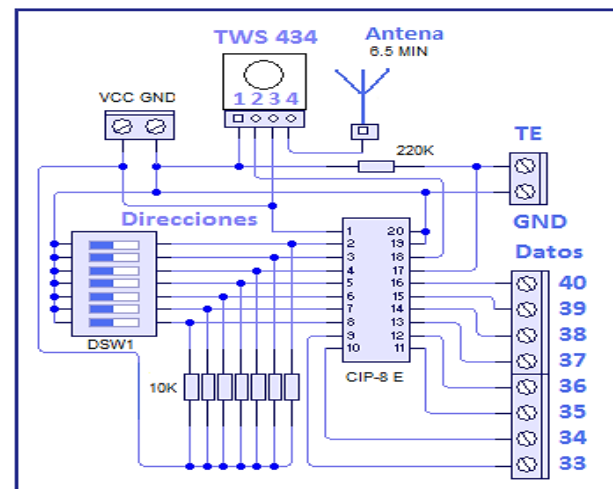


Fig. 4 Transmisor TWS-434.

Durante el desarrollo del presente proyecto se implemento un protocolo de comunicaciones en los micro controladores 16F877A y 18F4550. Esto con el objeto de evitar confusiones durante la recepción de los valores de las dos variables a medir. El protocolo consiste en enviar una clave antes del valor de la medición, con un retardo de 50 milisegundos entre cada número. El primer código consiste de tres números decimales (2, 5, 20). Después de estos números se envía la medición de la primera variable. El segundo código es (0, 7, 30), y después se envía la medición de la segunda variable. Durante la recepción el micro controlador identifica las claves y almacena las mediciones en distintas

variables para después enviarlas a través del puerto USB.

La función del receptor RWS-434, Fig. 5, es recibir los paquetes a través de la antena y transferirlos al decodificador CIP-8D, el cual se encarga de verificar la existencia de errores. El funcionamiento del decodificador consiste en entrar en un bucle de tiempo de espera al byte de sincronización. Si ningún byte válido de sincronización se recibe en 65.5mseg y el decodificador está configurado para funcionar en modo momentáneo, entonces un temporizador interno de 16 bits se desborda provocando una salida del bucle de espera para poner los pines de salida a cero. Si un byte de sincronización válido es recibido, el temporizador es desactivado y se inicia el proceso de verificación de la integridad del paquete. Una vez que el paquete ha sido considerado como libre de errores, el dato descifrado es colocado en los pines de salida y el decodificador regresa al bucle de espera.

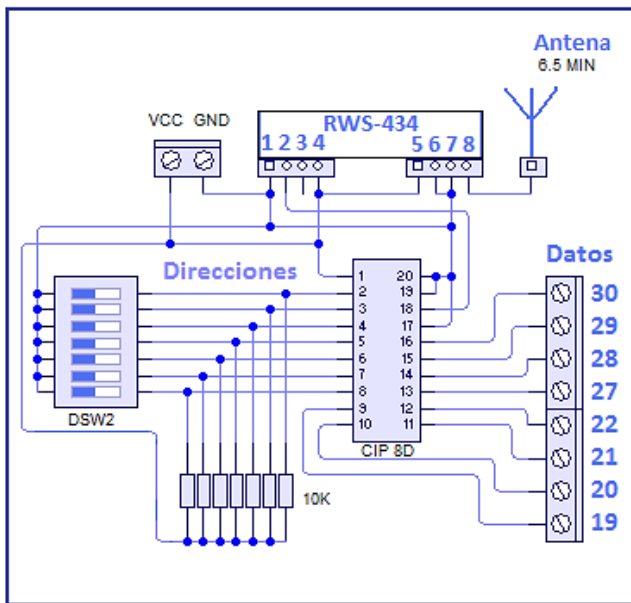


Fig.5 Receptor RWS-434.

Comunicación de datos vía USB.- Cuando la PC y el módulo de recepción se conectan, éste indica su detección por la misma a través de diodos LED. Los datos entregados por el decodificador son leídos por un micro controlador 18F4550, el

cual se encarga de mostrarlos con ayuda de un LCD. Y después los envía a través del puerto USB hacia la computadora, Fig. 6.

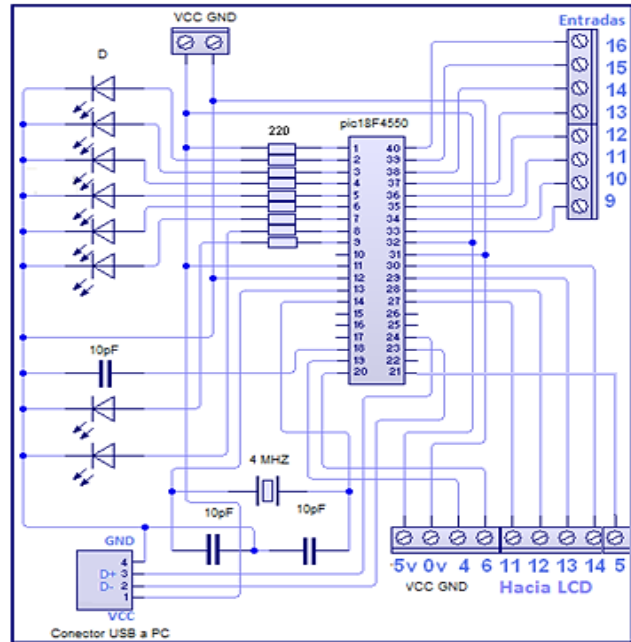


Fig. 6 Comunicación USB con el 18F4550.

Para la comunicación USB se programó al micro controlador 18F4550 en lenguaje CCS basado en el compilador PIC C. Entre el micro controlador y la PC se implemento una comunicación bidireccional serie [8], *Communications Devices Class*. El sistema consiste en un puerto COM serie estándar emulado vía USB 2.0. Se hacen ajustes de configuración en el Firmware del micro controlador, driver.inf, para que sea reconocido por Windows como dispositivo serie estándar y lo registre como puerto COM Virtual. En Windows se usa el mchpcdc.inf como intermediario entre el micro controlador y la PC, para definir los VID&PID que se utilizaran y que el micro controlador sea reconocido por Windows.

La interfaz USB implementada fue instalada en un sistema Windows 7 a 64 bits. La aplicación permite la asignación de una dirección de puerto COM a un puerto USB. Una vez reconocido el dispositivo USB el sistema lo indica con un diodo LED rojo que significa dispositivo conectado en

espera de ser enumerado, Fig. 7. Después de ser enumerado se enciende el diodo LED verde. Esto indica: dispositivo listo para la comunicación. En el administrador de dispositivos aparece el micro controlador reconocido y enumerado como un puerto COM, debido a la clase USB CDC.



Fig. 7 Diodos LED visualizadores.

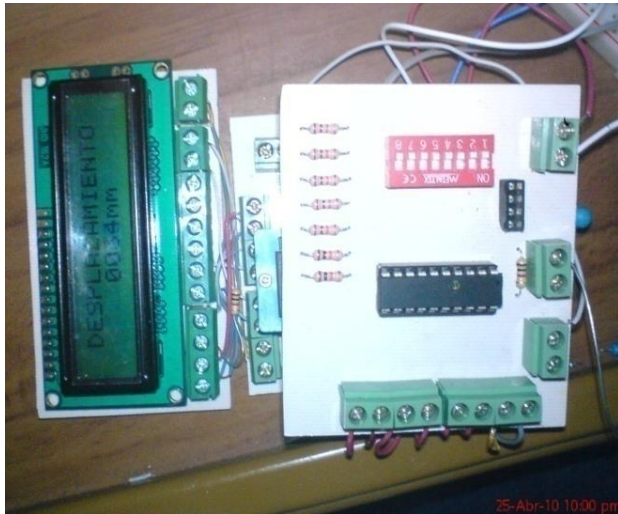


Fig. 8 Conversión A/D y transmisor RF.

Las placas de los módulos de conversión y transmisión se presentan en la Fig. 8. En la imagen se puede identificar al LCD, el cual se usa para corroborar visualmente que el dato transmitido es el mismo que el dato recibido en la PC. La implementación de los módulos de recepción y comunicación USB se presentan en la Fig. 9. En la imagen se puede identificar a un segundo LCD. Éste también sirve para corroborar

de manera visual que la comunicación esté libre de errores.

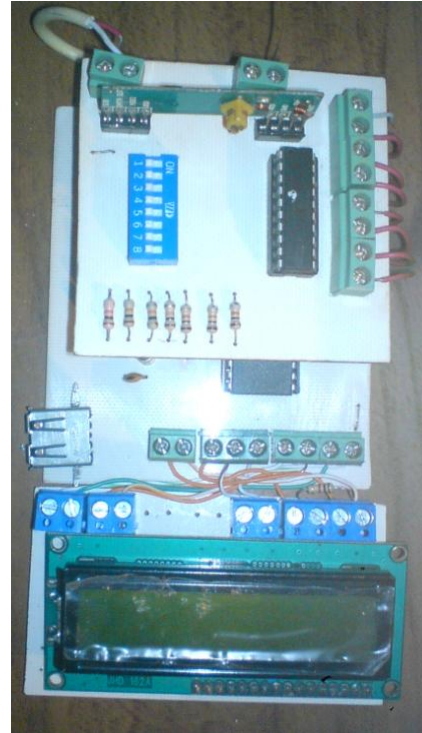


Fig. 9 Receptor RF e interfaz USB.

III. Resultados

Se implementó un programa en Visual Basic 2008 de nombre LG TRIAXIAL. Este programa ayuda a hacer pruebas triaxiales rápidas. Antes de iniciar la ejecución del programa es necesario que los sistemas transmisor y receptor estén funcionando y conectados. LG TRIAXIAL está estructurado por pestañas que representan a los pasos a seguir para efectuar un ensayo. Dos de las pestañas están relacionadas con la introducción de los datos necesarios para efectuar una prueba.

Al abrir el programa LG-TRIAxIAL, Fig.10, como primera se encuentra la pestaña de inicio, en donde el usuario direcciona al puerto USB a través de la selección de un puerto COM. En las cejas que siguen se introducen los datos del ensayo, se introducen: el nombre del operador, los datos de la muestra, y los pesos de la cápsula a usar. En la pestaña Cálculos se calibran a ceros por software los sensores a usar antes de iniciar la medición del proceso compresión - deformación.

The screenshot shows the 'LG-TRIAxIAL' software window with a menu bar (Inicio, Datos prueba, Datos muestra, Cálculos, Gráfica 1, Gráfica 2) and a main area titled 'TRIAxIAL RÁPIDA'. Under 'Datos de la muestra', there are two columns of input fields:

Parameter	Value	Unit
Vel. de aplicación de la carga. V_a	0.040	[mm/min]
Presión de confinamiento. σ_3	1	[Kg/Cm ²]
D_s	3.50	[Cm]
D_c	3.50	[Cm]
D_i	3.40	[Cm]
W_m	160.60	[gr]
H_m	8.36	[Cm]
A_s	9.6211	[Cm ²]
A_c	9.6211	[Cm ²]
A_i	9.0792	[Cm ²]
A_m	9.5308	[Cm ²]
V_m	79.6775	[Cm ³]
Y_m	2.0156	[t/m ³]

Buttons for 'Calcular' and 'Borrar' are also visible.

Fig. 10 Interfaz gráfica LG-TRIAxIAL.

También se selecciona en la pestaña Cálculos la frecuencia de captura. Una vez colocada la muestra en la cámara triaxial se pone en marcha el sistema. Al pulsar el botón *iniciar* el programa empieza a capturar datos. Y al final del ensayo calcula la deformación unitaria y el esfuerzo desviador. Al mismo tiempo que captura los datos de los sensores el programa los grafica con respecto al tiempo. El programa adquiere continuamente los valores de las señales provenientes de la celda de carga y del LVDT y presenta sus gráficas en tiempo real, Fig. 11.

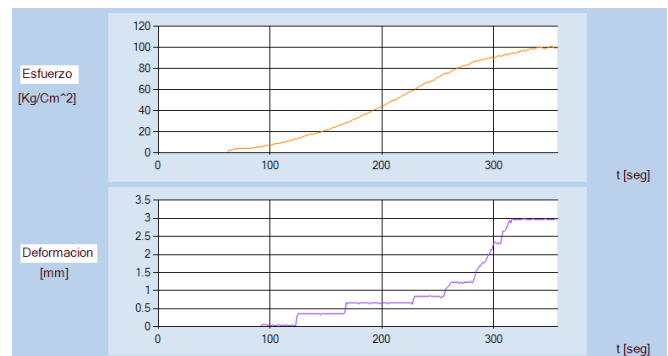


Fig. 11 Gráficas de las mediciones.

TRIAxIAL RAPIDA						
Obra:	prueba final			Fecha:	07/06/2011	
Localización:	Tuxtla Gutierrez			Operador:	Boris solis	
Sondeo No.	1			Ensaye No.	1	
Muestra No.	1	Prof. (m)	2.0	Calculó	1	
Descripción:	arcilla limosa					
Medidas de la muestra:						
Ds (cm).	2.20	As (cm ²)	3.80	Wm (gr)	132.00	Notas: la muestra ensayada en esta prueba fue una muestra sin humedad
Dc (cm).	2.50	Ac (cm ²)	4.91	Vm (cm ³)	33.05	
Di (cm).	2.70	Ai (cm ²)	5.73	γ_m (t/m ³)	3.99	
Hm (cm).	6.80	Am	4.86	(cm ²)		
VELOCIDAD DE APLICACION DE LA CARGA:				2 mm/min	σ_3	(kg/cm ²)
Tiempo transcurrido	Carga	Deformación Total	Deformación Unitaria	1- Deformación Unitaria	Area corregida	Esfuerzo Desviador
min	kg	mm			cm ²	kg/cm ²
	2.41	0.01	0.0001	0.9999	4.8610	0.4959
	13.69	0.25	0.0037	0.9963	4.8782	2.8054
	29.37	0.50	0.0074	0.9926	4.8963	5.9987
	59.27	0.75	0.0110	0.9890	4.9145	12.0609
	75.45	1.00	0.0147	0.9853	4.9328	15.2954
	85.25	1.25	0.0184	0.9816	4.9513	17.2184
	87.21	1.50	0.0221	0.9779	4.9699	17.5485
	88.20	1.75	0.0257	0.9743	4.9887	17.6790
	90.16	2.00	0.0294	0.9706	5.0076	18.0039
	90.16	2.25	0.0331	0.9669	5.0266	17.9357
	92.61	2.50	0.0368	0.9632	5.0458	18.3532
	93.59	2.75	0.0404	0.9596	5.0651	18.4767
	93.08	3.00	0.0441	0.9559	5.0846	18.3057
Contenido de agua						
Cápsula No.						15
Peso cáp + suelo húmedo						110
Peso cáp + suelo seco						91.9
Peso cápsula						18.8
Peso agua						18.1
Peso suelo seco						73
w (%)						24.70
OBSERVACIONES:						
Cápsula No.						
Peso cáp + suelo húmedo						
Peso cáp + suelo seco						
Peso cápsula						
Peso agua						
Peso suelo seco						
w (%)						
OBSERVACIONES:						

	Nota: $Area\ corregida = (Am / (1 - def. unit))$
	<p>ESQUEMA DE LA MUESTRA EN LA FALLA</p>
OBSERVACIONES	

Fig. 12 Valores registrados durante una prueba triaxial rápida y gráfica esfuerzo deformación.

El programa adquiere continuamente los valores de las señales del LVDT, pero registra solo aquellos valores que tengan un desplazamiento escalonado de 0.25mm.

Cuando el operador observa que los valores del esfuerzo desviador se repiten ó disminuyen en lugar de aumentar, significa que la muestra ha sufrido una falla. En ese momento el operador debe dar por terminado el ensayo y podrá observar las gráficas completas de las respuestas de los sensores.

Al finalizar la etapa de medición automática con LG TRIAXIAL el programa exporta los datos a una hoja de cálculo en Excel. En esta hoja se presentan el esfuerzo desviador de falla y la gráfica de esfuerzo vs deformación, Fig. 12.

Este procedimiento se ensaya con tres ó cuatro muestras de un mismo suelo. El objetivo del sistema concluye al obtener los esfuerzos máximos de desviación para cada una de las pruebas. Los esfuerzos máximos de desviación, llamados también esfuerzos de falla, sirven de base para el cálculo del ángulo de fricción " ϕ " y el coeficiente de cohesión " c ". El operador obtiene éstos a través de un procedimiento gráfico basado en los llamados círculos de Mohr.

La distancia máxima de comunicación entre el TWS-434 y el RWS-434 fue 40m en lugares cerrados y 50m en lugares abiertos, con una alimentación VCC +5V. Esta distancia satisface por completo las necesidades del sistema. La antena fue un trozo de cable UTP de 5cm de longitud. Esta distancia corresponde a un medio de la longitud requerida por el tipo de frecuencia de trabajo.

IV. Conclusiones

Se desarrolló un sistema, LG TRIAXIAL, el cual es una herramienta útil y de fácil manejo, más práctica, menos laboriosa, y ocupa menos tiempo del operador. Todo lo mencionado representa una ayuda para el área de ingeniería civil.

La prueba triaxial rápida fue sintetizada en su mayor parte a través de la adquisición y registro automático de datos. Además, se implementó una comunicación de datos inalámbrica.

V. Referencias

- [1] W. Powrie, "*Soil Mechanics: Concepts and Applications*", Ed. 2, 2004, Spon Press.
- [2] R. Whitlow, "*Basic Soil Mechanics*", Ed. 4, 2000, Prentice hall, 312-340.
- [3] I. Ishibashi, H. Hazarika, "*Soil Mechanics Fundamentals*", Ed. 1, 2010, Crc Press, 112-138.
- [4] I. Smith, "*Smith's Elements of Soil Mechanics*", Ed. 8, 2006, Crc Press, 57-68.
- [5] D. E. Radcliffe, J. Simunek, "*Soil Physics with HYDRUS: Modeling and Applications*", Ed.1, 2010, Crc Press, 89-115.
- [6] J. H. Atkinson, "*The Mechanics of Soils and Foundations*", Ed. 2, 2007, Spoon Press.
- [7] E. Juárez Badillo, "*Fundamentos de la mecánica de suelos*", Ed. 3, 2005, Limusa.
- [8] J. Axelson, "*USB Complete the Developer's Guide*", Ed. 4, 2009, Lakeview Research.
- [9] <http://www.controls.it>

VI. Autores

Dr. Rubén Herrera Galicia obtuvo el título de Doctor en Ciencias Técnicas con especialidad en electrónica por la *Technical University of Warsaw*, Polonia.

Jonnathan Rasgado Jiménez, Boris Solís Zavaleta egresados de la carrera de Ing. en electrónica por el *Instituto Tecnológico de Tuxtla Gutiérrez*.

Ecuaciones de Maxwell de Orden Fraccionario

J. Juan Rosales García, **José Francisco Gómez Aguilar, ***J. Jesús Bernal Alvarado, *Manuel Guía Calderón, *****A. González Parada, *****R. Guzmán Cabrera.**

* Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato, Carretera Salamanca-Valle de Santiago km 3.5+1.8 km. Comunidad de Palo Blanco. Tel. (464) 647 9940, ext. 2464. rosales@salamanca.ugto.mx

** Departamento de Ingeniería Física, División de Ciencias e Ingenierías del Campus León, Universidad de Guanajuato, 37150, León Gto. México. Tel. +52 (477) 788-5100, ext. 8449. jfga@fisica.ugto.mx

*** Departamento de Ingeniería Física, División de Ciencias e Ingenierías del Campus León, Universidad de Guanajuato, 37150, León Gto., México. Tel. +52 (477) 788-5100, ext. 8455. bernal@fisica.ugto.mx

**** Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato, Carretera Salamanca-Valle de Santiago km 3.5+1.8 km. Comunidad de Palo Blanco. Tel. (464) 647 9940, ext. 2369. guia@salamanca.ugto.mx

***** Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato, Carretera Salamanca-Valle de Santiago km 3.5+1.8 km. Comunidad de Palo Blanco. Tel. (464) 647 9940, ext. 2369. gonzaleza@salamanca.ugto.mx

***** Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato, Carretera Salamanca-Valle de Santiago km 3.5+1.8 km. Comunidad de Palo Blanco. Tel. (464) 647 9940, ext. 2369. guzmanc@salamanca.ugto.mx

Resumen

En el presente trabajo se analiza el comportamiento de las ondas electromagnéticas obtenidas de las ecuaciones de Maxwell fraccionarias. El parámetro que representa el orden de la derivada fraccionaria es $0 < \gamma \leq 1$. Se estudian los casos en que la derivada fraccionaria es respecto al tiempo y al espacio. Se consideran los casos particulares $\gamma=1$ y $\gamma=1/2$.

Palabras Clave: Cálculo Fraccionario, Ecuaciones de Maxwell, Ondas Electromagnéticas.

I. Introducción

Las ecuaciones de Maxwell son un conjunto de cuatro ecuaciones que describen por completo los fenómenos electromagnéticos. La gran contribución de James Clerk Maxwell fue reunir en estas ecuaciones largos años de resultados experimentales, debidos a Coulomb, Gauss, Ampere, Faraday y otros, introduciendo los conceptos de campo y corriente de desplazamiento, y unificando los campos eléctricos y magnéticos en un solo concepto: el campo electromagnético. Las ecuaciones de Maxwell para las ondas electromagnéticas en un medio se escriben de la siguiente manera

$$\vec{\nabla} \cdot \vec{E} = \frac{4\pi}{\epsilon} \rho(\vec{r}, t), \quad (1)$$

$$\vec{\nabla} \cdot \vec{B} = 0, \quad (2)$$

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c} \frac{\partial \vec{B}}{\partial t}, \quad (3)$$

$$\vec{\nabla} \times \vec{B} = \frac{4\pi\mu}{c} \vec{j}(\vec{r}, t) + \frac{\epsilon\mu}{c} \frac{\partial \vec{E}}{\partial t}. \quad (4)$$

donde $\rho(\vec{r}, t)$ y $\vec{j}(\vec{r}, t)$ son, la densidad de carga y la densidad de corriente, respectivamente. En (4),

$\vec{D} = \epsilon \vec{E}$ es el desplazamiento eléctrico proporcional al campo eléctrico \vec{E} , donde ϵ es la permitividad eléctrica, y el campo magnético $\vec{B} = \mu \vec{H}$, donde \vec{H} es la intensidad del campo magnético y μ la permeabilidad magnética. En presencia de materia la intensidad del campo magnético \vec{H} se reemplaza por la inducción magnética \vec{B} . En el vacío estas dos cantidades son iguales. En el caso de medios homogéneos e isotrópicos los parámetros ϵ y μ son constantes, en caso contrario son tensores.

Introduciendo el potencial vectorial $\vec{A} = (x_i, t)$ y el potencial escalar $\phi(x_i)$ de la siguiente manera

$$\vec{B} = \vec{\nabla} \times \vec{A} \quad (5)$$

$$\vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \vec{\nabla} \phi \quad (6)$$

La representación de los campos por potenciales se sigue directamente de las ecuaciones de Maxwell (2) y (3). Sustituyendo estas expresiones en las ecuaciones (1) y (4), se obtienen dos ecuaciones diferenciales acopladas para los potenciales

$$\Delta \phi + \frac{1}{c} \frac{\partial}{\partial t} (\vec{\nabla} \cdot \vec{A}) = -\frac{4\pi}{\epsilon} \rho \quad (7)$$

$$\Delta \vec{A} - \frac{\epsilon \mu}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \vec{\nabla} \left(\vec{\nabla} \cdot \vec{A} + \frac{1}{c} \frac{\partial \phi}{\partial t} \right) = \frac{4\pi}{c} \vec{j} \quad (8)$$

Para desacoplar estas ecuaciones se usa una calibración apropiada del potencial. El campo magnético (5) no cambia si el vector potencial se transforma de la siguiente manera

$$\vec{A}' = \vec{A} + \vec{\nabla} \alpha \quad (9)$$

De la ecuación (6) se calcula el potencial escalar

$$\phi' = \phi - \frac{1}{c} \frac{\partial \alpha}{\partial t} \quad (10)$$

Los campos electromagnéticos representan

cantidades medibles; los potenciales son construcciones auxiliares. A tales transformaciones se les llama transformaciones gauge (de calibración).

Las ecuaciones de Maxwell no cambian bajo las transformaciones de calibración; es decir, son invariantes de calibración.

Usando la calibración de Lorentz

$$\vec{\nabla} \cdot \vec{A} + \frac{\epsilon \mu}{c} \frac{\partial \phi}{\partial t} = 0 \quad (11)$$

las ecuaciones (7) y (8) se escriben como

$$\Delta \vec{A}(\vec{r}, t) - \frac{\epsilon \mu}{c^2} \frac{\partial^2 \vec{A}(\vec{r}, t)}{\partial t^2} = -\frac{4\pi}{c} \vec{j}(\vec{r}, t) \quad (12)$$

$$\Delta \phi(\vec{r}, t) - \frac{\epsilon \mu}{c^2} \frac{\partial^2 \phi(\vec{r}, t)}{\partial t^2} = -\frac{4\pi}{\epsilon} \rho(\vec{r}, t) \quad (13)$$

donde $\frac{\epsilon \mu}{c^2} = \frac{1}{v^2}$ y v es la velocidad de la luz en el medio. Las ecuaciones (12) y (13) son ecuaciones desacopladas para los potenciales $\vec{A} = (x_i, t)$ y $\phi(\vec{r}, t)$.

Calculo Fraccionario

El cálculo fraccionario es tan antiguo como el cálculo ordinario, sin embargo no es tan popular en la ciencia ni en la ingeniería como el ordinario. La peculiaridad del cálculo fraccionario radica en que las derivadas e integrales fraccionarias tienen propiedades no locales, por lo que considera la historia y los efectos no locales distribuidos de cualquier sistema físico. Por ésta propiedad tan peculiar, se espera que el cálculo fraccionario pueda proporcionar información más exacta de la naturaleza.

Las aplicaciones del cálculo fraccionario al desarrollo científico y problemas de ingeniería han crecido considerablemente en los últimos años [1]-[2]. El reciente interés por el cálculo fraccionario y en particular por las ecuaciones diferenciales fraccionarias, se ve estimulado por

las aplicaciones en varias áreas de la física, química, biología e ingeniería, [3]-[10]. Sin embargo, la derivación de estas ecuaciones de algunas leyes fundamentales no es una tarea fácil. El operador fraccionario refleja procesos disipativos intrínsecos de la naturaleza, los cuales son complicados de interpretar. Su relación teórica con el cálculo fraccionario aún no es del todo comprendida.

Desde el punto de vista matemático no hay ningún problema hacer el cambio directo de la derivada ordinaria a la fraccionaria, es decir,

$$\frac{\partial}{\partial t} \rightarrow \frac{\partial^\gamma}{\partial t^\gamma}, \quad 0 < \gamma \leq 1 \quad (14)$$

Sin embargo, para la física e ingeniería es importante tomar en cuenta las dimensiones de las magnitudes. Por tal motivo no podemos hacer el cambio (14) formalmente.

Definiendo el operador derivada temporal de la siguiente manera

$$\frac{\partial}{\partial t} \rightarrow \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma}{\partial t^\gamma}, \quad n-1 < \gamma \leq n \quad (15)$$

donde γ es un parámetro arbitrario que representa la derivada temporal fraccionaria, $0 < \gamma \leq 1$ y σ es un parámetro dimensional. Las dimensiones son

$$\left[\frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma}{\partial t^\gamma} \right] = \frac{1}{s}, \quad 0 < \gamma \leq 1 \quad (16)$$

tal que cuando $\gamma=1$ la expresión (16) se transforme en derivada ordinaria.

$$\frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma}{\partial t^\gamma} = \frac{\partial}{\partial t} \quad (17)$$

En este trabajo se va a usar la definición de derivada fraccionaria de Caputo [11], dada por:

$${}_0^c D_t^\gamma f(t) = \frac{1}{\Gamma(n-\gamma)} \int_0^t \frac{f^{(n)}(\tau)}{(t-\tau)^{\gamma+1-n}} d\tau, \quad (18)$$

donde $n=1,2,\dots \in \mathbb{N}$ y $n-1 < \gamma \leq n$. Una de las principales ventajas de la derivada de Caputo es que las condiciones iniciales para las ecuaciones diferenciales fraccionarias están dadas de la misma manera que en el cálculo ordinario con una interpretación física conocida. Este hecho es importante desde el punto de vista práctico.

Por ejemplo, para la función $f(t) = t^k$, donde k es un número arbitrario $0 < \gamma \leq 1$, la derivada fraccionaria es

$$D_t^\gamma t^k = \frac{\Gamma(k)k}{\Gamma(k+1-\gamma)} t^{k-\gamma}, \quad (19)$$

si $\gamma=1$,

$$D_t^1 = \frac{\partial t^k}{\partial t} = kt^{k-1}. \quad (20)$$

II. Desarrollo

Ecuaciones de Maxwell Fraccionarias

Las ecuaciones de Maxwell fraccionarias tienen la forma

$$\vec{\nabla} \cdot \vec{E} = \frac{4\pi}{\epsilon}, \quad (21)$$

$$\vec{\nabla} \cdot \vec{B} = 0, \quad (22)$$

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c} \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma \vec{B}}{\partial t^\gamma}, \quad (23)$$

$$\vec{\nabla} \times \vec{B} = \frac{4\pi\mu}{c} \vec{j} + \frac{\epsilon\mu}{c} \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma \vec{E}}{\partial t^\gamma}. \quad (24)$$

Usando la relación $\vec{B} = \vec{\nabla} \times \vec{A}$ y sustituyendo en (23), se tiene

$$\vec{\nabla} \times \left(\frac{1}{c} \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma \vec{A}}{\partial t^\gamma} + \vec{E} \right) = 0 \quad (25)$$

de esta condición resulta que se puede representar el campo eléctrico E mediante el campo escalar ϕ .

$$E = -\frac{1}{c} \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma \vec{A}}{\partial t^\gamma} - \vec{\nabla} \phi \quad (26)$$

donde $\phi(x_i, t)$ es el potencial escalar. La expresión (26) es invariante respecto a las transformaciones de calibración $U(1)$ con la derivada fraccionaria

$$\vec{A}' = \vec{A} + \vec{\nabla} \alpha \quad \phi' = \phi - \frac{1}{c} \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma \alpha}{\partial t^\gamma} \quad (27)$$

Tomando la condición de calibración como en el caso ordinario se tiene

$$\frac{\partial}{\partial t} \rightarrow \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma}{\partial t^\gamma}, \quad n-1 < \gamma \leq n \quad (28)$$

Entonces, se tiene

$$\vec{\nabla} \times \vec{A} + \frac{\varepsilon\mu}{c} \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma \phi}{\partial t^\gamma} = 0 \quad (29)$$

reemplazando (29) en (21), resulta

$$\vec{\nabla} \left(-\frac{1}{c} \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma \vec{A}}{\partial t^\gamma} - \vec{\nabla} \phi \right) = \frac{4\pi\rho}{\varepsilon} \quad (30)$$

Sustituyendo $\vec{B} = \vec{\nabla} \times \vec{A}$ en (24) se obtiene

$$\vec{\nabla} \times \vec{\nabla} \times \vec{A} - \frac{\varepsilon\mu}{c} \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma}{\partial t^\gamma} \left(-\frac{1}{c} \frac{1}{\sigma^{1-\gamma}} \frac{\partial^\gamma \vec{A}}{\partial t^\gamma} - \vec{\nabla} \phi \right) = \frac{4\pi\mu}{c} \vec{j} \quad (31)$$

Sustituyendo en (30) y (31) la condición de calibración (29) se obtienen las siguientes ecuaciones fraccionarias desacopladas para los potenciales, vectorial y escalar

$$\Delta \vec{A} - \frac{\varepsilon\mu}{c^2} \frac{1}{\sigma^{2(1-\gamma)}} \frac{\partial^{2\gamma}}{\partial t^{2\gamma}} \vec{A} = -\frac{4\pi\mu}{c} \vec{j} \quad (32)$$

$$\Delta \phi - \frac{\varepsilon\mu}{c^2} \frac{1}{\sigma^{2(1-\gamma)}} \frac{\partial^{2\gamma}}{\partial t^{2\gamma}} \phi = -\frac{4\pi}{\varepsilon} \rho \quad (33)$$

Estas ecuaciones se transforman en (12) y (13) si $\gamma=1$.

Si $\rho=0$ y $\vec{j}=0$, se tienen ondas en un medio fractal

$$\Delta \vec{A} - \frac{\varepsilon\mu}{c^2} \frac{1}{\sigma^{2(1-\gamma)}} \frac{\partial^{2\gamma}}{\partial t^{2\gamma}} \vec{A} = 0 \quad (34)$$

$$\Delta \phi - \frac{\varepsilon\mu}{c^2} \frac{1}{\sigma^{2(1-\gamma)}} \frac{\partial^{2\gamma}}{\partial t^{2\gamma}} \phi = 0 \quad (35)$$

Considerando que la onda electromagnética se propaga solo en la dirección x , se pueden escribir las ecuaciones (34) y (35) en una sola, de la siguiente manera

$$\frac{\partial^2 z(x,t)}{\partial x^2} - \frac{1}{v^2} \frac{1}{\sigma^{2(1-\gamma)}} \frac{\partial^{2\gamma} z(x,t)}{\partial t^{2\gamma}} = 0 \quad (36)$$

donde $z(x,t)$ representa a $\vec{A}(x,t)$ y $\phi(x,t)$. Si se considera que la onda electromagnética es polarizada, se tiene $A_x(x,t) = 0, A_y(x,t) \neq 0,$

$A_z(x,t) \neq 0$. La ecuación (36) es lineal y la solución particular es de la siguiente forma

$$z(x,t) = z_0 e^{-ikx} \square u(t) \quad (37)$$

donde k es la componente del vector de onda en la dirección x . Sustituyendo (37) en (36), se obtiene la ecuación diferencial fraccionaria respecto a t

$$\frac{\partial^{2\gamma} u(t)}{\partial t^{2\gamma}} + v^2 k^2 \sigma^{2(1-\gamma)} u(t) = 0 \quad (38)$$

Redefiniendo

$$\omega^2 = v^2 k^2 \sigma^{2(1-\gamma)} = \omega_0^2 \sigma^{2(1-\gamma)} \quad (39)$$

donde ω_0 es la frecuencia de la onda electromagnética, la ecuación (38) se puede escribir como

$$\frac{\partial^{2\gamma} u(t)}{\partial t^{2\gamma}} + \omega^2 u(t) = 0 \quad (40)$$

La solución de esta ecuación se busca en forma de series de potencias. Usando la definición (19) para derivadas fraccionarias de funciones en potencias, se tiene la solución

$$u(t) = E_{2\gamma}(-\omega^2 t^{2\gamma}) \quad (41)$$

donde

$$E_{2\gamma}(-\omega^2 t^{2\gamma}) = \sum_{n=0}^{\infty} \frac{(-\omega^2 t^{2\gamma})^n}{\Gamma(2n\gamma + 1)} \quad (42)$$

es la función de Mittag-Leffler.

III. Resultados

Análisis de los Casos $\gamma=1$ y $\gamma=1/2$.

En el primer caso $\gamma=1$, la función de Mittag-Leffler (41) se transforma en cosenos hiperbólicos $(\sigma^{1-\gamma})=1$

$$u(t) = E_2(-\omega_0^2 t^2) = \cosh\left(\sqrt{-\omega_0^2 t^2}\right) = \cosh(i\omega_0 t) = \cos(\omega_0 t) \quad (43)$$

La expresión (43) representa una función periódica respecto a t . Entonces, se puede escribir como

$$z(x, t) = \text{Re } z_0 e^{-i(\omega_0 t - kx)} \quad (44)$$

la cual define una onda periódica monocromática en x y t . Este resultado es muy bien conocido en la teoría de ondas electromagnéticas ordinarias.

Para el segundo caso $\gamma=1/2$, la ecuación (36) se transforma en

$$\frac{\partial^2 z(x, t)}{\partial x^2} - \frac{\epsilon\mu}{c^2} \frac{1}{\sigma} \frac{\partial z(x, t)}{\partial t} = 0 \quad (45)$$

La solución puede ser escrita en la forma (37), y se obtiene la siguiente ecuación para la función $u(t)$

$$\frac{du}{dt} + \omega^2 u(t) = 0 \quad (46)$$

donde ω está definida en (39). La solución puede ser obtenida de (41) para $\gamma=1/2$, o por integración directa de (46). El resultado es

$$u(t) = e^{-\omega^2 t}. \quad (47)$$

En este caso, se tiene

$$z(x, t) = z_0 e^{-\omega^2 t} e^{-ikx} \quad (48)$$

donde $\omega^2 = \omega_0^2 \sigma$. Para este caso la solución es periódica respecto a x y aperiódica respecto a t . La solución representa una onda plana con una amplitud que decae en el tiempo. El tiempo en el cual la amplitud z_0 decae e veces es

$$t_0 = \frac{1}{\omega^2} = \frac{1}{\omega_0 \sigma}. \quad (49)$$

Es importante hacer notar que se tiene una magnitud adimensional γ la cual caracteriza el orden de la derivada fraccionaria y la otra magnitud σ con dimensiones de tiempo que caracteriza la presencia de fractales en el medio.

Estas dos magnitudes se relacionan de la siguiente manera

$$\gamma = \sigma^2 \omega_0^2 = \frac{\sigma^2}{T_0^2} = \frac{\sigma_x^2}{\lambda^2}, \quad 0 < \sigma \leq T_0, \quad (50)$$

donde T_0 es el periodo de la onda, λ , su longitud y $\sigma = \frac{\sigma_x}{v}$.

Tomando en cuenta esta relación se puede escribir la solución (41) de la siguiente manera

$$u(t) = E_{2\gamma} \left(-\gamma^{(1-\gamma)} \tilde{t}^{2\gamma} \right) \quad (51)$$

donde $\tilde{t} = \frac{t}{T_0}$ es una magnitud adimensional, es decir, el tiempo dividido por el periodo de la onda electromagnética.

Derivada Fraccionaria Espacial

Ahora se analizara la ecuación (36) suponiendo que la derivada espacial es fraccionaria y la temporal es ordinaria. Entonces se tiene

$$\frac{1}{\sigma_x^{2(1-\delta)}} \frac{\partial^{2\delta} \tilde{z}(x, t)}{\partial x^{2\delta}} - \frac{1}{v^2} \frac{\partial^2 \tilde{z}(x, t)}{\partial t^2} = 0 \quad (52)$$

si $\delta=1$ se obtiene la solución periódica respecto a t y a x . Para el caso $\delta=1/2$ la onda solo es periódica respecto a t .

La solución de la ecuación (52) se busca de la siguiente manera

$$\tilde{z}(x, t) = \tilde{z}_0 e^{i\omega t} u(x) \quad (53)$$

Sustituyendo (53) en (52), obtenemos la ecuación

$$\frac{\partial^{2\delta} u(x)}{\partial x^{2\delta}} + \tilde{k}^2 u(x) = 0 \quad (54)$$

donde $\tilde{k}^2 = \frac{\omega^2}{v^2} \sigma_x^{2(1-\delta)} = k^2 \sigma_x^{2(1-\delta)}$ es el vector de onda en el medio en presencia de fractales y k es el vector de onda electromagnética en el medio. La solución de la ecuación (54) es la función de Mittag-Leffler, la cual tiene la forma

$$u(x) = E_{2\delta}(-\tilde{k}^2 x^{2\delta}) = \sum_{n=0}^{\infty} \frac{\left(\tilde{k}^2 x^{2\delta}\right)^n}{\Gamma(2n\delta + 1)} \quad (55)$$

Para el caso $\delta=1$, se tiene

$$E_2(-k^2 x^2) = \cosh(\sqrt{-k^2 x^2}) = \cosh(ik^2 x^2) = \text{Re}(e^{-ikx}) \quad (56)$$

Usando la fórmula $\cosh(-ikx)=\cos(-kx)=\text{Re}(e^{-ikx})$, para la función de onda se tiene

$$\tilde{z}(x, t) = \text{Re} \tilde{z}_0 e^{-i(\omega t - kx)} \quad (57)$$

con $\tilde{k} = \frac{\omega}{v} \sigma_x^{1-1} = \frac{\omega}{v} = k$, donde k es la componente del vector de onda en la dirección x y está relacionado con la longitud de onda electromagnética por la siguiente expresión $k = \frac{1}{\lambda}$.

Para el segundo caso $\gamma=1/2$, se tiene, $\tilde{k}^2 = \frac{\omega^2}{v^2} \sigma_x^{2(1/2)}$ y $[\tilde{k}^2] = \frac{1}{m}$ tiene dimensiones de longitud inversa. La solución para este caso tiene la forma

$$u(t) = E_{(2)(1/2)}(-\tilde{k}^2 x^{2(1/2)}) = e^{-\tilde{k}^2 x} \quad (58)$$

Es decir, la función de onda es periódica solamente respecto a t . La distancia en la cual la amplitud \tilde{z}_0 se reduce en e veces es igual a

$$x_0 = \frac{1}{\tilde{k}^2} = \frac{1}{k^2 \sigma_x} \quad (59)$$

En este caso se tiene que δ es una magnitud adimensional y σ_x se relaciona con la medida del fractal. Entre estas dos magnitudes existe la relación

$$\delta = k^2 \sigma_x^2 = \frac{\sigma_x^2}{\lambda^2}. \quad (60)$$

Se puede usar esta relación para escribir la ecuación (55) de la siguiente manera

$$u(t) = E_{2\delta}(-\delta^{(1-\delta)} \tilde{x}^{2\delta}) \quad (61)$$

donde $\tilde{x} = \frac{x}{\lambda}$.

Se puede observar que las soluciones (51) y (61) tienen la misma estructura y en el caso $\delta=\gamma$ tendrían la misma forma. Entonces se puede graficar la función $u(s)$ donde $s = (\tilde{x}, \tilde{t})$, (ver Fig. 1)

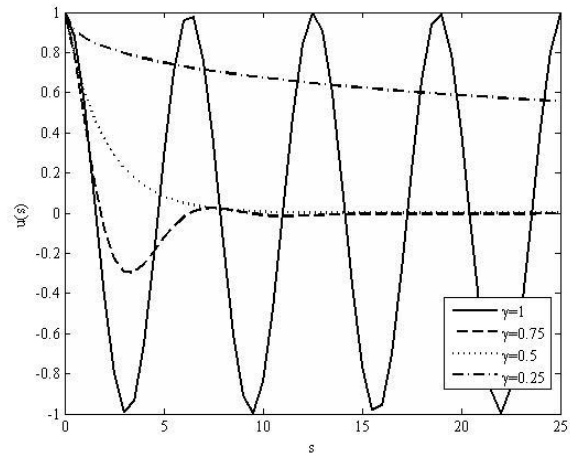


Fig. 1. Grafica correspondiente a las ecuaciones (51) y (61).

IV. Discusión

Las ecuaciones de Maxwell permiten ver en forma clara que la electricidad y el magnetismo son dos manifestaciones de un mismo fenómeno físico, el electromagnetismo. La representación generalizada de estas ecuaciones permite obtener más información acerca del fenómeno físico que ellas describen. En interacciones con células se sabe que una célula viva se compone de aproximadamente 80 por ciento de agua y cerca del 20 por ciento de su estructura la forman las macromoléculas, por ejemplo ADN y ARN. Estas macromoléculas tienen una dimensión de aproximadamente de 100-300 μm . En la célula también es posible la existencia de agrupamientos del mismo orden. Tales estructuras con dimensiones, $v=100-300 \mu\text{m}$, se pueden ver como fractales cuando interactúan con una onda electromagnética de longitud $\lambda \geq v$.

V. Conclusiones

Se ha presentado un ejemplo donde pueden aparecer estructuras fractales. Es de interés estudiar los efectos que causan estas ondas electromagnéticas al interactuar con sistemas biológicos, en particular con células.

VI. Referencias

- [1] D. Cafagna, “Past and present—fractional calculus: a mathematical tool from the past for present engineers”, *IEEE Industrial Electronics Magazine*, vol. 1, no. 2, 2007, pp. 35–40.
- [2] J. A. T. Machado. “A probabilistic interpretation of the fractional-order differentiation”. *Fractional Calculus and applied Analysis*, vol. 6, no. 1, 2003, pp. 73–80.
- [3] Bisquert J. *Phys. Rev. Lett.* 91, 010602. 2003.
- [4] Boldyrev S., Gwinn C.R. *Phys. Rev. Lett.* 91(131101). 2003.
- [5] Chechkin A.V., Gonchar V.Y., Szydlowsky M. *Phys. Plasma* 9(78). 2002.
- [6] Lacasta A.M., Sancho J.M., Romero A.H., Sokolov I.M., Lindenberg K. *Phys. Rev. E* 70(051104). 2004.
- [7] Mainardi F., “Fractals and Fractional Calculus in Continuum Mechanics”, Edited by A. Carpinteri, F. Mainardi (Springer-Verlag, New York), 1997, pp. 291-348.
- [8] Seredynska M., Hanyga A. *J. Math. Phys.* 41(2135). 2000.
- [9] Stanislavsky A.A. *JETP* 98(705). 2004.
- [10] Tarasov V.E., Zaslavsky G.M., *Physica A* 354(249). 2005.
- [11] Podlubny I. “The Laplace transform method for linear differential equations of the fractional

order”. Tech. Rep., Slovak Academy of Sciences, Institute of Experimental Physics. 1994.

VII. Autores

Dr. J. Juan Rosales García obtuvo su título de Doctor en Ciencias en la UAM Iztapalapa. Actualmente trabaja en el Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato. Sus intereses son efectos electromagnéticos en sistemas biológicos y aplicaciones del cálculo fraccionario.

M. I. José Francisco Gómez Aguilar obtuvo su título de Maestría en Ingeniería Eléctrica en la Facultad de Ingeniería Mecánica Eléctrica y Electrónica de la Universidad de Guanajuato con especialidad en Instrumentación y Sistemas Digitales. Actualmente estudia su Doctorado en el Departamento de Ingeniería Física, División de Ciencias e Ingenierías del Campus León, Universidad de Guanajuato. Sus intereses son el bioelectromagnetismo, el modelado matemático de sistemas biológicos y las aplicaciones del cálculo fraccionario.

Dr. J. Jesús Bernal Alvarado obtuvo su título de Doctor en Ciencias en la Universidad de Guanajuato. Actualmente trabaja en el Departamento de Ingeniería Física, División de Ciencias e Ingenierías del Campus León, Universidad de Guanajuato. Sus intereses son la espectroscopia de impedancia eléctrica, el modelado matemático de sistemas biológicos y bioelectromagnetismo.

M. I. Manuel Guía Calderón obtuvo su título de Maestría en Ingeniería Eléctrica en la Facultad de Ingeniería Mecánica Eléctrica y Electrónica de la Universidad de Guanajuato con especialidad en Instrumentación y Sistemas Digitales. Actualmente trabaja en el Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de

Guanajuato. Sus intereses son efectos electromagnéticos en sistemas biológicos.

Dr. Adrian González Parada obtuvo el grado de Doctor en Ingeniería Eléctrica en la Universidad Politécnica de Cataluña España. Actualmente trabaja en el Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato. Sus intereses son aplicaciones electrotécnicas de la superconductividad en equipos de generación, transmisión y distribución de la ingeniería eléctrica y bioelectromagnetismo.

Dr. Rafael Guzmán Cabrera obtuvo el grado de Doctor en Inteligencia Artificial por la Universidad Politécnica de Valencia, España. Actualmente trabaja en el Departamento de Ingeniería Eléctrica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato. Sus intereses son inteligencia artificial, reconocimiento de patrones, métodos numéricos aplicados a la ingeniería.

Transmisor Indicador de Temperatura (TIT) con Interfaz para LABVIEW.

***Manuel Vázquez Briones, **Javier Garrido Meléndez,
*** Juan Carlos Hernández Cruz, **** Wendy Ventura Aragón.**

*Universidad Tecnológica del Sureste de Veracruz (UTSV), Av. Universidad Tecnológica Lote Grande No. 1, Sin Colonia, Nanchital, Veracruz, CP 96360, (921) 2110165, profe.briones.utsv@hotmail.com.

**UTSV, Av. Universidad Tecnológica Lote Grande No. 1, Sin Colonia, Nanchital, Veracruz, CP 96360, (921) 2110165, odirragmj@hotmail.com.

***UTSV, Av. Universidad Tecnológica Lote Grande No. 1, Sin Colonia, Nanchital, Veracruz, CP 96360, (921) 2110165, jc_hernandezcruz@hotmail.com; profe.briones.utsv@hotmail.com.

****UTSV, Av. Universidad Tecnológica Lote Grande No. 1, Sin Colonia, Nanchital, Veracruz, CP 96360, (921) 2110165, wendy_vianey81@hotmail.com.

Resumen

Se presenta la continuación del artículo publicado en el SENIE 10, titulado “Transmisor Indicador de Temperatura (TIT) con Comunicación RS-232 con base en el Microcontrolador PIC16F877A”. Al diseño del transmisor anterior se le realizan, entre otras mejoras, la capacidad de conectar dos diferentes tipos de sensores: un RTD PT100 y un Termopar tipo J; Se modificó la interfaz de comunicación con la computadora, la cual pasó de ser RS-232 a comunicación por interfaz USB; y se implementó una interfaz gráfica de monitoreo con LabVIEW®. La razón de diseñar un transmisor es la necesidad del monitoreo de las variables que intervienen en los diversos procesos productivos, lo cual es requerido para el control de las mismas en los procesos. Una de las cuatro variables de proceso básicas es la temperatura, en este artículo se presenta un transmisor indicador de temperatura (TIT) con interfaz para LabVIEW® a través de USB y que es capaz de operar con los dos sensores de temperatura más usuales en el sector industrial, además de contar con una salida estándar de 4 – 20 mA.

Palabras claves: Transmisor, termopar, RTD PT100.

I. Introducción

La temperatura es la magnitud física que expresa el grado o nivel de energía térmica de los cuerpos o del ambiente^[1]. La temperatura suele medirse en grados Celsius (°C), grados Fahrenheit (°F) y grados Kelvin (°K).

Los transmisores son instrumentos que captan la variable de proceso y la transmiten a distancia a un instrumento receptor indicador, registrador, controlador o una combinación de estos^[2].

El diseño del transmisor de temperatura propuesto cuenta con un indicador que despliega la variable temperatura dentro de un rango de 0 a 150 grados centígrados emitiendo una salida de corriente de 4 a 20 mA, además se

cuenta con una interfaz para la comunicación serial por USB y es capaz de operar tanto con RTD como con termopares tipo J. En instrumentación industrial, los termopares y RTD son ampliamente usados como sensores de temperatura. Son económicos, intercambiables, tienen conectores estándar y son capaces de medir un amplio rango de temperaturas.

II. Desarrollo

2.1 Funcionamiento del transmisor

El diagrama de bloques que se muestra en la Figura 1, describe el funcionamiento del transmisor. En este diagrama los bloques de

entrada corresponden a los sensores de temperatura: el PT100 y el termopar.

El RTD PT100 varía su valor resistivo en función de la temperatura, tiene la característica de que a 0°C presenta una resistencia de $100\ \Omega$. El termopar tipo J mide un rango de valores de -270° y hasta 1370° , el cual genera un milivoltaje en función de la temperatura. A un valor de 100°C genera $4.6\ \text{mV}$ por lo que es necesario amplificar dicho voltaje a valores manipulables en el microcontrolador.

A la salida de los sensores requiere implementar una etapa de acondicionamiento de señal, la cual se encargue de convertir la señal de un valor óhmico o de milivoltaje en una señal de voltaje.

Para lo anterior es necesaria una etapa de acondicionamiento que se divide en dos partes, la primera corresponde al convertidor de V/I, el cual es el encargado de dar una salida de corriente en función del valor de voltaje de la etapa anterior, es importante resaltar que los transmisores de tipo industrial requieren que la salida sea un estándar en corriente de 4 a 20 mA.

La señal de salida de la etapa de acondicionamiento se envía al microcontrolador, éste es el encargado de realizar el procesamiento necesario para mostrar la temperatura a través de un LCD, además de enviarla a través de la interfaz USB.

La interfaz USB es la que permite llevar a cabo la comunicación con una computadora por medio de una interfaz gráfica diseñada en

LabVIEW®, en la cual se mostrará el valor de la temperatura y de corriente generados por el transmisor.

2.2 Sensores

Los RTD (Resistance Temperature Detector) son sensores de temperatura basados en la variación de la resistencia de un conductor con la temperatura. Su símbolo indica una variación lineal con coeficiente de temperatura positivo.

Al calentarse un metal habrá una mayor agitación térmica, dispersándose más los electrones y reduciéndose su velocidad media, aumentando la resistencia. A mayor temperatura, mayor agitación, y mayor resistencia.

La variación de la resistencia puede ser expresada de manera polinómica como se muestra a continuación. Por lo general, la variación es bastante lineal en márgenes amplios de temperatura.

$$R = R_0(1 + \alpha \Delta T) \quad (1)$$

donde:

R resistencia medida.

R_0 resistencia a la temperatura de referencia T_0 .

ΔT desviación de temperatura respecto a T_0
($\Delta T = T - T_0$)

α es el coeficiente de temperatura de la resistencia.

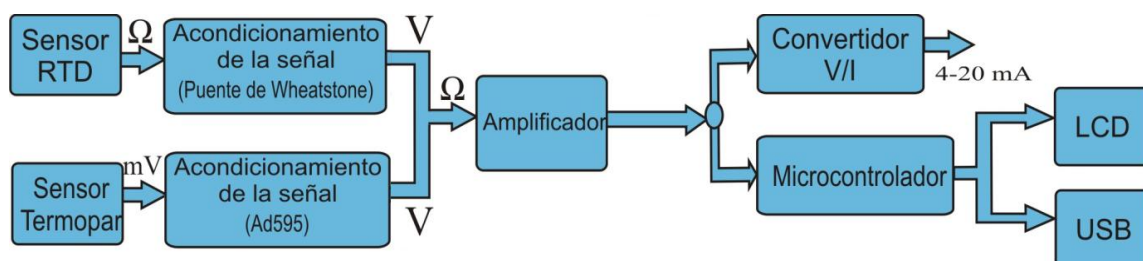


Fig.1 Diagrama a bloques del transmisor.

Los materiales empleados para la construcción de sensores RTD suelen ser conductores tales como el cobre, el níquel o el platino. De todos ellos es el platino el que ofrece mejores prestaciones. El sensor empleado es el Pt100 (RTD de platino con $R=100 \Omega$ a 0°C) con un valor $\alpha = 0.00385 \Omega / \Omega / \text{K}$.

Con la fórmula 1, conociendo el valor de α es posible calcular el valor óhmico del sensor.

Los termopares

El termopar se basa en el efecto descubierto por Seebeck en 1821, de la circulación de una corriente en un circuito formado por dos metales diferentes cuyas uniones (unión de medida o caliente y unión de referencia o fría) se mantienen a distinta temperatura. Esta circulación de corriente obedece a dos efectos termoeléctricos combinados, el efecto Peltier que provoca la liberación o absorción de calor en la unión de dos metales distintos cuando una corriente circula a través de la unión y el efecto Thomson que consiste en la liberación o absorción de calor cuando una corriente circula a través de un metal homogéneo en el que existe un gradiente de temperaturas.¹

En un sistema programable, se pueden emplear polinomios que aproximan la curva de calibración de los termopares con una exactitud dependiente de su orden.

$$T = a_0 + a_1 E + a_3 E^3 + \dots + a_n E^n \dots (2)$$

E = voltaje termoeléctrico (microvoltios).

a_n = Coeficientes del polinomio (diferentes para cada tipo de termopar).

T = temperatura ($^\circ\text{C}$).

n = orden del polinomio.

donde E es la tensión leída en el termopar. Los coeficientes del polinomio están disponibles para cada tipo de termopar y para un determinado rango de medida. La precisión del

polinomio depende del número de coeficientes utilizados.

0°C a 760°C con un rango de error de -0.9°C a 0.7°C
$a_0 = 0.0$
$a_1 = 1.9323799 \times 10^{-2}$
$a_2 = -1.0306020 \times 10^{-7}$
$a_3 = 3.7084018 \times 10^{-12}$
$a_4 = -5.1031937 \times 10^{-17}$

Tabla 1. Coeficientes del polinomio para Termopar Tipo J

Para aplicar el efecto Seebeck a la medida de temperaturas, es necesario mantener una de las uniones a la temperatura de referencia.

Una solución consiste en introducir la unión de referencia en hielo fundente. Esta solución, aunque de gran exactitud, es poco práctica y costosa por el uso de hilo extra en uno de los metales. Otra solución más factible es la denominada compensación electrónica de la unión de referencia. Consiste en dejar la unión fría a la temperatura ambiente y medir ésta con otro sensor dispuesto en sus cercanías. A la tensión del termopar se le suma la tensión que corresponde a la temperatura ambiente. De forma que es como si la unión fría estuviera a 0°C . El diagrama a bloques se muestra en la figura 1.

La principal limitación de los termopares es la exactitud ya que los errores del sistema inferiores a un grado Celsius son difíciles de obtener.

2.3 Acondicionamiento de la señal

El acondicionamiento de la señal se realiza de modo distinto para cada uno de los sensores. En el caso del RTD se conecta un puente de Wheatstone el cual emite una señal de milivoltaje como resultado de la variación de la resistencia en el sensor. La figura 2 muestra la configuración de puente de Wheatstone utilizado.

¹ Creus, Antonio Soler. Instrumentación Industrial.

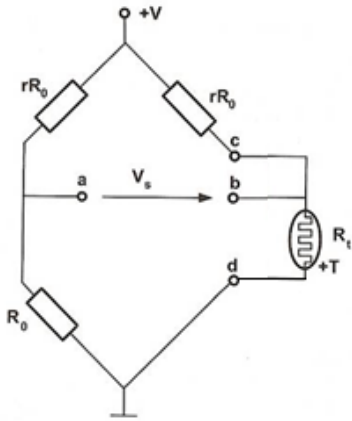


Fig.2 Puente de Wheatstone con RTD de dos hilos.

El puente de Wheatstone da un voltaje V_s en el orden de los milivolts. Para distancias largas es recomendable utilizar un puente que permita configurar un RTD de tres o cuatro hilos.

$$V_s = V \left[\frac{r(\alpha t)}{(r + 1)(r + 1 + \alpha t)} \right] \quad (3)$$

Con la ecuación 3 se obtiene un voltaje de salida de 680.19 mV. Dado que este valor es muy pequeño, se requiere un amplificador.

Para el acondicionamiento del termopar se utiliza un AD594 el cual es un amplificador de instrumentación y compensador de unión en frío para termopar.

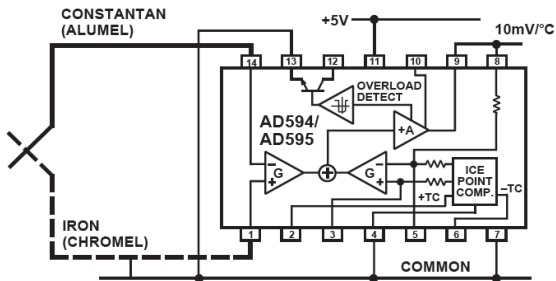


Figura 3. Diagrama del AD594.

Este integrado produce una salida de $10\text{mV}/^\circ\text{C}$. Para volver proporcional esta salida, la ganancia es ajustada para que coincida con la característica de transferencia del termopar a 25

$^\circ\text{C}$, la cual es de $40.44 \mu\text{V}$, en esta circunstancia la ganancia resultante es de 247.3.

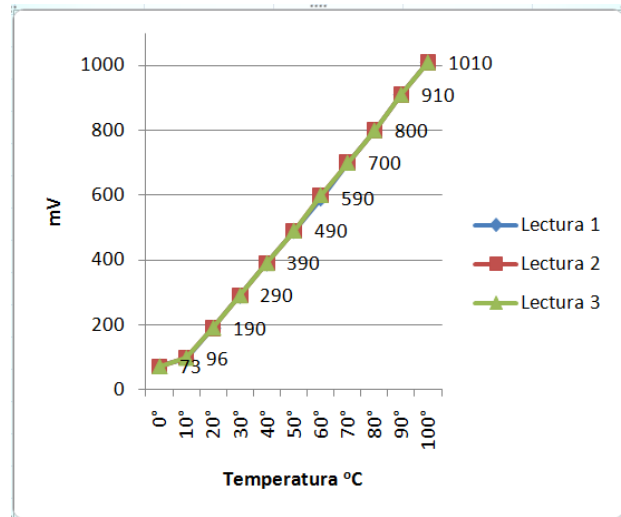
$$\text{Ganancia} = \frac{V_{out}}{V_{in}} = \frac{10 \frac{\text{mV}}{^\circ\text{C}}}{40.44 \frac{\mu\text{V}}{^\circ\text{C}}} = 247.3 \quad (4)$$

Además, un ajuste de precisión absoluta induce una compensación de entrada a la salida del amplificador característica de $11 \mu\text{V}$ para el AD594.

$$\text{salida AD595} = (\text{voltaje del termopar tipo K} + 11 \mu\text{V}) * 247.3 \quad (5)$$

Con la fórmula 5, si el rango de operación es de 0 a $100 \text{ }^\circ\text{C}$ la salida es de 0 a 1015 mV.

A continuación se muestran las medidas tomadas con un calibrador de proceso de la marca Fluke® modelo 744, el cual se utiliza para simular la señal del termopar tipo J y se miden al mismo tiempo la salida del AD594. Se muestran los resultados en la gráfica 1.



Gráfica 1. Salida del AD594.

Para amplificar las señales de los sensores se utiliza un amplificador instrumental monolítico AD620 de la marca Analog Device®. Este integrado hace una diferenciación de dos señales de referencia y amplifica la señal

resultante en un rango establecido por el circuito de 0 a 5 V.

Se seleccionó el amplificador AD620A, el cual se muestra en la figura 4.

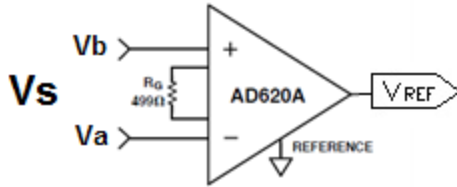


Fig.4 Amplificador instrumental AD620A.

$$R_G = \frac{49.4 \text{ k}\Omega}{G - 1} \quad (6)$$

Se calcula la ganancia considerando el voltaje de 680.19 mV para el sensor RTD y voltaje de salida de 5 V, obteniendo una ganancia de 7.35, con esta ganancia se determina la resistencia necesaria de 7.77 kΩ de la misma manera para el termopar considerando el voltaje de 1015 mV y voltaje de salida de 5 V, obteniendo una ganancia de 4.926, con esta ganancia se determina la resistencia necesaria de 12.58 kΩ con la ecuación 6.

2.4 Convertidor de V/I.

Para convertir la señal de corriente, se utiliza un convertidor de voltaje a corriente de la marca Texas Instruments® para la salida de 4 a 20 mA.

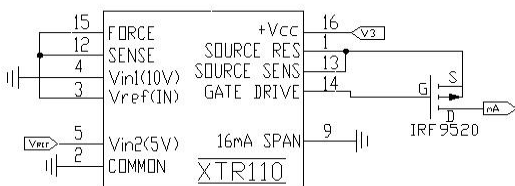


Fig.5 Configuración del XTR110.

La configuración del XTR110 permite configurar la entrada de 0 a 5 V y devuelve una salida estándar de 4 a 20 mA.

Como se muestra en la figura 5, la salida del XTR110 puede ser configurada y es lineal.

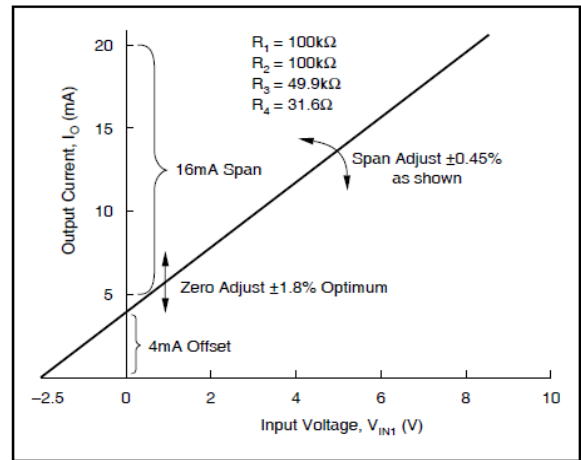


Fig. 6. Gráfica de V Vs I.

El IRF9520 es un MOSFET de canal P que se encarga de conducir la salida de corriente, determinada por la siguiente fórmula:

$$I_o = \frac{10 \left[\frac{V_{REF IN}}{16} + \frac{V_{IN1}}{4} + \frac{V_{IN2}}{2} \right]}{R_{SPAN}} \quad \dots\dots\dots(7)$$

Para un rango de entrada de 0 a 5 V y un rango de salida de 4 a 20 mA se debe conectar el dispositivo como se muestra en la figura 5.

2.5 Microcontrolador.

Con el microcontrolador se realiza la conversión de la señal analógica a digital, además de visualizar los valores de temperatura través de un LCD en la salida, y permite enviar los valores de la temperatura a través de una interfaz USB. El diagrama de conexión del microcontrolador se muestra en la figura 6.

El microcontrolador tiene un convertidor analógico digital de 10 bits con resolución de 4.88 mV con una referencia de 5 V.

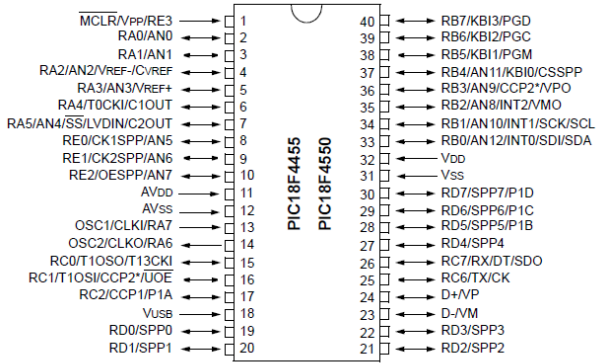


Fig. 7. Microcontrolador.

El valor de la temperatura con respecto al valor digital del voltaje de entrada está dado por la fórmula:

$$T = \frac{Vd * 150^\circ}{1023} \quad (8)$$

donde:

- T Temperatura mostrada en el display.
- Vd Valor de la señal analógica convertida a digital.

Para mostrar el valor de corriente en el LCD se utiliza la siguiente fórmula:

$$mA = \frac{(V)(16mA)}{5V} + 4mA \quad (9)$$

El microcontrolador cuenta con una interfaz para protocolo USB, lo cual permite, sin necesidad de contar con otro dispositivo, la conexión inmediata con el CPU.

El software utilizado para la programación del microcontrolador es el PCW HW Compiler 4.110.

El diagrama de flujo del programa del microcontrolador se muestra en la figura (8). El primer bloque de este diagrama consiste en leer el valor de voltaje del sensor a utilizar ya sea RTD o Termopar, una vez seleccionado se procede a convertir la señal de voltaje en un

valor digital el cual es de tipo entero. El tercer bloque consiste en calcular la temperatura y corriente utilizando las fórmulas (8) y (9), estos cálculos nos devuelven variables de tipo flotante. El cuarto bloque imprime el resultado de estas variables en el LCD. El quinto bloque envía el valor de la variable temperatura y el valor de la corriente a través de un protocolo USB.

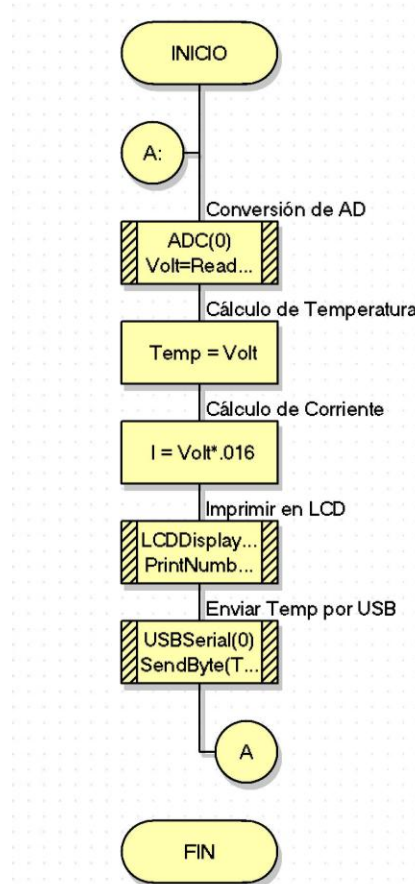


Fig. 8. Diagrama de flujo.

El módulo LCD es un sistema de visualización basado en un display de cristal líquido; los pines 33 al 38 del microcontrolador se conecta el display LCD tal y como se indica en (9).

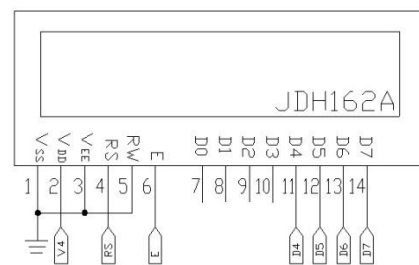


Figura 9. Conexión del LCD.

El programa en el microcontrolador envía la información por los puertos correspondientes a la transmisión y recepción que maneje los voltajes correspondientes al estándar USB. Esta información es monitoreada en LabVIEW® como se muestra en la figura 10.

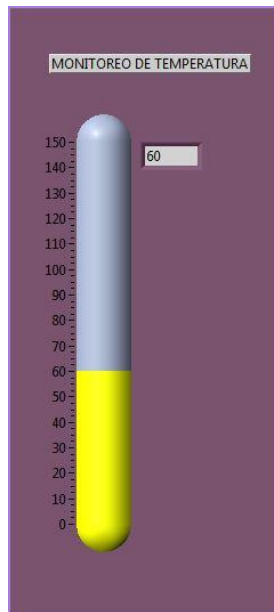


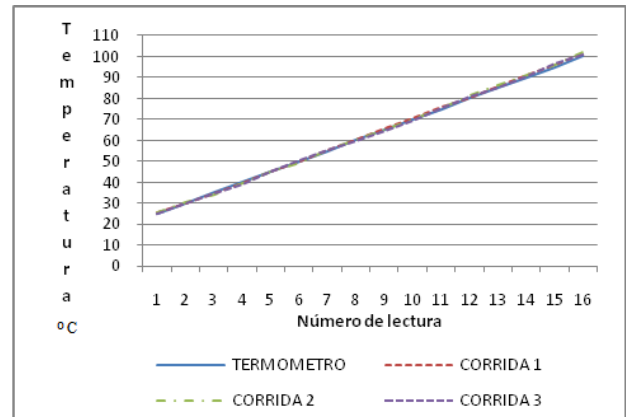
Figura 10. Pantalla de monitoreo en LabVIEW.

III. Resultados

Una vez construido el transmisor, se realizaron varias pruebas utilizando un calibrador de procesos de la marca Fluke® 744 para simular los sensores.

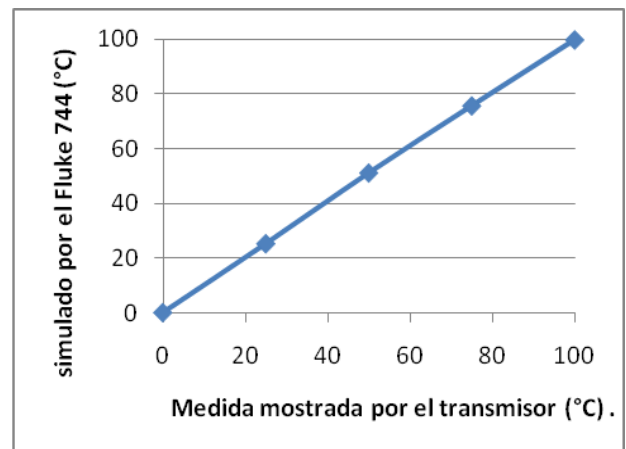
El dispositivo también se probó introduciendo los sensores en un recipiente con agua a diferentes temperaturas y comparando el valor entregado por el transmisor con el medido con un termómetro de mercurio, obteniendo los resultados mostrados en la gráfica siguiente.

La diferencia entre los valores mostrados en el termómetro y los obtenidos por el transmisor (ver gráfica 2) demuestran que el transmisor funciona adecuadamente al devolver valores muy aproximados o iguales a los leídos en el termómetro.



Gráfica 2. Comportamiento del transmisor con el sensor PT 100.

A continuación se muestra la salida generada por el calibrador de procesos para simular un termopar tipo J.



Gráfica 3. Comportamiento del transmisor con el termopar tipo J.

IV. Discusión

Para comprobar el funcionamiento del transmisor se han utilizado un RTD y un termopar. Entre las características del RTD encontramos: alta resistividad, lo cual otorga una mayor variación de resistencia por grado de temperatura (sensibilidad); relación casi lineal entre resistencia y temperatura; estabilidad de las características durante la vida útil del material; tamaño pequeño y con rapidez de respuesta.

El termopar es más usado en la industria pero presenta la desventaja de no ser tan exacto

como lo es un RTD ya que generalmente su medición varía un grado con respecto al valor real. Una de sus principales ventajas es su bajo costo en relación a los RTD.

V. Conclusiones

El TIT se desempeña correctamente en el rango de temperatura para el cual fue diseñado y cumple con las principales características de un transmisor de tipo industrial con la ventaja de presentar un bajo costo además de que el transmisor tiene una resolución en centésimas, es lineal y no presenta histéresis.

El hecho de contar con una interfaz con protocolo USB facilita la conexión con las PC y es posible su monitoreo desde cualquier software de programación que permita configurar los puertos COM seriales, ya que el diseño del transmisor está basado en un puerto serial virtual tal como se utilizaba con los puertos RS-232, pero con la ventaja de utilizar ahora USB.

Como trabajo futuro se propone el diseño de un transmisor que permita al usuario seleccionar un rango de temperatura de trabajo, a la vez que permita manejar otros estándares de transmisión digital tales como protocolo CAN, ProfiBus, RS-485.

VI. Referencias.

- [1] E. Garcia Breijo, "Compilador C CCS y simulador Proteus para microcontroladores PIC". Edit. Alfa-Omega, 1ª. Edic. México, D.F 2008.
- [2] Jan Axelson, "USB Complete". Cuarta Edición. Edit. Lake View Research LLC. E.E.U.U. 2009.
- [3] A. Creus Solé, "Instrumentación industrial", sexta edición, 1997, editorial Alfaomega-Marcobo, España, pagina 51.
- [4] J. M. Angulo Usategui, S. Romero Yesia, I. Angulo Martínez, "Microcontroladores

PIC diseño práctico de aplicaciones segunda parte", segunda edición, 2006, editorial Mc-Graw Hil, España.

- [5] R. Pallas Areny, "Sensores y Acondicionado-res de Señal", 3ª. edición, Edit. Alfa-Omega Marcombo, México, pág. 68.
- [6] AD595 - <http://pdf1.alldatasheet.com/datasheet-pdf/view/48077/AD/AD595.html>.
- [7] AD620 – Hoja de datos. <http://pdf1.alldatasheet.com/datasheet-pdf/view/48090/AD/AD620.html>. Febrero 2010.
- [8] XTR110 – Hoja de datos. <http://pdf1.alldatasheet.com/datasheet-pdf/view/56889/BURR-BROWN/XTR110.html>. Febrero 2010.
- [9] PIC16F87A. Hoja de datos. **¡Error! Referencia de hipervínculo no válida.**<http://pdf1.alldatasheet.com/datasheet-pdf/view/234440/MICROCHIP/PIC16F874.html>. Febrero 2010.
- [10] JHD162A – Hoja de datos. **¡Error! Referencia de hipervínculo no válida.**<http://pdf1.alldatasheet.com/datasheet-pdf/view/127934/ETC/JHD162A.html>. Febrero 2010.
- [11] Real academia española. Diccionario de la lengua española, XXII edición.

VII. Autores

M. en C. Manuel Vázquez Briones. Ingeniero en Electrónica y Comunicaciones egresado de la Facultad de Ingeniería de la Universidad

Veracruzana, pasante de Maestro en Ciencia en Ingeniería en Telecomunicaciones. Profesor de Tiempo Completo de la carrera de Mecatrónica de la UTSV.

M. en C. Javier Garrido Meléndez. Recibió el grado de Ingeniero Electrónico, en el Instituto Tecnológico de Minatitlán en 2003 y el grado de Maestro en Ciencias en Ingeniería Eléctrica con Especialidad en Control por el CINVESTAV. Profesor de Tiempo Completo de la carrera de Mecatrónica de la UTSV. Sus áreas de interés son el análisis de sistemas eléctricos industriales y modelado de equipos que integran electrónica de potencia.

TSU. Juan Carlos Hernández Cruz. TSU de la carrera de Electrónica y Automatización de la Universidad Tecnológica del Sureste de Veracruz.

Ing. Wendy V. Ventura Aragón. Ingeniera en Sistemas Computacionales egresada del Instituto Tecnológico de Minatitlán. Profesor de Tiempo Completo de la carrera de Mecatrónica de la UTSV.

Sistema Domótico de Red Inalámbrica para Monitoreo y Control del Gasto Electrodoméstico utilizando FPGA y LabVIEW.

* Anna Leonor Aguilar Torres,* Juan Sifuentes Mijares,* Oscar Hernández Reyes,
*Martín G. Vázquez Rueda.

*Departamento de Posgrado e Investigación del Instituto Tecnológico de la Laguna,
Revolución y Calzada Cuauhtémoc, Torreón, Coahuila, C.P. 27000, México,
Tel. 8717051324, juansm@faraday.itlalaguna.edu.mx

Resumen

En este artículo se presentan el diseño y los resultados referentes a la implementación en LabVIEW de un sistema para la medición del gasto eléctrico basado en una FPGA Spartan3E, el cual combina la domótica y las redes inalámbricas para el monitoreo del consumo electrodoméstico, y el cual está dividido de la siguiente forma: Módulo de Sensores analógicos y acondicionadores de señal, módulo de adquisición análogo – digital usando una FPGA, módulo transmisor y receptor inalámbrico XBee y módulo de monitoreo, cálculo y presentación de las mediciones de gasto electrodoméstico.

Palabras Clave: FPGA, LabVIEW, domótica, red inalámbrica, medición.

I. Introducción.

En los últimos años ha tomado mucha fuerza y popularidad la programación gráfica, siendo LabVIEW uno de sus principales precursores, que con el fin de seguir a la tecnología ha ido desarrollando módulos para tecnologías específicas, uno de estos es el módulo de programación para FPGA's (Field Programmable Gate Array).

LabVIEW FPGA utiliza el diseño gráfico para extender el poder de diseño a los sistemas embebidos y circuitos a la medida, tales como el FPGA. LabVIEW es claramente adecuado para programación del FPGA ya que se puede representar paralelismos y flujo de datos independientes. Con estas características se brinda al usuario la oportunidad de crear hardware personalizado a la medida, sin uso de los lenguajes de descripción de hardware y de bajo nivel o diseño a nivel de tarjeta. [1] [2]

Esta tecnología puede aplicarse para crear unidades de temporización y control de alta velocidad, construir interfaces con diferentes protocolos de comunicación, procesamiento

digital de señales (DSP), RF, comunicaciones y muchas otras aplicaciones.

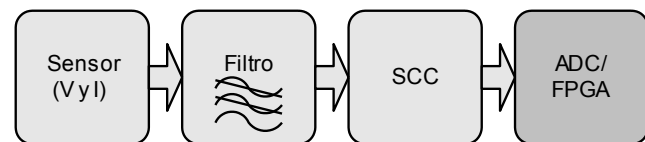


Fig. 1 Diagrama de la primera etapa del sistema.

Este trabajo está basado en el uso de LabVIEW y de la tarjeta de desarrollo SPARTAN 3E para la implementación en hardware de un sistema de medición de gasto electrodoméstico conectado en un red inalámbrica tipo Zigbee. El objetivo es detectar puntos críticos de consumo eléctrico.

II. Desarrollo.

1) Módulo de sensores de señales analógicas y acondicionamiento.

Este módulo realiza el muestreo de la señal de voltaje y corriente, así como también el escalamiento y filtrado. Adecuando la señal antes de poder ser digitalizada. Fig. 1.

Medición de Voltaje: Se utiliza un divisor de tensión por la facilidad y bajo costo de implantación, además de que también ocupa un espacio mucho menor al de un transformador de voltaje. Para el aislamiento eléctrico se utiliza un amplificador de instrumentación, Fig. 2. [3]

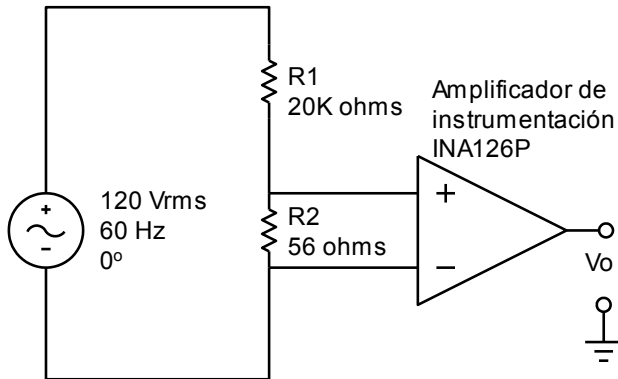


Fig. 2 Sensor de voltaje.

La resistencia R1 se selecciona mucho mayor que R2, de tal forma que la caída de voltaje en R2 sea aproximadamente del 1% del voltaje de de la línea, tomando en cuenta también la tolerancia de las resistencias.

La señal alterna de voltaje es medida en R2 y se aísla con el amplificador de instrumentación INA126P, el cual amplifica 5 veces la salida Vo.

Medición de Corriente: Se eligió un sensor de efecto Hall, dada su facilidad de uso. El sensor es el modelo HTP25 del TELCON (figura 3), que funcionando bajo el principio del efecto Hall y aunado a un amplificador operacional, otorga una salida casi lineal y muy fácil de manejar, al ser no invasivo para el circuito de prueba, no requiere aislamientos extras. [3] [4]

Calculo del Factor de potencia: El factor de potencia es el coseno del ángulo de desfase entre la señal voltaje que alimenta a una carga y la corriente que circula por esta, ver Fig. 4. [5]

El factor de potencia se obtiene del cruce por cero de las señales provenientes de los sen-

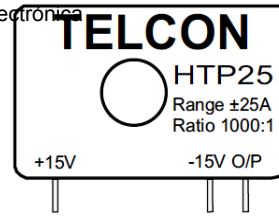


Fig. 3 Vista frontal del sensor de corriente.

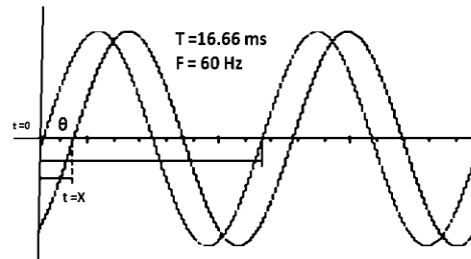


Fig.4 Representación del factor de potencia.

sores, este detector de cruce por cero se implementa haciendo uso de un comparador LM339, de esta manera se igualan los niveles de voltaje a LVTTTL que es el estándar que utiliza la tarjeta Spartan 3-E en la que se implemento el sistema de adquisición, véase figura 5.

Las salidas de los detectores de cruce por cero van hacia las entradas digitales de la tarjeta Spartan 3-E, en donde se obtiene la diferencia de tiempos entre el cruce de la señal de voltaje y la señal de corriente.

Con esta diferencia de tiempos y conociendo la frecuencia, que es contante, se puede obtener el ángulo de desfase, como se sabe bien para México la frecuencia de la señal es 60Hz.

Filtrado y acondicionamiento de la señal: Se utiliza un filtro pasa bajas de segundo orden con una frecuencia de corte de 60 Hz, cabe mencionar que esto reduce la señal al 70.7% de su valor, por lo que se debe considerar este efecto a la hora de calcular el valor final tanto de voltaje como de corriente.

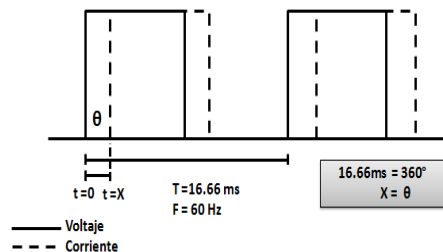


Fig. 5 Señales de salida de los detectores de cruce por cero.

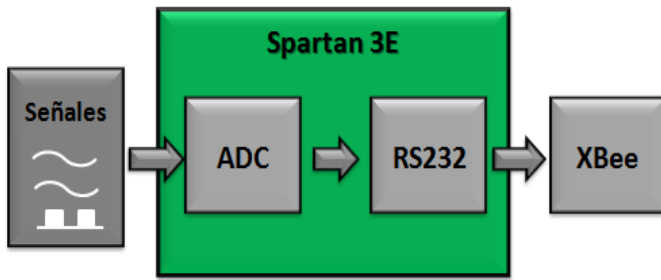


Fig. 6 Esquema general de la etapa de adquisición de datos.

Después del filtro, la señal pasa por un circuito acondicionador, de amplificadores operacionales, los cuales ajustan los niveles de voltaje para ser adquiridas por el ADC. [6] [7]

2) Modulo de adquisición análogo – digital usando una FPGA.

El primer paso para la adquisición de datos es la conversión de la señal analógica a digital. El convertidor utilizado es el incluido en la tarjeta Spartan 3-E. Una vez realizado el muestreo de voltaje, corriente y calculado el factor de potencia, se ponen los datos en el circuito transmisor construido en la FPGA y que implementa el protocolo RS-232 para el envío de datos en red ZigBee, usando un radio XBee, para las comunicaciones inalámbricas. Un esquema general se muestra en la figura 6.

Es importante mencionar que en el preamplificador, el muestreo del convertidor análogo a digital (ADC) y la implementación de la comunicación serial, todo esto fue realizado utilizando el software de programación gráfica LabVIEW con el módulo para FPGAs.

Conversión análoga a digital de señal: La señal de salida de los circuitos acondicionadores son las entradas del ADC. El circuito de captura analógica utiliza un preamplificador programable LTC6912-1, de *Linear Technology*, cuya salida está conectada al convertidor LTC1407A-1ADC del mismo fabricante, el cual cuenta con dos canales que se ajustan perfectamente a las señales de corriente y voltaje, ambos son programados y/o controlados por el FPGA. [8] El circuito de captura analógica convierte el voltaje en el canal A y B en una representación digital de 14 bits, tal como se expresa en la ecuación 1, dada por el fabricante. [9]

$$D = \text{Ganancia} \times \frac{(V_{in} - 1.65)}{1.25} \times 8192 \quad (1)$$

Donde:

D es el valor digital de 14 bits en representación decimal.

V_{in} es el voltaje de entrada.

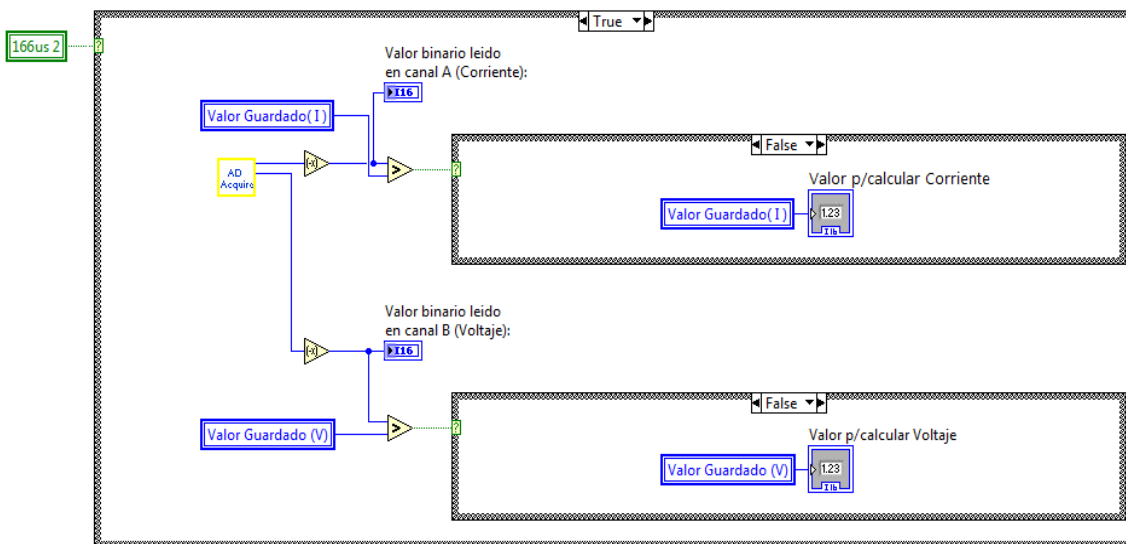


Fig. 7 Código de muestreo de las señales de voltaje y corriente.



Fig. 9 Esquema general de la transmisión y recepción inalámbrica.

4) Módulo de monitoreo, cálculo y presentación de las mediciones de gasto electrodoméstico.

Se desarrolló una aplicación en ambiente visual (Visual C#) para realizar la recepción de los datos provenientes de los módulos configurados como remotos, los cuales pasan a la PC vía USB. Se nombro a la aplicación “Sistema de Monitoreo de Gasto Electrodoméstico”, nos referiremos a esta aplicación por sus siglas SMGE.

El programa tiene la función de recibir los datos de las mediciones en forma serial, recordemos que estos datos son los valores digitales de las mediciones realizadas en la primera etapa del proyecto. A estos valores digitales de corriente y voltaje se les aplica la ecuación 1.

El cálculo del valor de factor de potencia se obtiene de multiplicar el dato recibido en número de cuentas por el tiempo de muestreo y transformando ese valor de tiempo a un ángulo que representa el $F.P.$, de la siguiente forma

$$\theta = \frac{(\text{Número de cuentas} * 166\mu s) * 360^\circ}{16.66ms} \quad (3)$$

$$F.P. = \cos \theta \quad (4)$$

Por ejemplo, si la diferencia de tiempo entre una señal y otra es de 10 cuentas obtendremos el factor de potencia como se muestra a continuación:

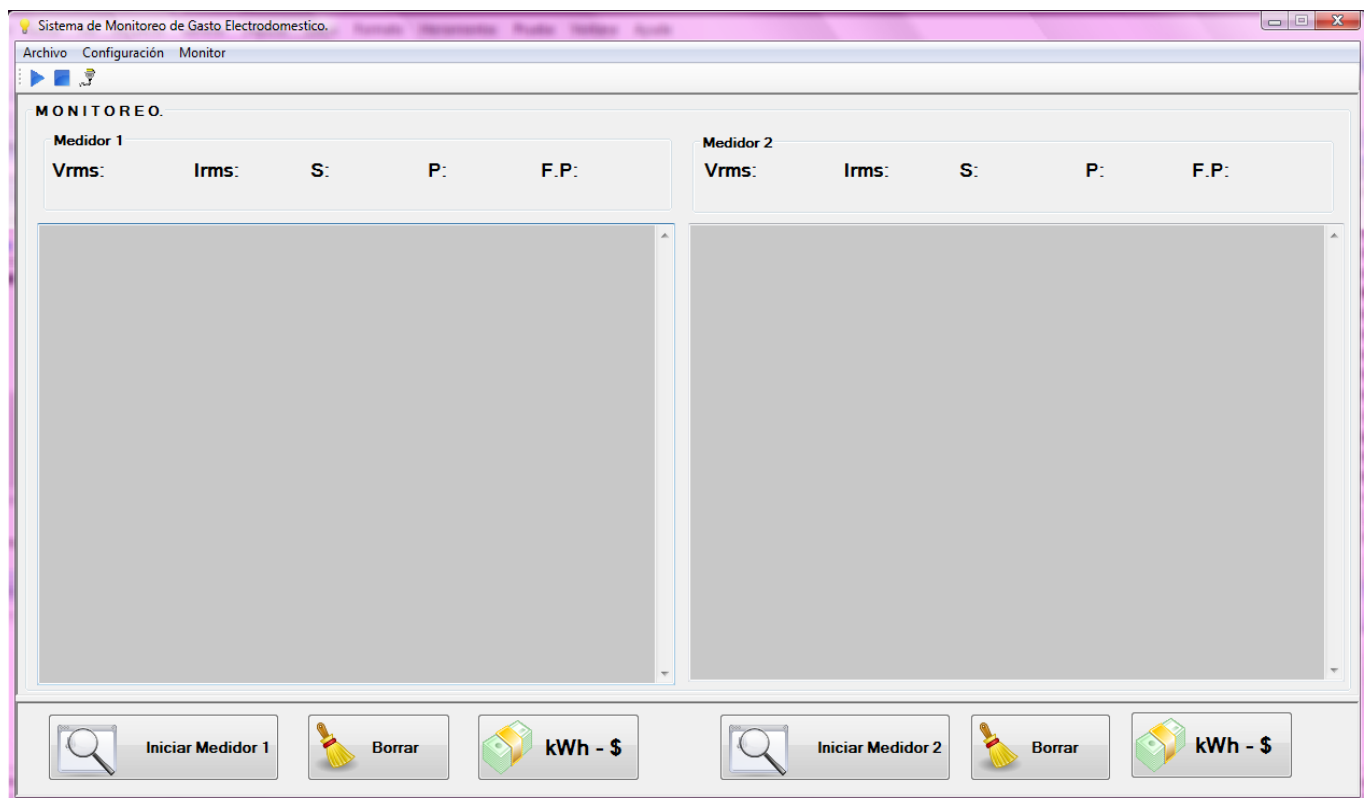


Fig. 10 Ventana principal del software de monitoreo de las mediciones SMGE.

$$\theta = \frac{(10 * 166 \mu s) * 360^\circ}{16.66 ms} = 35.87^\circ$$

$$F.P. = \cos 35.87^\circ = 0.8103$$

Una vez obtenidos los datos de valor de corriente, voltaje y factor de potencia se obtienen los valores de potencia aparente (S), potencia activa (P) y consumo económico de potencia (kWh). La interfaz de usuario diseñada, (figura 10), muestra el monitoreo de los datos de V_{rms} , I_{rms} , P, S, y F.P., al menos cada segundo.

El usuario puede borrar las mediciones, y si así lo desea, guardar los datos monitoreados o bien puede visualizar datos guardados anteriormente, detener el monitoreo y hacer cambio del aparato a medir, así mismo puede visualizar un estimado del gasto que tendrá en una hora y del gasto que se ha generado desde que se inicio el monitoreo de consumo del aparato, véase figura 11.

5) Red implementada.

Para tener una red con al menos dos puntos remotos y un coordinador se procede a implementar la red mostrada en la figura 12, cabe hacer notar que cada radio previo al envío de los valores de voltaje, corriente y factor de potencia, envía el número de medidor.



Fig. 11 Ventana donde se muestra el consumo de potencia (kWh)

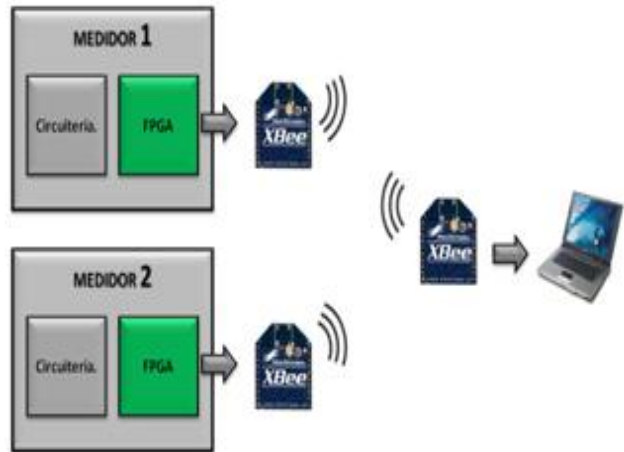


Fig. 12 Red inalámbrica Xbee.

III. Pruebas de funcionamiento.

Pruebas de Medición: Una de los detalles más importantes a verificar en el medidor es su funcionamiento para la medición de cualquier tipo de carga (resistiva, inductiva, y capacitiva).

En la tabla 1 se muestran las mediciones para un foco de 100W, que aunque prácticamente ninguna carga es puramente resistiva podemos considerarlo como tal, también se muestran las mediciones hechas en un transformador de 120/70Vac siendo este una carga inductiva, y para cubrir la carga capacitiva se toman mediciones a un monitor de PC que tiene el efecto de desfase la señal de corriente tal y como lo hiciera un capacitor debido a las fuentes de voltaje con las que cuenta dicho aparato, dichas mediciones se muestran en la tabla 1.

Así mismo se comprobó el funcionamiento de la red inalámbrica y del programa de recepción con los dos medidores a fin de verificar que no se pierdan o crucen datos.

Carga	V_{rms} V	I_{rms} A	S VA	F.P.	P W
Resistiva.	124.12	0.8	99	0.9688	96.2
Inductiva.	124.12	0.51	63.3	0.2533	16.03
Capacitiva.	124.12	0.51	63.3	0.6384	40.4

Tabla 1. Mediciones obtenidas en la fase de pruebas.

Pruebas de transmisión: Se implemento una red de dos puntos remotos y un coordinador. En este apartado se prueba el funcionamiento de cada

uno de los medidores en lo individual, así como el funcionamiento simultáneo de estos, en donde lo primordial es cuidar que los datos de un radio no perturben o detengan el envío del otro (figura 13), obteniendo pruebas favorables.

Alcance de la red: Se realizan simples pruebas de alcance de la red para determinar la distancia a la que seguirán funcionando los medidores. Cabe mencionar que el uso del sistema está orientado a interiores (sistema domótico). Los alcances de los radios XBee se muestran en la tabla 2, la cual nos indica un alcance de 30 metros para interiores y 100 metros para exteriores. Todas las pruebas se realizaron ubicando los medidores y el receptor en diferentes edificios del área de postgrado del Instituto Tecnológico de la Laguna. En la tabla tres se muestran los resultados de estas pruebas de alcance de la red.

Especificaciones	XBee
Rango en interiores	30 m
Rango en exteriores	100 m
Velocidad de datos (RF)	250,000 bps
Velocidad de datos (Serial)	1,200 bps - 115,200bps
Alimentación	2.8 - 3.4 V
Corriente de operación (transmitiendo)	0 dBm - 45 mA (3.3V,)
Corriente de operación (Recibiendo)	50 mA (3.3 V)
Banda de operación	ISM 2.4 GHz
Temperatura de operación	-40°C a 85°C (industrial)

Tabla 2. Alcance de los radios XBee.

Distancia	Calidad de Recepción.
25 m	Buena
28 m	Aceptable
43 m	Nula

Tabla 3. Resultados de la red implementada.

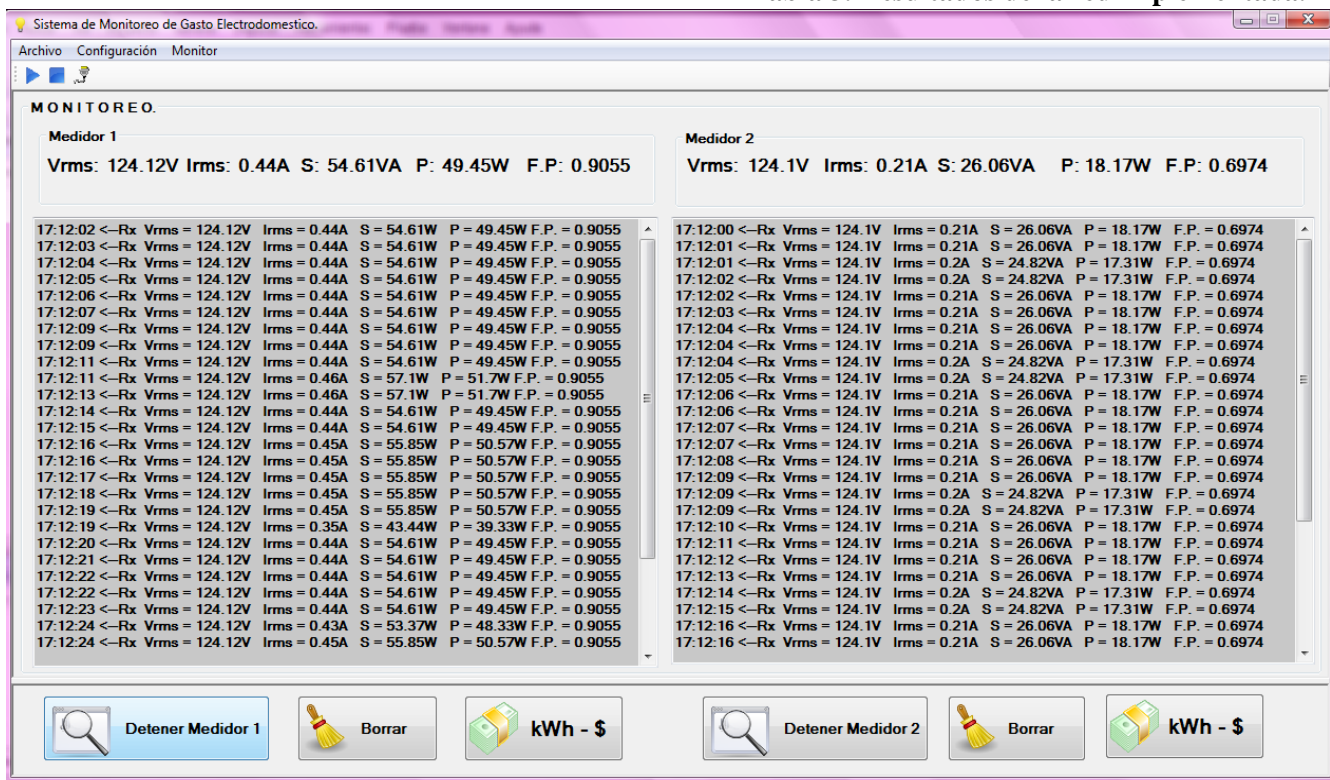


Fig.13 Recepción de los medidores simultáneamente.

<i>Aparato</i>	<i>Wattmetro Comercial</i>					<i>Sistema de Medición Implementado</i>				
	V_{rms} (V)	I_{rms} (A)	S (VA)	P (W)	F.P.	V_{rms} (V)	I_{rms} (A)	S (VA)	P (W)	F.P.
Foco 100W	124.3	0.79	99.81	90.82	0.9100	124.12	0.8	99.3	96.2	0.9688
Transformador	124.3	0.5	62.15	16.19	0.2605	124.12	0.51	63.3	16.03	0.2533
Monitor	124.3	0.5	62.15	41.9	0.6700	124.12	0.51	63.3	40.41	0.6384
Radio	124.3	0.029	3.6	1.51	0.4200	124.12	0.030	3.72	1.56	0.4186
Foco 25W	124.3	0.18	22.37	20.85	0.9323	124.12	0.20	24.82	24.08	0.9701
Ventilador	124.3	0.4	49.72	28.75	0.5801	124.12	0.42	52.13	30.00	0.5756
Secadora	124.3	12.32	1531	1493	0.9749	124.12	12	1489	1444	0.9698

Tabla 4. Resumen de resultados.

IV. Análisis de resultados.

En todas las pruebas realizadas de las mediciones se obtuvieron resultados positivos, tal como se observa en los valores de la tabla 4. Sin embargo se observa una pequeña desviación, respecto al medidor comercial, conforme aumenta la corriente. Se deben considerar las oscilaciones entre $120V_{ac}$ y $127V_{ac}$ que proporcionan la línea y también el acondicionamiento de la señal.

Se observa que los resultados obtenidos respecto al alcance de recepción son muy cercanos a los especificados por el fabricante, lo cual nos indica que mientras se mantenga el receptor a una distancia no mayor a 28 metros se seguirá captando las mediciones, como se observa en la tabla 3.[14]

V. Conclusiones.

La adquisición de las señales implementada con la tarjeta Spartan 3E y con programación gráfica son uno de los puntos más importantes de este trabajo, obviamente esta forma de programación trae consigo ventajas y desventajas. Dentro de las ventajas podemos recalcar que se tiene un ambiente más amigable e intuitivo, tal y como lo ofrece LabVIEW, también podemos mencionar la notable minoría de tiempo invertido en programación del FPGA. En la parte contraria la desventaja de la programación gráfica para FPGA's es el consumo, un tanto exagerado, de

recursos del mismo dispositivo así como los costos de las licencias del software.

Se podría mejorar en la construcción del hardware, la agrupación y distribución de los dispositivos utilizando elementos miniatura y esto podría fácilmente orientarse a ser comercializado, ya que realmente resulta útil, tanto para el ahorro de energía y dinero, el monitorear nuestra red eléctrica doméstica y así detectar los dispositivos que están teniendo un mayor consumo o un consumo desmedido, de la misma manera con la ayuda de la aplicación SMGE el usuario puede realizar los cálculos de gasto económico para tomar medidas preventivas así como crear una conciencia del ahorro de energía.

Cabe mencionar también que una foto del sistema completo no sería representativa del sistema que en realidad se ha desarrollado ya que el 97% del hardware es un ensamble de los módulos anteriormente mencionados, los cuales además contienen otras aplicaciones que no son utilizadas.

VI. Bibliografía.

- [1] National Instruments, "FPGA-Based Control: Millions of Transistors at Your Command", White paper, Junio 2010.
- [2] Martín Domínguez, Hugo. Sáez Vacas, Fernando. "Domótica: Un enfoque socio-

técnico”, E.T.S.I. de Telecomunicación, Universidad Politécnica de Madrid, primera edición, Junio 2006, ISBN: 84-7402-335-1.

[3] Cooper, William. Helfrick, Albert. “Instrumentación electrónica moderna y técnicas de medición”, Prentice Hall, 1991, ISBN: 968-880-236-0.

[4] Ward D.A Exon J. “Using Rogowski coils for transient current measurement”, Engineering Science and Educational Journal. 1993, pp 105-113.

[5] Pereira Rizzi, Álvaro. “Medidas eléctricas. Potencia-Energía-Factor de potencia-Demanda.” Editora de libros técnicos y científicos, Junio 1980, ISBN: 85-216-0089-5.

[6] Hwang, Enoch. “Digital logic and microprocessor design with VHDL”, Cengage editorial, primera edición 2006, ISBN: 053-446-59-35.

[7] Wakerly, Jonh F. “Digital Design, principles and practices”, Prentice Hall, cuarta edición 2005, ISBN: 0-13-186389-4.

[8] Manual del producto; “Spartan-3E Starter Kit Board User Guide UG230 (v1.0) March 9, 2006.

[9] National Instruments, “Using LabVIEW FPGA with the Xilinx SPARTAN-3E XUP” Video Seminario Web en vivo, disponible en www.zone.ni.com, marzo 2010.

[10] National Instruments, “Introducción al Módulo de LabVIEW FPGA”, Video Seminario Web en vivo, disponible en www.zone.ni.com, marzo 2010.

[11] Brown, Stephen. Rose, Jonathan. "Architecture of FPGAs and CPLDs: A Tutorial," IEEE Design and Test of Computers, Vol. 13, No. 2, 1996, pp. 42-57.

[12] Barr, Michael. "Programmable Logic: What's it to Ya?," Embedded Systems Programming, Junio 2006, pp. 75-84.

[13] Sun, Yichuang. “Wireless communication circuits and systems”, Institution of Engineering and Technology, 2004, ISBN: 085-296-443-9.

[14] Manual del producto; “Digi XBee™/XBee-PRO™ OEM RF Modules 802.15.4 –v1.xAx”, Minnetonka, Minnesota, EUA. Abril 2007.

VII. Autores.

Anna Aguilar obtuvo su título de Ingeniera en Electrónica en el año 2007, en el Instituto Tecnológico de la Laguna. Durante el año 2008 trabajó para la empresa SENSE Control Digital. Obtuvo su grado de Maestra en Ciencias de la Ingeniería Eléctrica en la Línea de Investigación en Instrumentación Electrónica en Marzo de 2011 en la División de Estudios de Posgrado e Investigación del Instituto Tecnológico de la Laguna, con el tema de Tesis: “Sistema domótico para control y monitoreo del gasto electrodoméstico”, actualmente trabaja en la empresa KONE Servicios Industriales como Ingeniero de Tecnología.

Juan Sifuentes obtuvo su grado de Maestro en Ciencias en Ingeniería Eléctrica en el año de 1995, en el Instituto Tecnológico de la Laguna. También estuvo realizando estudios de doctorado en diseño de sistemas digitales en la Universidad Politécnica de Valencia en España, actualmente se encuentra estudiando el Doctorado en Ciencias de la Ingeniería Eléctrica en el Instituto Tecnológico de la Laguna.

Oscar Hernández obtuvo su grado de Maestría en Ciencias en Ingeniería Eléctrica en la Línea de Investigación en Instrumentación Electrónica en Junio de 2010 en la División de Estudios de Posgrado e Investigación del Instituto Tecnológico de la Laguna, con el tema de Tesis: Diseño de módulos de propósito general para procesamiento de señales usando FPGA's.

Martin Vázquez obtuvo el grado de Maestría en Ciencias en Ingeniería Eléctrica, en el año 1989, en el Instituto Tecnológico de la Laguna, ha sido presidente de la academia de posgrado en el instituto tecnológico de la Laguna de 1995-1997, Jefe de laboratorio de propulsión ITLa Laguna 1998-2000, Jefe del departamento de Ingeniería Eléctrica y Electrónica del IT La Laguna 2001-2006, áreas de interés, Procesamiento digital de Imágenes, Instrumentación virtual, Inteligencia artificial e Instrumentación Médica.

Tarjeta de Adquisición de Datos para el Registro de Señales Mioeléctricas

*Celia Martínez-Chiñas, *S. Vergara, **M. Vargas, *F. Reyes, *M. A. Vargas, A. Palomino,
***J.M. Rodríguez-Lelis

*Facultad de Ciencias de la Electrónica, Grupo de robótica, Maestría en Ciencias de la Electrónica opción Automatización, Benemérita Universidad Autónoma de Puebla, Av. San Claudio y 18 Sur S/N C.U., Edificio 109A, C.P. 72570, Puebla Pue., México, Tel: (045) 2223582566, cely73@hotmail.com, svergara@ece.buap.mx, recf62@gmail.com, mavargas@ece.buap.mx, palomino@ece.buap.mx.

**Escuela de Ciencias, Universidad Autónoma Benito Juárez de Oaxaca, Ciudad Universitaria, Avenida Universidad s/n, Ex Hacienda “Cinco Señores”, Oaxaca de Juárez, Oaxaca, México, marcianoec@gmail.com.

***Centro Nacional de Investigación y Desarrollo Tecnológico Departamento de Ingeniería Mecánica - Tribología Interior Internado Palmira s/n, Col. Palmira, Apartado Postal 5-164 Cuernavaca, Morelos, México, jmlelis@cenidet.edu.mx.

Resumen: En la medicina existen diversos tipos de herramientas que son utilizadas como auxiliares en el tratamiento de ciertas patologías. En el presente trabajo se desarrolla una tarjeta PCI (del inglés Peripheral Component Interconnect) que funcione como un sistema de adquisición de datos para el registro de señales mioeléctricas y en un futuro evaluar la mejora de pacientes con discapacidad motriz bajo tratamiento de rehabilitación, la cual será de gran ayuda para los médicos del área ya que se proporcionará una herramienta que permita evaluar las diferentes técnicas y equipo de rehabilitación, con el fin de poder establecer cuáles de éstas son las más adecuadas en función del problema a tratar. Con lo cual se espera indirectamente contribuir a la mejora de la calidad de vida de los pacientes.

Palabras clave: electromiógrafo, señal mioeléctrica, sistema de adquisición de dato, técnicas de rehabilitación.

I. Introducción

En los últimos años la medicina ha experimentado grandes avances, tanto en el diagnóstico de enfermedades como en el tratamiento de las mismas. La electrónica es en parte responsable de ello, pues la medicina se sirve de numerosos aparatos electrónicos que permiten realizar múltiples tareas.

La tecnología avanza rápidamente y a cada momento sorprende con nuevas técnicas de diseño y medición. Un dispositivo de los más utilizados por sus características es la computadora, la cual es una herramienta que nos abre las puertas para implementar en ella muchas otras aplicaciones gracias a su capacidad de almacenamiento y gran velocidad en el procesamiento de datos.

El desarrollo de instrumentos controlados por computadora está permitiendo automatizar

procesos de caracterización y medición de diferentes sistemas usados en diversas áreas del conocimiento, reduciendo los tiempos en estas aplicaciones. El resultado de esta tendencia es la aparición de distintas interfaces de comunicación; un ejemplo de estas son las tarjetas PCI comerciales [1]. Actualmente en el país se han implementado diversas técnicas que son utilizadas para el tratamiento de rehabilitación de la discapacidad motriz, pero no se cuenta con un dispositivo que evalúe la mejora de los pacientes para determinar cuál de todas estas técnicas es la que brinda los mejores resultados. A pesar de que todas las personas con esta discapacidad deben tener las mismas oportunidades para acceder a estos tratamientos y evaluar su mejora, esto no es posible, ya que en el país no se desarrollan instrumentos para tratar este tipo de padecimientos; por lo tanto, tienen que ser importados, lo cual hace más difícil su adquisición.

En éste trabajo se desarrolla una tarjeta de adquisición de datos que cuenta con un bus de comunicación PCI. Esta se utilizará para la obtención, registro, procesamiento y análisis de señales mioeléctricas, con el fin de proporcionar al personal médico una herramienta de trabajo para evaluar la mejora de pacientes con discapacidad motriz bajo tratamiento de rehabilitación. En ésta tarjeta se implementa un electromiógrafo, que por su aplicación, es una herramienta con la cual se realiza el estudio de problemas a nivel muscular que están determinados por potenciales bioeléctricos. Este sistema consta de varias etapas, las cuales están en proceso de desarrollo. A continuación, se explican las etapas que hasta el momento se han abordado.

II. Electromiografía

La electromiografía (EMG) consiste en la adquisición, registro y análisis de la actividad eléctrica generada en nervios y músculos a través de la utilización de electrodos (superficiales, de aguja, implantados). Las mediciones extraídas proporcionan una información valiosa acerca de la fisiología y los patrones de activación muscular.

Estas señales son generadas por el intercambio de iones a través de las membranas de las fibras musculares debido a una contracción muscular. Lo que la gente considera normalmente un músculo (por ejemplo, el bíceps de los brazos) consta de miles de células individuales envueltas en tejido conectivo. Como las células musculares tienen una forma elongada, con frecuencia se llaman fibras. Las fibras musculares están dispuestas en haces llamados fascículos, que se encuentran envueltos en tejido conectivo [2, 3].

III. Características de la señal EMG

Las características importantes de la señal EMG son las siguientes:

- Espectro de importancia 0-500Hz.
- Energía dominante entre 50-150Hz.

- La adquisición de señales EMG se ve fuertemente afectada por el ruido de línea (60Hz).
- Amplitudes entre 0-6mV pico y 0-1.5mVrms.
- Los potenciales de acción de cada músculo presentan una magnitud fija.

Dicha información refleja las fuerzas que son generadas por los músculos y la temporización de los comandos motores. Además, se usa en el diagnóstico de patologías que afectan al Sistema Nervioso Periférico, las alteraciones funcionales de las raíces nerviosas, de los plexos y los troncos nerviosos periféricos, así como de patologías del músculo y de la unión neuromuscular. La amplitud de las señales EMG varía desde los microvolts (μV) hasta un bajo rango menor de 10 milivolts (mV).

IV. Modo de medición de la señal EMG

La contracción de fibras musculares genera actividad eléctrica que puede ser medida por electrodos fijados a la superficie de la piel próxima al grupo muscular de interés. Para la adquisición de las señales EMG se requiere de una correcta identificación de las regiones musculares comprometidas en la ejecución de los movimientos a clasificar. Debido a la elevada resistencia eléctrica natural de la piel, se recomienda la aplicación de un gel que mejore la conductividad además de lograrse una buena superficie de contacto y adherencia con los electrodos. A pesar de estas disposiciones, las señales recogidas serán demasiado débiles, por lo que es necesario un proceso previo de filtraje y amplificación antes de su análisis (Fig. 1).



Fig. 1. Etapa de preamplificación y filtrado.

En el registro de las señales EMG de superficie existen inconvenientes tales como el ruido ambiente, el ruido debido al movimiento de los electrodos y el crosstalk (definido como la interferencia producida por la actividad eléctrica de músculos próximos), entre otros [4, 5].

V. Diseño de un sistema de adquisición de datos basado en una tarjeta PCI.

El diagrama a bloques del sistema se muestra a continuación.

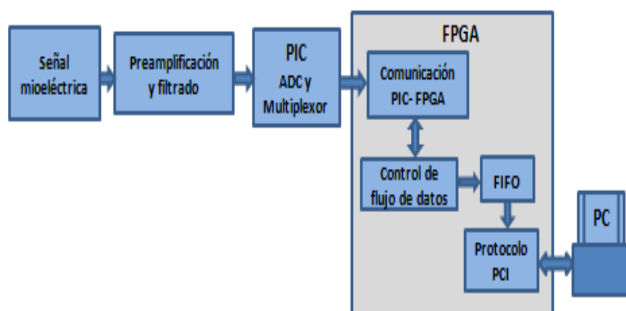


Fig.2. Diagrama a bloques general.

La primera parte del diagrama a bloques es la señal analógica de entrada, la cual se pre-amplifica y filtra debido a que la señal proveniente de los electrodos es muy débil. Para digitalizar la señal de entrada utilizamos el Convertidor Analógico Digital (ADC) con una resolución de 10 bits del microcontrolador 16F877A, es decir, convertir el nivel de la señal variable captada por los electrodos en una combinación de unos y ceros. El sistema consta de cuatro canales analógicos; si se utilizan los cuatro canales la señal es muestreada a 10 KHz, en caso de que se utilicen dos, esta puede ser muestreada a 20 KHz, con esto, nuestra señal es interpretada por el FPGA (del inglés Field Programmable Gate Array), el cual es un dispositivo semiconductor que contiene bloques de lógica cuya interconexión y funcionalidad se puede programar. En este se concentra el control del electromiógrafo y los filtros digitales. Esta etapa consta de varios módulos, en primer lugar

tenemos la comunicación entre el PIC y el FPGA seguido del control de flujo de datos; la memoria FIFO, en donde se guardan todos los datos que en un futuro reproducen la señal, y por último tenemos el protocolo, encargado de comunicar a la computadora con la tarjeta electrónica que opera como electromiógrafo.

VI. Diseño de las etapas de preamplificación y filtrado.

El punto importante del diseño de filtros es implementar un sistema con un bajo nivel ruido para la adquisición de la señal mioeléctrica.

La interferencia en la señal se origina principalmente por la línea de corriente alterna, por electrodos indebidamente fijados y otros factores, siendo la interferencia de la línea de corriente alterna la que destaca entre ellas, y por consecuencia, la eliminación de sus efectos es de suma importancia [6].

La siguiente imagen muestra el diagrama a bloques de la etapa de preamplificación y filtrado (Fig.3).

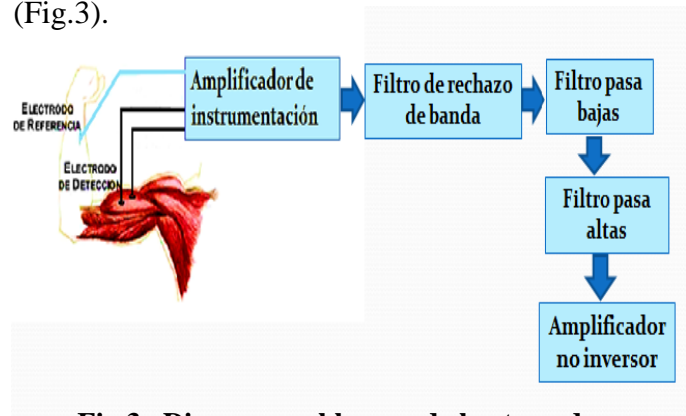


Fig.3. Diagrama a bloques de la etapa de preamplificación y filtrado.

a. Etapa de preamplificación

Para la etapa de preamplificación se utiliza el amplificador de instrumentación AD8221 con una ganancia de 500. El propósito de este es amplificar las débiles señales biológicas hasta obtener un nivel apropiado para que puedan ser registradas, debe admitir altas tensiones de DC de

modo diferencial a su entrada e importantes desbalances en las impedancias del generador de señal. Todo esto manteniendo buenas características; entre ellas, un alto rechazo de modo común (CMRR) y un bajo nivel de ruido [7].

La configuración que se utilizó la podemos observar en la siguiente figura:

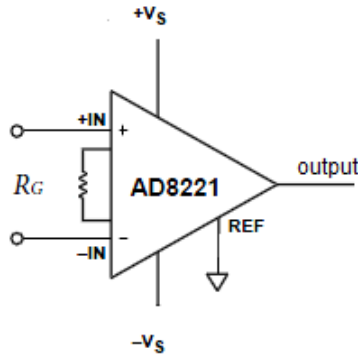


Fig.4. Amplificador de instrumentación AD8221.

b. Etapa de filtrado

Para la implementación de los filtros utilizamos el circuito integrado TL084, que contiene 4 amplificadores operacionales [8].

Los amplificadores operacionales son utilizados en diversas aplicaciones en la electrónica, desde cálculos matemáticos hasta como dispositivos que trabajan con señales de bajo nivel en medios ruidosos, como los que se observan en los organismos biológicos.

c. Filtro de rechazo de banda

El filtro implementado en este trabajo es un filtro de rechazo de banda de estructura VCVS (Voltage Controlled Voltage Source) con una frecuencia de corte de 60Hz y un ancho de banda de 6Hz (Fig. 5) para eliminar el ruido provocado por la línea de tensión. Se conectaron cuatro de estos filtros en cascada logrando una atenuación de -44 dB [9].

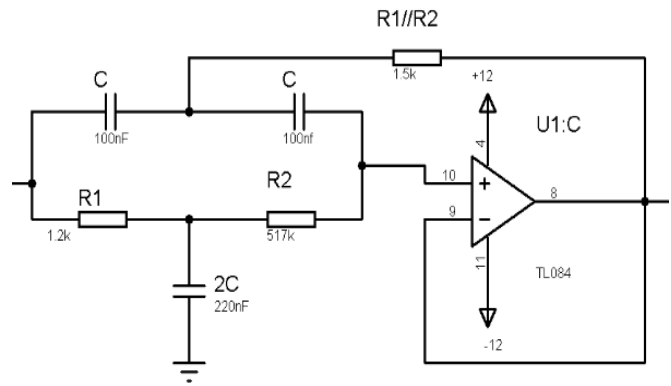


Fig.5. Filtro de rechazo de banda de estructura VCVS.

d. Filtro pasa bajas

Para limitar la frecuencia entre 0-500Hz utilizaremos un filtro pasa bajas de primer orden con una frecuencia de corte de 1kHz (Fig. 6).

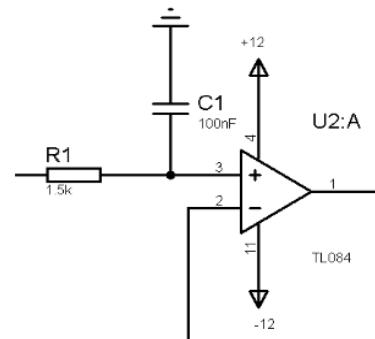


Fig.6. Filtro pasa bajas de primer orden

e. Filtro pasa altas

Para eliminar la componente en directa y el ruido por movimiento de los electrodos y cables, se implementó un filtro pasa altas de primer orden con una frecuencia de corte de 10Hz (Fig. 7).

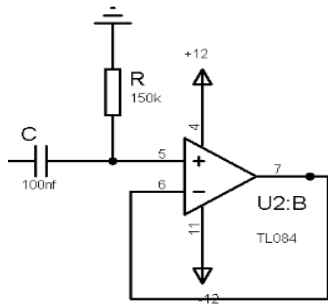


Fig.7. Filtro pasa altas de primer orden.

f. Amplificador no inversor

Para amplificar nuevamente la señal se utiliza un amplificador no inversor para obtener una ganancia de 11 (Fig. 8) [10].

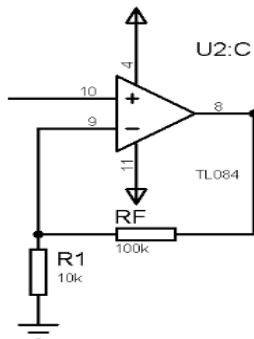


Fig.8. Amplificador no inversor.

VII. Diseño de la PCB

El diseño del PCB de la tarjeta electrónica consta de cuatro capas de cobre y tres de material dieléctrico. Las capas superior e inferior contienen todos los componentes de la tarjeta y el ruteo de pistas, la segunda capa es el plano de tierra y la tercera capa es el plano de voltaje. La tarjeta es modular, de tal manera que consta de una tarjeta hija que contiene solamente el FPGA; esto con la finalidad de no esclavizarlo a que realice únicamente esta función, sino que se pueda utilizar para otros fines. La tarjeta hija se muestra a continuación (Fig.9).



Fig.9. Modulo tarjeta Hija.

Los componentes a utilizar son de montaje superficial. En la etapa de entrada utilizaremos conectores LEMO para cuatro canales analógicos. En la etapa de conversión analógico/digital utilizaremos el microcontrolador 16F877A así como resistencias y capacitores con valores específicos para la aplicación. Luego tenemos el FPGA que será insertado en la tarjeta, por lo que consta de cuatro conectores. Por último tenemos el protocolo PCI encargado de la comunicación tarjeta-computadora. Este se encuentra también en el FPGA, pero se interconecta en la tarjeta.

El diseño de las tarjetas se realizó en el software Altium Designer 6.8, la cual se muestra a continuación.



Fig. 10. Tarjeta de adquisición de datos MIOCARD.

Se puede observar la colocación de vías alrededor de la etapa analógica para producir un escudo conectado a tierra y así disminuir la interferencia electromagnética producida por los circuitos

digitales. De igual forma se colocaron en puntos específicos de la tarjeta capacitores de 100nF, creando puntos de baja impedancia para las señales de corriente alterna.

VIII. Desarrollo del firmware y software

El firmware hace referencia a un bloque de instrucciones de programa para propósitos específicos, grabado en una memoria de tipo no volátil, que establece la lógica de más bajo nivel que controla los circuitos electrónicos de un dispositivo de cualquier tipo. El firmware es el intermediario (interfaz) entre las órdenes externas que recibe el dispositivo y su electrónica.

Para la programación del firmware se ocupa el software Quartus II versión 5.0 de Altera y el FPGA EP2C8Q208C7 de la familia Cyclone II.

El bus PCI se utiliza como vía de comunicación entre la computadora y la tarjeta de adquisición de datos, cuyo código se genera en Quartus II para la lectura y escritura de datos (Fig.11)

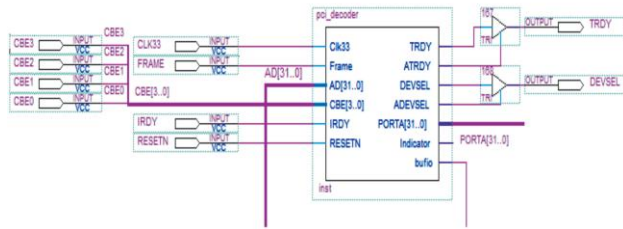


Fig. 11. Bloque del protocolo PCI.

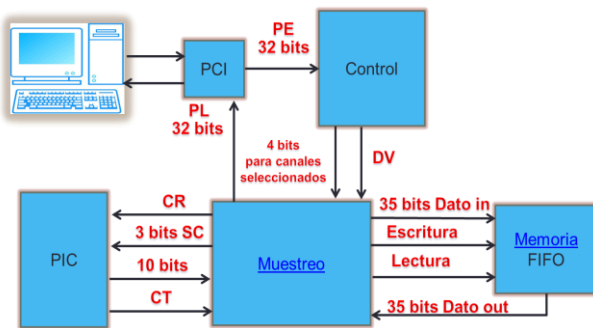


Fig. 12. Bloques del control de flujo de datos.

El bloque de control de flujo de datos se encarga de la selección de los canales que se desean utilizar y la frecuencia de muestreo de cada uno, en un rango de 1Hz-20KHz, así como el envío de datos a la memoria FIFO en donde se guardan para posteriormente reproducir la señal.

Para la visualización de la señal en la computadora se utilizó un software de instrumentación virtual.

IX. Resultados experimentales

a. Etapa de preamplificación y filtrado.

Los electrodos superficiales adheridos a la piel se colocaron, dos en el bíceps y uno neutro o de referencia en el tendón del sujeto de prueba aplicando gel conductor. Una vez montado el sistema se obtuvo la siguiente señal:

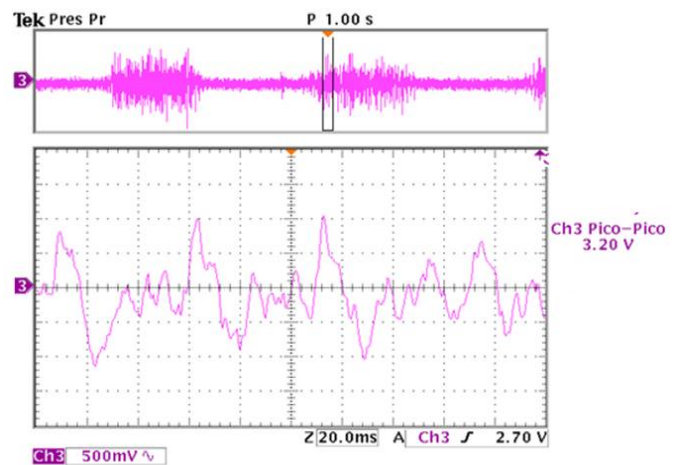


Fig.13. Señal mioeléctrica obtenida en las pruebas experimentales.

La señal fue tomada en un periodo de tiempo de 10 segundos. Para ver la forma de onda de la señal se realizó un acercamiento en el osciloscopio en el rango de 20 ms obteniendo la señal que se muestra en la figura 13.

En la figura 14 se observa el cambio de la señal cuando el músculo se encuentra contraído y en reposo y viceversa.



Fig. 14. Cambios en la señal durante la contracción muscular.

b. Firmware y software

En la siguiente figura se visualiza la señal mioeléctrica en la pantalla de la computadora, implementando el firmware y software desarrollados.

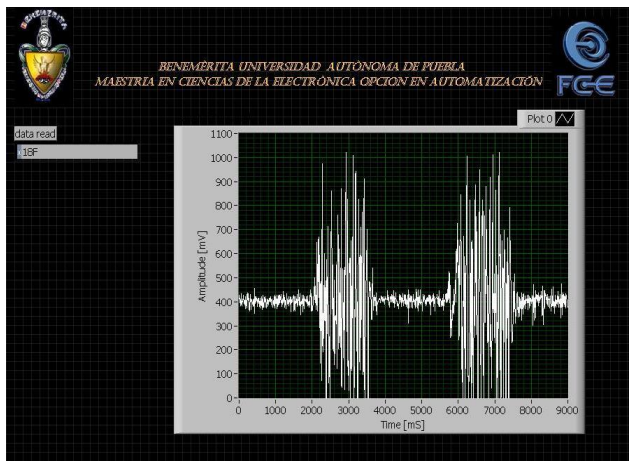


Fig.15. Panel frontal del sistema.

X. Conclusión

Es importante mencionar que el presente trabajo consta de varias etapas. La etapa de filtros digitales se encuentra en proceso de selección e implementación, continuando con el perfeccionamiento de la interfaz de usuario y la realización de pruebas experimentales con pacientes con discapacidad motriz bajo la supervisión de expertos en el área.

Considerando lo anterior, esta tarjeta de adquisición de datos será un sistema especializado en la medición y visualización de señales

mioeléctricas provenientes de la contracción muscular. Por lo tanto, es un dispositivo que capta información, la almacena en forma de datos en una memoria, la transforma adecuadamente y presenta en un elemento de visualización los resultados obtenidos; que para este caso es la pantalla de una computadora. Con el desarrollo de este tipo de instrumentos se piensa beneficiar a la comunidad científica y tecnológica, puesto que será de mucha utilidad en el campo de la medicina para evaluar la mejora de los pacientes en rehabilitación con discapacidad motriz.

La presente tarjeta de adquisición de datos fue diseñada de tal forma que puede ser utilizada para diferentes aplicaciones según las características que requiera el usuario para el sistema a implementar, tomando en cuenta que la tarjeta es modular y no se limita al FPGA.

XI. Referencias

- [1] R. J. Tocci “Sistemas Digitales” 6ª edición 1996, Pearson Educación, Pp690-691.
- [2] E. Mandado. “Instrumentación Electrónica”, 1995, Alfaomega marcombo. Pp. 41-96.
- [3] STUAR IRA FOX, “Fisiología humana”, séptima edición 2003, Mc Graw-Hill interamericana. Pp.150-370
- [4] Andrés Felipe Ruiz Olaya, “Sistema Robótico Multimodal para Análisis y Estudios en Biomecánica, Movimiento Humano y Control Neuromotor”, Tesis doctoral, Octubre de 2008. Leganés, Madrid, España.
- [5] Angélica Ramírez, Diego Alexander Garzón, “Análisis de sensibilidad por la colocación de los electrodos en la electromiografía de superficie”, Antioquia no.46 Oct./Dec. 2008, Medellín
- [6] Brown WF. “The physiological and technical basis of electromyography”. American Association of Electrodiagnostic Medicine, 198, Bstn: Butterworth-Heinemann,
- [7] Analog Devices, AD8221 ;EE.UU.2010 Disponible: <http://www.datasheetcatalog.net/es/datashee>

- ts_pdf/A/D/8/2/AD8221.shtml [Consulta 10 de Enero del 2011],
- [8] STmicroelectronics , TL084; 1999, Disponible:http://www.datasheetcatalog.net/es/datasheets_pdf/T/L/0/8/TL084.shtml [Consulta 10 de Enero del 2011].
- [9] Rolf Schauman, Mac E. Van Valkenburg, “*Desing of analog filters*”, 3ra. Edición Oxford university Press, 2001, New York.
- [10] L. Laura Álvarez, Ricardo Henao, Edison Duque. “*Análisis de esquemas de filtrado análogo para señales ECG*”. Scientia et Technica Año XIII, No 37, Diciembre de 2007.

Ciencias de la Electrónica de la BUAP, Miembro del SNI nivel Candidato.

José María Rodríguez Lelis, Doctor en Mecánica, Profesor Investigador, Centro Nacional de Investigación y Desarrollo Tecnológico Departamento de Ingeniería Mecánica - Tribología, Miembro del SNI nivel I.

XII. Autores

Celia del C. Martínez Chiñas, Ingeniera Electrónica egresada del Instituto Tecnológico de Tuxtla Gutiérrez en el 2008; actualmente es estudiante de la Maestría en Ciencias de la Electrónica opción Automatización en la Benemérita Universidad Autónoma de Puebla (BUAP).

Sergio Vergara Limón, Doctor en Ciencias con especialidad en Optoelectrónica, Profesor Investigador Titular A en la Facultad de Ciencias de la Electrónica de la BUAP, Miembro del SNI nivel I.

María Aurora Diozcora Vargas Treviño, Doctora en Ciencias con especialidad en Optoelectrónica, Profesor Investigador Titular A en la Facultad de Ciencias de la Electrónica de la BUAP, Miembro del SNI nivel I.

Fernando Reyes Cortés, Doctor en Electrónica, Profesor Investigador Titular C en la Facultad de Ciencias de la Electrónica de la BUAP, Miembro del SNI nivel I.

Marciano Vargas Treviño, Doctor en Física especialidad en Instrumentación, Profesor Investigador, Escuela de Ciencias, Universidad Autónoma Benito Juárez de Oaxaca, Miembro del SNI nivel I.

Amparo Palomino Merino, Doctor en Electrónica especialidad Control Automático, Profesor Investigador Titular A en la Facultad de

Construcción de un Sistema RFID para Aplicaciones en Control de Procesos

Miguel Magos Rivera, Ricardo Godínez Bravo, Ariana Carmona Velázquez, Aarón Hernández Torres

Universidad Autónoma Metropolitana-Azcapotzalco. Depto. de Electrónica
San Pablo 180 Col. Reynosa. 02200 México D.F.
mrm@correo.azc.uam.mx rgb@correo.azc.uam.mx

Resumen

En este artículo se presentan las etapas de construcción de un sistema electrónico que permite la lectura de tarjetas RFID. El sistema desarrollado se basa en el lector modelo ID-12 de la compañía *ID Innovations* y, vía un circuito electrónico, envía la información de la tarjeta a una computadora para la visualización de la información contenida en las tarjetas. Los resultados de la evaluación del sistema respecto a la distancia de detección son presentados en este documento.

Palabras Clave: Sistemas RFID, Identificación de Objetos, Control de Procesos, Sensores, Distancia de Detección.

I. Introducción

Empleados desde la segunda guerra mundial [1], los sistemas de identificación basados en radio frecuencia (RFID por sus siglas en inglés) han presentado un desarrollo importante en las últimas dos décadas debido, entre otros factores, a la miniaturización y a la reducción de los costos de fabricación de los elementos que lo conforman. Un sistema RFID permite identificar objetos por medio de la transmisión de señales de radiofrecuencia de un conjunto antena-microchip (tag), colocado sobre el objeto a identificar, a un equipo lector. El tag contiene información única, la cual es transmitida hacia el lector encargado de identificar el objeto, lo anterior a partir de una consulta en una base de datos [2], [3].

La información contenida en la memoria de los tags más comunes es grabada por el fabricante y no puede ser modificada por el usuario, además de no contar con fuente de energía propia. La alimentación necesaria para su funcionamiento es generada a partir de un campo electromagnético producido desde el lector. A estos se les denomina: tags pasivos de sólo lectura, figura 1. Existen otros tags que cuentan con fuente de energía integrada, así como con la opción de poder cambiar el código de la memoria; tanto el

tamaño como el costo de estos últimos es considerablemente superior a los tags pasivos de sólo lectura.

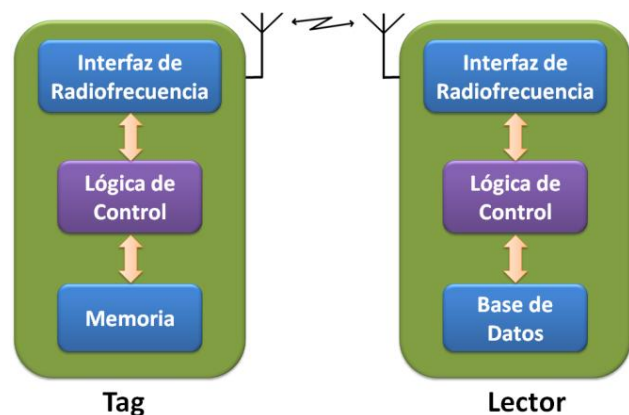


Fig. 1. Estructura de un tag pasivo de sólo lectura.

Como puede deducirse, los sistemas RFID permiten la identificación de objetos a distancia y sin contacto. La tecnología de identificación por radiofrecuencia presenta diversas ventajas con respecto a los sistemas lectores de códigos de barras ampliamente utilizados para la identificación de productos. Al emplear señales de radio para la transmisión de la información, los sistemas RFID no requieren de una línea de

vista como es el caso entre el escáner laser y la etiqueta impresa con el código de barras. Lo anterior lo convierte en una excelente solución en ambientes sucios o en los casos en que, debido a manipulación del producto, la impresión del código en la etiqueta no es clara. La tecnología de código de barras permite la identificación de un grupo de productos, por ejemplo, cajas de cereales de un peso y marca específicos, pero sin poder diferenciar entre elementos del mismo grupo. Por su parte, los sistemas de RFID tienen la ventaja de poder diferenciar un elemento de otro, lo cual le da un enorme valor agregado a esta tecnología [4].

La gran desventaja de los sistemas basados en radiofrecuencia radica en su costo; mientras que un código de barras puede ser impreso directamente sobre el objeto, los tags de los sistemas de RFID son dispositivos electrónicos los cuales, a pesar de haber reducido su costo de producción en la última década, implican una inversión de unas decenas de centavos de dólar. Lo anterior limita su aplicación actual a procesos de un valor económico elevado en el cual la inversión del sistema RFID pueda justificarse.

Existen diversas aplicaciones de la tecnología RFID a nivel comercial e industrial entre las cuales podemos mencionar: sistemas de identificación de personas [5], [6], control de acceso a instalaciones [7], autenticación de productos [8], [9], sistemas contra-robo de automóviles [10], cadenas de distribución de productos [11], [12], pago de peaje en autopistas [13], manejo de equipaje en aeropuertos [14], [15] y control de préstamos e inventarios en bibliotecas [16], [17], por mencionar sólo algunas.

La aplicación de esta tecnología en tareas de automatización de procesos industriales ha crecido durante la última década gracias a que los fabricantes de sistemas RFID han desarrollado componentes con características apropiadas a esta área. Uno de los sistemas de producción en el cual esta tecnología ha encontrado múltiples aplicaciones es en el de ensamblado de productos, tales como: motores, automóviles,

autobuses, motocicletas, etc. [18]. En estos procesos se alimenta con piezas al sistema encargado de ensamblar el producto, dependiendo del modelo a fabricar, las piezas pueden variar. Las líneas de ensamblado tradicionales contemplan la fabricación de un mismo producto durante un intervalo de tiempo, esto reduce costos asociados a modificaciones en los equipos y por lo tanto paros de producción, pero a cambio se tienen sistemas poco flexibles y altos costos en la fabricación de modelos con poca demanda. La posibilidad de identificar cada elemento dentro de un grupo de productos en una línea de ensamblado le proporciona a esta una gran versatilidad. En [12] se describe una aplicación implementada en la planta de Ford en Cuautitlán, en la cual se fija un tag a cada carrocería y mediante una serie de receptores ubicados en distintos puntos de la línea de ensamblado es posible determinar las operaciones específicas que deben realizarse sobre cada unidad.

Los sistemas descritos parten de la utilización de equipos RFID comerciales disponibles en el mercado, entre los cuales se pueden mencionar el modelo XR450 de la compañía Motorola [19] o el ALR9900 de la compañía Alien [20], por mencionar sólo algunos. Estos equipos están diseñados para aplicaciones en las cuales la detección puede alcanzar varios metros, cuentan con una gran variedad de puertos de comunicación así como posibilidad de realizar acciones de control directamente. La antena del lector generalmente se encuentra separada del equipo y están diseñados para una amplia gama de aplicaciones.

Respecto al desarrollo de sistemas, en [21] se presenta el diseño y construcción de un equipo RFID con alcance de 1 m. Este equipo fue concebido para la localización, mediante triangulación, de objetos y personas al interior de edificios, aplicación en la cual los equipos no se encuentran en condiciones agresivas.

El trabajo que aquí se presenta, describe la construcción de un sistema de identificación por radiofrecuencia enfocado a aplicaciones de

control e instrumentación de procesos. El sistema diseñado tiene como objetivo la identificación sin contacto de objetos en un rango de distancia pequeño, unos cuantos centímetros. El sistema construido se basa en el lector modelo ID-12 de la compañía *ID Innovations*. Se trata de un lector de tarjetas RFID con frecuencia de operación de 125 kHz y de alcance máximo de 12 cm.

En la primera sección de este artículo se describe brevemente la tecnología RFID y se mencionan diversas aplicaciones; asimismo, se presentan algunas diferencias con respecto a otros sistemas de identificación como son los lectores de códigos de barras. En la segunda sección del documento se plantean y describen los grandes bloques que conforman el sistema construido. Posteriormente se presentan los resultados de la evaluación del equipo respecto a la distancia de detección. En la última sección del trabajo, se presentan las conclusiones del mismo y se plantea una serie de trabajos a futuro.

II. Desarrollo

La base de este desarrollo es el dispositivo modelo ID-12 de la compañía ID Innovations, se trata de un módulo lector de tarjetas RFID con antena integrada, figura 2. Este elemento realiza la lectura del tag que se encuentra dentro de su espacio de detección y proporciona la información de la misma vía serial para que un sistema de procesamiento de información realice la identificación del objeto sobre el cual fue colocada.

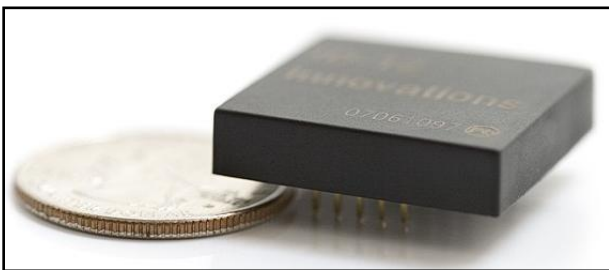


Fig. 2. Módulo lector empleado.

La frecuencia de operación del lector es de 125 kHz y es compatible con tarjetas de 64 bits (Estándar EM4001). El rango de lectura nominal, señalado por el fabricante, es de 12 cm y la comunicación con el sistema de procesamiento de información es mediante niveles TTL serial a 9600 bps [22]. El procesamiento de la información lo realiza un sistema digital basado en un microcontrolador, mismo que finalmente envía la información a cualquier otro dispositivo de despliegue, una computadora en este caso.

Los elementos que conforman el sistema pueden ser agrupados en tres grandes bloques mismos que se muestran en el diagrama de la figura 3 y que son descritos en el resto de esta sección.



Fig. 3. Diagrama de bloques del sistema construido.

Lector RFID ID-12

Este elemento es el encargado de decodificar la información que envían las tarjetas de identificación (Tag's) y transmitirla hacia el sistema digital para ser procesada. El lector ID-12 está configurado para transmitir la información de las tarjetas con formato ASCII, 9600 bps, sin paridad, 8 bits de datos, 1 bit de parada. Cuando el módulo lector detecta una tarjeta, envía una cadena de datos hacia el sistema digital bajo el formato que se muestra a continuación.

02		1 Byte
10	ASCII Hex Data Characters	10 Bytes
2	ASCII Char's Checksum	2 Bytes
CR		1 Byte
LF		1 Byte
03		1 Byte

Como puede observarse, la longitud total de la cadena es de 16 bytes. Los datos y los códigos de control se envían al circuito electrónico en formato ASCII hexadecimal (2 caracteres ASCII por byte).

Circuito Electrónico

Tiene como función principal establecer la comunicación con el lector RFID, recibir los códigos de las tarjetas de identificación (TAG's) y controlar el envío de información hacia una PC vía alguno de los puertos RS232 ó USB.

El circuito electrónico puede ser dividido para su explicación en cuatro grandes bloques: Microcontrolador, Circuito Oscilador, Convertidor TTL a RS-232/USB y Programa.

La figura 4 muestra los bloques que componen el circuito electrónico.

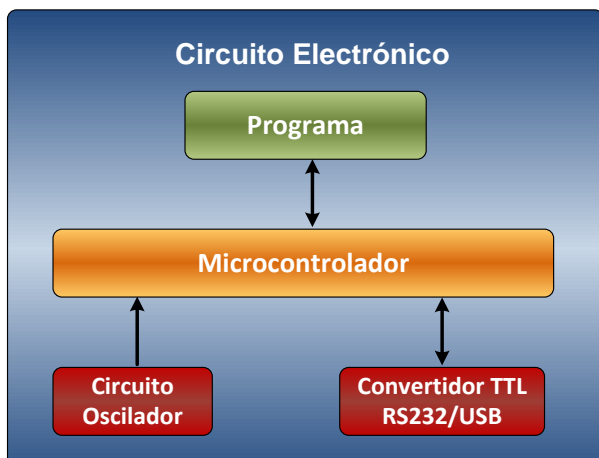


Fig. 4. Diagrama de bloques del circuito electrónico desarrollado.

Microcontrolador. La base del circuito electrónico es el microcontrolador AT89C2051 de ATMEL. Este dispositivo contiene internamente una CPU de 8 bits, 2 puertos de entrada y salida paralelo los cuales pueden ser direccionados por bit. El circuito genera señales de control para la comunicación serial (UART), posee 2 entradas para Timer/Contador de 16 bits, terminales para detectar interrupciones externas y un comparador analógico de precisión.

Circuito Oscilador. Este bloque considera a los elementos asociados en la generación de la señal de reloj necesaria para la operación del sistema digital. El circuito oscilador basa su funcionamiento en un cristal de cuarzo con frecuencia de oscilación de 11.059 MHz.

Convertidor TTL a RS-232/USB. La comunicación con el PLC se realiza de modo serial, para esto se agregaron circuitos convertidores para los estándares RS-232 y USB. Cabe mencionar que la alimentación del sistema se realiza vía el puerto USB, por lo que en caso de emplearse el puerto serial RS-232, es necesario también conectar el cable USB. La conversión de los niveles lógicos TTL proporcionados por los puertos del microcontrolador a los exigidos por la norma RS-232, se implementó mediante el dispositivo ICL232. Este elemento dispone de dos canales para la conversión de igual número de vías de comunicación. En esta aplicación sólo se empleó uno de ellos. Por su parte, el protocolo USB es implementado mediante el dispositivo FT232BL, el cual se encarga de acondicionar los niveles TTL del microcontrolador al estándar de comunicación USB.

La posibilidad de poder emplear cualquiera de estos dos tipos de puertos de comunicación, amplía la capacidad de comunicar al sistema con distintos tipos de equipos, tales como PLCs, controladores industriales y equipos de computo, dándole versatilidad en su uso.

Programa. Por último tenemos el programa elaborado para el microcontrolador. Este se encarga de coordinar la recepción de la información de las tarjetas de identificación enviada por el lector ID-12. Esta información es procesada para obtener los códigos grabados y enviarlos mediante comunicación RS232 y USB hacia una PC.

El diagrama de flujo del programa para el microcontrolador, se muestra en la figura 5. Como puede observarse el sistema se encuentra en un ciclo infinito de lectura de datos provenientes del lector de tarjetas y de envío de información a los puertos de comunicación.



Fig. 5. Diagrama de flujo del programa principal del microcontrolador.

En los diagramas de flujo de la figura 6, se muestran las secuencias de operaciones de los dos bloques importantes del programa: lectura de datos del módulo lector de tarjetas y transmisión de información a la PC.

El elemento lector se encapsulo en resina y se encuentra separado del gabinete que contiene al sistema electrónico. El intercambio de información se realiza mediante un cable de 1.5m de longitud el cual se conecta al sistema vía un conector DB-9, la figura 7 muestra una imagen del dispositivo.

Las imágenes que se presentan en la figura 8, muestran algunas vistas del sistema construido. Se observan los conectores para la comunicación con la PC en un extremo del gabinete, mientras que en el extremo opuesto está la conexión del módulo lector.

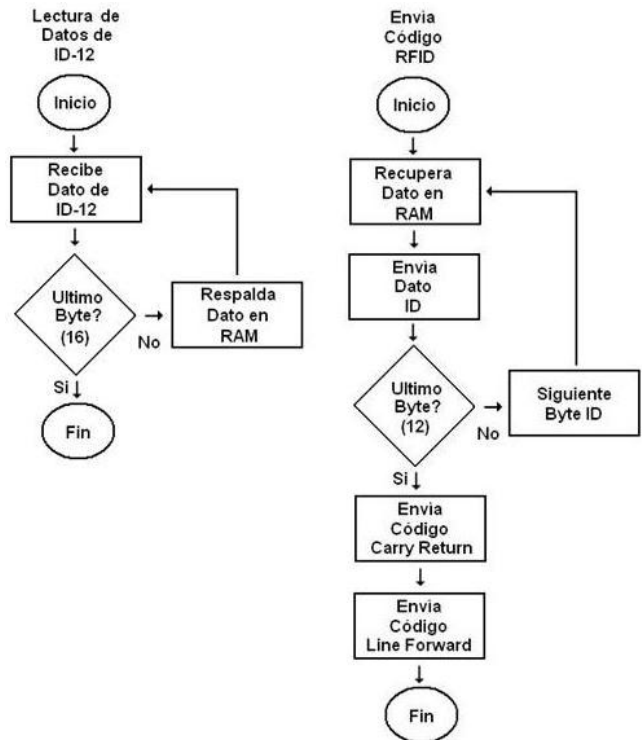


Fig. 6. Diagramas de flujo de las subrutinas de lectura de tarjetas y envío de información.

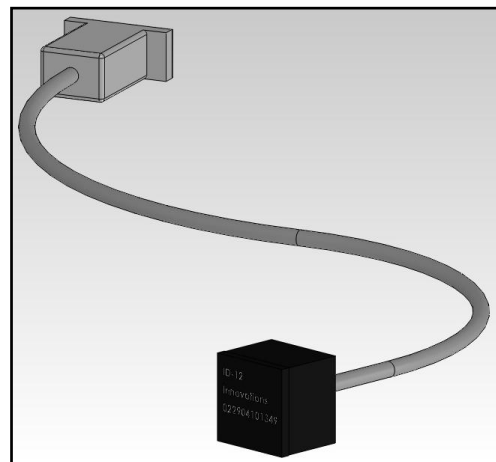


Fig. 7. Lector RFID con cable de conexión.



Fig. 8. Vistas del sistema construido.

Programa de visualización.

El programa elaborado para la computadora, es un pequeño interfaz desarrollado en Visual Basic que permite visualizar en pantalla la información contenida en el tag leído. El programa básicamente recibe mediante comunicación serial desde el circuito electrónico la información de las tarjetas de identificación. Esta información es

desplegada sin ningún tratamiento en la pantalla de la PC, el diagrama de flujo de la figura 9, muestra la secuencia de operaciones que realiza el programa.

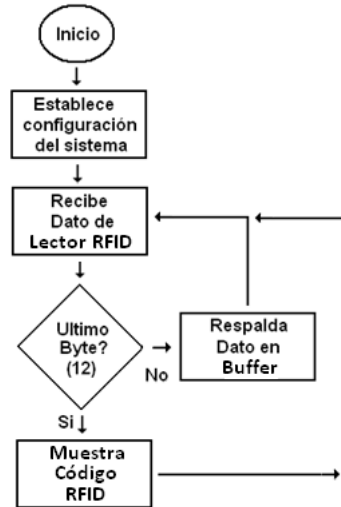


Fig. 9. Diagrama de flujo del programa en Visual Basic.

La figura 10 muestra los estados de la ventana en la cual se despliega la información.



Fig. 10. Interfaz de visualización.

III. Evaluación del funcionamiento

Una vez construido el sistema se realizaron pruebas simples de operación. Básicamente se verificó el funcionamiento del equipo respecto a distancias de detección sin obstáculos empleando tarjetas RFID tipo tarjeta de crédito.

La figura 11 muestra el dispositivo construido para realizar las pruebas. Se trata de un sistema que permite desplazar linealmente, tanto el lector como el tag para determinar las distancias de detección en dos dimensiones.

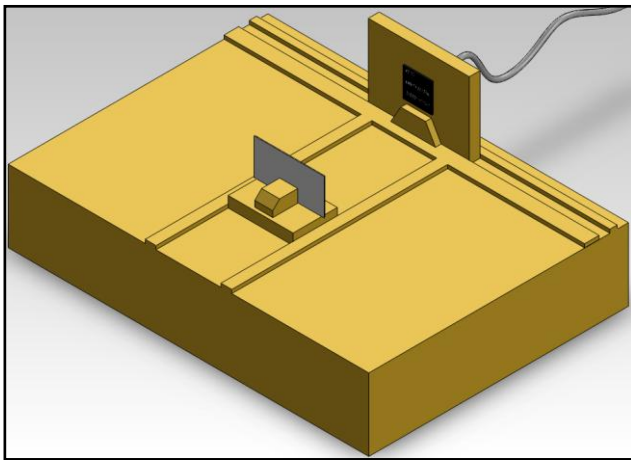


Fig. 11. Equipo para evaluación del sistema.

La figura 12, muestra el comportamiento del sistema cuando no existe ningún obstáculo entre el tag y el detector. El experimento fue realizado 20 veces y la gráfica muestra los valores promedio. La distancia promedio máximo de detección fue de 11.9 cm.

IV. Conclusiones

En este trabajo se presentó el desarrollo así como una evaluación simple del funcionamiento de un sistema de identificación por radio frecuencia para aplicaciones en control e instrumentación de procesos. Podemos observar de los datos que el alcance es similar al señalado por el fabricante, 11.9 cm contra 12.0 cm. Los resultados obtenidos permiten determinar los alcances del sistema, sus posibles aplicaciones y restricciones.

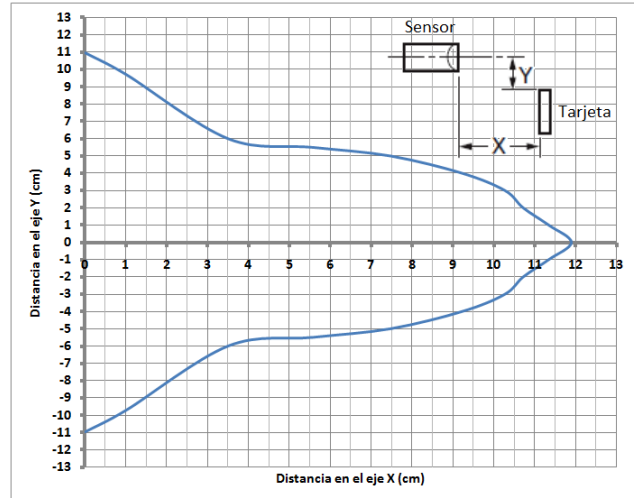


Fig. 12. Área de detección sin obstáculos.

La utilización de este tipo de sistemas en aplicaciones de control e instrumentación de procesos depende fuertemente de su capacidad a trabajar en condiciones agresivas. Presencia de polvo, humedad elevada y/o temperaturas elevadas, son algunas de las condiciones normales de trabajo de los sistemas de detección en la industria. Por lo anterior, una de las tareas a futuro está orientada a la evaluación del sistema construido bajo condiciones de operación más agresivas. Respecto a aplicaciones futuras del sistema, se plantea su implementación para identificar y/o ubicar objetos que se desplazan sobre trayectorias definidas, mismas que pueden ir desde bandas transportadoras hasta trenes.

Agradecimientos

Este trabajo se desarrolló en el marco del proyecto *Diseño y Construcción de un Prototipo de un Nuevo Sistema de Puertas para el Metro de la Ciudad de México con Dispositivo Inhibidor de la Apertura de las Mismas del Lado Contrario*, apoyado por CONACYT-Gobierno del Distrito Federal con clave 120694.

V. Referencias

- [1] W. Buesser, V. Fuentes, P. Goguillot, N. Nandra, T. Roz, J. Rudin, "A contactless read write transponder using low power

- EEProm techniques”. Proceedings of the 22nd European Solid-State Circuits Conference. Neuchâtel, Switzerland. September 1996.
- [2] D. Kou, K. Zhao, Y. Tao y W. Kou, “RFID Technology and applications”. Enabling Technologies for Wireless E-Business. pp. 486-489. 2006.
- [3] R. Want, “An introduction to RFID technology”. Pervasive computing. Vol.5. No.1. 2006. pp. 25-33.
- [4] “Auto ID in the material handling industry”. Accu-Sort systems. 2007.
- [5] V. Lockton, R. Rosenberg, “RFID: The next serious threat to privacy”. Ethics and Information Technology. Vol. 7. No. 4. December 2005. pp. 221-231.
- [6] J. Eagleson, “Security applications of RFID”. Proceedings of the Fifth Annual Wireless Symposium. Santa Clara, U.S.A. February 1997.
- [7] M. Kaur, M. Sandhu, N. Mohan, P. Sandhu, “RFID Technology principles, advantages, limitations & its applications”. International Journal of Computer and Electrical Engineering, Vol.3. No.1. February 2011. pp. 151-157.
- [8] T. Staake, F. Thiesse, E. Fleisch, “Extending the EPC Network – The Potential of RFID in anti-counterfeiting”. Proceedings of the 2005 ACM Symposium on Applied Computing. Santa Fe, U.S.A. March 2005.
- [9] Z. Nocht, T. Staake, E. Fleisch, “Product specific security features based on RFID technology”. International Symposium on Applications and the Internet Workshops. Phoenix, U.S.A. January 2006.
- [10] M. Ollivier, “RFID- A new solution technology for security problems”. European Convention on Security and Detection. Brighton. U.K. May 1995.
- [11] R. Angeles, “RFID technologies supply-chain applications and implementation issues”. Information Systems Management. Vol.22. No.1. 2005. pp. 51-65.
- [12] G. Gaukler, R. Seifert, “Applications of RFID in supply-chains”. Trends in supply chain design and management. pp. 29-48. Springer 2007.
- [13] B. Nath, F. Reynolds, R. Want, “RFID technology and applications”. Pervasive Computing. Vol.5. No.1. 2006. pp. 22-24.
- [14] Motorola, “Baggage Tracking RFID Solutions”. Industry Brief. 2008.
- [15] AeroAssist, “RFID in aviation: airport luggage control”. White paper. June 2008.
- [16] C. Kern, “Radio-frequency-identification for security and media circulation in libraries”. The electronic library. pp. 317-324. Emerald Group 2004.
- [17] K. Coyle, “Management of RFID in libraries”. The Journal of Academic Librarianship. Vol. 31. No. 5. 2005. pp. 486-489.
- [18] M. Baudin, A. Rao, “RFID applications in manufacturing”. http://www.mmt-inst.com/RFID_applications_in_manufacturing_Draft_.pdf. April 2005.
- [19] XR Series RFID Readers. Integrator Guide. Motorola Inc. U.S.A. December. 2008.
- [20] ALR-9900 Hardware Setup Guide. Alien Technology. U.S.A. September. 2008.
- [21] J. Hightower, G. Borriello, R. Want, “SpotON: An indoor 3D location sensing technology based on RF signal strength”. UW CSE Technical Report. U.S.A. February. 2000.
- [22] "EM module series V21". ID Innovations. March, 2005.

VI. Autores

Dr. Miguel Magos Rivera es Ingeniero en Electrónica por la Universidad Autónoma Metropolitana, Maestro y Doctor en Control Automático por la Universidad Claude Bernard de Lyon (Francia). Actualmente es profesor-investigador en la Universidad Autónoma Metropolitana unidad Azcapotzalco.

Ing. Ricardo Godínez Bravo es Ingeniero en Electrónica por la Universidad Autónoma Metropolitana. Actualmente se encuentra cursando la Maestría en Mecatrónica en el Instituto Tecnológico de Ecatepec además de ser profesor-investigador en la Universidad Autónoma Metropolitana unidad Azcapotzalco.

Ing. Ariana Carmona Velázquez es Ingeniero Industrial por la Universidad Autónoma Metropolitana unidad Azcapotzalco.

Aarón Hernández Torres es estudiante de 11^{avo} trimestre de la Licenciatura en Ingeniería Electrónica en la Universidad Autónoma Metropolitana unidad Azcapotzalco.

Mejora del Desempeño de un Equipo de Corte a través del Análisis de Vibración

* **María Magdalena Hernández Ramos**, ** **Francisco Javier López Jáquez**

* Departamento de Ingeniería Industrial y Manufactura, Instituto de Ingeniería y Tecnología
Universidad Autónoma de Ciudad Juárez
Av. Del Charro 450 Nte. Col. San Lorenzo, Ciudad Juárez Chihuahua, México
Tel. 656-6884843, magdys_hr@hotmail.com.

** Departamento de Ingeniería Industrial y Manufactura, Instituto de Ingeniería y Tecnología
Universidad Autónoma de Ciudad Juárez
Av. Del Charro 450 Nte. Col. San Lorenzo, Ciudad Juárez Chihuahua, México
Tel. 656-6884843, frlopez@uacj.mx

Resumen

Este artículo trata sobre los resultados obtenidos durante el análisis de vibración a través de estadística aplicada empleando pruebas de hipótesis y análisis de diseño de experimentos (análisis de varianzas), a un equipo de corte de placa metálica por medio de plasma. Se discuten diferentes áreas de conocimiento permitiendo establecer el mejor método a emplear y así lograr reducir la vibración del equipo a estándares establecidos de acuerdo a las normas ISO 2372 y VDI 2056. Por medio de la aplicación de aislante entre componentes del equipo y rediseños efectuados, mismos que se manufacturaron empleando prototipos rápidos. La utilización de estos prototipos permite tener un diseño exacto de los componentes en poco tiempo por muy difíciles que sean de manufacturar.

Palabras Clave: CAD/CAM, DOE, Norma ISO, Prototipo rápidos, Vibración.

I. Introducción

Hoy en día muchos equipos de manufactura existentes experimentan o producen vibraciones. Por ejemplo, cualquier desequilibrio en las máquinas que poseen partes giratorias, como ventiladores, separadores centrífugos, lavadoras, tornos, bombas centrífugas, prensas rotatorias y turbinas, ocasionan vibraciones. Para estas máquinas, las vibraciones son indeseables [1].

La referencia de vibración en su forma más sencilla, una vibración se puede considerar como la oscilación o el movimiento repetitivo de un objeto alrededor de una posición de equilibrio. La posición de equilibrio es a la que llegará cuando la fuerza que actúa sobre él sea cero. Este tipo de vibración se llama vibración de cuerpo entero, lo que quiere decir que todas las partes del cuerpo se mueven juntas en la misma dirección en cualquier momento [1].

Uno de los parámetros importantes empleados en los estudios de vibración es la frecuencia la cual utiliza los CPM (ciclos por segundo) o HZ (hercios). La mayoría de las máquinas y estructuras experimentan vibración hasta cierto grado y, su diseño, requiere generalmente consideraciones de su conducta oscilatoria [2].

Es por ello que existen estándares o normas que permiten ciertos niveles de vibraciones en las máquinas de acuerdo a su estructura, lo que es parte fundamental para llevar a cabo una realización de un estudio detallado de esto, permitiendo llevar a cabo la reducción de vibración.

Actualmente existen investigaciones sobre el tema de vibración, aplicado a sistemas de monitoreo por medio de transformadas wavelet y Fourier empleando técnicas de procesamiento de señales. Las transformadas wavelet se emplean

como una herramienta en un sistema de monitoreo de vibración para detectar con precisión y localizar las fallas en este sistema. Dentro de las transformadas wavelet existe la WPT como mejor herramienta para análisis de señales de vibración y detección de fallas en máquinas rotativas. La vibración en la investigación se ha orientado hacia el campo de equipos de diagnóstico en máquinas rotativas [3, 4].

La investigación de sistemas expertos de vibración ha permitido la verificación de las técnicas de análisis de vibraciones habituales en la industria. Empleando algoritmos de detección de falla robusta y así lograr identificarlas [5].

El uso de análisis de vibración en control estadístico de calidad se considera un problema común de calidad en algunos procesos de fabricación. Empleando para este un método de análisis ANOVA multifactorial realizado con variables controlables y sus efectos sobre las variables influyentes obteniendo como resultado que la vibración no se presenta como una dependencia directa como el desplazamiento de la variación de los parámetros del proceso por tal razón no debe ser elegida como elemento característico del proceso. Para conocer las diferentes características de las variables se emplea el control de estadístico de calidad por medio del análisis de vibraciones [6].

Considerando lo anterior, se tomará como base para encontrar una solución al problema que a continuación se describe el cual se detalla más ampliamente para su mejor comprensión.

II. Desarrollo

Prueba preliminar para el análisis

Se realiza una prueba preliminar para conocer la situación actual de la máquina donde se obtienen los primeros resultados de la vibración existente en la misma tomando varios puntos de referencia.

De los cuales los más críticos para tomar en consideración son los que se muestran en la Fig. 1 y Fig. 2.



Fig. 1- Tornillo eje Y.

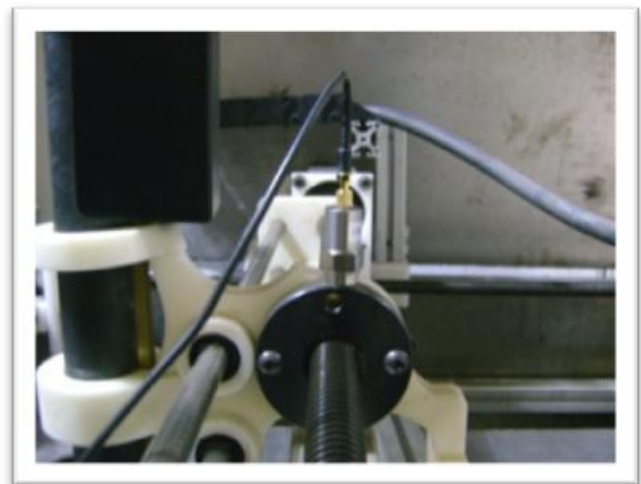


Fig. 2- Tornillo eje X.

Estos puntos muestran una vibración máxima de 4.2mm/s para la Fig. 1 Tornillo eje Y y 4.1mm/s para la Fig. 2 Tornillo eje X. Se tiene que el punto de la Fig. 1 será evaluado más a detalle en un segundo análisis de vibración. También se determinó el diseño de corte que se empleará en el siguiente análisis en este análisis se realizan solo líneas.

Análisis de vibración

Los parámetros a emplear para este estudio determinados en la sección anterior son:

- Clasificación de la maquinaria, la cual se ubica en el grupo K de maquinarias pequeñas hasta 15KW, ya que cuenta con motores de este rango [7, 8].
- Diseño de corte, un círculo. Ya que en este intervienen los dos ejes al momento de estar realizando el proceso y se tiene una mejor lectura sobre la vibración.
- Diámetro, 150mm.
- Datos, 30.
- Placa metálica de 420mm por 920mm.

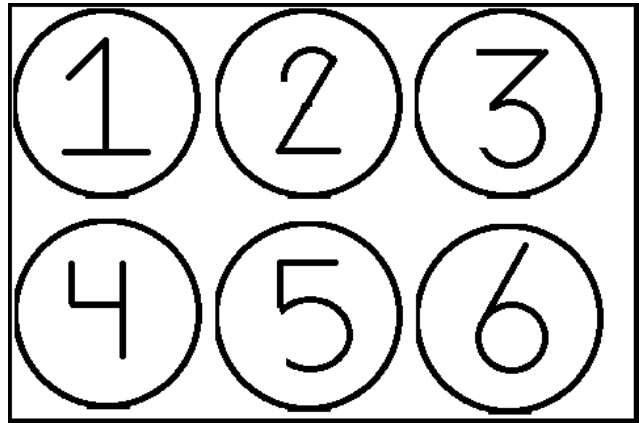


Fig. 3- Posición en la placa.

Teniendo la referencia del punto y los parámetros, se realiza un estudio basado en diseño de experimentos, el cual es de tipo de mejoramiento donde se utilizan menos factores que el del tipo exploratorio, esto con el propósito de identificar mejor la relación entre los factores críticos y la respuesta. A continuación se listan los pasos para realizar un diseño de experimentos [9].

1. Comprensión y planteamiento del problema.
2. Elección de factores y niveles.
3. Selección de las variables de respuestas.
4. Elección del diseño experimental.
5. Realización del experimento.
6. Análisis de datos.
7. Conclusiones y recomendaciones.

De acuerdo a los pasos antes mencionados los cuales se toman como base para el análisis de la situación se tiene que el diseño de experimentos a realizar es el método Anova unilateral. El método es un análisis de varianzas que compara las medias.

Prueba preliminar para posiciones

En la Fig. 3 se realizan posiciones en la placa metálica con el objetivo de llegar a una hipótesis nula, se plantea en este caso como que el nivel de vibración en cada una de las posiciones es igual.

Para este estudio se toman los siguientes datos que se muestran en la Tabla 1 en las unidades de mm/s.

Vibración			
Posición	Nivel Vib.	Posición	Nivel Vib.
6	5.2	1	8.7
6	4.8	1	7.3
6	6.3	1	112.9
6	7.1	1	6.5
6	93.2	1	6.1
2	4.6	5	5.1
2	12.9	5	6.2
2	7.5	5	6.8
2	6.5	5	5.4
2	6.9	5	7.3
3	5.0	4	6.5
3	5.5	4	5.8
3	4.9	4	12.8
3	16.0	4	7.2
3	1.0	4	6.1

Tabla 1- Grupo de datos.

De los datos anteriores se tienen algunos resaltados de los cuales se emitirán en el estudio ya que estos son muy elevados para el análisis y la posible causa de esta lectura no necesariamente tiene relación con el proceso del corte si no a otro factor externo que se relaciona con el cable del instrumento.

Prueba preliminar nivel de vibración

Otra de las hipótesis planteadas para determinar el nivel de vibración existente en el equipo de corte, en resumen la hipótesis se plantea de la siguiente forma, la media del análisis es igual o menor a 4.5mm/s valor que se encuentra dentro del rigor de vibración permitido de acuerdo a la norma ISO 2372 y VDI 2056 [7, 8].

Para el análisis de una muestra z es necesario tener, una distribución normal, desviación estándar conocida, el número de muestras sea menor de 30.

Aislamiento de la vibración

El término aislamiento de la vibración hace referencia a la inserción de un elemento relativamente blando o elástico. Esto entre dos componentes que están conectados entre sí, donde uno de ellos vibra y el otro se protege de esta vibración [1].

En la Fig. 4 se puede observar las partes donde se colocó aislante. En la Fig. 4- a) y d) son las partes donde está el motor conectado con otros componentes, principales objetivos para poner el corcho ya que el motor es considerado una fuente de vibración junto con su apoyo. En la Fig. 4- b) y f) existe el tornillo de conexión con los ejes por ello se determina poner corcho entre estos dos elementos ya que a la Fig. 4- b) es el punto de análisis de medición de vibración. Para la Fig. 4- c), e) fue necesario aplicar aislante, en estos casos como función de soporte para equilibrar el eje ya que al motor del eje X fue necesario aplicar aislante entre el perfil en el que se encuentra soportado así como en la Fig. 4- d) entre la pieza de soporte motor y el perfil.

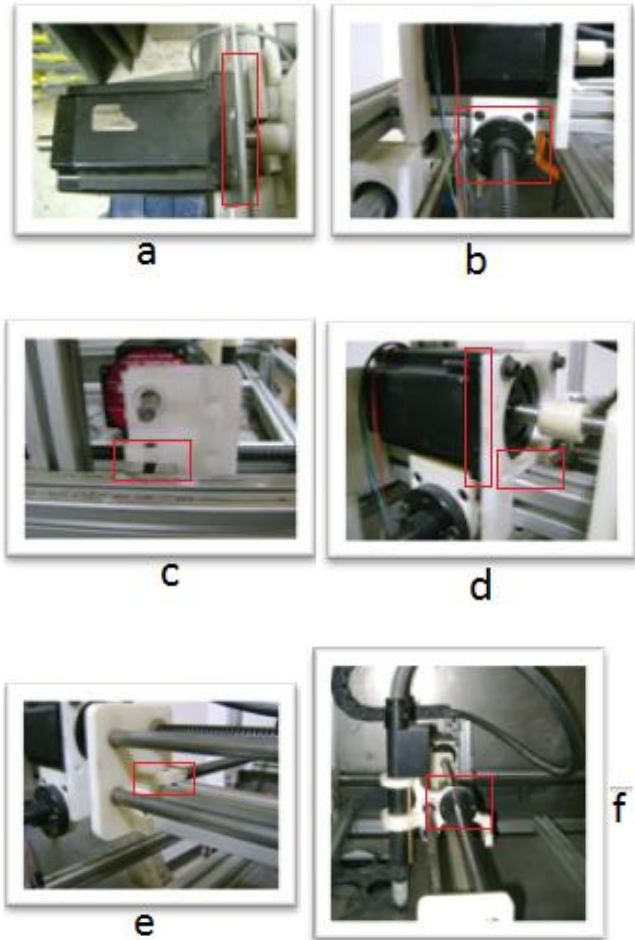


Fig. 4- a) motor eje Y b) tornillo eje Y c) soporte 1 eje X d) soporte motor e) soporte 2 eje X f) soporte 3 eje X g) tornillo eje X.

Otra de las mejoras realizadas en esta etapa es el alineamiento de los ejes y poner el torque, como fue establecido en un proyecto anterior, a los tornillos M6 un torque de 110 libras por pulgada de acuerdo a su material en este caso acero y a los tornillos M8 un torque de 120 libras por pulgada [10]. Que han sido modificados de muchas de las partes de la máquina razón por la cual el objetivo es regresarlo a ese torque ya que fue un determinante para la reducción de la vibración en el análisis de ese proyecto. Se lleva a cabo una nueva prueba de hipótesis con base en un solo factor. Para este análisis solamente se toman 573 datos del vibrómetro.

Con los datos anteriores se procede a realizar el estudio de prueba de hipótesis de un factor donde la hipótesis nula se plantea de la siguiente forma:

Hipótesis nula; la media de la muestra es igual a 6.2 mm/s.

Hipótesis alternativa; la media de la muestra es menor que 6.2 mm/s.

El siguiente procedimiento, efectuada la prueba de hipótesis, es realizar un estudio de capacidad del proceso para determinar la variabilidad existente comparada con la preliminar.

Rediseño de componentes de eje

Se realizan mejoras a los componentes del equipo como son Fig. 5 los sujetadores del eje X así como Fig. 6 sujetador de antorcha con el objetivo de hacer que el eje sea más estable al momento de estar en proceso de tal forma que se reduzca la vibración permitiendo mantenerla dentro de las especificaciones dadas.



Fig. 5- Sujetador eje X.



Fig. 6- Sujetador de antorcha.

Una vez efectuados lo rediseños de las piezas del eje x e instaladas las mismas se realiza una prueba más para ver cómo funciona con los nuevos cambios y también ver cómo se encuentra la vibración con estos, cabe mencionar que se utiliza el mismo punto de análisis que se definió anteriormente así como la hipótesis nula anterior.

III. Resultados

A continuación se tiene el resultado del análisis Anova mismo que se realiza en el programa de Minitab. En la Fig. 7 se muestra información de la descripción estadística realizada a los datos tomados, los cuales se muestran en la Tabla 1. En esta información se muestra el total de datos analizados, los cuales son 24 ya que se eliminaron unos datos por causas conocidas y ajenas al proceso, para poder realizar la prueba de normalidad, satisfaciendo así esta prueba, la cual indica que se puede llevar a cabo el estudio. De igual forma se puede observar la media, su desviación estándar, varianza, mediana, su valor mínimo y máximo.

Variable	N	Mean	SE Mean	StDev	Variance
Nivel Vib.	24	6.221	0.211	1.031	1.063

Variable	Minimum	Median	Maximum	Range
Nivel Vib.	4.600	6.250	8.700	4.100

Fig. 7- Descripción estadística.

En la Fig. 8 Se muestra el análisis Anova para el problema presentado donde se puede observar que el P-value es mayor a 0.05 por lo tanto existe evidencia de un intervalo de confianza del 95% de que la hipótesis nula es significativa, así como también se muestra el intervalo de cada posición en la pequeña gráfica donde se observa que las medias no están tan dispersas una de otra y que las varianzas son semejantes.

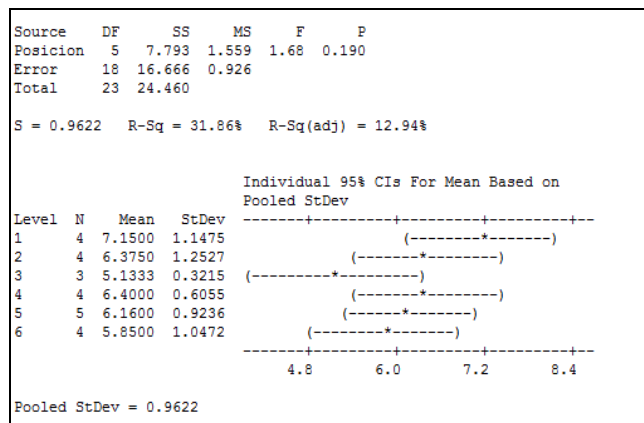


Fig. 8- Análisis Anova.

En la Fig. 9. Se muestra el resultado del análisis de una muestra z donde la hipótesis nula es igual o menor a 4.5mm/s, como se puede observar, así como se asume que la desviación estándar es 1.031, la cual fue determinada en el análisis Anova, un p-value menor a 0.05 lo que indica que existe la suficiente evidencia a un 95% de confianza de no aceptar la hipótesis nula planteada.

Test of mu = 4.5 vs > 4.5							
The assumed standard deviation = 1.031							
Variable	N	Mean	StDev	SE Mean	95% Lower Bound	Z	P
Nivel Vib.	24	6.221	1.031	0.210	5.875	8.18	0.000

Fig. 9- Prueba para un factor.

En la Fig. 10 se muestran los resultados de la prueba de hipótesis efectuada.

Test of mu = 6.2 vs < 6.2							
The assumed standard deviation = 0.8079							
Variable	N	Mean	StDev	SE Mean	95% Upper Bound	Z	P
Nivel Vib.	573	2.2941	0.8079	0.0338	2.3496	-115.73	0.000

Fig. 10- Segunda Prueba un factor.

De estos resultados arrojados en Minitab del estudio efectuado de un factor se puede concluir que dado un p-value menor a 0.05 con un intervalo de confianza del 95% se tiene la certeza de decir que existe la suficiente evidencia de no aceptar la hipótesis nula y aceptar la alternativa la cual dice que la media de la muestra es menor a 6.2 mm/s.

En resultado al análisis realizado a la nueva muestra de datos de nivel de vibración se obtiene una media de 2.29 mm/s.

En la Fig. 11 se muestran los resultados de la prueba de ajuste realizada para la prueba de normalidad donde se puede observar que tiene un Anderson-Darling pequeño y un p-value grande lo que indica que la distribución que mejor se ajusta a los datos es la Lognormal.

Distribution	AD	P	LRT	P
Normal	5.798	<0.005		
Box-Cox Transformation	0.821	0.034		
Lognormal	0.821	0.034		
3-Parameter Lognormal	0.783	*		0.485
Exponential	115.061	<0.003		
2-Parameter Exponential	64.824	<0.010		0.000
Weibull	5.096	<0.010		
3-Parameter Weibull	1.409	<0.005		0.000
Smallest Extreme Value	22.425	<0.010		
Largest Extreme Value	0.886	0.023		
Gamma	1.071	0.009		
3-Parameter Gamma	0.668	*		0.013
Logistic	4.227	<0.005		
Loglogistic	1.468	<0.005		
3-Parameter Loglogistic	1.488	*		0.745

Fig. 11- Prueba de ajuste.

En la Fig. 12 se muestran dos gráficas comparativas de la capacidad del proceso de las cuales la Fig. 12- a) muestra como se encontraba el proceso al realizar un análisis preliminar de la situación del equipo y la Fig. 12 b) muestra como se encuentra el proceso ahora.

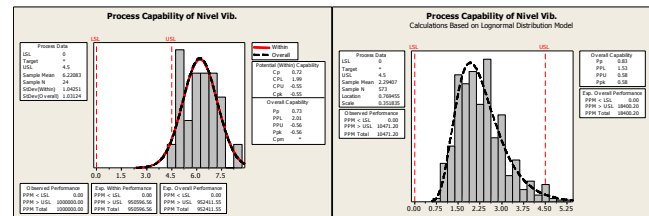


Fig. 12- a) capacidad del proceso, antes; b) capacidad del proceso, después.

La Fig. 12- b) se muestra un índice de capacidad menor a 1.33 lo que indica que aún existe demasiada variabilidad en el proceso. Con este nuevo índice de capacidad se muestra que el proceso mejoró comparado con la Fig. 12- a) donde existe mayor variabilidad.

Se puede observar que el Ppk (Cpk) de la Fig. 12- b) es mucho más pequeño que el Pp (Cp), esto indica que la media del proceso está alejada del centro de las especificaciones.

En la Fig. 13 se puede ver el análisis de la prueba de hipótesis el cual tiene un p-value menor a 0.05 lo que indica que no existe la suficiente evidencia para aceptar la hipótesis nula planteada de tal forma que se puede decir con un 95% de

confianza que existe evidencia para aceptar la hipótesis alternativa.

Test of $\mu = 6.2$ vs < 6.2						
The assumed standard deviation = 0.717						
				95% Upper		
Variable	N	Mean	StDev	SE Mean	Bound	Z P
Niv. Vib.	359	1.9986	0.7170	0.0378	2.0609	-111.03 0.000

Fig. 13- Prueba de hipótesis final.

En la Fig. 14 se muestra el estudio realizado para ver cuál es la distribución que mejor se ajusta a la prueba de normalidad obteniendo como resultado la distribución de Gamma ya que esta cuenta con un Anderson-Darling más pequeño y un p-value más grande. El estudio de análisis de capacidad se lleva a cabo en base a esta distribución el cual da como resultado el gráfico siguiente.

Distribution	AD	P	LRT	P
Normal	3.091	<0.005		
Box-Cox Transformation	0.868	0.026		
Lognormal	1.013	0.011		
3-Parameter Lognormal	0.631	*	0.033	
Exponential	69.504	<0.003		
2-Parameter Exponential	31.015	<0.010	0.000	
Weibull	2.487	<0.010		
3-Parameter Weibull	0.822	0.036	0.000	
Smallest Extreme Value	12.273	<0.010		
Largest Extreme Value	0.789	0.042		
Gamma	0.642	0.097		
3-Parameter Gamma	0.640	*	0.506	
Logistic	2.350	<0.005		
Loglogistic	0.955	0.008		
3-Parameter Loglogistic	0.843	*	0.303	
Johnson Transformation	0.628	0.101		

Fig. 14- Prueba de ajuste final.

En la Fig. 15- b) nuevo análisis, se muestra el gráfico de Cpk realizado con las nuevas lecturas dadas. Al igual que se muestra el gráfico del análisis anterior para realizar un comparativo donde se puede apreciar que en la Fig. 15- b) nuevo análisis, los datos se encuentran dentro de las especificaciones dadas comparado con el gráfico de la Fig. 15-, a) análisis anterior. El nuevo Cpk incremento pero no lo suficiente para decir que el proceso es capaz ya que sigue existiendo variabilidad cabe mencionar que se

puede observar que se mejoró comparado con el que se tenía anteriormente.

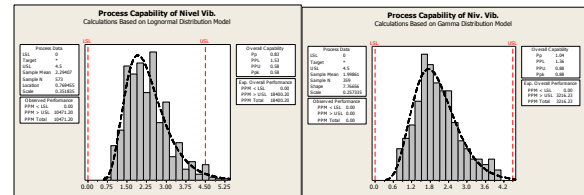


Fig. 15- Comparativo de capacidad de proceso a) análisis anterior b) nuevo análisis.

IV. Discusión

Uno de los puntos más discutido al inicio de todo análisis son los parámetros o requisitos necesarios para dicho análisis empezando con el estadístico que mejor se ajuste a la prueba que se efectúa como son el número de datos si estos deben o no seguir la curva de la normalidad.

Satisfacer la prueba de normalidad implica uno de los métodos empleados que es el de análisis de capacidad del proceso ya que si no se cumple la normalidad se realiza un estudio que busque la distribución que mejor se ajusta a los datos y así poder llevar a cabo el análisis con esa distribución. Para los otros estadísticos no fue necesario cumplir con este requisito.

V. Conclusiones

En el análisis Anova se puede concluir que, se tiene un intervalo de confianza del 95% de probabilidad de que el nivel de vibración es el mismo en cualquier posición, lo que indica que no se tiene una gran diferencia de vibración de una posición a otra. Se genera la misma vibración realizar un corte en la posición tres a realizarlo en la posición cinco.

Con el análisis preliminar se llega a la conclusión que actualmente el equipo de corte no se encuentra dentro del rigor de vibración aceptado por la norma ISO antes mencionada. Teniendo una media de 6.2mm/s valor que no está dentro del rango, menor o igual a 4.5mm/s, el cual es permisible para este tipo de equipo sino más bien

se encuentra en un rigor, mayor de 4.5mm/s, lo que indica que se encuentra en un estado de prueba peligroso.

Los siguientes análisis efectuados se pueden concluir, que se cumple con el objetivo de reducir el nivel de vibración al que anteriormente tenía el sistema en el análisis efectuado de la situación actual. Cumpliéndose así establecer la media dentro de las especificaciones de la norma ISO 2372 y VDI 2056 la cual dice que una maquinaria de este tipo debe estar dentro de un rigor de vibración (1.81 a 4.5) mm/s.

VI. Referencias

- [1] Balakumar Balachandran, Edward B. Magrab, Bruna Anzures. Vibraciones, International Thomson, 2006, México.
- [2] Robert F. Steidel Jr., Introducción a las vibraciones mecánicas, Compañía Editorial Continental, S.A. de C.V., 1991, México. Robert W. Soutas-Little, Daniel J. Inman, Daniel S. Balint, Ingeniería Mecánica Dinámica, Cengage Learning Editores, S.A. de C.V., 2009
- [3] F. Al-Badour, M. Sunar, L. Cheded, Vibration analysis of rotating machinery using time–frequency analysis and wavelet technique. Mechanical Systems and Signal Processing, 25, 2011, 2083-2101.
- [4] T.H. Loutas, D. Roulias, E. Pauly, V. Kostopoulos, The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery, Mechanical Systems and Signal Processing, 25, 2011, 1339-1352.
- [5] S. Ebersbach, Z. Peng, Expert system development for vibration analysis in machine condition monitoring, Expert Systems with Applications, 34, 2008, 291-299.

- [6] M. C. Carnero, R. González-Palma, D. Almorza, P. Mayorga, C. Lopez-Escobar, Statistical quality control through overall vibration analysis. Mechanical Systems And Signal Processing, 24, 2010, 1138-1160.
- [7] Brian Nesbitt, Pumping Manual International Handbook of Pumps and Pumping, B.H. 2006.
- [8] Guía del usuario, Extech Instruments Corporation, México, 2007.
- [9] Humberto Gutiérrez Pulido, Román de la Vara Salazar, Control Estadístico de Calidad y Seis Sigma, Mc Graw Hill, 2009.
- [10] Ing. Mariza del Carmen Martínez contreras, Dr. Francisco Javier López Jáquez, Caracterización de la firma acústica de una máquina de corte.

VII. Autores

María Magdalena Hernández Ramos obtuvo su título de Ingeniería en Manufactura y actualmente esta en proceso de obtener el grado de Maestría en Ingeniería en Manufactura Opción Diseño del Producto por la Universidad Autónoma de Ciudad Juárez,

Dr. Francisco Javier López Jáquez obtuvo su título de Doctorado en filosofía en el área de Manufactura por la Universidad de Liverpool en el Reino Unido, Inglaterra. Actualmente está adscrito como profesor investigador en la Universidad Autónoma de Ciudad Juárez.

Comentarios finales:

El análisis muestra que aunque la media se encuentre dentro de las especificaciones aún existe variabilidad en este dejando como resultado una puerta abierta a futura mejora. Ya que el objetivo del proyecto es hacer que la media se encuentre dentro de las especificaciones.

Análisis del Ensamble de una Máquina Cortadora de Plasma Utilizando la Metodología de Diseño para Ensamble

Paola Elizabeth López Ríos, Francisco Javier López Jáquez

Universidad Autónoma de Ciudad Juárez. Instituto de Ingeniería y Tecnología, departamento de Ingeniería Industrial y Manufactura. Av. Del Charro 450 Nte. Colonia Partido Romero C.P. 32310, Ciudad Juárez, Chih., México. Teléfono: 656 206 05 31 / E-mail: lrpaolae@hotmail.com.
Tel: 656 6884843, frlopez@uacj.mx

Resumen

Una máquina cortadora de plasma ubicada en el laboratorio de Ingeniería de Manufactura, en la Universidad Autónoma de Ciudad Juárez, este es el punto de inicio del presente escrito, donde se describe el proceso para realizar el análisis del ensamble de esta máquina, utilizando la metodología de Boothroyd: Diseño para Ensamble (DFA). Esta metodología inicia con la identificación de cada una de las secciones de la máquina, como segundo paso se caracteriza cada uno de los componentes que integran estas secciones, el tercer paso consiste en evaluar cada uno de los componente basados en base a las tablas de Boothroyd. Estos valores de las tablas consideran los movimientos e inserción para cada componente de la máquina. Como siguiente paso en la metodología, se realiza el cálculo del indicador DFA para cada de las secciones de la máquina cortadora de plasma, este valor determina el nivel de cumplimiento a las guías para el diseño, además permite identificar oportunidades de mejora en el ensamble de la máquina.

Palabras Clave: Diseño, Ensamble, Máquina, Análisis, DFA, Componentes, Reducción, Combinación.

I. Introducción

Diseñar es sinónimo de creatividad, innovación o modificación, esta derivado del término italiano designio: un dibujo, definido como el proceso previo a una configuración mental en la búsqueda de una solución [1, 2].

DFA es una metodología que se comenzó a utilizar en la época de los 60's. Diseño para ensamble o DFA se define como la acción de modificar, unir o combinar componentes para dar forma a un producto. El utilizar este método se obtiene un 70% de reducción de costos, de acuerdo a las estadísticas de Boothroyd [3, 4, 5].

El concepto DFA está basado en reducir y combinar componentes para mejorar de ensamble de un producto [6].

El estado del arte para este proyecto hace referencia a la aplicación de un DFA en una compañía mundial denominada Motorola, donde se realizaron 12 análisis del ensamble en productos como laptop y LCD TV, obteniendo

como resultado una reducción de componentes que va desde un 30% a 50% [7, 9].

El presente proyecto se basa en el análisis del ensamble de una máquina cortadora de plasma, mediante DFA. Una máquina cortadora de plasma, se utiliza para corte de metales ferrosos, y acero inoxidable principalmente. El equipo basa su funcionamiento de corte en base a descarga de iones y aire [5, 7, 8].

Actualmente esta cortadora de plasma, se encuentra instalada en laboratorio de Ingeniería de Manufactura, en el Instituto de Ingeniería y Tecnología, de la Universidad Autónoma de Ciudad Juárez.

La máquina cortadora de plasma en cuestión ha sido ensamblada una forma única, no encontrando un punto de referencia para determinar la eficiencia del ensamble actual.

La propuesta para este proyecto es observar, documentar, analizar y evaluar las secciones y los

componentes que integran la máquina cortadora de plasma, utilizando DFA.

El objetivo principal de este proyecto es analizar cada una de las secciones y componentes de la máquina cortadora de plasma, utilizando las guías de DFA. Además se pretende obtener el valor de eficiencia del ensamble mediante las fórmulas de DFA índex.

II. Desarrollo

El primer paso es segmentar la máquina cortadora de plasma, la estrategia es dividir para facilitar el entendimiento y comprensión de cada uno de los componentes que integran la máquina.

Los segmentos en que se divide la máquina cortadora de plasma son tres: la estructura, un eje X y un eje Y. Ver figura 1.

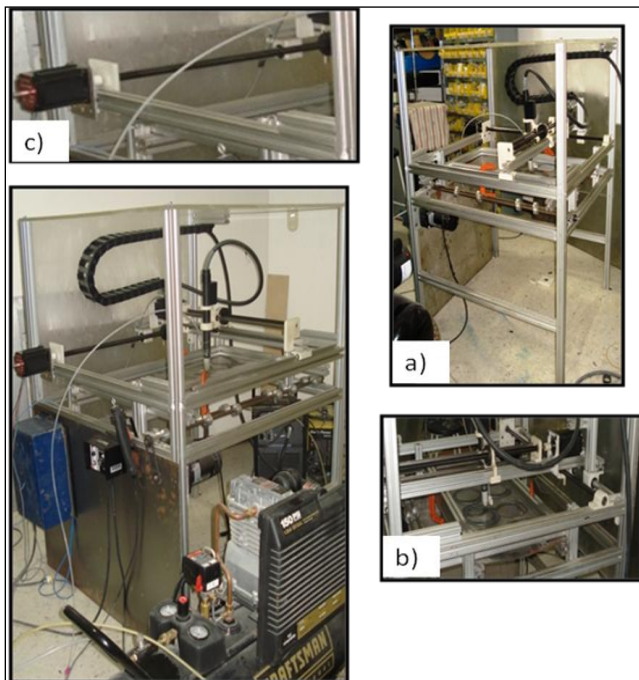


Fig. 1 Secciones de maquina cortadora de plasma.

El segundo paso consiste en generar una lista de los componentes que integran cada sección de la máquina. A cada componente se le asigna un código de identificación, así como una descripción del componente. Ver tabla 1.

Para el tercer paso se identifican los componentes que son únicos en el ensamble, así como los que

se requieren dos o más. A este valor se le denomina RP. Ver tabla 1.

El cuarto paso es caracterizar cada componente, empezando por describir las dimensiones de largo, ancho y espesor, además indicar el tipo de material del cada componente. Ver tabla 1.

CARACTERISTICAS DE LOS COMPONENTES							
PASO 1	PASO 2		PASO 3	PASO 4			
Seccion	Codigo de Identificacion	Descripcion	Cantidad (RP)	Largo cm	Ancho cm	Espesor cm	Material
Estructura	1A	Base inferior	4	90	4	4	Aluminio Extruido
	2A	Base superior	4	50	4	4	Aluminio Extruido
	3A	Lado A	5	98	4	4	Aluminio Extruido
	4A	Lado B	4	94	4	4	Aluminio Extruido
	5A	Soporte A	1	15	4	4	Aluminio Extruido
	6A	Soperte B	1	20	4	4	Aluminio Extruido
	7A	Sopрте C	1	25	4	4	Aluminio Extruido
	XA	Tornillos	10	25	13	8	Acero
	YA	Seguro	10	35	12	10	Aluminio

Tabla 1. Tabla de características de los componentes.

En la tabla 1, se muestran las características básicas de cada uno de los componentes del ensamble de la sección 1.

A continuación se desarrolla el análisis efectuado para el componente 1 de la tabla 1, utilizando las tablas de Boothroyd, 2002. Componente número uno, es denominado base de soporte inferior con dimensiones 90 cm de largo, 4 cm de ancho, con una altura de 4 cm de altura, el material de esta pieza es aluminio extruido. Para el paso tres, la cantidad de partes se determina en base a la cantidad de piezas iguales que se requieren en el proceso de ensamble para la base inferior de la estructura. En la tabla 2 se muestran las características del componente 1, sección 1.

ANÁLISIS DFA : Estructura										
		PASO 3	PASO 5	PASO 6	PASO 7	PASO 8	PASO 9	PASO 10	PASO 11	Paso 12
	Descripcion	Cantidad piezas (RP)	Tiempo adquisicion herramienta (TA)	Codigo de movimiento	Tiempo de movimiento (TH)	Codigo de inserccion	Tiempo de insercion (TI)	Tiempo total TA+RP(TH+ TI)	Posible eliminacion	Observaciones
1	Base inferior	4		00	1.13	30	3.6	18.92	1	Tomar y colocar

Tabla 2. Base inferior análisis DFA.

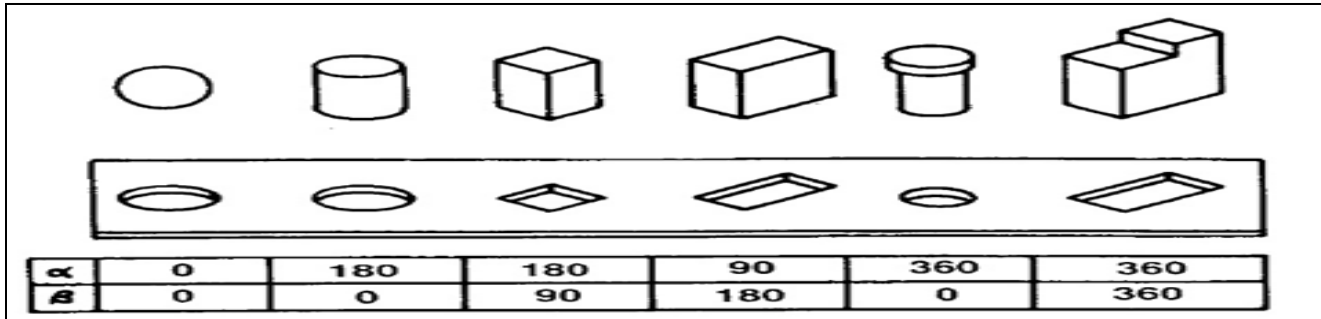


Fig. 2. Alpha y Beta rotación simétrica. Boothroyd, 2002.

Continuando con el desarrollo de la metodología, en el quinto paso se debe determinar el tiempo en tomar una herramienta, este puede ser Nulo (0) o 2.9. El valor 2.9 es asignado cuando se requiere utilizar una herramienta para realizar el ensamble, y se encuentra especificado en la tabla del código de inserción de Boothroyd, 2002, es un valor estándar. Para el componente uno se asigna el valor nulo debido a que no requiere herramienta para realizar el ensamble [3].

Para el sexto paso se requiere utilizar la tabla de código de movimiento de Boothroyd. El ángulo total del componente, se inicia con el primer valor de rotación de acuerdo a la figura 2, se determinan el ángulo de rotación del componente en alpha y beta (vertical y horizontal, respectivamente). La base inferior de acuerdo a sus dimensiones es un rectángulo, por lo que el valor de rotación asignado es menor o igual a 360, por lo tanto el primer código es el cero (0). Para determinar el segundo código se debe definir en base a dificultad para tomar y espesor del mismo. Se analiza que este componente es fácil de tomar, y tiene un espesor mayor a 15mm, por lo que segundo número del código será cero (0) [3].

En el séptimo paso el tiempo de movimiento se debe tomar el valor del número que hace referencial código doble cero (00), de la tabla de Boothroyd. En este análisis el valor de movimiento tiene un tiempo definido 1.13 segundos, como lo muestra la tabla 2.

Para el octavo paso se deben determinar el código de inserción utilizando la tabla Boothroyd basado en dos aspectos, el primero de ellos está enfocado al movimiento para colocar el componente en la máquina, considerando visibilidad, restricciones o área obstruida, el segundo se determina en la rapidez con la que se ensambla el componente en la máquina. Para el primer valor del código, analizando la base inferior, no se encuentra obstruida, sin visibilidad o con dificultades para realizar el ensamble por lo que el primer valor es tres (3). Para el segundo valor de rapidez en la inserción se debe analizar si la pieza requiere alineación, en este caso la base inferior es fácil de alinear, por lo que se asigna el valor de cero (0), se concluye que el código de inserción es igual a tres con cero (30) [3].

En el noveno paso, el tiempo de inserción se debe tomar el valor del número que hace referencial código tres con cero (30), de la tabla de

Boothroyd para los valores de movimiento el cual tiene como tiempo definido 3.6 segundos [1].

Para el decimo paso, es la suma de los tiempos de movimiento e inserción, multiplicados por la cantidad de partes que se tiene de ese componentes en ensamble, mas el tiempo de tomar la herramienta. Para el componente uno, tenemos un valor de movimiento de 1.13 mas un valor de inserción de 3.6, esto se repite por las cuatro bases del ensamble, es igual a 18.92, mas el tiempo de tomar herramienta que este componente es nulo. El total para la base inferior es 18.92 segundos [3].

En el paso once se debe anotar el valor de cero o uno, dependiendo de las posibilidades de eliminar o combinar esta componente, a lo que este caso se encuentra la posibilidad de unir la base inferior con la base superior a lo que el valor de una posible eliminación es uno. Si es un componente factible para combinar o eliminar [3].

Para el paso doce, se anotan los movimientos generales que realiza el operador para ensamblar este componente, es esta base inferior solo debe tomar y colocar. Estos mismos pasos se llevaron a cabo para completar las cuatro secciones de la máquina cortadora de plasma [3].

III. Resultados

Para cada sección se genero una tabla de información para obtener el valor DFA. En base a este análisis se procede a enlistar las oportunidades de mejora de cada segmento y realizar la evaluación teórica de esta propuesta.

En el primer segmento denominado estructura desarrollado en la tabla 3, ver en siguiente página. En esta tabla encontramos que cuenta con 9 elementos, 70 componentes, con un tiempo de ensamble total de 266.6 segundos y encontramos siete componentes que pueden ser combinados o reemplazados.

Basados en los datos de la tabla 3 se calcula el valor del DFA índice o eficiencia de ensamble para la estructura:

$$Ema = \frac{(Nmin)(ta)}{tma} = \frac{7(3)}{441.05} = .047 \text{ o } 4.7\% \quad (1)$$

El resultado para el segmento uno es de 4.7%. A continuación se definen la oportunidad de mejora de este segmento.

La estructura vertical de la máquina está compuesta por cuatro postes de aluminio extruido con una altura de 140 cm. Estos postes son el resultado de un subensamble entre la base inferior de 90 cm y la base superior de 50 cm, mediante tornillos y sujetadores, ambas bases de aluminio extruido.

Considerando que el material de los componentes es el mismo, se define como oportunidad de mejora el fabricar base única con una altura de 140 cm. Las acciones necesarias serian: 1) Eliminar cuatro bases superior del ensamble, 2) Eliminar soporte A y soporte C, debido a que estos sostenían la base superior en la base superior, 3) Eliminar seis tornillos y seis sujetadores al combinar las bases, 4) Diseñar la base vertical (única) con un largo de 140 cm.

Adicional a lo antes mencionado, se encontró oportunidad en unificar las dimensiones de las bases horizontales denominadas lado A con un largo de 98 cm y lado B con un largo de 94 cm, las cuales evaluando la funcionabilidad del equipo no requiere ser de diferentes dimensiones.

El beneficio principal al utilizar medidas estándar es reducir la cantidad de variables al usuario final, así como facilitar el proceso de ensamble.

A continuación en la tabla 4, se muestra el análisis para los mejoramientos teóricos al segmento uno.

ANÁLISIS DFA: Estructura										
		PASO 3	PASO 5	PASO 6	PASO 7	PASO 8	PASO 9	PASO 10	PASO 11	Paso 12
	Descripcion	Cantidad piezas (RP)	Tiempo adquisicion herramienta (TA)	Codigo de movimiento	Tiempo de movimiento (TH)	Codigo de inserccion	Tiempo de insercion (TI)	Tiempo total TA+RP(TH+ TI)	Posible eliminacion	Observaciones
1	Base inferior	4		00	1.13	30	3.6	18.92	1	Tomar y colocar
2	Base superior	4		00	1.13	30	3.6	18.92	1	Tomar y colocar
3	Lado A	5		00	1.13	41	8.0	45.65	0	Tomar, alinear y colocar
4	Lado B	4		00	1.13	41	8.0	36.52	0	Tomar, alinear y colocar
5	Soporte A	1		00	1.13	41	8.0	9.13	1	Tomar, alinear y colocar
6	Soperte B	1		00	1.13	41	8.0	9.13	1	Tomar, alinear y colocar
7	Sopрте C	1		00	1.13	41	8.0	9.13	1	Tomar, alinear y colocar
8	Tornillos	25	2.9	10	1.5	61	4.5	152.9	1	Tomar y ajustar
9	Sujetador	25		00	1.13	61	4.5	140.75	1	Tomar y ajustar
		70						441.05	7	

Tabla 3. DFA: Estructura Actual.

ANÁLISIS DFA: Estructura Teórico										
		PASO 3	PASO 5	PASO 6	PASO 7	PASO 8	PASO 9	PASO 10	PASO 11	Paso 12
	Descripcion	Cantidad piezas (RP)	Tiempo adquisicion herramienta (TA)	Codigo de movimiento	Tiempo de movimiento (TH)	Codigo de inserccion	Tiempo de insercion (TI)	Tiempo total TA+RP(TH+ TI)	Posible eliminacion	Observaciones
1	Base vertical	4		00	1.13	30	3.6	18.92	1	Tomar y colocar
2	Lado A	9		00	1.13	41	8.0	82.17	1	Tomar, alinear y colocar
3	Sopрте C	1		00	1.13	41	8.0	9.13	1	Tomar, alinear y colocar
4	Tornillos	19	2.9	10	1.5	61	4.5	116.9	1	Tomar y ajustar
5	Sujetador	19		00	1.13	61	4.5	106.97	1	Tomar y ajustar
		52						334.09	5	

Tabla 4. DFA: Estructura Teórico.

En base a la información de la tabla 4, encontramos que teóricamente la cantidad de elementos se reduce a 5, esto corresponde a un 55% de mejoramiento sobre la estructura actual, asimismo la cantidad de componentes presentó una reducción del 25%, al mostrar teóricamente que se requieren 52 componentes en lugar de los 70, por lo que el tiempo total de ensamble teórico es de 334.09, se redujo un 22%, el tiempo total de ensamble del segmento uno, en el cálculo de DFA teórico, encontramos que no se incremento, sustituyendo en la ecuación (1).

$$Ema = \frac{5(3)}{334.09} = .044 \text{ o } 4.4\%$$

Este mismo procedimiento se realizó para las restantes tres secciones de la máquina cortadora de plasma.

IV. Conclusiones

De acuerdo al análisis de resultados presentado anteriormente se determina que el valor de DFA índice del ensamble actual con respecto al valor teórico, no se incrementa en proporción a los mejoramientos recomendados, en contraste la cantidad de elementos y componentes si muestra un mejoramiento significativo de reducción, lo cual facilita el ensamble, y se muestra reflejado en las reducciones de tiempo de ensamble de las secciones.

A lo que se concluye que el valor DFA index no es el punto de referencia que determina el impacto del mejoramiento. A continuación la figura 3, muestra los valores actuales y teóricos del DFA Índice.

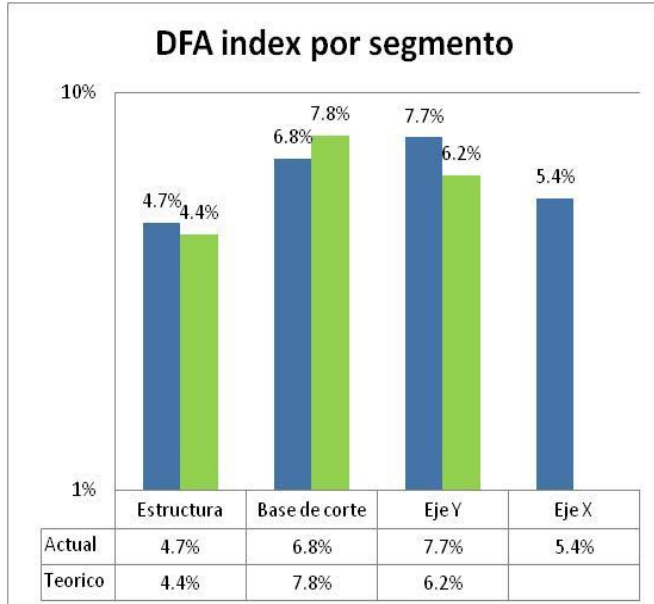


Fig. 3. Gráfica de Valores DFA Índice.

En contraste con este gráfico, se encuentran la gráfica 4: de elementos y componentes y la gráfica 5: los tiempos de ensamble actual contra el tiempo teórico evaluado. Esta muestra el impacto de los mejoramientos recomendados, los cuales fueron evaluados de una forma teórica.

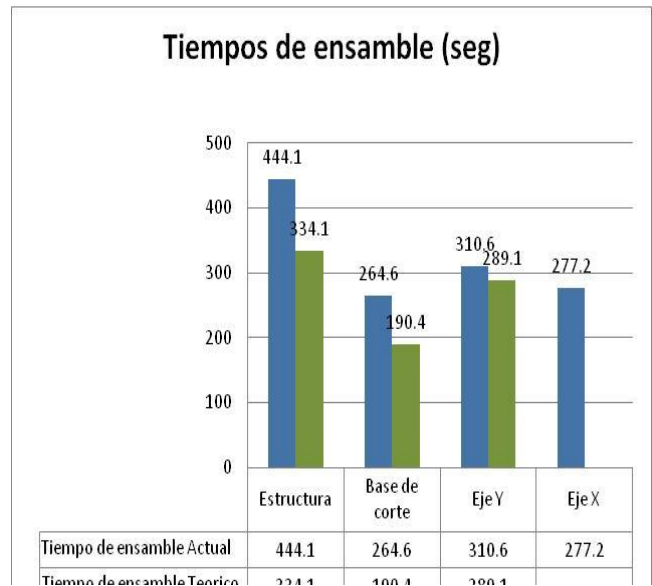


Fig. 5. Gráfica de tiempos de ensamble.

En la gráfica 4, el porcentaje de reducción de componentes se encuentra en 10% para la sección 3, un 25% para la sección 1, y un 30% de mejoramiento en la sección 3. En la gráfica 5, se muestra como en cada de las sección presentó un mejoramiento en la misma proporción para cada sección.

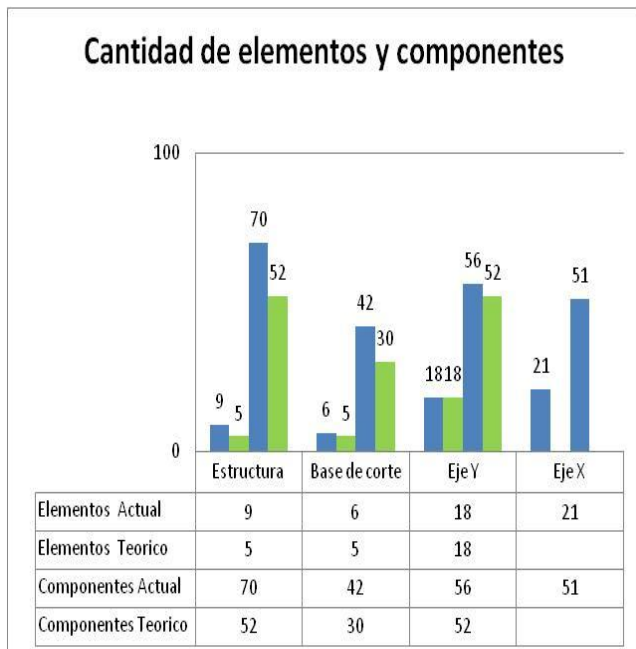


Fig. 4. Gráfica elementos y componentes.

En conclusión se establece que el valor de eficiencia del ensamble o DFA índice, no es representativo para determinar si el ensamble actual está diseñado dentro de las guías de diseño para ensamble, esto basado en que los valores de las evaluaciones actuales en el caso de sección 1 y 3 mostradas en la gráfica 3, fueron similares a los valores teóricos. Pero estos últimos están basados en mejoramientos recomendados en base a las guías para el diseño para ensamble.

En cambio de una manera más sencilla, fácil de interpretar y rápida de comparar las características cuantitativas mostradas para los elementos y componentes indican que el valor actual contra el valor teórico, si obtuvo un mejoramiento constante en las cuatro secciones de la máquina cortadora de plasma, como lo muestra la figura 4.

De igual forma los tiempos de ensamble actuales con respecto a los teóricos, también indican que se presentó un mejoramiento del ensamble actual en comparación con el teórico, es decir estas tres

características del ensamble determinan y afirman que el ensamble actual tiene área de oportunidad.

Con este análisis se cumple el objetivo del proyecto, determinando que el ensamble actual puede ser mejorado mediante la aplicación de las guías de diseño para el ensamble, aunque el DFA index no muestre un valor que lo respalde, considerando las características cuantitativas como tiempo, componentes y secciones lo demuestran.

Algunas las recomendaciones, basados en las guías de diseño para ensamble se puede recomendar al diseñador los siguientes aspectos:

Estandarizar el material utilizado en la fabricación de máquina, esto permite utilizar el mismo método de sujeción de componentes.

Utilizar tornillería estándar para evitar mezclas de componentes y cambios de herramienta durante el ensamble.

Combinar componentes para reducir la cantidad y así reducir el tiempo de ensamble.

Diseñar componentes forma única para ensamblar, evitando así errores al momento de ensamblar por el usuario.

Estandarizar forma y tamaño de los componentes para facilitar al usuario la identificación fácil y sencilla de los mismos.

Adicional a estas recomendaciones para mejorar el diseño para ensamble de la máquina cortadora de plasma, también es importante mencionar las recomendaciones generales para una continuación del presente proyecto, estas son:

1) Utilización de software DFA, para reducir tiempo de análisis. Con esto se evalúan las alternativas de forma fácil, sencilla y rápida.

2) Re-diseñar los componentes para sujeción (tornillos), analizando y evaluando las alternativas de las guías para diseño para ensamble.

V. Referencias

[1] B Huthwaite 2004. Lean Desing Solution. Editorial Words Plus Desing. 2004. Estados Unidos.

[2] F Meyers, M Stephens. Diseño de instalaciones de manufactura y manejo de materiales. Editorial Pearson. 2006 México.

[3] G Boothdroyd. Design for Manufacture and Assembly. Editorial Marcel Dekker Inc. 2002 New York.

[4] A Everett, E Ronald. Administración de la producción y las operaciones: Conceptos, modelos y funcionamientos. Editorial Pearson. 1992. México

[5] M Groveer. Fundamentos de manufactura moderna: materiales, proceso y sistemas. Editorial Prentice-Hall 1997. México.

[6] S Schmid, S Kalpakjani. Manufactura, Ingeniería y Tecnología. Editorial Pearson. 2002. México.

[7] J Bralla. Design for Manufacturability. Design for Manufacturability. Editorial McGraw-Hill. 1999. Estados Unidos

[8] J Lesko. 2004, Diseño industrial, Guía de materiales y procesos de manufactura. Editorial Limusa SA de CV. 2004. México.

[9] <http://www.dfma.com/news/motorola.htm> fecha de última consulta 1/09/2011.

VI. Autores

Paola Elizabeth López Ríos obtuvo su título de Licenciatura en Ingeniería Industrial y de Sistemas en el Instituto Tecnológico de Sonora y actualmente está en el proceso de obtener el grado de Maestría en Ingeniería en el área de Manufactura opción en Procesos de Fabricación por la Universidad Autonoma de Ciudad Juárez.

Francisco Javier López Jáquez obtuvo su título de Doctorado en filosofía en el área de Manufactura por la Universidad de Liverpool en el Reino Unido, Inglaterra. Actualmente está adscrito como profesor investigador al departamento de Ingeniería Industrial y Manufactura de la Universidad Autónoma de Ciudad Juárez.

Robot Humanoide ZERO: Diseño, Análisis y Prototipo

*,** Efraín Hernández Franco, ** Ramiro Velázquez Guerrero

* Universidad Panamericana campus Guadalajara, Calzada Circunvalación Poniente 49, Ciudad Granja, CP 45010, Zapopan, Jal., México, (33) 3679 07 08, ehernandezf@up.edu.mx.

** Universidad Panamericana campus Bonaterra, Josemaría Escrivá de Balaguer 101, Fracc. Rústicos Calpulli, CP 20290, Aguascalientes, Ags., México, (44) 9910 62 00, rvelazquez@up.edu.mx.

Resumen

Este artículo presenta los aspectos técnicos del robot humanoide ZERO. ZERO es una plataforma bípeda con 21 grados de libertad (GDL) accionado por servomotores que proporcionan un torque confiable y movimientos de alta precisión. Construido enteramente de aluminio, ZERO es ligero (5.5 kg) y compacto (altura: 60 cm, ancho: 33 cm) al mismo tiempo que fuerte y robusto. ZERO es capaz de moverse hacia adelante, hacia atrás, a los lados, girar en cualquier dirección, acostarse, y levantarse. Además, ZERO realiza el seguimiento y reconocimiento de objetos usando una videocámara montada en su misma estructura. La principal motivación detrás de este proyecto es crear nuestra propia plataforma de investigación en control de la marcha bípeda. En este artículo se presenta el diseño, la implementación y el análisis mecánico de este prototipo.

Palabras Clave: análisis cinemático, espacios de trabajo, robot humanoide, robótica móvil.

I. Introducción

En la última década, se ha presenciado un creciente interés en robots humanoides y mecanismos que permiten la marcha bípeda. Esto se debe seguramente a las nuevas necesidades y expectativas que los humanos tienen de los robots. Al parecer las necesidades de automatización, ensamblaje y procesos de control de la industria han sido razonablemente cumplidas. Sin embargo, nuevos campos como los robots de servicio y asistencia, requieren diferentes tipos de robots con los que puedan interactuar los humanos. Los robots humanoides se han posicionado fuertemente como sistemas amigables para el humano por su antropomorfismo, apariencia e interacción amigable, alta capacidad de integración en ambientes humanos, etc.

Dentro de este contexto, muchos robots humanoides han sido desarrollados en los últimos años. Algunos ejemplos representativos son los robots H6 y H7 de la Universidad de Tokio [1],

Johnnie de la Universidad Técnica de Múnich [2], el SDR-3X de Sony [3], la serie HRP de

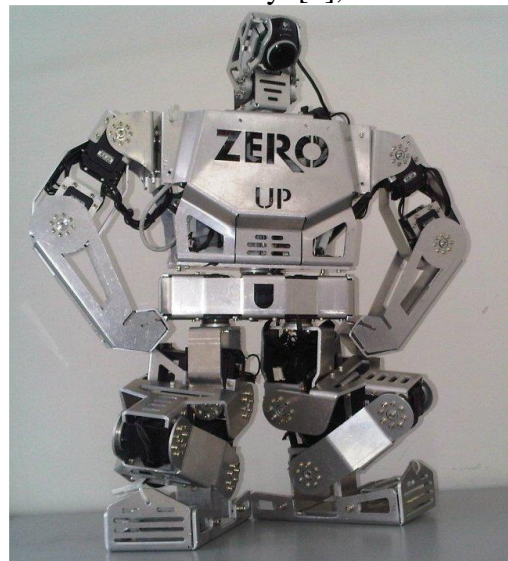


Fig. 1. Robot humanoide ZERO.

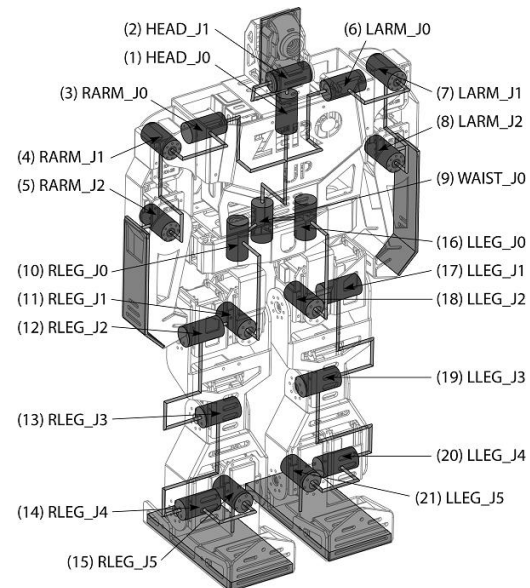
Industrias AIST/Kawada [4], y el bien conocido robot ASIMO de Honda [5]. Todas estas impresionantes plataformas son sin duda sistemas complejos de alta tecnología que representan una fuerte inversión económica de universidades y

centros de investigación. En consecuencia, pocos laboratorios bien financiados en el mundo tienen



(a)

uno de estos prototipos. Los investigadores que



(b)

Fig. 2. Robot humanoide ZERO: (a) diseño y (b) su estructura cinemática con etiquetas identificando todos los segmentos.

trabajan en los diferentes aspectos de la robótica humanoide tienen que buscar colaboraciones con estos laboratorios con la finalidad de probar sus trabajos. Desafortunadamente, la colaboración y los horarios no son siempre exitosos.

A partir del 2003, han aparecido en el mercado robots humanoides en miniatura a precios muy accesibles (por ejemplo, el robot Kondo [6], el robot Nao [7], el Robotinno [8], etc.). Estos prototipos miden generalmente de 30 a 40 cm de altura y tienen entre 16 y 19 GDL. Estos productos son definitivamente plataformas atractivas y robustas para aficionados y principiantes de la robótica humanoide. Su principal inconveniente es que no son completamente adecuados para investigación porque sus procesos de manufactura no fueron llevados a cabo con alta precisión, el número y tipo de sensores a bordo no son suficientes para permitir una retroalimentación efectiva para caminar y balancearse, y no ofrecen suficiente capacidad de procesamiento y cómputo.

Para superar las limitaciones antes mencionadas, la Universidad Panamericana decidió diseñar y

desarrollar un prototipo propio de robot humanoide. ZERO es una plataforma completamente accionada por servomotores que ofrece 21 GDL (Fig. 1). Su estructura consiste de eslabones complejos de aluminio manufacturados con alta precisión para reducir juegos mecánicos/desalineaciones y lograr movimientos precisos. ZERO es más alto que los robots miniatura disponibles comercialmente (60 cm) e incluye una videocámara, una amplia capacidad computacional y un número de sensores a bordo que permiten el desarrollo de estrategias de control para balancearse y caminar.

Este artículo presenta los aspectos técnicos de ZERO. El resto del artículo está organizado como sigue: La Sección 2 describe el diseño, implementación y prototipo, mientras que la Sección 3 presenta un análisis cinemático completo del robot ZERO. Finalmente, la conclusión resume las contribuciones principales y perspectivas de trabajo futuro.

II. Desarrollo e Implementación

A) Diseño

ZERO es una plataforma robótica humanoide diseñada para ser compacta, ligera, robusta y de alto rendimiento.

La Fig. 2(a) muestra su diseño completo. Note que ZERO es un prototipo que contiene una cabeza, torso, cintura, dos brazos y dos piernas. ZERO exhibe 21 GDL accionados por un total de 21 actuadores. Su estructura cinemática se detalla en la fig. 2(b). La cabeza consiste de una estructura pan & tilt que permite 2 GDL. Los brazos agregan 6 GDL, 3 GDL cada uno: hombro, brazo y codo. La cabeza y los brazos están montados en el torso y rotan en la cintura (1 GDL). Finalmente, las piernas agregan 12 GDL, 6 GDL cada una: uno al nivel de la cintura que permite la rotación entera de la pierna, dos en la parte superior, uno en la rodilla y dos en el pie.

Una videocámara a color está montada en la estructura pan & tilt de la cabeza. Tanto baterías como unidad de procesamiento se encuentran a bordo de la estructura.

B) Implementación

ZERO está construido completamente de dos tipos de aleaciones de aluminio: 3105-H22 y 3003-H23. La primera es maleable y se usa en las piezas más complejas que contienen dobleces. La segunda es rígida y se usa para las piezas más simples. Las piezas de aluminio H22 fueron cortadas con láser con ± 2 micrómetros de precisión, doblados con prensa hidráulica, y soldados usando soldadura de arco de gas tungsteno. Las piezas de aluminio H23 fueron manufacturadas con máquina de control numérico CNC de 3 GDL.

Los actuadores que accionan las articulaciones de ZERO son servomotores de tipo RX-10, RX-28, y RX-64 de Dynamixel [9]. Los tres proporcionan una precisión de paso de 0.29° mientras que el torque es 10, 28, y 64 kgf-cm, respectivamente. Los servomotores RX-10 son usados en la cabeza, los RX-28 en las extremidades superiores, y los RX-64 en la cintura y extremidades inferiores.

Como unidad de procesamiento, una fit-PC2 [10] ejecuta Windows embebido y C++ es usado para el procesamiento de imágenes y el control de los servomotores. Este módulo está localizado dentro del torso de ZERO (Fig. 3(a)). Un juego de

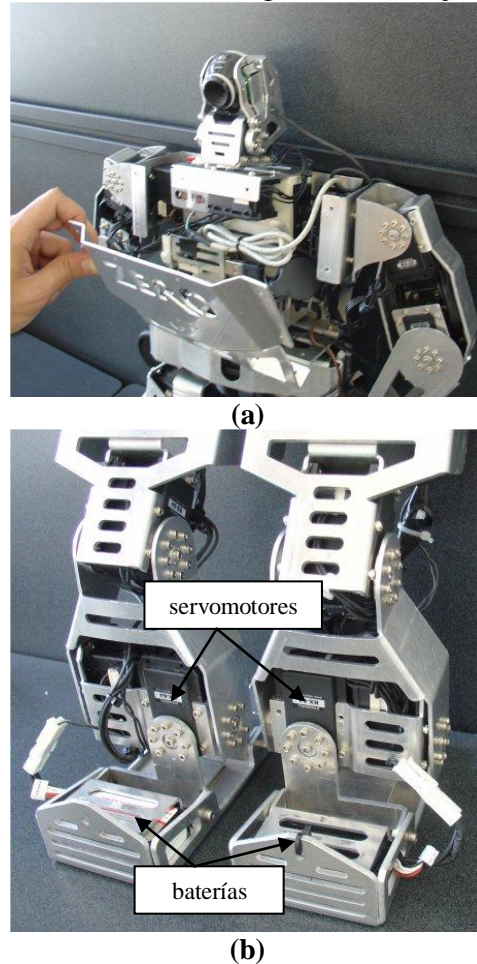


Fig. 3. (a) Unidad de procesamiento abordo y (b) baterías abordo.

baterías recargables de Litio-Polímero de 14.8 V, 850 mAh permiten 30 min de autonomía. Las baterías se encuentran en el pie (Fig. 3(b)).

C) Prototipo

El robot humanoide ZERO se muestra en la Fig. 1. Sus principales especificaciones físicas se enlistan en la Tabla 1.

Altura	60 cm
Ancho (brazos abiertos)	33 cm
Ancho (brazos cerrados)	72 cm
Profundidad	16 cm
Longitud de los brazos	24 cm
Longitud de las piernas	32 cm
Masa total	5.5 cm

Tabla 1. Especificaciones físicas del robot ZERO.

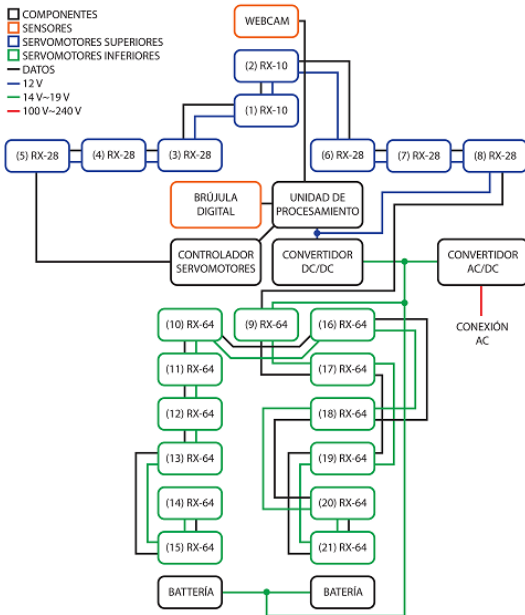


Fig. 4. Diagrama de bloques del Robot ZERO.

La Fig. 4 muestra el diagrama de bloques de conexiones de ZERO. La estructura principal comprende cuatro tipos de bloques conectados entre sí por líneas eléctricas y de datos. Estos bloques representan componentes, sensores, y los servomotores superiores e inferiores, que forman juntos la estructura mecánica del robot.

En la parte superior, hay seis servomotores RX-28 (azul), tres para cada brazo (3 GDL) y dos servomotores RX-10 (azul) para la cabeza (2 GDL), mientras que en la parte inferior hay 13 servomotores RX-64 (verde): seis para cada pierna (6 GDL) y uno para la cintura (1 GDL).

Ya que los servomotores de la parte inferior requieren mayor torque (debido al peso que tienen que mover), un voltaje mayor necesita ser suministrado (línea verde), mientras que en la parte superior, el torque es mínimo y un voltaje menor (línea azul) puede ser suministrado.

Dos tipos de sensores (naranja) son utilizados para tener retroalimentación del mundo exterior: una webcam colocada en la cabeza y una brújula digital localizada dentro del pecho. Ésta última incluye sensores magnéticos y acelerómetros que envían información a la unidad de procesamiento para mantener el equilibrio del robot y permitirle saber si está acostado boca abajo o boca arriba.

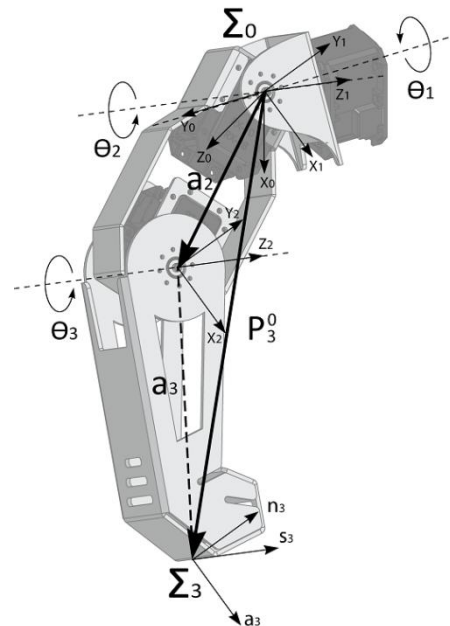


Fig. 5. Representación cinemática del brazo del Robot ZERO.

Asimismo, las instrucciones para los servomotores son enviadas a través del controlador de servomotores por la unidad de procesamiento. Los datos son enviados en conexión serial (línea negra) por los 21 servomotores, aún si algunos a la mitad de la cadena no se requieren mover. La fuente de alimentación principal del robot es tomada ya sea de las baterías o conectándolo directamente a la toma de corriente alterna (línea roja).

Análisis Mecánico

Para controlar el movimiento de ZERO, es importante analizar primero su cinemática. Esta sección presenta el modelo cinemático de los brazos y piernas de ZERO.

A) Modelo del brazo

La Fig. 4 muestra la estructura del brazo. Como se mencionó anteriormente, cada brazo es una estructura de 3 GDL accionada por tres servomotores: uno en el hombro, uno en el brazo y uno en el codo. Note que los motores que accionan el hombro y el brazo comparten un eje común de rotación.

En la fig. 4(a), la triada Σ_0 representa los ejes x_0, y_0, z_0 y es la referencia de coordenadas del brazo completo. Las triadas Σ_1, Σ_2 y Σ_3 son ejes locales de los eslabones. Juntas rotoidales conectan todos los eslabones. Sus rotaciones se denotan con θ_i . La longitud de cada eslabón se denota con a_i . El objetivo es determinar la posición y la orientación del efector final con respecto a Σ_0 . Eso se puede lograr estableciendo un modelo de cinemática directa que estime el vector resultante P_3^0 .

Inspirado en la cinemática de manipuladores robóticos, la posición P del efector final puede ser expresada en ambas Σ_0 y Σ_3 por la ec. (1):

$$\Sigma_0 = R \Sigma_3 \tag{1}$$

Donde R es la matriz de rotación que representa la transformación de coordenadas en la triada Σ_3 con respecto a las coordenadas de la triada Σ_0 . Para obtener R , se puede utilizar la matriz de transformación homogénea de la ec. (2):

$$A_n^{n-1}(\theta_n) = \begin{bmatrix} R_n^{n-1} & P_n^{n-1} \\ 0^T & 1 \end{bmatrix} \tag{2}$$

Donde A_n^{n-1} es la matriz de transformación desde el eslabón $n - 1$ hasta el eslabón n . La matriz A está en función de θ_n . R_n^{n-1} es una matriz de 3x3 que representa la rotación mientras P_n^{n-1} es una matriz de 3x1 que representa la traslación. Los parámetros de Denavit-Hartenberg para el brazo de ZERO se enlistan en la Tabla 2.

Eslabón	θ	d	a	α
1	0	0	0	$\pi/2$
2	0	0	a_2	0
3	0	0	a_3	0

Tabla 2. Parámetros de Denavit-Hartenberg para el brazo de ZERO.

Así, la ec. (2) se convierte en la ec. (3):

$$A_n^{n-1}(\theta_n) = \begin{bmatrix} C_n & -S_n C_{an} & S_n S_{an} & a_n C_n \\ S_n & C_n C_{an} & -C_n S_{an} & a_n S_n \\ 0 & S_n & C_{an} & d_n \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

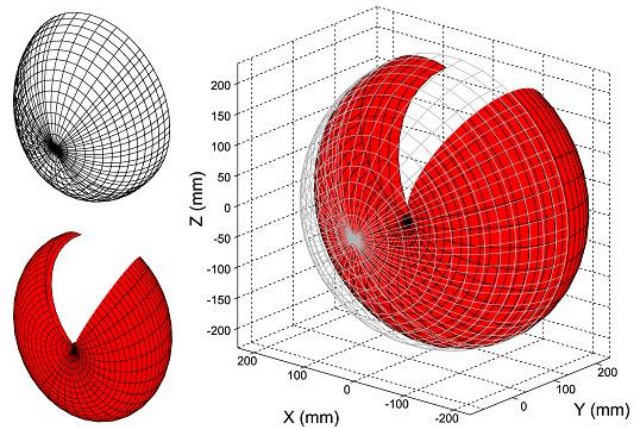


Fig. 6. Espacio de trabajo definido por el espacio entre los puntos límites máximos y mínimos del brazo. Figuras anexas: (arriba) superficie máxima y (abajo) mínima de alcance.

Donde $C_n = \cos(\theta_n)$, $S_n = \sin(\theta_n)$, $C_{an} = \cos(\alpha_n)$ y $S_{an} = \sin(\alpha_n)$.

Substituyendo los parámetros de la Tabla 2 en la ec. (3) da para $n = 1$:

$$A_1^0(\theta_1) = \begin{bmatrix} C_1 & 0 & S_1 & 0 \\ S_1 & 0 & -C_1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

y para $n = 2,3$:

$$A_n^{n-1}(\theta_n) = \begin{bmatrix} C_n & -S_n & 0 & a_n C_n \\ S_n & C_n & 0 & a_n S_n \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$T_3^0 = A_1^0 A_2^1 A_3^2 = \begin{bmatrix} C_1 C_{23} & -C_1 S_{23} & S_1 & C_1(a_2 C_2 + a_3 C_{23}) \\ S_1 C_{23} & -S_1 S_{23} & -C_1 & S_1(a_2 C_2 + a_3 C_{23}) \\ S_{23} & C_{23} & 0 & a_2 S_2 + a_3 S_{23} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

La cinemática directa T_3^0 se establece considerando las matrices de transformación consecutivas A_n^{n-1} (ec. (6)):

El vector P_3^0 está definido por los componentes de translación de la matriz T_3^0 . Éstos son (ec. (7)):

$$P_3^0 = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} C_1(a_2 C_2 + a_3 C_{23}) \\ S_1(a_2 C_2 + a_3 C_{23}) \\ a_2 S_2 + a_3 S_{23} \end{bmatrix} \quad (7)$$

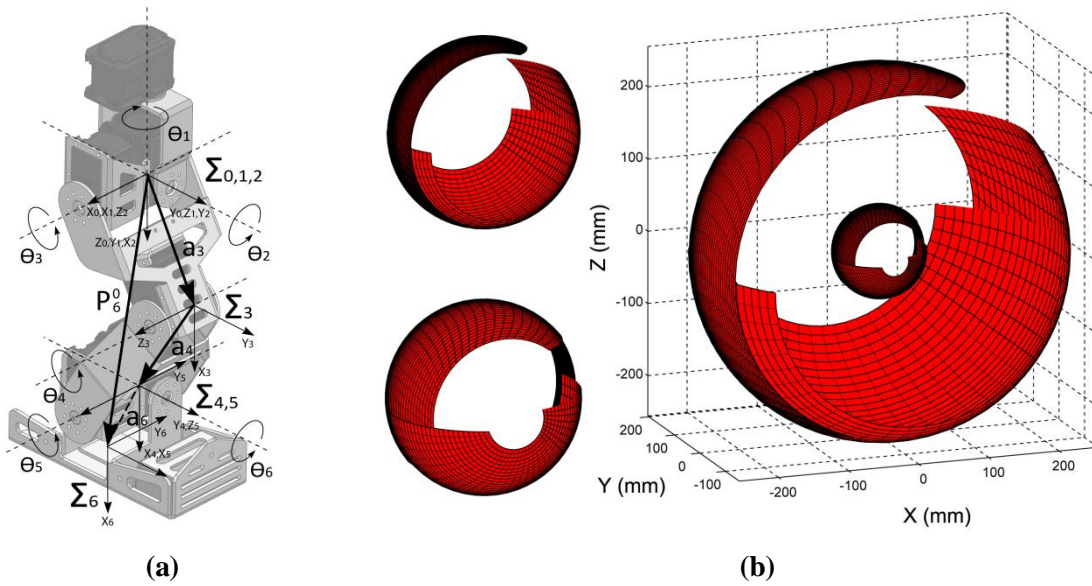


Fig. 7. Modelo de la pierna de ZERO: (a) representación cinemática y (b) el espacio de trabajo definido por el espacio entre los puntos límites máximos y mínimos de la pierna. Figuras anexas: (arriba) superficie máxima y (b) mínima de alcance.

La simulación de la ec. (7) provee una visualización del espacio de trabajo del brazo. La Fig. 6 muestra una gráfica en 3D del espacio de trabajo del brazo. Note que hay 2 gráficas anexas: la superior representa los puntos máximos que el brazo alcanza en el espacio, mientras que la inferior representa los mínimos. Esto es debido a una restricción de 300° de rotación en el servomotor del hombro. El espacio entre ambas gráficas es el espacio de trabajo efectivo.

B) Modelo de la pierna

La Fig. 7(a) muestra la estructura de una pierna. Cada pierna es una estructura de 6 GDL accionada por seis servomotores, creando las articulaciones de la cadera, ingle, muslo, rodilla, tobillo y pie. La pierna puede ser considerada un

manipulador esférico en las 3 articulaciones superiores y un manipulador antropomorfo en las 3 inferiores. Note que los motores que accionan la cadera, ingle y muslo comparten un eje de rotación en común, mientras que el tobillo y el pie comparten otro, obteniendo así una modelo cinemático simplificado con menos variables a resolver.

El objetivo es determinar la posición y la orientación de la triada Σ_6 en la planta del pie con respecto a Σ_0 en las coordenadas de referencia de la pierna entera. Como en el brazo, esto se puede lograr estableciendo un modelo cinemático directo que estime el vector resultante P_6^0 expresado en ambas triadas Σ_0 y Σ_6 (ec. (8)).

$$\Sigma_0 = R \Sigma_6 \quad (8) \quad \text{con:}$$

Para obtener R, la matriz de transformación homogénea A_n^{n-1} descrita en la ecuación (2) puede ser considerada de nuevo. Los parámetros de Denavit-Hartenberg para la pierna de ZERO se enlistan en la Tabla 3.

Eslabón	θ	d	a	α
1	0	0	0	$\pi/2$
2	$\pi/2$	0	0	$-\pi/2$
3	0	0	$-a_3$	0
4	0	0	$-a_4$	0
5	0	0	0	$\pi/2$
6	0	0	$-a_6$	0

Tabla 3. Parámetros de Denavit-Hartenberg para la pierna de ZERO.

Substituyendo los parámetros de la Tabla 3 en la ec. (3) da para $n = 1,5$:

$$A_n^{n-1}(\theta_n) = \begin{bmatrix} C_n & 0 & S_n & 0 \\ S_n & 0 & -C_n & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

para $n = 2$:

$$A_2^0(\theta_2) = \begin{bmatrix} C_n & 0 & -S_n & 0 \\ S_n & 0 & C_n & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10)$$

y para $n = 3,4,6$:

$$A_n^{n-1}(\theta_n) = \begin{bmatrix} C_n & -S_n & 0 & -a_n C_n \\ S_n & C_n & 0 & -a_n S_n \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11)$$

La cinemática directa para la parte superior de la pierna (T_3^0) se define entonces por la ec. (12):

$$T_3^0 = A_1^0 A_2^1 A_3^2 = \begin{bmatrix} U_{11} & U_{12} & U_{13} & U_{14} \\ U_{21} & U_{22} & U_{23} & U_{24} \\ U_{31} & U_{32} & U_{33} & U_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (12)$$

$$\begin{aligned} U_{11} &= C_1 C_2 C_3 - S_1 S_3 \\ U_{12} &= -C_1 C_2 C_3 - S_1 S_3 \\ U_{13} &= -C_1 S_2 \\ U_{14} &= a_3 (S_1 S_3 - C_1 C_2 C_3) \\ U_{21} &= S_1 C_2 C_3 + C_1 S_3 \\ U_{22} &= -S_1 C_2 C_3 - C_1 C_3 \\ U_{23} &= -S_1 S_2 \\ U_{24} &= a_3 (C_1 S_3 + S_1 C_2 C_3) \\ U_{31} &= S_3 C_3 \\ U_{32} &= S_2 S_3 \\ U_{33} &= C_2 \\ U_{34} &= a_3 S_2 C_3 \end{aligned}$$

Mientras que la cinemática directa de la parte inferior de la pierna (T_6^3) se define por la ec. (13):

$$T_6^3 = A_4^3 A_5^4 A_6^5 = \begin{bmatrix} L_{11} & L_{12} & L_{13} & L_{14} \\ L_{21} & L_{22} & L_{23} & L_{24} \\ S_6 & C_6 & 0 & -a_6 S_6 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (13)$$

con:

$$\begin{aligned} L_{11} &= C_6 (C_4 C_5 - S_4 S_5) \\ L_{12} &= -S_6 (C_4 C_5 - S_4 S_5) \\ L_{13} &= C_4 S_5 - C_5 S_4 \\ L_{14} &= -a_4 C_4 - a_6 C_6 (C_4 C_5 - S_4 S_5) \\ L_{21} &= C_6 (C_4 S_5 + S_4 C_5) \\ L_{22} &= -S_6 (C_4 S_5 + S_4 C_5) \\ L_{23} &= S_4 S_5 + C_4 C_5 \\ L_{24} &= -a_4 S_4 - a_6 C_6 (C_4 S_5 - S_4 C_5) \end{aligned}$$

Ya que la pierna entera puede ser considerada como un doble manipulador (un manipulador esférico en la parte superior y un manipulador antropomorfo en la parte inferior), la cinemática directa de la pierna se define entonces por la combinación de las ecs. (12) y (13):

$$T_6^3 = T_3^0 T_6^3 \quad (14)$$

La solución de la ec. (14) es de hecho difícil de expresar. Sin embargo, como en el brazo, el vector P_6^0 está definido por los componentes de la matriz de translación T_6^0 . Éstos son (ec. (15)):

$$P_6^0 = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} T_{6\ 14}^0 \\ T_{6\ 24}^0 \\ T_{6\ 34}^0 \end{bmatrix} \quad (15)$$

La simulación de la ec. (15) provee una visualización del espacio de trabajo de la pierna. La Fig. 7(b) muestra una gráfica en 3D del espacio de trabajo de la pierna. Similarmente, hay 2 gráficas agregadas: la superior representa los puntos máximos que la pierna alcanza en el espacio, mientras que la inferior representa los mínimos. Los puntos entre ambas gráficas definen el espacio de trabajo efectivo.

Es posible visualizar el espacio de trabajo total de ZERO simulando las ecs. (7) y (15) en diferentes referencias de coordenadas y traslapando las gráficas resultantes. La Fig. 8 muestra una gráfica 3D del espacio de trabajo total para las cuatro extremidades del robot humanoide ZERO. Las gráficas internas en color sólido, son los límites de los puntos mínimos alcanzables de los brazos y piernas, mientras que las gráficas externas con transparencias, son los puntos máximos alcanzables para las extremidades superiores e inferiores de ZERO. El espacio entre las cuatro gráficas juntas define el espacio de trabajo efectivo del robot.

III. Conclusión

En este artículo, se presentó el robot humanoide ZERO, una plataforma propia desarrollada para investigación de la marcha bípeda.

ZERO es un robot de 21 GDL, cuidadosamente manufacturado, capaz de desarrollar un amplio número de movimientos. ZERO incluye una poderosa unidad de procesamiento a bordo, baterías y un grupo de sensores que serán explotados en un futuro cercano para balancearse de pie, control de la marcha y movimiento autónomo.

Este artículo también presentó detalles de su cinemática. Inspirados en manipuladores robóticos, se han establecido modelos de cinemática directa para determinar los espacios de trabajo de los brazos y las piernas. Estos

modelos serán utilizados para establecer esquemas de control para el posicionamiento preciso del robot.

IV. Referencias

- [1] S. Kagami, K. Nishiwaki, J. Kuffner, Y. Kuniyoshi, M. Inaba, and H. Inoue, "Online 3D vision, motion planning and bipedal locomotion control coupling system of humanoid robot: H7", *Memorias de: IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp 2557-2563, 2002.
- [2] K. Loffler, M. Gienger, and F. Pfeiffer, "Sensor and control design of a dynamically stable biped robot", *Memorias de: IEEE International Conference on*

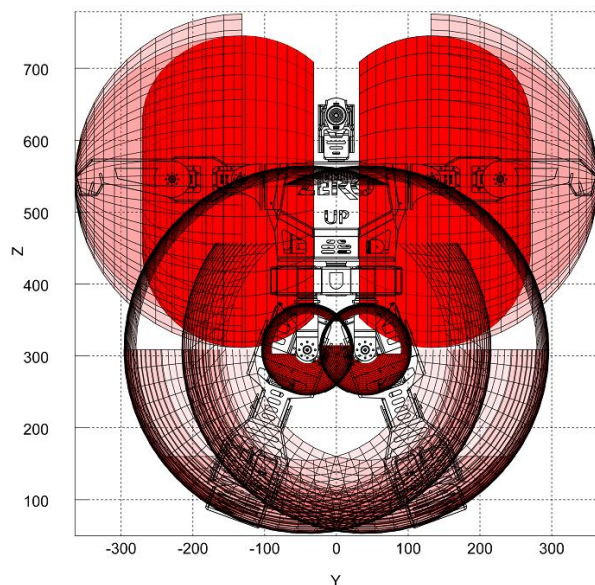


Fig. 8. Espacio de trabajo total de ZERO

- Robotics and Automation, pp 484-490, 2003.
- [3] Y. Hiroki, T. Ishida, J. Yamaguchi, M. Fujita, and T. Doi, "A small biped entertainment robot", *Memorias de: IEEE International Conference on Humanoid Robots*, pp 181-186, 2001.
- [4] K. Kaneko, F. Kanehiro, S. Kajita, H. Hirukawa, T. Kawasaki, M. Hirata, K. Akachi, and T. Isozumi, "Humanoid robot HRP-2", *Memorias de: IEEE International*

Conference on Robotics and Automation, pp 1083-1090, 2004.

- [5] K. Hirai, M. Hirose, Y. Haikawa, and T. Takenaka, “The development of Honda humanoid robot”, Memorias de: IEEE International Conference on Robotics and Automation, pp 1321-1326, 1998.
- [6] Kondo Kagaku Co. Información actualizada disponible en: <http://www.kondo-robot.com/>
- [7] Aldebaran Robotics. Información actualizada disponible en: <http://www.aldebaran-robotics.com/>
- [8] Innovati Inc. Información actualizada disponible en: <http://en.innovati.com.tw/>
- [9] Dynamixel Robotis. Información actualizada disponible en: <http://www.robotis.com/xe/dynamixel>
- [10] CompuLab. Información actualizada disponible en: <http://www.fit-pc.com/we>

V. Autores

Ing. Efraín Hernández Franco es Ingeniero en Mecatrónica por la Universidad Panamericana campus Guadalajara. Actualmente se encuentra cursando la maestría en ciencias en la Universidad Panamericana de Aguascalientes.

Dr. Ramiro Velázquez Guerrero es Ingeniero en Electrónica por la Universidad Panamericana, Maestro en Ciencias por el INSA de Lyon (Francia) y Doctor en Robótica por la Universidad Pierre et Marie Curie (Francia). Actualmente es profesor-investigador en la Universidad Panamericana de Aguascalientes.

RS232 para aplicaciones y conectores para manejar hardware externo.

En este trabajo, se desarrolló un sistema para manipular un brazo robótico comercial, comandándolo por medio de un teclado de computadora estándar, dado que la tarjeta Nexys2 incluye un puerto PS/2 [5]. El sistema resultante se compone de varios módulos, ya que se le incorporó la capacidad de guardar movimientos, con su correspondiente duración, para posteriormente reproducirlos en forma automática.

Se utilizó el brazo desarrollado y distribuido por la empresa Steren, con número de parte K-682 [6], cuya imagen se muestra en la figura 2.



Fig. 2. Brazo robótico a manipular.

El brazo incluye 5 motores, 4 de movimiento y 1 para la apertura y cierre de unas pinzas. Sin embargo, únicamente se manipularon 4 motores, 3 de movimiento y el de las pinzas, esto porque con 3 motores es suficiente para alcanzar cualquier posición en un espacio finito [7].

El control alámbrico incluido con el brazo se basa en interruptores, por lo que se desarrolló una etapa de acondicionamiento para que pudiera ser manipulado desde la tarjeta Nexys2.

II. Desarrollo

El sistema que se diseñó tiene la capacidad de controlar libremente cada uno de los 4 motores que le proporcionan movimiento al brazo; para esto se utilizan ocho teclas diferentes desde un teclado PS/2, dos para cada motor; también se emplea una tecla más, para iniciar el grabado de

una secuencia y para su finalización e inicio de la reproducción.

III. Módulos del Sistema

El sistema se desarrolló bajo un enfoque modular [8], proponiendo un diagrama a bloques con los módulos necesarios, para luego codificar a cada uno con VHDL. En la figura 3 se muestra la propuesta modular del sistema; la función de cada módulo se describe a continuación.

1. El **Divisor** genera un “reloj” con una frecuencia inferior a la de la tarjeta; esto es porque la frecuencia de la tarjeta es de 50 MHz, la cual resulta inadecuada para un sistema de esta naturaleza, dado que el teclado opera en frecuencias entre 20 y 30 KHz.
2. El módulo **Teclado** es el responsable de hacer la comunicación entre la tarjeta y el teclado de la computadora. Lee la trama enviada por éste y entrega un código de rastreo por cada tecla presionada. Es importante mencionar que el módulo mantiene el código de rastreo de la última tecla presionada, aunque ya no lo esté. Por ello, también genera una señalización de cuando hay un nuevo código de rastreo (CR_Listo).
3. El módulo **deco_comp** acepta el código de rastreo que le entrega el módulo del **teclado** e identifica si es una de las teclas asignadas para el movimiento del brazo, de ser así, lo que hace es transformar este código en uno que se encargará de mover el brazo. También detecta si se ha presionado la tecla que indica el inicio de la grabación de movimientos, para ello se destinó a la tecla G.
4. El módulo **salida_a_motor** acepta el código que le entrega el módulo **deco_comp** y genera las salidas que mueven a los motores del brazo. También acepta los códigos provenientes de la memoria, donde se almacenará la secuencia a ejecutar, la decisión de cual código ejecutar se toma con base en las entradas de control: act1 y act2. Siendo act2 de mayor jerarquía, pues indica

que se tiene que ejecutar la secuencia grabada, es decir, mientras act2 esté en alto, no importa que se presione otra tecla de

activación de motores, puesto que será ignorada.

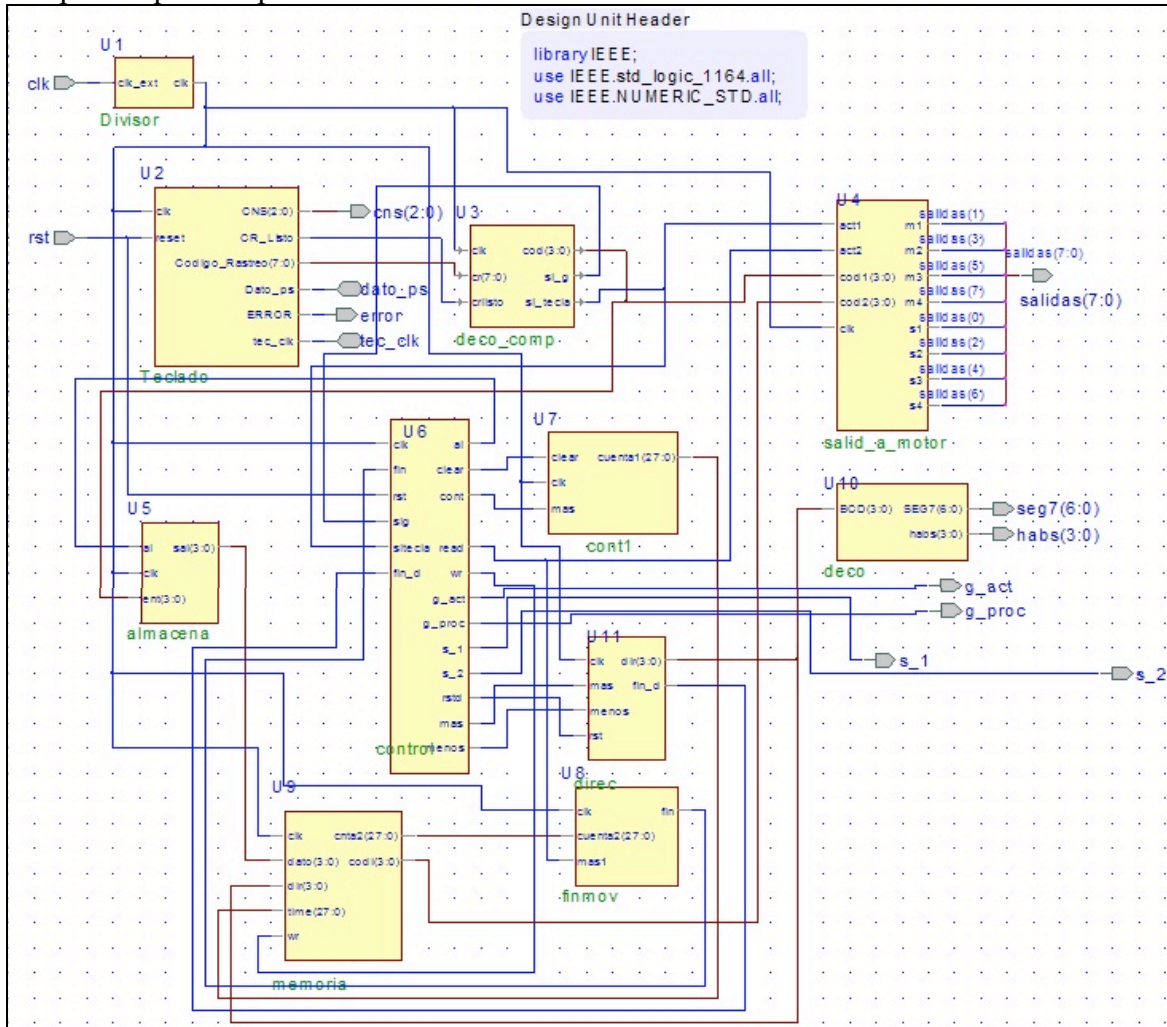


Fig. 3. Diagrama a bloques del sistema para la manipulación de un brazo.

5. El módulo **almacena** se encarga de asegurar que el código que se guarda en la memoria es un código de movimiento, y no un código en blanco; es necesario este módulo por los tiempos perdidos en la máquina de estados.
6. El módulo **control**, como su nombre lo indica, es el responsable de comandar a los demás módulos del sistema.
7. El módulo **cont1** cuenta el tiempo que se mantiene presionada una tecla de activación de motores.
8. El módulo **fin_mov** es empleado durante la reproducción de movimientos. Del módulo **memoria** obtiene una duración registrada y con base en un contador determina si ya se alcanzó el tiempo recibido, indicándolo con la activación de su salida.
9. En el módulo **memoria** se guardarán claves que identifican a los motores y los sentidos de movimiento, siempre que se haya activado la opción de grabar movimientos. También se guardará en la misma dirección, pero en un espacio diferente, el tiempo que duró la

activación de esta tecla. Estos datos los obtiene de los módulos **almacena** y **cont1**.

10. El módulo **deco** es el manejador de los displays de 7 segmentos incluidos en la tarjeta. Básicamente se emplea para mostrar la dirección de acceso en la memoria, para que el usuario conozca el movimiento que se está grabando o reproduciendo.
11. El módulo **direc** es el contador responsable de la dirección de la memoria a la que se va a acceder.

IV. Control del Sistema

El control es el módulo principal del sistema, ya que es quien coordina la operación de los demás módulos. El control se basa en una máquina de estados y su operación principal la realiza mientras graba o recupera una secuencia de movimientos. En las figura 4 se muestra el comportamiento del control, en donde puede notarse un estado jerárquico, el cuál es expandido en la figura 5.

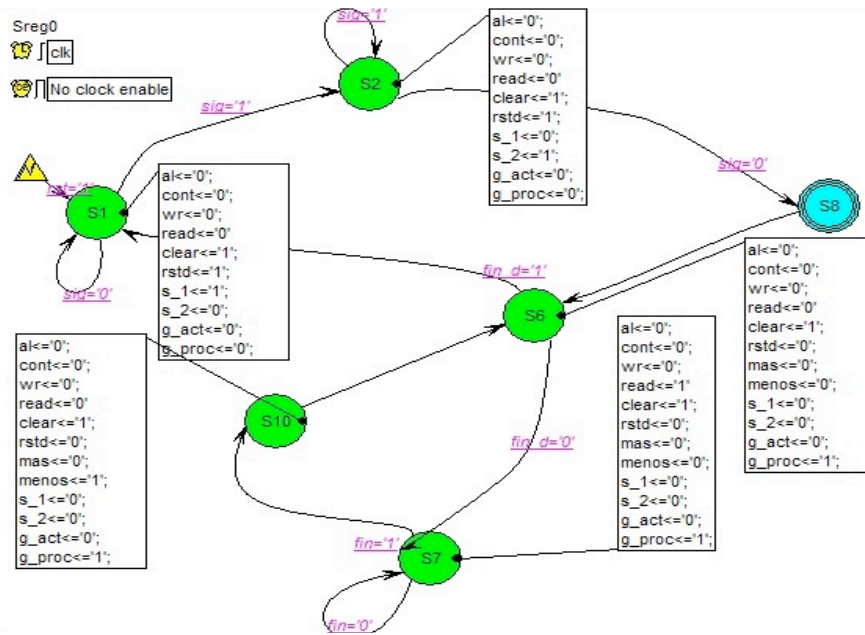


Fig. 4. Máquina de Estados del Control.

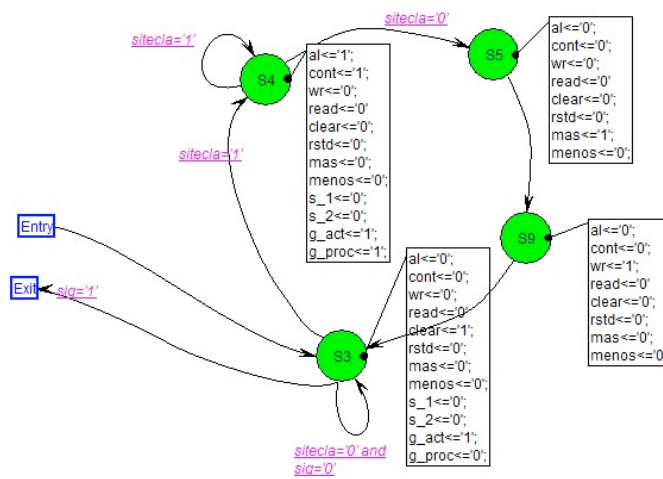


Fig. 5. Estados que conforman al Estado Jerárquico de la figura 4.

La grabación de la secuencia inicia y termina con la tecla G, mientras el usuario no presione esa tecla, el control se mantendrá en el estado 1, en el cual se han acondicionado sus salidas para que las acciones en las teclas de los movimientos se vean reflejadas directamente en el brazo robótico (la activación de la tecla G se indica con la señal sig).

En el estado 2 el sistema espera a que ya no esté presionada la tecla G. En los 4 estados que conforman al estado jerárquico, el control generará las señales para ir grabando la secuencia de movimientos. La grabación también concluye con la tecla G.

En los últimos 3 estados, después del estado jerárquico, el sistema iniciará con la reproducción de movimientos, la cual se realiza en el orden inverso en que se fueron generando.

V. Acondicionamiento para el Manejo del Robot

Las salidas del módulo **salida_a_motor** están directamente conectadas a los pines de uno de los conectores de la tarjeta Nexys2, por lo tanto, es necesario desarrollar una etapa de acoplamiento. En la figura 6 se puede ver como el control alámbrico incluido básicamente cierra unos interruptores, con lo que suministrará voltaje a los motores. Los interruptores serán remplazados por puentes H, para que los motores sean manejados desde la tarjeta Nexys2.

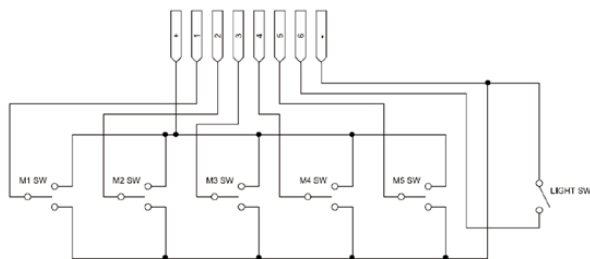


Fig. 6. Control alámbrico del Brazo Robótico.

La etapa de acondicionamiento se implementó con base en Puentes H integrados, modelo L293, en la figura 7 puede verse su diseño [9], en donde se observa que se incluyeron puentes de diodos como protección [10].

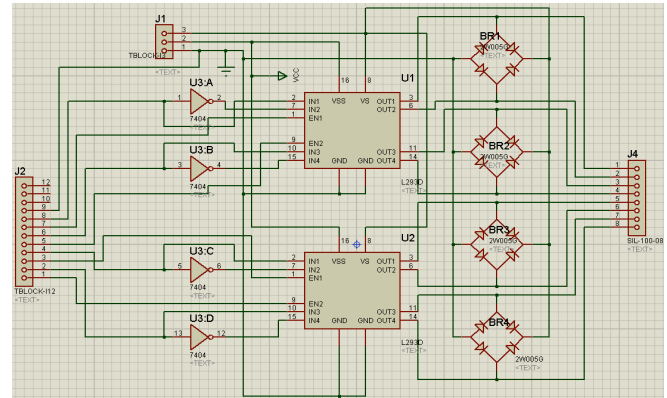


Fig. 7. Etapa de acoplamiento para el manejo de los motores.

VI. Resultados

El sistema cumplió los requerimientos planteados y además, resultó muy robusto. Las grabaciones y reproducciones de secuencias funcionaron adecuadamente.

En la figura 8 se exhibe al brazo conectado con la tarjeta Nexys2, por medio de la etapa de acoplamiento. Además de mostrar al teclado de computadora conectado en el puerto PS/2 de la misma tarjeta.



Fig. 8. Sistema implementado.

En la figura 9 se pueden ver detalles en las conexiones, mostrando a la tarjeta Nexys2 completa, para ilustrar como en los displays de 7 segmentos aparece la dirección de acceso. Se conectó un teclado marca Dell, aunque no hay restricciones, se podría conectar un teclado de cualquier marca.

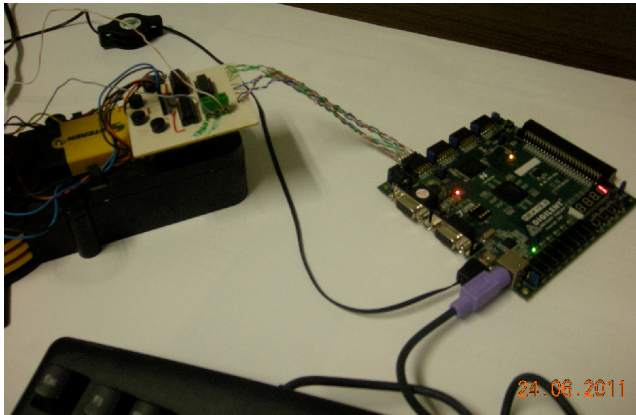


Fig. 9. Conexiones en la tarjeta Nexys2.

En la figura 10 está el primer nivel del modelo de hardware, generado por la herramienta ISE de Xilinx al momento de realizar la implementación. Corresponde con la interfaz del sistema, las señales `dato_ps` y `tec_clk` no son salidas, realmente son señales bidireccionales.

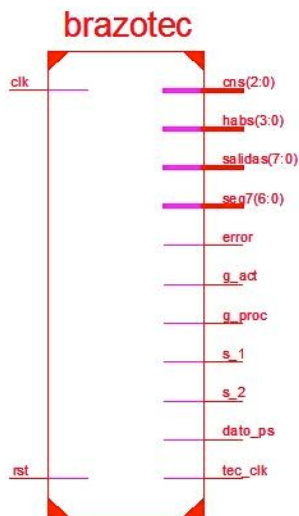


Fig. 10. Modelo de hardware resultante.

En la tabla 1 se muestra un resumen con los recursos empleados, se observa una amplia disposición de recursos, para complementar la funcionalidad del sistema.

VII. Conclusiones

En este proyecto se usó un módulo ya elaborado con anterioridad, el controlador de un teclado de computadora [11]. En el diseño de sistemas basados en VHDL es común y conveniente reutilizar módulos, como en este caso.

Recurso	Usado	Disponibile	Uso
Slices	205	4, 656	4 %
Bloques I/O	31	232	13 %
Bloques RAM	0	20	0 %
GCLKs	2	24	8 %

Tabla 1. Resumen de los recursos.

Para el correcto funcionamiento de todos los módulos fue preciso tomar en consideración la operación del módulo del teclado, ya que como se explicó con anterioridad, mantiene el código de rastreo aunque la tecla ya no esté presionada.

Para ello, se realizó un cambio en la máquina de estados que afectó al módulo de memoria: cuando se reconoce un código de movimiento y éste se va a grabar, primero se incrementa la dirección de memoria y posteriormente se guarda el dato en esa localidad. De la misma manera, al momento de leer la memoria, primero se realiza la lectura y posteriormente se disminuye la dirección. Se hace la comparación con la dirección de inicio de la memoria, y de haber una coincidencia, se da por terminada la secuencia grabada, de lo contrario se repetirá el ciclo. Esto garantiza un correcto funcionamiento de la máquina de estados, que además está sincronizada con el módulo que lee el teclado.

Para trabajos futuros se podría proponer el guardar varias secuencias de movimientos en la memoria y tener un menú para decidir cuál de las secuencias se ha de ejecutar. Por ejemplo, una secuencia puede mover un objeto de un lugar a otro, mientras que una secuencia diferente puede hacer que el brazo regrese a la posición inicial, para tomar nuevamente otro objeto.

También sería interesante acondicionar un sistema de sensores para conocer la posición del brazo en tiempo real, y así poder acondicionar algún controlador automático tradicional, aunque a este respecto, primero sería necesario introducir en la tarjeta ciertos módulos de operaciones matemáticas para poder llevar a cabo esta tarea.

VIII. Referencias

- [1] S. Brown, Z. Vranesic, “Fundamentos de Lógica Digital con Diseño VHDL”. 2ª Ed., 2006, Editorial McGraw-Hill.
- [2] F. Pardo, J. A. Boluda, “VHDL, Lenguaje para síntesis y modelado de circuitos”, Ed. Alfaomega, 2000
- [3] Spartan-3E FPGA Family: Data Sheet, http://www.xilinx.com/support/documentation/data_sheets/ds312.pdf, última visita: Junio de 2011.
- [4] Digilent Inc. - Digital Design Engineer's Source, <http://www.digilentinc.com/>, última visita: Junio de 2011.
- [5] Digilent Nexys2 Board Reference Manual, Doc: 502-107, Copyright Digilent, Inc., June 21, 2008.
- [6] Brazo Mecánico con Control Remoto Alámbrico, modelo K-682, Manual de Instrucciones V0608, <http://www2.steren.com.mx/doctosMX//K-682-instr.pdf>, última visita: Junio de 2011.
- [7] R. Kelly, V. Santibáñez and A. Loría, “Control of Robot Manipulators in Joint Space”, Ed. Springer-Verlag, 2005.
- [8] R. Romero Troncoso, “Sistemas Digitales con VHDL”; Legaria Ediciones.
- [9] P. Miles, T. Carroll, “Build Your Own Combat Robot”, Ed. McGraw-Hill/Osborne, 2002.
- [10] A. P. Malvino, “Principios de Electronica”. 6ª Ed., 1999, Editorial McGraw-Hill.
- [11] J. C. Tepozán Ríos, “Diseño de controladores para la tarjeta de desarrollo XSA-100”. Tesis de Ingeniería, Universidad Tecnológica de la Mixteca, Mayo de 2004.

IX. Autores

M. C. Felipe Santiago Espinosa es Maestro en Ciencias con especialidad en Electrónica por parte del INAOE, Licenciado en Electrónica por parte de la BUAP, desde 1998 labora como Profesor –Investigador en la Universidad Tecnológica de la Mixteca, adscrito al instituto de Electrónica y Mecatrónica.

Yukio Rosales Luengas es estudiante de Ingeniería Mecatrónica en la Universidad Tecnológica de la Mixteca.

Alejandro Rivera García es estudiante de Ingeniería Mecatrónica en la Universidad Tecnológica de la Mixteca.

Adrián Humberto Martínez Vásquez es estudiante de Ingeniería Mecatrónica en la Universidad Tecnológica de la Mixteca.

Aplicación de Robots Industriales Coordinados

F. Chávez-Montejano, Carlos Pizano-Martínez, J. Gudiño-Lau y Abel Delino-Silva.

*Facultad de Ingeniería Electromecánica, Universidad de Colima Km. 20.5 Carretera Manzanillo-Cihuatlán,
Tel. (314) 3311207, jglau@uacol.mx.

Resumen

Este artículo muestra una aplicación con dos robots industriales, que consiste en la manipulación de dos brazos robot para destapar un refresco y servir el contenido en un vaso, esta actividad resulta muy cotidiana para la mayoría de las personas, sin embargo hacerlo de manera automática no es una labor sencilla, requiere de una sincronización de los robots para que puedan trabajar coordinadamente y evitar posibles colisiones. Es una introducción del proyecto de investigación “Control de Robots cooperativos”.

Palabras Clave: Robots coordinados, control punto a punto, planeación de trayectorias.

I. Introducción

El empleo de robots manipuladores permite elevar la calidad de los procesos industriales que requieren de gran precisión y ser repetitivos, por lo que resulta demasiado complicado para una persona.

En los últimos años, la mayoría de las investigaciones estaban enfocadas a mejorar los controladores de robot manipuladores y móviles individuales, sin embargo en los años recientes ha tenido especial atención la parte de la coordinación y cooperación de los robots, esto con la finalidad de realizar principalmente tareas más complejas.

Una de las razones es que este tipo de sistemas poseen alta capacidad de manipulación de movimientos finos y flexibilidad en tareas de ensamblado, manufactura, ambientes de trabajo complejos, manipulación de objetos, etc. En particular, cuando un objeto es demasiado grande o pesado, no es suficiente un sólo robot para manipularlo; por esta razón se emplean dos o más robots.

Existen muchas labores industriales en las que pueden ser llevadas a cabo por un solo robot, es conveniente en algunos casos el empleo de dos o más manipuladores para realizar una tarea específica. En el presente trabajo se habla de

robots coordinados a un grupo de robots realizando una tarea específica. La coordinación de robots consiste que un grupo de robots realicen una tarea específica sobre un objeto sin que lo sujeten simultáneamente.

A los manipuladores que no interaccionan con el ambiente se le conoce como robots en movimiento libre, y pueden ser modelados por un conjunto de ecuaciones diferenciales ordinarias no lineales. El caso contrario son los que interaccionan con el ambiente y se conocen como robots en movimiento restringido, los cuales pueden ser modelados por un conjunto de ecuaciones diferenciales algebraicas no lineales. Una gran parte de los manipuladores industriales se encuentran en este grupo [1, 2].

Para el caso de sistemas cooperativos no sólo se tiene una restricción, sino un conjunto de restricciones, las cuales describen la sujeción del objeto en cada punto de contacto y son impuestas a nivel cinemático sobre cada uno de los manipuladores. En conjunto se tiene un sistema de ecuaciones diferenciales algebraicas no lineales. [5, 6]

Por otro lado, es importante hacer notar que la teoría que se emplea para robots coordinados es independiente del tamaño de los mismos, por lo que se puede incluir en esta área el estudio de manos mecánicas, cuyos dedos pueden ser

considerados como grupos de robots en miniatura. El estudio de manos mecánicas es importante no sólo porque estas pueden ser utilizadas como prótesis en seres humanos, sino también porque aumentan considerablemente la capacidad de manipulación de un robot al sustituir la pinza que poseen la mayoría de ellos por una mano con dos o más dedos. [3, 4]

Este trabajo tiene como propósito mostrar la coordinación de dos robots industriales con una aplicación, es un proyecto de investigación que la Facultad actualmente está iniciando, que consiste en sustituir el controlador de fábrica por una tarjeta de adquisición de datos, es decir convertir el sistema actual de arquitectura cerrada a una arquitectura abierta; así poder controlar ambos manipuladores con un solo controlador.

II. Desarrollo

Existe una gran diversidad de aplicaciones que podemos desarrollar con dos o más robots, estas actividades las podemos adecuar a nuestras necesidades dependiendo de la tarea a realizar. Los robots coordinados requieren de un grado de sincronización elevado, para evitar las posibles colisiones. Por lo cual a continuación se describe un ejemplo de una actividad coordinada con robots. La Fig. 1 muestra la coordinación de dos robots manipuladores industriales.



Fig. 1. Coordinación de robots.

Un problema que debe ser considerado cuando se habla de robots coordinados es la planeación de

las trayectorias para cada uno de los robots, para evitar las colisiones, la planeación de la trayectoria consiste en programar un conjunto puntos para cada manipulador, se empleó el software del fabricante.

En la Fig. 2 se indica la tarea a realizar por parte de los robots de manera coordinada, dicho proceso trata de imitar a un ser humano al servir un refresco, como se puede imaginar esto resulta complicado poder hacerlo con un sólo brazo.

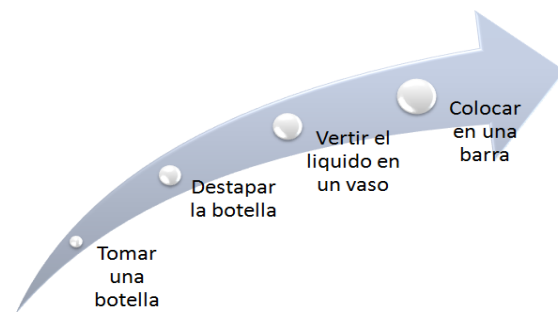


Fig. 2. Diagrama de la tarea a realizar.

Elaboración del efector final

Una vez definida la tarea a realizar, se diseñan y se elaboran dos efectores finales para cada robot manipulador, de acuerdo a la aplicación. Para este caso se requiere sujetar la botella como se observa en la Fig. 3(a) y se adiciona material poroso y suave para aumentar la fricción y disminuir los riesgos de daño de la botella de cristal. La Fig. 3(b) se muestra el efector final con dos aplicaciones, que consiste en un destapador y un sujetador de vasos, los dos efectores finales acoplados a cada robot, fueron diseñados y realizados en las instalaciones de la Facultad de Ingeniería Electromecánica de la Universidad de Colima, por los propios estudiantes. La apertura y cierre de las pinzas son controlados neumáticamente, no tiene pasos intermedios, es un control on-off.



Fig. 1(a) Destapador y sujetador de vaso



Fig. 2 (b) Sujetador de botella

Debido a que este es un proyecto de investigación que todavía está en el proceso de recopilación de información, no es manipulado por un sólo controlador. Para la coordinación empleando los controladores y programación de fábrica, es indispensable la comunicación entre los dos robots manipuladores industriales, se requiere un bus de datos que comunica a cada controlador, lo que permite controlar los tiempos, así como las señales digitales que darán la pausa para que un robot espere instrucción del otro y poder seguir la ejecución de sus movimientos.

III. Descripción del equipo experimental

La Facultad de Ingeniería Electromecánica tiene un equipo experimental, que consiste en dos robots manipuladores industriales de la firma alemana KUKA KR-16, cada robot tiene su propia unidad de control con una computadora con Windows xp embedded, un panel de control

KCP, un bus de datos que comunica a cada control y un software KSS V5.0.

Cada robot tiene seis grados de libertad y un efector final con capacidad de carga de 16 Kg, tiene un alcance de 1,611mm, la repetitividad es de $\pm 0,05$ mm, cuenta con resolvers como sensores de posición siendo estos de una mayor precisión. En la Fig. 4 se muestra el equipo experimental empleado para la aplicación, con sus respectivos efectores finales. El proyecto de investigación consiste en coordinar los robots industriales con un sólo controlador, en el que se pueda implementar controladores no lineales y planeación de trayectorias.



Fig. 4. Equipo experimental.

IV. Programación del sistema coordinado

Los robots industriales KUKA KR-16 con los que se trabaja en este proyecto tienen un tipo de programación a la cual se le conoce como programación punto a punto, que consiste en grabar una serie de posiciones de tal manera que formen una trayectoria deseada y el robot puede seguirlos para realizar una tarea y de esta manera culminar un proceso.

En la Tabla 1 se describen los comandos o tipos de movimientos empleados para esta coordinación, donde describe los tipos de movimientos que los manipuladores realizan.

Movimiento	Descripción
PTP	Punto a Punto el cual el robot realiza el mejor movimiento que a él le convenga para llegar de un punto a otro, sin importar seguir un patrón.
LIN	Este es un movimiento <i>lineal</i> lo cual ocasiona que el robot mantenga una línea recta de un punto a otro, realizando las maniobras necesarias para que el efector final mantenga la posición en que se encuentra.
CIR	Con este comando se pueden obtener movimientos de manera circular, permitiendo poder evadir obstáculos.

Tabla 1. Comandos empleados en la tarea.

La correcta coordinación y cooperación de este par de robots se debe a un sistema de comunicación entre los dos manipuladores, es por ello que se realiza un enlace entre los módulos de entradas/salidas digitales, para evitar colisiones de los robots o del entorno.

El algoritmo de programación empleado es, que el primer manipulador realiza su movimiento hasta la posición deseada, después envía una señal (salida digital) al segundo manipulador que ya terminó su tarea, y lo mismo sucede con el segundo manipulador, hasta que termine su tarea manda una señal al primer manipulador; y así sucesivamente, esto debido a que cada robot tiene su propio controlador.

La unidad de control del robot puede administrar un máximo de 4096 de entradas digitales y 4096 salidas digitales [3], mediante la cual se puede organizar una lógica tan extensa como sea posible, y así mantener la seguridad de los robots y de las personas. En la Fig. 5 se ilustra el puerto de comunicación entre los 2 robots.

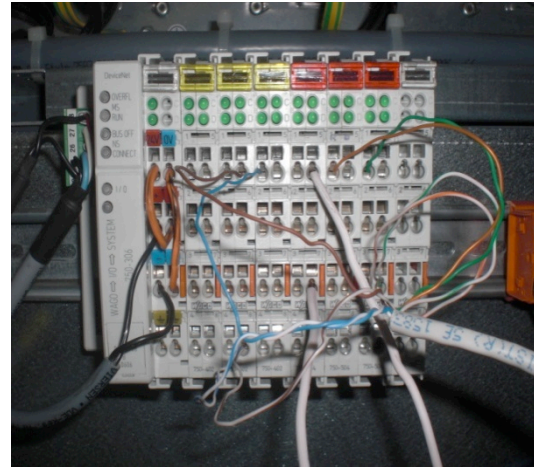


Fig. 5. Bus de conexiones de entradas/salidas digitales.

Al momento de trabajar con dos o más robots manipuladores se vuelve complicado debido a establecer ciertas condiciones que puedan dar los tiempos necesarios para evitar choques o accidentes entre ellos, para evitar las posibles colisiones se mandan señales digitales de un controlador a otro, esto con la finalidad de hacer que los robots dependan uno del otro, lo que pareciera que se dan ordenes entre sí, esto se evita al tener un solo controlador, esto es parte del proyecto de investigación que se trabaja actualmente en la Facultad.

Para iniciar con la programación es necesario conocer un comando fundamental, el cual un robot espera una señal del otro para continuar ejecutando sus movimientos, en la Tabla 2 se describen algunos comandos [3].

Comando	Descripción
WAIT FOR	Este comando permite decirle al robot que espere una señal para poder continuar.
WAIT	Permite dar tiempos entre un movimiento y otro dentro de un mismo programa.

Tabla 2. Comandos empleados en la coordinación.

Con esto se vuelve más fácil poder trabajar de manera coordinada con los dos robots industriales, debido que se crea una dependencia entre ellos, uno del otro. Por otra parte, son

indispensables para evitar cualquier tipo de colisión o accidente entre ellos, por esta razón es importante conocer los comandos mostrados en la Tabla 2.

Por otra parte, es importante declarar las herramientas para los efectores finales, ya que permite mover fácilmente el robot respecto al efector final, con la finalidad de precisar los movimientos en el programa, para destapar la botella. Se trabaja dentro de las coordenadas globales del robot, un robot es el maestro y el otro esclavo.

Una vez declaradas las herramientas en el controlador, se inicia la programación que consiste en grabar los movimientos necesarios para poder sujetar la botella y llevarla a una posición donde el otro robot pueda llegar a destapar, el otro robot graba sus movimientos necesarios para llegar y retirar la ficha de la botella. En esta parte se programa el comando WAIT FOR mencionado en la Tabla 1 para mandar que el robot que va a destapar la botella no realice ningún movimiento hasta que el otro termine la tarea asignada destapar la botella.

Una vez que el robot destape la botella, este realiza los movimientos para ir a sujetar un vaso y llevarlo a una posición donde el otro robot vacía el líquido, para que el otro robot inicie este proceso primero tiene que recibir una señal donde el otro robot le dice que ya tiene listo el vaso y pueda continuar.

Una vez concluido este proceso se graba los pasos para que un robot retire la botella y el otro coloque el vaso con el refresco en una mesa y el humano pueda deleitar de su refresco.

V. Resultados experimentales

Este es un estado del arte del proyecto de investigación que la Facultad desarrolla actualmente, por lo que no se muestran gráficas del error de posición y del desempeño del controlador, debido a que el sistema es de arquitectura cerrada y no permite visualizar la señal de los sensores. Sin embargo, se aplica un ejemplo de lo que puede hacer el proyecto una

vez concluido, los resultados de la coordinación es solamente visual.

Una vez que se concluye la programación se realizan las pruebas para rectificar que se cumpla el objetivo, destapar una botella y servir en un vaso y entregar a un ser humano; se obtuvieron resultados satisfactorios como se puede observar en las Fig. 6 y 7 las partes del proceso para la coordinación.

Ejecutando cada robot por separado con lo que evitamos ocasionar algún tipo de colisión entre ellos, realizando una validación visual.



Fig. 6 Manipulador sujetando botella



Fig. 7 Robot manipulando el vaso

Una de las parte importantes del proceso de la validación experimental, es ejecutar el programa paso a paso debido a que la velocidad del robot en modo de ejecución T1, T2 y AUTOMATICO son muy distintas es por ello que al momento de trabajarlo en automático se manipula dicha velocidad a una conveniente y adecuada, para no dañar la botella.

VI. Conclusiones

Como se menciona en parte de este artículo, este trabajo es una aplicación futura del proyecto de investigación que la Facultad de Ingeniería Electromecánica actualmente desarrolla, que consiste en coordinar dos o más robots a través de controladores no lineales y la planeación de trayectorias.

Coordinar el trabajo de dos o más manipuladores para realizar una tarea conjunta no es una labor fácil. Este trabajo presenta un resumen de los robots coordinados y una aplicación con robots industriales.

Debido a la gran variedad de aplicaciones de los robots coordinados, se hace cada vez más necesario la elaboración de sistemas de interacción entre robots, esto para desarrollar actividades en forma conjunta, ya que de otra manera sería muy difícil que un sólo robot la realice.

Como se mostró en esta aplicación, la realización de esta tarea para una persona resulta ser muy sencilla, pero sin embargo, se necesitó de una visión e imaginación para poder hacer que este par de robots realicen de manera exitosa la actividad planteada.

Estos sistemas pretenden mejorar la calidad de los servicios y productos que se pueden realizar con ellos, y así contribuir con el desarrollo de nuestro entorno.

VII. Referencias

- [1] **Cole, A.** 1990. Constrained motion of grasped objects by hybrid control. *Proceedings of the 1990 IEEE International Conference on Robotics and Automation* 1954–1960.
- [2] **Liu, G., J. Li y Z. Li.** 2002. Coordinated manipulation of objects by multifingered robotic hand in contact space and active joint space. *IEEE International Conference on Robotics and Automation* 4. 3743–3748.

- [3] **Mason, M. y J. Salisbury.** 1985. *Robot Hands and the Mechanics of Manipulation*. London: The MIT Press.
- [4] **Montana, D.** 1988. The kinematics of contact and grasp. *International Journal of Robotics Research* 7(3). 743–748.
- [5] **Schlegl, T., M. Buss, T. Omata y G. Schmidt.** 2001. Fast dextrous regrasping with optimal contact forces and contact sensor based impedance control. *IEEE Int. Conf. on Robotics and Automation* 103–109.
- [6] **Sciavicco, L. y B. Siciliano.** 2000. *Modeling and Control of Robot Manipulators*. 2th. Edition: McGraw-Hill.

VIII. Autores

Fidel Chávez Montejano, estudiante de la Facultad de Ingeniería Electromecánica de la Universidad de Colima, generación 2007-2012, he recibido cursos de actualización disciplinar tales como manipulación y programación del Robot KUKA KR16, un diplomado en procesos de fabricación y una estancia laboral en la empresa DEVILLE INTERNACIONAL S.A.

Carlos Pizano Martínez, estudiante de la Facultad de Ingeniería Electromecánica de la Universidad de Colima, generación 2007-2012, he recibido cursos de actualización disciplinar tales como manipulación y programación del Robot KUKA KR16.

Dr. Jorge Gudiño Lau, Ingeniero Electricista por la Facultad Ingeniería Mecánica y Eléctrica (FIME), Universidad de Colima (1992). Maestro en Ciencias Computacionales por la FIME, Universidad de Colima (2000), Doctor en Ingeniería por la Facultad de Ingeniería, UNAM en el área de Control con especialidad en Robótica (2005). Actualmente profesor investigador de la Facultad de Ingeniería Electromecánica de la Universidad de Colima. El área de interés es control de robots cooperativos, control de fuerza, diseño de observadores y automatización de procesos industriales.

Simulación de un Sistema Robótico de 3 DOF Orientado a la Rehabilitación Neuromuscular del Brazo.

*Wilebaldo Martínez Velazco, **Felipe de Jesús Trujillo Romero

Universidad Tecnológica de la Mixteca, Laboratorio de Robótica Inteligente. Km 2.5 Carretera a Acatlima C. P. 69000, Huajuapán de León, Oax., Teléfono +52 953 5320214 Ext. 300, *wmartinez@mixteco.utm.mx, **ftrujillo@mixteco.utm.mx

Resumen

En este artículo se presenta la simulación de un sistema robótico de 3 Grados de Libertad (DOF Degree Of Freedom), que ayude a la rehabilitación del miembro superior en personas que sufrieron una lesión neuromuscular en dicha parte del cuerpo. Para que el sistema pueda cumplir con esta tarea, el efector final del robot de 3 DOF guiará el brazo del paciente para que éste realice ejercicios repetitivos. Estos ejercicios están de acuerdo a ciertos patrones de movimientos empleados en terapias de rehabilitación neuromuscular del brazo. El cálculo de la posición del efector final está basado en el análisis de la cinemática directa del sistema propuesto. El diseño del sistema robótico, al estar basado en sólo 3 DOF, pretende la construcción de un dispositivo de bajo costo, tanto de producción como de mantenimiento, que pueda ser adquirido con mayor facilidad por centros de rehabilitación en el país.

Palabras Clave: Cinemática Directa, Entorno de Simulación, Rehabilitación Neuromuscular, Sistema Robótico.

I. Introducción

Según datos de la Organización Mundial de la Salud (OMS) se estima que actualmente existen 650 millones de personas con algún tipo de discapacidad, lo cual representa cerca del 10% de la población Mundial. Según el censo del año 2000, en México el 2.31% de la población sufre alguna discapacidad, de los cuales el 45.3% del total de discapacitados presenta limitaciones relacionadas con sus piernas o brazos [1].

Oaxaca ocupa el 15° lugar entre los estados que presentan mayor población con esta problemática, y de acuerdo con el número total de habitantes considerando que el estado tiene un total de 3, 438,765 habitantes y de ellos el 1.92% son considerados con alguna discapacidad se puede afirmar que 29,710 personas viven con algún tipo de discapacidad en alguno de sus miembros [2].

Esta discapacidad generalmente se da como consecuencia de alguna enfermedad cerebral o

algún traumatismo por accidente. Estos problemas son atendidos por medio de terapias impartidas por neurólogos o licenciados en terapia física u ocupacional. Sin embargo existen herramientas que pueden ayudar a acelerar el tiempo de rehabilitación. Una de estas herramientas son los llamados robots para rehabilitación, tales como *Armin Robot* [3], el proyecto *MIT-Manus* [4] o el exoesqueleto para rehabilitación y entrenamiento del brazo [5], y los sistemas de realidad virtual, los cuales ofrecen la posibilidad de proporcionar una retroalimentación [6].

Debido al alto costo de adquisición, en Oaxaca solamente existe un centro de rehabilitación con un sistema robótico especializado para rehabilitación; este es el Centro de Rehabilitación Infantil Teletón (CRIT), que cuenta con el sistema para rehabilitación de miembro inferior Lokomat [7], el cual ayuda a pacientes impedidos a realizar los movimientos de marcha sobre una cinta rodante y combina una terapia de

locomoción intensiva con herramientas de evaluación.

La propuesta de éste proyecto es simular un sistema robótico que pueda proporcionar los movimientos comunes en rehabilitación neuromuscular y que utilice solamente 3 DOF para reducir el costo de fabricación y mantenimiento del sistema físico. De esta manera se atienden directamente los músculos involucrados con motricidad gruesa del brazo, dejando fuera de la rehabilitación la motricidad fina que se realiza con la muñeca y los dedos.

En la sección II se presenta el algoritmo Denavit-Hartenberg para el cálculo de la posición del efector final respecto del origen del sistema de coordenadas, para un robot de n grados de libertad, basado en su configuración física. En la sección III se detalla el principio de funcionamiento del sistema propuesto, se calcula la matriz de transformación directa del sistema y se da una explicación del programa de simulación utilizado. En la sección IV se muestran los resultados y finalmente en la sección V se hace mención de las conclusiones.

II. Fundamento Teórico

II.1 Algoritmo Denavit-Hartenberg

Denavit-Hartenberg propusieron en 1955 un método matricial que permite establecer de manera sistemática un sistema de coordenadas (S_i) ligado a cada eslabón i de una cadena articulada, pudiéndose determinar a continuación las ecuaciones cinemáticas de la cadena completa [8]. Según la representación D-H, escogiendo adecuadamente los sistemas de coordenadas asociados para cada eslabón, será posible pasar de uno al siguiente mediante 4 transformaciones básicas. Estas transformaciones básicas consisten en una sucesión de rotaciones y traslaciones que permitan relacionar el sistema de referencia del elemento i con el sistema del elemento $i-1$ (ver Fig. 1).

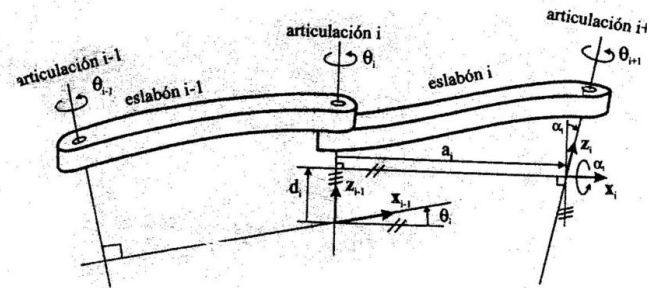


Fig. 1. Parámetros D-H con sus referencias.

Las transformaciones mencionadas en el párrafo anterior utilizan cuatro parámetros que dependen únicamente de las características geométricas de cada eslabón y de las articulaciones que le unen con el anterior y el siguiente. Estos Valores se obtienen mediante las siguientes reglas:

- q_i : Es el ángulo que forman los ejes X_{i-1} y X_i medido en un plano perpendicular al eje Z_{i-1} , utilizando la regla de la mano derecha. Es un parámetro variable en articulaciones giratorias.
- d_i : Es la distancia a lo largo del eje Z_{i-1} desde el origen del sistema de coordenadas $(i-1)$ -ésimo hasta la intersección del eje Z_{i-1} con el eje X_i . Es un parámetro variable en articulaciones prismáticas.
- a_i : Es la distancia a lo largo del eje X_i que va desde la intersección del eje Z_{i-1} con el eje X_i hasta el origen del sistema i -ésimo, en el caso de articulaciones giratorias. En el caso de articulaciones prismáticas, se calcula como la distancia mas corta entre los ejes Z_{i-1} y Z_i .
- α_i : Es el ángulo de separación del eje Z_{i-1} y el eje Z_i , medido en un plano perpendicular al eje X_i , utilizando la regla de la mano derecha.

II.2 Problema Cinemático Directo e Inverso

El problema cinemático directo consiste en determinar cuál es la posición y orientación del extremo final del robot, con respecto a un sistema

de coordenadas que se toma como referencia, conocidos los valores de las articulaciones y los parámetros geométricos de los elementos del robot [9]. En el problema cinemático inverso se calculan los valores de las articulaciones para que el extremo del robot se posicione en determinado punto. La solución de este último problema tiene un costo computacional mayor, debido al mayor número de cálculos que se realizan.

Un robot de n grados de libertad está formado por n eslabones unidos por n articulaciones, de forma que cada par articulación-eslabón constituye un grado de libertad. Cada eslabón tiene asociado un sistema de referencia solidario a él y, utilizando las transformaciones homogéneas, es posible representar las rotaciones y traslaciones relativas entre los distintos eslabones que componen el robot. Teniendo ya los valores de θ_i , d_i , a_i , α_i , que son los denominados parámetros D-H del eslabón i , la matriz de transformación que relaciona los sistemas de referencia $\{S_{i-1}\}$ y $\{S_i\}$ es la siguiente:

$$T = T_{z,\theta} T_{z,d} T_{x,a} T_{x,\alpha} \quad (1)$$

Desarrollando esta expresión en términos de los parámetros D-H, nos queda:

$${}^{i-1}A_i = \begin{pmatrix} \cos \theta_i & -\cos \alpha_i \sin \theta_i & \sin \alpha_i \sin \theta_i & a_i \cos \theta_i \\ \sin \theta_i & \cos \alpha_i \cos \theta_i & -\sin \alpha_i \cos \theta_i & a_i \sin \theta_i \\ 0 & \sin \alpha_i & \cos \alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2)$$

Del mismo modo, la matriz 0A_n , resultante del producto de las matrices ${}^{i-1}A_i$ con i desde 1 hasta n , es la que representa de forma total la cadena cinemática que forma el robot con respecto al sistema de referencia inercial asociado a la base. A la matriz 0A_n se le denomina T , matriz de transformación que relaciona la posición y orientación del extremo final del robot respecto del sistema fijo situado en la base del mismo. Así, dado un robot de 3DOF, se tiene que la posición y orientación del eslabón final vendrá dado por la ecuación (3).

$$T = {}^0A_3 = {}^0A_1 \cdot {}^1A_2 \cdot {}^2A_3 \quad (3)$$

Con los parámetros definidos en esta sección se construyeron los sistemas de referencia del robot de 3 DOF.

III. Desarrollo

III.1 Descripción del proyecto.

El sistema simulado consiste de un brazo robótico de 3 DOF que en el efector final sostiene un lápiz, el cual a su vez guiara la mano del paciente para realizar desplazamientos sobre un plano, simulando el dibujo de figuras sobre un papel (ver Fig. 2).

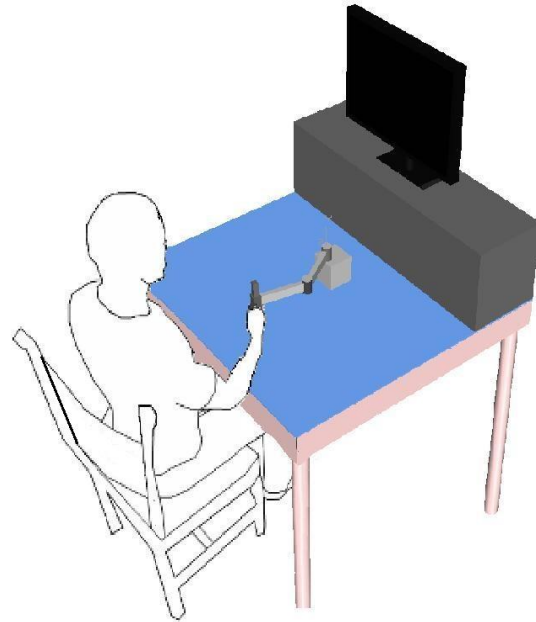


Fig. 2. Sistema de rehabilitación propuesto.

Consiste de 3 elementos con dos articulaciones Giratorias y una Prismática. El primer elemento está unido a la base a través de una articulación giratoria, de igual manera que el segundo elemento está unido al primer elemento. El tercer elemento se une al segundo por medio de una articulación prismática (ver Fig. 3).

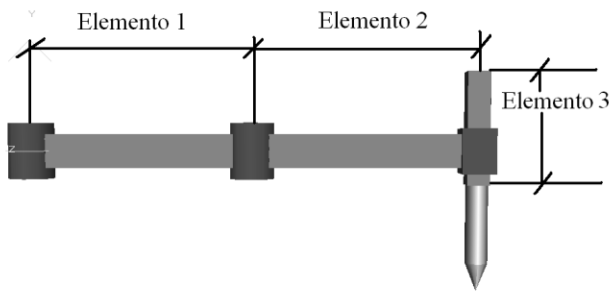


Fig. 3. Elementos del sistema robótico propuesto.

Los dos primeros grados de libertad del sistema robótico propuesto, que corresponden con las articulaciones giratorias, permiten posicionar el lápiz en cualquier punto dentro de un radio de l_1+l_2 desde la base del brazo, donde l_1 y l_2 son la longitud de los elementos 1 y 2 respectivamente (ver Fig. 4). El tercer grado de libertad sube el lápiz para separar su punta de la superficie.

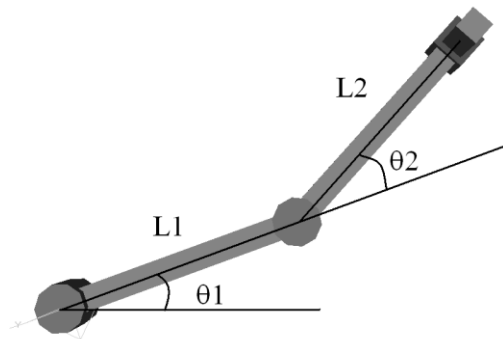


Fig. 4. Articulaciones giratorias del sistema.

III.2 Definición de los parámetros D-H y las matrices de Transformación.

A continuación se definen cada uno de los 3 sistemas de referencia, uno por cada grado de libertad. Con el Método descrito en el apartado II.1 obtenemos los parámetros Denavit-Hartenberg del Robot los cuales quedan como en la tabla (1).

i	α_{i-1}	a_{i-1}	d_i	θ_i
1	0	0	0	θ_1
2	0	l_1	0	θ_2
3	0	l_2	d_3	0

Tabla 1. Parámetros D-H del sistema.

Con estos parámetros, se encuentra la matriz de transformación para cada articulación, sustituyendo los valores en (2).

De esta manera las matrices de transformación quedan como:

$${}^0A_1 = \begin{pmatrix} \cos\theta_1 & -\text{sen}\theta_1 & 0 & 0 \\ \text{sen}\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4)$$

$${}^1A_2 = \begin{pmatrix} \cos\theta_2 & -\text{sen}\theta_2 & 0 & l_1\cos\theta_2 \\ \text{sen}\theta_2 & \cos\theta_2 & 0 & l_1\text{sen}\theta_2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5)$$

$${}^2A_3 = \begin{pmatrix} 1 & 0 & 0 & l_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6)$$

Finalmente en (7) se obtiene la matriz de transformación directa del sistema sustituyendo (4), (5) y (6) en (3):

$${}^0T_3 = \begin{pmatrix} r_{11} & r_{12} & 0 & x_1 \\ r_{21} & r_{22} & 0 & x_2 \\ 0 & 0 & 1 & d_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (7)$$

Donde:

$$r_{11} = \cos\theta_1 \cos\theta_2 - \text{sen}\theta_1 \text{sen}\theta_2$$

$$r_{21} = \text{sen}\theta_1 \cos\theta_2 + \cos\theta_1 \text{sen}\theta_2$$

$$r_{12} = -\cos\theta_1 \text{sen}\theta_2 - \text{sen}\theta_1 \cos\theta_2$$

$$r_{22} = -\text{sen}\theta_1 \text{sen}\theta_2 + \cos\theta_1 \cos\theta_2$$

$$x_1 = \cos\theta_1(l_2\cos\theta_2 + l_1\cos\theta_2) - \text{sen}\theta_1(l_2\text{sen}\theta_2 + l_1\text{sen}\theta_2)$$

$$x_2 = \text{sen}\theta_1(l_2\cos\theta_2 + l_1\cos\theta_2) + \cos\theta_1(l_2\text{sen}\theta_2 + l_1\text{sen}\theta_2)$$

Con estas matrices de transformación podemos obtener la posición y orientación del efector final, conociendo los valores de los ángulos. El

algoritmo para realizar un trazo es el siguiente:

1. Se define el trazo a realizar.
2. Se propone una configuración adecuada para alcanzar el punto inicial del trazo.
3. Por Cinemática Directa se calcula la matriz de transformación del sistema.
4. Se varían los valores de x_1 y x_2 que corresponde con los valores de las coordenadas x y del plano, según sea la dirección del movimiento.
5. Por cinemática inversa se calcula la nueva configuración de las articulaciones y se mandan a visualizar.
6. Regresamos al paso 3 para calcular nuevamente la matriz de transformación del sistema. Esto hasta alcanzar el punto final del trazo.

Estos resultados fueron probados mediante la simulación mostrada en la siguiente sección.

III.3 Entorno de Simulación

La visualización de las trayectorias propuestas para el efector final del sistema, requiere de un entorno de simulación en el que puedan ser observados los movimientos de cada uno de los elementos del sistema, así como de la trayectoria calculada para la punta del lápiz en el efector final.

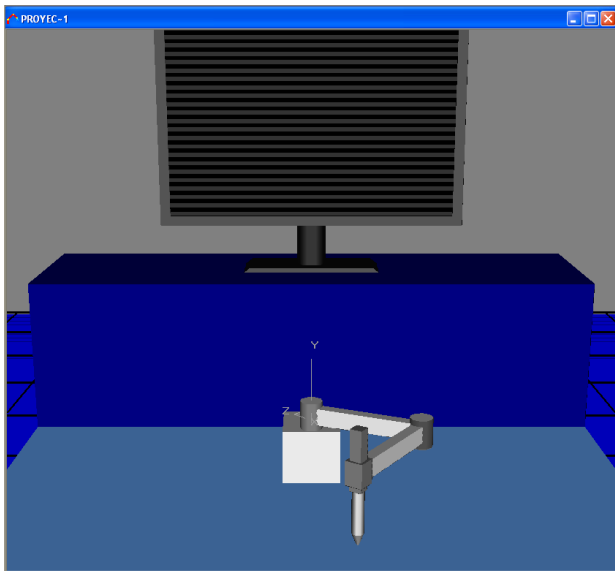


Fig. 5. Simulación del sistema propuesto en Roboworks.

El cálculo de las trayectorias deseadas se realizó en lenguaje C++ utilizando la plataforma Visual Studio 2008 y para su visualización se utilizó el programa Roboworks [10] desarrollado por Newtonium, el cual es un modelador tridimensional para la simulación de elementos mecánicos (ver Fig. 5). Este programa permite al usuario realizar un modelo en 3D y animarlo desde un archivo ejecutable creado en C++, utilizando librerías de enlace las cuales pueden ser descargadas desde la página del fabricante al igual que el programa en su versión de evaluación.

IV. Resultado de Simulación.

Se propusieron los dos primeros elementos del sistema con un largo de 20 cm. cada uno por lo que el radio máximo de posición del efector final es de 40 cm., y alcanza un espacio de trabajo de 0.2513m^2 (ver Fig. 6).

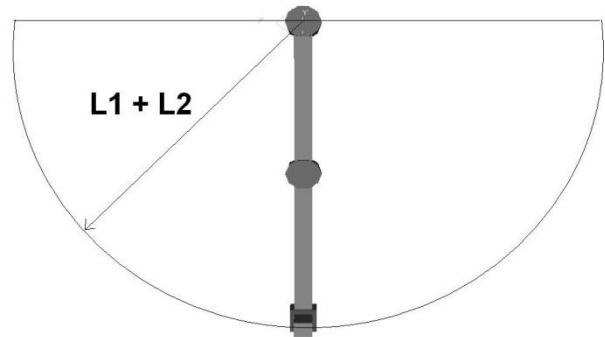


Fig. 6. Espacio de trabajo del sistema propuesto.

De esta manera se comprobaron las posiciones calculadas para el sistema propuesto visualizándolo con distintos patrones de trayectorias para el efector final. Para esto se desarrolló una interfaz gráfica donde se pueden seleccionar entre nueve patrones de movimiento y tres figuras más elaboradas que son construidas a través del trazo de patrones básicos. En el caso de los patrones de movimiento se pueden elegir el tamaño y la velocidad de los trazos entre cinco posibles, además del número de repeticiones de los movimientos (ver Fig. 7).

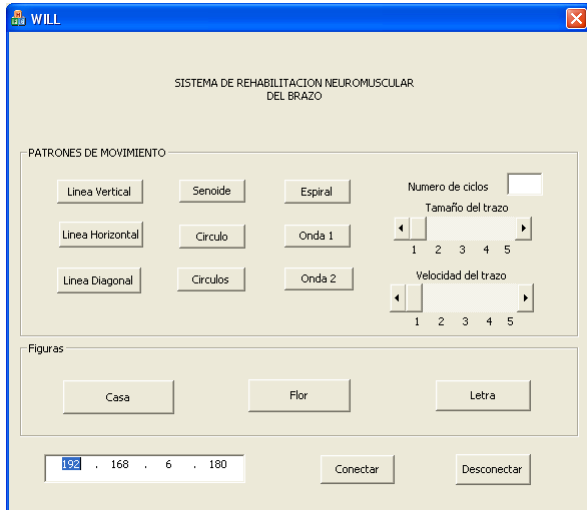


Fig. 7. Interfaz Grafica.

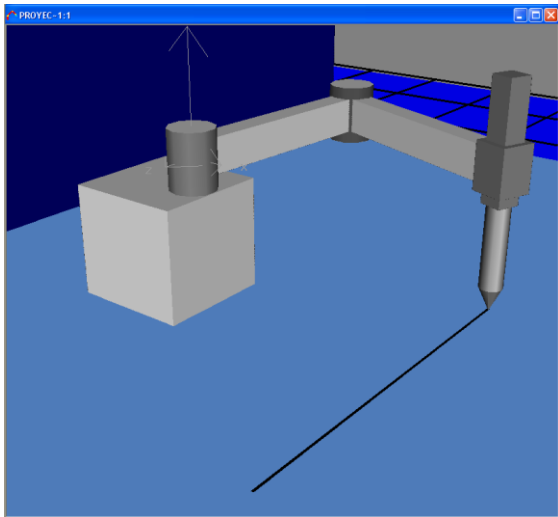


Fig. 8. En el punto de inicio de la trayectoria lineal.

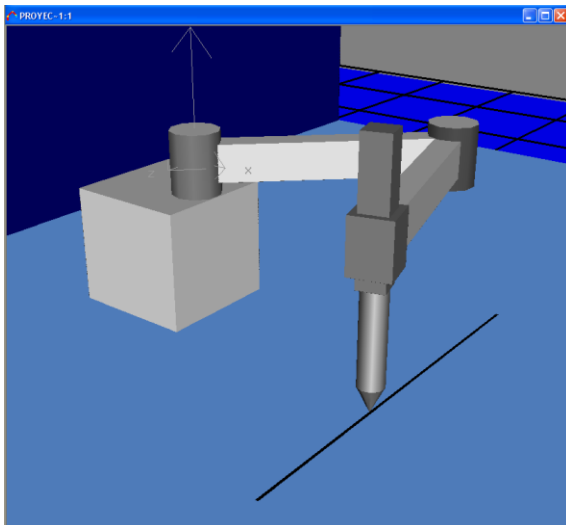


Fig. 9. A la mitad de la trayectoria lineal.

Los patrones de movimiento propuestos, son movimientos utilizados en terapias de rehabilitación como círculos, líneas y la combinación de estos (ver Fig. 8, 9, 10 y 11).

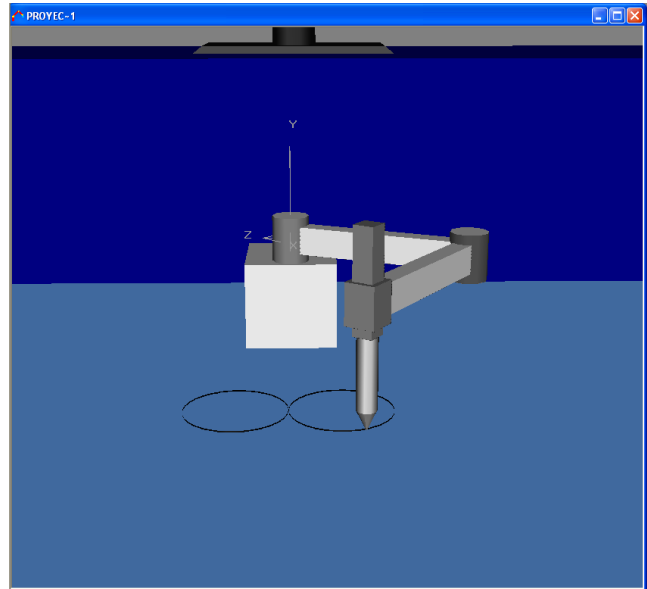


Fig. 10. Trazo del patrón de círculos.

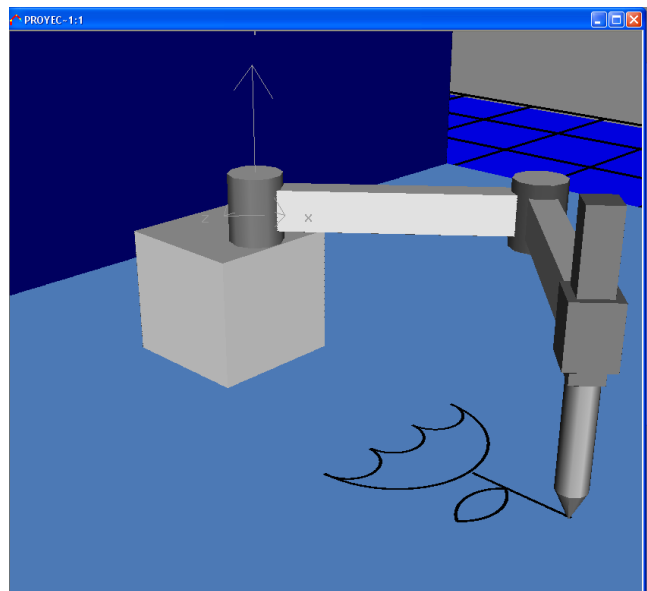


Fig. 11. Trazo de figuras con patrones combinados.

V. Conclusiones

El desarrollo de este simulador nos permite evaluar los tipos de movimientos que se pueden generar con un Brazo robótico de 3 DOF y contrastarlos con los movimientos que los

médicos aplican en las terapias, de esta manera puede existir una retroalimentación preliminar antes de construir un sistema físico.

El visualizar las trayectorias nos proporciona información acerca de los tipos de motores, material de los elementos y el sistema de control necesario para generarlas.

El sistema de rehabilitación simulado en este artículo puede combinarse con otras técnicas de rehabilitación como la visualización en pantalla de las trayectorias generadas, creando de esta manera una retroalimentación.

Cualquier ayuda robótica que implique la interacción con el ser humano, necesita ser evaluada y aprobada por los médicos especialistas en rehabilitación.

VI. Referencias

- [1] Programa de Acción Específico para la Atención Integral a la Salud de las Personas con Discapacidad, 2007-2012. http://www.conadis.salud.gob.mx/descargas/pdf/atn_intgrl_salud_pcd.pdf. Última fecha de consulta 8 de Junio 2011.
- [2] Diagnostico de la Discriminación en el Estado de Oaxaca. http://cedoc.inmujeres.gob.mx/lgamv/v/CO_NAPRED/conapred04.pdf. Última fecha de consulta 8 de Junio 2011.
- [3] T. Nef, R. Riener, "Armin, Design of a Novel Arm Rehabilitation Robot". 9th International Conference on Rehabilitation Robotics. 0-7803-9003-2/05. Junio 2005.
- [4] H.I. Krebs, B.T. Volpe, M.L. Aisen, N. Hogan, "Increasing productivity and quality of care: Robot-aided neuro-rehabilitation". Journal of Rehabilitation Research and development. Volumen 37. Numero 6. Noviembre/Diciembre 2000. Páginas 639-652.
- [5] A. Gupta, M.K. O'Malley, "Design of a Haptic Arm Exoskeleton for Training and Rehabilitation". IEEE/Asme Transactions on Mechatronics. Volumen 11. Número 3. Junio 2006.
- [6] M.S. Cameirao et al, "Neurorehabilitation using the virtual reality based Rehabilitation Gaming System: Methodology, desing, psychometrics, usability and validation". Journal of NeuroEngineering and Rehabilitation. Volumen 7. Numero 48. 2010.
- [7] Robot para rehabilitación de los miembros inferiores Hocoma Lokomat. <http://www.hocoma.com/en/products/lokomat>. Última fecha de consulta 8 de Junio de 2011.
- [8] J.Denavit, R.S. Hartenberg, "A kinematic Notation for lower-pair Mechanisms based on Matrices". Journal of Applied Mechanics. Volumen 77. 1955. Páginas 215-221.
- [9] K.S.Fu et al. Robótica. Control, detección, visión e inteligencia. McGraw-Hill, Madrid. 1998.
- [10] Simulador tridimensional Roboworks, <http://www.newtonium.com>. Última fecha de consulta 8 de Junio 2011.

VII. Autores

Ing. Wilebaldo Martínez Velazco, se graduó en la Universidad Tecnológica de la Mixteca, de la Ingeniería en Electrónica en el año 2004. Actualmente es estudiante de la Maestría en Robótica en la Universidad Tecnológica de la Mixteca y se desempeña como técnico de tiempo completo en el laboratorio de Robótica Inteligente en la misma Universidad.

Felipe de Jesús Trujillo Romero, Dr. En Sistemas Informáticos por el Instituto Nacional Politécnico de Toulouse, Francia, actualmente profesor investigador de la División de Estudios de Posgrado de la Universidad Tecnológica de la Mixteca.

Sistema de Ojo Robótico y Aplicación al Seguimiento de Objetos

***José Arce y de la Borbolla, *Juan Carlos Ruiz Cruz, *Mario Acevedo Alvarado, **Ramiro Velázquez Guerrero**

* Universidad Panamericana campus México. Augusto Rodín 498, Col. Insurgentes Mixcoac, CP 03920, México DF, México. Email: macevedo@up.edu.mx

**Universidad Panamericana campus Bonaterra. Josemaría Escrivá de Balaguer 101, Fracc. Rústicos Calpulli, CP 20290, Aguascalientes, Ags., México. Email: rvelazquez@ up.edu.mx

Resumen

En este artículo se presenta el diseño e implementación de un ojo robótico capaz de seguir la trayectoria de un objeto particular mediante técnicas de procesamiento de imágenes y visión por computadora. El artículo se subdivide en las tres grandes áreas que abarca el proyecto: mecánica, visión y orientación automática. En la sección mecánica se discuten las principales razones por las que se llegó al diseño final del ojo robot, se muestran tablas y figuras de la geometría y ensamblado de los componentes y finalmente se analiza la cinemática del eslabonamiento que conforma al prototipo. En la parte de visión se justifican los filtros y transformaciones que se aplican a las imágenes con el fin de localizar rápida y eficientemente un objeto deseado. Por último, en la sección de orientación automática se describe cómo interactúan mecánica y visión para lograr el funcionamiento del ojo robot.

Palabras Clave: análisis cinemático, ojo robótico, robot esférico, seguimiento de objetos.

I. Introducción

El movimiento ocular tiene por objetivo optimizar la percepción visual de nuestro entorno [1]. Por ello, no es de extrañar que los sistemas de visión busquen imitar el movimiento del ojo para la orientación de las cámaras de video.

Fisiológicamente, el movimiento del ojo se debe a seis músculos que proporcionan al ojo tres grados de libertad (GDL). Éstos son el oblicuo superior e inferior, el recto superior, lateral, medial e inferior (Fig. 1(a)). Estos seis músculos trabajan en grupo para mover los ojos arriba/abajo, en sentido lateral y en rotación (Fig. 1(b)).

Un sinnúmero de robots y dispositivos mecatrónicos utilizan hoy en día sofisticados sistemas de visión para tareas de navegación autónoma, posicionamiento, localización, control, seguimiento e identificación, entre otras. Independientemente del tipo y la eficiencia del algoritmo computacional utilizado, muchas veces

resulta conveniente orientar la cámara para optimizar el rango visual en la aplicación.

El sistema de orientación de cámaras más utilizado es sin duda el PT (pan & tilt) que permite dos GDL (Fig. 2). En el pasado, hemos explorado sistemas de orientación de cámaras de video con mecanismos PT [2]. En este artículo, se propone el uso de mecanismos esféricos para el posicionamiento rápido y preciso de una cámara de video. Inspirados en la fisiología del ojo humano, hemos desarrollado un ojo robótico que imita el movimiento ocular humano. El resto del artículo está organizado de la siguiente manera: la Sección 2 presenta el diseño, la implementación y el análisis cinemático del prototipo. La Sección 3 describe los principios de visión computacional para una aplicación particular: seguimiento de objetos, mientras que la Sección 4 muestra como la estructura mecánica

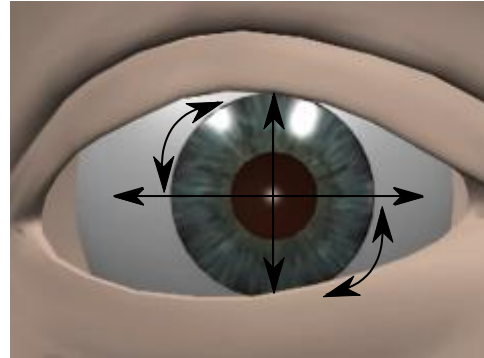
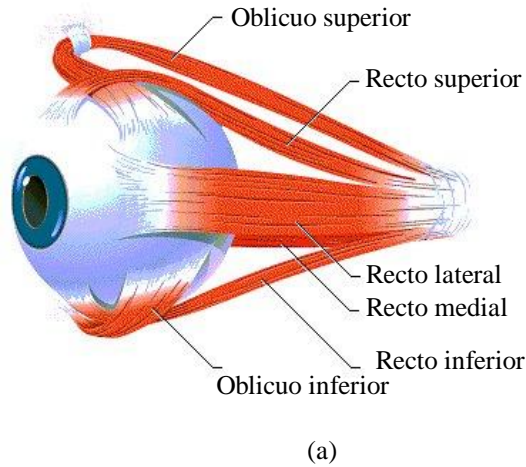


Fig. 1. (a) Los seis músculos que permiten el movimiento ocular y (b) los tres GDL resultantes del ojo.

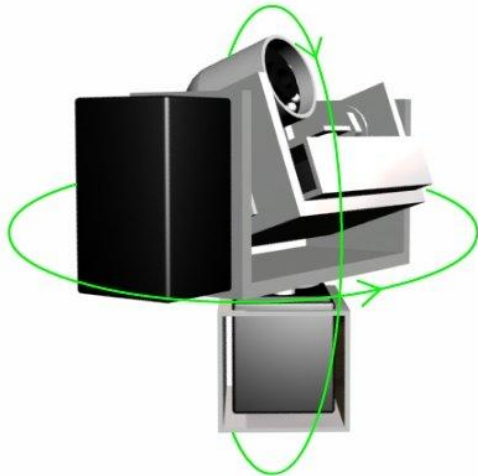


Fig. 2. Un mecanismo PT típico para orientación de cámaras de video.

se orienta automáticamente a partir de una imagen. Finalmente, la Sección 5 resume los conceptos principales y perspectivas de trabajo futuro.

II. Mecánica

A) Diseño

Al igual que el ojo humano, el diseño mecánico del ojo robótico debe permitir a la cámara modificar su orientación dentro de un sistema coordinado esférico. De esta manera, la cámara puede ser reorientada sin tener traslación relativa a un punto fijo en el espacio, más sí contar

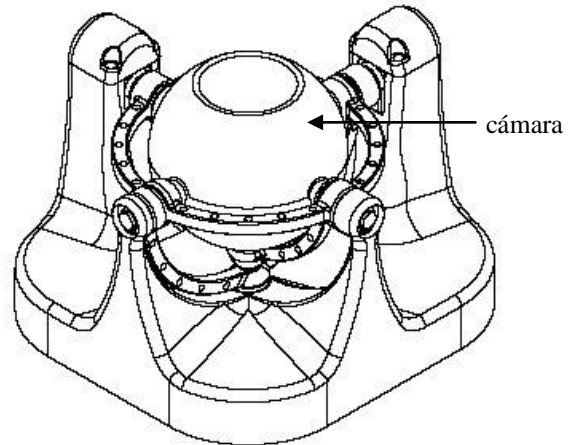


Fig. 3. Diseño mecánico del ojo robótico.

con al menos dos GDL rotacionales alrededor de dicho punto.

Con este objetivo, se decidió utilizar una arquitectura de robot paralelo, la cual permite generar movimientos del actuador final con mayor precisión, velocidad y aceleración que una arquitectura en lazo abierto.

La estructura principal del eslabonamiento de nuestro ojo robótico se basa en el 2 DOF-Agile Eye de la Universidad Laval en Quebec, Canadá [3]. El mecanismo en cuestión consta de un eslabonamiento de cinco barras dispuestas en una configuración de suspensión de Cardán que sostiene una cámara en su centro (Fig. 3).

A pesar de existir otras configuraciones para un orientador esférico (por ejemplo [4], [5]) se optó por la configuración de la fig. 3 por su simplicidad en cuanto a menor número de componentes y solución cinemática.

La fig. 4 muestra en detalle la estructura del ojo robótico mientras la tabla I enlista sus componentes.

El diseño y dimensionamiento de los componentes dependió de las dimensiones de los motores y de la cámara, siendo éstos los únicos elementos adquiridos en el comercio.

Dado que el sensor y el sistema óptico de la mayoría de las cámaras se encuentran descentrados con respecto al circuito impreso de la misma, las dimensiones del eslabón central deben ajustarse para que tanto el lente y el sensor se mantengan dentro de un eje diametral de la esfera y no excéntricos con respecto a éste.

Alrededor del eslabón esférico central se diseñaron los anillos de la suspensión de Cardán, procurando que tuvieran pares de revolución diametralmente opuestos (como se aprecia en los elementos A, B y C de la fig. 4) para así tener un mecanismo más balanceado, con apoyo a ambos lados y por consiguiente menores esfuerzos de flexión y menor carga para los motores.

B) Manufactura

Todos los componentes fabricados, a excepción de los ejes, fueron diseñados para ser elaborados en plástico de ingeniería, por lo que se utilizan espesores de pared delgados (entre 1.5 y 3 mm) y se agregaron soportes estructurales como costillas para rigidizar las piezas finales. Además, éstas cuentan con todas las características de sujeción necesarias para ser unidas entre sí mediante tornillos auto-roscantes.

Estas características hacen posible la fabricación de las piezas mediante procesos de inyección de plástico. En particular, el prototipo fue realizado mediante un proceso de impresión 3D en ABS con material de soporte.

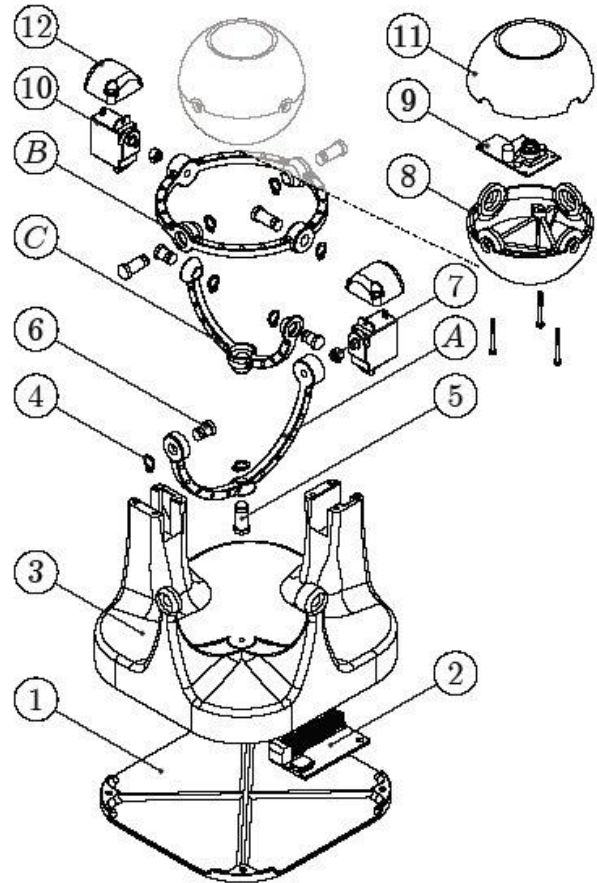


Fig. 4. Diseño mecánico del ojo robótico (vista explosiva de componentes).

No.	Descripción	Cantidad
A	Anillo de cardán externo	1
B	Anillo de cardán medio	1
C	Anillo de cardán interno	1
1	Cubierta inferior	1
2	Pololu 16-Servo USB Controller	1
3	Base y soportes	1
4	Circlip DIN 471 - Ø6 × 0.7mm	7
5	Eje Ø6 × 14mm	4
6	Eje Ø6 × 8mm	3
7	Cople	2
8	Cubierta inferior del ojo	1
9	Vimicro USB 2.0 PC Camera	1
10	Servomotor Vigor VS-5	2
11	Cubierta superior del ojo	1
12	Cubierta del motor	2

Tabla I. Componentes de la estructura del ojo robótico.

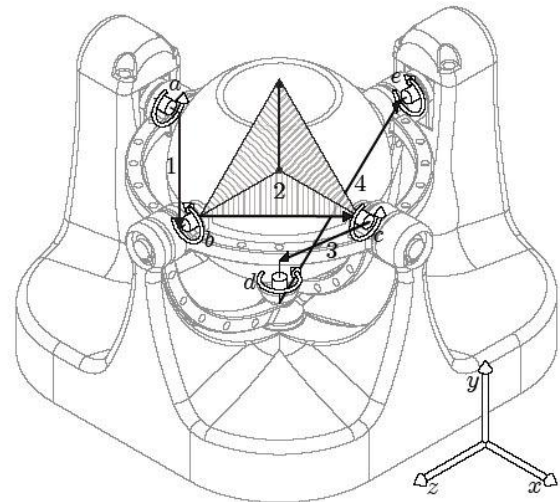
Por otra parte, los ejes se fabricaron de acero rolado en frío y maquinados hasta su geometría y dimensiones mediante un proceso de torneado CNC que permite obtener una gran precisión y repetitividad.

Los barrenos que conforman los pares de revolución fueron ajustados manualmente para obtener un buen soporte de los ejes e impedir así rotaciones no axiales y desplazamientos (juego) entre los componentes. El diseño de las piezas pretende que su ensamble y desarmado sean relativamente sencillos y con acceso a cualquier elemento de sujeción con facilidad. Para ello, en el diseño se contempla que debe realizarse primero el ensamblado de los elementos que conforman el mecanismo (incluyendo los motores) para después insertar dicho subensamblaje de forma vertical en la base. Subensamblaje y base se sujetan entre sí mediante las cubiertas de los motores y los ejes restantes.

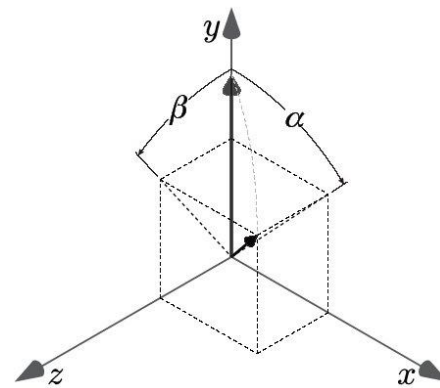
C) Cinemática

La fig. 5(a) muestra la cadena cinemática que conforma el mecanismo del ojo robótico. Los vectores con líneas gruesas representan los eslabones que conforman la suspensión de Cardán, los cilindros con flechas representan los pares de revolución entre los componentes y el eslabón central se representa por las tres caras sombreadas, donde el punto negro indica el centro y punto de rotación de la esfera y el vector vertical apunta en la dirección de la cámara.

En lo subsecuente, la notación será de la siguiente forma: los eslabones se denominarán por el número correspondiente mostrado en la fig. 5(a), siendo el eslabón tierra el elemento 0; los pares cinemáticos se denotarán mediante las letras correspondientes a la figura; y por último los ángulos formados entre los elementos en cada par cinemático se denotarán como θ_i donde i representa la literal de la junta y se medirá en sentido positivo desde el elemento con número menor hasta el de mayor número. Las variables de control serán los ángulos aplicados a los elementos 1 y 4, respectivamente: θ_a y θ_e .



(a)



(b)

Fig. 5. (a) Cadena cinemática para el ojo robótico y (b) definición de los ángulos α y β .

Por otra parte, las variables de interés son los ángulos que forman las proyecciones del vector vertical de la cámara con los planos xy y zy con respecto al eje y del sistema de referencia global. En adelante se denominarán dichos ángulos como α y β , los cuales se ilustran en la fig. 5(b).

De la fig. 5(b) se puede escribir:

$$\alpha = \tan^{-1} \left(\frac{z}{y} \right) \quad (1)$$

$$\beta = \tan^{-1} \left(\frac{-x}{y} \right) \quad (2)$$

Los ángulos α y β se relacionan con las coordenadas esféricas de elevación y acimut

(considerando un radio unitario) a través de las ecuaciones (3) y (4):

$$\theta = \cos^{-1} \left(-\frac{1}{\sqrt{\tan^2(\alpha) + \sec^2(\beta)}} \right) \quad (3)$$

$$\varphi = \tan^{-1}(\tan(\alpha) \cot(\beta)) \quad (4)$$

donde θ es el ángulo de inclinación o elevación medido desde el plano \mathbf{xz} y φ es el ángulo acimutal medido a partir del eje \mathbf{x} .

La solución del problema de posiciones no se obtuvo dado que, para esta arquitectura en particular, se puede considerar que los valores de salida (α , β) son iguales a los de entrada (θ_a , θ_e). Lo anterior acarrea consigo un error, mayor cuando α y β están más alejados de 0, pero el algoritmo de control corregirá dicho error en cada iteración.

Por otro lado, dado que para resolver la posición del robot es necesario resolver un sistema de ecuaciones no lineales, y que ello requiere de un alto tiempo de procesamiento, se optó por no incluir la solución cinemática en la programación en pro de un menor tiempo de procesamiento y por consiguiente tener una mayor frecuencia de muestreo del sistema.

III. Visión

El procesamiento de imágenes obtenidas por la cámara del ojo robótico se lleva a cabo en una computadora externa mediante el programa *Matlab*. Una aplicación demostrativa para este prototipo es la detección y seguimientos de objetos. En esta sección se detallan los principios de detección y seguimiento para una pelota.

A) Filtros y transformaciones

El primer paso a realizar durante el procesamiento de las imágenes adquiridas por la cámara es el de filtrar la imagen en cuanto al color que se desea buscar. Por ejemplo, consideremos una pelota roja (Fig. 6(a)). Dado que el objeto posee una gran proporción de rojo en comparación con los demás componentes de color, es necesario localizar en la imagen

únicamente los objetos de color semejante. Para ello, no basta sólo con tomar los valores altos de la capa roja, siendo que muchísimos colores pueden tener un valor elevado de rojo mas no ser en sí rojos sino morados, anaranjados o incluso blancos.

Por lo anterior, y conociendo que la proporción aproximada de colores que conforman el promedio del objeto es de 2:1:1 (R:G:B), se propone la transformación de la ecuación (5) a la imagen:

$$I = 2R - G - B \quad (5)$$

donde I es la imagen resultante y R , G y B son las capas roja, verde y azul de la imagen original, respectivamente.

Con la ecuación (5) se obtendrán resultados fuera del rango permitido por una imagen, por lo que dichos valores deben cerrarse a los límites permitidos que admite una imagen (0-255 para el caso de una imagen de 8 bits). El resultado será entonces una “radiografía” del color rojo (Fig. 6(b)), en particular el color propio del objeto, mostrando más claros los puntos que se asemejen más al color del objeto y más oscuros los pixeles que se asemejen poco.

Una vez obtenida la imagen de los tonos del color del objeto se le aplica la operación morfológica opening [6] utilizando un elemento estructural circular. Con ello se eliminan o atenúan los pixeles libres que no pertenecen a una geometría circular (Fig. 6(c)).

Finalmente se binariza la imagen utilizando un umbral, dado que la representación del objeto contrasta altamente con respecto al fondo.

Prácticamente cualquier valor umbral es capaz de mostrar únicamente el objeto deseado (Fig. 6(d)).

B) Seguimiento de objeto

Para el seguimiento del objeto deseado existen diversos métodos que pueden utilizarse. Entre ellos destacan la transformada de Hough circular y la segmentación-etiquetación de regiones de la imagen.

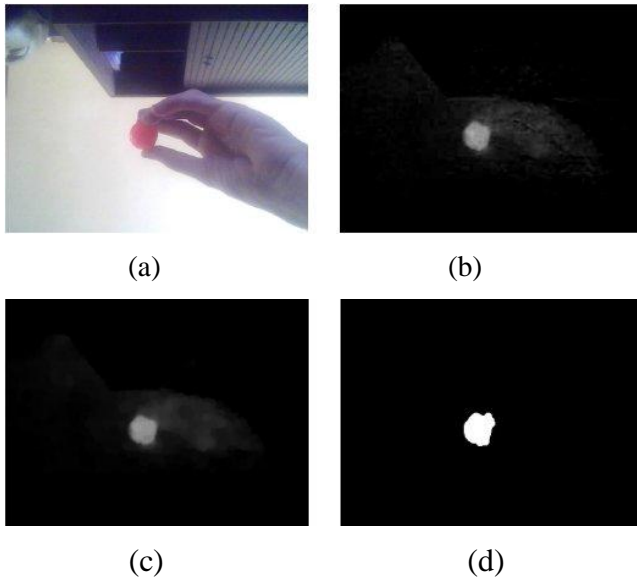


Fig. 6. Etapas de procesamiento de una imagen:
(a) imagen original, (b) capa roja en la imagen,
(c) imagen después de operación morfológica,
(d) imagen binaria.

Dado que la transformada de Hough circular requiere, en su algoritmo más simple, de barrer toda la imagen en busca de posibles círculos y que a su vez debe barrerse de nuevo una matriz del tamaño de la imagen pero con una profundidad igual al rango de radios que se desea localizar con tal de encontrar el mayor número de votos a favor de un cierto círculo [7], se descartó por completo dado que dicha operación tardaba mucho tiempo en su ejecución: por ejemplo, buscando círculos de radio menor o igual a 150 píxeles, el programa debía realizar más de 19 millones de iteraciones. Incluso algoritmos altamente optimizados tomaban demasiado tiempo, memoria o procesamiento.

En cuanto al método de segmentación, etiquetación y cálculo de las propiedades de las áreas resultantes, se optó también por no usarse. Aún cuando dicho método requiere de un tiempo mucho menor al de la transformada de Hough, igual consume valioso tiempo de procesamiento. Además ambos métodos traen consigo otro problema: una vez localizado el círculo, es necesario realizar un proceso de toma de decisión para determinar cuál es el objeto deseado.

Aunque esto pueda realizarse mediante el uso de funciones de probabilidad, se prefirió utilizar un método más simple: la imagen obtenida por las transformaciones descritas en la sub-sección anterior es de por sí ya bastante precisa, y diversas pruebas experimentales demostraron que ni siquiera otras tonalidades de rojo la afectan considerablemente. Por ello, se optó por mejor localizar los límites horizontales y verticales del área blanca de la imagen final. Así es posible determinar la posición de las esquinas del rectángulo que encerrará por completo al objeto deseado.

El código siguiente muestra como es la implementación de este concepto en Matlab:

```
[row, col] = find(I);
xmax = max(col);
xmin = min(col);
ymax = max(row);
ymin = min(row);
```

Encuentra las filas y columnas donde se encuentran los píxeles blancos en la imagen.

Obtiene las coordenadas máximas y mínimas del objeto.

El radio y las coordenadas del centro del círculo se estiman mediante las ecuaciones (6), (7) y (8):

$$r = \frac{y_{max} - y_{min} + x_{max} - x_{min}}{4} \quad (6)$$

$$c_x = \frac{x_{max} + x_{min}}{2} \quad (7)$$

$$c_y = \frac{y_{max} + y_{min}}{2} \quad (8)$$

La fig. 7 muestra una imagen con detección y seguimiento de la pelota roja. Este método es el más simple y rápido de los tres, y su error depende totalmente de la forma de sujetar el objeto (si se obstruye con la mano la mayor parte de su geometría evidentemente no funcionará).

IV. Orientación automática de la cámara

Una vez detectado el objeto deseado surge el problema de relacionar el movimiento de la estructura del ojo robótico con la imagen de la cámara. La fig. 8 ilustra este concepto.

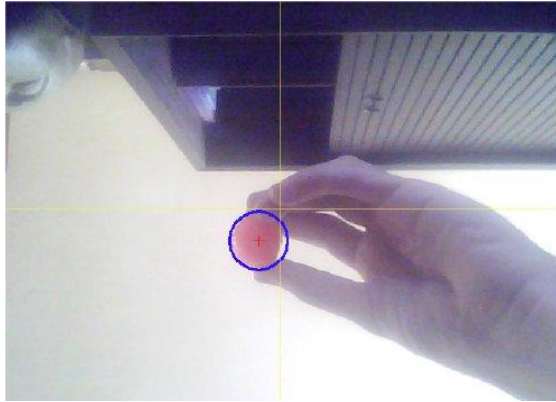


Fig. 7. Seguimiento de objeto.

El reto consiste en trasladar las coordenadas c_x y c_y obtenidas por el algoritmo de visión desde el sistema coordenado xy de la imagen al sistema $x'y'$ de la estructura mecánica que sujeta la cámara.

Las traslaciones se hacen añadiendo a los valores calculados de c_x y c_y un offset igual a la mitad del largo y alto del sensor respectivamente. De igual modo, como se aprecia en la fig. 8, el valor de c_y debe reflejarse sobre el eje x invirtiendo su signo. Por último, el valor de c_x también debe reflejarse sobre el eje y y dado que la imagen que produce la cámara se encuentra espejeada en esa dirección. De esta forma y considerando una cámara de 642 x 482 pixeles, los valores de c_x y c_y referenciados al sistema coordenado $x'y'$ serán de la forma:

$$c'_x = s_p(321 - c_x) \quad (9)$$

$$c'_y = s_p(241 - c_y) \quad (10)$$

donde cx' y cy' son las coordenadas del centro referidas al sistema $x'y'$, c_x y c_y son las coordenadas del centro del objeto referidas al sistema coordenado xy de la imagen en Matlab, y s_p es el tamaño del pixel (en nuestro caso de 0.005mm).

Para estimar el ángulo que debe rotar el ojo en cada eje, considere las figs. 9(a) y (b). En la fig. 9(a) se aprecia que el ángulo que forma la proyección del centro del objeto con respecto al eje de la cámara es distinto al ángulo que debe girar el ojo para centrar dicho objeto en la

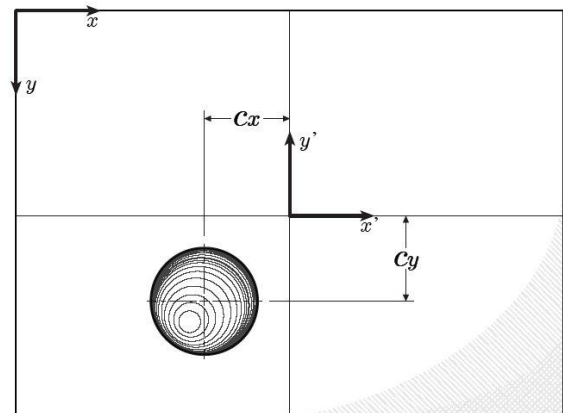
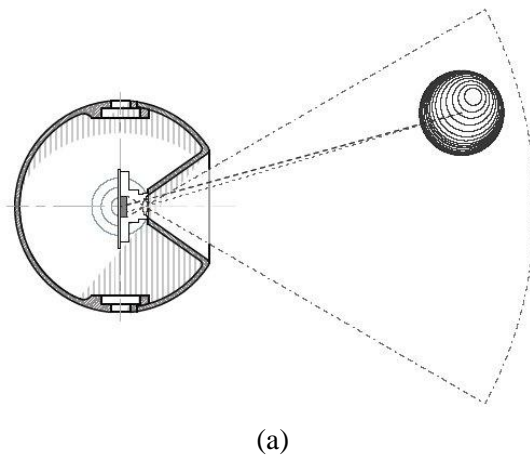
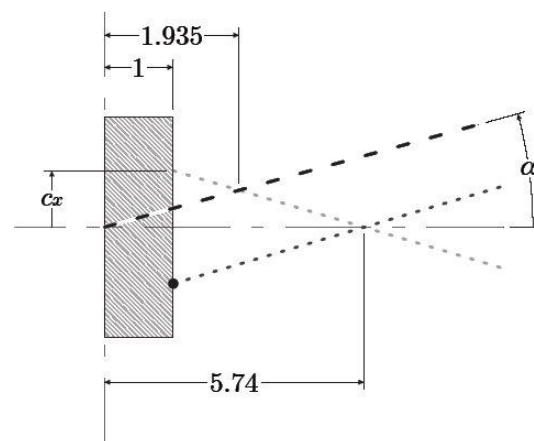


Fig. 8. Representación del objeto en la imagen y su relación con la estructura del ojo robótico.



(a)



(b)

Fig. 9. Estimación del ángulo α .

imagen. Lo anterior debido a que existe una diferencia entre la posición del foco óptico de la cámara con respecto al centro de la esfera.

Por eso, para obtener el ángulo α que debe girar el ojo, hay que considerar la fig. 9(b), donde el recuadro sombreado representa el sensor CMOS de la cámara, la línea rayada más oscura representa la distancia angular que debe rotar el ojo para centrar el objeto deseado y las líneas punteadas representan la proyección del centro del objeto en el sensor, siendo la más oscura la proyección real, la más clara la proyección reflejada y el punto donde se cruzan representa el foco.

La cota c_x es el valor de la posición del centro del objeto con respecto al centro del sensor. Basados en la fig. 9(b), el valor del ángulo α se puede estimar mediante triángulos semejantes como lo expresa la ecuación (11):

$$\alpha = \tan^{-1} \left(\frac{5.74 - 1.935}{(5.74)(1.935)} c_x' \right) \quad (11)$$

De forma análoga es posible obtener β a partir de c_y' .

Finalmente, dado que los ángulos se obtienen a partir de un sistema coordenado dependiente de la posición de la cámara, los valores globales se obtienen sumando los ángulos obtenidos en la iteración presente más los ángulos de la iteración anterior. Esto es válido si y sólo si se conoce la posición inicial del robot, por lo que en nuestro caso es necesario inicializar el ojo robótico con $\alpha = \beta = 0$.

V. Conclusiones

Este artículo introdujo un mecanismo esférico con una cámara de video montada en su centro que asemeja al globo ocular humano.

Con el fin de emular el movimiento del ojo, este prototipo robótico ofrece dos GDL rotacionales y capacidades de procesamiento digital de imágenes para aplicaciones en el campo de visión por computadora.

En este artículo, el diseño, implementación y cinemática de la estructura fueron descritos así como una aplicación demostrativa de procesamiento de imágenes: detección y seguimiento de objetos (pelotas). El prototipo es

compacto, ligero y de bajo costo lo que lo hace viable para aplicaciones en robótica móvil, teleoperación, seguridad y vigilancia entre otros.

Las perspectivas de trabajo futuro se concentran en la parte de visión computacional: optimización de algoritmos para minimizar la carga computacional y la detección de rostros para aplicaciones de seguridad y biometría.

VI. Referencias

- [1] G. Cannata and E. Grosso, "On perceptual advantages of active robot vision", *Journal of Robotic Systems*, 16(3), pp 163-183, 1999.
- [2] J.A. Aguilera, R. Velázquez y J. Varona, "Plataforma de teleoperación y telepresencia para aplicaciones de seguridad y vigilancia con robots móviles", *Memorias de: 6° Semana Nacional de Ingeniería Electrónica*, Huajuapán de León, Oax., México, 2010.
- [3] C. Gosselin and F. Caron, "Two degree-of-freedom spherical orienting device", *US Patent No. 5966991*, 1999.
- [4] M. Karouia, "Conception structurale de mécanismes parallèles sphériques", *Tesis doctoral de la Escuela Central de Paris*, 2003.
- [5] C. Gosselin and J. Angeles, "The optimum kinematic design of a planar three-degree-of-freedom parallel manipulator", *Journal of Mechanisms, Transmissions, and Automation in Design*, 110, 1988.
- [6] R. Gonzalez, R. Woods, "Digital image processing", 2° edición, Prentice Hall, New Jersey, USA, 1992.
- [7] M. Rizon et al., "Object detection using circular hough transform", *American Journal of Applied Sciences*, 2(12), pp 1606-1609, 2005.

VII. Autores

Ing. José Arce y de la Borbolla es Ingeniero en Mecatrónica y Especialista en Robótica por la Universidad Panamericana.

Ing. Juan Carlos Ruiz Cruz es Ingeniero Electromecánico y Especialista en Robótica por la Universidad Panamericana.

Dr. Mario Acevedo Alvarado es Ingeniero Electromecánico por la Universidad Panamericana y Doctor en Mecánica por la Universidad de Navarra, España. Actualmente es profesor-investigador de la Universidad Panamericana campus México.

Dr. Ramiro Velázquez Guerrero es Ingeniero en Electrónica por la Universidad Panamericana, Maestro en Ciencias por el INSA de Lyon y Doctor en Robótica por la Universidad Pierre et Marie Curie. Actualmente es profesor-investigador de la Universidad Panamericana campus Bonaterra.

Diseño de un Control Óptimo para un Robot Móvil

*David Rodríguez Ortiz, **José. E. Moisés Gutiérrez Arias,
María Montserrat Morín Castillo, *José E. Flores Mena,
*****Gerardo Mino Aguilar

Facultad de Ciencias de la Electrónica.
Benemérita Universidad Autónoma de Puebla.
Av. San Claudio y 18 Sur, Col. San Manuel, C.P. 72570, Puebla, Pue., México.

*drodriguez@ece.buap.mx
** jmgutierrez@ece.buap.mx
***mmorin@ece.buap.mx
****eflores@ece.buap.mx
*****gerardo.mino@ece.buap.mx

Resumen

En este trabajo presentamos un robot móvil autónomo provisto de dos ruedas activas y una pasiva, así como dos algoritmos de control para la estabilización de las trayectorias programadas; la programación dinámica es el fundamento para determinar ambas leyes de control. La primera ley de control óptimo la obtenemos al solucionar una ecuación diferencial matricial del tipo Riccati, la segunda ley se deduce al aprovechar una disertación hecha por Kalman, la cual permite reducir una ecuación diferencial matricial a una ecuación algebraica matricial. La simulación del algoritmo se realiza cuando la trayectoria programada es una línea inclinada 45° , la cual permite observar que la ley de control óptimo, objetivo primordial de este artículo, presenta una calidad superior en la estabilización que la ley de control obtenida mediante la disertación de Kalman.

Palabras Clave — Control óptimo, Ecuación diferencial matricial de Riccati, Programación dinámica, Sistemas lineales variantes en el tiempo, Robot móvil.

I. Introducción

Como rama de la inteligencia artificial, la robótica móvil tiene grandes avances vistos durante las décadas pasadas, principalmente adentro la formalización matemática de diverso determinista y algoritmos no deterministas, así como la creación de nuevas teorías que complementan los conceptos que existieron ya dentro de la inteligencia artificial. Estas teorías hacen posible alcanzar metas que mejoran la autonomía y la inteligencia del robot móvil [1].

Las máquinas que llamamos robots están tomando cada vez más la mayor importancia en la vida del hombre, los seres humanos diseñan y construyen estas máquinas para ayudarnos a

realizar varias actividades, por ejemplo la manipulación de los materiales peligrosos, tareas que son más allá de la capacidad natural de seres humanos y de actividades en los ambientes donde se pone en peligro la vida humana. Cuando hablamos acerca de los robots móviles, nos referimos a una clase particular de agentes inteligentes para el que su entorno es la interacción del contexto físico que rodea al robot con el material objeto. Los estímulos que provienen de la medida de los sensores de las propiedades físicas (distancia, tamaño, color, intensidad luminosa, etc.) de los objetos, y de las respuestas son actos físicos sobre este entorno (el movimiento propio del robot) [2].

Esto causa algunas dificultades en los robots móviles, consistiendo principalmente en la incertidumbre y el error implicados en la transformación de la comprobación de las magnitudes medidas por los sensores, porque estos valores se utilizan como entrada al sistema que controla al robot.

Particularmente, como el área de control es un subsistema que gobierna la actividad del robot móvil autónomo, ha aumentado su gama del estudio, generando los algoritmos que son más y más robustos. La autonomía de un robot móvil se basa sobre el sistema de navegación automático.

En estos sistemas, las tareas del planeamiento, de la opinión, y del control son incluidas. El problema de la planificación global de la trayectoria consiste en el hacer la trayectoria de menos longitud para alcanzar la meta. Esto implica un problema en el diseño del control que regula la misión del robot, que se considera en este artículo [3], [4]. En orden para hacer esto, consideramos la clase de robot móvil autónomo que consiste en tres ruedas, dos activos y una pasiva, con restricciones no holonómica, que están presentes debido al supuesto de que no se desliza.

El estudio y diseño de los sistemas de control para los robots móviles es muy amplio, así como también la cantidad de algoritmos de control con los que se puede encontrar una aplicación interesante. Aunque el robot móvil es un sistema con movimientos físicamente simples, matemáticamente no lo es, debido a que su modelo dinámico es no lineal. Considerando minimizar un índice de desempeño cuadrático para solucionar el problema, se aplica la solución estándar al problema de control óptimo para un sistema lineal, la cual puede demostrarse mediante la técnica de programación dinámica aplicada a la ecuación de Hamilton-Jacobi-Bellman, para obtener la ley de control óptima para un sistema lineal en tiempo continuo; en la estructura del control óptimo resultante, está inmersa la solución de la ecuación diferencial matricial de Riccati la cual contiene la información del sistema lineal y del índice de

desempeño cuadrático; esta matriz es la clave para la estabilización del sistema lineal [9], [10],[3]. La elección de las matrices de peso del índice de desempeño, juega un papel muy importante en la calidad de estabilización del sistema lineal. El control óptimo diseñado para el sistema linealizado alrededor de alguna trayectoria deseada, es posteriormente aplicado al sistema de ecuaciones de movimiento no lineal del robot móvil para obtener la trayectoria deseada.

Aunque el sistema en lazo cerrado construido usando la teoría de control óptimo tiene algunas ventajas, hay todavía muchos problemas que resolver. Uno de los más serios es que es bastante difícil especificar el desempeño de control descrito en términos de un índice de desempeño cuadrático; las matrices de peso usualmente serán especificadas en base a prueba y error para obtener un desempeño satisfactorio. Para el caso de sistemas grandes y sobre todo cuando no se conocen suficientemente las reacciones del sistema, resulta muy complicado asignar valores a todos los elementos de las matrices de peso; por esta razón es una práctica común elegir estas matrices como matrices diagonales.

En la literatura de control óptimo para robots móviles, es posible mencionar a la referencia [5], donde se plantea el problema de encontrar esta ley de control considerando el modelo matemático a partir del error de orientación y el error de posición para una trayectoria deseada, obteniendo una ecuación lineal de cuatro variables de estado, además haciendo una transformación sobre los valores característicos se obtiene un segundo sistema de dos variables de estado las cuales representa los modos rápidos y lentos del sistema; el control óptimo es calculado solo para este segundo sistema sin hacer referencia a la elección de las matrices de peso.

Para la solución del control óptimo del sistema lineal invariante en el tiempo, se puede mencionar el trabajo presentado en [11], donde se plantea el problema de encontrar un controlador óptimo cuadrático tal que el sistema óptimo en

lazo cerrado tiene valores característicos dentro de una región vertical en el semiplano izquierdo del plano complejo s .

Con el objetivo de mejorar la estabilidad del sistema el método de diseño no necesita la especificación de las matrices de peso. Este método también es utilizado en [12] y [13] donde es aplicado para mejorar la estabilidad dinámica de sistemas de estabilización de potencia, obteniendo resultados muy satisfactorios

En la referencia [6], se considera un robot móvil de conducción diferencial de dos ruedas situadas en el centro, el modelado se hace usando la formulación de Lagrange. En este artículo, la dinámica de la parte eléctrica (los motores) puede generalmente descuidarse, como constantes de tiempo eléctricas suelen ser mucho más pequeñas que las constantes de tiempo mecánico. No presentan cualquier estrategia de control. Este artículo tiene por objeto, la deducción de un modelo dinámico, estos son muy accesibles pero al mismo tiempo, la captura de las no linealidades esenciales del robot; sintetizar dos estrategias de control óptimo para planear una trayectoria del robot, teniendo en cuenta un funcionamiento de un indicador cuadrático. La primera ley que se obtiene al resolver la ecuación diferencial matricial de Riccati, la segunda ley de control se dedujo del trabajo previo realizado por Kalman.

La segunda estrategia de control la conseguimos para la comparación entre las dos y demostramos que la estrategia obtenida solucionando la ecuación diferencial matricial tiene una estabilización más alta calidad.

Este trabajo se organiza de la siguiente manera. Sección. 2 es dedicado a la declaración del problema de control. En la sección. 3, se deduce el modelo matemático del robot móvil. En Sección. 4, se describe las trayectorias deseadas mediante las ecuaciones lineales obtenidas del modelo dinámico del robot móvil. En la sección. 5, se deduce el control mediante la herramienta de la programación dinámica y de esta mediante ello obtener las trayectorias programadas. En la sección. 6, las soluciones y resultados obtenidos en el tiempo normal y el

tiempo infinito además de sus respectivas simulaciones. En la sección 7 se presentan las conclusiones obtenidas. Por último, en la sección. 8, se presentan las referencias.

II. Establecimiento del Problema

Considerando el siguiente proceso controlable

$$\begin{aligned} \dot{y} &= f(y, u), \\ u(\cdot) &\in U = \{u: u(t) \in \Omega \subseteq \mathcal{R}^r\}; \end{aligned} \quad (1)$$

donde $y \in \mathcal{R}^n$ contiene los estado del sistema, $u \in \mathcal{R}^m$ representa los controles de entrada. Las ecuaciones diferenciales lineales que gobiernan las desviaciones $x(t) = y(t) - y^d(t)$ para algún movimiento deseado $y^d(t)$ y un control deseado $u^d(t), \in [t_0, t_1]$ pueden escribirse como:

$$\dot{x} = A(t)x + B(t)u; \quad (2)$$

donde

$$A(t) = \frac{\partial f[y^d(t), u^t]}{\partial y}, B(t) = \frac{\partial f[y^d(t), u^t]}{\partial u};$$

y considerando también el criterio

$$\int_{t_0}^{t_1} [x^T Q(t)x(t) + u^T R(t)u(t)]dt + x^T(t_1)F_x(t_1), \quad (3)$$

donde F y Q son matrices simétricas definidas no negativas ($F = F^T \geq 0; Q(t) = Q(t)^T \geq 0$); y $R(t)$ es una matriz simétrica definida positiva ($R(t) = R(t)^T > 0$) para $t_0 \leq t \leq t_1$. Entonces el problema de determinar un vector de entradas admisibles $u^0(t), t_0 \leq t \leq t_1$, para el cual el criterio es *mínimo* es llamado *problema del control óptimo lineal*.

III. Ecuaciones de Movimiento del Robot Móvil

El tipo de diseño del robot móvil que se utilizará a lo largo de este trabajo será el de *control diferencial*, el cuál como ya se mencionó, consta de dos ruedas activas para tracción y dirección, y una rueda pasiva (ver figura 1).

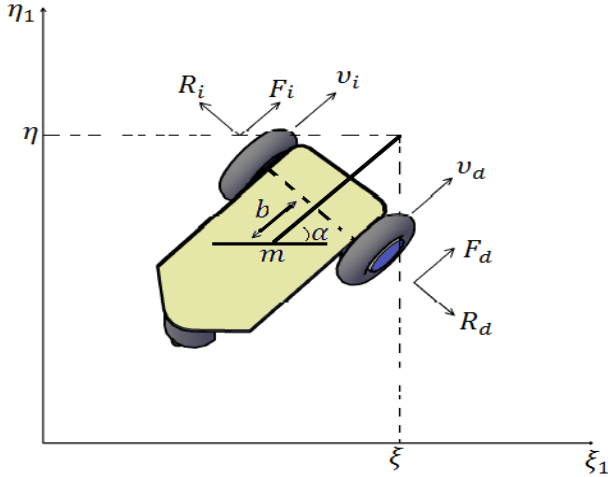


Fig. 1. Diagrama físico del Robot Móvil.

Considerando que el robot sólo presenta movimientos planeares translacionales y rotacionales, y que no presenta ningún tipo de deslizamientos lo que lleva a las restricciones no holonómicas; podemos, utilizando principios básicos de la mecánica, obtener las siguientes ecuaciones que describen la cinemática del robot móvil dentro del marco de referencia fijo ξ y η :

$$\dot{\xi} = v \cos(\alpha) - h w \sin(\alpha), \quad (4)$$

$$\dot{\eta} = v \sin(\alpha) - h w \cos(\alpha), \quad (5)$$

$$\dot{\alpha} = w, \quad (6)$$

El robot es un sistema no-holonómico porque las velocidades de las ruedas activas deben tener la dirección P_η . Utilizando el *método de Newton*, y sustituyendo las restricciones por las fuerzas reactivas \vec{N}_d y \vec{N}_i , se obtienen las ecuaciones dinámicas que involucran las fuerzas de empuje \vec{F}_d y \vec{F}_i :

$$\begin{pmatrix} F_i + F'_i \\ N_i + N'_i \end{pmatrix} = m_{wi} \begin{pmatrix} \ddot{X}_{r1} - \frac{T}{2} \ddot{\theta} \\ \ddot{X}_{r2} - \frac{T}{2} \dot{\theta}^2 \end{pmatrix} \quad (7)$$

$$\begin{pmatrix} F_d + F'_d \\ N_d + N'_d \end{pmatrix} = m_{wd} \begin{pmatrix} \ddot{X}_{r1} - \frac{T}{2} \ddot{\theta} \\ \ddot{X}_{r2} - \frac{T}{2} \dot{\theta}^2 \end{pmatrix} \quad (8)$$

es necesario además, asociar las fuerzas F_d y F_i , las masas y aceleraciones asociadas a cada una de las ruedas que determinan el giro de cada uno de los motores del robot móvil; para esto, se sabe que el torque o momento de un motor está dado por

$$\left(m_c + 2m_w + 2 \frac{I_{2w}}{r^2} \right) \ddot{X}_{r1} - m_c C \dot{\theta}^2 = \frac{\tau_d - \tau_i}{r}, \quad (9)$$

$$\begin{aligned} \left(I_{3c} + 2I_{3w} + m_w \frac{T^2}{2} + m_c C^2 + \frac{T}{2r} I_{2w} \right) \ddot{X}_{r2} \\ = \frac{T}{2r} (\tau_d - \tau_i) \end{aligned} \quad (10)$$

es necesario también, un modelo matemático que describa el movimiento del motor, para esto el modelo más simple conocido para un motor de DC es:

$$(\tau_d + \tau_i) = x(u_d - u_i) - \sigma \frac{2v}{r}, \quad (11)$$

$$(\tau_d + \tau_i) = x(u_d - u_i) - \sigma \frac{T_w}{r}, \quad (12)$$

donde τ es el torque del motor en cada rueda (Nm); $(u_d - u_i)$ es el voltaje de control aplicado a las terminales del motor (V) y σ es la fuerza contra-electromotriz del motor, y considerando las ecuaciones (9), (10) y sustituyendo en las expresiones (11) y (12) se pueden obtener una expresión asociada a las ruedas derecha e izquierda en las ecuaciones de movimiento

$$\begin{aligned} \left(m_c + 2m_w + 2 \frac{I_{2w}}{r^2} \right) \dot{v} - m_c b w^2 = \\ \frac{x}{r} (u_d - u_i) - \sigma \frac{T_w}{r}, \end{aligned} \quad (13)$$

$$\begin{aligned} \left(I_{3c} + 2I_{3w} + m_w \frac{T^2}{2} + m_c b^2 + \frac{T}{2r} I_{2w} \right) \dot{w} = \\ \frac{T}{2r} x(u_d - u_i) + \frac{T}{2r} \sigma w, \end{aligned} \quad (14)$$

Por lo que las ecuaciones (4), (5), (6), (13) y (14) describen el movimiento del robot móvil:

$$\begin{aligned} \dot{\xi} &= v\cos(\alpha) - hwsin(\alpha), \\ \dot{\eta} &= vsin(\alpha) + hwc\cos(\alpha), \\ \dot{\alpha} &= w, \\ (J_1)\dot{v} &= m_cbw^2 + \frac{x}{r}(u_d - u_i) - \sigma\frac{T_w}{r}, \\ (J_2)\dot{w} &= \frac{T}{2r}x(u_d - u_i) + \frac{T}{2r}\sigma w, \end{aligned} \quad (15)$$

IV. Trayectorias deseadas y ecuaciones lineales.

Por medio del análisis de los posibles movimientos del robot en el plano, se pueden obtener las trayectorias deseadas o programadas que este puede realizar. Considerando que el robot presenta movimiento de plano inclinado 45° obtenemos por derivadas parciales la siguiente matriz de Estado (16) para obtener la trayectoria deseada (ver figura 2).

$$A = \begin{pmatrix} 0 & 0 & v\cos(\alpha_0) & \sin(\alpha_0) & h\cos(\alpha_0) \\ 0 & 0 & vsin(\alpha_0) & \cos(\alpha_0) & -h\cos(\alpha_0) \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \frac{-2\sigma}{r^2J} & 0 \\ 0 & 0 & 0 & 0 & \frac{T}{2lr}\sigma \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{x}{rJ_1} & \frac{x}{rJ_2} \\ \frac{T_x}{2lr} & -\frac{T_x}{2lr} \end{pmatrix} \quad (16)$$

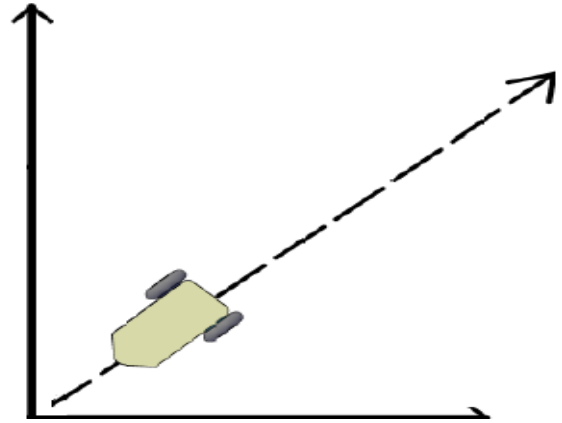


Fig. 2. Trayectoria inclinada con ángulo de 45° del robot móvil.

V. Deducción del Control mediante la programación dinámica.

Considerando el problema de control óptimo

$$\begin{aligned} \min_{u(t)} J &= \int_{t_0}^{t_1} [x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)]dt \\ &+ x^T(t_1)Fx(t_1), \end{aligned}$$

$$\begin{aligned} \text{suje}to \ a : \dot{x}(t) &= A(t)x(t) + B(t)u(t), \\ \text{con} \quad : x(t_0) &= x_0; \end{aligned} \quad (17)$$

se define la función valor $W^*(x, t)$, que proporciona el mínimo valor del funcional objetivo del problema de control, para solucionar éste problema de control óptimo, primeramente se escribe la ecuación de Hamilton-Jacobi-Bellman:

$$0 = W_t(x, t) + \min_{u(t)} \{x^T(t)Qx + u^T Ru(t)\nabla_x W(x, t)dt[Ax + Bu]\}$$

$$\text{con} : W(x, t_1) = x^T Fx, \quad (18)$$

ahora, se establece la suposición de que la solución a la ecuación (18) es de la forma:

$$W(x, t) = x^T P(t)x, \quad (19)$$

donde $P(t)$ es una matriz simétrica n-dimensional. Obteniendo:

$$\frac{\partial W(x, t)}{\partial t} = W_t(x, t) = x^T \frac{\partial P(t)}{\partial t} x, \quad (20)$$

$$\frac{\partial W(x, t)}{\partial x} = \nabla_x(x, t) = 2x^T P(t), \quad (21)$$

y sustituyendo (20) y (21) en la ecuación (22) de Hamilton-Jacobi-Bellman, se tiene que:

$$0 = x^T \frac{\partial P(t)}{\partial t} x + \frac{\min}{u(t)} \{x^T Q x + u^T R u + 2x^T P(t)[Ax + Bu]\}, \quad (22)$$

derivando parcialmente con respecto a u la expresión entre llaves e igualando a cero se obtiene:

$$\begin{aligned} 2u^T R &= -2x^T P(t)B, \\ u &= -R^{-1}B^T P(t)x. \end{aligned} \quad (23)$$

Sustituyendo (23) en la ecuación (22), se tiene:

$$-\frac{\partial P(t)}{\partial t} = Q - P(t)BR^{-1}B^T P(t) + P(t)A(t) + A^T P(t),$$

$$\text{con : } P(t_1) = 0, \quad (24)$$

que es una ecuación matricial de Riccati, donde $P(t)$ se encuentra resolviendo ésta ecuación.

Por lo tanto, el control óptimo del problema (24) es

$$u(t) = -Kx(t), \quad K = R^{-1}B^T P(t); \quad (25)$$

donde esta ley de control resuelve el problema de estabilización; la solución $P(t) \geq 0$ de la ecuación de Riccati (24) es llamada solución de estabilización; y la matriz K definida en (25) es la matriz de ganancia del control óptimo o la matriz de estabilización óptima.

Finalmente, la ecuación del sistema del control óptimo en lazo cerrado está dada por

$$\dot{x} = [A(t) - B(t)K]x(t), \quad (26)$$

VI. Solución y Resultados

El valor de cada uno de los parámetros del sistema necesarios para obtener los resultados se presenta en la tabla (1).

Variable	Valor	Descripción
v_0	0,5	Velocidad deseada, (m/s).
w_0	0	Velocidad angular deseada, (m/s).
h	0,25	Distancia que hay entre el frente del robot móvil y el eje de las ruedas, (m).
m	3,0	Masa del Robot Móvil, (Kg).
b	0,04	Distancia del centro de masa al eje de las ruedas, (m).
T	0,27	Distancia media entre las dos ruedas, (m).
r	0,05	Radio de las ruedas activas, (m).
σ	0,00196	Fuerza contra-electromotriz
J	0,00549	Momento de inercia,
x	0.00419	Fricción Viscosa.

Tabla 1. Parámetros físicos del Robot Móvil.

El algoritmo de solución para la obtención del control

1. Obtener el sistema lineal $\dot{x}(t) = A(t)x(t) + B(t)u(t)$, alrededor de alguna trayectoria deseada.
2. Verificar si el par de matrices (A, B) del sistema lineal $\dot{x}(t) = A(t)x(t) + B(t)u(t)$ es completamente controlable, si no, no es posible obtener el control óptimo.
3. Desarrollo y solución de la ecuación diferencial matricial de Riccati con condiciones finales (24), una vez que se han escogido valores adecuados para las matrices de peso Q y R . Para la solución de la ecuación diferencial de Riccati se realiza un cambio de variable para el tiempo ($\tau = t_f - t$), para convertir este problema en uno de condiciones iniciales.

4. Obtener mediante polinomios una aproximación a cada una de las curvas de solución de la matriz diferencial de Riccati; posteriormente se hace la inversión de estas curvas en el tiempo para obtener el resultado a condiciones finales buscado.
5. Obtener el control óptimo (25).
6. Se obtienen las gráficas del comportamiento del sistema bajo el control óptimo diseñado mediante la simulación de la solución de la ecuación del sistema en lazo cerrado (26).

A continuación, se presenta la solución de la ecuación diferencial matricial del tipo Riccati, para obtener el control óptimo para el robot móvil cuando el tiempo es finito con una trayectoria correspondiente a una inclinada 45°. Teniendo el conjunto de ecuaciones lineales, $\dot{x}(t) = A(t)x(t) + B(t)u(t)$ se puede obtener utilizando los parámetros de la tabla (1) para dicha trayectoria de 45 _ en donde las matrices A y B del sistema $\dot{x}(t) = A(t)x(t) + B(t)u(t)$ tienen la siguiente forma:

$$A = \begin{pmatrix} 0 & 0 & 0,2627 & 0,8509 & 0,0263 \\ 0 & 0 & 0,4255 & 0,5253 & -0,0425 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -285,6102 & 0 \\ 0 & 0 & 0 & 0 & 0,0054 \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 15,2641 & 15,2641 \\ 0,0115 & -0,0115 \end{pmatrix}$$

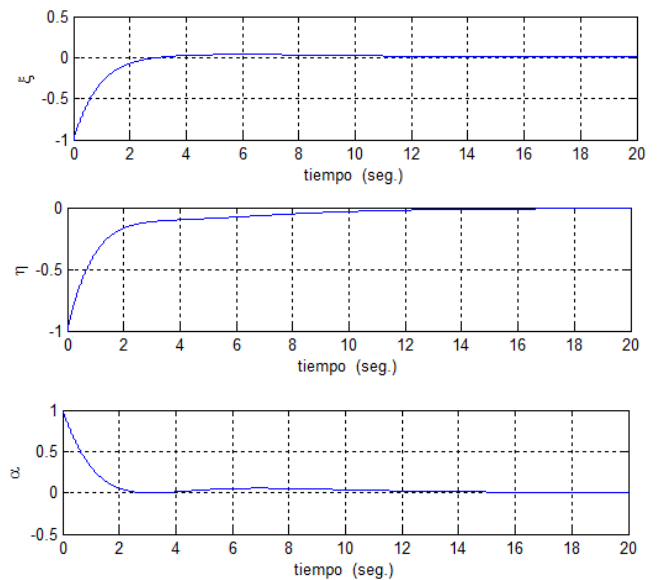
En la solución y diseño de un control óptimo, la matriz K, se calcula con una apropiada selección de las matrices de peso Q y R, y usualmente se determinan a prueba y error hasta lograr el comportamiento deseado. Cuando se trabaja con sistemas grandes este proceso puede no ser muy fácil, por esta razón es común elegir a estas matrices como diagonales.

Para el diseño del control óptimo, se tomará la matriz de peso $R^{2 \times 2}$ como una matriz identidad, con lo cual se asigna igual peso a las señales de control y se logra que cada una de ellas actúe en forma independiente. Para el caso de la matriz de peso $Q^{5 \times 5}$, se tomará como matriz identidad.

$$R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad Q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Los resultados de la simulación de cada variable de estado del sistema para el procedimiento descrito anteriormente, se muestran en la figura (3) para el tiempo finito, y tomando en consideración cuando el tiempo es infinito figura (4). Tomando en ambos casos condiciones iniciales para la simulación del sistema en lazo cerrado $x(0) = [-1; -1; 1; 1; -1]$.

Se presentan las graficas como resultado para cuando el tiempo es finito:



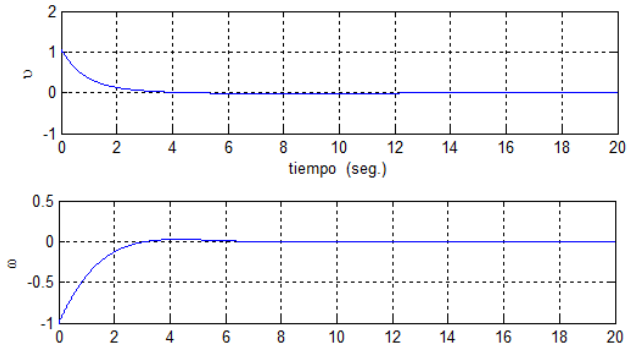


Fig. 3. Simulación del sistema en lazo cerrado para tiempo finito.

Así entonces, para obtener la solución de la ecuación diferencial de Riccati en tiempo infinito, se hace uso de la herramienta en Matlab $[P, L, G] = \text{care}(A, B, Q)$ ya que realiza la solución única P de la ecuación algebraica matricial de Riccati.

$$0 = Q - PBR^{-1}B^T P + PA + A^T P, \quad (28)$$

sujeto al sistema lineal invariante en tiempo

$$\dot{x} = Ax + Bu(t), \quad (29)$$

La función *care* también retorna la matriz de ganancia,

$$K = R^{-1}B^T P,$$

Se presentan las graficas como resultado para cuando el tiempo es infinito la matriz P se vuelve constante:

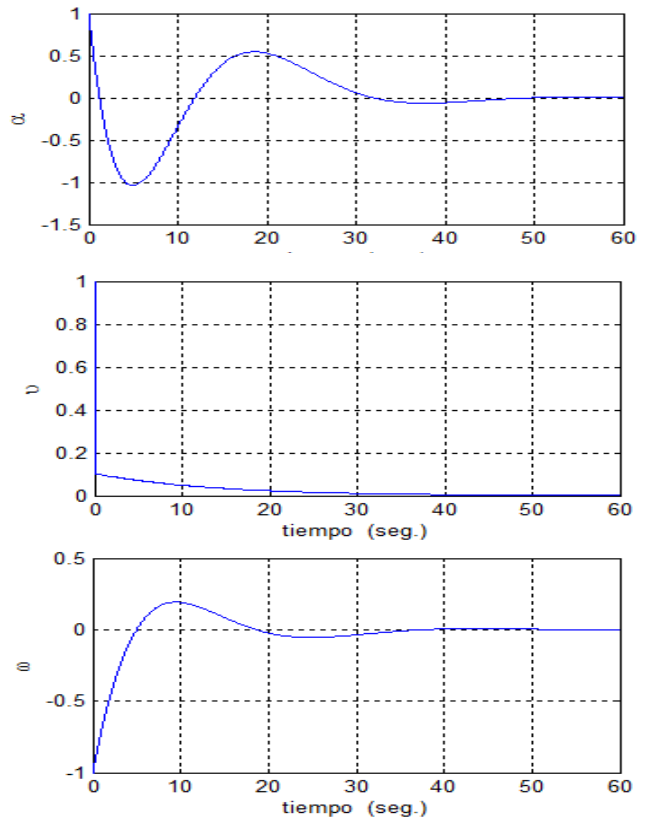
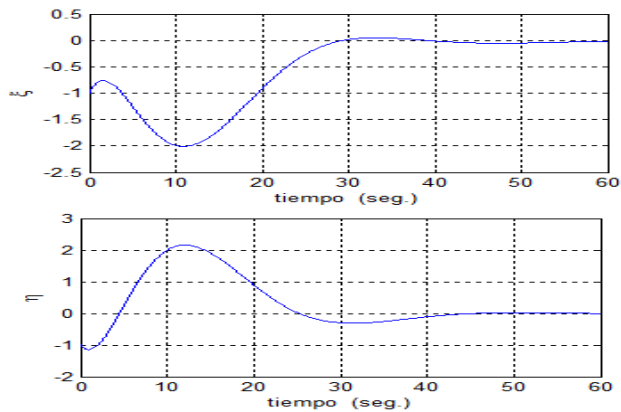


Fig. 4. Simulación del sistema en lazo cerrado para tiempo infinito.

VII. Conclusiones

Las matrices de peso en el funcional objetivo son un elemento primordial en el desarrollo de un control óptimo; ya que estas determinan que tan estable será el sistema dinámico.

Se obtienen simulaciones para el tiempo finito y se observa claramente que el sistema se estabiliza en menor tiempo, pero esto requiere mayor coste computacional y por otra parte en la simulación con el tiempo infinito ∞ se observa que el sistema se estabiliza en un tiempo mucho mayor y el coste computacional es mínimo.

VIII. Referencias

- [1] A. Ollero Baturone, *Robótica Manipuladores y Robots Móviles* (Marcobombo, 2001).
- [2] R. Siegwart and I.R. Nourbakhsh, *Introduction to Autonomous Mobile Robots* (2004).
- [3] D. E. Kirk, *Optimal Control Theory An Introduction* (Printice-Hall, 1970).
- [4] G. Knowles, *An Introduction to Applied Optimal Control* (New York and London Academic Press, 1981).
- [5] E A. Hemani, M.G. Mehrabi, and R.M.H. Cheng *Automatica* (1991).
- [6] G. Klancar, B. Zupancic, R. Karba *Modelling and simulation of a group of mobile robots (Simulation Modelling Practice and Theory 15, ELSEVIER, 2007)*. pp. 647.
- [7] T. Bräunl, *Embedded Robotics*, Springer, 2a Ed., 2006, págs. 3-5, 41-44, 97-105.
- [8] F. Beer, R. Johnston, W. Clausen, *Mecánica Vectorial para Ingenieros: Dinámica*, McGraw Hill, 8_ Ed., 2007.
- [9] Kwakernaak H., Sivan R., 1972. *Linear Optimal Control Systems*. WileyInterscience, New York.
- [10] Cerda E., 2001. *Optimización Dinámica*. Prentice Hall, España.
- [11] Shieh L.S., Dib H.M., 1986. *Linear Quadratic Regulators with Eigenvalue Placement in a Vertical Strip*. IEEE Trans. on Automatic Control Vol. AC-31 No. 3, 241-243.
- [12] Araya P., Eddie A., 1999. *Coordinación de PSS y SVC para mejorar la estabilidad dinámica*. Ingeniería.
- [13] Hardiansyah, Furuya S., Irisawa J., 1999. *Optimal Power System Stabilization via Output Feedback Excitation Control*.

Journal of Nagaoka University of Technology, Japan, 21-28.

IX. Autores

Ing. David Rodríguez Ortiz es Ingeniero en Electrónica por el Instituto Tecnológico de Minatitlán, Veracruz; en el año 2008.

Dr. José Eligio Moisés Gutiérrez Arias obtuvo el grado de Doctor en la Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla en el año 2003.

Dra. María Montserrat Morín Castillo obtuvo el grado de Doctora en la Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla.

Dr. José Eladio Flores Mena obtuvo el grado de Doctor en la Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla.

Dr. Gerardo Mino Aguilar obtuvo el grado de Doctor en la Universidad Politécnica de Catalunya; España, en el año 2007.

Estudio de Viabilidad de un Nuevo Sistema de Puertas para el Transporte Colectivo Metro

¹Antonio Gómez González, ¹Ramiro Velázquez Guerrero, ²Miguel Magos Rivera

¹ Universidad Panamericana campus Bonaterra
Josemaría Escrivá de Balaguer 101, Fracc. Rústicos Calpulli,
CP 20290, Aguascalientes, Ags., México

Email: {asgomez, rvelazquez}@up.edu.mx

² Universidad Autónoma Metropolitana – Azcapotzalco
Ave San Pablo 180, Col. Reynosa Tamaulipas,
CP 02200, México, DF., México
Email: mrm@correo.azc.uam.mx

Resumen

La constante modernización del Metro de la ciudad de México es una actividad apremiante para satisfacer las necesidades de transporte de las cinco millones de personas en promedio que lo utilizan a diario. Entre los muchos rubros a mejorar se encuentran las puertas que permiten a los usuarios el acceso a los vagones. En este trabajo se presenta un estudio de las características y los puntos críticos más importantes en el sistema de apertura y cierre de puertas existente y se propone un nuevo mecanismo que permitirá responder más eficazmente a las necesidades actuales del Metro.

Palabras Clave: análisis de fallas, sistema de apertura/cierre de puertas, mecanismo sinfín, mecanismo con banda, análisis con método de elementos finitos.

I. Introducción

El sistema de transporte colectivo Metro de la ciudad de México es un sistema público que atiende en promedio a cinco millones de personas diariamente [1]. Por el número de usuarios que mueve a lo largo y ancho de una de las metrópolis más grandes del mundo, es considerado uno de los sistemas de transporte más importantes de la ciudad capital. Esta gran cantidad de usuarios requiere un servicio eficiente y seguro para cumplir con sus actividades cotidianas.

El Metro fue inaugurado hace más de 40 años y, aunque se ha ido modernizando gradualmente a lo largo de este tiempo, varios sistemas tanto mecánicos como eléctricos aun están basados en los conceptos originales. Según reportes internos del Metro, las averías más frecuentes se dan en las puertas de los vagones por el constante uso y

contacto con los usuarios. Otras causas de fallas comunes se mencionan en la Tabla 1.

Fallas	No. de fallas
Sistema de puertas	77
Mando y control	24
Causas ajenas	13
Generación de aire	18
Pilotaje automático	26
Tracción frenado	8
Equipos mecánicos	1
Generación de energía	3
Total	170

Tabla 1. Relación de ocurrencia de fallas en el Metro de la Ciudad de México de enero a diciembre 2009. (Línea 3), [2].

Cuando el sistema de apertura de alguna de las puertas del vagón deja de funcionar, es necesario retirar de circulación todo el vagón en cuestión, afectando tanto la logística del Metro como

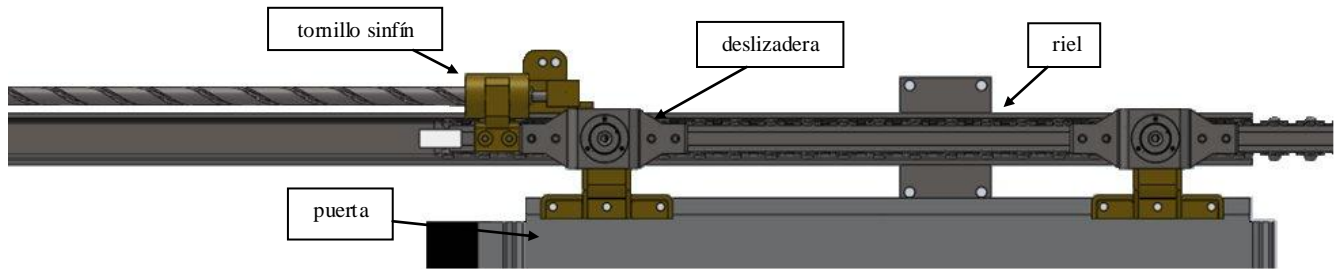


Fig. 2. Detalle del mecanismo tornillo sinfín del sistema de apertura de puertas actual del Metro.

a los usuarios por la alteración del servicio. Una vez en mantenimiento, muchas veces es necesario reemplazar las piezas por repuestos suministrados por proveedores extranjeros exclusivos. Esto implica gastos para el gobierno de la ciudad de México que se podrían utilizar en otras áreas y servicios del Metro.

En este contexto, este artículo presenta una evaluación detallada del mecanismo actual de apertura y cierre de puertas del Metro, se analizan todos los factores de falla y se propone un nuevo sistema con el objetivo de mejorar el servicio, bajar los costos de reparación, tiempos de mantenimiento así como generar conocimiento práctico para el sistema de transporte colectivo Metro del Distrito Federal.

II. Análisis de mecanismo actual

El sistema actual de apertura y cierre de puertas de los vagones del Metro de la Ciudad de México se muestra de manera general en la fig. 1 y se detalla en la fig. 2.

Éste consiste básicamente de un mecanismo de tornillo sinfín que desplaza simultáneamente y en sentido contrario dos pares de deslizaderas a la cuales están sujetas las puertas. Cada par corre por un riel y su movimiento se traduce en la apertura y cierre de puertas. El actuador utilizado es un pistón neumático mismo que es alimentado por un compresor ubicado en el mismo vagón. Cada vagón cuenta con ocho puertas, de las cuales solo cuatro se activan para permitir el acceso al vagón. Las otras cuatro permanecen bloqueadas por seguridad de los pasajeros ya que dan acceso al vacío (vías del metro).

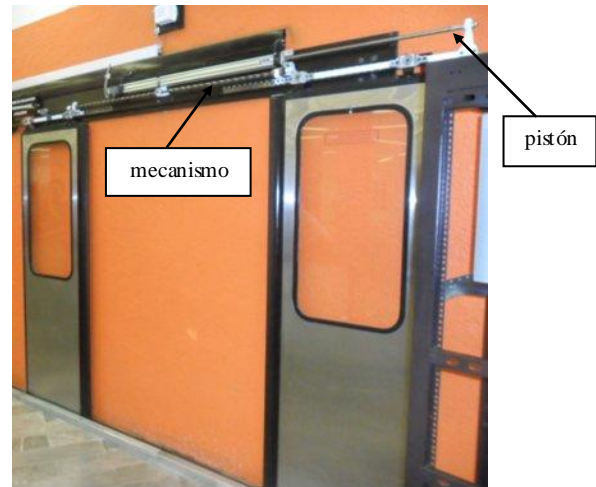


Fig. 1. Sistema de apertura y cierre de puertas.

Como se puede intuir, este diseño de mecanismo es muy sencillo y ha demostrado su eficiencia por más de 40 años.

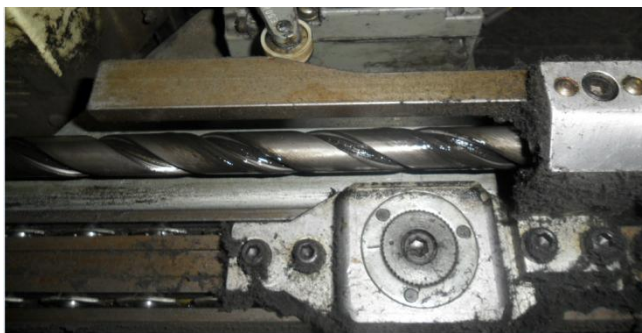
Sin embargo, las condiciones de uso del Metro han cambiado desde entonces. En efecto, la población de la ciudad se ha incrementado exponencialmente y ahora hay más gente que lo usa. En consecuencia:

- Más gente entra a los vagones y éstos viajan constantemente saturados. Los usuarios irremediamente ejercen un esfuerzo en las puertas que repercute en el mecanismo de apertura: las piezas se flexionan, no se mueven con la naturalidad con que debieran y el pistón comienza a presentar fugas en sus empaques.
- La suciedad (polvo y mugre) aumenta naturalmente con el número de usuarios. Si consideramos que los vagones cuentan con sistema de ventilación para abastecerlos de aire fresco, entonces la

suciedad se remueve por el aire del piso y viaja hacia las puertas del vagón ensuciando gradualmente el mecanismo de apertura. La fig. 3(a) muestra los efectos de la suciedad en el mecanismo.

Adicionalmente, recordemos que el mecanismo sinfín involucra el contacto entre dos sólidos (metales) por lo que existe un desgaste que se puede minimizar o retardar con el uso de lubricantes. Para este mecanismo que se utiliza constantemente, el no aplicar frecuentemente pastas lubricantes tiene consecuencias de desgaste y abrasión como las que se muestran en la fig. 3(b). Incluso, en ocasiones se tiene que substituir el mecanismo completo por esta causa.

El servicio de mantenimiento del Metro reporta que es necesario lubricar cada mecanismo de cada puerta por lo menos una vez cada tres meses. Si consideramos los 3042 vagones en funcionamiento en las 11 líneas que componen la red del Metro de la ciudad de México, el proceso de lubricación es una inversión en horas-hombre muy importante.



(a)



(b)

Fig. 3. Causas más comunes de falla en el sistema de apertura de puertas: (a) acumulación de suciedad y (b) signos de desgaste y abrasión en el husillo del mecanismo sinfín.

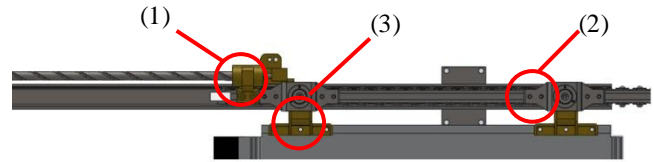


Fig. 4. Tres puntos críticos del mecanismo de apertura y cierre de puertas.

El propósito de este proyecto es hacer un estudio de las fallas en el mecanismo de apertura de puertas y proponer una solución tecnológica que permita eliminarlas o cuando menos que permita retardar la frecuencia de mantenimiento a una vez cada seis meses.

En lo que concierne al desgaste mecánico, es necesario identificar los puntos críticos del mecanismo. La fig. 4 revela tres puntos críticos: (1) el contacto entre el husillo y la tuerca con bolas del mecanismo sinfín, (2) el contacto entre la deslizadora y el riel y (3) la excéntrica, que es la pieza que une a la puerta con la deslizadora.

Para mejor apreciar los esfuerzos y los puntos exactos de desgaste en las piezas, se modelaron en computadora las piezas y se aplicó un análisis de esfuerzos y deformación con el método de elementos finitos. Paquetería especializada como Solidworks [3] y Ansys [4] se utilizó para estas tareas respectivamente.

La fig. 5 muestra el análisis con elementos finitos efectuado a la tuerca con bolas del mecanismo sinfín. La fig. 5(a) muestra la malla aplicada a la pieza mientras que la fig. 5(b) muestra la distribución de esfuerzos a lo largo de la pieza. Note que los puntos más críticos son el ensamble con el husillo y el ensamble con la deslizadora. Este análisis revela que los esfuerzos aplicados están dentro del rango de tolerancia de resistencia del material por lo que basta con lubricar esta pieza para asegurar un funcionamiento correcto.

La fig. 6 muestra el análisis efectuado al ensamble deslizadora-excéntrica. La fig. 6(a) muestra la malla aplicada al ensamble mientras que las figs. 6(b) y 6(c) muestran la distribución de esfuerzos y la deformación del ensamble respectivamente.

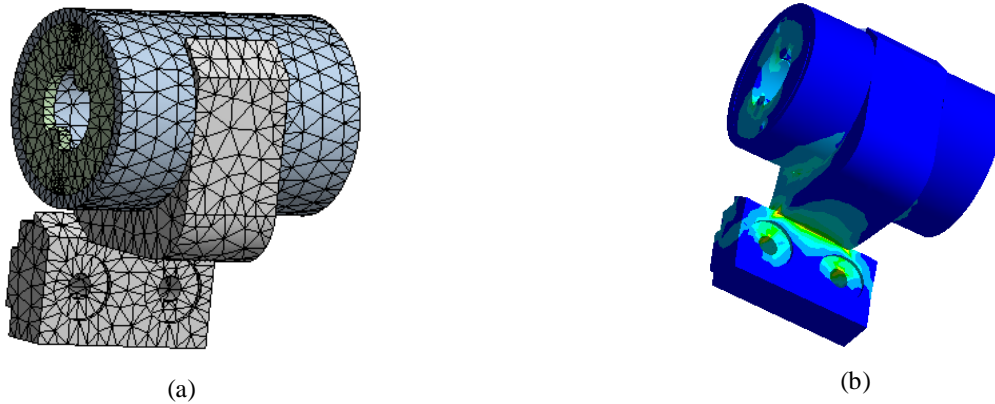


Fig. 5. Análisis de la tuerca con bolas del mecanismo sinfín con elementos finitos: (a) malla y (b) distribución de esfuerzos.

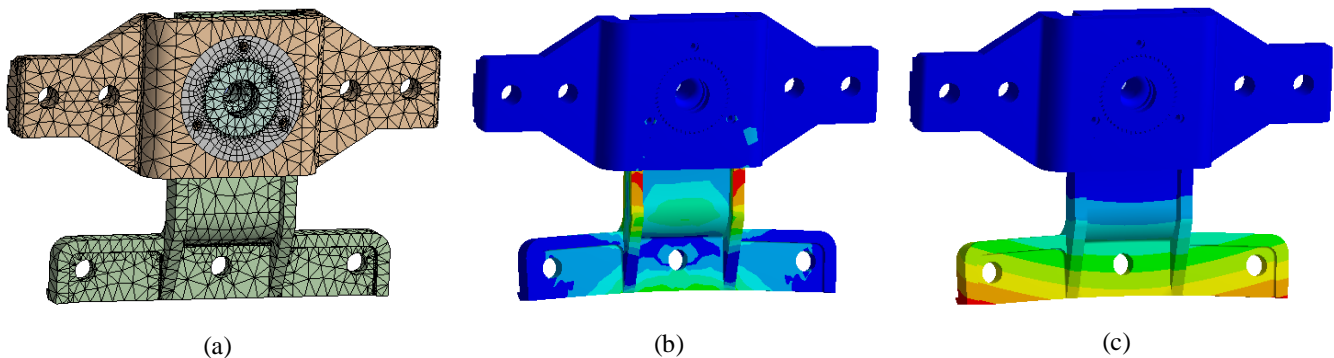


Fig. 6. Análisis de la excéntrica con elementos finitos: (a) malla y (b) distribución de esfuerzos y (c) deformación.

En la fig. 6(b) se observa que la parte del ensamble donde se aplica mayor esfuerzo es precisamente el punto de unión entre la deslizadora y la excéntrica. Este es el punto de fatiga una vez que se aplican esfuerzos perpendiculares al ensamble. Este tipo de esfuerzos son los que se ocasionan cuando los usuarios se apoyan o recargan en las puertas del Metro. Situación no ideal pero inevitable en el uso diario del Metro.

La fig. 6(c) muestra las deformaciones resultantes en el ensamble. Note que el área de acoplamiento con la puerta es la que presenta una mayor deformación una vez que se aplican fuerzas perpendiculares a la puerta.

Con el uso cotidiano, esta deformación se va incrementando hasta afectar los ensambles

deslizadora-riel y deslizadora-turca con bolas que bloquean el movimiento transmitido por el pistón. Una consecuencia del desalineamiento del mecanismo es que los empaques del pistón comienzan a presentar fugas.

III. Hacia un nuevo mecanismo de apertura y cierre de puertas

Con el fin de minimizar la fatiga en el mecanismo se han diseñado dos soluciones tecnológicas: una de corte conservador y otra más radical.

La primera consiste simplemente en sustituir el pistón neumático por un motor de corriente directa y reforzar el ensamble deslizadora-excéntrica.

En efecto, los argumentos de desechar completamente un mecanismo que lleva

funcionando 40 años puede no ser una idea viable de principio si se piensa términos industriales.

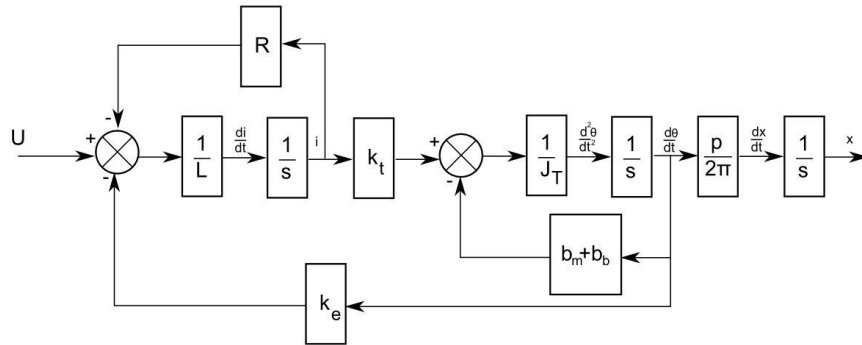


Fig. 7. Diagrama de bloques del mecanismo de puertas con un motor de DC como actuador.

Con el fin de seleccionar el motor adecuado en torque y velocidad es necesario establecer el modelo dinámico del mecanismo. Éste puede expresarse por las ecuaciones (1)-(3):

$$L \frac{di}{dt} = U - Ri - k_e \frac{d\theta}{dt} \quad (1)$$

$$J_T \frac{d^2\theta}{dt^2} = k_t i - (b_m + b_b) \frac{d\theta}{dt} \quad (2)$$

$$\text{con: } J_T = J_m + J_h + \frac{mp^2}{8\pi^2}$$

$$\frac{dx}{dt} = \frac{p}{2\pi} \frac{d\theta}{dt} \quad (3)$$

La ecuación (1) describe la parte eléctrica del motor de DC. Aquí U es el voltaje aplicado al motor, R , L y k_e son la resistencia eléctrica, la inductancia y la constante de fuerza contra-electromotriz del motor, respectivamente. Las variables i y θ son la corriente y el desplazamiento angular del rotor.

La ecuación (2) describe la parte mecánica del motor, aquí J_T es la inercia total que mueve el motor, k_t es la constante de torque del motor, b_m es el coeficiente de rozamiento interno en las balastras del motor mientras que b_b es el rozamiento entre el husillo y la tuerca con bolas. La inercia total J_T está compuesta por la inercia del rotor (J_m), la inercia del husillo (J_h) y la inercia equivalente de la puerta que está definida

por su masa m y por el paso de rosca del husillo del mecanismo de tornillo sin fin p .

Finalmente, la ecuación (3) describe la transformación de movimiento angular a lineal por el mecanismo sin fin.

El diagrama a bloques de la fig. 7 resume la dinámica del mecanismo de apertura de puertas. Con los datos físicos del mecanismo y la velocidad de apertura y cierre deseada, se procederá a la identificación de los parámetros del motor para seleccionarlo adecuadamente.

Por otro lado, la propuesta radical consiste en diseñar un mecanismo sencillo, completamente nuevo que requiera de un mínimo de materiales y que sea robusto para evitar que el polvo afecte su funcionamiento. Esta propuesta se considera radical puesto que de adaptarse, implicaría cambiar todos los mecanismos de las puertas de todos los vagones del Metro, acción que naturalmente se daría de manera gradual, implicaría una inversión considerable de inicio pero que a la larga retribuiría en vida operacional y costos de mantenimiento.

El nuevo diseño que se propone para abrir y cerrar las puertas del Metro se muestra en la fig. 8.

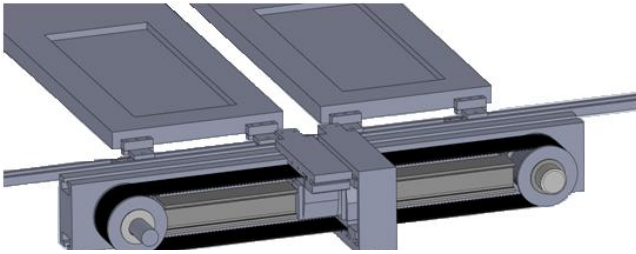


Fig. 8. Concepto de mecanismo con bandas (vista superior)

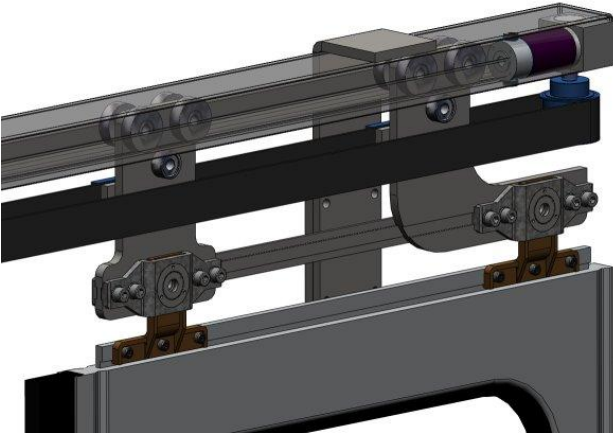


Fig. 9. Detalle del mecanismo con bandas.

Note que este diseño consta de una banda perfectamente bien tensada. Al moverse por medio del motor eléctrico, esta banda desplaza linealmente un par de deslizaderas que mueven las puertas (Fig. 9). Note que las deslizaderas están compuestas en su parte superior por cuatro ruedas que se acoplan por la parte interior y exterior de un carril, esta característica le da estabilidad y fuerza al diseño.

Una de las ventajas más significativas de este diseño es que las bandas pueden durar miles de ciclos, no usan lubricantes, requieren de poco mantenimiento, son de bajo costo [5] y si se encapsulan pueden evitar la acumulación de polvo y suciedad.

IV. Conclusiones

Este artículo abordó la problemática de desgaste en los mecanismos de apertura y cierre de las puertas del Metro de la Ciudad de México.

La observación cotidiana del mecanismo revela que las fallas principales se dan debido a esfuerzos no contemplados en el diseño original y a la acumulación de polvo y suciedad.

En este artículo se presentó un análisis con el método de elementos finitos de las piezas y ensambles críticos del mecanismo y se propusieron dos soluciones tecnológicas: 1) el uso de un motor de DC y el reforzamiento de los materiales que componen dichas piezas y ensambles y 2) un nuevo diseño basado en bandas.

Este último presenta ventajas atractivas para el sistema de transporte colectivo ya que es de bajo costo, larga vida de operación y requiere prácticamente de un mantenimiento nulo. Como trabajo futuro se contempla el análisis de este nuevo mecanismo con el método de elementos finitos y su implementación en prototipo.

V. Agradecimientos

Este trabajo se desarrolla dentro del marco del proyecto *Diseño y Construcción de Prototipo de un Nuevo Sistema de Puertas con Dispositivo Inhibidor de la Apertura de las Mismas del Lado Contrario*, apoyado por CONACYT-Gobierno del Distrito Federal con clave 120694.

VI. Referencias

- [1] Sistema de Transporte Colectivo-Metro de la Ciudad de México. Información actualizada disponible en: <http://www.metro.df.gob.mx/>
- [2] Reporte interno de actividades 2009, Metro de la Ciudad de México, México DF.
- [3] Dassault Systèmes. Información actualizada disponible en: <http://www.solidworks.com/>
- [4] ANSYS Inc. Información actualizada disponible en: <http://www.ansys.com/>

[5] R.L. Mott, Diseño de elementos de máquinas, Cuarta edición, Pearson Education, México, 2006.

VII. Autores

Ing. Antonio Gómez González es Ingeniero en Mecatrónica por la Universidad Panamericana. Actualmente se encuentra cursando la maestría en ciencias en la Universidad Panamericana de Aguascalientes.

Dr. Ramiro Velázquez Guerrero es Ingeniero en Electrónica por la Universidad Panamericana, Maestro en Ciencias por el INSA de Lyon (Francia) y Doctor en Robótica por la Universidad Pierre et Marie Curie (Francia). Actualmente es profesor-investigador en la Universidad Panamericana de Aguascalientes.

Dr. Miguel Magos Rivera es Ingeniero en Electrónica por la Universidad Autónoma Metropolitana, Maestro y Doctor en Control Automático por la Universidad Claude Bernard de Lyon (Francia). Actualmente es profesor-investigador en la Universidad Autónoma Metropolitana unidad Azcapotzalco.

Clasificación de Piezas Metálicas mediante Perfilometría Láser

Gonzalo Elías Blanco Silva, Alejandro González Álvarez, José A. Padilla Medina, Juan Prado Olivares, Agustín Ramírez Agundis.

Instituto Tecnológico de Celaya, Av. Tecnológico y A. García Cubas S/N, A.P.57, C.P. 38010, Celaya, Gto.
Tel. 01(461) 611 75 75, zaki_ottoyama@hotmail.com, agaitc@hotmail.com

Resumen

En el presente artículo se describe el desarrollo de un sistema de visión para la adquisición y clasificación de piezas metálicas sobre una banda transportadora mediante perfilometría láser. El sistema se basa en un programa que captura la imagen de forma segmentada para posteriormente reconstruirla y reconocerla para su clasificación de entre un conjunto de ocho piezas. Las técnicas utilizadas para la identificación de las piezas son: en primera instancia la binarización de la imagen, la técnica de sobel para la identificación de contornos, la transformada de hough para detectar el ángulo de inclinación de la pieza y por último la correlación para comparar las imágenes patrón con la imagen escaneada. Se ha utilizado como criterio de clasificación el factor de correlación r . Obteniendo un nivel de eficiencia del 90% y una repetibilidad del 95%.

Palabras Clave: Factor de correlación, perfilometría laser, sistema de visión.

I. Introducción

En el área de automatización de procesos son comunes las tareas de selección, identificación y clasificación de objetos utilizando diversos métodos que van desde la descripción hasta la caracterización de dichos objetos.

Una alternativa de automatización es el uso de sistemas de visión, los cuales permiten, de forma rápida, segura y confiable, la supervisión de procesos industriales que implican tareas de identificación y/o clasificación de objetos [1].

Sin embargo, los sistemas de visión tradicionales son muy sensibles a los cambios de iluminación y esto hace que su etapa de calibración sea muy tediosa y se lleve a cabo frecuentemente. Una alternativa es el uso de un sistema de perfilometría laser calibrado para operar en la frecuencia del laser utilizado, esta característica evita largas y frecuentes etapas de calibración [2,3].

Una de las grandes ventajas de utilizar un sistema de visión de tipo escaneo laser y no un sistema tradicional para la identificación de objetos es la

independencia del factor iluminación ya que un sistema tradicional es muy sensible a los cambios de iluminación pudiendo llegar a ser inservible si estos se presentan mientras que un sistema laser no se ve afectado por dichos cambios.

El objetivo principal del presente trabajo es la identificación y clasificación de piezas metálicas, las cuales entre si son muy similares tanto en construcción como en forma pero con características particulares diferentes. Para la selección se utiliza una cámara monocromática, un láser, una banda transportadora y una computadora que contiene el software encargado del reconocimiento y selección de las piezas.

El algoritmo antes mencionado ha sido implementado para llevar a cabo las tareas de detección, captura, corrección del ángulo de error utilizando la transformada Hough y clasificación de las piezas de un conjunto de ocho en total.

El ambiente de trabajo en el que se desarrolló el sistema de visión propuesto es Matlab, sobre el cual se diseñaron los procesos básicos para el

tratamiento y análisis de las imágenes capturadas utilizando perfilometría láser [4,5]. La razón por la que se optó por utilizar la técnica láser antes mencionada se debe a que éste tipo de sistemas no requieren de sistemas de iluminación complejos comúnmente utilizados en los sistemas de visión convencionales [1-3].

II. Desarrollo

El sistema de visión desarrollado comprende aspectos de hardware y software; sin embargo, se presta especial atención a la descripción del software diseñado para la captura, análisis y clasificación de la perfilometría laser de 8 piezas metálicas utilizadas en los mecanismos de cinturones de seguridad de automóviles (ver Fig. 1).



Fig. 1. Piezas metálicas a identificar.

II.1 Sistema Óptico

Para la proyección de una línea láser sobre las piezas metálicas, cuyas medidas son 13cm de alto por 6cm de ancho, se ha utilizado un láser de alta intensidad modelo LAS-SNF-XXX-660-35 con potencia de salida de 35mW a una longitud de onda de 660nm, al láser se le acopla una lente positiva, comúnmente denominada *single line*, con la finalidad de modificar el puntero laser de forma circular a una forma lineal que nos permita el escaneo de los objetos, ubicándose este sobre una plataforma fija montada sobre la banda transportadora, la cual nos proporciona la velocidad de escaneo de los objetos.

Para la captura de las imágenes se utilizó una cámara monocromática tipo MV-D640 CMOS VGA que proporciona imágenes monocromáticas

de 256 niveles de gris a una resolución de 640x480 pixeles.

La distancia a la cual se ubicaron el láser y la cámara del objeto fue de 50 cm obteniendo un ángulo visual de 14.81° asegurando con esto que las imágenes presentaran el 100% de las piezas metálicas. La Figura 2 muestra el sistema óptico utilizado para la captura de las imágenes de piezas metálicas ubicadas sobre una banda transportadora.

La captura de las imágenes se realiza utilizando una tarjeta digitalizadora conectada a una computadora personal en la cual se ha instalado previamente el software que permite configurar dicha cámara y realiza el procesamiento de las imágenes.

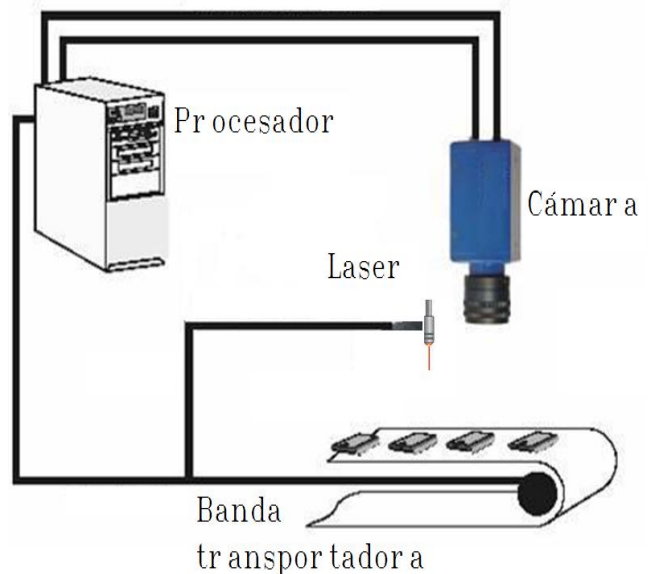


Fig. 2. Sistema óptico.

II.2 Captura y reconstrucción de la imagen

Como primer paso se establece la comunicación entre la cámara y el algoritmo desarrollado en Matlab mediante el archivo `hardware.m` que es un archivo pregenerado desde *imaqtool*, una herramienta de trabajo de Matlab utilizada para la configuración y control de tarjetas digitalizadoras de video. Posteriormente, se inicia con la captura de video utilizando el comando `getsnapshot` para tomar una imagen durante el video y poder así convertirla a una imagen binarizada mediante la

función *bwlabel* obteniendo umbrales de operación de 0.21 a 0.35.

El sistema contiene procedimientos de software para identificar el momento en que alguna de las piezas esta cruzado el área de visión de la cámara hasta cuando esta ha cruzado por completo dicha área. Esto es efectuado mediante un ciclo de espera que depende de los cambios en la perfilometría de la línea láser, donde el sistema compara la imagen actual con la imagen capturada con anterioridad haciendo uso de la operación *xor* entre ambas muestras y generando la media de la imagen obtenida. Cabe señalar que la banda transportadora puede contener manchas o uniones que modifiquen la perfilometría de la línea laser por lo que fue necesario considerar un cambio mínimo a fin de que estos no sean interpretados como objeto dando inicio a la captura del mismo.

Una vez que el sistema detecta cambios significativos en la perfilometría de la línea láser, es decir que existe un objeto presente en el área de visión de la cámara, se inicia un nuevo ciclo de captura de imágenes muy similar al anterior.

Si la muestra ha variado es concatenada con la anterior hasta poder al final de éste ciclo obtener una representación 2-D de la pieza metálica. En la Figura 3 se muestra un diagrama de flujo de la forma en cómo funciona el procedimiento de captura de la imágenes hasta obtener la representación 2-D de las piezas.

II.3 Corrección del ángulo de error

La imagen reconstruida y capturada de la pieza pudo haberse capturado con un ángulo de inclinación respecto de su eje principal, esto como consecuencia de la ubicación física de la pieza. Este ángulo puede ocasionar deficiencias en la clasificación de las piezas por lo que es necesario identificarlo y corregirlo. La técnica utilizada para conocer el ángulo de inclinación de las piezas fue la transformada Hough, la cual implica obtener el espacio rho-theta de los pixeles de la imagen utilizando la ecuación 1, una vez que se ha identificado el ángulo de inclinación de la pieza se utilizó un algoritmo de

rotación de imágenes conocido como interpolación del vecino más cercano ecuación 2.

$$\varphi = x\cos(\rho) + y\sen(\rho) \quad (1)$$

$$R(x, y) = Ax(\cos\alpha - \sen\alpha) + Ay(\cos\alpha + \sen\alpha) \quad (2)$$

El sistema puede corregir ángulos de inclinación de las piezas en el intervalo de $[-45,45^\circ]$. La Figura 4 muestra un ejemplo de una imagen reconstruida que presenta un ángulo de inclinación de aproximadamente 20° .

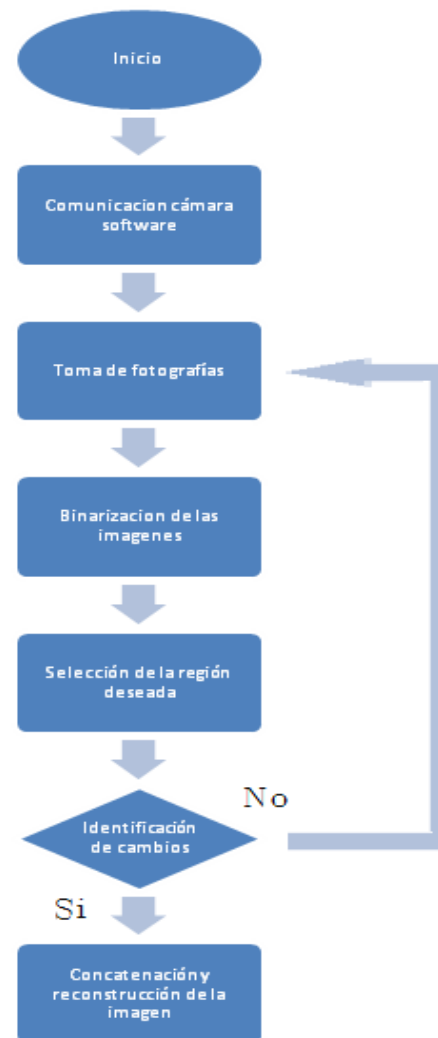


Fig. 3. Pasos para la captura y reconstrucción de la imagen.

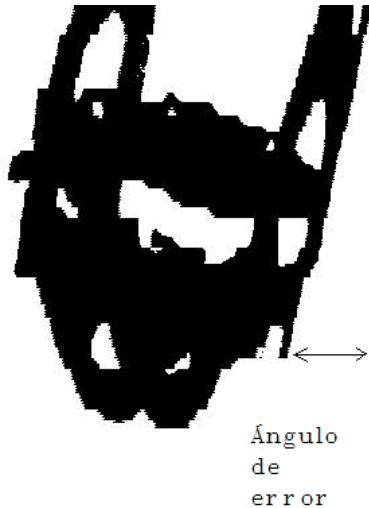


Fig. 4. Figura con ángulo de error.

II.4 Procesamiento de imágenes

Previo a la tarea de clasificación se extraen los bordes, cabe mencionar que existen diversas técnicas de extracción de bordes como pueden ser: Prewitt, Gradiente o Sobel por mencionar algunas de ellas, se decidió utilizar la técnica de Sobel que aplica un kernel o máscara considerando una dirección vertical en los píxeles mediante la ecuación 3, esta orientación es debido a la posición en la cual es capturada la imagen para su análisis, pudiendo cambiar dicha orientación si es que se decidiera cambiar la posición de captura de la pieza. Sin embargo, en la imagen resultante del algoritmo de Sobel se presentan zonas de píxeles que reducen la eficiencia del sistema y que se presentan debido a la reflexión del láser sobre las piezas metálicas, se decidió eliminar áreas de píxeles menores a 200 de un total de 307200 píxeles.

$$G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \text{Imagen} \quad (3)$$

II.5 Extracción

Una vez corregido el ángulo de inclinación de la figura, se extraen únicamente de la imagen los píxeles que contienen información de la pieza, con el fin de eliminar los píxeles de la periferia de la imagen que no contienen datos importantes para la identificación de la misma.

Para poder realizar el proceso de identificación de la pieza es necesario modificar su tamaño (120x120) para que este sea igual al de las imágenes de referencia de la base de datos previamente obtenidas y así poder correlacionarlas entre sí.

II.6 Clasificación

En el proceso de clasificación se utiliza de una base de datos que son las imágenes de referencia que se han conformado para fines de comparación con las imágenes de prueba. Las imágenes de prueba y las de referencia son utilizadas para obtener un factor de correlación utilizando la ecuación 4. La identificación de las piezas depende del valor obtenido en el factor de correlación en el intervalo [0,1], un valor cercano a uno entre dos imágenes indica que hay una alta posibilidad de que ambas imágenes sean similares; caso contrario, un valor de factor de correlación cercano a cero indica que se trata de dos imágenes con características espaciales muy diferentes.

$$r = \frac{N \sum xy - \sum x \sum y}{\sqrt{[N \sum x^2 - (\sum x)^2][N \sum y^2 - (\sum y)^2]}} \quad (4)$$

Cabe señalar que la carga de las imágenes de referencia para la base de datos debe realizarse de forma independiente al programa de identificación de la pieza, para evitar que esto se realice repetidamente cada vez que se ejecute el programa de reconocimiento del objeto, evitando con esto retardos y mejorando el tiempo de respuesta del algoritmo (ver Fig. 5).

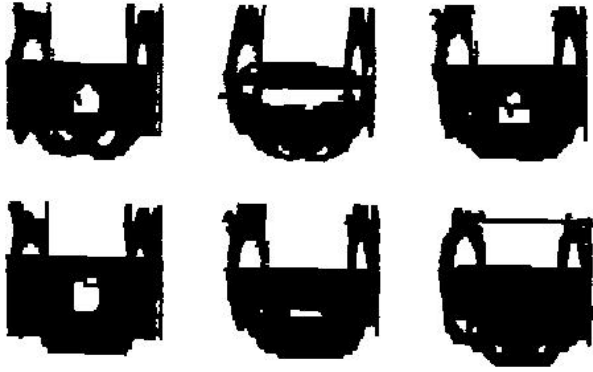


Fig. 5. Algunas piezas patrón.

El diagrama a bloques del funcionamiento del sistema se muestra en la Figura 6.

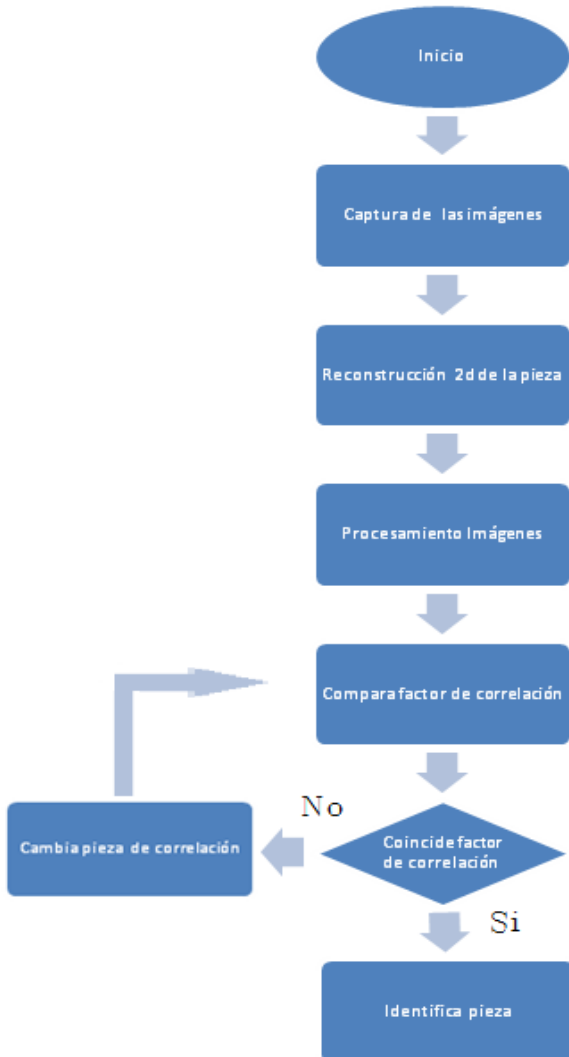


Fig. 6. Diagrama de procesamiento y reconocimiento de la imagen.

III. Resultados

La resolución de las imágenes capturadas es de 640x480; sin embargo, las imágenes reconstruidas utilizando perfilometría láser resultaron con una resolución máxima de 240x480, lo que permitió obtener menores tiempo de procesamiento de las imágenes.

Refiriéndonos a dos piezas la 1 y 5 (ver Fig. 7) y comparándolas con las figuras obtenidas (ver Fig. 8), podemos observar diferencias entre la imagen real y la obtenida debido que al tratarse de una superficie capaz de reflejar el haz de luz, la figura resultante no será del todo exacta, aun mas cuando se trata una imagen con un ángulo crítico la luz es parcialmente reflejada perdiendo detalles en la forma de la imagen (ver Fig. 9).

Sin embargo al tratarse de un sistema de clasificación podemos decir que la forma de la imagen no se ha perdido del todo y es aun clasificable si estos cambios se añaden a la base de datos, es decir, debido a que nuestro sistema de reconocimiento, está basado en una base de datos conformada por las piezas escaneadas (ver Fig. 8) y no por las imágenes sin escanear (ver Fig. 7).

Podemos tomar como referencia o patrón no solo una sino varias imágenes de la pieza en la misma posición, con la finalidad de captar todas las posibles diferencias mínimas entre patrón y patrón de la misma pieza y así asegurar su reconocimiento, esto claro bajo la condición de que al aumentar el número de patrones disminuye el margen de error mejorando la eficiencia pero aumentando el tiempo de procesamiento.



Fig. 7. Pieza 1 y 5.



Fig. 8. Pieza 1 y 5 escaneadas.

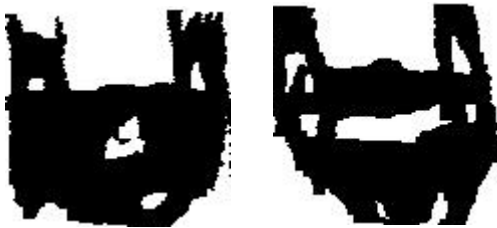


Fig. 9. Pieza 1 y 5 escaneadas con ángulo crítico.

Los tiempos de captura y clasificación de las piezas se muestran en la Tabla 1, es posible notar que dichos tiempos son considerablemente pequeños. Sin embargo, lo más relevante de la técnica propuesta, es que no es significativamente dependiente del sistema de iluminación y el sistema tiene un alto nivel de eficiencia aun cuando la pieza se coloque sobre la banda con un ángulo de inclinación en el intervalo $[-45,45]$.

Pieza	Tiempo (segundos)
01	5.226651
05	5.683736
07	5.817868
08	5.979817
09	5.362135
10	5.367965
12	5.696779
13	5.352354

Tabla 1. Tiempos de cómputo para escaneo de piezas.

Los tiempos promedio de reconstrucción y clasificación para cada una de las piezas evaluadas son mostrados en la Tabla 1. Es posible notar que el tiempo para cada pieza fue muy similar y que el tiempo promedio de tratamiento en cada pieza fue de 5.589seg.

Con una base de datos de hasta tres imágenes a correlacionar por pieza se logró obtener una precisión del 70% y con cinco 90%. Se puede tener en cuenta que el tiempo de cómputo efectuado en la carga de la base de datos es ignorado para el cálculo de estos tiempos y así mismo el tiempo de cómputo en la función de correlación es muy pequeño .000033 segundos por correlación entre piezas por lo que se puede añadir una base de datos mayor sin comprometer los tiempos.

Estadísticamente, tenemos que de un conjunto de 8 piezas a clasificar y teniendo un total de una base de datos de 40 patrones a correlacionar, y de los cuales cinco corresponden a cada pieza tenemos que: la probabilidad de que una imagen de referencia corresponda a la pieza que se está evaluando es de 0.20; es decir, por cada correlación entre la imagen de referencia y la imagen a evaluar, existe un veinte por ciento de posibilidades de reconocer la pieza.

En la etapa de prueba del sistema encontramos que de un total de 80 pruebas de reconocimiento realizadas, diez se le realizaron a una sola pieza, y se logró en promedio nueve reconocimientos acertados y solo una respuesta equivocada esto indica un porcentaje de error del 10% con un 90% de eficiencia en el reconocimiento de las piezas. Encontrando también que al comparar con pruebas posteriores y repitiendo todo el proceso de reconocimiento para las ocho piezas el porcentaje de repetibilidad es del 95% ya que solo se presentaron 4 cambios significativos del total de las pruebas realizadas.

IV. Discusión

Hoy en día el software mas empleado para tratamiento de imágenes es Matlab que permite tanto configuración de hardware (cámara) como de software además de poseer una gran documentación sobre el mismo y una interfaz amigable de programación de alto nivel [1] presenta una gran variedad de ejemplificaciones así como una amplia referencia a comandos y

funciones para el tratamiento de imágenes lo cual nos permite adentrarnos más al entorno de trabajo del presente artículo.

En [2] se presenta un sistema de reconocimiento de caracteres y análisis mediante Matlab en el que se reconocen, segmentan y extraen para poder reconocerlos mediante una base de datos similar a la empleada por nuestro sistema, aportando así la idea central de tratar el sistema de una manera muy similar, aterrizando esta idea de acuerdo a nuestras necesidades.

En [3] se presenta una manera de tratar imágenes binarias para reconocimiento dactilar, así como el tratado de filtros para corrección y detección de discontinuidades para su reconocimiento y asimilación en bases de datos, empleando procesos enteramente binarios permitiéndonos una serie de ideas que anteceden a la implementación de hardware y software.

Asimismo [4, 5] presentan resultados que permiten suponer los alcances posibles de obtener, cuando se aplica la técnica de perfilometría laser a la representación de objetos en 2D y 3D. Además, [3, 4] sugieren el uso de técnicas de inferencia en base a redes neuronales para obtener representaciones de objetos de buena calidad; sin embargo, estas técnicas requieren un significativo tiempo de procesamiento lo que implicaría un excesivo tiempo de ejecución de nuestro sistema (Ver Tabla1), consecuentemente, se optó por recurrir a la extracción de características geométricas de las piezas y así evitar tiempos de procesamiento mucho mayores a los obtenidos.

Finalmente, en [6] se muestra una técnica de visión tridimensional basada en metrología láser y algoritmos de inteligencia artificial, en donde la visión se efectúa mediante proyección de una línea láser y procesamiento de imágenes.

V. Conclusiones

Al utilizar un sistema automático de visión artificial para la identificación y clasificación de

piezas mecánicas se mejora considerablemente el proceso de selección ya que se reducen costos y tiempos además de que el sistema de clasificación propuesto no se ve afectado en su eficiencia ante cambios de iluminación externos debido a que se ha calibrado el sistema para extraer información de las piezas únicamente en la zonas donde la línea láser es proyectada. Además, el criterio de clasificación propuesto en base al factor de correlación evaluado entre las imágenes de prueba y las imágenes consideradas en la base de datos se presenta como una buena opción para los sistemas de visión con aplicaciones industriales en línea.

En el desarrollo del sistema se han identificado dos factores que afectan su nivel de eficiencia y repetibilidad, estos son: la velocidad de giro del motor que mueve la banda transportadora y el tamaño de la base de datos que está en función de los valores del ángulo de inclinación en que pueden llegar las piezas al campo visual de la cámara y de las divisiones consideradas en el plano rho-theta sugerido por el método de la transformada Hough.

Por último al tratar con piezas metálicas que poseen un coeficiente de reflexión muy cercano a la unidad, se corre el riesgo de saturación de los sensores de la cámara si no se considera una adecuada potencia del láser y un correcto ángulo de captura de la cámara. Además, al tratar con piezas de diversas características que implicaron diversos ángulos de inclinación con respecto al láser y a la cámara de video, fue necesario incluir cinco imágenes por pieza en la base de datos para una clasificación en condiciones de reflexión ante ángulos críticos, en las que se incluyen los efectos de dichos ángulos para así lograr obtener una más alta eficiencia en el sistema, así mismo ampliar la base de datos mejora la eficiencia del sistema a costa de tiempo de computo que como antes se menciona es aproximadamente .000033 segundos por pieza añadida a la base.

VI. Referencias

- [1] R. C. González, R. E. Woods, “Digital Image Processing”. Second Edition. 2002.

Prentice Hall. pp. 146-169. Printed in the United States of America; Upper Saddle River, New Jersey 07458.

- [2] A. M. Rodríguez, V. Rodríguez. "Evaluation of the light line displacement location for object shape detection". Revista: Journal of Modern Optics, 2003. Vol. 50, No. 1, Páginas 137-154.
- [3] J. Apolinar Muñoz Rodríguez, R. Rodríguez Vera. "Representation of Shape Based on Laser Stripe Projection and Modelling with Bezier Approximation" Revista: Journal Lasers in Engineering, 2007. Vol. 17. No. 1, Páginas 21-43.
- [4] D. Barragán, P. Vallejo, "Reconocimiento de caracteres ópticos (OCR) usando Matlab". www.matpic.com Fecha de consulta 3 de mayo 2011. 5 Páginas.
- [5] D. Barragán, P. Vallejo, "Reconocimiento de Huellas Digitales Con Matlab". www.matpic.com Mayo 2009. 40 Páginas.
- [6] J. Apolinar Muñoz Rodríguez, Edgar Fernando Velázquez Pedroza, Diego Alonso Castañeda Aguilar. "Visión tridimensional basada en metrología láser y algoritmos de inteligencia artificial". León Guanajuato, Revista académica semestral: Cuadernos de Educación y Desarrollo, 2009. Vol. 1. No. 2. Páginas 1-7.

VII. Autores

Ing. Gonzalo Elías Blanco Silva nació en Querétaro, Qro. El 25 de Noviembre de 1986. Obtuvo el grado de Ingeniero en Mecatronica en Abril de 2010 en el Instituto Tecnológico de Celaya.

Ing. Alejandro González Álvarez nació en Puruándiro, Mich. El 29 de Julio de 1984. Obtuvo el grado de Ingeniero en Electrónica en Junio de 2010 en el Instituto Tecnológico de Celaya.

Dr. José Alfredo Padilla Medina nació en Iguala, Gro. El 21 de Abril de 1969. Obtuvo el

grado de Ingeniero en Electrónica en junio de 1992 en el Instituto Tecnológico de Celaya. En 2003 obtuvo el grado de Doctor en Ciencias (Óptica) por parte del Centro de Investigaciones en Óptica. Ha realizado cerca de 20 publicaciones en revistas indizadas y en congresos internacionales. A partir del 2002 es profesor-investigador en el Instituto Tecnológico de Celaya donde desarrolla investigaciones en las áreas de controladores difusos, sistemas de visión y teoría ROC.

Dr. Juan Prado Olivares nació en Cortázar, Gto. México el 25 de diciembre de 1969. Recibió el grado de ingeniero en el Instituto Tecnológico de Celaya en 1998. En 2001 recibió el grado de maestro en Ingeniería Eléctrica (Instrumentación y control digital) por la Universidad de Guanajuato y el grado de Doctor en Instrumentación y Microelectrónica en 2006 en la Université Henry Poincaré. En 2007 fue contratado como profesor investigador por el Instituto Tecnológico de Celaya dentro del cuerpo académico de Adquisición y procesamiento de Señales donde actualmente labora. Su principal interés es la espectroscopia de impedancia eléctrica y la Instrumentación Electrónica.

Dr. Agustín Ramírez Agundis nació en San Miguel de Allende Gto. Obtuvo el grado de Maestro en Ingeniería en el año de 1977 por la Universidad de Guanajuato. En 2008 obtuvo el grado de Doctor en Diseño de Sistemas Digitales por la Universidad Politécnica de Valencia. Desde 1986 ha sido profesor del departamento de Ingeniería Electrónica del Instituto Tecnológico de Celaya. Ha publicado una decena de artículos en revistas y congresos internacionales.

Reconocimiento del Lenguaje de Señas Mexicano para su Interpretación y uso en Sistemas Robóticos

***Felix Emilio Luis-Pérez, ** Felipe de Jesús Trujillo-Romero**

Universidad Tecnológica de la Mixteca,
División de Estudios de Posgrado
Carretera a Aclatlíma Km. 2.5 Hujuapan de León Oaxaca,
Tel.: 9535320214,
*eluis@mixteco.utm.mx
**ftrujillo@mixteco.utm.mx

Resumen

En este trabajo se presentan los resultados obtenidos al reconocer el alfabeto del lenguaje de señas mexicano como un elemento de control para un sistema robótico. Para el reconocimiento de las señas se utiliza una segmentación de las imágenes mediante el uso de contornos activos (snakes). Una vez segmentado se procede a obtener la firma de la seña correspondiente y se entrena una red neuronal para su reconocimiento posterior. A cada símbolo del lenguaje de señas se asigna una tarea que el sistema robótico debe de ejecutar, en esta primer etapa se definen solo seis tareas diferentes. El reconocimiento que presenta el sistema es del 89.5%. Se realizaron pruebas para validar el funcionamiento del sistema de reconocimiento con el simulador Roboworks con el modelo de la plataforma móvil PowerBot y PowerCube el cual es un brazo de 6 grados de libertad.

Palabras Clave: Contornos activos, Firma del objeto, Lenguaje de señas mexicano, Redes neuronales, Sistema robótico.

I. Introducción

Cada vez son más presentes los robots en nuestra vida cotidiana. Por ello es necesario que estos evolucionen para que el ser humano pueda tener una mejor interacción con ellos.

Una rama de la robótica llamada robótica social [1] se enfoca a desarrollar algoritmos, metodologías y robots que cumplan ciertos requisitos para que sean aceptados por el ser humano y desde luego se tenga la interacción hombre-máquina ideal.

Ejemplos de ello es la gran cantidad de robots como el Robosapien[6] que se encuentran en el mercado y que ofrecen un primer acercamiento, principalmente a los niños, con estos sistemas robóticos. Este tipo de robots son “programables” y el usuario puede hacer que desarrollen

diferentes rutinas. Estas rutinas están orientadas más hacia el campo lúdico de la robótica, pero ofrecen un campo de experimentación con una plataforma más o menos inteligente y a bajo costo [7, 8].

Por otra parte existen desarrollos de universidades y/o laboratorios que trabajan en la realización del robot ideal. Podemos mencionar al robot Kissmet [1, 2] desarrollado por el MIT. Este robot tiene la particularidad de poder realizar expresiones faciales de acuerdo a lo que está diciendo o incluso de acuerdo a lo que escucha. Dicho de otra forma es un robot emocional que reacciona a lo que sucede en su entorno.

Otro ejemplo es Justin, quien es un robot que posee dos brazos articulados y es capaz de realizar algunas tareas domésticas, como preparar

café. Este robot ha sido desarrollado en el instituto de robótica y mecatrónica del centro aeroespacial alemán [3].

Si bien los ejemplos anteriores son robots que tratan de parecerse al ser humano no todos tienen un aspecto totalmente humanoide. Citemos el caso de Rackham[4, 5], robot desarrollado en por el LAAS en Francia y que sirvió de guía de museo “la ciudad del espacio” en la ciudad de Toulouse. Este robot era capaz de recibir órdenes mediante una pantalla táctil mediante la cual el visitante le indicaba el lugar del museo que quería visitar y el robot lo guiaba manteniendo siempre contacto visual con el mediante cámaras.

Otro desarrollo europeo es Cogniron[10, 11], en este proyecto se pretendía realizar un robot de compañía que fuera, entre otras cosas, curioso. Es decir, el sistema debía entrar en una habitación y comenzar a explorarla para poder así tener un conocimiento de la misma. Esto le permitiría saber en qué lugar se encontraban los objetos que pudiera pedirle un usuario humano.

La lista de robots que intentan sociabilizar con el hombre es muy grande por lo que finalmente comentaremos un desarrollo realizado en la UNAM y el cual es el robot Golem [9]. Este robot entre otras cosas puede reconocer comandos vocales con cierta precisión lo cual le permite interactuar con el ser humano de una forma más natural.

Si observamos a detalle los ejemplos de robots mencionados se verá que los avances han sido significativos pero aun así el problema sigue siendo la manera en la que los robots se comunican con el ser humano.

Lo anterior se debe principalmente a la forma en la cual se le van a indicar al robot una acción que queremos que realice. Por lo tanto se necesita de algún mecanismo que nos permita darle los comandos necesarios o las indicaciones necesarias para que el robot pueda llevar a cabo de manera aceptable la tarea encomendada.

Por lo cual en este trabajo se plantea el control de un sistema robótico mediante la utilización de los signos del alfabeto del lenguaje de señas mexicano (LSM). La lengua de señas mexicana es la lengua natural de las comunidades de sordos. Una de las características de las lenguas de señas en general es su forma de transmisión visual, además de que estas lenguas han surgido en las comunidades de sordos de manera espontánea, son parte de su cultura y reflejo de su visión del mundo y se encuentran en un proceso productivo constante, como sucede con cualquier otra lengua natural.

En la figura 1 se muestra el alfabeto de lenguaje de señas mexicano. El LSM es, como toda lengua natural, algo complejo y que está en constante evolución. Razón por la cual nuestro análisis se basará solo en el alfabeto de señas del LSM y su utilización como sistema de control de un sistema robótico. El sistema robótico que va a ser controlado a partir de señas está formado por una plataforma móvil equipada con un manipulador de 6 grados de libertad.

El interés en la utilización del LSM es, principalmente, que son casi nulos los trabajos que se han desarrollado usando esta codificación. Ya que la mayoría utiliza el Lenguaje de señas Americano, en [20] se puede ver un ejemplo de ello. Cabe mencionar que cada país posee su propia codificación para su alfabeto [20, 21, 22, 23]. Aún en países que se habla la misma lengua existen variantes en los símbolos que representan un mismo sonido.

II. Desarrollo

A fin de poder llevar a cabo de manera satisfactoria el sistema propuesto se realizaron diferentes etapas. Estas etapas se enumeran a continuación:

1. Obtención de imágenes.
2. Segmentación
3. Obtención del descriptor de forma
4. Entrenamiento de la red neuronal

5. Reconocimiento de formas
6. Validación del sistema en un entorno de simulación.

En los párrafos siguientes se explicará de forma precisa cada una de las 4 primeras etapas, dejando para la sección de resultados la parte de reconocimiento y validación.

II.1. Obtención de imágenes

Se utilizaron 23 símbolos (ver Fig. 1) de los 27 que componen el alfabeto de señas mexicano. Eligiéndose aquellos cuya forma era diferente unos de otros. Además que en esta etapa no se consideran secuencias de imágenes, descartándose de esta forma los signos que tienen movimiento.

En la Fig. 1 se puede observar que las imágenes se tomaron usando un fondo uniforme y teniendo en cuenta que la persona que hacía los signos usaba una prenda que solo dejaba descubierta la mano.

Estas condiciones eran para evitar la interferencia de objetos en la muñeca (reloj, pulseras, etc.) y asegurar que la región de interés que se obtiene de la segmentación pertenece solamente a la mano.



Fig. 1. Conjunto de signos del alfabeto del LSM utilizados.

Se tomaron tres diferentes conjuntos de imágenes cada uno de ellos con diferencias en la iluminación del entorno.

Un ejemplo de las imágenes usadas para el desarrollo del presente trabajo se muestra en la Fig. 2. En esta figura se pueden observar tres diferentes imágenes, una de cada conjunto, para uno de los signos utilizado.

Cada una de las tres imágenes tiene una variación en la intensidad luminosa presente al momento de tomar las imágenes. Esta variación en la iluminación nos permite incorporar, en cierta manera, un espectro más amplio y poder discriminar este parámetro de la iluminación.

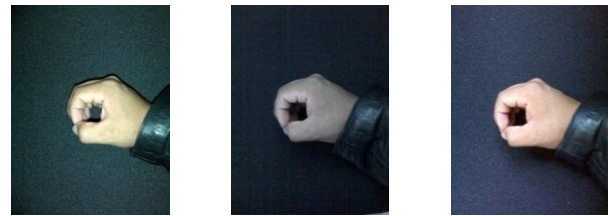


Fig. 2. Imágenes tomadas con diferentes condiciones de iluminación.

Ya una vez capturado el conjunto de imágenes se procede a realizar el procesamiento necesario para poder entrenar la red neuronal. Este procesamiento comienza con la segmentación de las imágenes, proceso que se explicara en la subsección siguiente.

II.2. Segmentación

El primer paso para la realización de ese sistema de reconocimiento de señas fue la segmentación de imágenes mediante contornos activos [13]. Estos contornos modelan las fronteras entre un objeto, el fondo y el resto de objetos de la imagen.

Además, permiten extraer los contornos de los objetos de interés basándonos en modelos que utilizan información a priori de la forma de los objetos. Estas técnicas son mucho más robustas frente a la presencia de ruido y otros elementos espúreos y permiten segmentar imágenes mucho

más complejas que las imágenes para las que eran aplicables los métodos de segmentación de bajo nivel, como en el caso de las imágenes médicas.

Los contornos activos se pueden clasificar en snakes (serpientes), patrones deformables y contornos dinámicos. En este trabajo se utilizaron snakes para llevar a cabo la segmentación.

Un snake se puede definir como una curva spline minimizadora de energía, guiada por fuerzas restrictivas externas e influenciada por fuerzas de la imagen, que tiende a localizarse en características de ésta como líneas y bordes. Es, por tanto, un contorno activo que evoluciona de forma dinámica hacia los contornos relevantes de la imagen.

El snake también posee una serie de fuerzas internas que sirven para imponer restricciones de suavidad, es decir, para regularizar la solución.

Las fuerzas de la imagen empujan al snake hacia características de la imagen como líneas, bordes y contornos subjetivos, mientras que las fuerzas restrictivas externas añaden información de alto nivel para hacer que el snake se vaya hacia el mínimo local deseado.

El snake se representa como una curva paramétrica mediante $r(s) = (x(s), y(s))$, con $s \in (0, 1)$. Su funcional de energía se puede poner como:

$$E_{snake}^*(r) = \int_0^1 E_{int}(r(s))ds + \int_0^1 E_{img}(r(s))ds + \int_0^1 E_{res}(r(s))ds \quad (1)$$

Donde E_{int} representa la energía interna del snake, E_{img} las fuerzas de la imagen y E_{res} las fuerzas restrictivas externas.

Existen variantes del algoritmo original como el presentado en [12, 14]. En los cuales se realiza la snake sin bordes [14] y en [12] se implementa usando un vector de campo.

El algoritmo empleado en este trabajo también es una variante la cual realiza la segmentación por regiones. Este algoritmo está basado en el presentado en [25].

Con esta implementación se procedió a segmentar las imágenes que iban a ser usadas para entrenar la red neuronal. Se puede observar en la Fig. 3 diferentes transiciones del snake para la imagen del símbolo de LSM que representa a la letra “w”.

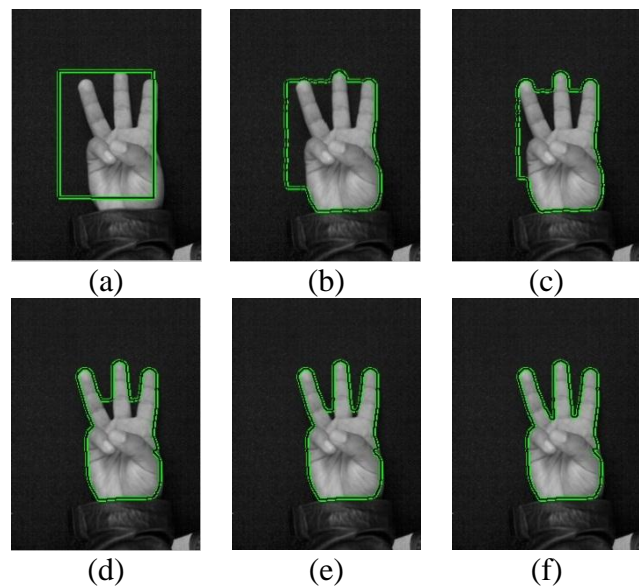


Fig. 3. Diferentes transiciones del snake (a) snake inicial y (f) contorno final.

II.3. Descriptor de forma

Una vez obtenida la segmentación de las imágenes se procedió a obtener los descriptores de forma. Existen gran cantidad de descriptores que utilizan los contornos como información para generar el descriptor correspondiente. Dos descriptores de este tipo son *shape context* [18] el cual explota los contornos obtenidos a partir de un operador de gradiente y *shape signature* [15] que utiliza solo el contorno exterior de los objetos a clasificar.

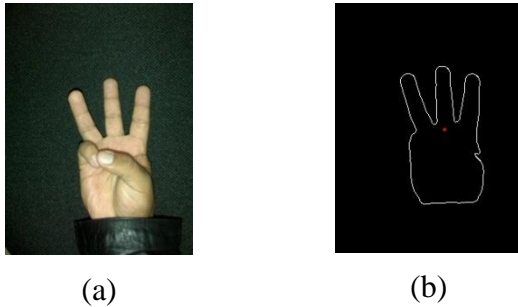


Fig. 4. (a) imagen original, (b) bordes de (a) obtenido a partir de la snake.

Dado que la snake nos entrega como resultado el borde del objeto se utilizó *shape signature* como descriptor de forma. En este caso el contorno de los objetos era la mano con la forma del signo correspondiente a un símbolo del alfabeto LSM (véase Fig. 4).

De manera general la obtención de la firma del objeto se realiza mediante el cálculo de las distancias desde el centro de gravedad del objeto hacia cada uno de los puntos que forman el contorno. Esto nos genera un histograma de distancias como el de la Fig. 5 el cual es el histograma para el contorno del símbolo mostrado en la Fig. 4a.

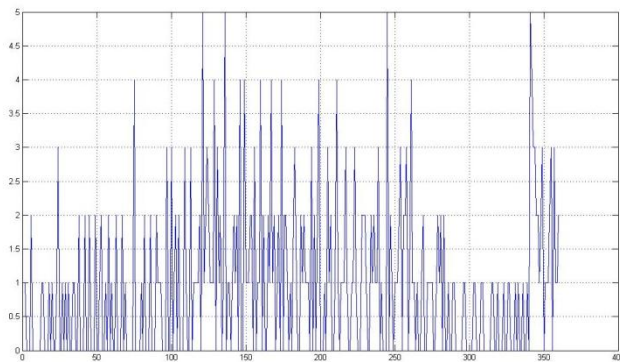


Fig. 5. Histograma de distancias para el símbolo que representa la letra "w".

Este histograma está formado por 360 valores diferentes y forma el vector de entradas para realizar el entrenamiento de la red neuronal. La cual se describirá a continuación.

II.4. Entrenamiento de la red neuronal

Con la obtención de los histogramas que forma la descripción de los diferentes símbolos del

alfabeto LSM se tiene los datos de entrada para poder entrenar la red neuronal.

Se utilizó una red neuronal compuesta por tres capas con la configuración siguiente:

- Una capa de entrada de 360 neuronas una capa intermedia de 23 neuronas y con una neurona en la capa final. La estructura de la red neuronal se puede observar en la Fig. 6.

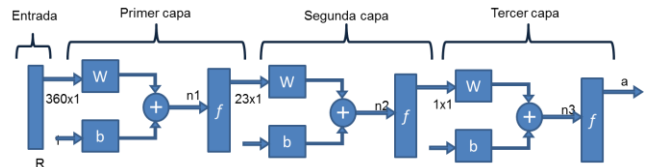


Fig. 6. Estructura de la red neuronal empleada.

Al escoger esta estructura para la red neuronal se obtiene una mayor rapidez para el entrenamiento, sobre todo por el tamaño del vector de entrada. En [21] se utiliza un vector de entrada de 18000 valores y por consecuencia definen una capa de entrada con la misma cantidad de neuronas. Esto hace que el procesamiento sea más lento.

En este caso, los vectores de entrada son formados por los histogramas obtenidos del descriptor de forma *shape signature*. A cada vector de entrada se le asocia una etiqueta de salida que va desde el número 1 hasta el número 23. Por lo tanto se tiene una salida de 23 valores diferentes uno para cada símbolo del LSM utilizado.

Para el entrenamiento de esta red se usó el algoritmo de *backpropagation*. La ventaja que presenta *backpropagation* es rapidez de convergencia y robustez respecto a otro tipo de entrenamiento.

Algunos de los parámetros utilizados para este entrenamiento son:

- Índice de aprendizaje de 0.01,
- Un máximo de 9000 iteraciones, y
- Un error mínimo de $1e-10$.

El tiempo que le toma a la red neuronal aprender los 69 patrones de referencia y asociarlos con los 23 patrones objetivo es de 187 segundos.

III. Resultados

En esta sección se mostrarán los resultados obtenidos al reconocer e interpretar el conjunto de símbolos que forman el LSM.

Para validar los resultados se usaron imágenes en las cuales las condiciones de iluminación se simularon usando una función de potencia del tipo:

$$f(x) = ce^x \quad 0.91 \leq x \leq 1.09 \quad (2)$$

Donde c es una constante y x es el parámetro que hay que variar para obtener diferentes intensidades en la imagen que se procesa. Al evaluar (2) se tiene que si x es menor que 1 la imagen es más oscura y si x es mayor que 1 se hace más brillante (ver Fig. 7).

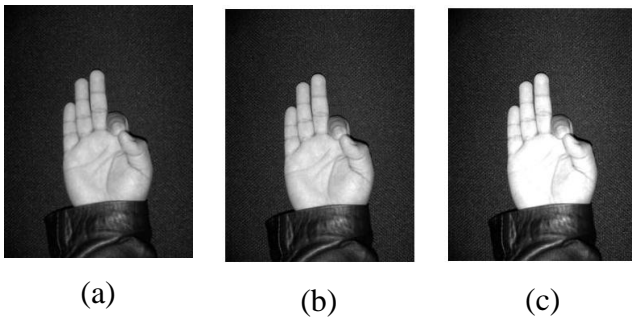


Fig. 7. Ejemplos para diferentes valores de potencia.

En la Fig.7 la imagen (b) se le aplicó un valor del parámetro x de 0.9 y en la imagen (c) de 1.1, la imagen (a) es la imagen original.

La función descrita en (2) se aplica de forma aleatoria a las imágenes de prueba de forma independiente a cada una de ellas. Es decir no se tiene el mismo valor del parámetro x para cada serie de experimentos ni para cada símbolo en la serie.

El objetivo de aplicar (2) a cada una de las imágenes de prueba es tener un conjunto diferente a las imágenes usadas para el entrenamiento.

Al evaluar para uno de los símbolos se obtuvo el siguiente error respecto al símbolo aprendido. En la Fig. 8 se muestra el histograma para el símbolo

que representa la letra **w** modificada mediante la función mostrada en (2).

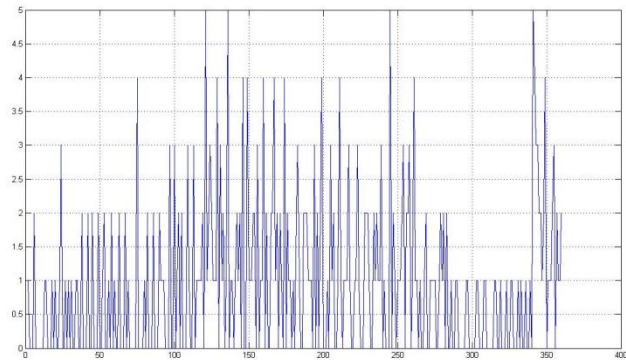


Fig. 8. Histograma de distancias para el símbolo que representa la letra “w” en la fase de prueba.

Al obtener la diferencia entre los histogramas de las Figs. 5 y 8 se obtiene que la diferencia es casi nula ya que solo presentan pocas diferencias entre los dos histogramas. Esto se puede apreciar en la Fig. 9.

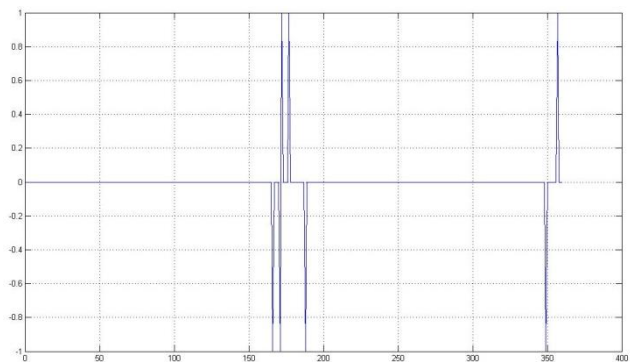


Fig.9 Error obtenido de la diferencia de histogramas de distancias de las Figs. 5 y 8 para el símbolo que representa la letra “w”.

Con esta variación de las imágenes para la etapa de verificación de reconocimiento por parte de la red neuronal se obtuvieron los resultados mostrados en la matriz de confusión, por referencia cruzada, que se puede apreciar en la Tabla 1.

De la información de la matriz de confusión se obtiene que:

- Las símbolos que siempre se reconocen de manera adecuada son los que corresponden a las letras b, c, d, e, f, g, i, n, o, u, v, x, y.
- Si bien el sistema confunde algunos símbolos casi siempre los relaciona con el mismo

símbolo. Como en el caso del símbolo de la letra l el cual lo clasificó como la g de forma errónea.

- c. El porcentaje de reconocimiento es del 89.5%.

	a	b	c	d	e	f	g	h	i	l	m	n	o	p	q	r	s	t	u	v	w	x	y
a	29	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
c	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
d	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
e	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
f	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
g	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
h	0	0	0	0	0	0	1	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
i	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l	0	0	0	0	0	0	12	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0
m	0	0	0	0	0	0	0	0	0	27	3	0	0	0	0	0	0	0	0	0	0	0	0
n	0	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0
o	0	0	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0
p	0	0	0	0	0	0	0	0	0	0	0	0	18	1	4	7	0	0	0	0	0	0	0
q	0	0	0	0	0	0	0	0	0	0	0	0	0	4	25	1	0	0	0	0	0	0	0
r	0	0	0	0	0	0	0	0	0	0	0	0	0	2	28	0	0	0	0	0	0	0	0
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	16	0	0	0	0	0	0
t	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24	0	0	0	6	0	0
u	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0
v	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30	0	0	0	0
w	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	15	0	0
x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30	0	0
y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30

Tabla 1. Matriz de Confusión para el reconocimiento de los símbolos del LSM.

Este porcentaje de reconocimiento se obtuvo de las pruebas de validación que se realizaron utilizando todos los símbolos aprendidos por el sistema.

Cabe mencionar que al realizar el mismo experimento con las imágenes de entrenamiento el sistema siempre clasificaba de forma correcta todos los símbolos. Dicho de otra manera se tiene un porcentaje de reconocimiento del 100%.

Finalmente se probó la interpretación de los símbolos y se evaluó en un contexto de simulación robótica. Para ello se diseñó el robot en el entorno de simulación Roboworks [24].

Símbolo	Acción
B	Detener robot
C	Configuración de estacionado del brazo
D	Configuración de homing del brazo
E	Girar robot
G	Avanzar el robot
U	Incrementar velocidad de movimiento
N	Decrementar velocidad de movimiento
Y	Rutina de servir copa

Tabla 2. Asociación de símbolos con una acción a desarrollar por el robot.

Para ello se asignó una tarea a un grupo de signos del alfabeto LSM. Se utilizaron los símbolos que reconocía mejor el sistema. Dicho de otra forma aquellos que durante la etapa de evaluación de reconocimiento siempre fueron reconocidos. Esos símbolos y las acciones correspondientes asignadas a cada uno de ellos se muestran en la Tabla 2.

Se verificaron cada una de las acciones anteriores usando como medio de evaluación el entorno de simulación robótica Roboworks[24]. En la Fig.10 se muestran evoluciones de la simulación realizada para que el robot sirva una copa.

La posición de inicio de la simulación se muestra en la imagen (a) de la Fig. 10. Las imágenes de (b)-(e) son diferentes capturas en la progresión de la tarea asignada. Y en la última imagen (f) es la posición final de la asignación “sirve la copa”.

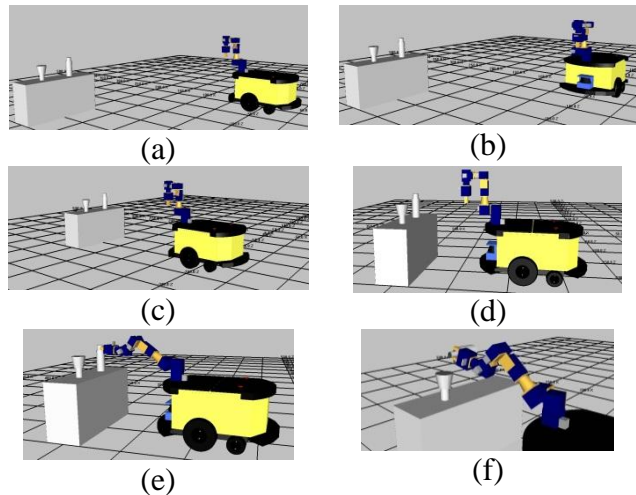


Fig. 10. Evolución de la tarea “sirve copa” desarrollada por un robot móvil equipado con un manipulador de 6 DOF.

Para la simulación mostrada en la Fig. 10 el sistema reconoce la seña que representa el símbolo “Y” del alfabeto LSM. Una vez que el sistema identifica la seña la interpreta y envía el conjunto de subtareas que debe de ejecutar el sistema a fin de poder servir la copa. En la Tabla 2 se puede observar que es la acción asignada a el símbolo “Y”.

El código que se utilizó para realizar las diferentes etapas de este trabajo se programó en Matlab® usando una computadora portátil Dell XPS con microprocesador Intel Core 2 Duo 2.0GHz y 3 Gb de Ram. Además se utilizó Visual Studio 2008 Express para interactuar con el entorno de simulación RoboWorks.

IV. Conclusiones

El desarrollo del presente trabajo nos permitió obtener un sistema robótico capaz de reconocer e interpretar un conjunto de señas del lenguaje de señas mexicano para la realización de diversas tareas. Si bien el sistema es capaz de reconocer 23 símbolos del alfabeto de lenguaje de señas mexicano solo se asignaron tareas a ocho de ellos.

Se obtuvo un 89.5% de reconocimiento de los símbolos aprendidos, esto debido principalmente a cambios en la iluminación en las imágenes de prueba y a la similitud entre algunos símbolos.

Las tareas que desarrolló el robot están orientadas hacia la robótica de servicio.

Como trabajo futuro se prevé asignar una tarea a cada símbolo del alfabeto de señas. Esta asignación de tareas podrá ser a partir de una sola seña o mediante una secuencia de ellas.

Otra consideración será la utilización de imágenes con fondo variable y sin importar si la persona que hace las señas usa una prenda en particular, como en las imágenes usadas para este trabajo. Además de tratar problemas de rotación, escala y de oclusiones.

También se implementará sobre el sistema robótico real para lo cual se utilizará la plataforma móvil y el brazo de seis grados de libertad, PowerBot y PowerCube respectivamente.

Otro campo de aplicación de este reconocedor podría ser como traductor del lenguaje de señas mexicano al lenguaje hablado.

V. Referencias

- [1] C. Breazeal, *Designing Sociable Robots*, 2002, The MIT Press.
- [2] C. Breazeal, *Sociable Machines: Expressive Social Exchange Between Humans and Robots*, Sc.D. dissertation, Department of Electrical Engineering and Computer Science, MIT, 2000.
- [3] Sitio web del Instituto de robótica y mecatrónica del centro aeroespacial alemán, http://www.dlr.de/rm/en/desktopdefault.aspx/tabid-5471/8991_read-16694/
- [4] A. Clodic, S. Fleury, R. Alami, M. Herrb, R. Chatila., *Supervision and Interaction: Analysis of an Autonomous Tour-Guide Robot Deployment*, 12th International Conference on Advanced Robotics, ICAR05, Seattle, USA, 2005.
- [5] Sitio web de Rackham en el museo espacial, <http://spiderman-2.laas.fr/robots/rackham/data/en/rackham.php>
- [6] Sitio web de Wow Wee: Robosapien, <http://www.wowwee.com/en/products/toys/robots/robotics/robosapiens/robosapien>
- [7] J. Samans, *The Robosapien Companion: Tips, Tricks, and Hacks (Technology in Action)*, 1 edition, 2005, Springer, USA.
- [8] S. Behnke, J. Müller, M. Schreiber, *Playing Soccer with RoboSapien, RoboCup-2005: Robot Soccer World Cup IX*, pp. 36-48, *Lecture Notes in Artificial Intelligence*, LNAI 4020, Springer, 2006.
- [9] L. Pineda, I. Meza, A. Espinosa, *Direct Interpretation of Speech Acts*, Reporte Interno DCC-IIMAS, 2003.
- [10] Sitio web del proyecto Cogniron, <http://www.cogniron.org/final/Home.php>
- [11] R.Alami, L.Brethes, R.Chatila, A.Clodic, M.Cottret, G.Covo, P.Danes, M.Devy, N.Do Huu, X.Dollat, F.Fleury, M.Fontmarty, M.Herrb, I.Herrera Aguilar, C.Lemaire, F.Lerasle, E.Lopez Damian,

- J.Manhes, P.Marcoul, L. F.Marin Urias, V.Montreuil, R.Philippsen, D.Sidobre, E.A.Sisbot, F.J.Trujillo Romero, Mid-term key experiment specification and implementation status Rappports de Contrat, COGNIRON, Project FP6-IST-002020, Juin 2006, 30p.
- [12] S. Lankton, Sparse Field Methods, Technical Report, Georgia Tech, July, 2009.
- [13] A. Blake, M. Isard, Active Contours, Springer, Cambridge, 1998.
- [14] T. Chan and L. Vese. Active contours without edges. IEEE Transactions, on Image Processing, 10(2):266–277, February 2001.
- [15] K. Fujimura, Y. Sako, Shape Signature by Deformation, In Shape Modeling International (SMI 99), IEEE Computer Society, Aizu, Japan, March, 1999.
- [16] S. Giannarou, T. Stathaki, Shape Signature Matching for Object Identification Invariant to Image Transformations and Occlusion, Computer Analysis of Images and Patterns, Lecture Notes in Computer Science, 4673: 710-717, Springer, 2007.
- [17] R. Gal, A. Shamir, D. Cohen-Or, Pose-Oblivious Shape Signature, IEEE Transactions on Visualization and Computer Graphics, 13, 2, 261-271, March, 2007.
- [18] S. Belongie, J. Malik, J. Puzicha, Shape Matching and Object Recognition Using Shape Contexts, IEEE Transactions on Pattern Analysis and Machine Intelligence (24): 509–521, April, 2002.
- [19] M. T. Hagan, H.B. Demuth, M. H. Beale, Neural network design, PWS Pub, 1995.
- [20] Q. Munib, M. Habeeb, B. Takruri, H. A. Al-Malik, American sign language (ASL) recognition based on Hough transform and neural networks, Expert Systems with Applications, Vol. 32, issue 1, pages 24-37, ISSN 0957-4174, January, 2007.
- [21] L. P. Vargas, L. Barba Jiménez, L. Mattos, Sistema de identificación de Lenguaje de Señas usando Redes Neuronales Artificiales, Revista Colombiana de Física, Vol 42, No 2, 2010.
- [22] A. Karami, B. Zanj, A. K. Sarkaleh, Persian sign language (PSL) recognition using wavelet transform and neural networks, Expert Systems with Applications, Volume 38, Issue 3, Pages 2661-2667, ISSN 0957-4174, March 2011.
- [23] M.P. Paulraj, S. Yaacob, M.S. bin Zanar Azalan, R. Palaniappan, A phoneme based sign language recognition system using skin color segmentation, 6th International Colloquium on Signal Processing and Its Applications (CSPA), pp.1-5, May 2010.
- [24] Sitio web de los desarrolladores de Roboworks©, <http://www.newtonium.com>
- [25] S. Lankton, A. Tannenbaum, Localizing Region Based Active Contours, IEEE Transaction on Image Processing. Vol. 17, No. 11, pp. 2029-2039, November 2008.

VI. Autores

Felix Emilio Luis Pérez es egresado de la carrera de Ingeniería en Computación de la Universidad Tecnológica de la Mixteca actualmente estudia la maestría en robótica en la misma institución.

Felipe de Jesús Trujillo Romero es Dr. en Sistemas Informáticos por el Instituto Nacional Politécnico de Toulouse, Francia, actualmente es profesor investigador de la División de Estudios de Posgrado de la Universidad Tecnológica de la Mixteca.

Herramienta de Clasificación de Información Basada en Métodos de *Clustering*: Aplicación a la Segmentación de Imágenes

*Ing. José Miguel Torres Garfias, **Dr. Antonio Orantes Molina

Universidad Tecnológica de la Mixteca, Carretera a Acatlima km 2.5, Huajuapán de León, Oaxaca, México.

* jmtg139@hotmail.com

** tonito@mixteco.utm.mx

Resumen

El presente documento muestra los resultados al aplicar la herramienta genérica **GUILSI** (*Generic tool of classification Using the algorIthm Lamda for Supervision of functional states and assIgnment of elements*) basada en el algoritmo de clasificación **LAMDA** (*Learning Algorithm for Multivariable Data Analysis*) para la segmentación y distinción de partes en una imagen.

LAMDA es una herramienta de clasificación secuencial con aprendizaje y reconocimiento de datos multivariados basada en la Lógica Difusa (*Fuzzy Logic*). Permite el aprendizaje secuencial con o sin clases iniciales. Permite diferentes particiones del mismo conjunto de datos gracias al concepto de “grado de exigencia”. Existe una clase NIC (Clase No informativa), la cual actúa como un límite para asignar un objeto a una clase significativa. Esta clase NIC es asignada automáticamente por el contexto. El algoritmo **LAMDA**, la herramienta **GUILSI**, el proceso de clasificación y los resultados obtenidos son descritos a lo largo del documento.

Palabras clave: Clasificación, clustering, GUILSI, LAMDA, segmentación.

I. Introducción

Los objetivos perseguidos en esta investigación se enfocan en el desarrollo de herramientas de fusión de la información que ofrezca a los operadores de procesos industriales u otro tipo de sistema, una información fácilmente interpretable, tanto en funcionamiento normal como en presencia de fallas.

Para alcanzar estos objetivos proponemos una herramienta que pueda ajustarse a necesidades particulares de clasificación y en la que el tipo de datos de la aplicación no sea una limitante, es decir una herramienta de tipo genérica, por lo que nos enfocamos a los métodos de clasificación de datos y reconocimiento de patrones para caracterizar los estados funcionales de un proceso. Se propone la utilización de un nuevo método de clasificación con aprendizaje no supervisado y supervisado denominado **LAMDA**, que presenta las características útiles para combinar los aspectos cuantitativos y cualitativos, así como dirigir los estudios de diagnóstico.

Actualmente, nuestras principales aplicaciones se sitúan en el dominio de los procesos complejos industriales, segmentación de imágenes, detección temprana de enfermedades como el cáncer, pronóstico de catástrofes ambientales e incluso clasificación de estados psicológicos.

II. El algoritmo LAMDA

LAMDA es un método de clasificación desarrollado por N. Piera y J. Aguilar [1] [2]. Como primera concepción de **LAMDA** se puede decir que es una metodología de base radial (datos tipo *ratio*), que se fundamenta en clases prototipo. Cada conjunto de descriptores de objetos es comparado con un prototipo a partir de la medición de la distancia para determinar su clase.

Las principales características de la metodología **LAMDA** son las siguientes [3].

- Distinción de los objetos (indistinguibles) mediante el concepto de la Clase No Informativa (NIC).

- Permite realizar clasificaciones supervisadas (clases iniciales existentes) y no supervisadas (sin clases iniciales existentes).

El algoritmo de *clustering* **LAMDA** permite desarrollar un sistema que clasifique datos multivariados mediante dos tipos de grados de adecuación: Grado de Adecuación Marginal (MAD) y Grado de Adecuación Global (GAD). Los puntos más importantes con que el algoritmo **LAMDA** realiza la clasificación se describen a continuación.

Colección de objetos y descriptores

La metodología **LAMDA** consiste en que se tiene una colección de objetos \vec{X} y un conjunto finito de n descriptores cualitativos o cuantitativos. Un descriptor es una característica tomada de una señal y que se caracteriza por su aporte de información sobre algún evento.

Un objeto \vec{X} es representado por un vector de n componentes llamados descriptores. La información que cada descriptor aporta contribuye al grado de pertenencia de un objeto a una clase mediante el grado de adecuación marginal (MAD). La información que aportan en conjunto los valores MAD de cada elemento, definen el grado de adecuación global a cada clase. El grado de adecuación global (GAD) a cada clase se obtiene combinando los MAD mediante un operador lógico difuso. Ver tabla 1.

Operadores lógicos		Intersección (T-Norma) T(x,y)	Unión (S-Conorma) S(x,y)
Conectivo	Probabilísticos	Producto	Suma
	Difusos	Mínimo	Máximo
	Mixtos	$\alpha = 1$... 0.5	0

Tabla 1. Operadores lógicos y conectivos

En la tabla podemos observar que la variable α puede tomar valores de entre 1 (donde se considera 100% un operador lógico T-Norma) y cero (donde se considera 100% operador lógico T-Conorma). Además existen tres tipos de

conectivos: probabilísticos, difusos y mixtos (una combinación de ambas). Esto se representa con la siguiente ecuación:

$$CM(ix, y) = \alpha * T(x, y) + (1 - \alpha) * S(x, y) \quad (1)$$

Donde α es el grado de exigencia.

Normalización de los objetos

Antes del cálculo del GAD y la MAD, es necesario normalizar los vectores que representan los objetos. La normalización se realiza con el propósito de suavizar los vectores y establecer una métrica estándar para los objetos. Cada vector \vec{X} posee un valor mínimo (x_{min}) y un valor máximo (x_{max}). Cada componente (x_i) del vector es normalizado mediante la siguiente fórmula:

$$x_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (2)$$

Cálculo del MAD y la GAD

Después de normalizar los valores de los componentes, se forma el vector \vec{p} , que se calcula obteniendo la media aritmética de los valores de los componentes de cada clase existente. Cada elemento de este vector, es considerado como el valor más representativo de cada descriptor existente.

El MAD, el cual es un vector, se puede calcular mediante cuatro funciones diferentes, atendiendo a la naturaleza de la distribución de los elementos:

MAD binomial, generalmente es la función utilizada, y se calcula de la siguiente forma:

$$\mu_{cj}(x) = \rho_{i,j}^{x_j} (1 - \rho_{i,j})^{(1-x_j)} \quad (3)$$

MAD binomial-central, es una modificación de la anterior:

$$par = \rho_{i,j}^{x_j} (1 - \rho_{i,j})^{(1-x_j)} \quad (4)$$

$$des = \rho_{i,j}^{x_j} (1 - x_j)^{(1-x_j)} \quad (5)$$

$$\mu_{cj}(x) = \frac{par}{des} \quad (6)$$

MAD binomial-distancia, basado en la distancia entre dos puntos:

$$a = \max \left[\rho_{i,j}^{x_j}, (1 - \rho_{i,j}) \right] \quad (7)$$

$$x_{dist} = 1 - \text{abs}(x_j - \rho_{i,j}) \quad (8)$$

$$\mu_{cj}(x) = (a)^{x_{dist}} (1 - a)^{(1-x_{dist})} \quad (9)$$

MAD Gaussiana, basada en la función de Gauss:

$$\mu_{cj}(x) = e^{-\frac{1}{2\sigma^2_{ij}}(x_j - \mu_{i,j})^2} \quad (10)$$

Donde:

$\mu_{cj}(x)$ es el valor de adecuación marginal (MAD) con respecto al descriptor i de la clase j .

$\rho_{i,j}$ es el valor promedio de los valores correspondientes al descriptor i pertenecientes a la clase j .

x_j es el valor del descriptor j del objeto \vec{X} .

El valor de GAD (el cual es un escalar) del objeto \vec{X} con respecto a la clase k dependerá del vector MAD y del operador lógico y conectivo utilizado para su cálculo, como se muestra en la siguiente ecuación:

$$GAD(\vec{X}/C_k) =$$

$$\alpha * T(MAD_{k_1}(x_1), \dots, MAD_{k_n}(x_n)) + (1 - \alpha) * S(MAD_{k_1}(x_1), \dots, MAD_{k_n}(x_n)) \quad (11)$$

Por ejemplo, si se utiliza un operador lógico 100% T-Norma y un conectivo probabilístico, la GAD se obtiene mediante la siguiente fórmula:

$$GAD(\vec{X}/C_k) = \prod_{i=1}^{nd} \mu_i(x_i/c_k) \quad (12)$$

Donde:

$GAD(\vec{X}/C_k)$ es el valor de la GAD del objeto \vec{X} a la clase k .

nd es el número de descriptores de un objeto.

$\mu_i(x_i/c_k)$ es el valor de adecuación marginal (MAD) del descriptor i a la clase k .

Asignación mediante GAD y NIC

El algoritmo LAMDA no utiliza la similitud entre los elementos para la clasificación, sino el cálculo de grado de adecuación global entre el objeto y las clases formadas.

Cada objeto se verá asignado a una de las clases existentes con la condición de que su grado de adecuación global (GAD) sobrepase el valor de la clase NIC. Cuando el elemento es asignado a una de las clases existentes, se realiza una modificación de los parámetros de esa clase, utilizando la ecuación de la media general:

$$\rho_{i,j} = \rho_{i,j} + \frac{x_j - \rho_{i,j}}{N+1} \quad (13)$$

Donde:

N es el número de elementos asignados a la clase.

Si el valor de la GAD no sobrepasa el valor de la NIC (valor definido por el usuario), significa que esta es la clase más próxima al elemento y que no pertenece a alguna de las clases existente. La clase NIC es no informativa, por lo tanto se crea una clase nueva y se le asigna el nuevo objeto. En la figura 1 se muestra el cálculo del GAD del objeto \vec{X} con respecto a la clase C_j .

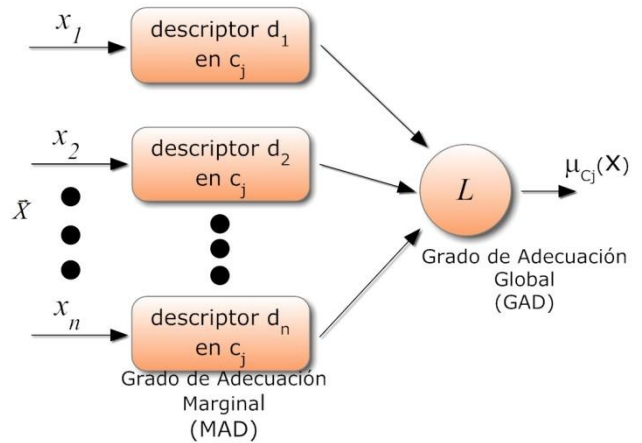


Fig. 1. Esquema de clasificación

Este proceso se repite para todas las clases existentes en su momento. El objeto \vec{X} será asignado a la clase cuyo GAD sea el mayor.

Algunas aplicaciones en donde LAMDA ha sido utilizado son: clasificación y localización de fallas en sistemas eléctricos [4] [5], modelado del comportamiento de clientes de bancos [6], alternativas para el mejoramiento de la continuidad del suministro de energía [7], clasificación de glóbulos rojos [8], clasificación de la personalidad y sus trastornos [9],

validación y reconstrucción de los datos de medidores de flujo en redes de distribución de agua [10], control y supervisión de plantas tratadoras de aguas residuales [11], entre otros.

III. Una nueva herramienta genérica de clasificación: GUILSI

A continuación se describe a la herramienta **GUILSI**. **GUILSI** es la herramienta que proponemos y es desarrollada de acuerdo a nuestras necesidades de investigación.

La herramienta GUILSI

GUILSI es una herramienta que fue desarrollada en el lenguaje **Matlab**. Este lenguaje puede ser más lento que los lenguajes de alto nivel como C, C++ o Fortran, pero los fines de utilizar **Matlab** son de experimentación e investigación, el cual cuenta con funciones que permiten un rápido desarrollo de una nueva aplicación, cuenta con capacidades gráficas, permite el uso de funciones y *toolbox* para el uso de matrices, vectores, estructuras de datos, almacenar, recuperar, comparar datos y acceder a muchas rutinas útiles.

GUILSI como herramienta genérica

El objetivo de desarrollar la herramienta **GUILSI**, es aplicar el algoritmo de clasificación **LAMDA**, de manera que resulte una herramienta de tipo genérica, que pueda ser ajustada a necesidades de investigación, y que pueda ser acoplada con otras herramientas desarrolladas también en el lenguaje **Matlab** (como por ejemplo, un pre-tratamiento de datos brutos). **GUILSI** permitirá analizar y comparar los resultados de clasificación mediante gráficas, archivos de salida, y visualización de indicadores en pantalla, como un medio de apoyo para los usuarios tratadores de imágenes y supervisores de procesos complejos.

Modo de operación

GUILSI cuenta con dos modalidades de operación: *Offline* y *Online*. El modo de operación es seleccionado de acuerdo a los

archivos de entrada. Cuando se cuenta con un archivo de historial se considera que el modo de trabajo es *Online*, y se clasificarán en línea los datos de un archivo nuevo. Cuando no se selecciona un archivo de historial, se considera que el tipo de operación es *Offline*, donde se clasifican los datos de un archivo nuevo. En la figura 2 se puede observar la modalidad *Offline*.

En la parte **HISTORIAL** se presentan los datos de entrenamiento (conjunto de datos conocidos del sistema y/o operador) y de prueba. Cada dato se le conoce como objeto o elemento.

En la parte de **CONFIGURACION** se definen los parámetros de clasificación, selección del modo de aprendizaje y tipo de ecuación a utilizar en el cálculo de la MAD. En la sección de **CLASIFICACIÓN** se obtiene las clases o *clusters* después de aplicar la metodología **LAMDA** y la matriz de las GAD, con el propósito de poder validar la clasificación resultante. En caso de que el operador valide la clasificación, se continúa a la siguiente etapa, que es, **ASIGNACIÓN DE LOS ESTADOS FUNCIONALES**, donde se agrupan diversas clases para formar un estado funcional. Finalmente, al conjunto de estados conforman al modelo de comportamiento de todo el sistema.

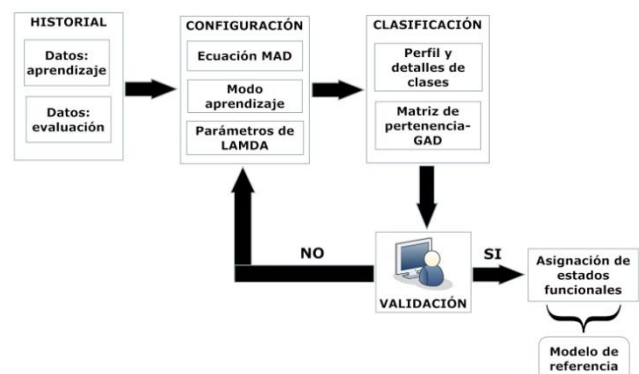


Fig. 2. GUILSI: estructura fuera de línea

En la figura 3 se puede observar la modalidad *Online*. En esta modalidad se tiene como objetivo obtener el diagnóstico de un sistema complejo. Se trata de comparar los datos en línea procedentes del sistema a diagnosticar con el modelo de referencia o de comportamiento obtenido en la modalidad *Offline*. Como

resultado de esta comparación se obtiene el reconocimiento del estado actual del sistema (estado normal o de fallo). Finalmente se somete a un análisis por parte del operador para la toma de decisión, la cual depende si el sistema se encuentra en un estado de operación normal o si ocurre una falla.



Fig. 3. GUILSI: estructura en línea

También se puede generar un modelo de comportamiento en base a una o varias clasificaciones previas, donde se pueden crear nuevos estados, modificar los existentes, y fusionar clases y estados.

La herramienta **GUILSI** fue implementado para identificar los estados funcionales de una planta química y diagnosticar fallas, obteniendo muy buenos resultados. También se utilizó para clasificar datos estadísticos de población. Pero dado que es una herramienta de tipo genérica, se puede utilizar esta en otras aplicaciones o tipos de datos como una imagen.

IV. Segmentación de imágenes

Cada dato de entrada para el algoritmo **LAMDA** consiste de un vector de n variables. Cada variable representa un descriptor o señal característica del dato. Sin embargo, clasificar una imagen se requiere de un pre-tratamiento de los píxeles para formar el vector de entrada (llamado objeto).

Para esto se realizó una aplicación en la herramienta **Matlab** que permite obtener a partir de una imagen un archivo de texto que contiene los datos de entrada a la herramienta **GUILSI**.

Posteriormente, la herramienta clasifica esos datos y genera un archivo de texto que contiene los datos ya tratados (en este caso píxeles) y la correspondiente clase a la que fue asignado cada uno.

La clasificación de la imagen ha sido realizada por medio de sus vectores representantes. Cada vector representa las características de su correspondiente píxel. El resultado entregado por esta clasificación es un archivo de texto. Sin embargo, se tiene que convertir los datos de este archivo a imágenes, dado que el formato del archivo es diferente y los datos que contiene no son los reales de la imagen original.

Una aplicación más de la herramienta de **Matlab** se encarga de generar a partir del archivo de texto de salida, una imagen para cada clase.

Aplicación en MATLAB para generar los datos de entrada a GUILSI a partir de una imagen

Esta aplicación recibe como entrada una imagen y genera como salida dos archivos de texto. Un archivo de texto contiene los datos de entrada a la herramienta **GUILSI**, que representan algunas características de cada píxel. El segundo archivo contiene la relación de cada dato del primer archivo con su correspondiente píxel y posición en la imagen original.

En los siguientes puntos se describe cómo trabaja esta aplicación.

Lectura de la imagen

La aplicación lee una imagen, la cual está en formato .tiff (sin compresión). Si la imagen está en colores se encarga de convertirla a escala de grises.

Posteriormente la imagen es almacenada en una matriz.

Es recomendado que la imagen sea de dimensiones moderadas, puesto que la magnitud del archivo de entrada a la herramienta **GUILSI** depende de ésta.

Formación de los vectores representantes de la imagen

En esta etapa se forman los vectores de entrada a la herramienta **GUILSI**. Cada uno de estos vectores se forma de cada píxel y de cuatro de sus vecinos adyacentes. De esta manera, el píxel actual y sus vecinos adyacentes representan los

descriptores. El número de vectores será de $x*y$, donde “ x ” es el número de columnas y “ y ” el número de filas de la imagen. Por lo tanto se puede decir que la imagen ha sido linealizada.

Relación de los vectores representantes de la imagen con los de la imagen original

Cada uno de los vectores formados representa un pixel y sus características de acuerdo a la imagen completa, con los que esta ha sido linealizada. Para no perder la correspondencia de estos vectores con respecto a los pixeles y sus posiciones de la imagen original, se forma una matriz de $x*y$ filas y dos columnas. Donde cada fila representa un pixel, la primera columna indica la posición en x y la segunda columna indica la posición en y del pixel respecto a la imagen original.

Aplicación en MATLAB para la generación de las imágenes resultantes de la clasificación en GUILSI

Después de haber generado la clasificación en GUILSI de los datos que representan la imagen, es necesario generar una imagen por cada clase, es decir, visualizar la segmentación de la imagen original. Para ello, se utilizan los datos de cada clase del archivo de salida de la herramienta GUILSI y el archivo que guarda la relación de los pixeles de la imagen original. El propósito de este archivo es permitir reconstruir las imágenes correspondientes a los resultados de clasificación. De esta manera se generan tantas imágenes como clases se hayan obtenido.

El número de imágenes obtenidas, la cantidad de pixeles de cada una y la sensibilidad de la segmentación dependerán del método seleccionado (por ejemplo, la función binomial) para la clasificación y del grado de exigencia utilizado (α).

V. Resultados

Aplicación 1: Separación de unas monedas del fondo de una imagen

La imagen mostrada en la figura 4 se introdujo a la aplicación en **Matlab** para la generación de los vectores representantes de la imagen.



Fig. 4. Monedas

Los vectores formados que representan la imagen fueron introducidos a la herramienta GUILSI y como resultado se obtuvieron tres clases principales (denominadas: clases 1, 2 y 3), cada una agrupando los pixeles correspondientes, como se muestra en la figura 4.



Fig. 5. Clases y pixeles después de la clasificación

Posteriormente las clases y sus elementos obtenidos pasaron por la aplicación en **Matlab** para la reconstrucción de las imágenes, una por cada clase. Con esto se pudo identificar, de la imagen original, el fondo, el contorno y el objeto. En la figura 5 se muestra la imagen original y la imagen segmentada. En la parte izquierda de la ventana se seleccionan los parámetros de clasificación. La gráfica superior muestra los valores normalizados de los pixeles de la imagen, y la gráfica inferior muestra la clase a la que se asignó cada uno de esos pixeles.

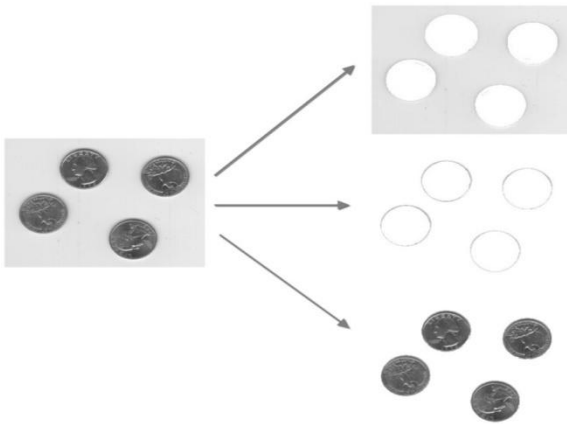


Fig. 6. Imagen original y segmentada (fondo, contorno y objetos)

Aplicación 2: Separación del núcleo y citoplasma en una célula

Como segunda implementación, para probar el funcionamiento de la herramienta **GUILSI** con datos provenientes de imágenes, se utilizó la imagen de una célula del epitelio vaginal. Con lo que se buscaba separar de la imagen, el núcleo y el citoplasma, como se realizó en [12]. La imagen introducida se muestra en la figura 7, luego esa imagen fue convertida a escala de grises, mostrada en la figura 8.



Fig. 7. Célula a color



Fig. 8. Célula en escala de grises

En la figura se puede apreciar las partes de la célula, como el núcleo y citoplasma.

La aplicación en **Matlab** para generar los vectores representantes de la imagen fue introducida a **GUILSI**. El resultado de clasificar los datos de la imagen se muestra en la figura 9.

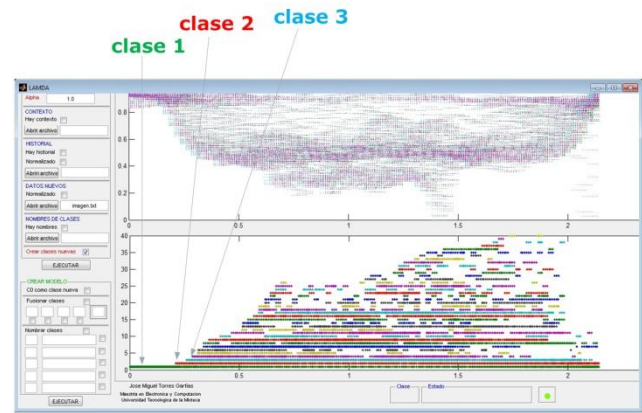


Fig. 9. Resultado en GUILSI de la clasificación de la imagen

La herramienta **GUILSI** generó varias clases, sin embargo de acuerdo a su similitud algunas fueron fusionadas con el fin de no tener demasiado segmentada la imagen.

De las clases obtenidas, cinco son las principales, con los que posteriormente se formó una imagen de cada una. Estos resultados (características de la imagen y respectivas clases) fueron introducidas a la aplicación en **Matlab** para reconstruir las imágenes, mostradas en la figura 10, donde se puede apreciar las partes de la célula, habiendo también separado el fondo.

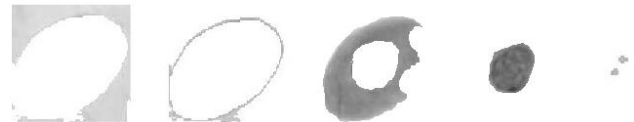


Fig. 10. Partes de la célula

Modificando el método de clasificación y grado de exigencia se pueden obtener distintos resultados a los mostrados en la figura 10. Es decir, se puede sintonizar el algoritmo de clasificación con el propósito de aproximar los resultados esperados por el experto.

VI. Conclusiones

- Se diseñó una herramienta genérica para comprender fácilmente la información proveniente de diversos sistemas, como son la segmentación de imágenes.
- En la segmentación de imágenes, se utilizó la herramienta **GUILSI** como aprendizaje para

crear las clases (*clusters*) que caracterizan distintas porciones de la imagen.

- **GUILSI** ha sido utilizado para identificar procesos complejos en una planta química, pero se deseó probar su funcionamiento con imágenes, y se obtuvieron los resultados mencionados. El tratamiento de imágenes se irá refinando para obtener mejores resultados y darle otras aplicaciones.
- La segmentación de imagen o identificación de partes en tejido celular, puede permitir identificar anomalías de crecimiento, identificar tumores y detección temprana de enfermedades como el cáncer de mama.
- Los descriptores a utilizar dependerán del experto, según él considere cuales son aquellos que representen las señales más características y representativas del proceso.
- En la modalidad *offline* el número de clases formadas no puede determinarse antes del proceso de clasificación, sin embargo después de que esta ha sido realizada si el experto lo considera necesario puede fusionar clases o incluso omitir algunas de ellas si lo cree conveniente, por ejemplo si una de ellas agrupa muy pocos elementos.
- Una ventaja de la herramienta **GUILSI** es que no acumula todos los datos de la base histórica, sino que a partir de ellos realiza una abstracción y forma una nueva base de datos, con característica evolutiva, que consiste de un conjunto de vectores representativos que forman el modelo de comportamiento.
- El algoritmo **LAMDA** resultó ser un algoritmo de tipo genérico, los tipos de datos que permite no son restringidos ni limitados, así como de tipo cuantitativo y cualitativo simultáneamente.

VII. Referencias

- [1] N. Piera, P. Desroches, J. Aguilar Martin, "LAMDA: An incremental conceptual clustering method". Rapport du LAAS No. 89420. Laboratoire d'analyse et d'architecture des systèmes du CNRS, décembre 1989.
- [2] Marie-Véronique LE LANN, J. Aguilar Martin, A. Orantes. "Outil d'aide à la sélection des capteurs pour la détection de fautes". Paris, France. Récents Progrés en Génie des Procédés no. 92. 2005. ISBN 2-910239-66-7.
- [3] T. Kempowsky, J. Aguilar Martin, M-V Lelann, Subias, "A Learning Methodology of a supervision System using LAMDA Classification Method". LAAS/CNRS. Iberamia. 2002.
- [4] J. Mora Flórez, V. Barrera-Núñez, G. Carrillo Caicedo. "Fault Location in Power Distribution Systems Using a Learning Algorithm for Multivariable Data Analysis". IEEE transactions on power delivery, vol. 22, no. 3, Julio 2007.
- [5] J. Mora Flórez, D. Llanos, J. Meléndez, J. Colomer, J. Sánchez, M. Castro. "Classification of disturbances in electrical systems". 2007.
- [6] J. Mora Flórez, J. Colomer Llinás. "Statistical analysis and a conceptual clustering method to rank clients of a financial bank". Scientia et Technica Año X, No 25. Agosto 2004. UTP. ISSN 0122-1701.
- [7] V. Barrera Núñez. "Una solución económica a los problemas de calidad del servicio del suministro de energía eléctrica". Revista Ingeniería y Ciencias. Septiembre 2006. ISSN 1794-9165.
- [8] C. Allayous, S. Regis, A. Bruel, D. Schoevaert, R. Emilion, T. Marianne-Pepin. "Velocity allowed red blood cell classification". 10th International IFAC Symposium on Computer Applications in Biotechnology. Vol. 1. Junio 2007. Cancún, México.
- [9] J. Aguilar Martin, M. Galindo de Ariza. "Clasificación de la personalidad y sus

trastornos, con La herramienta LAMDA de inteligencia artificial en una muestra de personas de origen hispano que viven en Toulouse –Francia”. Revista de estudios sociales no. 18. Agosto 2004.

- [10] J. Aguilar Martin, J. Quevedo, V. Puig, G. Cembrano, J. Blanch, D. Saporta, G. Benito, M. Hedó, A. Molina. “Validation and Reconstruction of Flow Meter Data in the Barcelona Water Distribution Network”. Abril 2010.
- [11] J. Waissman, C. Youssef, G. Vásquez. “Fuzzy Automata Identification Based on Knowledge Discovery in Datasets for Supervision of a WWT Process”. SETIT 2005. IEEE. Tunisia. Marzo 2005.
- [12] K. Muñoz Arcos. “Identificación de células displásicas en el epitelio vaginal”. Umbral científico no. 9. ISSN: 1692-3375. 2006.

VIII. Autores

Ing. José Miguel Torres Garfias obtuvo su título de Ingeniería en Electrónica en la Universidad Tecnológica de la Mixteca (UTM) en Huajuapán de León, Oaxaca, México.

Dr. Antonio Orantes Molina obtuvo su título de Doctor en Sistemas Automáticos en el Instituto Nacional de Ciencias Aplicadas (INSA) en Toulouse, Francia. Un diplomado en Sistemas Automáticos en el LAAS, Toulouse, Francia. Miembro del SNI desde 2008. Obtuvo el perfil PROMEP desde 2007. Miembro del Consejo Técnico del EGEL CENEVAL de Ingeniería en Mecatrónica. Actualmente jefe de carrera de Ingeniería en Mecatrónica de la UTM.

Reconocimiento de Voz para Comunicación y Diagnóstico de Personas con Disartria en México

***Santiago Omar Caballero Morales, **Gladys Bonilla Enríquez**

Universidad Tecnológica de la Mixteca, Carretera a Acatlima Km 2.5, Huajuapán de León, Oaxaca, 69000

* División de Posgrado +52 (953) 532 02 14, Ext. 300, scaballero@mixteco.utm.mx

** Instituto de Computación +52 (953) 532 02 14, Ext. 200, gbonillae@live.com.mx

Resumen

En México hay un sector significativo de la población que sufre de algún tipo de discapacidad motora. En el caso de personas con Disartria, que es una discapacidad motora del habla, se ve afectada la articulación, la resonancia y la respiración, reduciendo la inteligibilidad del individuo. En la mayoría de estas personas se observa también una discapacidad motora corporal, la cual limita el uso de dispositivos mecánicos, como teclados o apuntadores, para su comunicación. En estos casos, aún cuando la voz se encuentra afectada, resulta ser un medio más factible para comunicación que el uso de dispositivos mecánicos. En este artículo se presenta el desarrollo de un sistema de reconocimiento de voz con fines de comunicación y asistencia para un usuario del Estado de Oaxaca con Disartria y discapacidad motora corporal y visual. Con este sistema, que hace uso de Modelos Ocultos de Markov (HMM's), fue posible reconocer el habla del usuario con tasas de precisión mayores al 85% para frases continuas. De igual manera fue posible obtener información para un diagnóstico preliminar del nivel de Disartria del usuario.

Palabras Clave: disartria, discapacidad en el habla, reconocimiento automático de voz

I. Introducción

El ser humano tiene diferentes formas de comunicarse e interactuar con su medio ambiente, por ejemplo, por señas, por tacto, por voz, etc. Sin embargo en algunos casos existen discapacidades que limitan a una persona el interactuar mediante alguna de estas formas.

Tan solo en México, aproximadamente el 5.1% de la población total tiene alguna discapacidad (física, mental, intelectual o sensorial). De esta población, aproximadamente el 58.3% tiene una discapacidad motora corporal (limitación de la movilidad), y el 8.3% una discapacidad para hablar o comunicarse [1]. Enfermedades como el Mal de Parkinson, Esclerosis Múltiple, o daños neuronales ocasionados por un infarto, una embolia o trauma cerebral, pueden causar ambos tipos de discapacidad.

La disartria se refiere a un grupo de desórdenes motores del habla que resulta del déficit en el

control muscular de los mecanismos del habla debido a un daño en el sistema nervioso periférico o central [2, 3]. Por lo tanto se puede presentar un estado neuromuscular anormal de los músculos articuladores de la voz (parálisis, atrofia, espasticidad), o una disrupción de los movimientos de los mismos (debilidad o falta de coordinación).

A menudo hay un rango reducido en los movimientos y flexibilidad de los articuladores. La hipernasalidad, las consonantes imprecisas, las distorsiones vocales y los problemas en el control de la velocidad del habla, son las características que generalmente se observan en el habla disártrica [4]. Por lo tanto la disartria comprende disfunciones motoras de la respiración, fonación, resonancia, articulación y prosodia.

Adicionalmente, en la mayoría de las personas afectadas con disartria, la discapacidad se

extiende al control de los músculos de las extremidades, limitando sus capacidades para realizar actividades físicas como caminar, escribir, comer, asearse, etc. Sin embargo la capacidad cognitiva de la persona se mantiene completa en la mayoría de los casos.

Para apoyar a estas personas a comunicarse y desplazarse, se han desarrollado herramientas tecnológicas, por ejemplo, control por medio de voz y tacto de sillas de ruedas o dispositivos en el hogar, teclados especiales para marcado telefónico o escritura, apuntadores ópticos, etc. [5]. Sin embargo la mayoría de estos dispositivos pueden ser inaccesibles dado el costo que pueden tener. Por ejemplo, para mejorar la comunicación por voz se tienen el EchoVoice de Saltillo Corporation con un costo de \$495 USD, o el Speech Enhancer de VoiceWave Technology Inc. con \$7500 USD [6]. Es importante señalar que dichos sistemas sólo amplifican la voz pero no hacen reconocimiento y, por lo tanto, no implementan algún proceso de corrección sobre las anomalías que presenta la misma. Es por esto que se ha avanzado en el desarrollo de software con el cual se pueden implementar algoritmos complejos para la corrección de voz, los cuales pueden ser instalados en computadoras portátiles de bajo costo.

El uso de sistemas comerciales de dictado, como Dragon Naturally Speaking, Microsoft Dictation, VoicePad Platinum, e Infovox RA [7,8,9,10,11,12] han mostrado niveles variables de reconocimiento en el rango del 50% al 95% para usuarios con diferentes niveles de disartria y esquemas de uso (palabras discretas o frases continuas), obteniendo los mejores desempeños cuando se usaron vocabularios pequeños (10 – 78 palabras). Proyectos de investigación se han desarrollado para mejorar estos sistemas. En [13] se hizo uso de Redes Neuronales Artificiales (ANN's), las cuales tuvieron mejor desempeño que el sistema comercial IntroVoice. Desempeños significativos también fueron obtenidos con Modelos Ocultos de Markov (HMM's) [14, 15, 16]. En [14] se obtuvieron tasas de precisión en el reconocimiento de voz del 86.9% en usuarios con disartria severa y un

vocabulario de 7-10 palabras para control de dispositivos electrónicos (Radio, TV, etc.).

Es de interés propio el desarrollar en México un sistema de reconocimiento de voz para estas personas, que pueda usarse para habla continua (frases) en lugar de palabras discretas (comandos) con un vocabulario más amplio (>100 palabras). Esto con el fin de que la comunicación de la persona sea más completa y con niveles de precisión mayores al 85%. En el presente artículo se presentan los detalles de desarrollo de este sistema de la siguiente manera: la metodología de diseño, la cual comprende la especificación del contexto de uso, selección del usuario, el diseño de los corpora de voz, la construcción del reconocedor de voz, y la métrica de desempeño, se presenta en la Sección II; en la Sección III se presentan los resultados del sistema propuesto y un análisis de los mismos; en la Sección IV se realiza una discusión acerca del análisis de los resultados; en la Sección V se presentan nuestras conclusiones y comentarios acerca del trabajo a futuro.

II. Metodología de Desarrollo

II.A. Contexto de Uso del Sistema

Como se presentó en la Sección I, se han desarrollado sistemas de reconocimiento de voz para diferentes propósitos, por ejemplo, para redacción de frases cortas [15, 16] o comandos [14]. Para el desarrollo del sistema propuesto se consideraron las necesidades del usuario y sus capacidades para usar dicho sistema. Se observó que los usuarios con ambas discapacidades (vocal y de movilidad) necesitaban de mayor asistencia por parte de sus familiares, por lo tanto el contexto de conversación se identificó como el de solicitar alimentos y cuidados.

El esquema de comunicación se definió por medio de frases cortas por dos razones: (1) la disartria afecta el control sobre la respiración, por lo tanto los usuarios no pueden entablar conversaciones continuas de varias palabras; (2) con 4 o 6 palabras se puede dar la mayor información necesaria para la solicitud de un servicio, algo que en el caso del uso de comandos es muy restrictivo.

El vocabulario y el tipo de frases de uso son dependientes del usuario, ya que cada uno tiene sus necesidades en específico. De igual manera cada usuario, al presentar anomalías únicas en su voz, se puede beneficiar más de un reconocedor personalizado (Dependiente del Usuario, DU) [14, 15, 16]. La metodología presentada se enfoca al desarrollo de un sistema DU, la cual se puede aplicar para el desarrollo de sistemas para otros usuarios.

II.B. Selección de Usuario

La búsqueda de candidatos para el proyecto se realizó en el DIF de la Ciudad de Huajuapán de León en el Estado de Oaxaca. Con asistencia de terapeutas se pudo establecer colaboración con un usuario que actualmente recibe Terapia Ocupacional debido a un accidente vascular cerebral. Por motivos de confidencialidad identificaremos a este usuario como GB.

En la Tabla 1 se presenta el cuadro clínico general del usuario. Debido al accidente vascular cerebral, GB muestra discapacidad motora en sus piernas y brazos (movilidad nula en piernas, y muy limitada en sus brazos). Ajeno a este incidente el usuario tiene visibilidad nula en el ojo derecho y muy limitada en su ojo izquierdo. El nivel de disartria se evaluó por parte de los terapeutas como leve (baja) para la articulación de algunas palabras pero moderada para otras.

Nombre:	GB	
Edad:	67 Años	Sexo: M
Patologías:	Disartria leve-moderada. Accidente Vascular Cerebral. Hemiplejía Izquierda (problemas para caminar y moverse, especialmente del lado izquierdo del cuerpo). Pérdida casi total de la vista.	

Tabla 1. Cuadro clínico del usuario con voz disártrica.

II.C. Vocabulario

• Para Comunicación

Para el diseño del vocabulario y de las frases de comunicación se utilizó la metodología de Diseño Centrado en el Usuario [17]. Se tuvieron visitas semanales con el usuario, observando la manera en cómo llevaba a cabo sus terapias corporales y del habla, tomando nota de la forma en cómo pedía las cosas y entablaba una conversación.

En efecto, GB utilizaba frases cortas, por ejemplo: “quiero ir afuera” o “quiero una taza de café caliente”. Su vocabulario para estas expresiones y lo que usualmente necesitaba era de aproximadamente 100 palabras.

En base a estos hallazgos se documentaron alrededor de 100 frases las cuales se usaron como base para el **Corpus de Prueba - Uso** del sistema. Un corpus es una recopilación de ejemplos reales del uso de una lengua, y puede consistir sólo de textos o de muestras de voz (transcritas a niveles ortográficos y/o fonéticos).

Las 100 frases documentadas se utilizaron como texto de lectura para que el usuario al leerlas proporcionara muestras de voz, integrando así el Corpus de Prueba. Las lecturas fueron grabadas usando un equipo Sony ICD-BX800 con una velocidad de muestra de sonido de 8 kHz monoaural en formato WAV. Estas se etiquetaron (transcripción) manualmente a nivel ortográfico utilizando la herramienta *WaveSurfer*.

Para la construcción del reconocedor de voz se realizó un **Corpus de Entrenamiento** con una estructura diferente. Los detalles se muestran en la siguiente sección.

• Para Entrenamiento del Reconocedor de Voz

El **Corpus de Entrenamiento** se realizó para el entrenamiento supervisado de los modelos acústicos del reconocedor de voz. Por lo tanto este corpus fue etiquetado en ambos niveles, ortográfico y fonético.

La Figura 1 muestra gráficamente el uso de los corpora en las tareas de construcción del

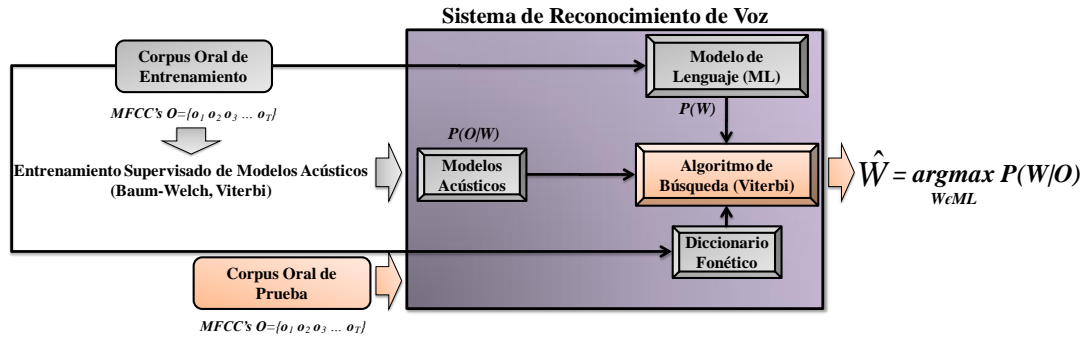


Fig. 1. Esquema de Entrenamiento y Prueba del Reconocedor de Voz.

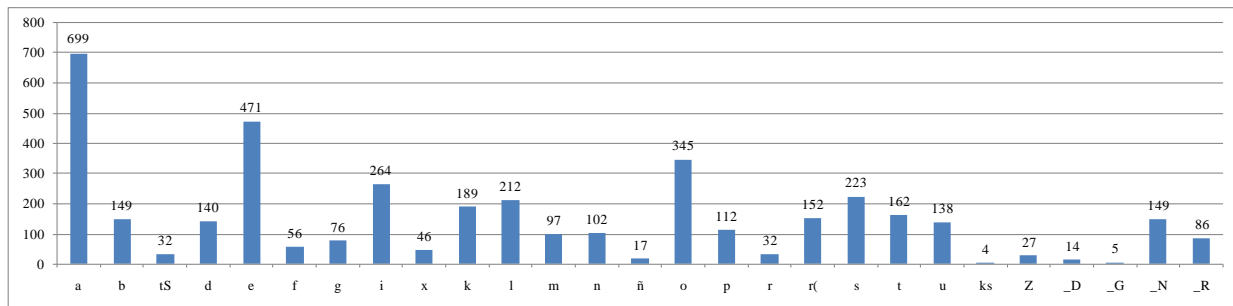


Fig. 2. Distribución de fonemas del español mexicano para el reconocedor de voz.

reconocedor y evaluación del mismo. A manera preliminar también se muestran los elementos del reconocedor de voz, los cuales serán explicados en la siguiente sección.

Este corpus debe tener la característica de estar fonéticamente balanceado. Esto es, que contenga todos los fonemas existentes en la lengua de uso y en cantidades suficientes. Para esto se utilizaron tres diferentes fuentes de texto representativo: (1) un relato de 102 palabras de un Corpus Textual [18]; (2) una lista de 53 palabras utilizadas por los terapeutas para análisis de fonemas; y (3) 16 frases de 65 palabras diseñadas para obtener al menos 5 pronunciaciones de cada fonema en el español mexicano. En la Tabla 2 se muestra el número de lecturas de cada fuente de texto.

Fuente	Repeticiones
Relato/Narración	3
Lista de Palabras	5
Frases Balanceadas	1

Tabla 2. Fuentes para Corpus de Entrenamiento.

Mientras que para el etiquetado ortográfico sólo es necesario escribir la palabra que le corresponde a cada segmento acústico de la voz, el etiquetado fonético consiste de la secuencia de fonemas que forman dicha palabra (por ejemplo, la palabra “hola” está formada por los fonemas /o/, /l/, /a/). Para esto es necesario conocer la fonética del lenguaje de uso, en este caso el español mexicano, y para ello se utilizó la herramienta *TranscribeMex* [19, 20]. Esta herramienta se desarrolló para el etiquetado del Corpus de Español Mexicano DIMEx100 [21].

De esta manera se pudo obtener la transcripción fonética del Corpus de Entrenamiento, el cual tiene la distribución de fonemas mostrada en la Figura 2. Esto se consideró suficiente para el modelado acústico de la voz del usuario.

II.D. Sistema de Reconocimiento del Habla

Como se mostró en la Figura 1, el reconocedor de voz consta de cuatro elementos principales, los cuales se construyeron de la siguiente manera:

- **Modelos Acústicos:** La técnica de Modelos Ocultos de Markov (HMM's) [22] fue utilizada para el modelado de la voz (esto es, estimar $P(O|W)$, la probabilidad de observar la señal O dada una palabra W), siendo la herramienta de implementación el *HTK Toolkit* de la Universidad de Cambridge [23]. Los modelos acústicos se construyeron siguiendo una topología estándar de tres estados con secuencia izquierda-a-derecha y 10 componentes gaussianos por estado. Esta topología, mostrada en la Figura 3, favorece el modelado de fonemas y las variaciones de los mismos en cuanto a duración.

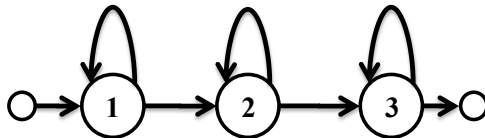


Fig.3. Topología de tres estados izquierda-a-derecha.

Las señales de voz fueron muestreadas cada 10 milisegundos con una ventana Hamming de 25 milisegundos para su codificación, la cual consistió en Coeficientes Cepstrales de Frecuencia Mel (MFCC's). Se utilizaron 12 MFCC's con un coeficiente adicional que representa la energía de la señal acústica [23]. A cada uno de estos coeficientes (13 en total) se les añadió coeficientes delta y de aceleración, dando una codificación de 39 coeficientes. Esto mejora la detección de fonemas [23,24].

Utilizando los módulos *HInit* (para inicialización de modelos) y *HRest/HERest* (para re-estimación de modelos) del *HTK Toolkit*, se realizó el entrenamiento supervisado de los modelos acústicos mediante los algoritmos de Baum-Welch y Viterbi [23,24]. Estos módulos estiman los parámetros de los modelos mediante realineamientos temporales de las muestras de voz del Corpus de Entrenamiento y las respectivas transcripciones a nivel fonético.

Dado que el modelado se hizo a nivel fonético, se realizó un HMM para cada fonema de la Figura 2.

- **Modelo de Lenguaje (ML):** Representa un conjunto de reglas o probabilidades que determinan las secuencias de palabras permisibles en un lenguaje $P(W)$. Por ejemplo, es correcto decir “quiero café caliente”, pero no “caliente café quiero”. En este caso un ML le asignaría una mayor probabilidad a la primera secuencia que a la segunda. El ML ayuda al algoritmo de búsqueda a encontrar la secuencia de palabras que mejor describe la señal de voz. A partir de los etiquetados ortográficos de los Corpus de Prueba y de Entrenamiento se estimaron bigramas (pares de palabras) como el ML para el reconocedor. Para esto se usaron los módulos *HLStats* y *HBuild* del *HTK Toolkit*.
- **Diccionario Fonético:** Consiste en un documento de texto en donde se especifica la palabra y la secuencia de fonemas que la conforman. El diccionario para este caso consistió del vocabulario completo de ambos corpora (244 palabras) y las secuencias obtenidas mediante el *TranscribeMex*.
- **Algoritmo de Búsqueda:** El módulo *HVite* se utilizó para implementar el algoritmo de Viterbi [23, 24] para la función de reconocimiento de voz, la cual consiste en la búsqueda de la secuencia de palabras en el ML y el diccionario que mejor describen la señal de voz. Esto se expresa como:

$$\widehat{W} = \operatorname{argmax}_{W \in ML} P(O|W)P(W) \quad (1)$$

III. Resultados

La métrica de desempeño para el reconocedor de voz fue el Porcentaje de Precisión de Palabras (**Word Accuracy, WAcc**) [23], la cual se calcula como:

$$WAcc = \frac{N-D-S-I}{N} \quad (2)$$

Esta métrica se calcula mediante un alineamiento entre la transcripción correcta W (referencia) de la señal de voz, y la secuencia de palabras decodificada \widehat{W} (reconocida) por el sistema para la misma señal. De esta manera se tiene que, para la Ecuación 2, N es el número de palabras en W , D el número de palabras en W que no aparecen

en \hat{W} (eliminaciones), S el número de palabras en W que fueron confundidas con otras en \hat{W} (sustituciones), e I el número de palabras extra que aparecen en \hat{W} pero no tienen correspondencia con ninguna en W (inserciones).

Una métrica paralela al $WAcc$ que se utiliza ampliamente para medir el desempeño de estos sistemas es la Tasa de Error de Palabras (**Word Error Rate, WER**) [23], la cual se expresa como:

$$WER = 1 - WAcc \quad (3)$$

Para nuestro reconocedor de voz se utilizaron ambas métricas.

En la Tabla 3 se muestran los resultados de $WAcc$ y WER para cada sub-grupo y el total del Corpus de Entrenamiento del usuario GB. La precisión es mayor al 90% y el WER menor al 10% en todos los casos. Esto se debe a que los modelos acústicos están ajustados a las muestras de voz de este corpus.

Corpus de Entrenamiento	%WAcc	%WER
Narración	90.87	9.13
Lista de Palabras	99.02	0.98
Frases Balanceadas	94.44	5.56
Total	94.08	5.92

Tabla 3. Desempeño del Reconocedor de Voz con el Corpus de Entrenamiento.

Mediante el uso de un ML fonético (bigramas de fonemas) se pudo obtener la respuesta del sistema a nivel fonético. En este caso, al no haber restricciones del diccionario fonético, se obtiene una respuesta a bajo nivel, en la cual se pueden observar patrones de confusiones fonéticas (ver Figura 4).

Como se muestra en la Figura 4, la mayoría de los fonemas son reconocidos correctamente, lo cual explica los resultados de la Tabla 3. Es importante mencionar que no se puede esperar que este desempeño se mantenga cuando el sistema se pruebe con muestras de voz diferentes (por ejemplo, que no pertenezcan al Corpus de Entrenamiento). Sin embargo esta información si proporciona un panorama general acerca de: (1)

la precisión inicial del modelado, y (2) variaciones significativas en las pronunciaciones del usuario. Por ejemplo: se observan confusiones significativas del fonema /a/ con /e/, /o/; /d/ con /b/; /f/ con /k/; /s/ con /b/, /d/, /e/, /t/; y /_R/ con /e/, /r/, /s/. También hay una proporción significativa de eliminaciones (*Del*).

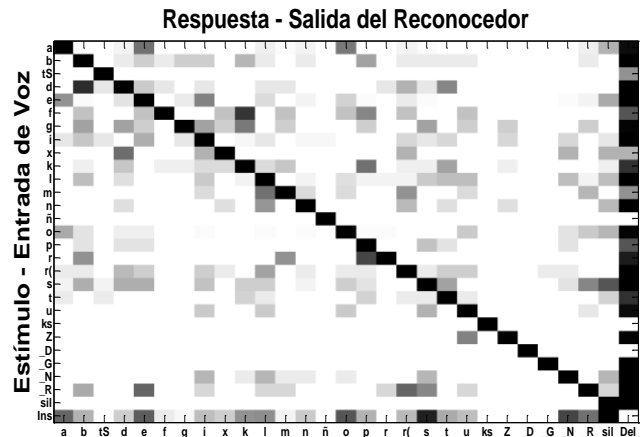


Fig. 4. Matriz de confusión fonética del Corpus de Entrenamiento.

Como se presenta en la Figura 5, estos patrones difieren de los observados en una matriz de confusión fonética de un usuario con disartria severa. En dicho caso, la capacidad del usuario para pronunciar sonidos diferentes está más limitada, por lo que varios sonidos que intenta pronunciar son confundidos por aquellos pocos que puede articular.

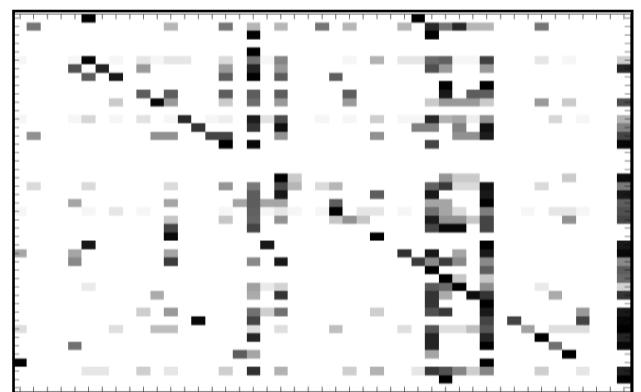


Fig. 5. Matriz de confusión fonética de un usuario norteamericano con disartria severa del Corpus Nemours [25].

Los patrones observados en la Figura 4 son coherentes con el diagnóstico presentado en la Tabla 1, por lo tanto esta información puede

Nº. Frases	N	D	S	I	%WAcc	%WER	Nº. Frases	N	D	S	I	%WAcc	%WER
1 COMER FRIJOLES CON HUEVO	4	0	0	0	100.0	0.0	11 UNA REBANADA DE PAPAYA	4	1	0	0	75.0	25.0
2 COMER GELATINA DE PIÑA	4	0	0	0	100.0	0.0	12 UN TACO DE HUEVO	4	0	0	0	100.0	0.0
3 COMER LA SOPA DE POLLO	5	1	0	0	80.0	20.0	13 UNA REBANADA DE SANDÍA	4	1	0	0	75.0	25.0
4 COMER UNA TAJADA DE PAPAYA	5	1	1	0	60.0	40.0	14 UNA TAZA DE TE CALIENTE	5	0	0	0	100.0	0.0
5 TOMAR UN VASO DE AGUA DE HORCHATA	7	0	0	0	100.0	0.0	15 UN VASO DE REFRESCO DE MANZANA	6	0	0	0	100.0	0.0
6 TOMAR UN VASO DE JUGO DE MANZANA	7	1	1	0	71.4	28.6	16 UN TACO DE HUEVO	4	0	0	0	100.0	0.0
7 TOMAR UN REFRESCO DE NARANJA	4	0	0	0	100.0	0.0	17 UN PAN CON HUEVOS Y FRIJOLES	6	1	0	0	83.3	16.7
8 TOMAR UN VASO DE LECHE FRIA	6	0	0	0	100.0	0.0	18 UN TACO DE POLLO	4	1	0	0	75.0	25.0
9 TOMAR UNA TAZA DE CHOCOLATE CALIENTE	6	0	1	0	83.3	16.7	19 UN TACO DE POLLO Y CAFÉ CON PAN	8	2	1	0	62.5	37.5
10 UNA TAJADA DE MELÓN	4	0	1	0	75.0	25.0	20 SALIR A TOMAR AIRE	4	0	0	0	100.0	0.0

Tabla 5. Frases de Prueba para el Reconocedor de Voz

ocuparse también para el diagnóstico del nivel de disartria del usuario.

En la Tabla 4 se muestran los resultados de desempeño del sistema al probarse con 20 frases del Corpus de Prueba del usuario GB en tiempo real. En la Tabla 5 se muestran en detalle los resultados para cada frase.

Corpus de Prueba	%WAcc	%WER
20 Frases	86.14	13.86

Tabla 4. Desempeño del Reconocedor de Voz con frases del Corpus de Prueba.

En el 50% de las pruebas se reconoció la frase de manera exacta ($%WAcc = 100.0$), y en el 65% la precisión fue mayor o igual al 80%. En el 20% la precisión fue de 75%, esto quiere decir que de 4 palabras una tenía error, sin embargo en su mayoría eran eliminaciones de alguna preposición (DE, CON) y no de una palabra significativa. En el 15% de las pruebas se obtuvo una precisión mayor al 60% pero menor que el 75%. Sin embargo hay que señalar que se trataban de frases de 5, 7, y 8 palabras que requieren de mayor esfuerzo para articularse. En la Figura 6 se muestra la matriz de confusión fonética respectiva, en donde se observa un patrón intermedio entre aquellos mostrados en las Figuras 4 y 5.

Esto es consistente con el uso de muestras de voz diferentes a aquellas muestras utilizadas para el entrenamiento del sistema. Aún así, la tasa de precisión es mayor que la de errores, aunque hay confusiones significativas. Por ejemplo: /b/ con /p/; /tS/ con /Z/; /d/ con /b/; /e/ con /a/, /o/; /f/ con /b/, /u/; /g/ con /a/; /l/ con /e/; /x/ con /b/, /k/; /r/

con /t/; /s/ con /t/; /u/ con /o/; /Z/ con /i/, /p/, /t/; /_R/ con /o/, /s/.

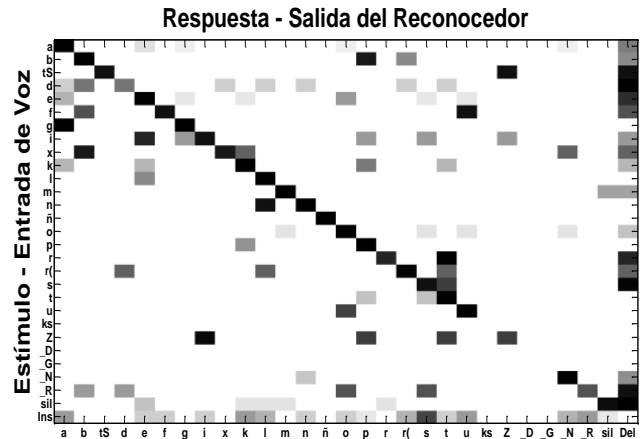


Fig. 6. Matriz de confusión fonética de las Frases de Prueba.

En ambas matrices de confusión (Figuras 4 y 6) se observan patrones que se mantienen entre ellas, por ejemplo, la confusión entre /d/ y /b/, /a/ y /e/, /s/ y /t/, lo cual puede indicar una deficiencia constante sobre la pronunciación de dichos fonemas.

Por lo tanto ambas matrices pueden ser utilizadas para medir un grado de discapacidad en el habla independientemente de la secuencia de palabras que genera el reconocedor.

IV. Discusión

Los resultados presentados en las Tablas 3 y 4 son consistentes con el diagnóstico de una persona con disartria leve o baja. Sin embargo un análisis al nivel fonético, como se muestra en las matrices de confusión de las Figuras 4 y 6, da un panorama de un nivel de disartria un poco más alto. Mientras que hay fonemas que se reconocen

correctamente, hay otros fonemas que se confunden más que otros o se eliminan, lo cual puede ser indicio de un problema particular en cuanto a la articulación de los mismos. El uso de este reconocedor y las matrices de confusión fonética se pueden utilizar para un diagnóstico automático del nivel de disartria de un usuario.

En cuanto a la metodología de modelado, los niveles de *WAcc* y *WER* son comparables con aquellos obtenidos en sistemas desarrollados para usuarios de habla no-española [7-16]. Por lo tanto el desarrollo de un sistema de reconocimiento y diagnóstico de voz para personas con disartria en México es posible.

V. Conclusiones y Trabajo Futuro

En este trabajo se presentó el desarrollo de un sistema de reconocimiento de voz para una persona con disartria, la cual es una discapacidad del habla frecuentemente asociada con otras discapacidades motoras. Con el propósito de que este sistema pueda ser utilizado para la comunicación del usuario, el sistema reconoce frases menores de 8 palabras, lo cual es más flexible que el uso de comandos. La tasa de precisión en el reconocimiento de frases fue del 94% en la etapa de entrenamiento, y del 86% en la etapa de prueba. Al utilizar el sistema para obtener una respuesta al nivel fonético se observaron patrones de errores que dan soporte al diagnóstico de disartria leve del usuario. Sin embargo también se observaron algunas deficiencias en las articulaciones de algunos fonemas, lo cual puede indicar un nivel de disartria más alto.

El desempeño del sistema es consistente con el de otros sistemas desarrollados en el extranjero para usuarios de habla no-española. El vocabulario utilizado de 244 palabras es mayor que el utilizado por dichos sistemas, lo cual lo hace más flexible para comunicación e interacción con otros sistemas.

El trabajo a futuro consiste en el desarrollo de una interface para este reconocedor, incluyendo la función de diagnóstico, al igual que una técnica que pueda mejorar el desempeño del

mismo. También se considera el trabajar con usuarios con diferentes tipos de disartria para evaluar el alcance de los HMM's como técnica de modelado acústico para estos usuarios, y estudiar alternativas como ANN's (Artificial Neural Networks) y SVM's (Support Vector Machines).

VI. Agradecimientos

Los autores desean agradecer a la Coordinadora, Doctora María de la Luz Gutiérrez, y a las Terapeutas, Rocío Bazán y Diana Pérez, del DIF municipal de la Cd. de Huajuapán de León por su apoyo para la realización de la presente investigación. De igual manera al Sr. GB y su familia por su participación activa en este proyecto.

VII. Referencias

- [1] Discapacidad en México. <http://cuentame.inegi.gob.mx/poblacion/discapacidad.aspx?tema=P>. Junio 2011.
- [2] J. Prater, R. Swift, Manual de Terapia de Voz. 1989. Salvat.
- [3] F. Darley, A. Aronson, J. Brown, Motor Speech Disorders. 1975. Saunders.
- [4] Desórdenes Motores del Habla y PROMPT (Parte II). http://www.espaciologopedico.com/articulos2.php?Id_articulo=1692. Junio 2011.
- [5] AbleData, Your Source for Assistive Technology Information: Communication. http://www.abledata.com/abledata.cfm?pag_eid=19327&top=10825&deep=2&trail=22&ksectionid=19327. Junio 2011.
- [6] Speech Enhancer. Voicewave Technology Inc. <http://www.speechenhancer.com/equipment.htm>. Junio 2011.
- [7] L.J. Ferrier, H.C. Shane, H.F. Ballard, T. Carpenter, A. Benoit, "Dysarthric speaker's intelligibility and speech characteristics in relation to computer speech recognition".

- Augmentative and Alternative Communication. 11. 1995. 165-175.
- [8] A. Kotler, N. Thomas-Stonell, "Effects of speech training on the accuracy of speech recognition for an individual with a speech impairment". *Augmentative and Alternative Communication*. 13. 1997. 71-80.
- [9] N.J. Manasse, K. Hux, J. Rankin-Erickson, "Speech recognition training for enhancing written language generation by a traumatic brain injury survivor". *Brain Injury*. 14. 2000. 1015-1034.
- [10] N.J. Manasse, K. Hux, J. Rankin-Erickson, E. Lauritzen, "Accuracy of three speech recognition systems: Case study of dysarthric speech". *Augmentative and Alternative Communication*. 16. 2000. 186-196.
- [11] P. Raghavendra, E. Rosengren, S. Hunnicutt, "An investigation of different degrees of dysarthric speech as input to speaker adaptive and speaker dependent recognition systems". *Augmentative and Alternative Communication*. 17. 2001. 265-275.
- [12] C. Goodenough-Trapagnier, M.J. Rosen, "Towards a method for computer interface design using speech recognition". *Proc. of Rehabilitation Engineering and Assistive Technology Society of North America (RESNA)*. 1991. 328-329.
- [13] G. Jayaram, K. Abdelhamied, "Experiments in dysarthric speech recognition using artificial neural networks". *Journal of Rehabilitation Research and Development*. 42. 1995. 162-169.
- [14] M.S. Hawley, P. Enderby, P. Green, S. Cunningham, S., Brownsell, J. Carmichael, M. Parker, A. Hatzis, P. O'Neill, R. Palmer, "A speech controlled environmental control system for people with severe dysarthria". *Medical Engineering & Physics*. 29. 2007. 586-593.
- [15] H. Strik, E. Sanders, M. Ruiter, L. Beijer, "Automatic recognition of dutch dysarthric speech: a pilot study". *ICSLP*. 2002. 661-664.
- [16] P. Green, J. Carmichael, A. Hatzis, S. Cunningham, R. Palmer, M. Parker, P. O'Neill, "An integrated toolkit deploying speech technology for computer based speech training with application to dysarthric speakers". *Proc. of European Conference on Speech Communication Technology*. 2003. 2213-2216.
- [17] J. Prence, *Human-Computer Interaction*. Addison-Wesley Longman. 1997. Gran Bretaña.
- [18] Corpus del Español. <http://www.corpusdelespanol.org>. Junio 2011.
- [19] L. A. Pineda, H. Castellanos, J. Cuétara, L. Galescu, J. Juárez, J. Llisterri, P. Pérez, L. Villaseñor, "The Corpus DIMEx100: Transcription and Evaluation". *Language Resources and Evaluation*. 44. 4. 2010. 347-370.
- [20] J. Cuétara. *Fonética de la ciudad de México. Aportaciones desde las tecnologías del habla*. Tesis de Maestría: Universidad Nacional Autónoma de México. 2004. México.
- [21] L.A. Pineda, L. Villaseñor, J. Cuétara, H. Castellanos, I. López, "DIMEx100: A new phonetic and speech corpus for Mexican Spanish". *Advances in Artificial Intelligence, Iberamia-2004*. 2004. 974-983.
- [22] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition". *Proc. of IEEE*. 37. 1989. 257-286.
- [23] S. Young, P. Woodland, *The HTK Book (for HTK Version 3.4)*. Cambridge University Engineering Department. 2006. Gran Bretaña.

- [24] D. Jurafsky, J.H. Martin, Speech and Language Processing. 2009. Pearson: Prentice Hall.
- [25] H.T. Bunnell, J.B. Polikoff, "The Nemours Database of Dysarthric Speech: A perceptual analysis". Proc. International Congress of Phonetic Sciences. 1. 1999. 783-786.

VIII. Autores

Dr. Santiago Omar Caballero Morales es Doctor en Ciencias Computacionales con Especialidad en Reconocimiento del Habla de la University of East Anglia (UEA) en Reino Unido. Actualmente es profesor-investigador de tiempo completo en la Universidad Tecnológica de la Mixteca (UTM) en Huajuapán de León, Oaxaca. Áreas de investigación: reconocimiento de habla, control automático por voz, traducción en tiempo real, interacción humano-computadora, optimización combinatoria.

C. Gladys Bonilla Enríquez actualmente se encuentra realizando su tesis profesional para obtener el título de Lic. en Ingeniería en Computación en la Universidad Tecnológica de la Mixteca (UTM). Sus áreas de interés son reconocimiento del habla e inteligencia artificial.

Entrenamiento de Perceptrón Unicapa para la Detección Aproximada de Bordes en una Imagen Digital.

***Andrea Santillana Fernández, **Carlos Delgado-Mata**

Universidad Panamericana, Campus Bonaterra

José María Escrivá de Balaguer No. 101 Aguascalientes, Aguascalientes 20290, México

Tel. 9106210, correo electrónico: *andreasantillana@gmail.com, **cdelgado@up.edu.mx

Resumen

Este artículo explica el desarrollo de un algoritmo que detecta bordes aproximados en una imagen digital. El algoritmo utiliza una Red Neuronal Artificial (ANN) que es entrenada por medio de nuestra implementación del algoritmo de aprendizaje por corrección de error. Adicionalmente, se presentan los resultados obtenidos por este algoritmo así como los obtenidos por conocidos métodos en el ámbito de detección de bordes, como lo son Canny, Sobel, Laplaciano, contornos activos.

Palabras Clave: Detección de bordes, Red Neuronal Artificial (ANN), Perceptrón unicapa, algoritmo de aprendizaje de corrección de error.

I. Introducción

Para que una criatura, biológica o artificial, desempeñe una interacción efectiva con su medio ambiente, necesita saber la localización de los objetos a su alrededor. La visión computacional provee un método primario para entender cómo tomar decisiones inteligentes respecto al medio ambiente, basado en entradas sensoriales [1].

El Procesamiento Digital de Imágenes mejora la calidad de imágenes tomadas bajo condiciones extremadamente desfavorables en varios aspectos: ajuste de brillo y contraste, detección de bordes, reducción de ruido, ajuste de enfoque, reducción de señales de movimiento, etc. La ventaja es que el Procesamiento de Imágenes permite un rango mucho más amplio de algoritmos aplicables a la información de entrada para reducir problemas como el ruido y la distorsión de la señal durante el procesamiento de ésta [2].

La detección de bordes ha sido una parte esencial de muchos sistemas de visión computacional, ya que ayuda a simplificar el análisis de las imágenes al reducir la cantidad de información que se debe procesar, y al mismo tiempo se mantiene la estructura básica de la imagen [3]. Se

usa principalmente cuando se necesita ubicar objetos amontonados o muy cercanos entre sí [1].

La detección de bordes es particularmente utilizada para el procesamiento de imágenes médicas, para reconocer y localizar sub-patrones dentro de una imagen [4].

Actualmente se conocen diversos algoritmos para la detección de bordes. El algoritmo de Canny es considerado como un método óptimo ya que toma en cuenta varios criterios. Uno evita la eliminación de bordes importantes y evita falsos bordes, otro establece que la distancia entre la posición real y la localizada del borde se debe minimizar, el último integra las respuestas múltiples correspondientes a un único borde.

El algoritmo de Sobel es también muy utilizado en este ámbito. Este algoritmo se basa en el gradiente de una imagen, que es el cambio de intensidad o color de una imagen; cuando el gradiente es muy grande, existe un borde [2][5].

El operador Laplaciano está basado en la segunda derivada, en ocasiones se utiliza para detección de bordes. Si se compara con el operador Sobel (basado en la primera derivada) el Laplaciano puede obtener mejores resultados en la

localización de bordes. Desafortunadamente, el Laplaciano es muy sensible al ruido [6].

En el presente artículo se propone la utilización de aprendizaje para encontrar bordes en imágenes digitales.

El objetivo de aprender a partir de ejemplos es encontrar una definición aproximada para una función desconocida $f(x)$ dados archivos de entrenamiento de la forma $(x_i, f(x_i))$; para casos en las que la función solo toma valores de cero o uno [7].

El perceptrón es un sistema inteligente integrado por componentes simples, y éste es inspirado en el modelo de un cerebro biológico. El perceptrón traza una línea que los divide un conjunto de datos de entrada en dos subconjuntos, para determinar la salida [8].

En este artículo se aborda el problema de detección de bordes, usando un perceptrón unicapa. El resto del artículo está ordenado de la siguiente manera: En la sección II se describe el desarrollo del algoritmo, mientras que en la sección III se presentan la pruebas realizadas y se describen los resultados obtenidos. Finalmente en la sección IV se presentan las conclusiones.

II. Desarrollo

1. Perceptrón unicapa (ver Fig. 1)

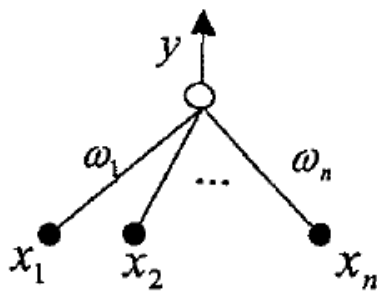


Fig. 1. Estructura de perceptrón unicapa [9].

El perceptrón es un tipo de Red Neuronal Artificial (ANN) de una capa con una sola neurona. Es una red en la que la neurona calcula la combinación lineal de su valor real o entradas

booleanas y lo manda por el umbral de una función de activación [8].

Las ecuaciones 1 y 2 expresan la salida de la neurona [9].

$$y = f(\sum_{i=1}^n \omega_i x_i - \theta) \tag{1}$$

$$f(\mu) = \begin{cases} 1 & \mu \geq 0 \\ 0 & \mu < 0 \end{cases} \tag{2}$$

En ocasiones se refiere al perceptrón como una Unidad Lógica de Umbral (TLU) dado que discrimina la información dependiendo si la suma es más grande que el valor de umbral (3) o la suma es menor que el valor de umbral (4). En las fórmulas 3 y 4 suponemos que el valor de umbral w_0 es el peso de una conexión adicional que se mantiene constante a $x_0 = 1$.

$$\sum_{i=1}^n w_i x_i > w_0 \tag{3}$$

$$\sum_{i=1}^n w_i x_i < w_0 \tag{4}$$

El perceptrón es estrictamente equivalente a un discriminante lineal (ver Fig. 2) y es usado frecuentemente como un dispositivo que decide a cuál de dos clases pertenece un patrón de entrada.

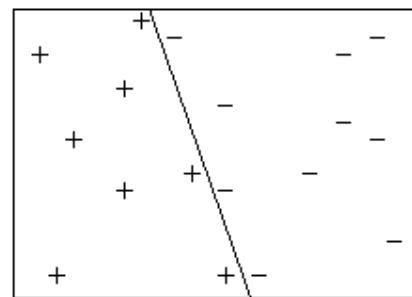


Fig. 2. Discriminante lineal [8].

La tarea principal, en una aplicación como la descrita en el presente artículo, es aprender los pesos de conexión entre los nodos. Existen varios algoritmos de aprendizaje que se pueden implementar en una neurona. El del algoritmo

presentado en este artículo se basa en el algoritmo de corrección de error.



Fig. 3. Diagrama de bloques del algoritmo.

1. 1 Algoritmo de aprendizaje por corrección de error

En el algoritmo de aprendizaje los parámetros a definir son el límite de iteraciones, el tamaño de paso de aprendizaje y el margen de error.

La neurona generalmente debe aprender los pesos de conexión basándose en los patrones de entrenamiento establecidos.

Se calcula el valor de salida de cada línea de entrada sumando el producto del peso de cada variable y la variable. El error actual es la diferencia entre este valor calculado y la salida establecida por el archivo de entrenamiento.

En cada iteración el algoritmo intenta minimizar el error actual.

El peso de cada variable se modifica en base al paso de aprendizaje, así como del error actual y la variable [10].

Este proceso se realiza para cada una de las líneas de entrada y se realiza mientras el error actual sea mayor que el margen de error o mientras el número de iteraciones actuales sea menor que el límite establecido.

2. Desarrollo del algoritmo

El algoritmo consta de tres pasos principales (Ver Fig. 3):

- Entrenamiento del perceptrón unicapa
- Evaluación de la imagen
- Generación de la imagen de salida a partir de los resultados del perceptrón unicapa.

2.1 Entrenamiento del perceptrón unicapa

Para entrenar el perceptrón, se utilizó una imagen y su resultado pre-calculado.

Las imágenes se convirtieron a escala de grises y después se normalizaron, para obtener dos matrices de unos y ceros.

Para transferirle esta información al perceptrón, se tomó cada pixel de la imagen de entrada y sus ocho pixeles adyacentes, como respuesta se envió el mismo pixel de la imagen de salida pre-calculada.

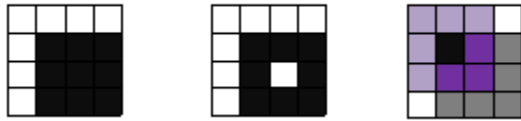


Fig. 4. Imagen de entrada, imagen de salida, adyacencias del pixel (2,2).

Si tomamos como ejemplo la Figura 4, para construir la línea de entrada del pixel en la posición (2,2), se envía primero el pixel a evaluar y los ocho pixeles adyacentes a éste empezando por el pixel (3,2) y recorriendo los demás en sentido contrario a las manecillas del reloj, de tal modo que la línea de entrada para el pixel (2,2) es:

1 1 0 0 0 0 0 1 1

La respuesta a esta entrada es el pixel (2,2) de la imagen de salida pre-calculada, que en este caso es uno (ver Fig. 5).

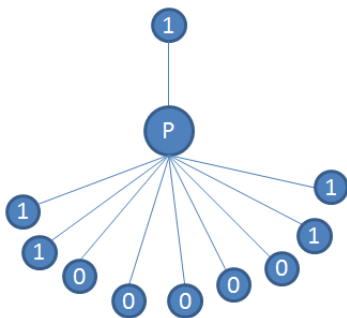


Fig. 5. Esquema de la línea de entrada de ejemplo con su salida en el perceptrón unicapa.

Una vez que juntamos la línea de entrada con su resultado, se envía a la neurona para que de ahí aprenda. Este procedimiento se realiza con cada uno de los pixeles de la imagen de entrada.

2.2 Evaluación de la imagen

Para mandar la imagen al perceptrón y obtener los resultados, el procedimiento fue igual que para la imagen de entrenamiento sin incluir la respuesta. Se tomó cada pixel y sus ocho pixeles adyacentes, en los casos en los que el pixel

formaba parte del límite de la imagen, se consideraron los pixeles adyacentes inexistentes como ceros.

Después de que se generó cada línea, se envió esta información al perceptrón unicapa.

El perceptrón ya solamente calculó la salida utilizando el último vector de pesos obtenido durante el entrenamiento.

2.3 Generación de la imagen de salida a partir de los resultados del perceptrón unicapa.

El perceptrón arrojó como salida una matriz de unos y ceros que después se multiplicó por 255 para obtener ya la matriz de la imagen resultante.

III. Pruebas y resultados

Para entrenar el perceptrón se utilizaron dos imágenes una con la imagen de entrada y otra con la imagen de salida. El par de imágenes de entrenamiento iniciales fueron de 10x10 pixeles, el tamaño de las imágenes de entrenamiento se fue incrementando hasta conseguir los mejores resultados. Los tamaños probados fueron 10x10, 20x20, 50x50, 100x100. Se encontró que no había diferencia perceptible entre los resultados del entrenamiento con 50x50 y 100x100 y por tanto se optó por usar las imágenes de entrenamiento de 50x50 pixeles para reducir el tiempo de ejecución. Los pares de imágenes de entrenamiento se muestran en las Figs. 6-9.

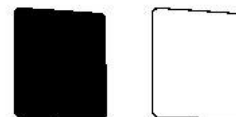


Fig. 6. Figuras de entrenamiento de 100x100 pixeles.

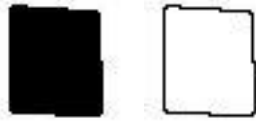


Fig. 7. Figuras de entrenamiento de 50x50 pixeles.



Fig. 8. Figuras de entrenamiento de 20x20 pixeles.



Fig. 9. Figuras de entrenamiento de 10x10 pixeles.

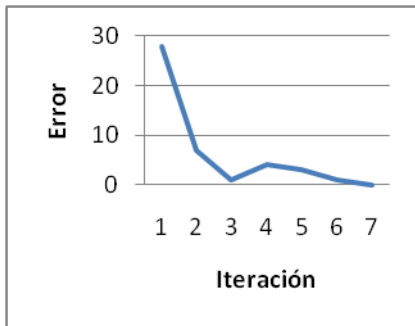


Fig. 10. Relación error vs iteración, usando una imagen de entrenamiento de 10x10 pixeles.

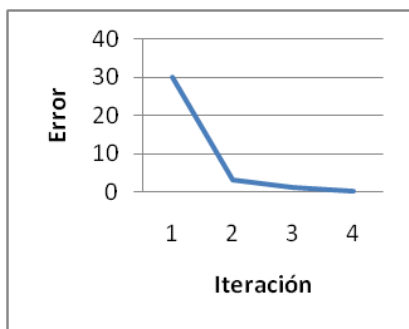


Fig. 11. Relación error vs iteración, usando una imagen de entrenamiento de 20x20 pixeles.

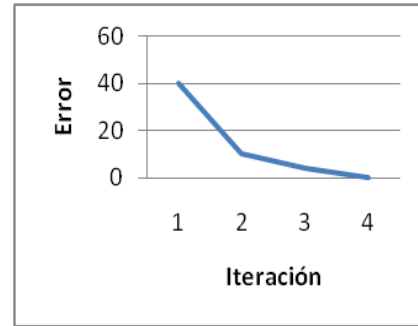


Fig. 12. Relación error vs iteración, usando una imagen de entrenamiento de 50x50 pixeles.

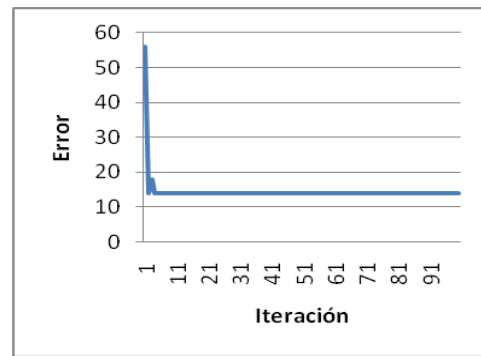


Fig. 13. Relación error vs iteración, usando una imagen de entrenamiento de 100x100 pixeles.

Como se puede apreciar en las Figs. 10-13 para los casos de entrenamiento de 10x10 pixeles, 20x20 pixeles y 50x50 pixeles el entrenamiento Termina con un error 0 con pocas iteraciones. Sin embargo, para el caso de 100x100 pixeles, no se llega a obtener un error 0 y el entrenamiento se termina al realizar el máximo número de iteraciones. En la Fig. 14 se puede apreciar la imagen a evaluar, Lena (204x204 pixeles); la más común para realizar pruebas en procesamiento digital de imágenes. La Fig. 15 muestra la imagen convertida a escala de grises, mientras que la Fig. 16 muestra el resultado que se obtienen al utilizar el algoritmo Canny. La Fig. 17 muestra el resultado que se obtiene utilizando el algoritmo de Sobel. La Fig. 18 muestra el resultado del operador Laplaciano Las Figs. 19-21 muestran los resultados obtenidos por cada uno de los entrenamientos. Como se puede apreciar en el resultado obtenido, para el caso de 50x50 pixeles es similar al obtenido con Canny.



Fig. 14. Imagen a evaluar.



Fig. 17. Resultado utilizando algoritmo Sobel [2].



Fig. 15. Imagen en escala de grises.



Fig. 18. Resultado utilizando el operador Laplaciano.



Fig. 16. Resultado utilizando algoritmo Canny [3].

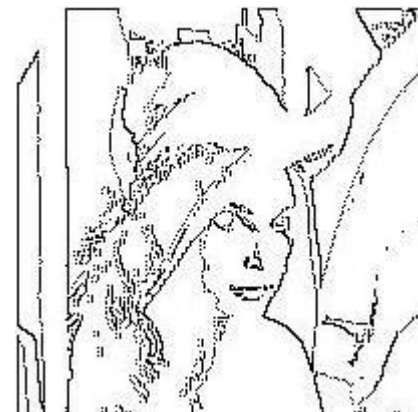


Fig. 19. Resultado del algoritmo con la imagen de entrenamiento de 10x10 pixeles.

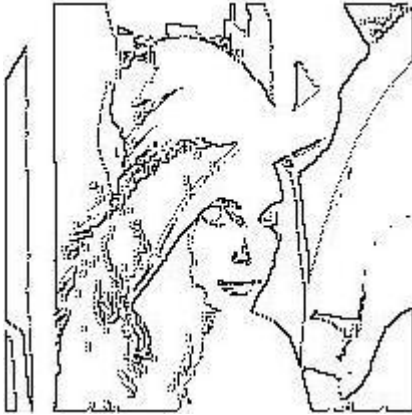


Fig. 20. Resultado del algoritmo con la imagen de entrenamiento de 20x20 píxeles.

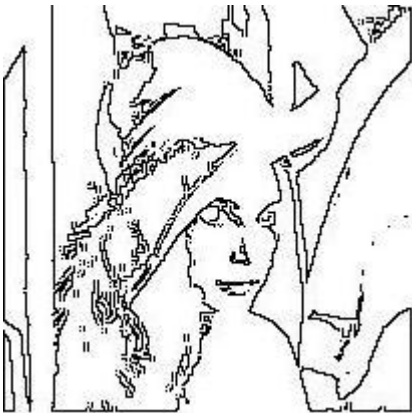


Fig. 21. Resultado del algoritmo con la imagen de entrenamiento de 50x50 píxeles.

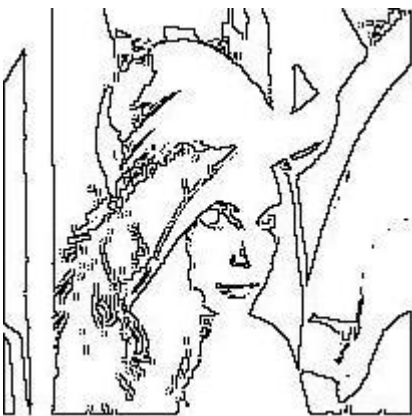


Fig. 22. Resultado del algoritmo con la imagen de entrenamiento de 100x100 píxeles.

Después de determinar que el mejor tamaño de las imágenes de entrenamiento es de 50x50 píxeles, se utilizó este entrenamiento para probar el algoritmo con imágenes de distinta resolución, una fotografía del futbolista Ryan Giggs (416 x 300 píxeles) (ver Fig. 23 y 24) y un paisaje (1600x1200 píxeles) (ver Fig. 25 y 26).



Fig. 23. Futbolista Ryan Giggs.

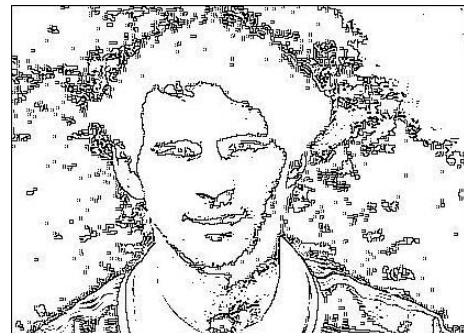


Fig. 24. Imagen de salida usando este algoritmo.



Fig. 25. Paisaje del Puente de Londres.

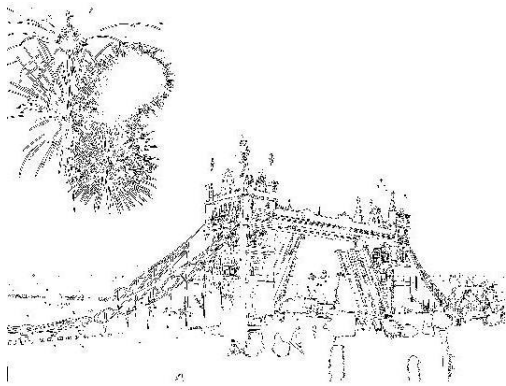


Fig 26. Bordes de la imagen del Puente de Londres.

IV. Conclusiones

Con este algoritmo descubrimos una forma muy sencilla de detectar los bordes aproximados en una imagen. La complejidad computacional de este algoritmo es $O(2n)$ mientras que la complejidad del Canny es $O(4n)$. Siendo computacionalmente más barato este algoritmo, sin embargo, arroja resultados muy similares a los del Canny, creando así una alternativa muy atractiva para procesos de visión computacional en los que se requiere minimizar el tiempo computacional de tareas recurrentes como lo es la detección de bordes.

Como trabajo futuro se planea considerar otros tipos de generación de entradas de cada pixel, buscando no eliminar bordes importantes y también evitar falsos bordes.

V. Referencias

- [1] W.E.L. Grimson, J.L. Mundy, "Computer vision applications". Commun, ACM 37, 3. 44-51. Marzo 1994.
- [2] O.R. Vincent, O. Folorunso. "A Descriptive Algorithm for Sobel Image Edge Detection". Proceedings of Informing Science & IT Education Conference (InSITE) 2009.
- [3] J. Canny. "A Computational Approach to Edge Detection". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 6. Noviembre 1986.
- [4] I. Aizenberg, N. Aizenberg, J. Hiltner, C. Moraga, E. Meyer zu Bexten. "Cellular neural networks and computational intelligence in medical image processing". Image and Vision Computing 19 (2001) 177-183.
- [5] M. Nixon, A. Aguado. "Feature extraction and Image Processing". 2ª edición. 2008. Elsevier Science. Londres. 123-126.
- [6] Xin Wang; , "Laplacian Operator-Based Edge Detectors," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.29, no.5, pp.886-890, May 2007 doi: 10.1109/TPAMI.2007.1027.
- [7] T.G. Dietterich, G. Bakiri. "Error-Correcting Output Codes: A General Method for Improving Multiclass Inductive Learning Programs". Oregon State University. Proceeding of National Conference on Artificial Intelligence AAAI-9. 1991.
- [8] N. Nikolaev. Single-Layer Perceptrons. "311: Neural Networks" [En línea] <http://homepages.gold.ac.uk/nikolaev/311perc.htm>
- [9] Z. Yanling. D. Bimin, W. Zhanrong. "Analysis and Study of Perceptron to Solve XOR Problem". 2nd International Workshop on Autonomus Decentralized System. Noviembre 2002.
- [10] A.K. Jain, J. Mao, K.M. Mohiuddin, "Artificial neuron networks: a tutorial". Computer , vol.29, no.3, pp.31-44, Marzo 1996.

Reconocimiento Automático del Habla para Pronunciación de la Lengua Mixteca

***Santiago Omar Caballero Morales, **Edgar De Los Santos Ramírez**

Universidad Tecnológica de la Mixteca, Carretera a Acatlima Km 2.5, Huajuapán de León, Oaxaca, 69000

División de Posgrado, +52 (953) 532 02 14, Ext. 300

*scaballero@mixteco.utm.mx

**edgarinteractivo@hotmail.com

Resumen

El Mixteco es una de las principales lenguas indígenas de México. La modernización, migración, discriminación, y poca difusión de la cultura, han disminuido la presencia de esta lengua en los últimos años. Como manera de contribuir a la preservación de esta lengua, se propone el uso del Reconocimiento Automático del Habla (RAH) ya que, mediante esta tecnología, se puede favorecer a que personas que deseen aprender la lengua puedan practicar su pronunciación. El Mixteco tiene la característica de ser una lengua tonal, por lo que el significado de una palabra depende de su entonación. Esto representa un reto para el desarrollo de un sistema RAH eficiente con baja tasa de confusiones. En el presente artículo se muestra el diseño de un sistema RAH para entrenamiento de la pronunciación de la variante del Mixteco de San Juan Diquiyú. Mediante la definición fonética de los diferentes tonos, y el uso de Modelos Ocultos de Markov, se desarrolló un sistema con una tasa de precisión en el reconocimiento de 94.84% para un usuario nativo. Al implementarse una técnica estándar de adaptación de usuario se obtuvo una tasa del 88.61% para un usuario no nativo. Esto ofrece un panorama de aplicación en donde el sistema propuesto pueda ser usado en tareas más avanzadas de aprendizaje de la lengua.

Palabras Clave: lenguas indígenas, reconocimiento del habla, interacción humano-computadora.

I. Introducción

En México hay alrededor de 89 lenguas indígenas las cuales son habladas por cerca de 6.6 millones de personas. Sin embargo, la mayoría de estas personas se encuentran en áreas alejadas de las ciudades, con pocos servicios de transporte y comunicación. Habitantes de la Mixteca de Oaxaca han mencionado que la urbanización, migración en busca de trabajo en las ciudades, y el rechazo social han afectado la preservación de su cultura. A los niños de estas comunidades se les enseña español en las escuelas, y como la mayoría del material didáctico se encuentra en este idioma, esto ocasiona que se vaya perdiendo la lengua desde temprana edad. De igual manera, hay poco material didáctico para aprender una lengua indígena, el cual es en su mayoría escrito,

lo cual limita la capacidad de alguien para aprender o practicar la pronunciación.

La población de hablantes nativos del Mixteco ha disminuido en los últimos años. En 1990 el 7.5% de la población total hablaba alguna lengua indígena, sin embargo en el 2005 esa cifra se redujo al 6.6% [1]. El Mixteco actualmente cuenta con alrededor de 480,000 hablantes, pero dentro de éste hay variantes que cuentan con menos de 10 (por ejemplo, el Mixteco de la Región Mazateca tiene 6 hablantes conocidos [1]).

Para ayudar a preservar un idioma por medio de la enseñanza y/o práctica es necesario considerar características como son las reglas gramaticales, la expresión escrita de la voz, su articulación / pronunciación, etc. Tecnología moderna en

Reconocimiento Automático del Habla (RAH) e Interfaces Humano-Computadora (IHC) ha sido utilizada recientemente para esta tarea, habiendo herramientas comerciales para el aprendizaje de idiomas como el inglés [2, 3], francés o español [4]. Sin embargo no se conocen herramientas similares para aprender o practicar una lengua indígena (ya sea en México o en el extranjero).

Esta fue la motivación para desarrollar un sistema RAH que pudiera reconocer la Lengua Mixteca, lo cual es un paso importante para una interface de aplicación ya sea de aprendizaje, práctica, o comunicación. Los avances logrados se presentan de la siguiente manera: en la Sección II se muestran los detalles de la fonología de la variante del Mixteco que se tomó como referencia, al igual que los detalles de la metodología para definir el vocabulario y la toma de muestras de voz; en la Sección III se presentan los detalles de la construcción del sistema RAH que incluyen la modelación acústica de fonemas mediante entrenamiento supervisado; en la Sección IV se muestran los resultados obtenidos los cuales son discutidos en la Sección V; finalmente en la Sección VI se presentan nuestras conclusiones y trabajo a futuro.

II. Lengua Mixteca

II.A. Características Fonéticas

La lengua Mixteca, también llamada “Tu’uñ Savi” (Palabra de la Lluvia) [5], es una de las lenguas indígenas más importantes de México. Actualmente sus diferentes variantes tienen presencia en los estados de Puebla, Guerrero y Oaxaca, aunque debido a la migración hay comunidades también en Sinaloa, y Jalisco. Hay muchas variantes de esta lengua, y el número varía desde 30 [6] hasta 81 [7].

Esta lengua es del tipo tonal, por lo que el significado de una palabra depende de su entonación. Debido a la dispersión entre hablantes hay diferencias en las entonaciones y pronunciaciones del Mixteco, por lo que no se puede hablar de una sola lengua.

Es por esto que cada variante se identifica con una región o pueblo, por ejemplo, el Mixteco de Tezoatlán [8], el Mixteco de Yosondúa [9], o el Mixteco de Xochapa [10], pudiendo haber grandes diferencias entre sus vocabularios y/o definiciones.

En general el Mixteco tiene tres tonos característicos: alto, medio, y bajo. En la Tabla 1 se muestran ejemplos de palabras en donde con solo cambiar el tono se modifica su respectivo significado. Para la representación de los tonos se han encontrado diferentes estilos. En algunos casos el tono medio no se representa, y el tono bajo se representa como una línea horizontal arriba de la vocal [11], o con una línea horizontal debajo de la vocal [5]. Usualmente el tono alto se representa con el diacrítico (^).

Palabra	Significado	Palabra	Significado
ñoó	noche	yukú	quién
ñoo	pueblo	yuku	montaña
ño <u>o</u>	palma	y <u>u</u> ku	hoja

Tabla 1. Ejemplos de palabras en Mixteco.

Para tratar de mantener una consistencia con las representaciones utilizadas en la literatura, se utilizó el diacrítico (^) para representar el tono bajo, y (ˆ) para el tono alto, el tono medio se dejó sin marcar. A partir de los fonemas identificados en [5, 8, 9, 10, 11], e integrando los diferentes tipos de tonos, se definió la lista de fonemas mostrada en la Tabla 2.

La fonética del Mixteco tiene una forma de definirse diferente a la del español. Por ejemplo, de la Tabla 2:

- el fonema /dj/ representa un sonido como el de la z del español de España, que es un poco más fuerte que el de la s. En el español de México, ambos sonidos (z y s) se representan con el mismo fonema /s/ [12];
- los fonemas /sh/ y /ch/, los cuales se pronuncian en español como *sh* en la palabra “*Xicoténcatl*”, y *ch* en “*chicle*”, representan a las consonantes del Mixteco **X** y **CH** respectivamente. En la fonética del español del

centro de México, /x/ (*j*) representa a la consonante *X*, en tanto que /tS/ (*ch*) representa

Para esto se estableció contacto con un profesor nativo del lenguaje que enseñaba la variante de

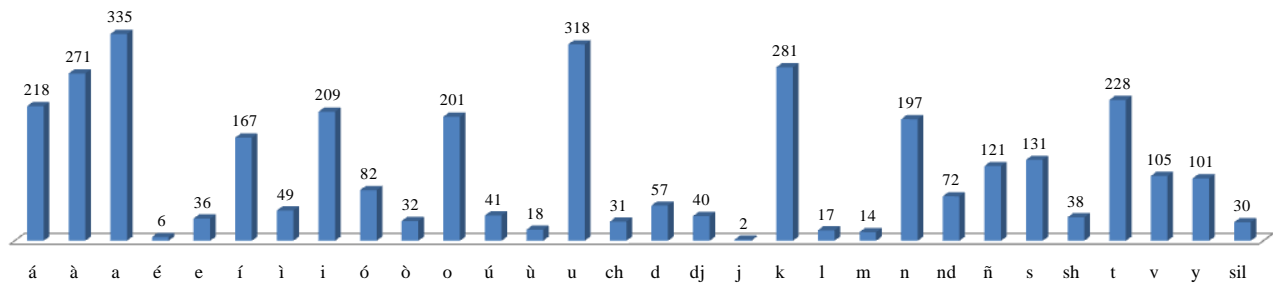


Figura 1. Frecuencia de fonemas en el Corpus de Entrenamiento Mixteco.

a *CH*, /sh/ no tiene representación;

- existen pequeñas pausas, un cierre glotal entre vocales de una misma palabra, la cual se representa por (‘) como en “tu’un” o “ndá’a”;
- la consonante *N* se pronuncia como *n* (/n/), igual que en el español cuando se presenta antes de una vocal, pero es muda si se escribe después de ella.

Estas consideraciones fueron muy importantes para el desarrollo del diccionario fonético del RAH.

Mixteco de su región de origen, San Juan Diquiyú, Oaxaca. En su dinámica de enseñanza utilizaba 15 lecciones que consistían de diferentes relatos tradicionales, desde básicos hasta más complejos.

Se seleccionaron 7 lecciones, las cuales fueron leídas dos veces por el profesor en una cabina de grabación. Este material al ser leído proporcionó una colección de muestras de voz que formaron el Corpus Oral de Entrenamiento para el sistema RAH. El total de palabras diferentes en estos relatos (y por ende del corpus) fue de 192 palabras.

Cada muestra de voz, grabada en formato WAV con una velocidad de muestreo de sonido de 44kHz monoaural, se etiquetó manualmente a niveles ortográfico y fonético usando el software *WaveSurfer*. La transcripción fonética es importante para la creación del diccionario fonético del RAH (ver Sección III).

De esta manera se pudo estimar la frecuencia (número de ocurrencias) de los 30 fonemas identificados en la Tabla 2 del corpus. La gráfica de frecuencias de cada fonema se presenta en la Figura 1. En base a esta distribución se consideró que el corpus estaba fonéticamente balanceado, y por lo tanto, había suficientes muestras de voz para el entrenamiento supervisado del RAH.

No.	Fonema	No.	Fonema	No.	Fonema
1	/á/	11	/o/	21	/m/
2	/à/	12	/ú/	22	/n/
3	/a/	13	/ù/	23	/nd/
4	/é/	14	/u/	24	/ñ/
5	/e/	15	/ch/	25	/s/
6	/í/	16	/d/	26	/sh/
7	/ì/	17	/dj/	27	/t/
8	/i/	18	/j/	28	/v/
9	/ó/	19	/k/	29	/y/
10	/ò/	20	/l/	30	/sil/

Tabla 2. Repertorio de Fonemas Mixtecos.

II.B. Vocabulario

Ya que el propósito de este sistema es que se pueda utilizar para prácticas de pronunciación, se optó por trabajar con relatos en Mixteco como fuente de vocabulario.

III. Construcción del Sistema RAH

Los elementos de un RAH se presentan en la Figura 2, siendo la herramienta *HTK Toolkit* [13]

la utilizada para la implementación de los mismos. HTK es un software de libre acceso desarrollado por la Universidad de Cambridge para la construcción de Modelos Ocultos de Markov para aplicaciones de reconocimiento de patrones, particularmente de la voz.

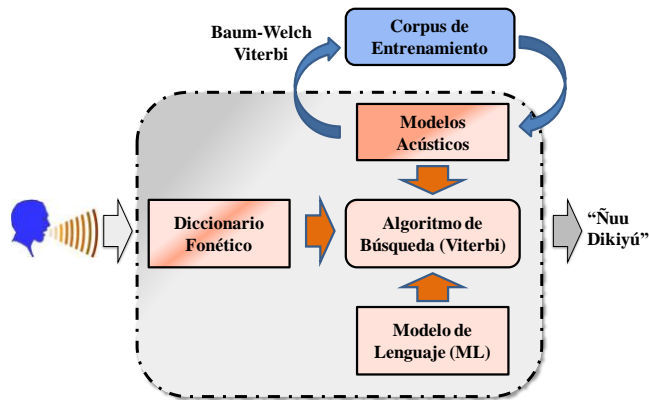


Figura 2. Estructura de un Sistema de Reconocimiento Automático del Habla.

Cada uno de los elementos de la Figura 2 se construyó de la siguiente manera:

- **Modelos Acústicos:** Modelos Ocultos de Markov (Hidden Markov Models, HMM's) [14, 15] fueron utilizados para hacer el modelado acústico de cada fonema existente en el repertorio del Corpus de Entrenamiento. La estructura de los HMM's fue una estándar de 3 estados con secuencia de izquierda a derecha, con 10 mixturas gaussianas por estado [13]. La codificación en MFCC's (Mel Frequency Cepstral Coefficients) fue utilizada para los archivos de voz, para lo cual se usaron 12 coeficientes cepstrales con coeficientes adicionales de energía, delta, y aceleración [13, 15].

Para el entrenamiento supervisado se usaron los archivos de sonido codificados y las transcripciones fonéticas de éstos, identificando así los segmentos de la voz que corresponden a cada fonema. El entrenamiento se realizó mediante los algoritmos de Baum-Welch y Viterbi.

El módulo de HTK, *HCopy*, se utilizó para la codificación de los archivos de voz (formato WAV a MFCC), en tanto que *HInit*, *HRest*, y *HERest* se utilizaron para la inicialización y

re-estimación supervisada de los modelos acústicos de los fonemas del RAH (esto es, el entrenamiento del sistema). Estos módulos estiman los parámetros de los modelos mediante realineamientos temporales de las muestras de voz del Corpus de Entrenamiento y las respectivas transcripciones a nivel fonético.

- **Diccionario Fonético:** este elemento presenta todas las palabras en el vocabulario y la secuencia de fonemas que constituye a cada una. Por ejemplo, la palabra “ká'ano” (“grande”) está formada por los fonemas /k/ /ah/ /am/ /n/ /om/. Esta secuencia se determinó de manera empírica mediante análisis perceptual, respetando las reglas definidas en la Sección II.A.
- **Modelo de Lenguaje (ML):** Su función es la de ayudar al proceso de reconocimiento mediante la restricción de búsqueda de palabras a secuencias que son permisibles dadas unas probabilidades o reglas gramaticales. Estas reglas o restricciones son el Modelo de Lenguaje, el cual consistió de bigramas (pares de palabras) estimados a partir del Corpus de Entrenamiento. Los módulos *HLStats* y *HBuild* se ocuparon para (1) la estimación de las probabilidades de ocurrencia de cada una de las secuencias de dos palabras encontradas en el Corpus de Entrenamiento, y (2) la construcción de la red de secuencias ponderadas con dichas probabilidades, lo que representa las secuencias permisibles de palabras.
- **Algoritmo de Búsqueda:** Este es el algoritmo de Viterbi para la decodificación de observaciones acústicas en palabras escritas. Este algoritmo compara las observaciones generadas por los modelos acústicos (HMM's) con las observaciones acústicas recibidas (voz). La secuencia de modelos que se van asociando a dichas observaciones se restringe a secuencias válidas dadas por el diccionario fonético (esto es, secuencias que van a formar palabras). Por último se integran estas probabilidades con las del ML para generar secuencias de palabras válidas.

III.A. Adaptación de Usuario

Como se presentó en la Sección II.B., el Corpus de Entrenamiento para el sistema RAH consiste de las muestras de voz de un hablante nativo del

de fonemas de la Figura 3, la cual tiene una correlación con la del corpus (Figura 1) de 0.69.

La adaptación MLLR calcula transformaciones que actúan sobre la media y la varianza de las mixturas gaussianas de los modelos acústicos del sistema para adaptarlos a las características

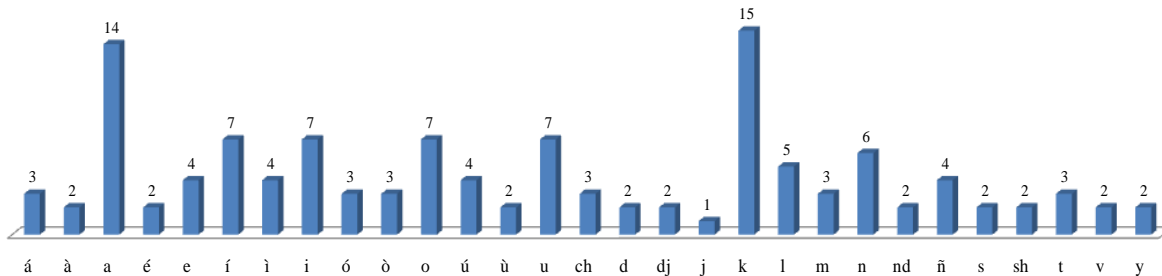


Figura 3. Frecuencia de fonemas en las palabras para adaptación de usuario.

lenguaje. Este sistema, del tipo Dependiente del Usuario (DU), sólo funcionará para el hablante cuya voz se utilizó para su entrenamiento. Para fines prácticos de uso por otros usuarios esto no es conveniente.

Sistemas comerciales son entrenados con las muestras de voz de cientos o miles de personas, lo cual da como resultado sistemas Independientes del Usuario (IU). Cuando un usuario nuevo desea utilizar este sistema, es común que tenga que leer algunas frases o palabras para que el sistema adapte sus modelos acústicos a los patrones de voz del nuevo usuario. En el caso de sistemas IU, el sistema es lo bastante robusto y técnicas como MAP o MLLR [13, 15] adaptan de manera eficiente el sistema, obteniendo tasas de reconocimiento hasta del 95%.

En el caso del Mixteco es más difícil el desarrollo de un sistema IU dadas las variantes y la disponibilidad de los usuarios. Es por esto que se estudió la posibilidad de aplicar adaptación para el caso de este sistema DU.

La técnica de adaptación fue MLLR (Maximum Likelihood Linear Regression), y para esto se hizo una selección de palabras en donde se tuvieran suficientes ocurrencias de todos los fonemas del repertorio del corpus. Estas palabras, mostradas en la Tabla 3, muestran la distribución

acústicas del usuario [13]. Estas transformaciones aumentan conforme haya más muestras de voz para la adaptación. La adaptación se realiza en dos pasos:

- Transformación global: se genera una transformación general, la misma para todas las clases (clasificaciones de fonemas). Esta sirve como entrada para generar transformaciones más específicas (transformación dinámica).
- Transformación dinámica: se generan transformaciones más específicas para cada clase. Las clases se establecieron mediante un árbol de regresión, el cual sirve para definir los fonemas (o grupos de fonemas) que comparten la misma transformación [13]. Este árbol de regresión, con 32 nodos terminales, utilizó las estadísticas mostradas en la Figura 1 para la asignación y estimación de transformaciones.

El usuario de prueba para la adaptación fue un estudiante del género masculino de 28 años de edad. A este usuario, antes de utilizar el RAH, se le dio una plática acerca de la lengua, una revisión de las lecciones, e información acerca de la pronunciación de las palabras (lecciones y palabras de adaptación). Esta sesión informativa duró aproximadamente 6 horas, las cuales se distribuyeron en tres días.

Una vez que el usuario terminó la sesión, se procedió a grabar las palabras de la Tabla 3 y tres relatos de las lecciones: uno básico, uno intermedio, y uno avanzado. La adaptación MLLR se implementó mediante el módulo *HERest*, el cual generó 3 transformaciones para el RAH base.

No.	Palabra	Fonemas
1	ÀNE'ECHOOS	à n e e c h o o s
2	ÁTOKÓ	á t o k ó
3	DIKIYÚ	d i k i y ú
4	DJAMA	d j a m a
5	DJÀVÌ	z à v ì
6	CHÁNÍ	ch á n í
7	CHI	ch i
8	ÍDJONA	í z o n a
9	ÍÑÒ	í ñ ò
10	KAMA	k a m a
11	KÌVÌ	k ì v ì
12	KOKUMI	k o k u m i
13	KÒÒÍÚN	k ò í ú
14	KOÚNI	k o ú n i
15	KUALÍ	k u a l í
16	KÛTAKU	k ù t a k u
17	KUTÓ	k u t ó
18	LAA	l a a
19	LULI	l u l i
20	NDAKONÓ	nd a k o n ó
21	NDEEYÉ	nd e y é
22	NÍKÉE	n í k é e
23	ÑA	ñ a
24	ÑUU	ñ u u
25	SÁXI	s á s h i
26	ÙXÌ	ù s h ì

Tabla 3. Palabras para adaptación de usuario.

IV. Resultados

La medida de desempeño para el reconocedor fue la de *Precisión de Palabras (Word Recognition Accuracy WAcc)*, la cual es un análogo de la *Tasa*

de Error de Palabras (Word Error Rate, WER), las cuales se calculan de la siguiente manera:

$$WAcc = \frac{N-D-S-I}{N} \quad (1)$$

$$WER = 1 - WAcc \quad (2)$$

en donde N es el número de palabras en las transcripciones correctas de la señal de voz, D el número de palabras eliminadas en la salida del reconocedor, I el número de palabras insertadas (adicionales) en la salida del reconocedor, y S el número de palabras en la transcripción correcta sustituidas por otras en la salida del reconocedor. Estos valores se estiman mediante un alineamiento que compara las transcripciones de referencia (correctas) con la salida del reconocedor. El módulo *HResults* se utilizó para esta función.

El sistema RAH fue probado primero con el Corpus de Entrenamiento, es decir, con el usuario nativo del Mixteco. Los resultados se presentan en la Tabla 4.

N	D	S	I	WAcc	WER
911	0	18	29	94.84	5.16

Tabla 4. Tasa de errores y precisión del RAH con el Corpus de Entrenamiento.

Una matriz de confusión fonética fue estimada a partir de la respuesta del RAH con este corpus, la cual se muestra en la Figura 4. Esta respuesta del RAH se obtuvo mediante la eliminación del diccionario fonético (el cual forma palabras) y del ML de palabras.

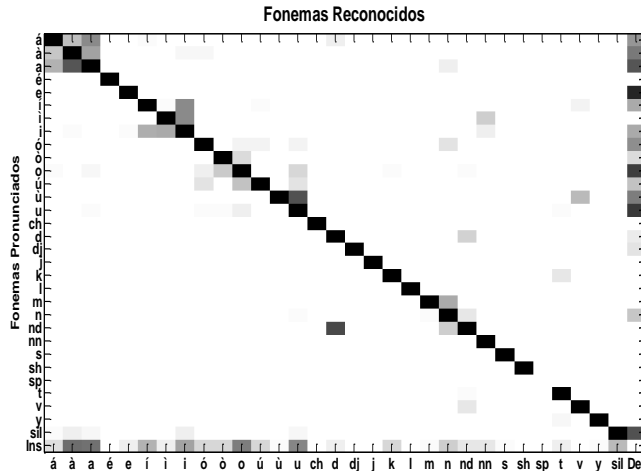


Figura 4. Matriz de confusión fonética del RAH con el Corpus de Entrenamiento Mixteco.

La Figura 4 se puede considerar como la respuesta a bajo nivel del reconocedor. Como se puede observar, la mayoría de los fonemas son reconocidos correctamente, con sólo algunas confusiones con otros fonemas: por ejemplo, entre las vocales /á/, /à/, /a/; /í/, /î/, /i/; y una significativa entre /nd/, /d/. La tasa de precisión para el reconocimiento de fonemas (Figura 4) se presenta en la Tabla 5.

N	D	S	I	WAcc	WER
3846	340	338	144	78.63	21.37

Tabla 5. Tasa de errores y precisión del RAH con el Corpus de Entrenamiento a nivel fonético.

Tanto en la matriz de confusión de la Figura 4 como en la Tabla 5 se observa que las tasas de eliminaciones y de sustituciones son aproximadamente del 10% del total de palabras, en tanto que las inserciones representan el 5%. La mayoría de los fonemas que son eliminados son vocales. Una tasa del 78% para la precisión es normal dado que no hubo restricción del diccionario fonético (el cual da una precisión del 94.84% para palabras).

Estos resultados muestran que el modelado se ha realizado de manera satisfactoria con el Corpus de Entrenamiento. Sin embargo esto es normal en la mayoría de los casos a menos que hubiera

muchas variaciones en la pronunciación del usuario.

En la Tabla 6 se muestran los resultados para el usuario de prueba una vez que se aplicó la adaptación con tres diferentes lecciones.

Lección	N	D	S	I	WAcc	WER
1	53	0	2	0	96.23	3.77
3	53	1	5	3	83.02	16.98
6	96	0	10	2	87.50	12.50
Total	202	1	17	5	88.61	11.39

Tabla 6. Tasa de errores y precisión del RAH con frases del usuario no-nativo

V. Discusión

Los resultados presentados en la Sección IV muestran que, para un vocabulario de 192 palabras, es posible desarrollar un sistema RAH con una precisión mayor al 94%. Se observaron patrones de confusiones entre las vocales con diferentes tonos, sin embargo las tasas son bajas. Por ejemplo, de 654 ocurrencias de la vocal /á/, 483 fueron reconocidas correctamente (74%), 36 fueron confundidas con /à/ (6%), y 66 con /a/ (10%), mientras que 57 ocurrencias fueron eliminadas (9%). Para la vocal /à/ se obtuvieron estadísticas similares: 73% correctas, 5% confusión con /á/, 8% confusión con /a/, 10% eliminadas. Las estadísticas para los demás fonemas se muestran en la Tabla 7.

Corroborando la información de la Figura 4, las consonantes son las que menos confusiones tienen y mayor porcentaje de reconocimiento correcto.

Por limitantes de tiempo sólo se grabó una vez el material de adaptación (Tabla 3) con el usuario no-nativo. Sin embargo, con sólo 26 palabras, se obtuvo el 88.61% de precisión, lo cual demuestra que MLLR puede ser aplicado a un sistema DU y mejorar el desempeño de un RAH. Con mayor material se podría incrementar esta tasa de precisión hasta un nivel cercano al del usuario nativo.

Es importante mencionar que existe una amplia variedad de técnicas en el ámbito del

reconocimiento de voz. En el presente trabajo se utilizaron HMM's dada la disponibilidad de la herramienta y el material didáctico. Sin embargo en el trabajo a futuro se contempla el uso de Redes Neuronales Artificiales (Artificial Neural Networks, ANN's), y SVM's (Support Vector Machines), para el modelado acústico y la construcción del sistema RAH.

Fonema	Correctas	S	D
/á/	0.74	0.17	0.09
/à/	0.73	0.15	0.11
/a/	0.60	0.24	0.16
/é/	1.00	0.00	0.00
/e/	0.75	0.00	0.25
/í/	0.80	0.13	0.07
/ì/	0.86	0.14	0.00
/i/	0.77	0.16	0.07
/ó/	0.84	0.06	0.10
/ò/	0.94	0.03	0.03
/o/	0.68	0.12	0.20
/ú/	0.85	0.10	0.05
/ù/	0.67	0.22	0.11
/u/	0.74	0.06	0.20
/ch/	1.00	0.00	0.00
/d/	0.95	0.04	0.02
/dj/	0.98	0.00	0.03
/j/	1.00	0.00	0.00
/k/	0.98	0.02	0.00
/l/	1.00	0.00	0.00
/m/	0.93	0.07	0.00
/n/	0.92	0.03	0.05
/nd/	0.78	0.22	0.00
/ñ/	1.00	0.00	0.00
/s/	1.00	0.00	0.00
/sh/	1.00	0.00	0.00
/t/	1.00	0.00	0.00
/v/	0.98	0.02	0.00
/y/	0.99	0.01	0.00
/sil/	0.78	0.05	0.17

Tabla 7. Tasa de errores y precisión del RAH para cada fonema del Corpus de Entrenamiento.

VI. Conclusiones

En el presente artículo se mostraron los avances logrados hacia la realización de un reconocedor del habla para la Lengua Mixteca. Usando como base a un usuario nativo y un corpus de 192 palabras fue posible obtener una tasa de precisión cercana al 95.00% y, mediante una técnica de adaptación, 88.61% para un usuario no-nativo. Estos resultados son significativos para futuras aplicaciones de aprendizaje de la lengua, sin embargo más trabajo es necesario. Dentro del trabajo a futuro se tienen los siguientes planes:

- desarrollar una técnica de mejora para incrementar el desempeño de la adaptación MLLR para el sistema DU.
- reclutar a más usuarios nativos de la misma variante para el desarrollo de un Corpus de Entrenamiento más amplio, lo cual conllevaría a la construcción de un sistema Independiente del Usuario (IU);
- medir el desempeño de MLLR conforme aumente el material de adaptación;
- probar el sistema con un usuario del género femenino, ya que ambos usuarios en este trabajo fueron del género masculino, y los patrones de voz femeninos tienen características particulares;
- realizar pruebas con más usuarios no-nativos;
- diseñar una interface interactiva para la realización de pruebas de usabilidad con un número más amplio de usuarios;
- incrementar el vocabulario y complejidad de las lecciones/relatos;
- implementar una métrica para medir el grado de avance/aprendizaje del usuario.

VII. Referencias

- [1] Hablantes de Lengua Indígena en México. Instituto Nacional de Estadística y Geografía (INEGI). <http://cuentame.inegi.org.mx/poblacion/lindigena.aspx?tema=P>. Junio 2011.
- [2] Pronunciation Power Speech Test. English Computerized Learning Inc. <http://www.english.computerizedlearning.com>

- englishlearning.com/products/pronunciation-power-speech-test/. Junio 2011.
- [3] Babbel. Lesson Nine GmbH. <http://es.babbel.com/#Reconocimiento-de-voz>. Junio 2011.
- [4] Rosetta Stone Version 4 TOTALe. Rosetta Stone. <http://www.rosettastone.com/content/rosettastonecom/en.html>. Junio 2011.
- [5] Bases para la Escritura de *tu'un savi* (Academia de la Lengua Mixteca). Colección Diálogos: Pueblos Originarios de Oaxaca. México. 2007.
- [6] Familia Mixteca: Instituto Lingüístico de Verano en México. <http://www.sil.org/mexico/mixteca/00e-mixteca.htm>. Junio 2011.
- [7] Catálogo de las Lenguas Indígenas Nacionales: Variantes Lingüísticas de México con sus autodenominaciones y referencias geoestadísticas. Instituto Nacional de Lenguas Indígenas. http://www.inali.gob.mx/pdf/CLIN_completo.pdf Junio 2011.
- [8] J. Ferguson de Williams, Gramática Popular del Mixteco del Municipio de Tezoatlán, San Andrés Yutatío, Oaxaca. Instituto Lingüístico de Verano, A.C. 2007. México D.F.
- [9] K. Beaty de Farris, P. García Sánchez, R. García Sánchez, J. Ojeda Sánchez, A. García, A. Santiago Jiménez, Diccionario Básico del Mixteco de Yosondúa, Oaxaca. Instituto Lingüístico de Verano, A.C. 2004. México, D. F.
- [10] S. Stark, A. Johnson, B. González de Guzmán, Diccionario Básico del Mixteco de Xochapa, Guerrero. Instituto Lingüístico de Verano, A.C. 2003. México, D.F.
- [11] R. M. Alexander, Mixteco de Atlatlahuca. 1980. Instituto Lingüístico de Verano. México, D.F.
- [12] L. A. Pineda, H. Castellanos, J. Cuétara, L. Galescu, J. Juárez, J. Llisterri, P. Pérez, L. Villaseñor, “The Corpus DIMEx100: Transcription and Evaluation”. *Language Resources and Evaluation*. 44:4. 2010. 347-370.
- [13] S. Young, P. Woodland, The HTK Book (for HTK Version 3.4). Cambridge University Engineering Department. 2006. Gran Bretaña.
- [14] L. Rabiner, “A tutorial on hidden markov models and selected applications in speech recognition”. *Proc. of IEEE*. 77:2. 1989. 257-286.
- [15] D. Jurafsky, J.H. Martin, *Speech and Language Processing*. 2nd Edition, 2009. Pearson: Prentice Hall, New Jersey, USA.

VIII. Autores

Dr. Santiago Omar Caballero Morales obtuvo su título de Doctor en Ciencias Computacionales con Especialidad en Reconocimiento del Habla en la University of East Anglia (UEA) en Reino Unido en 2009. Actualmente trabaja en la Universidad Tecnológica de la Mixteca en Oaxaca realizando investigación en el campo de reconocimiento de voz, traducción en tiempo real, interacción humano-computadora, y optimización combinatoria.

Ing. Edgar de los Santos Ramírez es Ingeniero en Diseño y actualmente realiza su tesis de maestría en Medios Interactivos en la Universidad Tecnológica de la Mixteca. Su investigación está enfocada al diseño gráfico de interfaces de aprendizaje.

Estimación Espectral de Potencia de Señales EEG con Métodos No Paramétricos.

***Norma Ramírez Hernández, **José Luis Ramos Quirarte**

*Universidad Politécnica de Madrid - Centro de Tecnología Biomédica, Madrid, España,
mirandapartida@terra.com.mx.

**Universidad de Guadalajara – Departamento de Electrónica (CUCEI) y Departamento de Ciencias naturales y Exactas (CUVALLES), Av. Revolución 1500, Col. Olímpica, C.P. 44430, Guadalajara, Jal., México, (33) 1378 5900 ext. 7726, jlquirarte@yahoo.com.mx, luis.ramos@valles.udg.mx

Resumen

En este trabajo se presentan los fundamentos y aplicación de algunos métodos de estimación espectral de potencia no paramétricos. La aplicación se realiza a señales sintéticas del tipo de sumas de senoides con y sin ruido para posteriormente aplicarlas a señales EEG. Se describen las características de las señales cerebrales y el porqué pueden representarse como procesos estocásticos para los cuáles se establecen algunas de sus propiedades estadísticas fundamentales como el sesgo y la consistencia. En particular se presentan las características propias de los métodos del Periodograma, de Blackman-Tukey, de Bartlett y de Welch. Se explica la ventaja de usar ventanas de ponderación y se presentan resultados aplicando estos métodos a señales sintéticas y a señales EEG.

Palabras Clave: Estimación Espectral de Potencia, Métodos No Paramétricos, Periodogramas, Señales EEG, Welch.

I. Introducción

La electroencefalografía (EEG) es el registro de la actividad eléctrica a lo largo del cuero cabelludo. EEG mide las fluctuaciones de tensión resultante de los flujos de corrientes iónicas dentro de las neuronas del cerebro. En contextos clínicos, EEG se refiere al registro de la actividad eléctrica espontánea del cerebro en un corto período de tiempo, generalmente 20-40 minutos, registrado con múltiples electrodos colocados en el cuero cabelludo, en reposo, durante la vigilia o el sueño. En neurología, las principales aplicaciones de diagnóstico del EEG son: epilepsia, coma, encefalopatías, la muerte cerebral, tumores, derrames cerebrales y otros trastornos cerebrales focales. También se utiliza ampliamente en la neurociencia, la ciencia cognitiva, la psicología cognitiva, y la investigación psicofisiológica [1,2].

Los potenciales eléctricos generados por las neuronas individuales son demasiado pequeños para ser recogidos por la actividad EEG por lo

tanto siempre refleja la suma de la actividad sincrónica de miles o millones de neuronas que tienen una orientación espacial similar. Si las células no tienen orientación espacial similar, los iones no se alinean para crear ondas que sean detectadas. Las posiciones de los electrodos y sus nombres son especificados por el sistema internacional 10-20 para la mayoría de las aplicaciones clínicas y de investigación.

EEG suele ser descrita en términos de actividad rítmica y transitorios. La actividad rítmica se divide en bandas de frecuencia: delta (hasta 4 Hz), theta (4-8 Hz), alfa (8-13 Hz), beta (13-30 Hz), gamma (30-100 Hz), estas bandas surgieron porque la actividad rítmica dentro de una determinada variación de frecuencias fueron observadas en cierta distribución sobre el cuero cabelludo o representando un significado biológico determinado. Las bandas de frecuencias para su interpretación, se suelen extraer utilizando métodos espectrales como el presentado en este trabajo.

Los datos de EEG están casi siempre contaminados por los artefactos que son señales provenientes de origen no cerebral. Algunos de los tipos más comunes de los artefactos biológicos son: artefactos inducidos por los ojos (los guiños, los movimientos oculares y la actividad de los músculos extraoculares), artefactos cardíacos, la activación muscular, etc., también hay muchos artefactos provenientes del exterior del cuerpo; movimientos del paciente, o incluso el electrodo puede causar estallidos, los picos procedentes de un cambio momentáneo en la impedancia de un electrodo, una conexión a tierra deficiente de los electrodos de EEG puede causar importantes artefactos a 50 o 60 Hz, dependiendo de la frecuencia del sistema eléctrico local.

El análisis cuantitativo del EEG permite la determinación objetiva de ciertos parámetros con el fin de identificar las conductas normales o patológicas. Un enfoque fenomenológico en lugar de una relación causal, basado en las características de la señal tales como frecuencias, formas de onda, fase, duración, etc., están relacionados con situaciones del tipo conductual, mental o fisiológico o incluso con algunas patologías específicas.

Debido a su complejidad, la señal EEG tiene un comportamiento plenamente aleatorio que contiene además señales transitorias, siendo representada como un proceso estocástico (PE) no estacionario, que tiene que ser caracterizado por sus propiedades probabilísticas y estadísticas. Así pues, el análisis de la señal EEG requiere entre otras acciones: la eliminación de artefactos, la extracción de transitorios para su caracterización temporal y morfológica, segmentación para considerar intervalos estacionarios, caracterización estadística y determinación de estimadores y clasificación de los segmentos para identificación de estados funcionales o patologías asociadas [3-6].

II. Desarrollo

A.- Métodos de Estimación Espectral.

El propósito de la estimación espectral es estimar la densidad espectral de potencia de un proceso estocástico a partir de las muestras de una realización. De este modo, es posible caracterizar el contenido espectral de una señal determinada.

Las propiedades asociadas con un estimador espectral son las mismas que se consideran en los estimadores estadísticos clásicos, como error cuadrático medio, varianza, desviación, sesgo, etc. Se llamará $\hat{\theta}$ a un estimador que estima un valor θ . Las propiedades más importantes que se consideran en un estimador espectral son las siguientes [7-9].

Sesgo.- El sesgo de un estimador, se define como la diferencia entre su valor esperado y el verdadero valor del parámetro a estimar,

$$B(\hat{\theta}) = E[\hat{\theta}] - \theta \quad (1)$$

Un estimador es insesgado si,

$$B(\hat{\theta}) = 0 \quad (2)$$

Consistencia.- Un estimador es consistente cuando converge en probabilidad a θ si el tamaño de la muestra aumenta. Hay que tener en cuenta que cuando el tamaño de la muestra varía, el estimador que se usa es otro. Así, una secuencia de estimadores para θ , $\{t_n\}$, converge en probabilidad hacia θ si, y solamente si, $\forall \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P_r\{|t_n - \theta| < \varepsilon\} = 1 \quad (3)$$

Eficiencia.- Se dice que un estimador es más eficiente que otro cuando su varianza es menor. El teorema de Cramer-Rao impone un valor mínimo para la varianza de un estimador, donde f es la función de verosimilitud de θ .

$$\text{Var}(\theta) \geq \frac{1}{E\left[\left[\frac{\partial}{\partial \theta} \log f(x; \theta)\right]^2\right]} \quad (4)$$

Robustez.- Un estimador es robusto si su salida no varía significativamente cuando las suposiciones sobre las que se basa no se cumplen.

Las técnicas utilizadas en estimación espectral suelen clasificarse en métodos paramétricos y no paramétricos.

Los métodos no paramétricos utilizan únicamente las muestras disponibles del PE para obtener su covarianza o espectro. No se asume ningún modelo generador de la señal observada.

Los métodos paramétricos parten de alguna suposición acerca del PE a analizar, estimando los parámetros que en teoría lo definen. Se asume algún modelo (AR, MA, ARMA...).

B.- Métodos No-Paramétricos de Estimación Espectral.

Los métodos no paramétricos más utilizados son: Periodograma, Blackman-Tuckey, Bartlett y Welch [10,11].

B.1.- Periodograma

Considerando a $x[n]$ una realización de un PS con intervalo de observación $0 \leq n < N$ (siendo N el tamaño de la ventana), se define el periodograma como:

$$P_x(\omega) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x[n] e^{-j\omega n} \right|^2 \geq 0 \quad (5)$$

Empleando la transformada discreta de Fourier (DFT) con longitud L y con $k = 0, \dots, L-1$,

$$P_x \left(\frac{2\pi k}{L} \right) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi}{L} kn} \right|^2 \quad (6)$$

También podemos definir al periodograma como el cuadrado del módulo de la transformada de Fourier (TF) de $x[n]$ que coincide con la TF del estimador de autocorrelación sesgado:

$$P_x(\omega) = \frac{1}{N} \left| X(e^{j\omega}) \right|^2 = \sum_{m=-N+1}^{N-1} \hat{R}[m] e^{-j\omega m} \quad (7)$$

Ya que,

$$\hat{R}[m] = \frac{1}{N} (x[m] \star x[-m]) \xrightarrow{TF} \frac{1}{N} X(e^{j\omega}) X^*(e^{j\omega})$$

$$\hat{R}[m] = \frac{1}{N} \left| X(e^{j\omega}) \right|^2$$

Se puede observar en el caso del periodograma, que el sesgo es

$$E\{P_x(\omega)\} = E\left\{ \sum_{m=-N+1}^{N-1} \hat{R}[m] e^{-j\omega m} \right\} \quad (8)$$

$$E\{P_x(\omega)\} = \sum_{m=-N+1}^{N-1} E\{\hat{R}[m]\} e^{-j\omega m}$$

$$E\{P_x(\omega)\} = \sum_{m=-N+1}^{N-1} \frac{N-|m|}{N} R[m] e^{-j\omega m}$$

$$E\{P_x(\omega)\} = TF \left\{ \frac{N-|m|}{N} R[m] \right\} \quad (9)$$

De donde:

$$w[n] = \frac{N-|m|}{N} \xrightarrow{TF} W(e^{j\omega}) = \left(\frac{\sin \frac{N\omega}{2}}{\sin \frac{\omega}{2}} \right)^2$$

Es decir,

$$E\{P_x(\omega)\} = S_x(e^{j\omega}) \star W(e^{j\omega}) \quad (10)$$

Por lo que se pierde resolución debido al enventanado y el estimador resulta sesgado pero asintóticamente insesgado ya que cuando,

$$N \rightarrow \infty \Rightarrow w[n] \rightarrow 1$$

También se observa que,

$$\lim_{N \rightarrow \infty} \text{var}[P_x(\omega)] = S_x^2(\omega) \quad (11)$$

Indica que el periodograma es inconsistente ya que su varianza no tiende a 0 al aumentar el tamaño de la ventana de observación, sino que oscila alrededor de su propio valor espectral $S_x^2(\omega)$. Es decir, la varianza es mayor cuanto mayor es el valor de la DEP.

B.2.- Periodograma ponderado

Para solucionar el problema de la inconsistencia del periodograma, se puede recurrir al uso de ventanas de promediación espectral que hagan al estimador consistente. Ponderando la secuencia de la señal por una ventana $[w_0, \dots, w_{N-1}]$, el periodograma modificado, quedaría como:

$$P_{xw}(\omega) = \frac{\frac{1}{N} \left| \sum_{n=0}^{N-1} w[n] x[n] e^{-j\omega n} \right|^2}{\frac{1}{N} \left| \sum_{n=0}^{N-1} w[n] \right|^2} \quad (12)$$

La aplicación del enventanado a la función de densidad espectral produce alisamiento o redondeamiento de los picos de la función (suavizado del espectro) y aparición de picos espectrales o fugas espectrales, no existentes en la función de densidad espectral real del PE.

Dependiendo del efecto a minimizar, será la elección del tipo de ventana.

B.3.- Método de Blackman y Tukey

Este método consiste en reducir la varianza, ponderando los valores de acuerdo con su cercanía al origen, enventanando al estimador de autocorrelación con $w[n]$, $-M \leq m \leq M \leq N$.

$$\hat{P}_{BT}(\omega) = \sum_{m=-M+1}^{M-1} \hat{R}[m]w[m]e^{-j\omega m} \quad (13)$$

En este caso, puede emplearse una longitud de ventana menor que la longitud del estimador de la autocorrelación, despreciando las muestras más alejadas que son las peor estimadas.

El estimador es sesgado y asintóticamente insesgado si $N, M \rightarrow \infty$.

$$E\{\hat{P}_{BT}(\omega)\} = S_x(e^{j\omega}) \star W_B(e^{j\omega}) \star W(e^{j\omega}) \quad (14)$$

La varianza puede aproximarse para $M \ll N$ por:

$$\text{var}\{\hat{P}_{BT}(\omega)\} \approx S_x^2(\omega) \frac{1}{N} \sum_{m=-M+1}^{M-1} w^2[m] \quad (15)$$

Se tiene pérdida de resolución, de acuerdo con la ventana $w[n]$. El estimador es inconsistente pero se vuelve consistente si $N \rightarrow \infty$ y M se mantiene constante. El compromiso entre sesgo (resolución) y varianza es: para $N \gg M$ se tiene menor varianza pero peor resolución, para $w[n]$ diferente a la rectangular, se mejora la varianza pero se pierde resolución.

B.4.- Método de Bartlett

En este caso, se reduce la varianza promediando varias realizaciones. Se divide la señal en K segmentos de M puntos ($N = KM$). Se calcula el periodograma para cada segmento y se promedian:

$$\hat{P}_x^i(\omega) = \frac{1}{M} \left| \sum_{n=0}^{M-1} x_i[n]e^{-j\omega n} \right|^2 \quad (16)$$

$$\hat{P}_B(\omega) = \frac{1}{K} \sum_{i=0}^{K-1} \hat{P}_x^i(\omega) \quad (17)$$

Dado que la ventana contiene N/K muestras, el estimador presenta mayor sesgo que el periodograma:

$$E\{\hat{P}_{BT}(\omega)\} = \frac{1}{K} \sum_{i=0}^{K-1} E\{\hat{P}_x^i(\omega)\} = S_x(\omega) \star W_B(e^{j\omega}) \quad (18)$$

Considerando segmentos independientes, el estimador es consistente si $N \rightarrow \infty$ con M constante:

$$\text{var}\{\hat{P}_B(\omega)\} = \frac{1}{K} \text{var}\{\hat{P}_x^i(\omega)\} \quad (19)$$

B.5.- Método de Welch

A diferencia del método de Bartlett, se utiliza enventanado de datos donde los valores extremos del intervalo se usan para determinar los valores extremos de la correlación y traslape entre segmentos, lo que permite promediar más realizaciones sin degradar la resolución.

Para K segmentos de longitud N traslapados D muestras, se tiene:

$$x_l[n] = x[n + lD]; \quad n = 0, \dots, N-1; \quad l = 0, \dots, K-1$$

Entonces el periodograma modificado de cada segmento es:

$$\hat{P}_{xl}(\omega) = \frac{1}{NU} \left| \sum_{n=0}^{M-1} x_l[n]w[n]e^{-j\omega n} \right|^2 \quad (20)$$

Con U , factor de normalización de potencia,

$$U = \frac{1}{N} \sum_{n=0}^{N-1} w^2[n] \quad (21)$$

El promediado de periodogramas modificados es:

$$\hat{P}_W(\omega) = \frac{1}{K} \sum_{l=0}^{K-1} \hat{P}_{xl}(\omega) \quad (22)$$

Este método también es sesgado y depende de la ventana empleada

$$E\{\hat{P}_W(\omega)\} = \frac{1}{K} \sum_{l=0}^{K-1} E\{\hat{P}_{xl}(\omega)\} \quad (23)$$

$$E\{\hat{P}_{xl}(\omega)\} = \frac{1}{NU} S_x(\omega) \star |W(e^{j\omega})|^2 = E\{\hat{P}_B(\omega)\} \quad (24)$$

Considerando independencia estadística entre los segmentos, su varianza está dada por:

$$\text{var}\{\hat{P}_W(\omega)\} \approx \text{var}\{\hat{P}_{xl}(\omega)\} \quad (25)$$

C.- Ventanas Espectrales.

Existen muchos tipos de ventanas factibles de emplear en los diversos métodos de estimación espectral. El MatLab cuenta con una herramienta "wintool" que cuenta con una interfaz gráfica de usuario (GUI) para el análisis y diseño de ventanas espectrales que incluye entre otras, las siguientes: Bartlett, Chebyshev, Gausiana,

Hamming, rectangular, etc. En la Fig. 1, se muestran dos casos.

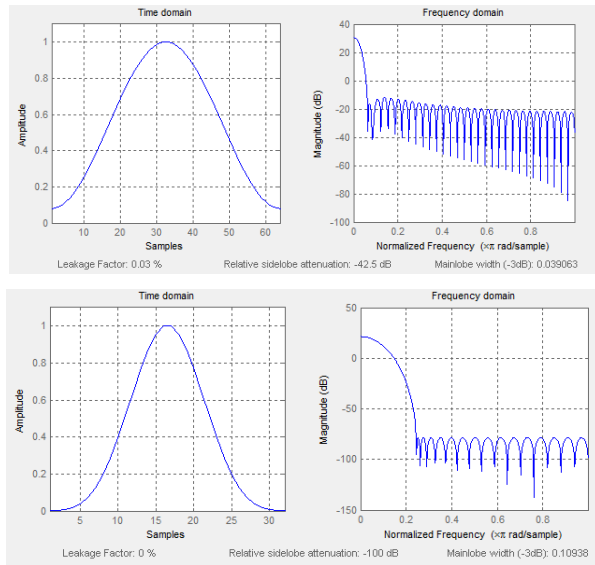


Fig. 1. Ventanas Hamming y Schebyshev.

III. Resultados

A.- Aplicación de los Métodos a Señales Simuladas.

A.1.- Estimación para señales sin ruido

Con el propósito de analizar el desempeño de los métodos de estimación espectral y su comparación, se implementan señales simuladas resultantes de la suma de senoides a 10 Hz y a 30 Hz, como se muestra en la Fig. 2.

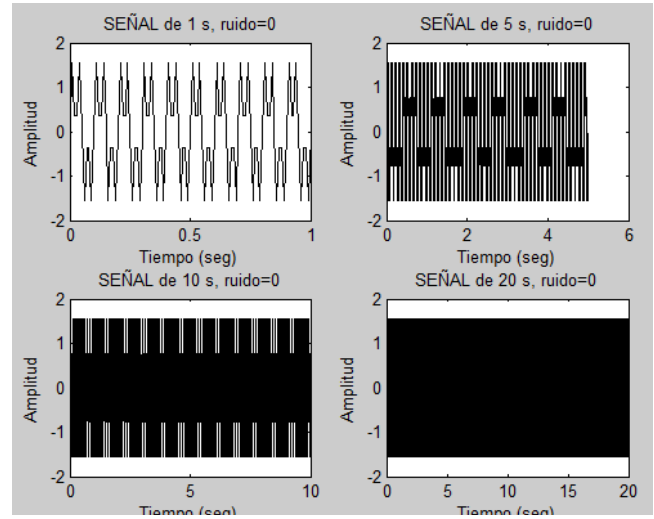


Fig. 2. Señales simuladas sin ruido.

Para las señales presentadas en la Fig. 2, se obtuvo el periodograma y los resultados se muestran en la Fig. 3.

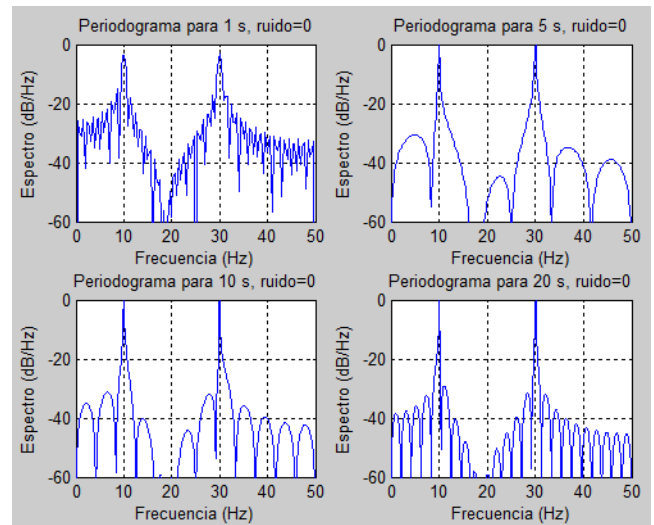


Fig. 3. Periodogramas para señales sin ruido.

Los resultados obtenidos aplicando el método de Welch a las señales de la Fig. 2, se presentan en la Fig. 4.

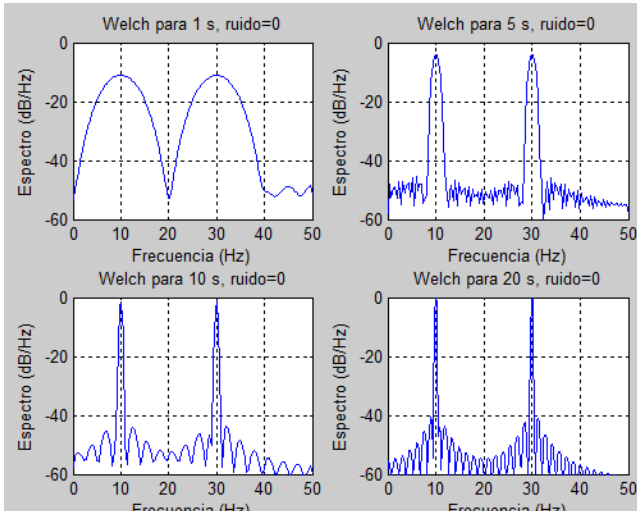


Fig. 4. Método de Welch para señales sin ruido.

A.2.- Estimación para señales con ruido

Para incorporar efectos aleatorios, a la señal senoidal con duración de 10 s, se le agrega ruido aleatorio con valores de: 0, 0.5, 1, y 2, tal como se muestra en la Fig. 5. Los resultados obtenidos aplicando el periodograma y el método de Welch a estas señales, se presentan en las Figs. 6 y 7, respectivamente.

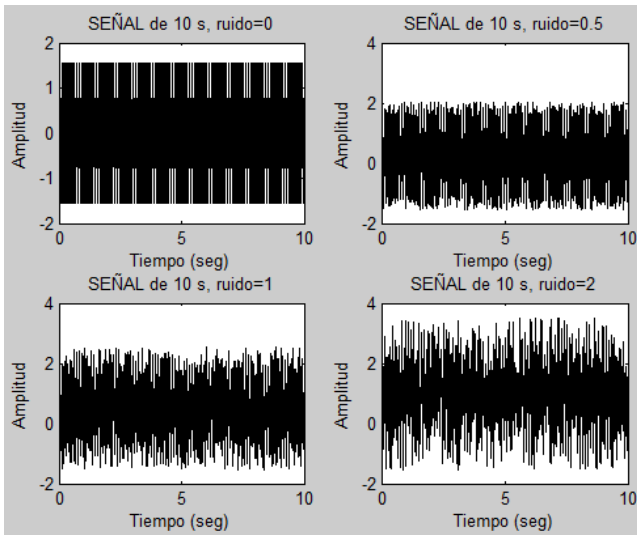


Fig. 5. Señales simuladas con ruido.

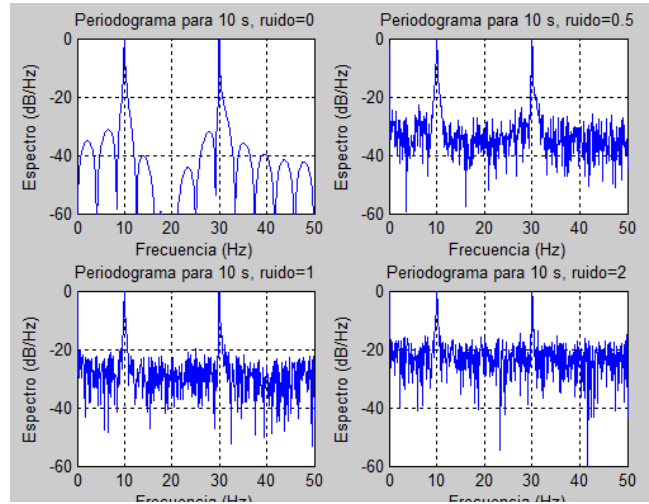


Fig. 6. Periodogramas para señales con ruido.

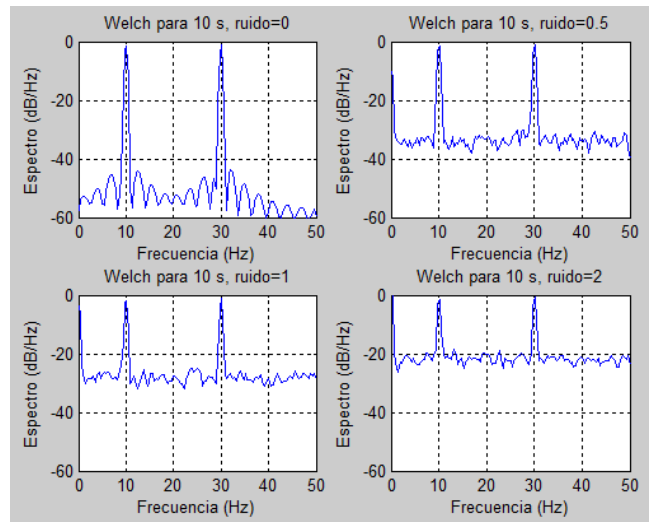


Fig. 7. Método de Welch para señales con ruido.

A.3.- Estimación con el uso de ventanas

Para analizar el efecto del uso de ventanas, se elige la señal con duración de 10 s y ruido de 0.5. Para el periodograma se eligen ventanas Hamming, Gauss y Tukey, todas con la misma extensión que la señal. Los resultados empleando estas ventanas, se muestran en la Fig. 8.

Para el método de Welch, empleamos ventanas con tamaños de: 32, 128, 256 y 512 con los resultados presentados en la Fig. 9 para ventanas Hamming y en la Fig. 10 para ventanas Tukey.

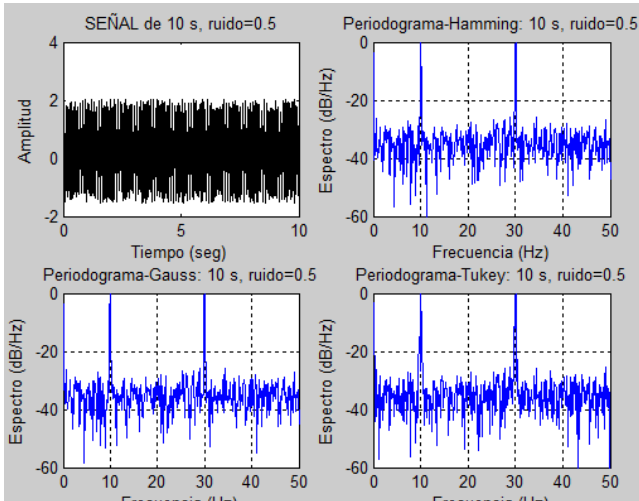


Fig. 8. Periodogramas con ventanas.

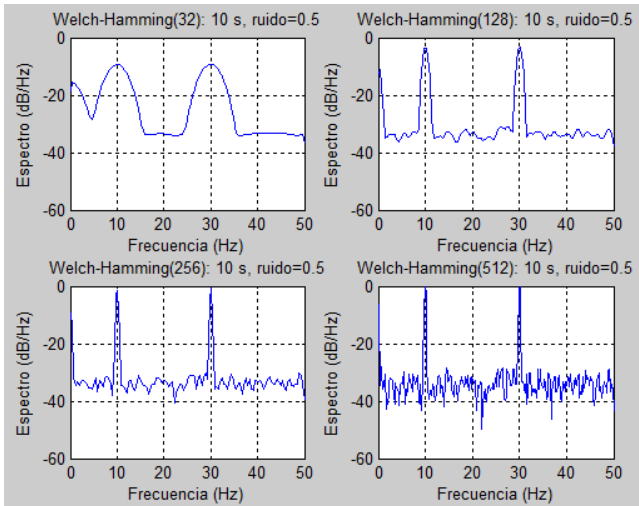


Fig. 9. Método de Welch con ventanas Hamming.

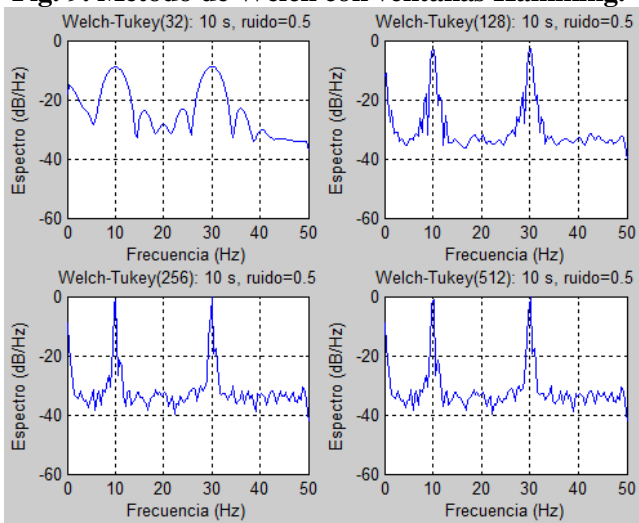


Fig. 10. Método de Welch con ventanas Tukey.

B.- Aplicación de los Métodos a Señales EEG.

Como un ejemplo de la aplicación de estimación espectral, se hace una comparación de espectros de señales de EEG recogidas en diferentes fases de conducción para evaluar el nivel de alerta de un sujeto. Si la actividad cerebral contiene mayor potencia en frecuencias altas (banda beta) que en bajas (banda alpha), entonces se tiene un estado de mayor alerta.

La Fig. 11, muestra una sección de la actividad eléctrica de un sujeto, capturada con un único electrodo EEG en la región occipital en tres estados distintos: 5 minutos en reposo con ojos cerrados, 5 minutos utilizando un simulador de conducción con dificultad baja, y 5 minutos con dificultad alta.

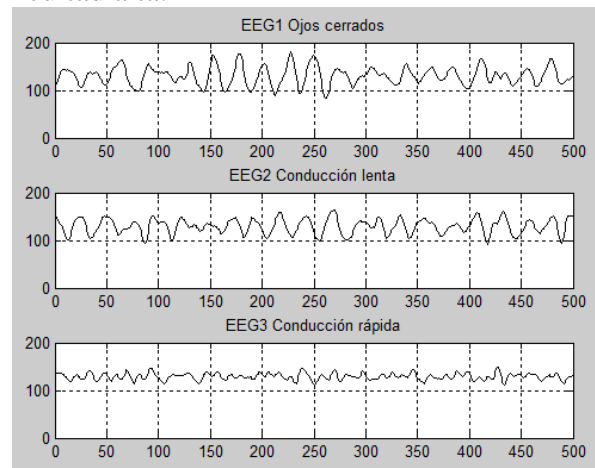


Fig. 11. Señal EEG de un sujeto en: reposo, baja actividad y alta actividad.

Los resultados obtenidos aplicando el método del periodograma, se muestran en la Fig. 12.

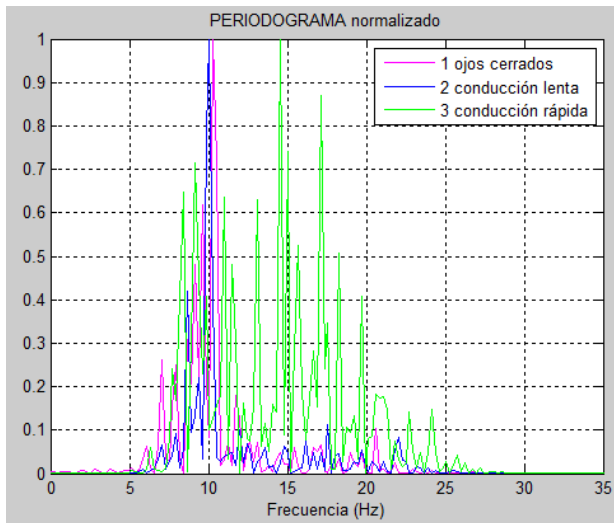


Fig. 12. Estimación espectral: Periodograma.

De manera semejante, los resultados obtenidos por el método de Welch bajo las tres condiciones de excitación, se presentan en la Fig. 13.

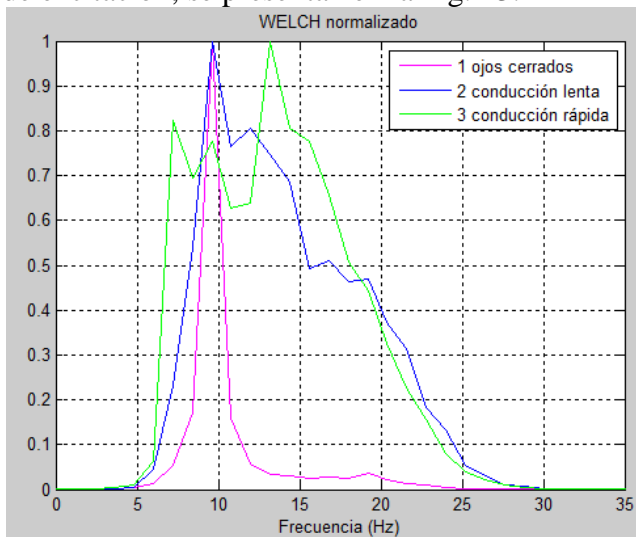


Fig. 13. Estimación espectral: Método de Welch.

IV. Discusión

A través de múltiples investigaciones [1-3] y como es enfatizado en este trabajo, las señales cerebrales son de una naturaleza tan compleja, que una aproximación para su estudio es considerarlas como procesos estocásticos que aunque no satisfacen plenamente algunas propiedades fundamentales como estacionaridad y ergodicidad, se aproximan lo suficiente bajo ciertas condiciones, lo que permite el poder aplicar metodologías para hacer estimaciones apropiadas para el estudio e investigación de

comportamientos conductuales y cierto tipo de enfermedades cerebrales.

Los resultados aquí presentados se basan en métodos no paramétricos que producen buenos resultados sin necesidad de modelos preestablecidos. Se puede apreciar que para señales sintéticas simples, su respuesta es muy buena, empeorando para situaciones aleatorias con mucho ruido y evidentemente para señales reales que además de su aleatoriedad, se encuentran contaminadas con artefactos de diferentes tipos. Los resultados pueden ser mejorados cuando las señales son limpiadas de artefactos a través del uso de filtros y otras técnicas pero aún así, se puede apreciar que ambos métodos de estimación son lo suficientemente buenos para obtener resultados sobre todo del tipo cualitativos.

V. Conclusiones

De los resultados presentados, es evidente que para investigar y comprender el funcionamiento cerebral bajo diversas condiciones o en caso de enfermedades, el desarrollo de diversas metodologías que se basen en la caracterización de señales eléctricas, magnéticas, electromagnéticas, imágenes, etc., obtenidas directa o indirectamente del cerebro son indispensables. Actualmente existen numerosos grupos de investigación en el mundo desarrollando y aplicando metodologías que permitan un mejor acercamiento a la comprensión de estos fenómenos. En particular los métodos de estimación espectral no paramétrica presentados en este trabajo, permiten el análisis del contenido de potencia en las señales de EEG para interpretar respuestas conductuales y posibles enfermedades. Existen además estimadores paramétricos que pueden emplearse asimismo y comparar resultados para diferentes situaciones para establecer ventajas y desventajas o bien para obtener información complementaria. Por lo anterior, los autores pretenden presentar resultados análogos empleando métodos paramétricos y realizar comparaciones pertinentes.

VI. Referencias

- [1] Shanbao Tung y Nitish V. Thakor, Quantitative EEG Analysis, Methods and Clinical Applications, Eds. Artech House, 2009.
- [2] Jaakko Malmivuo y Robert Plonsey, Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields, Oxford Univ Press, 1995.
- [3] William O Tatum, Aatif Husain, Selim R. Bembadis, Peter W. Kaplan, Handbook of EEG Interpretation, Demos, Medical Publishing, 2008.
- [4] J. G. Proakis, D. G. Manolakis, Digital Signal Processing: Principles, Algorithms, and Applications, Third Edition, 1996.
- [5] Alan V. Oppenheim, Alan S. Willsky, S. Hamid Nawab. Señales y Sistemas, 2o ed. Prentice Hall 1998.
- [6] Oppenheim, A.V., and R.W. Schafer, Discrete-Time Signal Processing, Prentice-Hall, 1989.
- [7] Emanuel Parzen, Teoría Moderna de Probabilidades y sus Aplicaciones, Editorial Limusa, 1979.
- [8] Athanasios Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, 1965.
- [9] Steven M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice Hall, 1993.
- [10] Stoica, P., and R.L. Moses, Introduction to Spectral Analysis, Prentice-Hall, 1997.
- [11] Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms", IEEE Trans. Audio Electroacoustics, Vol. AU-15, June 1967, pp.70-73.

VII. Autores

M. en C. Norma Ramírez Hernández obtuvo su título de Ingeniería en Comunicaciones y Electrónica y el grado de Maestría en Ciencias en Ingeniería Electrónica en la Universidad de Guadalajara, México y actualmente se encuentra realizando sus estudios doctorales en la Universidad Politécnica de Madrid, España, en la especialidad de Tecnología Biomédica.

Dr. José Luis Ramos Quirarte obtuvo su título de Ingeniería en Comunicaciones y Electrónica en la Universidad de Guadalajara, México, su grado de Maestría en Ciencias en Ingeniería Eléctrica (Telecomunicaciones) en el CINVESTAV, México y el Doctorado en Ingeniería de Alta Frecuencia en la Universidad de Chalmers, Suecia. Actualmente es Profesor-Investigador en la Universidad de Guadalajara, México.

GENERACION DE INDICADORES DE PRODUCTIVIDAD MEDIANTE LA IMPLEMENTACION DE UN DATA WAREHOUSE EN SQL SERVER ENTERPRISE APLICADO A HOSPITALES DE ALTA ESPECIALIDAD.

***Karina Cancino Villatoro, **Vanessa Benavides García**

*Universidad Politécnica de Tapachula 4ª. Norte entre 1ª y 3ª poniente, Tapachula, Chiapas, México
{kari_cancino@hotmail.com}

**Facultad de Contaduría Pública del Campus IV de la Universidad Autónoma de Chiapas. Carretera a Puerto Madero Km 1.5, Tapachula, Chiapas, México (962) 6251723 {vanebenavides@gmail.com}

Resumen

El presente trabajo tiene como objetivo presentar los resultados obtenidos de la creación e implementación de un Data Warehouse para la generación de indicadores de productividad para el área de Estadística del Hospital Regional de Alta Especialidad (HRAE) “Ciudad Salud” ubicado en la ciudad de Tapachula, Chiapas, México. Se utilizó como tecnología de desarrollo el Sistema Gestor de Base de Datos SQL Server Edición Empresarial, procedimientos almacenados de Transact SQL para la extracción y transformación de datos, jobs de SQL Server para la carga de datos y la suite de Microsoft .NET como plataforma de desarrollo de la aplicación de acceso a datos.

Palabras Clave: Data Warehouse, Indicadores de Productividad, Procedimientos Almacenados, Transact SQL, Jobs, Hospital de Alta Especialidad.

I. Introducción

Actualmente los sistemas de información son un factor muy importante para la competitividad de una organización, pero la mayoría de las veces son solo transaccionales, y los directivos requieren de sistemas que proporcionen verdaderas herramientas para la toma de decisiones, esto ha sido factor importante para que se desarrolle la tecnología de Data Warehouse que consiste en la integración de datos históricos, proveniente de múltiples y posiblemente diferentes fuentes de datos con el propósito de hacer análisis y a partir de este tomar decisiones. [1]

Las organizaciones gubernamentales no han sido la excepción y es por ello que el objetivo de este trabajo es mostrar la manera en cómo se implementó la tecnología de Data Warehouse en el Hospital de Alta Especialidad (tercer nivel) Ciudad Salud para la generación de indicadores de productividad.

Un hospital de tercer nivel se encarga de brindar atención médica a usuarios que presentan padecimientos de alta complejidad diagnóstica y de tratamiento a través de una o varias especialidades medicas, quirúrgicas o medico-quirúrgicas. La presencia de las tecnologías de información en los hospitales de tercer nivel es indispensable, ya que permite identificar y mejorar la calidad de servicios mediante una correcta toma de decisiones.

Dentro de la Red de Servicios de Alta Especialidad (REDSAES) a la cual pertenecen todos los Hospitales Regionales de Alta Especialidad (HRAES) del país se emplean diversos indicadores de productividad como medida de evaluación del desempeño de los servicios de salud que dichas instituciones gubernamentales ofrecen. Dichos indicadores son la base para la determinación de la productividad de estos nosocomios. [2]

El trabajo se organiza de la siguiente manera. La sección 2 presenta el concepto y características

de un Data Warehouse para el conocimiento del lector. La sección 3 describe la metodología de desarrollo del Data Warehouse Estadístico. La sección 4 se muestra los resultados obtenidos de la implementación. Finalmente en la sección 5 se presentan la conclusión y trabajos futuros de la investigación.

II. Datawarehouse

II.1. ¿Qué es un Almacén de Datos (Data Warehouse)?

Un almacén de datos es un repositorio de datos que está orientado a resolver la necesidad de toma de decisiones del usuario final dentro de una organización no solo en el soporte de la gerencia de alto nivel que debe tomar decisiones de objetivos y planificar estrategias; sino también en niveles más técnicos de toma de decisión.

Principales objetivos que persigue: [3]

1. Hacer que la información de una organización sea de fácil acceso.
2. Presentar la información de la organización consistente.
3. Ser adaptable y resistente a los cambios.
4. Ser un bastión seguro que proteja la información activa.
5. Servir como fundamento para la toma de decisiones.
6. Ser aceptado por la comunidad del negocio si se desea estimar como exitoso.

II.2 Características

Las características principales con las que cuenta un almacén de datos son: orientado al tema, integrado, de tiempo variante y no volátil.

- *Orientado a temas:* se basa en clasificar la información según los temas de interés de la organización y que a su vez estarán interrelacionados con los procesos que se llevan a cabo dentro de la organización. Como su nombre lo indica los datos deben ser organizados por los temas de estudio, no por aplicación.

- *Integrado:* es importante que la información que se encuentra almacenada coincida con la información que vendrá de entrada hacia el almacén. Debido a que el almacén de datos se alimenta de diversas fuentes de datos tanto internas como externas de la organización, deben integrarse con la mayor consistencia posible, es posible que se tenga que homologar los campos de los datos para su correcta integración.
- *De tiempo variante:* En los sistemas transaccionales la información es requerida desde el momento en que se inicia una sesión dentro del sistema, a diferencia de un almacén de datos en donde por lo regular la información que más interesa es la histórica para la identificación y evaluación de tendencias.
- *No volatil:* En el caso de los almacenes de datos la información se considera útil cuando es completamente estable, es decir, que los datos no cambian constantemente como en el caso de los sistemas operacionales, esta pasividad en el movimiento de datos permite una perspectiva más grande del análisis y la toma de decisiones.

En el ambiente operacional, la actualización (insertar, borrar y modificar), se hace regularmente sobre una base de registro por registro. En el almacén de datos es mucho más simple, hay dos únicos tipos de operaciones: la carga inicial de datos y el acceso a los mismos. No hay actualización de datos en el depósito, como una parte normal de procesamiento (ver Figura 1).

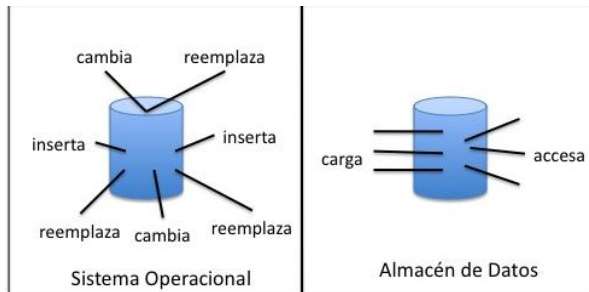


Figura 1. Operaciones sobre los datos en un Data Warehouse y un sistema operacional.

III. Desarrollo

III.1 Indicadores de Productividad

Un indicador es una variable estadística numérica que representa un valor que indica el comportamiento de la misma, la determinación de que variables se manejan como indicadores depende de la Institución de Salud en la cual se encuentre y en esta investigación se tomaron de los Lineamientos para la Integración de Carpetas de Junta de Gobierno de los Institutos Nacionales de Salud de la Secretaria de Salud.

III.2 Estrategia de Implementación.

La estrategia de implementación que se utilizó fue bottom-up (abajo hacia arriba) ya que la tecnología con la que se cuenta actualmente (Sistema de Cobros Hospitalario y Sistema de Información para la Gerencia Hospitalaria - SIGHO) se encuentra en sus primeras etapas de implementación.

III.3 Metodología de Desarrollo

La metodología de desarrollo fue de análisis y diseño estructurado en cascada el cual consiste en reunir los requerimientos, analizarlos, realizar un diseño preliminar en base a los resultados del análisis. El diseño inicia como un nivel abstracto y se divide sucesivamente en niveles más concretos.

III.4 Diseño del Data Warehouse

INGRESOS	EGRESOS
+ID_INGRESO: INTEGER	+ID_EGRESO: INTEGER
+FECHA: DATETIME	+FECHA: DATETIME
+SECCION: VARCHAR	+SECCION: VARCHAR
+SEXO: VARCHAR ()	+SEXO: VARCHAR ()
+NUM_EXP: VARCHAR ()	+NUM_EXP: VARCHAR ()
+ESPEC: VARCHAR ()	+ESPEC: VARCHAR ()
+NIVEL: VARCHAR ()	+NIVEL: VARCHAR ()
+NOMBRE: VARCHAR ()	+NOMBRE: VARCHAR ()
+APPELLIDOS: VARCHAR ()	+APPELLIDOS: VARCHAR ()

CONSULTAS
+ID_CONSULTA: INTEGER
+FECHA: DATETIME
+ESPECIALIDAD: VARCHAR ()
+MEDICO: VARCHAR ()
+HORA: DATETIME
+CONSULTORIO: VARCHAR ()
+SUBSECUENTE: BIT
+SEXO: VARCHAR ()
+DIAGNOSTICO: VARCHAR ()
+MUNICIPIO: VARCHAR ()
+ESTADO: VARCHAR ()
+DERECHOHABIENCIA: VARCHAR ()
+EXPEDIENTE: VARCHAR ()
+NOMBRE_PACIENTE: VARCHAR ()

Figura 2: Diseño del Data Warehouse.

III.5 Arquitectura

La arquitectura que se utilizó es de solo almacén de datos la cual consiste en construir un Data Warehouse centralizado aplicando las operaciones de depuración, integración, resumen y adición de datos. En este tipo de arquitectura este único Data Warehouse alimenta de información a los usuarios para el apoyo a la toma de decisiones. [4]

En la figura 3 se muestra la arquitectura lógica sobre la cual se desarrolló el Data Warehouse Estadístico.

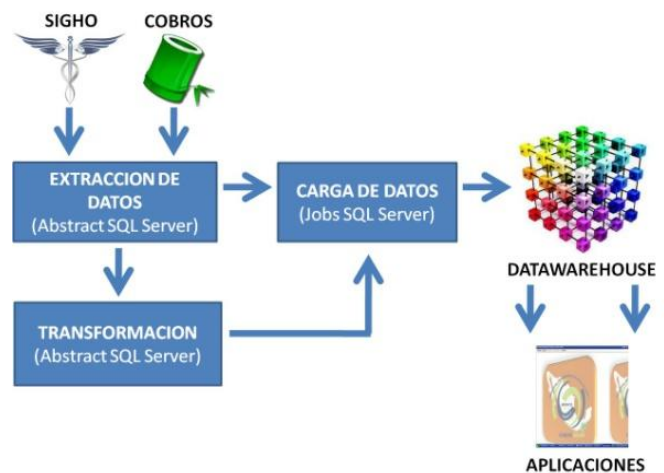


Figura 3: Arquitectura Lógica Data Warehouse

Se tomaron las bases de datos transaccionales de los sistemas principales que operan dentro del

HRAE Ciudad Salud que son el Sistema de Información para la Gerencia Hospitalaria y el Sistema de Cobros y a partir de ahí se realizó el proceso de extracción y transformación de datos mediante el desarrollo de procedimientos almacenados (ver Figura 4) utilizando el lenguaje procedural Abstract de SQL Server Enterprise.

```
ALTER procedure [dbo].[dwh_proc_egresos] as
declare @fecha1 varchar(20), @fecha2 varchar(20)
declare @fecini datetime, @fecfin datetime
begin

--obteniendo la fecha
select @fecha1=cast(getdate() as varchar(20))
set @fecha1=@fecha1 + ' 00:00:00 a.m.'
set @fecini=cast(@fecha1 as datetime)
select @fecha2=cast(getdate() as varchar(20))
set @fecha2=@fecha2+' 23:59:59 p.m.'
set @fecfin=cast(@fecha2 as datetime)

--insertando los datos en la tabla de estadística
INSERT INTO estadistica..egresos
select distinct
sh.cns_sec_hosp,sh.fh_egr_sec,cs.seccion,p.sexo,p.mum_exp,
ce.descripcion,cn.nivel,p.nom,p.ape
from sc_sec_hospitalizados sh,sc_cat_camras ce,
sc_cat_secciones cs,sc_hospitalizados h,
sc_cat_paciente p,sc_cat_medicos m,sc_cat_especialidades
ce,sc_cat_niveles cn
where sh.fh_egr_sec between DATEADD(day,-7,@fecha1)
and @fecha2 and cc.id_cama=sh.id_cama and
cs.id_seccion=cc.id_seccion and
h.id_hosp=sh.id_hosp and p.id_pac=h.id_pac and
m.id_medico=sh.id_medico and
ce.id_esp=m.id_esp and cn.id_niv=p.id_niv
end
```

```
ALTER procedure [dbo].[dwh_proc_ingresos] as
declare @fecha1 varchar(20), @fecha2 varchar(20)
declare @fecini datetime, @fecfin datetime
begin
```

```
--obteniendo la fecha
select @fecha1=cast(getdate() as varchar(20))
set @fecha1=@fecha1 + ' 00:00:00 a.m.'
set @fecini=cast(@fecha1 as datetime)
```

```
select @fecha2=cast(getdate() as varchar(20))
set @fecha2=@fecha2+' 23:59:59 p.m.'
set @fecfin=cast(@fecha2 as datetime)
```

```
--insertando los datos en la tabla de estadística
INSERT INTO estadistica..ingresos
select distinct sh.cns_sec_hosp,sh.fh_ing_sec,cs.seccion,
p.sexo,p.num_exp,ce.descripcion,cn.nivel,p.nom,p.ape
from sc_sec_hospitalizados sh,sc_cat_camras ce,
sc_cat_secciones cs,sc_hospitalizados h,
sc_cat_paciente p,sc_cat_medicos m,sc_cat_especialidades ce,
sc_cat_niveles cn
where sh.fh_ing_sec between DATEADD(day,-7,@fecha1) and
@fecha2 and cc.id_cama=sh.id_cama and
cs.id_seccion=cc.id_seccion and h.id_hosp=sh.id_hosp and
p.id_pac=h.id_pac and m.id_medico=sh.id_medico and
ce.id_esp=m.id_esp and cn.id_niv=p.id_niv
```

```
end
```

```
ALTER procedure [dbo].[dwh_proc_consulta] as
declare @fecha1 varchar(20), @fecha2 varchar(20)
declare @fecini datetime, @fecfin datetime
begin
```

```
--obteniendo la fecha
select @fecha1=cast(getdate() as varchar(20))
set @fecha1=@fecha1 + ' 00:00:00 a.m.'
set @fecini=cast(@fecha1 as datetime)
select @fecha2=cast(getdate() as varchar(20))
set @fecha2=@fecha2+' 11:59:59 p.m.'
set @fecfin=cast(@fecha2 as datetime)
```

```
--insertando los datos en la tabla de estadística
INSERT INTO estadistica..consultas
select distinct c.idconsulta,
cast(cast(datepart(yy,c.dfecha) as varchar(5)) + '/' +
cast(datepart(mm,c.dfecha) as varchar(5)) + '/' +
cast(datepart(dd,c.dfecha) as varchar(5)) as datetime) as fecha,
p.cNombre + ' + p.cPaterno + ' + p.cMaterno as medico,
ce.cDescripcion as especialidad,
cast(cast(datepart(hh,c.dfecha) as varchar(20)) + ':' +
cast(datepart(n,c.dfecha) as varchar(20)) + ':' +
cast(datepart(ss,c.dfecha) as varchar(20)) as datetime) as hora,
co.cdescripcion as consultorio,a.NSUBSECUENTE,cp.cSexo,
cc.cDescripcion as diagnostico,m.cNombre as Municipio,e.cNombre as Estado,
td.cDescripcion as DerechoHabiencia,exped.IDExpediente as Expediente,
cp.cNombre + ' + cp.cPaterno + ' + cp.cMaterno as paciente
from Consultas c INNER JOIN Personal p ON p.IDPersonal=c.IDMedico
INNER JOIN Det_Personal_Especialidad dpe ON dpe.IDPersonal=p.IDPersonal
INNER JOIN ctl_Especialidades ce ON ce.IDEspecialidad = dpe.IDEspecialidad
LEFT JOIN hgc_agenda a ON a.cFolioAgenda=c.IdFolioAgenda
LEFT JOIN ctl_consultorios co ON
co.CCVE_CONSULTORIO=a.CCVE_CONSULTORIO and
co.IDUMedica=a.IDUMedica
INNER JOIN Ctl_Pacientes cp ON cp.IDPaciente=c.IDPaciente
INNER JOIN Det_DiagnosticosConsulta dd ON dd.IDConsulta=c.IDConsulta
INNER JOIN Ctl_CIE10 ce ON ce.IDDiagnostico=dd.IDDiagnostico
INNER JOIN Det_DomicilioPacientes dp ON dp.IDPaciente=cp.IDPaciente
INNER JOIN Ctl_Municipios m ON m.IDMunicipio=dp.IDMunicipio and
m.IDEstado=dp.IDEstado and m.IDJurisdiccion=dp.IDJurisdiccion
INNER JOIN Ctl_Estados e ON e.IDEstado=m.IDEstado
INNER JOIN Det_Pacientes_Derechohabiencia pd ON
pd.IDPaciente=cp.IDPaciente
INNER JOIN Ctl_TipoDe00r0echohabiencia td ON
td.IDTipoDerechohabiente=pd.IDTipoDerechohabiente
INNER JOIN Expedientes exp ON exp.IDPaciente=cp.IDPaciente
where c.dFecha between DATEADD(day,-7,@fecha1) and @fecha2
order by c.idconsulta
end
```

Figura 4. Procedimientos almacenados del proceso de extracción y transformación de datos de Ingreso, Egreso y Consultas.

Estas etapas se consideran las más importante de la construcción del Data Warehouse ya que de ello depende que los datos cargados sean consistentes, seguros y resistentes, debido a esto se presenta el código fuente de los procedimientos almacenados que se encargan de realizar esta actividad.

Posteriormente se realiza la carga de datos mediante la implementación de jobs de SQL Server Enterprise los cuales fueron programados para almacenar los datos automáticamente cada día domingo de la semana. [5] Tomando en cuenta que una vez que se han cargado los datos en un Data Warehouse estos ya no deben ser

modificables, el almacenamiento se realiza con una semana de retraso. (ver Figura 5).

Para el acceso a los datos se construyó un Sistema de Información de Soporte a la Toma de Decisiones utilizando el lenguaje Visual Basic de la suite de Microsoft .NET en donde se generaron cada uno de los indicadores de productividad solicitados por la Comisión Coordinadora de Institutos Nacionales de Salud y Hospitales de Alta Especialidad (CCINSHAE). (ver Figura 6).

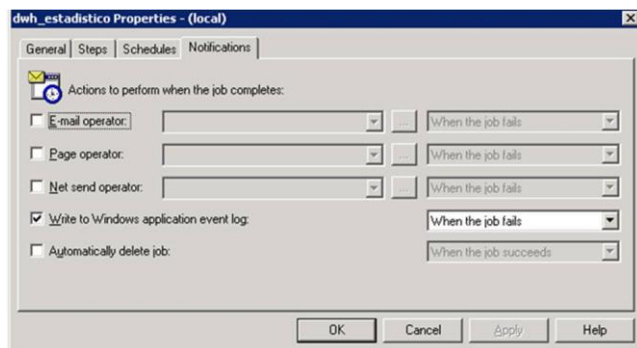
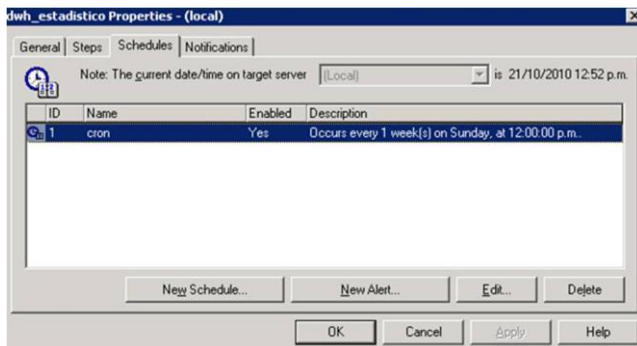
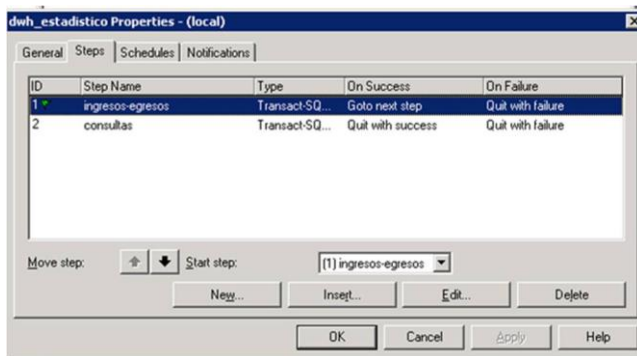
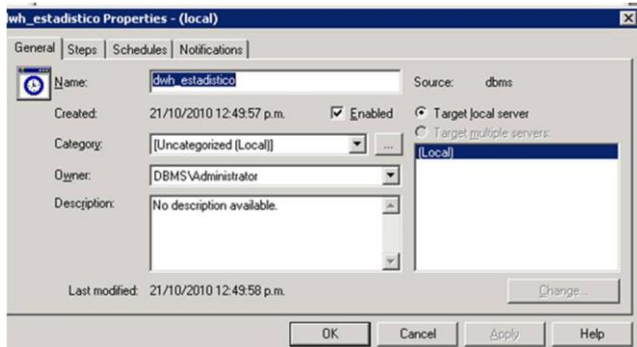


Figura 5. Pasos para la creación de jobs.

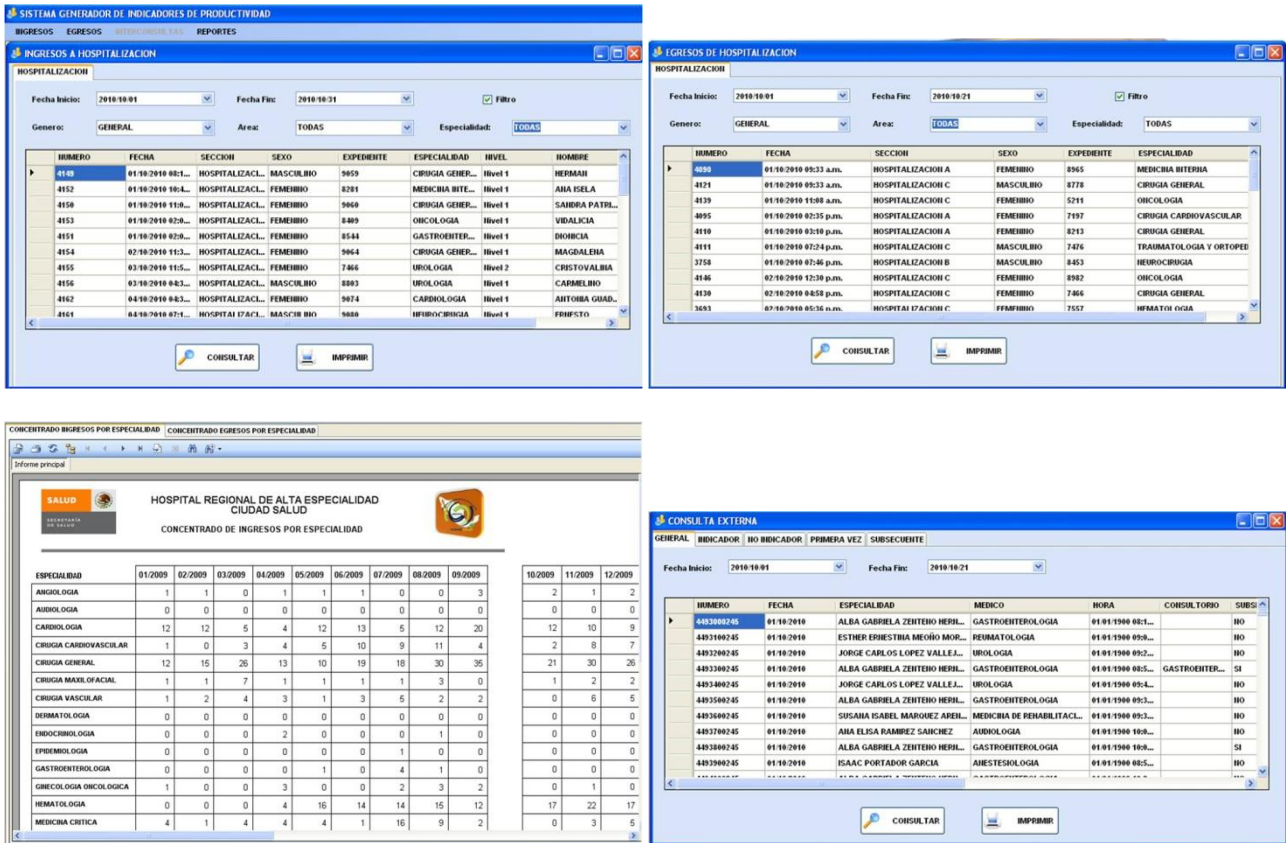


Figura 6. Sistema de Información de Soporte a la Toma de Decisiones en lenguaje Visual Basic .NET.

IV. Resultados

A continuación se presenta una tabla comparativa obtenida mediante una entrevista con los expertos estadísticos en donde se establecen las principales diferencias del proceso de generación de indicadores estadísticos que se tenía con el uso de hojas de cálculo y del actual mediante la implementación del Data Warehouse (ver Tabla 1).

TEMA	PROCESO ACTUAL	DATA WAREHOUSE
Recopilación de Datos	Manual y Automatizado(excel)	Automatizado
Problemática en recopilación de datos	* Registro Deficiente * Falta de *Responsabilidad de los Servicios	* Registro eficiente * Independencia de los servicios
Herramientas de análisis	Manual y Automatizado(excel)	Automatizado
Puntos Ciegos	Si detecta	Si detecta
Horas al día que les toma la recopilación, análisis y generación de información estadística	Más de 8 horas	Menos de 1 hora
Tiempo de generación de información para la dirección médica	De 6 a 10 días	1 día
Cantidad de indicadores estadísticos generados	35	35

Tabla1: Comparativa proceso anterior vs actual Data Warehouse.

En cuanto a la recopilación de datos actualmente todo se hace mediante la aplicación de acceso a datos construida a partir de la implementación del Data Warehouse lo cual elimina la problemática del registro ineficiente y la falta de responsabilidad de los servicios en la captura y entrega en tiempo y forma de la información. Así también mediante la herramienta de acceso a datos ya se permite la detección de puntos ciegos lo cuales eran detectados anteriormente pero hasta el cierre de información y entrega de reportes. Anteriormente la cantidad de tiempo invertida en el proceso de recopilación, análisis y generación de información era de 8 horas y actualmente es de 1 hora lo cual demuestra la eficiencia del acceso a datos. La generación de reportes para la Dirección Medica se hace en un solo día cuando anteriormente los estadísticos se tomaban de 6 a 10 días en generar un informe.

V. Conclusiones

La generación de indicadores de productividad es prioritaria dentro de los Hospitales de Alta

Especialidad ya que de eso depende determinar el comportamiento y la calidad de servicios médicos que se brinda a la población.

La implementación del Data Warehouse ha agilizado los procesos de recopilación, análisis y generación de información estadística, permite la entrega de información a tiempo a las áreas directivas y facilita la toma de decisiones ya que la consulta de los indicadores de productividad está disponible en tiempo real mediante la aplicación de acceso a datos.

En un futuro se pretende aplicar algoritmos de Minería de Datos (Datamining) para la extracción de conocimientos estadísticos o predictivos.

VI. Referencias

- [1] Farooq F., Mansoor S., **Real-Time Data Warehousing for Business Intelligence**, Diciembre 2010, Islamabad, Pakistán, Página 1, Consultado en ACM Digital Library.
- [2] Miranda L., Villalpando F., López G., M., López Uribe M. **Lineamientos para la Integración de carpetas de Junta de Gobierno de los Institutos Nacionales de Salud**. Secretaría de Salud (Enero 2002). Coordinación General de los Institutos Nacionales de Salud.
- [3] Kimball Ralph y Ross Margy (2002). **The Data Warehouse Toolkit: the complete guide to dimensional modeling**. Segunda Edición. 2002. Editorial Wiley. United States of America.
- [4] Harjinder G., Rao P., **Data Warehousing: la integración de información para la mejor toma de decisiones**, 1996, Prentice Hall.
- [5] Vieira R., **Professional SQL Server Programing**, Wiley Publishing, 2007, Indianapolis Indiana. Páginas 281-327, 699-735.

VII. Autores

M. en C. Karina Cancino Villatoro obtuvo su Maestría en Ciencias con especialidad en Base de

Datos. Es profesora de tiempo completo en la Universidad Politécnica de Tapachula. Analista y programador del Hospital Regional de Alta Especialidad Ciudad Salud.

M. en C. Vanessa Benavides García obtuvo su Maestría en Ciencias con especialidad en Base de Datos. Es profesora de asignatura de la Licenciatura en Sistemas Computacionales en la Facultad de Contaduría Pública del Campus IV en la Universidad Autónoma de Chiapas. Analista y programador del Hospital Regional de Alta Especialidad Ciudad Salud.

Pantalla Interactiva para Acceso a Material Didáctico Multimedia utilizando Mapas Conceptuales

Abraham Rodríguez-Jiménez, Alfredo Cristóbal-Salas, Silverio Pérez-Cáceres, Efrén Morales-Mendoza

Facultad de Ingeniería Electrónica y Comunicaciones. Universidad Veracruzana. Av. Venustiano Carranza Col. Revolución Poza Rica Veracruz México, 93390.

Correo electrónico: abramsk8@gmail.com, {acristobal, sperez, efmorales}@uv.mx

Resumen

Este trabajo visualiza a las TICs como potencial factor distractor de los usuarios en su búsqueda por el conocimiento. Esto se debe a que las TIC requieren conocimientos y habilidades técnicas que aún no se han desarrollado en todos los niveles educativos en México. Por eso, presentamos un sistema electrónico y computacional que permite a los usuarios acceder a contenidos educativos multimedia organizados como los mapas conceptuales. Nuestro sistema mejora la experiencia del usuario al no exigir, de este último, habilidades técnicas y que permite al usuario explorar el mapa conceptual utilizando solamente sus manos para encontrar el conocimiento deseado. En este trabajo, también presentamos los detalles técnicos del diseño y la implementación de este sistema así como los resultados obtenidos de la experimentación con los usuarios finales del prototipo desarrollado.

Palabras Clave: representación del conocimiento, manejo de documentos, educación de las ciencias, pantalla interactiva, activación por interrupción

I. Introducción

Hasta el día de hoy, la computación ha sido capaz de procesar y almacenar datos; sin embargo, con la consolidación de las Tecnologías de la información y las comunicaciones (TICs) como herramienta para transferir grandes cantidades de datos y la aplicación de técnicas de la minería en bases de datos se ha logrado que la computación también logre procesar y almacenar información.

Hoy en día, resulta importante encontrar algoritmos que permitan el procesamiento y el almacenamiento de conocimiento, entendiendo a este, como una generalización de la información o adaptación de la información a distintos contextos.

Este nuevo objetivo nos lleva a diseñar dichos algoritmos buscando definir el concepto conocimiento y definiendo nuevas formas de representarlo en formato digital para finalmente lograr su almacenamiento y recuperación.

El uso de tecnologías dentro del área educativa representa un reto constante debido a los cambios que la tecnología inyecta al proceso enseñanza-aprendizaje. Por ejemplo, si se quiere introducir el uso de la computadora como herramienta didáctica, los componentes de dicho equipo seguramente cambiarán en el corto plazo (nuevo software, nuevo hardware, cambios en las plataformas de administración de recursos). Estos cambios no permiten estabilizar la curva de aprendizaje de los usuarios debido a que cuando están por acostumbrarse al uso de una tecnología, ésta ha cambiado y está comenzando a ser reemplazada por otra nueva tecnología.

La representación gráfica del conocimiento es una herramienta didáctica valiosa para usuarios con poca experiencia en las TICs. Algunos ejemplos de representaciones gráficas de información que pueden derivar en conocimiento como lo son: los mapas conceptuales, mapas mentales, diagramas UML, redes semánticas, etc.

Todos ellos convergen en representar el conocimiento con símbolos particulares dependiendo de la relación que tienen los componentes.

El presente trabajo se sitúa en el área de electrónica debido a que se desarrolla un sistema interactivo computacional que hace uso de una malla láser para ayudar al usuario en la recuperación gráfica de la información.

II. La Interacción Humano-Computadora

Esta área de investigación busca desarrollar dispositivos que permitan la interacción entre el usuario final y los dispositivos electrónicos de una manera precisa, concreta, conceptualizada y contextualizada.

Actuales investigaciones en esta área describen formas innovadoras de interacción como por ejemplo en el artículo [13], Ríos propone un sistema de comunicación para cuadripléjicos, el cual funciona a través de un sensor colocado en una estructura de anteojos; este sensor cuenta el número de veces que el paciente parpadea para avisar a la enfermera a través de una pantalla LCD de la necesidad que este tiene. A su vez, Phung, en [11] propone el uso de un sistema que procesa una señal visual digital (DSP) para detectar la posición de la esquina del ojo, el centro de la pupila y luego calcular el movimiento de los ojos.

De igual forma, Deniz en [3] describe una cabeza de robot multimodal que puede detectar los rostros, el movimiento facial, la selección de la acción y la localización del sonido. Los resultados preliminares indican que la cabeza del robot potencialmente puede alcanzar la interacción humana y la asistencia.

Por su parte, en [18] se menciona el diseño de un lápiz óptico el cual es seguido en tiempo real para interactuar con un entorno virtual.

Similarmente, en [12] se menciona un sistema que comprende el seguimiento de la cabeza con seis grados de libertad que calcula la trayectoria del movimiento de la cabeza del usuario para

después juzgar el nivel de coincidencia con una de las interacciones predefinidas.

En área musical, [10] propone una interacción humano-computadora basada en el reconocimiento automático de dos niveles de gestos musicales. El modelo se aplica a la guitarra clásica y la interacción se logra sin el uso de sensores. El dispositivo es capaz de reconocer el tipo de gesto que el guitarrista ha tocado (sonoridad de la nota) y los matices de este gesto (la fuerza al tocar la cuerda) con sólo analizar los sonidos producidos por el intérprete.

Con todo lo anterior, es evidente entonces que el área de interacción humano-computadora sigue activa; sobre todo, con la aparición de dispositivos innovadores como Wiimote®, el Kinect®, el TouchPad multipunto.

Sin embargo, en la vida cotidiana se requiere de dispositivos que soporten alto tráfico de usuarios pero a la vez que cuiden aspectos más del corte social como por ejemplo: la higiene y la transmisión de enfermedades vía contacto directo con los dispositivos. Además, que los usuarios requieren de dispositivos para consulta personalizada de datos en tiempo real. Para ello, es necesario desarrollar dispositivos que les permitan interactuar con repositorios distribuidos de información.

III. Problemática

Esta investigación está centrada en el área educativa a la cual se pretende dotar de un dispositivo que permita la interacción entre los estudiantes y material didáctico multimedia almacenado en una red de repositorios los cuales se actualizan en tiempo real.

Como requerimientos principales tenemos la preservación de la salud de los usuarios puesto que tenemos dos situaciones de alto riesgo: usuarios con pobres hábitos de higiene y un alto uso del sistema.

Además, el costo de mantenibilidad del dispositivo juega un papel importante cuando se trata de usuarios con bajo presupuesto como podrían ser las instituciones educativas.

En [2] y [16] se presenta la idea de usar mapas conceptuales para representar el conocimiento; esto debido a la facilidad de expresar un tema en conceptos y la facilidad de expresión matemática del mapa conceptual como un grafo dirigido y etiquetado.

Los mapas conceptuales [1]-[9], [14], [15], [17] son un recurso para representar de manera gráfica conceptos en los cuales se asientan significados en forma de proposiciones simples.

Los mapas ofrecen un medio adicional para obtener información o conocimiento tomando en consideración la esencia de las partes que componen al nuevo conocimiento; así como, los componentes y las relaciones entre ellos.

Los mapas conceptuales permiten representar relaciones significativas entre los conceptos en forma de relaciones. Además, ayudan a que se refuerce el conocimiento al almacenar la información en la memoria de largo plazo.

Esta investigación considera que los mapas conceptuales pueden resultar en una interfaz más intuitiva para interactuar con bases de conocimiento mediante dispositivos electrónicos.

En la siguiente sección se muestran detalles del sistema electrónico y computacional necesario para cumplir los requerimientos del usuario final.

IV. Diseño de prototipo

El sistema electrónico y computacional desarrollado cae en la categoría de retro-proyección y es un sistema interactivo de notoque sino de activación por interrupción en una malla de 13x7 puntos generados por el cruce de rayos de luz láser (ver Fig. 1).

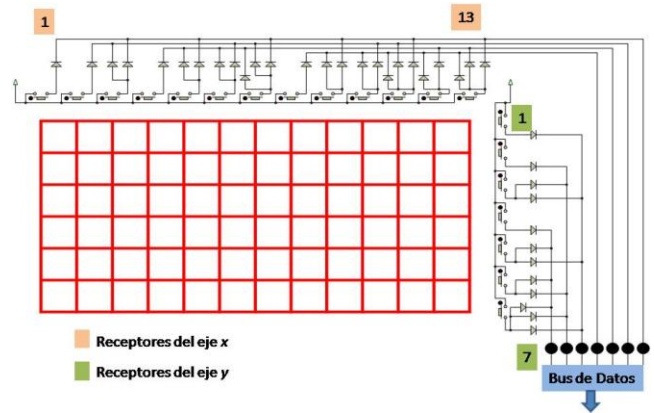


Fig. 1. Diseño de malla láser y bus de comunicación con la computadora.

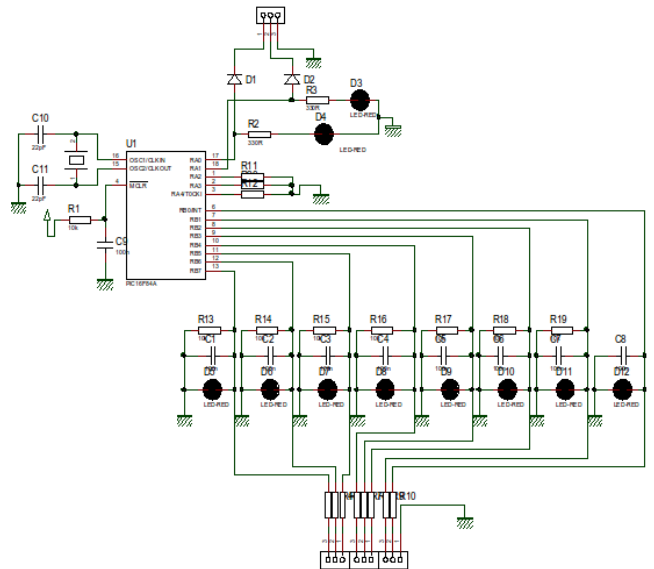


Fig. 2. Diseño de transductor de comunicación paralela a serial usando el microcontrolador PIC 16F84A

El prototipo consta de tres módulos: malla láser, tarjeta de adquisición de datos y comunicación RS-232.

El módulo *malla láser* recibe las interrupciones en la comunicación entre el emisor (haz láser) y el receptor (fototransistor) realizadas por el usuario y entonces ocurre la activación del circuito indicando las coordenadas (x,y) de la *malla láser* donde el usuario requiere de interacción. Esta coordenada es transmitida a la computadora a través de la *tarjeta de adquisición de datos* presentada en la Fig. 2. en el que se reciben 8 bits (4 bits por cada eje

coordinado) y son transferidos a la computadora de manera serial mediante *comunicación RS-232*.

El motor de búsqueda de material didáctico multimedia basado en mapas conceptuales tiene un menú de opciones de interacción disponibles para el usuario los cuales son: (a) mover el mapa conceptual a la izquierda, derecha, arriba y abajo, (b) regreso a la posición original del mapa y (c) seleccionar un concepto o documento multimedia. (d) Cerrar el documento multimedia y regresar al mapa conceptual (ver Fig. 3).

En la Fig. 4, se presenta una fotografía de la implementación de la pantalla interactiva que contiene la malla láser de 91 coordenadas. Mientras que en la Fig. 5. se muestra a un usuario utilizando la pantalla interactiva. Como es posible notar, el sistema es 'solo consulta'; esto es, el usuario solo puede ver el mapa conceptual y navegar en él, mas no puede modificarlo. Además, el usuario no interactúa con el sistema computacional que lo administra por lo que el uso del sistema resulta transparente.



Fig. 3. Captura de pantalla de interacción con base de conocimiento desplegado en la malla láser



Fig. 4. Implementación de la malla láser interactiva

Finalmente, nuestro diseño contempla que el usuario debe contextualizar la información contenida en los documentos multimedia presentados en el mapa conceptual. Es por ello que el sistema no tiene un motor de búsqueda de documentos; sino más bien, requiere que el usuario navegue en el mapa conceptual.

V. Resultados experimentales

El sistema electrónico fue sometido a pruebas técnicas de rendimiento con variación de temperatura y con variación de luminosidad. Como era de esperarse los dispositivos de generación de luz láser se calientan con el paso del tiempo lo que hace disminuir la intensidad de la luz y produce errores en la tasa de reconocimiento. También, se observó que el reconocimiento disminuye cuando la temperatura ambiente sobrepasa los 35 grados centígrados esto debido a que se acelera el calentamiento de los generadores de luz láser lo que disminuye su intensidad de luz generada y dificulta el reconocimiento de coordenadas.

También, no es de sorprender que pueden surgir falsos positivos en el reconocimiento de coordenadas producidos por la luz ambiental la cual se intensifica al mediodía.

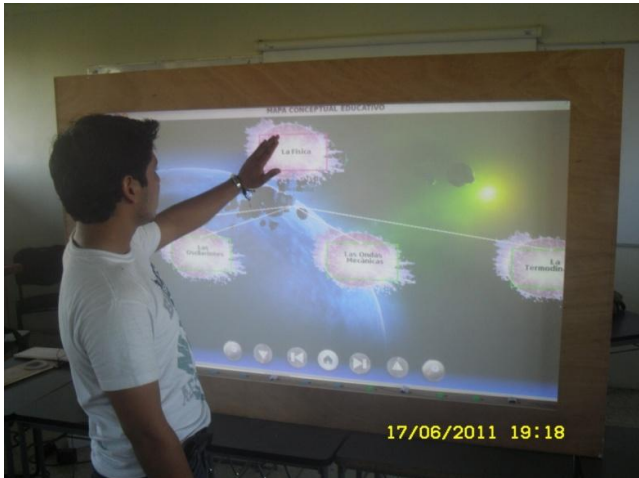


Fig. 5. Usuario interactuando con mapa conceptual

Los resultados experimentales nos muestran que el dispositivo electrónico requiere de no más de 40 lúmenes de luz ambiental para evitar falsos reconocimientos de coordenadas.

También, se llevaron a cabo experimentación con 28 Estudiantes de licenciatura de la Universidad Veracruzana Campus Poza Rica - Tuxpan quienes respondieron un cuestionario orientado a explorar la opinión de los estudiantes sobre el uso del prototipo desarrollado.

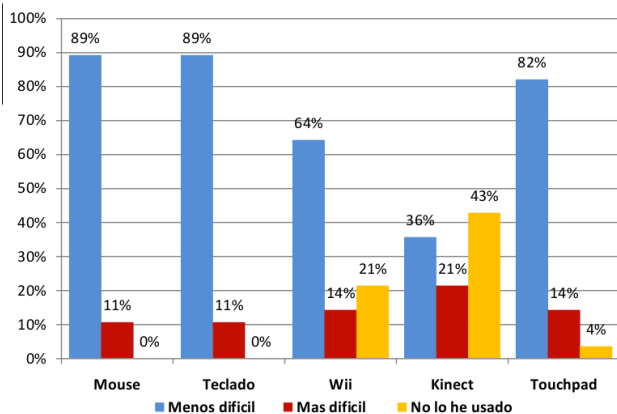


Fig. 6. Comparación de la dificultad entre el uso de la pantalla interactiva con las diferentes herramientas de Interacción computacional.

En la Fig. 6., se muestran los resultados de la comparación que se les pidió hacer a los usuarios de la pantalla interactiva con respecto a otros dispositivos de interacción con la computadora como el ratón, teclado, Wii Remote®, Kinect® o el Touchpad multipunto. En esta figura se puede

notar que la mayoría de los usuarios prefiere el uso de dispositivos simples de interacción.

VI. Conclusiones

De acuerdo los resultados de las pruebas técnicas aplicadas al prototipo podemos decir que la tasa de detección de interacciones del sistema electrónico no varía en condiciones de luminosidad promedio existente en un aula; los resultados muestran que la pantalla funciona correctamente en un rango de 0 a 40 lúmenes. También, los resultados de las pruebas nos arrojan que la tasa de reconocimiento no varía en los rangos de temperatura promedio existentes en un aula (24 a 31 grados centígrados).

De las pruebas de usabilidad del prototipo aplicadas junto con el usuario final del sistema se tiene que los estudiantes consideran más fácil la interacción cuando utilizan el sistema electrónico basado en una malla láser que cuando utilizan otros dispositivos de interacción humano-computadora.

Por otro lado, se tuvo cuidado de desarrollar el prototipo con materiales adquiribles en la zona de aplicación de la pantalla interactiva para aumentar la mantenibilidad del prototipo y reducir su costo el cual está por debajo de los 500 dólares; por lo que, podemos considerar al prototipo como de bajo costo cuando se le compara con equipos similares.

VII. Agradecimientos

Este trabajo de investigación es financiado, en parte, por la Universidad Veracruzana a través de la Facultad de Ingeniería en Electrónica y Comunicaciones y a través de la “Convocatoria Interna de Apoyo al Desarrollo de los Cuerpos Académicos 2010” y a través de la Dirección General de Investigaciones con el proyecto: “Interacción con Repositorios de Documentos Multimedia usando Mapas Conceptuales”. Así mismo, los autores agradecen a la Secretaría de Educación Pública que por medio de su convocatoria: “Apoyo a la Incorporación de Profesores de Tiempo Completo 2010” quien

apoyó al financiamiento de este trabajo de investigación a través del proyecto “Documentación digital y colaborativa del aprendizaje para el fortalecimiento de las competencias académicas” con número de registro #47676. Cualquier opinión, descubrimiento, conclusión o recomendación expresada en este trabajo de investigación es responsabilidad de los autores y no necesariamente reflejan el punto de vista de la Universidad Veracruzana o la Secretaría de Educación Pública.

VIII. Referencias

- [1] Alarcón, P., Sánchez, J. (2004). Diseñador de mapas conceptuales: una herramienta implementada con y para el usuario final. En libro “Concept Maps: Theory, Methodology, Technology Proc. of the First Int. Conference on Concept Mapping”. A. J. Cañas, J. D. Novak, F. M. González, Eds. Pamplona, Spain 2004
- [2] Cristóbal-Salas, A. Peralta-Ruiz, M., Cristóbal-Salas, A. “Sistema Computacional para el Almacenamiento y Recuperación de la Documentación del Aprendizaje”. XXIV Simposio Internacional de Computación en la Educación. Xalapa de Enríquez, Veracruz, México. 25-29 de octubre del 2008
- [3] Deniz, O., Castrillon, M., Lorenzo, J., Guerra, C., Hernandez, D., Hernandez, M. (2002). CASIMIRO: a robot head for human-computer interaction. Robot and Human Interactive Communication, Proceedings. 11th IEEE International Workshop on, 319- 324.
- [4] Edwards, J., Fraser, K. (1983). Concept maps as reflections of conceptual understanding. Research in Science Education, 13, pp. 19-26
- [5] Macnamara, J. (1982). Names for things: A study of human learning. Cambridge, MA: M.I.T. Press.
- [6] Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (2000). Assessing science understanding: A human constructivist view. San Diego: Academic Press.
- [7] Novak, J. D. (1990). Concept maps and vee diagrams: Two metacognitive tools for science and mathematics education. Instructional Science, 19, 29-52.
- [8] Novak, J. (1998). Learning, Creating, and Using Knowledge, Concept Maps as Facilitative Tools in Schools and Corporations. New York: Lawrence, Erlbaum Associate, Inc.
- [9] Novak, J.D., Cañas, A.J. (2006). The theory underlying concept maps and how to construct them. Technical report CmapTools 2006-01. Institute for Human and Machine Cognition. pp. 31.
- [10] Orio, N., (1999). A model for human-computer interaction based on the recognition of musical gestures. Systems, Man, and Cybernetics, IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on. 333-338
- [11] Phung, M.D., Quang, V. T.; Hara, K.; Inagaki, H.; Abe, M. (2008). Easy-setup eye movement recording system for human-computer interaction. Research, Innovation and Vision for the Future. IEEE International Conference on, 292-297.
- [12] Qiao, Y., Xiaofang, X., Tao, S., Yongsheng, L., (2008) "A Design of Human-Computer Interaction Based on Head Tracker," Computational Intelligence and Industrial Application. PACIA '08. Pacific-Asia Workshop on. 718-721.
- [13] Ríos C. (2011). “Interfaz de Comunicación amigable para rehabilitación de cuadripléjicos”. Tomado del <http://www.fundacionypf.org.ar/educacion/05/fypf12.html>. Última consulta: 1 de septiembre del 2011.
- [14] Sánchez, J. (1993). Concept mapping and educational software production in science.

Annals of Presented Papers, National Meeting of the National Association for Research in Science Teaching. Atlanta, Georgia, USA.

- [15] Sánchez, J. and Mallegas, A. (1994). Cognitive maps human-computer interface design tools for learning. In Baeza-Yates, R. y Manber, U. (Editors), Computer Science 2: Research and Applications, Nueva York: Plenum Press, 387-397.
- [16] Sánchez-Escobar, F.J., Cristóbal-Salas, A., Pérez-Cáceres, S., Morales-Mendoza, E. “Sistema Computacional y Electrónico para Acceso a Material Didáctico Multimedia utilizando Mapas Conceptuales”. En memorias del I Congreso Internacional de Ingeniería Electrónica y Computación. 22 al 24 de Junio del 2011, Minatitlán Veracruz, México.
- [17] Sowa J. F. (2006) Concept mapping. In proceedings of the 2006 Technology, Instruction, Cognition and Learning (TICL) at the AERA Conference, San Francisco, 10 April 2006.SEP
- [18] Wang, Y., Tie. (2009) Toward natural and efficient human computer interaction. IEEE International Conference on Multimedia and Expo. ICME 2009, pp. 1560-1561.

IX. Biografías

Abraham Rodríguez Jiménez. Estudiante de la Facultad de Ingeniería en Electrónica y Comunicaciones de la Universidad Veracruzana. Actualmente trabajando en su proyecto de tesis titulada: “Sistema Electrónico para Interacción Humano Computadora utilizando una Malla Láser”.

Alfredo Cristóbal-Salas. Maestría y doctorado obtenido del Centro de Investigación y de Educación Superior de Ensenada en 1997 y 2002 respectivamente. Posdoctorado en la Escuela Henry Samueli de la Universidad de California,

Irvine, USA. Actualmente profesor de tiempo completo en la Facultad de Ingeniería en Electrónica y Comunicaciones de la Universidad Veracruzana.

Silverio Pérez-Cáceres. Maestría en administración por la Universidad Veracruzana y candidato a doctor por el Instituto Veracruzano de Educación. Actualmente profesor de tiempo completo de la Universidad Veracruzana.

Efrén Morales-Mendoza. Maestría en ciencias por la Universidad Veracruzana y candidato a doctor por el Instituto Veracruzano de Educación. Actualmente director de la Facultad de Ingeniería en Electrónica y Comunicaciones.

La Háptica como Complemento a la Realidad Virtual para el Aprendizaje de un Idioma Extranjero

*Héctor Hugo Cortés Dueñas, **Miguel Ángel García Ruiz, ***Ricardo Acosta Díaz

*Universidad de Colima, Av. Universidad 333 C.P. 28040, (312)3161000, hugo_cortes@ucol.mx

**Universidad de Colima, Av. Universidad 333 C.P. 28040, (312)3161000, mgarcia@ucol.mx

***Universidad de Colima, Av. Universidad 333 C.P. 28040, (312)3161000, acosta@ucol.mx

Resumen

La Realidad Virtual ha demostrado ser una herramienta bastante útil en el proceso de aprendizaje de distintas disciplinas. Este artículo de investigación se enfoca en la construcción de un dispositivo háptico que funge como un controlador, el cual es un complemento a un entorno de Realidad Virtual que tiene por nombre Realtown.

Estudios diversos concluyen que el involucrar el sentido háptico (sentido del tacto de manera activa) de personas de un rango de edades muy amplio afecta de manera positiva la forma en que un individuo aprende, es por eso que en este artículo se propone la combinación de un Entorno de Realidad Virtual con un dispositivo háptico que retroalimenta a través de vibraciones a un grupo de estudiantes del idioma inglés de la Facultad de Telemática de la Universidad de Colima.

El proceso de diseño y desarrollo del dispositivo háptico, la interacción con el entorno virtual Realtown, así como los resultados que se han obtenido hasta el momento por parte de los estudiantes del idioma inglés, se describen de manera concisa en este trabajo.

Palabras clave: Entorno Colaborativo (multi-usuario) de Realidad Virtual, Háptica, Proceso de Aprendizaje, Realidad Virtual.

I. Introducción

Se define como Realidad Virtual a un entorno que es creado por computadora, en el cual uno o varios usuarios tienen la posibilidad de adentrarse de forma virtual (inmersión) por medio de herramientas distintas y, en donde además, los usuarios son capaces de navegar por su entorno y manipularlo [15]. La Realidad Virtual puede ser utilizada para promover el aprendizaje de un idioma extranjero, ya que permite una simulación que se asemeja a situaciones que ocurren en el mundo real [16]. Además, un entorno virtual presenta una variedad de estímulos visuales, auditivos y kinestésicos, lo cual provoca que un usuario involucre más de un sentido al momento de interactuar con un mundo virtual [5].

Si a los estímulos visuales, auditivos y kinestésicos añadimos un nuevo sentido, como lo es el háptico, que consiste en emplear el sentido del tacto de manera activa, es decir, con el

objetivo de obtener información del objeto que se esté tocando; la experiencia, el nivel de concentración y la atención prestada aumentan de manera significativa dando como resultado una mejora en el proceso de aprendizaje de una actividad [1]. El aprendizaje del idioma inglés representa la actividad en la que decidimos observar de qué forma y hasta qué grado un estudiante es capaz de mejorar su desempeño escolar.

Para que la medición cuantitativa y cualitativa de la mejora del aprendizaje del idioma inglés pueda ser llevada a cabo, se decidió utilizar el entorno de Realidad Virtual Realtown, además, se está desarrollando un dispositivo háptico que servirá como un controlador que permita a los usuarios navegar por Realtown. Es necesario también emplear herramientas de evaluación con el fin de obtener respuestas precisas de los usuarios y de expertos en las áreas de Realidad Virtual y

háptica. Las evaluaciones seleccionadas para este proyecto de investigación se describen en secciones posteriores de este artículo.

II. El proceso de aprendizaje de un idioma extranjero a través de la Realidad Virtual

La mejor manera para aprender un idioma es por medio de la participación en una comunidad cuyo lenguaje objetivo (en este caso el inglés) sea usado para comunicarse en un contexto real y cotidiano [17]. De esta forma es como cualquier ser humano aprende su idioma original (natal). Por consiguiente, un individuo se ve motivado y hasta cierto punto obligado a pensar, hablar y escribir en el idioma en el cual su familia, amigos y demás personas que lo rodean se comunican.

En una comunidad con estas características –que el lenguaje objetivo sea usado para comunicarse cotidianamente-, una persona no puede escapar de la necesidad de comunicarse, dando por consiguiente el aprendizaje secuencial de vocablos, palabras, frases e ideas completas en última instancia [14]. Dicho esto, se observa que una persona que necesita aprender un idioma por vez primera, se ve inmerso en un contexto natural, multi-sensorial y significativo culturalmente hablando, de modo que el lenguaje se adquiere de manera espontánea [17].

Es precisamente el fenómeno de la inmersión, el que le brinda tanta riqueza en estímulos a un entorno de Realidad Virtual, ya que dentro de un mundo que no existe realmente, una persona de hecho siente como si estuviera en un entorno concreto debido a la interacción que proviene directamente del entorno virtual [12]. Tanto la inmersión como el contexto multi-sensorial en Realidad Virtual son muy importantes, ya que los estudiantes obtienen estímulos en diferentes sentidos dentro del mundo virtual y, esto propicia de manera significativa el aprendizaje de acuerdo a ciertas teorías constructivistas [5].

III. Realtown, un entorno de Realidad Virtual orientado al aprendizaje del inglés como idioma extranjero

Partiendo de la premisa de que la Realidad Virtual fomenta e impulsa el aprendizaje de una actividad, fue que el Profesor de la Facultad de Telemática de la Universidad de Colima, Miguel Ángel García Ruíz, en conjunto con otros colaboradores, desarrolló Realtown, un entorno de Realidad Virtual que consiste en una ciudad convencional donde existen edificios, escuelas, bancos, farmacias, y otras construcciones que se pueden encontrar en una ciudad real [3].

Asimismo, en Realtown es posible escuchar sonidos comunes como tráfico vehicular, niños jugando y conversaciones que ocurren en inglés dentro de los edificios y en las escuelas, de esta forma, la riqueza visual y auditiva que el usuario experimenta al interactuar con Realtown, le da una sensación de inmersión muy realista.

Una característica importante de Realtown es el hecho de que es un Entorno Colaborativo (multi-usuario) de Realidad Virtual (CVRE, por sus siglas en inglés).

Un CVRE consiste en un entorno virtual compartido, en donde los usuarios pueden “conocerse” y comunicarse a través de un chat, en tiempo real, a través de gestos y, a su vez tienen la posibilidad de navegar por todo el entorno [2].

Realtown se implementa de la misma forma en que se ejecuta un CVRE, es decir, la aplicación se almacena en un servidor y se puede acceder a ella utilizando una computadora de escritorio, laptop o PDA a través de una red local inalámbrica (WLAN) o una red de cobertura amplia (WAN) dependiendo de los protocolos de red y de la configuración que tenga el equipo empleado [5].

La Figura 1 muestra dos laptops ejecutando de manera simultánea el entorno virtual Realtown.



Fig. 1. Dos computadoras portátiles ejecutando el Entorno Colaborativo de Realidad Virtual (CVRE) Realtown. Adaptado de [5].

Los dispositivos por los que se accede a Realtown pueden ser un Mouse convencional, un trackball y un joystick de juego genérico, sin embargo, se demostró a través de un estudio con estudiantes que la manera más sencilla de navegar por Realtown es usando las flechas de un teclado convencional.

Háptica, elemento importante en cualquier proceso de aprendizaje

El término háptica fue introducido por primera vez en 1931 por G. Revesz. El origen de la palabra nos lleva al término griego *haptikos*, que quiere decir “capaz de tocar” [13]. Hoy en día, en un sentido general, el término háptica engloba el estudio del sentido del tacto y la interacción del ser humano con el entorno externo por medio de este sentido [10]. Cualquier dispositivo con interfaz háptica tiene la capacidad de proveer un intercambio simultáneo de información entre el usuario y una aplicación por computadora [11].

En la Figura 2 se observa la interacción entre un usuario y un mundo virtual, utilizando un dispositivo háptico como controlador.

El entorno Realtown es definitivamente una aplicación por computadora, en este mundo virtual, un estudiante puede representarse a sí mismo a través de un avatar (término que en la

mitología Hindú significa encarnación de un dios), que en el contexto de Realidad Virtual se refiere a una personificación gráfica que puede transmitir los gestos del usuario real, así como su voz en tiempo real.

Como ya se mencionó, la riqueza en estímulos visuales y auditivos en el entorno Realtown es interesante; sin embargo, debido a investigación ya existente que concluye que el sentido háptico de una persona es importante para el proceso de aprendizaje, se decidió construir un dispositivo háptico a manera de controlador que pueda sustituir al teclado, Mouse convencional y joystick para navegar en Realtown y, que además brinde retroalimentación háptica para que el aprendizaje del idioma inglés se refuerce y complemente, ya que haciendo esto se involucra el sentido del tacto activo y, si lo sumamos a los sentidos de la vista y el oído, la experiencia del usuario, así como su nivel de inmersión, se enriquece aún más.

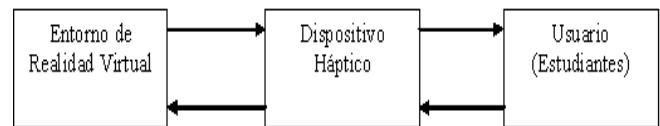


Fig. 2. Representación del intercambio bidireccional de información que es exclusivo de interfaces hápticas en combinación con un entorno de Realidad Virtual. Adaptado de [11].

Una razón importante para decidir incluir el sentido háptico en el entorno Realtown es que a través del tacto activo, una persona no solo obtiene sensaciones, sino que además obtiene información valiosa que le ayuda a inferir y tomar decisiones incluso si no está utilizando los sentidos de la vista y el oído [7].

IV. Dispositivo háptico propuesto para evaluar la mejora de desempeño en el aprendizaje del idioma inglés

Un hecho que sirve como justificación para este artículo de investigación es que personas de diferentes rangos de edades invierten una considerable cantidad de tiempo jugando videojuegos [3].

Según un estudio de la compañía Microsoft que lleva por nombre “Play Smart, Play Safe”, el 75% de padres de familia de varios países de Europa y Estados Unidos de Norteamérica, consideran que los juegos de vídeo pueden ser benéficos para sus hijos y para la familia en general, ya que por lo regular varios miembros de la familia (y no solo los más jóvenes) comparten el gusto por los videojuegos; sin embargo, solo un 50% de los padres de familia señala estar en control de la cantidad de tiempo que sus hijos emplean para jugar juegos de vídeo y, más del 60% de padres aceptan que el tiempo de juego supera los sesenta minutos por día.

Esta situación sirve como justificación ya que el entorno Realtown es en muchos aspectos un videojuego, por lo que si los estudiantes se ven atraídos de la misma forma hacia esta herramienta educativa, el tiempo que invertirán para “jugar” y estudiar a la vez será notablemente mayor, por lo que esperar una mejora en el desempeño del aprendizaje del idioma inglés es lógico hasta cierto punto.

En otro orden de ideas, el uso de controles de videojuego que vibran es relativamente nuevo, por lo que para los estudiantes sigue siendo una novedad, de manera que al combinar un “videojuego” como lo es Realtown con una retroalimentación háptica como lo es una vibración (este proceso se describe a continuación), los resultados que arrojen los estudios que se llevarán a cabo pueden ser muy interesantes, por lo que bien valdría la pena compartirlas en un futuro.

Con el objetivo de involucrar el sentido háptico de los estudiantes del idioma inglés en la Facultad de Telemática de la Universidad de Colima, se decidió crear un dispositivo que brinde retroalimentación háptica a los usuarios del entorno de Realidad Virtual Realtown. Esta retroalimentación será dada a los estudiantes a través de vibraciones de diferente intensidad. Es conveniente señalar que el dispositivo funciona únicamente cuando el usuario (por medio de un avatar) entra en alguna de las escuelas de

Realtown. Las actividades que tienen lugar en las escuelas, como las conversaciones que tienen lugar con los alumnos o profesores, así como el dar respuesta a preguntas que se muestran de manera visual al usuario, generan el funcionamiento del dispositivo háptico.

Un ejemplo sencillo de la manera en que trabaja el dispositivo es que cuando se muestre una pregunta de un tema determinado de inglés (como puede ser conjugación de verbos, estructuras gramaticales de acuerdo a los tiempos presente, pasado o futuro, entre otras), el usuario debe dar una respuesta a través del teclado, en caso de que la respuesta sea errónea, el dispositivo vibra en un nivel de intensidad bajo, además, se le muestra al usuario un mensaje visual y auditivo que le informa sobre el tema de inglés que debe reforzar para poder dar una respuesta correcta a la pregunta que falló en contestar.

El dispositivo háptico tiene tres niveles de vibración, conforme el estudiante ingrese respuestas incorrectas, la intensidad de vibración aumentará, este aumento va en función de la dificultad de las preguntas que se presenten, de manera que si el estudiante recibe la vibración más alta, significa que debe cambiar de tema, ya que el grado de dificultad de las preguntas irá en aumento, por lo que es más probable que el estudiante no responda de manera correcta las preguntas subsecuentes. De esta manera, además de dar retroalimentación háptica al usuario, se lleva un seguimiento del nivel en el que se encuentra el estudiante en un tema específico.

V. Desarrollo técnico del dispositivo háptico

El proceso de desarrollo del dispositivo háptico que complementa al entorno Realtown para el aprendizaje del idioma inglés se explica en esta sección.

La retroalimentación háptica que recibirán los usuarios del dispositivo háptico será a través de vibraciones, las vibraciones son generadas por un micro motor/vibrador que es capaz de generar

hasta ocho mil revoluciones por minuto, la intensidad de la vibración está determinada por la cantidad de voltaje que suministra una fuente de poder interna que genera un máximo de tres voltios. El voltaje de la fuente de poder es regulado por un controlador PIC 18F2550, el cual tiene una interfaz USB, de manera que el dispositivo puede ser conectado a través de cualquier puerto USB de la computadora que esté ejecutando el entorno Realtown.

La Figura 3 corresponde al controlador PIC 18F2550 y al micro motor/vibrador que se emplearon para desarrollar el dispositivo háptico.

De acuerdo al voltaje generado por la fuente de poder, el micro motor/vibrador (que es similar a los vibradores que utilizan los teléfonos celulares) trabaja a una cantidad diferente de revoluciones por minuto, entre mayor sean las revoluciones, mayor es la vibración que se produce.



Fig. 3. Controlador PIC 18F2550 y micro motor/vibrador empleados para desarrollar el dispositivo háptico.

Debido a que tanto el micro motor/vibrador como el controlador PIC 18F2550 trabajan con voltaje muy bajo, es necesario utilizar dos capacitores con el objetivo de proteger de altos voltajes (generados por la fuente de poder interna) al micro motor/vibrador y al controlador de voltaje.

Además de las consideraciones físicas, como por ejemplo, el material que se emplea para desarrollar el dispositivo háptico, así como los estudios de ergonomía realizados, es necesario elegir modelo de ciclo de vida de software que se adecue al proceso de fabricación del dispositivo a crear. Un ciclo de vida (en Ingeniería de Software) es un marco de referencia que contiene, los procesos, actividades y tareas involucradas en el desarrollo, la explotación y el

mantenimiento de un producto de software, abarcando la vida del sistema desde la definición de los requisitos hasta la finalización de su uso [6].

El modelo de ciclo de vida del software más apropiado para el desarrollo del dispositivo háptico propuesto en este artículo es el Modelo de Prototipado Evolutivo. Este modelo consiste en que el concepto del sistema (dispositivo háptico) se desarrolla conforme avanza el proyecto [9]. Una característica importante del Modelo de Prototipado Evolutivo es que se basa principalmente en la retroalimentación que los desarrolladores reciben por parte de los usuarios.

En la Figura 4 se observa el proceso del Modelo de Prototipado Evolutivo de manera gráfica.

El primer paso en el modelo de ciclo de vida Prototipado Evolutivo es el Concepto Inicial, es aquí donde se propuso la idea del dispositivo háptico de manera general, así como las funciones que éste llevaría a cabo. En el Diseño e Implementación del prototipo inicial se definió el material necesario para la construcción del dispositivo, el estudio de ergonomía también se realizó en este paso. Conforme se reciben observaciones y sugerencias por parte de los estudiantes respecto al dispositivo háptico, se realizan los cambios necesarios de manera iterativa y, es precisamente en este paso – refinación del prototipo hasta que sea aceptable– donde el dispositivo háptico propuesto en este trabajo se encuentra.

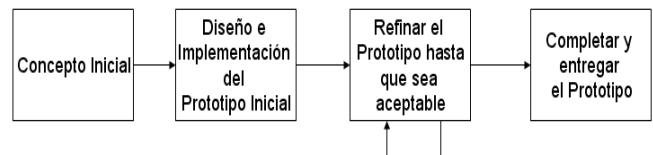


Fig. 4. Modelo de ciclo de vida Prototipado Evolutivo. Adaptado de [9].

Una vez que el dispositivo se termine de desarrollar se seguirá con el último paso del Modelo de Prototipado, que es el de Completar y entregar el prototipo. Cuando el modelo de ciclo de vida finalice, se comenzará con la evaluación

tanto del dispositivo háptico como del entorno de Realidad Virtual.

VI. Evaluación Heurística y Cuestionarios de Usabilidad

Para evaluar objetivamente al entorno de Realidad Virtual Realtown y al dispositivo háptico desarrollado para involucrar el sentido del tacto activo de los estudiantes del idioma inglés, se realizará una evaluación heurística. Esta evaluación consiste en solicitar a expertos en materia de Realidad Virtual y háptica que utilicen de manera intensiva ambas aplicaciones (Realtown y dispositivo háptico) y que den una opinión detallada de sus observaciones, las opiniones de los expertos pueden ser registradas de manera escrita o bien, por medio de entrevistas [4]. Para fines de este proyecto de investigación, se decidió escribir las opiniones de los expertos para que puedan ser analizadas posteriormente.

También, es recomendable utilizar el protocolo de pensar en voz alta (Think Aloud Protocol) con el objetivo de registrar las opiniones de los expertos al mismo tiempo en que están analizando Realtown y el dispositivo háptico [3]. Los puntos heurísticos a evaluar deben ser cuestiones relacionadas con las facilidad/dificultad de uso de las aplicaciones, la complejidad de las funciones que presentan tanto el dispositivo como el entorno virtual y la cantidad de tiempo que tuvieron que invertir para acostumbrarse al uso de ellos.

A pesar de que la evaluación heurística representa una buena herramienta para conocer si el dispositivo háptico y Realtown se desempeñan de manera adecuada, no es suficiente para conocer todos los aspectos de las aplicaciones. Para evaluar de una manera más completa tanto al dispositivo háptico como a Realtown, es conveniente aplicar cuestionarios de usabilidad a los estudiantes que utilicen ambas herramientas. Las preguntas que se harán deben estar relacionadas con la manera en que las aplicaciones se desempeñaron, también, se

deberá cuestionar de qué manera las vibraciones recibidas por los estudiantes estimularon de alguna manera sus sentidos y si esto afectó el tiempo de estudio y el esfuerzo para contestar más preguntas de manera correcta en las escuelas que están dentro del entorno virtual Realtown.

También preguntas relacionadas con la comodidad del uso del dispositivo háptico deben hacerse, para que en caso de que haya incomodidades al contacto con el dispositivo se realicen correcciones en el diseño del mismo. Cabe la aclaración de que para el diseño del dispositivo háptico se siguieron consideraciones y estudios sobre ergonomía con el fin de evitar pérdida de tiempo en un posible rediseño de la herramienta háptica debido a irregularidades físicas que incomoden al usuario.

VII. Grupos piloto de estudiantes para evaluar el dispositivo háptico

El dispositivo háptico que retroalimenta el sentido del tacto activo de los estudiantes se encuentra actualmente en la fase final de desarrollo. Una vez que se termine de construir, se planea utilizar dos grupos de estudiantes del idioma inglés de la Facultad de Telemática de la Universidad de Colima para su evaluación cualitativa y cuantitativa.

Uno de los grupos utilizará el entorno de Realidad Virtual Realtown junto con el dispositivo háptico, el otro grupo seguirá el proceso normal de aprendizaje de sus temas de inglés. Los temas que serán evaluados serán los mismos para ambos grupos, se analizará la mejora del aprovechamiento en función de la cantidad total de respuestas correctas de los alumnos en los temas específicos del idioma inglés.

Una vez que los datos en relación a la cantidad de respuestas correctas por alumno hayan sido recabadas, se procederá a aplicar cuestionarios (a los estudiantes) que tengan preguntas relacionadas directamente con la manera en que las vibraciones influyeron para que se esforzaran en tratar de responder correctamente las

preguntas conforme se fueron presentando. La evaluación de las respuestas será en un espacio de tiempo de dos semanas continuas, de manera que se pueda dar un seguimiento en la cantidad de respuestas correctas cada vez para así observar la mejora en el desempeño, si es que en realidad hay una mejora.

VIII. Trabajo Futuro

Analizar la manera en que la combinación de Realidad Virtual y háptica promueve y mejora el aprendizaje del idioma inglés en estudiantes de la Facultad de Telemática de la Universidad de Colima es en cierta forma limitado. Creemos que es posible extender el estudio a más disciplinas, como puede ser el aprendizaje de las matemáticas o cualquier otra asignatura académica. También es factible realizar estudios similares a los que se proponen en este artículo en estudiantes de inglés de menor edad, de esta manera se puede conocer si existe alguna variabilidad en los resultados para así determinar si el estar en un nivel académico inferior o superior afecta el proceso de aprendizaje de los alumnos utilizando las mismas herramientas.

Es importante también que se consideren estimulaciones hápticas distintas a las vibraciones, ya que cabe la posibilidad de que con otro tipo de retroalimentación del sentido activo del tacto, los resultados en el aprovechamiento del aprendizaje del idioma inglés sean diferentes, por lo que se puede buscar un tipo de señal háptica que sea la óptima para esta actividad.

IX. Conclusiones

En un proceso de aprendizaje, entre más sentidos involucre un estudiante para interactuar con su objeto de enseñanza, la riqueza de estímulos y, por tanto la calidad del aprendizaje siempre tiende a mejorar [8]. En este trabajo de investigación se da a conocer un estudio en proceso sobre la combinación de una herramienta de Realidad Virtual con un dispositivo háptico (sentido activo del tacto) con el propósito de analizar la manera en que estudiantes de inglés de

la Facultad de Telemática de la Universidad de Colima mejoran su desempeño en dicha asignatura.

La investigación y documentación realizada hasta el momento nos han dado la pauta para esperar resultados que confirmen que el sentido háptico de una persona es muy útil para aprender alguna actividad. Una vez que este trabajo de investigación se termine, se darán a conocer los resultados obtenidos, con el propósito de compartir experiencias en dos temas que han sido muy poco tratados en México, que es la Realidad Virtual y la Háptica con un enfoque meramente educativo y no solo de entretenimiento pasajero.

X. Referencias

- [1] J. M. Alexander, K. E. Johnson, J. B. Schreiber. "Knowledge is not everything: Analysis of children's performance on a haptic comparison task". *Journal of Experimental Child Psychology*, 82, 2002, 341-366.
- [2] G. Burdea, P. Coiffet, *Virtual Reality Technology*. 2nd ed. with CD-ROM. 2003. John Wiley and Sons, New York.
- [3] H. Cortés, M. García, R. Acosta, P. Santana, "Diseño y Desarrollo de un Dispositivo Háptico con Aplicaciones para Entornos Educativos". En *Memorias de la Novena Conferencia Iberoamericana en Sistemas, Cibernética e Informática (CISCI 2010)*.
- [4] J. S. Dumas, J.C. Redish, *A Practical Guide to Usability Testing*. Revised ed. Intellect, Ltd. Exeter. 1999.England.
- [5] M. A. Garcia-Ruiz, A. Edwards, S. A. El-Seoud, R. Aquino-Santos, "Collaborating and learning a second language in a Wireless Virtual Reality Environment". *Int. J. Mobile Learning and Organisation*. Vol. 2. No. 4. 2008 pp.369-377.
- [6] ISO/IEC 12207: 1995, "Information Technology – Software life cycle processes". International Organization for

Standardization/International Electrotechnical Commission. 1995.

- [7] J. M Kennedy, P. Gabias, M. A. Heller, "Space, haptics and the blind". Geoforum. 23. 1992. p.175.
- [8] S. Loucks-Horsley, R. Kapitan, M. Carlson, P. Kuerbis, R. Clark, G. Melle, "Elementary school science for the '90s". Alexandria, VA: Association for Supervision and Curriculum Development. 1990
- [9] S. McConnell, Desarrollo y Gestión de Proyectos Informáticos: Cómo dominar planificaciones ajustadas de software, McGraw Hill/Interamericana de España. 1997. Madrid, España.
- [10] J. Minogue, M. G. Jones, (2006). Haptics in Education: Exploring an Untapped Sensory Modality. Review of Educational Research. 76 (3). 2006. pp. 317-348.
- [11] A. M.Okamura, C. Richard, M. R Cutkosky, "Feeling is believing: Using a force-feedback joystick to teach dynamic systems". Journal of Engineering Education, 92, 2002. pp. 345-349.
- [12] E. Prasolova-Forland, M. Divitini, "Supporting learning communities with collaborative virtual environments: Different spatial metaphors". Proceedings of the Second IEEE International Conference on Advanced Learning Technologies (ICALT 2002). 2002. pp. 259-264.
- [13] G. Revesz, The psychology and art of the blind. London: Longmans Green. 1950.
- [14] M. Saville-Troike, The ethnography of communication: An introduction. Cambridge, MA: Basil Blackwell. 1982.
- [15] W. R. Sherman, A. B. Craig, Understanding Virtual Reality, Morgan Kauffman, San Francisco, CA. 2003.
- [16] Y.C. Shih, "3D virtual immersion English learning experiences: College student views of their needs and challenges". Proceedings

of the Seventh International Conference on Multimedia Language Education of ROCMELIA.The Crane Publishing. 2003. pp. 385-394.

- [17] Y. C. Shih, M.T. Yang, (2008). A Collaborative Virtual Environment for Situated Language Learning Using VEC3D. Educational Technology & Society, 11 (1). 2008. pp. 56-68.

XI. Autores

Ing. Héctor Hugo Cortés Dueñas es Ingeniero en Telemática por la Universidad de Colima. Es Estudiante del 3er semestre de la Maestría en Computación en la Universidad de Colima

Dr. Miguel Ángel García Ruiz obtuvo su título de Doctorado en Ciencias de la Computación e Inteligencia Artificial por parte de la Universidad de Sussex, Inglaterra. Es Profesor-Investigador de tiempo completo en la Facultad de Telemática de la Universidad de Colima.

M. en C. Ricardo Acosta Díaz obtuvo su grado de Maestro en Ciencias por el Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Baja California. Es actualmente Director de la Facultad de Telemática de la Universidad de Colima.

OkAR: Generador de cargas académicas para la Ingeniería en Tecnologías de la Información de la Universidad Tecnológica de Izúcar de Matamoros

Ivan Antonio Flores Trujillo, **Patricia Mendoza Crisóstomo, ***Gustavo Soriano Jiménez, *Sergio Valero Orea, *****Alejandro Salvador Vargas**

Universidad Tecnológica de Izúcar de Matamoros, Prol. Reforma 168 B. Santiago Mihuacán, Puebla,

*2431109900, iflores@utim.edu.mx.

**2431083904, pmendoza@utim.edu.mx.

***2431014244, gsoriano@utim.edu.mx.

****2431012198, svalero@utim.edu.mx.

*****2431076084, asalvador@utim.edu.mx.

Resumen

Dentro de las instancias educativas es común el proceso de creación y planeación de horarios académicos trimestrales, cuatrimestrales, semestrales, etc., buscando siempre la forma más eficiente posible. En cada periodo académico, se debe asignar la programación de horarios tanto de alumnos como de docentes, a ellos mismos, de acuerdo a una serie de restricciones. Las soluciones construidas de manera manual, además de tener días o semanas de elaboración, no garantizan la remoción de bloqueos. La programación de horarios académicos es una caracterización del problema general de Timetabling (TTP). En este trabajo, se presenta una propuesta de solución a tal problemática basada en algoritmos evolutivos, concretamente, la implementación de un algoritmo genético, aplicada el Programa Educativo de Ingeniería en Tecnologías de la Información correspondiente a la Universidad Tecnológica de Izúcar de Matamoros.

Palabras Clave: algoritmo evolutivo, algoritmo genético, carga académica, problemática timetabling.

I. Introducción

Algunas instituciones académicas presentan dificultades para resolver sus problemas de programación horaria de acuerdo a diversas entidades (constantes y variables), como por ejemplo: *profesores, alumnos, asignaturas, aulas*, etc. Este problema es conocido en la literatura con el nombre de Timetabling, que consiste en la asignación de entidades o recursos escasos dentro de un número limitado de periodos de tiempo y cuya complejidad radica en las restricciones asociadas a los recursos.

La problemática de tipo Timetabling ha sido tratada de manera internacional como nacional, empleando técnicas como Búsqueda Tabú [1,2], Algoritmos Genéticos [3,4,5], Sistemas Expertos [6], Programación Celular [7], entre otras.

Con la presentación de este tipo de trabajos se puede ver que la mayoría de las soluciones son presentadas y no implementadas al cien por ciento como un sistema institucional, sino más bien como un proceso de investigación y comparativa entre diversas herramientas y algoritmos que se dicen resolver la problemática timetabling.

Así mismo, la gran diversidad de modelos para este problema se debe al hecho de que cada uno tiene que ser resuelto a partir de un conjunto de características muy específicas, que dependen de cada institución, lo cual obliga a considerar restricciones distintas en cada caso.

Por tal motivo y con el afán de contribuir al Sistema de Gestión de la Calidad de la Universidad Tecnológica de Izúcar de Matamoros, se propone un sistema informático

que permita automatizar y agilizar el proceso de generación de cargas académicas cuatrimestrales para el Programa Educativo de Ingeniería en Tecnologías de la información.

II. Estado del Arte

John F. Baquero, Eliana Mirledy y Ramón Alfonso, presentan un modelo matemático para el problema de asignación de salones basándose en búsqueda Tabú [1].

Joseph Gallart, Fernando Alva, Anthony Alama y Gissella Bejarano, generan un sistema SchedulePowerSoft para comparar los resultados arrojados por la aplicación de la heurística GRASP y la mejora del mismo con la Búsqueda Tabú [2].

Zulema Furones, Joaquín Pacheco y Alberto Aragón, plantean la problemática de calendarizar exámenes y la solución mediante Algoritmos Genéticos [3].

Bruno Lopez y Jaime Johnston presentan un modelo computacional capaz de encontrar la asignación óptima de clases, maestros y horarios utilizando un algoritmo genético [4].

Pedro Flores, Ernesto Brau, Jazmín Monteverde, Norman Salazar, José Figueroa, Eliseo Cadena y Caleb Lizárraga, construyen un modelo con sus respectivas restricciones y definen una representación vectorial de toda la información que supone simplifica la implementación de los Algoritmos Genéticos [5].

Pablo Gervás y Beatriz San Miguel enfrentan su sistema experto a distintas colecciones de datos de entrada como asignaturas con distinto número de créditos, distintos prerrequisitos y distinta colocación de asignaturas fijas [6].

F. Martínez realiza un modelo de la problemática en su institución y finaliza con una arquitectura aproximada al proceso de generación de horarios académicos, todo el trabajo está basado en la programación celular [7].

III. Desarrollo

El Programa Educativo (PE) de Ingeniería en Tecnologías de la Información de la Universidad Tecnológica de Izúcar de Matamoros (UTIM), al igual que el resto de los PE de la misma y que la mayoría de las instituciones educativas, lleva a cabo el proceso de generación de cargas académicas, de manera particular, lo realiza de forma manual y con ayuda de la aplicación Microsoft Excel.

A grandes rasgos, el proceso se realiza de la siguiente manera:

1. Las cargas académicas se realizan cuatrimestralmente de acuerdo a un mapa curricular.
2. La dirección del PE entrega a los profesores una posible carga académica acorde a su perfil.
3. El profesor por asignatura tiene la posibilidad de modificar la carga académica con respecto a sus actividades extras y el profesor de tiempo completo no.
4. Concluidas las cargas académicas, se asignan las respectivas aulas para su impartición.

Este proceso, por muy corto que parece, se demora días para realizarlo, sobre todo cuando los horarios otorgados coinciden en algún o algunos periodos y es necesario entonces volver a capturar una nueva carga académica para el profesor o profesores involucrados, el proceso es iterativo y termina hasta que no existan colisiones entre las cargas académicas.

Por tal motivo, el principal objetivo de este trabajo es mostrar una solución al problema de asignación de recursos de horarios a través de una técnica no tradicional llamada Algoritmos Evolutivos con el afán de disminuir el tiempo de dicho proceso.

Para resolver tal problemática, se muestra un modelo matemático específico al PE que incluye parámetros, variables de decisión, restricciones, una función objetivo y criterio de optimización.

La problemática timetabling plantea un modelo general como solución, del cual basándose en él, se adaptó el modelo propuesto de este trabajo y entre los principales elementos que podemos citar son:

Parámetros contemplados para el modelo específico:

K_i = Capacidad del aula i . $\forall i \in \{1, \dots, 14\}$
 C_j = Cantidad de alumnos en la asignatura j .
 $\forall j \in \{1, \dots, 25\}$

De lo cual se detalla:

1. La cantidad de aulas igual a 14.
2. La cantidad de profesores de tiempo completo es igual a 9 y por asignatura 15 (cantidad variable)
3. La cantidad total de asignatura es igual a 59, distribuidas como: 1 a 8 corresponde al primer cuatrimestre, 9 a 15 al segundo cuatrimestre, 16 a 22 al tercer cuatrimestre, 23 a 28 al cuarto cuatrimestre, 29 a 34 al quinto cuatrimestre, 35 a 41 al séptimo cuatrimestre, 42 a 47 al octavo cuatrimestre, 48 a 53 al noveno cuatrimestre y 54 a 59 al décimo cuatrimestre
4. La cantidad total de periodos semanales (p) es 35, considerando cinco periodos diarios a partir de 07:00 hrs. hasta las 20:00 hrs.

La **variable de decisión** quedó definida como:

$$X_{ijk} = \begin{cases} 1 & \text{si el aula } i \text{ es asignada a la asignatura } j \text{ en el periodo } k \\ 0 & \text{en otro caso} \end{cases}$$

$$\forall i \in \{1, \dots, 14\}, \forall j \in \{1, \dots, 59\}, \forall k \in \{1, \dots, 35\}$$

Por otro lado, las **restricciones obligatorias para evitar colisiones:**

1. Un aula i puede tener a lo más una asignación en un periodo. $\sum_{j=1}^{59} X_{ijk} \leq 1 \forall i, k$
2. Una asignatura j tiene P_j periodos semanales. $\sum_{i=1}^{14} \sum_{k=1}^{35} X_{ijk} = P_j \quad \forall j$

3. Una asignatura debe tener asignada a lo más un aula en un periodo. $\sum_{i=1}^{14} X_{ijk} \leq 1$

$$\forall j, \forall k \in \left\{ \begin{matrix} (1, \dots, 8); (9, \dots, 15); (16, \dots, 22); (23, \dots, 28); (29, \dots, 34); \\ (35, \dots, 41); (42, \dots, 47); (48, \dots, 53); (54, \dots, 59) \end{matrix} \right\}$$

4. Los horarios de las asignaturas de un mismo cuatrimestre no deben coincidir en un mismo periodo.

$$\sum_{j=a}^b \sum_{i=1}^{14} X_{ijk} \leq 1 \quad \forall k, \forall (a, b) \in \left\{ \begin{matrix} (1, \dots, 8); (9, \dots, 15); (16, \dots, 22); \\ (23, \dots, 28); (29, \dots, 34); \\ (35, \dots, 41); (42, \dots, 47); \\ (48, \dots, 53); (54, \dots, 59) \end{matrix} \right\}$$

La **función objetivo global** queda definida como la siguiente ecuación:

$$f_{og}(x) = \sum_{x=1}^N f_{ol}(x) \quad \forall x \in \{1, \dots, 315\}$$

Así mismo, la **función objetivo local** queda definida como:

$$f_{ol}(x) = ((chsl(x) + chsg(x))/2) \times ((chpl(x) + chpg(x))/2)$$

Respecto al empleo del algoritmo genético, éste define un conjunto de elementos requeridos para su utilización tales como: *población, descripción del individuo, representación de operadores genéticos*, etc.

Así mismo, la información sobre los *cromosomas* necesarios requeridos en el algoritmo genético es representada como una matriz de tres dimensiones, cuyos ejes coordenados representan los periodos de clases (eje y), los días de la semana (eje x), y los años de carrera (eje z), ver Fig. 1.

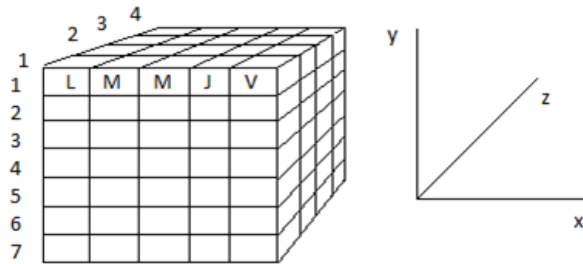


Fig. 1. Representación de cromosoma.

Con respecto al *gen* utilizado para cada cromosoma, está definido como los valores que contiene cada elemento de la matriz, los cuales se representan como una estructura de 3 elementos, ver Fig. 2.



Fig. 2. Representación del gen.

Durante la etapa de alteración del algoritmo genético, se usarán dos clásicos operadores genéticos: **Mutación** que consiste en la modificación del código genético del cromosoma, y **Recombinación** que es el intercambio de material genético entre dos cromosomas.

El procedimiento principal que se utilizó para resolver el problema de *cargas académicas* para el PE en cuestión es el siguiente:

```

t ← 0
Inicializar población P(t)
Evaluar población P(t)
Mientras llenar población hacer
    Ejecutar rueda de ruleta
    P(t) ← Seleccionar_pareja P(t)
    P'(t) ← Recombinación P(t)
    Seleccionar_individuo P(t)
    P'(t) ← Mutación P(t)
    Evaluar P'(t)
    P(t+1) ← selecc_entorno [P'(t) ∪ P(t)]
t ← t+1
    
```

Fin mientras

IV. Resultados

Tanto el algoritmo como los demás métodos o procedimientos han sido implementados en un entorno de desarrollo Java, particularmente se empleó la herramienta JCreator.

El sistema carece aún de una interfaz gráfica y parte de la respuesta del sistema se muestra en las siguientes figuras.

La captura y lectura de los datos correspondientes a la disponibilidad de horario de los docentes se realiza mediante un archivo en Microsoft Excel, ver Fig. 3, el formato es una propuesta previa.

	A	B	C	D	E	F	G	H	I	J
1	Materias	Cuatrimestre								
2	Desarrollo de habilidades del Pensamiento Lógico									
3	Expresión Oral y Escrita I	r	Identificación del Docente							
4	Formación Sociocultural I	i	Apellido Paterno	Apellido Materno	Nombre(s)					
5	Fundamentos de Redes	m	Rosas	Cabrera	Gonzalo					
6	Idioma Extranjero I	e	Asignatura: Metodología de la Programación							
7	Metodología de la Programación	r	Tabla de Valores							
8	Olimpíada	o	1	No puedo						
9	Soporte Técnico		2	Tal vez puedo						
10			3	Si Puedo						
11			Según la clasificación que se indica en la tabla de valores, designe los días y períodos en los cuales puede realizar sus clases (llenar todos los periodos y días)							
12										
13										
14										
15										
16			Periodo \ Días	Lunes	Martes	Miércoles	Jueves	Viernes		
17			07:00 - 09:00	3	3	3	3	3	3	
18			09:00 - 11:00	3	3	3	3	3	3	
19			11:00 - 13:00	3	3	3	3	3	3	
20			13:00 - 15:00	3	3	3	3	3	3	
21			15:00 - 17:00	3	3	3	3	3	3	
22			17:00 - 19:00	3	3	3	3	3	3	
23										
24										
25										
26			Nota:							
27			1	Completa los 5 cuatrimestres						
			2	Si das clases en Ingeniería, completa hasta el décimo cuatrimestre						

Fig. 3. Captura de datos

Una vez generado el archivo en Excel, el sistema lee los respectivos datos generando archivos de texto plano para realizar una posible modificación, ver Fig. 4; cuando el proceso ha terminado de iterar, es decir, ya no existen cambios en los horarios de los profesores respectivos, entonces el sistema genera la carga académica final del próximo cuatrimestre, como lo muestra la Fig. 5.

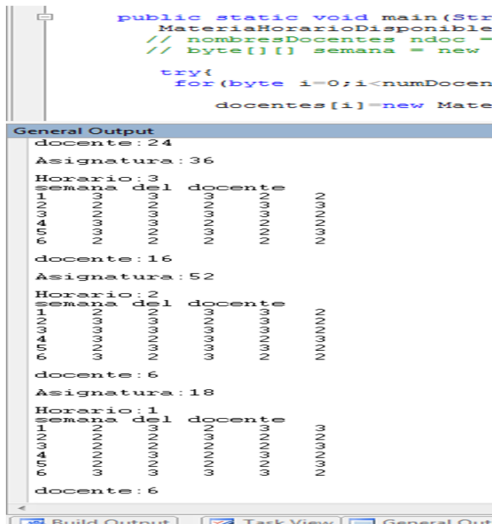


Fig. 4. Horarios de profesores

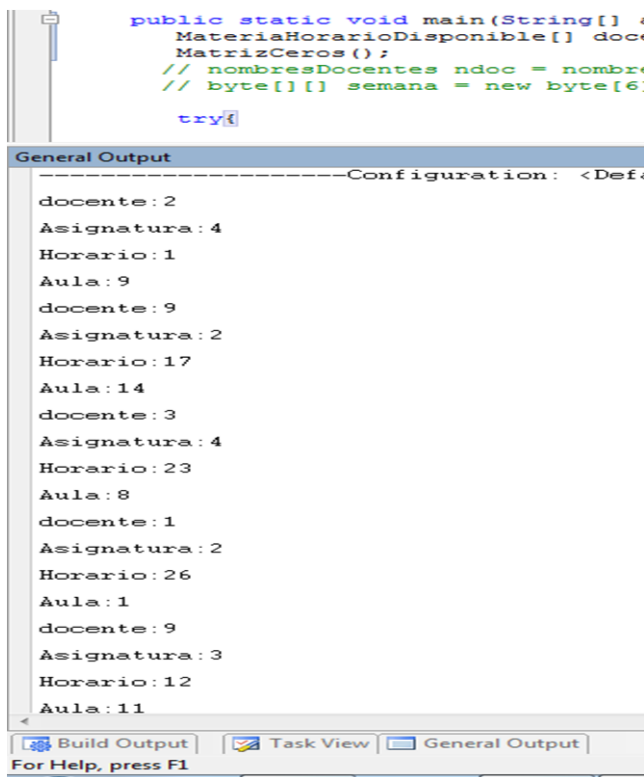


Fig. 5. Cargas académicas

Finalmente, parte de la información que arroja el sistema referente a la matriz principal de cromosomas para poder generar las cargas académicas, considerando las posibles colisiones se muestran en la Fig. 6.

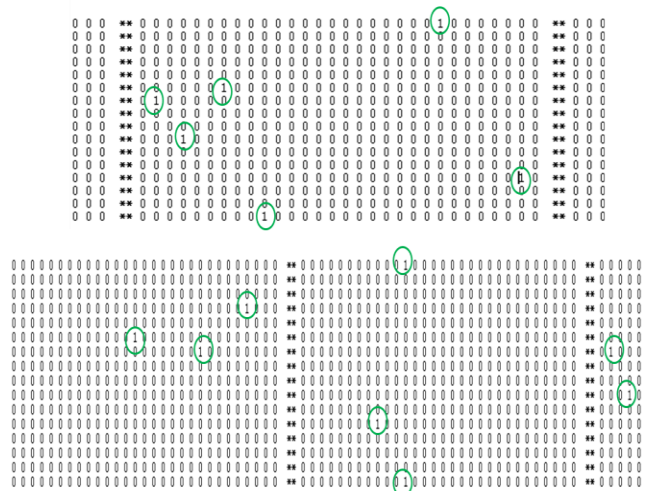


Fig. 6. Datos contenidos en la Matriz principal

V. Conclusiones

En la literatura se puede apreciar las escasas publicaciones relacionados con los Algoritmos Evolutivos referente al tema, así como su respectiva implementación.

El problema de asignación de horarios, aplicado a cualquier institución educativa es un desafío difícil de enfrentar, debido a la gran cantidad de variables y restricciones existentes, sin embargo, esto no es motivo para dejar pasar su estudio. Este trabajo tuvo como objetivo principal mostrar una solución al problema de asignación de horarios para uno de los cinco PE de la UTIM, a través de una técnica llamada Algoritmos Evolutivos perteneciente al grupo de Metaheurísticas.

El trabajo incluye además la implementación de un algoritmo *rueda ruleta* modificado, puesto que se supone que como éste trabaja con funciones de maximizar, al calcular el porcentaje acumulado del *fitness*, el algoritmo *rueda ruleta* nos proporcionó una mayor fracción de los individuos que tenían más choques, ya que estos poseen un valor de *fitness* mayor, y lo ideal era obtener una fracción menor a los que tienen menos choques.

Así mismo, existen investigaciones en las cuales se ha modificado la variable para el número de generaciones (nuestro caso es de 30), con el

objeto de conocer el comportamiento de las soluciones a través de un número de generaciones dado.

Actualmente se está diseñando una interfaz que permita mostrar los datos de los horarios solicitados mediante el lenguaje Visual C#, dentro de la plataforma .NET.

La idea es crear un sistema fácil y usable para el usuario final, contemplando también para su desarrollo las herramientas de Microsoft Expression Studio.

En lo futuro se pretende:

1. Realizar la implementación del diseño propuesto para la solución de timetabling para la UTIM empleando el modelo MVC de la plataforma .NET.
2. Comparar nuestros resultados con otros algoritmos propuestos así como con la herramienta LPSolve.
3. Adaptar el algoritmo para resolver las cargas académicas de cada uno de los PE de la UTIM.
4. Desarrollar un portal WAP que permita a los directivos realizar las acciones respectivas desde un dispositivo móvil.
5. Considerar bloques de pruebas para la aplicación, primero se trabajará con cincuenta generaciones, una segunda prueba con cien y por último una prueba con quinientas generaciones; de esta manera, al finalizar cada prueba se obtendrán las soluciones y con ello su comportamiento, para ser comparado con técnicas como Programación Lineal Entera Binaria (PLEB).

VI. Referencias

- [1] J. Franco, E. Mirledy, R. Alfonso, “Problema de asignación óptima de salones resuelto con Búsqueda Tabú”. Ingeniería & Desarrollo. Núm. 24. Julio-Diciembre 2008. Pp. 149-175.
- [2] J. Gallart, F. Alva, A. Alama, G. Bejarano, “Intelligent Generation of Schedules using GRASP heuristic with Tabu Search for the

Pontificia Universidad Católica del Perú”, Ingeniería Informática PUCP. Vol. 1. Núm. 1. Abril 2010, pp. 15-24.

- [3] Z. Furones, J. Pacheco, A. Aragón, “Diseño de calendario de exámenes en la universidad”, Anales de la economía aplicada, XIV Reunión ASEPELT, Junio 2000, pp. 151-162.
- [4] B. López, D. Jaime. “Modelo de asignación de carga académica usando Algoritmos Genéticos”. Sigma. Mayo 2000. Año 3. Núm. 5.
- [5] P. Flores, E. Brau, J. Monteverde, N. Salazar, J. Figueroa, E. Cadena, C. Lizarraga, “Experimentos con algoritmos genéticos para resolver un problema real de programación Maestros-Horarios-Cursos”, Universidad de Sonora, 2002.
- [6] P. Gervás, B. San Miguel, “Un sistema experto basado en reglas para la automatización de la elaboración de horarios para un conjunto de restricciones particulares”, Universidad Europea de Madrid, 1999.
- [7] F. Martínez, “Timetabling académico usando algoritmos genéticos y programación celular”, Universidad Autónoma de Zacatecas, 2005.

VII. Autores

LCC Ivan Antonio Flores Trujillo obtuvo su título de Licenciatura en la Benemérita Universidad Autónoma de Puebla en el año 2004, actualmente estudia una Maestría en sistemas Computacionales en la Universidad Popular Autónoma del Estado de Puebla y labora como profesor de tiempo completo en la Universidad Tecnológica de Izúcar de Matamoros; ha redactado artículos sobre la problemática Timetabling, Puntos Función en el desarrollo de Software, Diseño de estrategias de enseñanza sobre cursos de Programación; ha participado en eventos como SITOI 2008 y 2009, CONCyE 2011 y ha obtenido certificaciones en MOS y MCP por parte de Microsoft. Así mismo, cuenta

con un Diplomado en Herramientas Metodológicas para la Formación basada en Competencias Profesionales por parte del ITESM.

MSC. Patricia Mendoza Crisóstomo es maestra en Sistemas Computacionales por la Universidad Popular del Estado de Puebla, profesora de tiempo completo de la Universidad Tecnológica de Izúcar de Matamoros y colaboradora en diversos proyectos dentro de la institución relacionados con el control de procesos de producción en conjunto con la carrera de Agrobiotecnología y la Universidad Autónoma de Chapingo; ha participado en eventos como SITOI 2009, CONCyE 2011 y CIEIC 2011, tiene certificaciones en MOS, CS4 y IC3 por parte de Microsoft.

Lic. Gustavo Soriano Jiménez es egresado del Instituto Tecnológico de Puebla como Licenciado en Informática con especialidad en Sistemas Expertos. Cuenta con un Diplomado en Herramientas Metodológicas para la Formación basada en Competencias Profesionales (2010-ITESM), y actualmente cursa la Maestría en Sistemas Computacionales en la UPAEP (80%). Desde el año 1997 he trabajado primero como Agente de Ventas (en R.I.S.C. S.A. de C.V.) y luego como Independiente en la Compra-Venta de Equipo de Cómputo y en actividades de Soporte Técnico. De forma independiente ha desarrollado sitios Web y pequeñas aplicaciones para negocios. De 1999 a la fecha se ha dedicado al sector educativo. En el Instituto Tecnológico Superior de Libres desempeño los puestos de Jefe del Departamento de Estudios Superiores y Jefe del Departamento de Investigación. En la Universidad Tecnológica de Izúcar de Matamoros (UTIM) ha desempeñado puestos tales como Ingeniero en Sistemas de la Dirección de Administración y Finanzas, Profesor por Asignatura y actualmente Profesor de Tiempo Completo. Participo en la reestructuración del sitio web de la Universidad. Actualmente pertenece al Cuerpo Académico de “Sistemas de Información aplicando Minería de Datos y Algoritmos de Cómputo” y trabaja en el desarrollo del proyecto “Minería de Datos para

determinar el Estilo de Aprendizaje Dominante en alumnos del Programa Educativo de TIC-SI”. Ha participado como ponente en Congresos y como Instructor de diversos Talleres en otras Instituciones Educativas.

MSC. Sergio Valero Orea, egresado de la Universidad Madero como Ingeniero en Sistemas Computacionales, y como Maestro en Sistemas Computacionales por la UPAEP. Desde el año 2000 ha trabajado como consultor independiente desarrollando proyectos de sistemas y bases de datos. Actualmente, y desde el año 2000, labora como profesor de tiempo completo en la Universidad Tecnológica de Izúcar de Matamoros y como coordinador del departamento de sistemas de la misma Universidad. Ha desarrollado proyectos en el área de bases de datos, ingeniería del software y seguridad de la información. Ha participado como ponente en congresos nacionales e internacionales de computación, sistemas y tecnologías de la información.

MSC. Alejandro Salvador Vargas es Maestro en Sistemas Computacionales por la UPAEP. Maestro en Administración Escolar por el Colegio Latinoamericano de Posgrados (CLAP). Licenciado en Informática por el Instituto Tecnológico Superior de la Sierra Norte de Puebla, Zacatlán. Desde 2004 labora como Profesor de Tiempo Completo e Instructor CCNA de CISCO en la Universidad Tecnológica de Izúcar de Matamoros en el Programa Educativo de TIC-SI, donde ha implementado un Sistema de Aprendizaje Virtual, cuyo objetivo apoyar al proceso enseñanza – aprendizaje con el uso de las TIC’s. Ha laborado en el CBTis 184 de Izúcar de Matamoros (2005–2008) y en la Universidad Tecnológica de Xicotepec de Juárez (2003–2004).

Desarrollo de un Sistema para la Contabilidad de Costos basado en Tecnología Web (ExtJs-PHP), con la implementación de un DataWarehouse (MySQL).

Christian Mauricio Castillo Estrada, Mtro. David Ristori Cueto, Daniel Hernández de la Cruz

Facultad de Contaduría Pública del Campus IV de la Universidad Autónoma de Chiapas. Carretera a Puerto Madero Km 1.5, Tapachula, Chiapas, México (962) 6251723
{cmce@unach.mx, ristori@unach.mx, danielh1050@gmail.com}

Resumen

En este trabajo se presentan los resultados derivados del Desarrollo de un Sistema de Información para el control de la Contabilidad de Costos, mismo que ha sido implantado en la organización MoscaFRUT del Programa General de Cooperación Técnica y de Gestión de Proyectos SAGARPA-IICA. Para el desarrollo de este proyecto, se utilizó como herramientas tecnológicas de desarrollo el Sistema Gestor de Base de Datos MySQL, procedimientos almacenados de Transact SQL, y la plataforma de desarrollo de aplicaciones web Sencha (ExtJS-PHP). Las pruebas experimentales se realizaron en las instalaciones de la Planta MoscaFRUT en el área de Costos y Dirección, con la finalidad de evaluar el funcionamiento y validar la integridad de datos del Sistema, durante el proceso de toma de decisiones de los Directivos de la Organización. Por lo anterior, resulta interesante estudiar el impacto de la Tecnología Web y DataWarehouse en el desarrollo de Sistemas de Información que apoyen a la toma de decisiones y al procesamiento de grandes cantidades de información.

Palabras Clave: DataWarehouse, MySQL, ExtJS, PHP, Contabilidad de Costos.

I. Introducción

La globalización provoca como efecto inmediato que las micro y pequeñas empresas tengan que optimizar sus procesos para estar en condiciones de competitividad en un mercado tan convulsionado. En ese sentido, los Sistemas de Información resultan ser la clave del éxito de las empresas, si se toman las decisiones correctas con base a la información que proveen estos.

En la mayoría de las organizaciones, indistintamente de su naturaleza, la información constituye un elemento imprescindible para la planificación y el proceso de control. Para la planificación se requiere de información relacionada con las variables controlables y las no controlables sobre recursos financieros, humanos y materiales, y para el proceso de control es necesario información proveniente de las diferentes fases de los procesos de producción, donde se pueda detectar mejoras,

fallas y marcar tiempos para optimizar cada uno de ellos.

Un sistema de información es un conjunto de personas, datos y procedimientos que funcionan en conjunto. El énfasis en un sistema significa que los diversos componentes buscan un objetivo común para apoyar las actividades de la organización. Éstas incluyen las operaciones diarias de la empresa, la comunicación de los datos e informes, la administración de las actividades y la toma de decisiones [1].

La información contable debe cumplir con cualidades que involucren la satisfacción de las necesidades de los usuarios. Esto determina que este tipo de información debe ser comprensible, clara, confiable, transparente, oportuna y comparable; lo cual se resume en que la información contable debe representar fielmente los hechos económicos de la empresa. Bajo esta perspectiva, los Directivos de la Organización MoscaFRUT del Programa General De

Cooperación Técnica y de Gestión de Proyectos SAGARPA-IICA, decidieron realizar un convenio de colaboración con la Universidad Autónoma de Chiapas con la finalidad que dos docentes participaran como asesores en tecnología y análisis de requerimientos en colaboración con el área de informática de la Planta; ante la necesidad del Desarrollo de un Sistema de Información de Contabilidad de Costo; misma que se ha originado en los últimos años, debido a que la organización MoscaFRUT ha tenido que enfrentarse a diferentes cambios en su entorno, principalmente con una radical variación en su producción de pupas (moscas), lo que conlleva a la necesidad de tener sistemas de información confiables que les permita ayudar a mantener el control y mejorar sus procesos de producción.

Actualmente, MoscaFRUT produce tres tipos de pupas, siendo estas: Parasitoide, Anastrepha Ludens y Anastrepha Oblicua, las cuales se envían a diferentes Estados de la República Mexicana; la producción de pupas son controladas a través de un identificador denominado semana lote. Por tal motivo, los jefes de producción de cada uno de los tipos de pupas solicitan una cantidad de insumos para dar inicio al proceso de producción en cada una de las semanas lote que se establecen al inicio del año. Sin embargo, no se tiene un control histórico de las solicitudes de esos insumos, lo cual produce problemas para el área de almacén impactando en la entrega oportuna de insumos y la determinación de los costos unitarios de producción. Con base a lo anterior, se identifican problemas en sus presupuestos, afectando severamente la producción y tiempos de entrega. Uno de los procedimientos más complicados para las organizaciones es calcular sus costos unitarios de producción. En ese sentido, al no contar con Sistemas de Información que apoyen a este tipo de procesos, resulta ser un reto tratar de obtener información confiable y veraz que permita la toma de decisiones.

Con la finalidad y el interés de dar a conocer las experiencias de implementar y aplicar las tecnologías web para el control de presupuestos y

generación de costos unitarios, que permitan apoyar al proceso de tomas de decisiones; se elabora este artículo, en el cual se presentan la metodología para el desarrollo de un Sistema de Información basado en un entorno web para la Contabilidad de Costos.

II. Desarrollo

El desarrollo del proyecto, por su parte, se muestra constituido por dos grandes momentos: el análisis y desarrollo de la solución tecnológica basada en el construcción de estudio de caso y la determinación de las competencias profesionales que serán incorporadas a los programas educativos que oferta la Facultad de Contaduría Pública, Campus IV.

El Sistema de Información desarrollado que presenta en este artículo, fue diseñado de manera modular clasificando los diversos tipos de gastos que se generan en la Planta MoscaFRUT, mismos que se registran en un apartado específico de la base de datos denominado concentrado de Gastos, partiendo de la idea de la implementación de un DataWarehouse. Lo anterior, permitió facilitar el diseño de los algoritmos para la generación de los costos unitarios de producción.

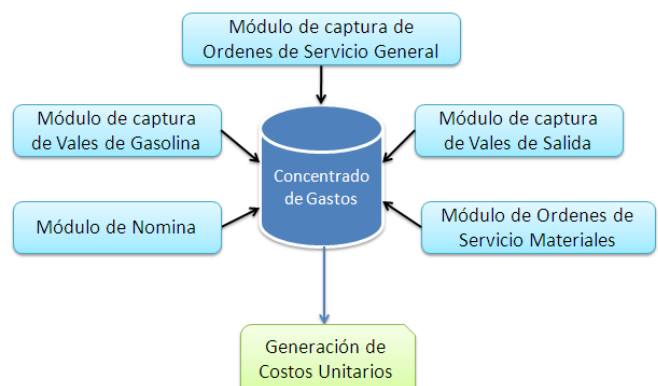


Fig. 1. Módulos que integran el Sistema de Información de Contabilidad de Costos

Para el desarrollo del Proyecto, se utilizó la metodología planteada para el desarrollo de Aplicaciones Web identificando de forma simultánea las competencias profesionales en cada una de las etapas que se establecen en la metodología. Una metodología no es otra cosa

que un conjunto de proposiciones lógicas, graduadas y jerarquizadas, destinadas a facilitar y mejorar el ejercicio intelectual, la capacidad creadora de la mente humana en cualquier rama del saber. [2]

Roger Pressman, el reconocido autor de libros sobre Ingeniería de Software, moderó una mesa redonda virtual con representantes de la ingeniería software tradicional y del desarrollo software basado exclusivamente en Internet. La conclusión de esta reunión fue que el proceso de ingeniería de software tradicional debería adaptarse a los requerimientos de cambio continuo y rapidez presentes en el proceso de desarrollo de Aplicaciones Web. En virtud de lo anterior, surge una nueva disciplina denominada Ingeniería Web, la cual se define como el proceso utilizado para crear, implantar y mantener aplicaciones y sistemas Web de alta calidad. Tomando como base el “Ciclo de vida clásico”. Pressman, “adapta” esta metodología para el desarrollo de Web, considerando las siguientes etapas: Formulación, Planificación, Análisis, Diseño, Implementación y Pruebas; misma que se utilizó para desarrollar la solución tecnológica presentada en este artículo.

2.1 Análisis de Requerimientos

Con la finalidad de realizar la identificación de los requerimientos, se realizaron entrevistas a los Directivos, Jefes de Área y al responsable del Departamento de Contabilidad de Costos. Posteriormente, se procedió con el análisis de los requerimientos para determinar y elaborar los casos de uso, diagrama de clases utilizando el lenguaje de modelado UML, para lograr obtener un esquema general de la problemática.

2.2 Diseño e Implementación del DataWarehouse

Un DataWarehouse es una base de datos corporativos que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con

grandes velocidades de respuesta [3]. Esta base de datos, representa un repositorio de información orientado a apoyar al proceso de toma de decisiones de los Directivos, permitiéndole establecer y planificar estrategias que mejoren el desempeño de la producción de moscas. Para ello, se diseñó un esquema de la base de datos correspondiente al DataWarehouse; lo cual permitió a los participantes desarrollar competencias profesionales enfocadas al uso de nuevas tecnologías de información, diseño y comunicación. En la figura 2, se muestra un Diagrama de las Entidades o Tablas principales que integran al DataWarehouse.

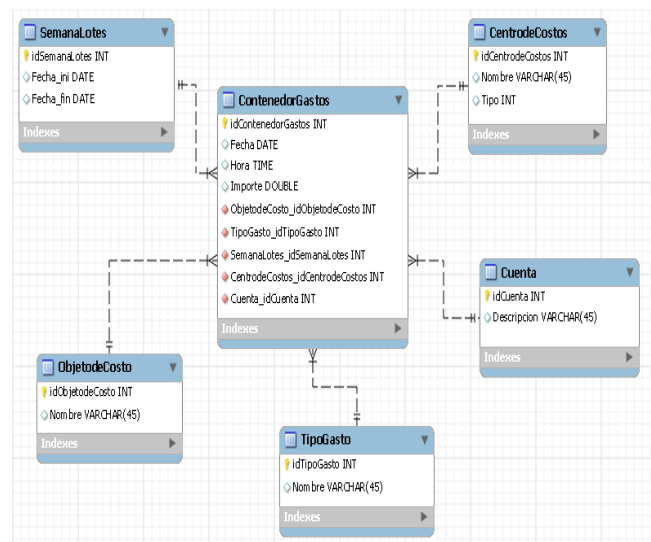


Fig. 2. Esquema del DataWarehouse

En relación a la arquitectura del DataWarehouse, la cual es una forma de representar la comunicación, la estructura de los datos, los procesos y la presentación de información al usuario final; podemos mencionar que a Nivel de Base de datos operacional, seleccionamos el Gestor de Base de Datos denominado MySQL debido a que posee un motor de almacenamiento pensado para trabajar con Datawarehouses, mismo que se le denomina MyISAM. A Nivel físico, MySQL optimiza el uso de la memoria y no requiere de altas especificaciones técnicas para el Servidor. Además, este Gestor posee características que lo hacen ideal para su implementación con Datawarehouses, como son: gestión de almacenamiento automático, soporte

indexación, carga paralela de datos, etc. Por otra parte, en los Niveles de directorio de datos (metadatos) y de acceso a los datos, se usó una herramienta ETL llamada MySQL Workbench, la cual facilitó la definición de los objetos de negocio y enlazar la base de datos con los diferentes orígenes, definir las transformaciones para la integración de los datos, y la programación de cargas periódicas. En lo que respecta al Nivel de acceso a la información, utilizamos la herramienta de administración llamada PhpMyAdmin, misma que permitió la interacción con la información almacenada de una forma fácil y transparente. Así mismo, para la generación de informes y gráficos se usó un componente de Graficación de la plataforma de desarrollo ExtJS en combinación con el lenguaje de programación PHP.

2.3 Diseño de la Interfaz de Usuario

Posteriormente, se realizó el Diseño de Interfaz de usuario, lo cual implicó identificar la competencia de educación continua puesto que se tuvieron que consultar los lenguajes de programación especializados para el desarrollo de aplicaciones Web que no son contemplados en la formación profesional del programa educativo.

Para diseñar la Interfaz de Usuario del Sistema de Contabilidad de Costos que se presenta, se usó la plataforma de desarrollo ExtJS, la cual ofrece un conjunto de librerías JavaScript para la creación de Rich Internet Applications (RIA), las cuales son aplicaciones web interactivas que usan tecnologías como AJAX, DHTML y DOM, mismas que permiten el acceso a datos de forma remota por parte de los usuarios. En la Fig.3 se puede observar la interfaz gráfica correspondiente al menú principal del sistema.



Fig. 3. Interfaz de Usuario

2.4 Implementación del Modelo de N-Capas

En la actualidad, las aplicaciones de Internet como noticias en línea, comercio, y sitios financieros se han vuelto bastante comunes; en ese sentido, las aplicaciones modernas de Internet son sistemas complejos de software que emplean una arquitectura de varios niveles. Cada nivel ofrece una cierta funcionalidad a su nivel anterior y hace uso de la funcionalidad proporcionada por su sucesor, para llevar a cabo su parte del procesamiento de la solicitud general [5].

La arquitectura multicapa se divide en tres partes: cliente, servidor de aplicaciones y servidor de base de datos. La capa cliente o de presentación, interactúa con el servidor enviando solicitudes para que sean procesadas, en este proyecto usamos la tecnología ExtJS, para diseñar el entorno visual de esta capa. Por otra parte, el servidor de aplicaciones es el responsable de proveer al cliente el acceso a los datos, procesar las peticiones y funcionar como una interfaz entre la capa cliente y la capa de datos; el servidor web usado fue Apache interactuando con el intérprete del lenguaje de scripts (PHP). La capa de datos, se encarga de procesar las solicitudes de datos enviadas por la capa cliente; además, es responsable de almacenar toda la información a través de tablas, índices, y otros objetos de datos. En el proyecto, se implementó el Sistema Gestor de Base de Datos MYSQL [6].

materiales, subcuentas directivos en el momento de reportes y unitarios. La que se características:

los indirectos una de las clasificados

para los jefes directivos de

unitarios de de insectos

(Ludens, Obliqua, Parasitoide).

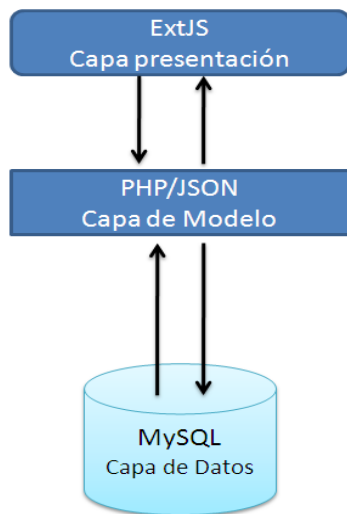


Fig. 4. Modelo de 3-Capas utilizado para el desarrollo de la Solución Tecnológica (Sistema de Contabilidad de Costos).

III. Resultados

Para comprobar la funcionalidad del Sistema de Información de Contabilidad de Costos, se realizó la instalación de un Servidor Web dentro de la Organización; así mismo, se configuró el Sistema Gestor de Base de Datos MySQL para realizar la carga de datos y los procedimientos almacenados. Posteriormente, se capturaron los datos de entrada relacionados con los catálogos

En la figura 5, se muestra una pantalla del módulo que permite la generación de los costos unitarios de los tres tipos de insectos que produce la planta, mismos que son representados como objetos de costos.

Así mismo, en la figura 6 se puede apreciar el reporte de concentrado de gastos; el cual considera todos los gastos efectuados en cada centro de costos, clasificándolos por cuenta presupuestal y los visualiza por periodos de semanas lote; facilitándole a los Directivos de la planta la toma de decisiones respecto a los gastos que influyen en la producción.

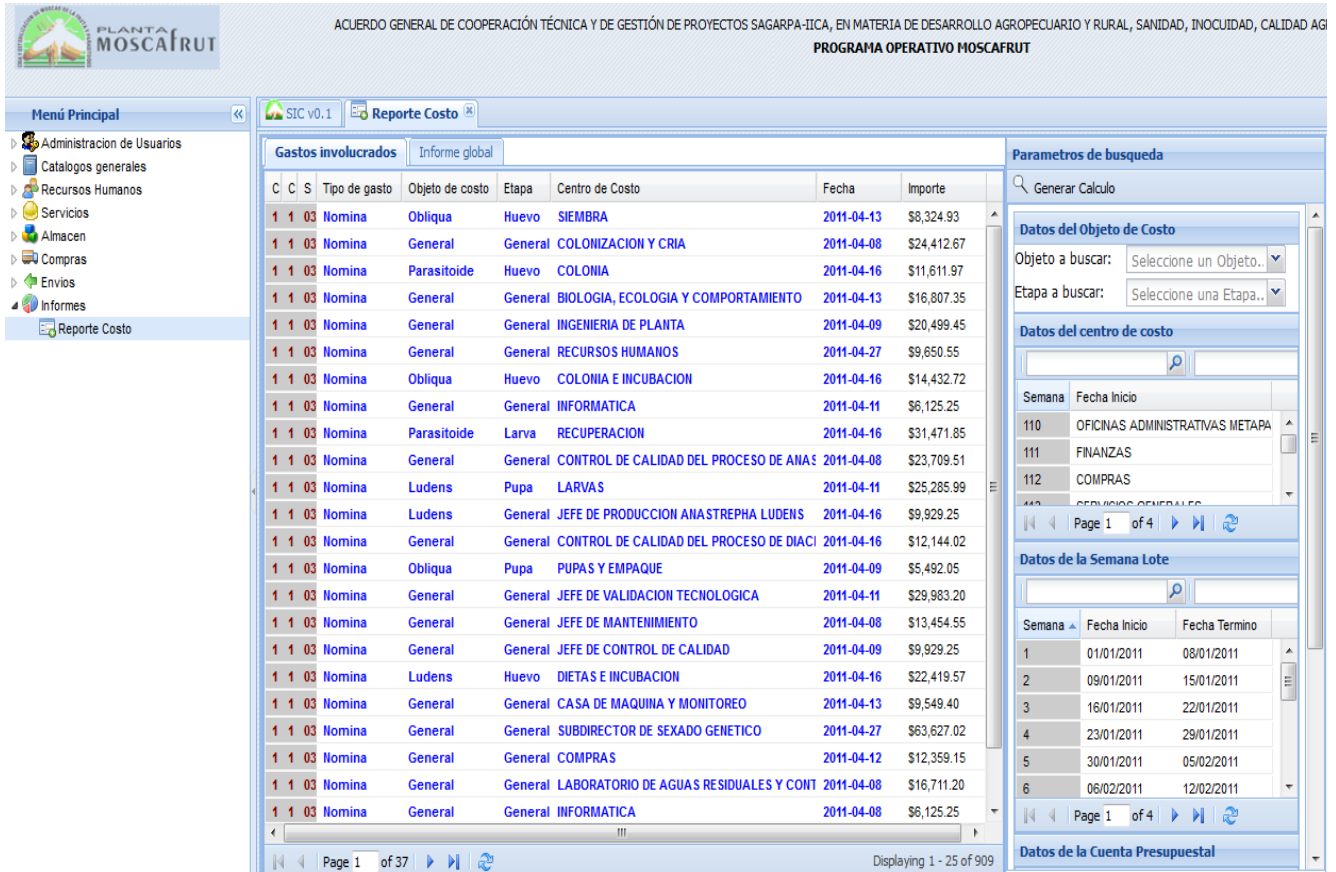


Fig. 5. Interfaz Gráfica del Módulo de generación de costos unitarios.

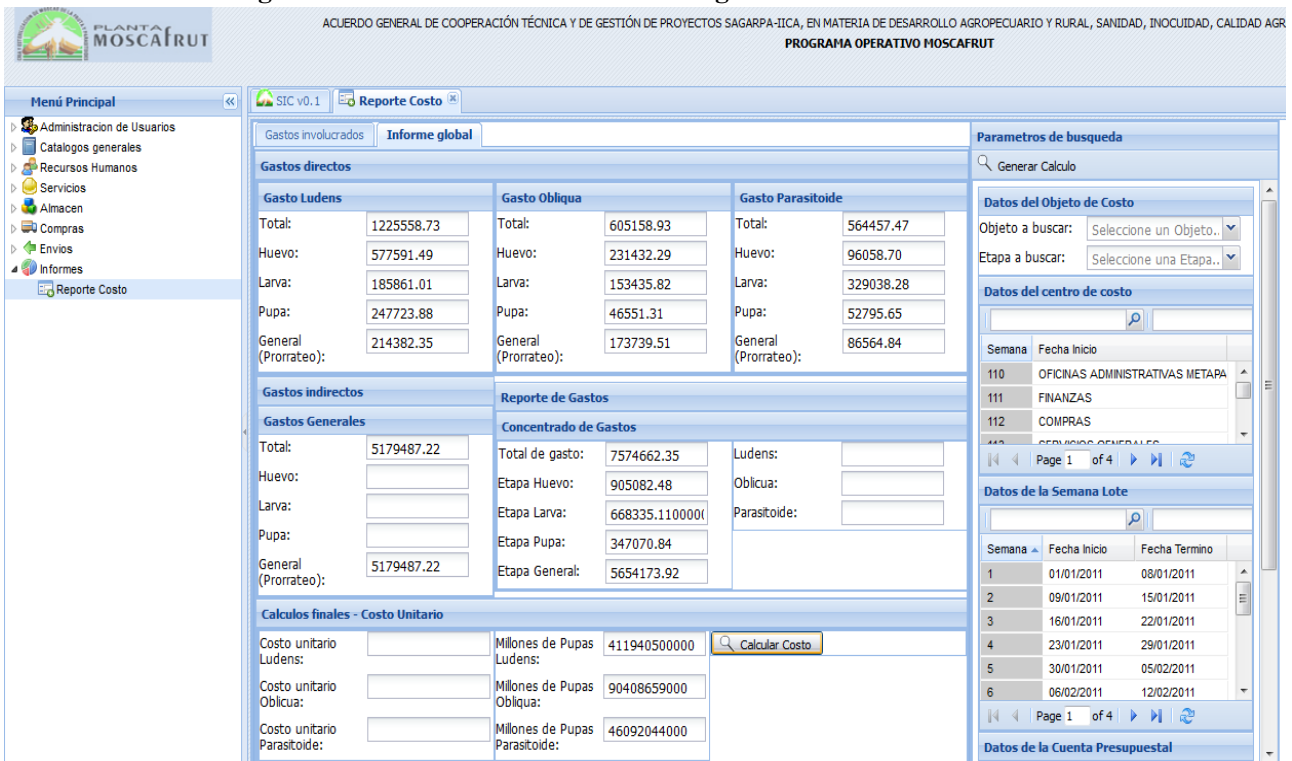


Fig. 6. Interfaz Gráfica del Módulo de generación del reporte de concentrado de gastos.

IV. Conclusiones

Hoy en día, la Internet se ha convertido en una plataforma interesante para la implementación de aplicaciones empresariales basadas en un entorno web, debido a la flexibilidad que ofrece. Las tendencias indican una rápida proliferación de este tipo de aplicaciones en la Web; muchos desarrolladores y profesionistas de TI, encuentran en la Web un entorno valioso para el desarrollo rápido y despliegue de software de calidad basado en una arquitectura multi-capa.

Para el diseño y la implementación del Sistema de Contabilidad de Costos, fue indispensable la adopción de las más recientes tecnologías de desarrollo e intercambio de datos, tal como lo son ExtJS y JSON; permitiendo lograr un alto desempeño de la aplicación en el proceso de captura y generación de información.

Actualmente, el Sistema de Contabilidad de Costos se encuentra en las etapas de prueba y adaptación; debido a la gran cantidad de información que maneja, es necesario capacitar a los empleados de las diferentes áreas de producción y departamentos relacionados, para poder centralizar toda la información relacionada a gastos, misma que es almacenada en el DataWarehouse.

La implementación del Sistema basado en un entorno Web, en combinación con el uso del DataWarehouse han permitido agilizar los procesos de recopilación, análisis y generación de información, para calcular los costos unitarios de los tres tipos de insectos que se producen en la Planta, provocando una disminución destacable del tiempo requerido para realizar el concentrado de gastos y el procedimiento de cálculo de costos unitarios. Así mismo, ha permitido la entrega de información en tiempo real a los Directivos, facilitando la toma de decisiones. Por otra parte, el desarrollo e implementación del Sistema, permitió comprobar que la arquitectura de tres capas, mejora considerablemente el desempeño de la aplicación web, debido que la tecnología usada en la capa cliente o de presentación, ejecuta algunas operaciones de la lógica de

negocio en la máquina del usuario cliente y libera al servidor de aplicaciones de ejecutar ciertas tareas, provocando una disminución del tráfico en la red.

V. Referencias

- [1] Gómez F.C. Fernando (1999) "**Sistemas administrativos estructuras y procesos**". Edit. Macchi grupo editor S.A.
- [2] Pacheco A., (2006). **Metodología crítica de la Investigación Lógica, Procedimiento y técnicas**, Edit.CECSA.
- [3] Sinnexus. "**Definición de DataWarehouse**". Consultado en el día 4 de Mayo de 2010 en http://www.sinnexus.com/business_intelligence/datawarehouse.aspx.
- [4] Mitecnológico. "**Metodologías Desarrollo Aplicaciones Web**". Consultado en el junio de 2011 en <http://www.mitecnologico.com/Main/MetodologiasDesarrolloAplicacionesWeb>.
- [5] Bhuvan Uргаonkar, Giovanni Pacifici, Prashant Shenoy. "Analytic Modeling of Multitier Internet Applications". Consultado en el mes de junio de 2011 en la ACM Digital Library.
- [6] Technologies Cadcim. "Oracle 11G: With Pl/Sql Approach". Person Education. Consultado en el mes de junio de 2011 en <http://books.google.com.mx>
- [7] Roger Barga, David Lomet. "Recovery guarantees for Internet applications". Consultado en el mes de junio de 2011 en la ACM Digital Library.
- [8] María A. Nieto. "**Ingeniería Web - Construyendo Web Apps**". Consultado en el mes de junio de 2011 en <http://www.informandote.com/jornadasIngWEB/articulos/jiw01.pdf>.
- [9] Domingo García. "**La contabilidad de costos y rentabilidad en la Pyme**". Consultado en el mes de junio de 2011 en <http://www.ejournal.unam.mx/rca/218/RCA21803.pdf>

- [10] Chen, Jim Q. "Building Web Applications". Consultado en el mes de Junio de 2011 en la base de datos Business Source Premier de EBSCO Host.

VI. Autores

M.A. David Ristori Cueto obtuvo su título de Maestro en Administración por la Universidad Autónoma de Chiapas. Actualmente, estudia el primer año del Doctorado en Gestión para el Desarrollo y es profesor de tiempo completo en la Universidad Autónoma de Chiapas.

L.S.C. Christian Mauricio Castillo Estrada es Licenciado en Sistemas Computacionales por la Universidad Autónoma de Chiapas y se encuentra cursando el último trimestre de la Maestría en Comercio Electrónico con área terminal en Tecnologías de Información en la Universidad Virtual del ITESM. Es profesor de asignatura de la Licenciatura en Sistemas Computacionales en la Facultad de Contaduría Pública del Campus IV en la Universidad Autónoma de Chiapas.

L.S.C. Daniel Hernández de la Cruz es Licenciado en Sistemas Computacionales por la Universidad Autónoma de Chiapas y se encuentra cursando el segundo cuatrimestre de la Maestría en Ciencias de la Computación con Especialidad en Bases de Datos en la Universidad Valle de Grijalva. Actualmente, colabora en la Planta MoscaFRUT del Programa General De Cooperación Técnica y de Gestión de Proyectos SGARPA-IICA como Desarrollador de Software en el Área de Informática.

Consideraciones en la Migración de Datos Locales hacia Entornos Distribuidos

***Servín Juárez Fidencio, **Cárdenas Robledo Leonor Adriana, ***Santiago Paz Daniel**

Universidad de la Sierra Juárez, Av. Universidad s/n, Ixtlán de Juárez, Oax. C.P. 68725, teléfono. 9515536362 ext. 201.

*servin_fidencio@juppa.unsij.edu.mx

** lcardenas@juppa.unsij.edu.mx

***li050026@juppa.unsij.edu.mx

Resumen

Los avances tecnológicos en comunicaciones y aplicaciones de software han facilitado la integración de sistemas distribuidos para el manejo de información ubicada en áreas geográficamente distantes. En el presente documento se expone una propuesta de migración paulatina de datos para una empresa tipo PyME, considerando aspectos esenciales en bases de datos distribuidas como la seguridad, robustez y consistencia en la información, así también la implementación del protocolo commit de dos fases como mecanismo de control de concurrencia. Con la finalidad de tener un esquema de distribución de recursos multiplataforma permitiendo el acceso seguro a los datos remotos, la compartición de los mismos en forma controlada y eficiente, garantizando siempre la integridad de la información.

Palabras Clave: Sistemas distribuidos, migración de datos, consistencia de datos, bases de datos.

I. Introducción

La independencia de los datos es una característica de las bases de datos actuales, esto implica que puedan ser utilizados por distintos sistemas, dependiendo de su objetivo. Para ello es necesario adecuarlos al nuevo entorno sin que pierdan su significado y esencia. Los datos heredados llevados hacia entornos distribuidos conllevan una reformulación de contenidos ampliados que representan una realidad diferente inmersa en la complejidad del intercambio de información.

El propósito del documento consiste en presentar una propuesta de migración de datos hacia un ambiente distribuido, diseñada para una empresa dedicada a la distribución y venta de productos farmacéuticos. En la propuesta se incorporan características de seguridad, robustez y consistencia de datos, factores claves para la óptima eficiencia de un sistema distribuido compatible en el manejo de datos con el esquema cliente-servidor.

La migración de datos desde sistemas heredados comerciales donde no se cuenta con la

documentación del sistema, implica un mayor esfuerzo, sin embargo a medida que los datos pueden ser adquiridos y analizados con herramientas de terceros, la tarea de incluirlos se facilita, sin embargo la validación continua aún después de la migración.

II. Descripción del problema

La empresa actualmente opera con el sistema de punto de venta Microsip [1], el cual no permite dar soporte a nuevos requerimientos de almacenamiento de datos, así mismo la actualización de información se realiza de forma asíncrona entre una sucursal de la empresa y el servidor. Al momento de concentrar la información generada en sus diferentes áreas se generan inconsistencias en la información enviada al servidor y por consiguiente el nivel en la integridad no es confiable.

Otro problema se presenta al realizar traspasos de productos entre diferentes sucursales, en donde se requiere de un enlace cliente-servidor pre-configurado para su ejecución, lo cual representa un riesgo elevado de inconsistencia en los datos

actuales y transferidos, debido a que el envío de información como catálogos digitales de artículos con precios, descripciones y detalles generales son actualizados constantemente desde el corporativo a través del servidor, en este caso las sucursales carecen de información actualizada, ya que el servidor no interviene en la transacción entre las sucursales donde los datos del traspaso son incorporados al sistema local.

Microsip se encuentra ejecutándose sobre el sistema operativo Windows, siendo éste un software propietario, vulnerable al ataque de software malicioso como virus, que alteran de manera total o parcial la información de la empresa. Al presente se utiliza el Sistema Manejador de Base de Datos (SMBD) Firebird 1.5 que es lento para manejar índices secundarios al aumentar el volumen de información, en cuanto a seguridad se refiere, el manejar la información en forma de archivos individuales se vulnera fácilmente debido a que son transferibles.

III. Solución propuesta

La propuesta incluye el análisis y diseño de un sistema que reemplace de manera paulatina los módulos de punto de venta e inventarios del sistema Microsip y su futura implementación en el Lenguaje Java [2]. Se recomienda en las sucursales de la empresa el cambio del sistema operativo Windows XP (software propietario) a Ubuntu (software libre) ya que representa una opción para reducir el pago de licencias por concepto de sistema operativo y uso de antivirus, así también se sugiere suplir el SMBD Firebird 1.5 por Oracle 10g. La metodología propuesta hace posible la migración de datos desde Firebird hacia Oracle[3], considerando el modelo cliente-servidor como una alternativa a la intercomunicación horizontal entre nodos distribuidos. Para el nodo central ubicado en la matriz de la empresa se recomienda instalar dos servidores (principal y de respaldo) con la distribución linux Centos y el SMBD Oracle 11g.

Para Goldberger *et al.* [4], Linux es un sistema operativo con dos particularidades que lo distinguen de los demás: por un lado su

condición de gratuito, lo cual significa que no deben pagarse regalías por los derechos de uso del software a ninguna empresa; por otro lado la disponibilidad de su código fuente para modificar el programa en función de las necesidades del usuario. Ambas características lo hacen ideal para ser utilizado en las pequeñas y medianas empresas (PyMEs), reduciendo los costos operativos, moldeándose a los requerimientos de cada una de ellas.

El modelo cliente-servidor de Microsip podría adaptarse hacia aplicaciones en entornos distribuidos, sin embargo las inconsistencias referentes a la información generadas a nivel de sucursales no permiten considerar al servidor como un ente integral. Por otra parte el manejar herramientas adicionales implicaría rediseñar el sistema Microsip adaptándolo a las necesidades del sector farmacéutico, perdiendo el sentido general actualmente manejado por los creadores.

Según Allen [5] el proceso de migración entre plataformas Windows y Linux es factible, ya que algunos de los sistemas de software libre cubren los requisitos tanto de seguridad, robustez y consistencia, haciendo de ellos una alternativa inigualable para las organizaciones en constante crecimiento como son las PyMES, que no pueden pagar costosas licencias por uso de software tampoco poseerlo por medios ilegales. La primera opción sería económicamente inviable mientras que la segunda derivaría en situaciones legales complejas para la empresa.

3.1. Sincronización distribuida

El Internet permite a los usuarios de todo el mundo acceder a diversos servicios, hoy en día los Sistemas Distribuidos (SD) están por todas partes Coulouris *et al.* [6]. Cada organización administra una intranet, que provee servicios locales e Internet en ese ámbito, así como información a usuarios fuera de la red local. De esta manera se consolida una red global de información acotada y disponible a la vez.

En este sentido el principal objetivo de un sistema distribuido es facilitar a los usuarios (y a

las aplicaciones) el acceso a los recursos remotos, y compartirlos de manera controlada y eficiente Tanenbaum et al. [7]. Las razones para desear compartir recursos son muchas, la más evidente es la económica. El hecho de conectar usuarios y recursos facilita también la colaboración y el intercambio de información.

Las bases de datos distribuidas (BDD) hacen posible que una transacción acceda a datos en varios sitios. Siendo necesario que se emita un COMMIT final una vez que todos los sitios hayan completado su parte de la transacción. Para Gutiérrez [8] las aplicaciones que utilizan BDD son una respuesta a las necesidades de enfocar los sistemas de información hacia arquitecturas descentralizadas y autónomas, logrando con ello la transparencia, fragmentación, replicación y actualización de la información.

La computación distribuida o grid hace que los recursos de software y hardware que tiene una empresa puedan visualizarse y usarse como un sistema único. La integración de recursos implica la comunicación de sistemas heterogéneos compatibles solo a través de capas y niveles, algunos de ellos imperceptibles al usuario, al que se le presenta la solución en términos del entorno cotidiano donde la información se adapta al entorno tecnológico. El reto principal de los nuevos sistemas es considerar la migración de datos empleados en sistemas anteriores como parte integral de las nuevas soluciones.

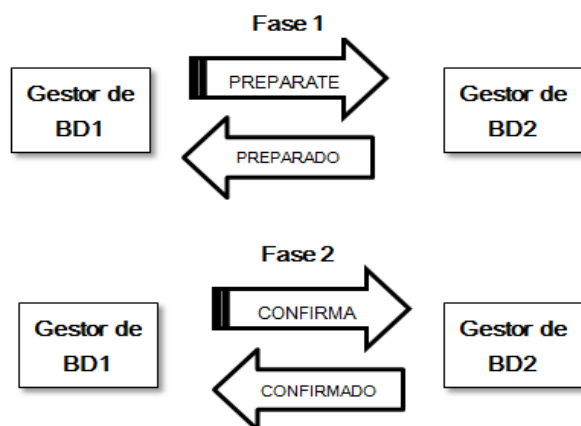


Fig. 1. Protocolo de confirmación en dos fases.

En la Fig. 1 se muestran dos gestores de bases de datos, uno local y otro remoto donde intervienen dos etapas: en la primera fase el gestor local se prepara para confirmar la operación, guardando la información para poder deshacerla si fuera necesario a través de un log histórico, y pide al gestor remoto que también se prepare para ello. Esta fase acaba cuando el segundo declara estar listo. La segunda fase se inicia cuando el gestor local confirma sus operaciones y solicita al remoto que haga lo mismo, y acaba cuando este comunica su finalización. Si el proceso se interrumpe en cualquiera de las fases por alguna razón, como la caída en las líneas de comunicación, ambos gestores deberán deshacer los cambios realizados y dejar las bases de datos (BD) como si no hubiera tenido lugar ninguna operación, garantizando la consistencia. La coordinación entre los diferentes gestores se realiza a través de los componentes de distribución cliente y servidor de los sistemas que intervengan. Será una aplicación de usuario la que inicie y dirija las diferentes fases del protocolo, éste debe confirmar la ejecución correcta de los mandatos, aunque es posible automatizar ciertos procesos por parte del administrador del SMBD.

El objetivo del protocolo COMMIT de dos fases es garantizar que todos los nodos completen su parte de la transacción, de lo contrario esta es abortada. Si uno de los nodos no completa su parte, entonces la información necesaria para restaurar la BD se encuentra en el registro de transacciones, y la BD puede ser recuperada con el protocolo DO-UNDO-REDO, dado que la información guardada en el registro se actualizó con el protocolo de escritura anticipada registrando cada transacción de manera individual y durable.

El protocolo de dos fases puede limitar la concurrencia, porque una transacción no realiza todos sus desbloques hasta que comienza la fase de contracción, y por tanto, otras transacciones que requieran alguno de los ítems bloqueados deben permanecer a la espera. Esto garantiza la seriabilidad, es decir la ejecución concurrente de varias transacciones, que generen el mismo

resultado que la ejecución en serie de las mismas sin necesidad de supervisión manual. El beneficio es la garantía de que los resultados obtenidos son los correctos.

IV. Metodología

El considerar a Linux CentOS v5 como sistema operativo a diferencia de otras distribuciones, tiene su origen en la compatibilidad y transparencia con el manejador de datos Oracle 11g; además de la comunidad activa que brinda soporte, actualización de seguridad y mejoras en el desempeño del sistema de uso libre, estable y su facilidad de configuración simplificada.

Para recabar la información y conocer la problemática real de la empresa, se usaron técnicas de metodología de investigación como cuestionarios y entrevistas, además del análisis que permitió identificar los recursos disponibles y requerimientos. Las preguntas contenidas en los cuestionarios y entrevistas fueron enfocadas a ciertos departamentos de la empresa, esto con la finalidad de detectar las capacidades técnicas y operativas del sistema actual e identificación de las áreas de oportunidad.

Se elaboró una propuesta de diseño de un software nuevo basado en herramientas de Software Libre (SL) y gratuitas para la empresa, haciendo uso del modelo Entidad-Relación-Extendido para el diseño conceptual de la BD [9], el cual permite describir la realidad mediante un conjunto de representaciones gráficas y lingüísticas, haciendo posible la representación de objetos en una BD relacional.

Para el diseño del sistema se utilizó UML, el cual permite especificar, visualizar y construir sistemas informáticos.

Se propone Java como lenguaje de programación, el cual facilita el manejo de objetos remotos y la conectividad con la BD, enfocado al uso de los SD, facilitando a los usuarios el acceso y compartición de recursos remotos de manera controlada y eficiente. Para evitar fallas en la transmisión de datos se implementó el protocolo de dos fases, garantizando que las operaciones y

transacciones entre la sucursal y el servidor se realicen de manera correcta.

De igual manera se realizaron pruebas de conectividad entre los nodos y el servidor a través de los protocolos establecidos. Esto permite simular las condiciones de eficiencia, seguridad y usabilidad de los sistemas operativos y elementos inmersos en el proceso de migración.

V. Resultados

El diseño del sistema propuesto incluye el modelado del módulo de inventarios y punto de venta, particularmente en los diagramas de casos de uso, de clases y entidad/relación-extendido (E/R-Ext) [10], expresando en lenguaje gráfico las relaciones que involucran al sistema [11], la interacción entre los módulos propuestos y sus tablas principales de forma integral.

Las rutinas probadas en esta fase involucran el envío y recepción de datos desde el manejador de datos anterior Firebird; su recepción, validación e incorporación a través de procedimientos, transacciones y triggers involucrados en la acumulación y sistematización de los datos, agrupados ahora por sucursal como clave de identificación adicional. En esta etapa los identificadores locales son redefinidos y los triggers y procedimientos contenidos en la base de Firebird son reinterpretados para confirmar las operaciones de su funcionamiento, considerando los módulos de punto de venta, inventarios, vendedores y catálogos en la base de datos en Oracle, esto permite comparar la consistencia de la información recibida versus enviada.

5.1. Autenticación entre gestores de datos

Para realizar una transacción entre dos SMD es necesario que el servidor remoto sea definido en el archivo *tnsnames.ora* e identificado como un servicio. La Fig. 2 muestra la configuración del archivo *tnsnames* en Oracle.

La variable *orcl* representa el nombre de la instancia o servicio referenciado y generado al momento de crear la base de datos, mientras que la variable *nacgen* es la liga que reconoce el

tnsames a nivel remoto, haciendo uso del puerto 1521 donde se ejecuta el socket (*listener*). Una vez establecida la conexión el envío de instrucciones SQL permitidas se establece y valida al mismo tiempo.

```

XE =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = farmal)(PORT = 1521))
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = XE)
    )
  )
NACGEN =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = nacgen.server)(PORT = 1521))
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = orc1)
    )
  )
EXTPROC_CONNECTION_DATA =
  (DESCRIPTION =
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = IPC)(KEY = EXTPROC_FOR_XE))
    )
    (CONNECT_DATA =
      (SID = PLSEXPProc)
      (PRESENTATION = RO)
    )
  )

```

Fig. 2. Configuración del archivo *tnsames*.

Una vez configurado el *tnsnames.ora*, es posible crear la liga para hacer la conexión al servidor remoto *nacgen* especificado en la Fig. 2 mediante la siguiente instrucción en el manejador de datos local Oracle 10g.

```

create public database link farmal_link
connect to usuario_farmal
identified by password_farmal
using 'nacgen';

```

Cada una de las instrucciones remotas es manejada a nivel local en la sucursal, anteponiendo la identificación de la liga generada anteriormente.

De esta manera queda ligada la información a través del modelo cliente-servidor de forma durable, formando parte de un objeto en la base de datos, permitiendo ejecutar transacciones distribuidas en dos fases a través de una referencia segura como se muestra en la Fig. 1. Para enviar peticiones se incorporan elementos seguros como la autenticación remota.

La comunicación asíncrona involucra un medio de almacenar información en caso de fallas, para

ello se utiliza una bandeja de log temporal, que se encarga de replicar la información en el servidor una vez se haya restablecido la comunicación, esta opción es reconfigurable por el administrador del sistema, pudiéndose agrupar por número de instrucciones acumuladas en la bandeja, para reintentar la sincronización de datos, no obstante estos hayan sido almacenados en la base local, la otra forma de enviar la información de la bandeja es a través de una petición directa del servidor que solicita el envío de las instrucciones pendientes de procesar desde la sucursal, verificando la seriabilidad.

5.2. Validación distribuida en dos fases

El procedimiento para validar la información distribuida consiste en interceptar las peticiones remotas de la aplicación por parte del cliente, empaquetándolas de acuerdo al formato requerido por el protocolo de dos fases, y las envía al sistema remoto. En ese punto, el servidor desempaqueta los mandatos recibidos y se los pasa a un gestor local de base de datos en el formato esperado por este.

Una vez realizadas las operaciones, los resultados son empaquetados por el servidor y enviados al cliente, que se encarga de pasarlos a la aplicación. Como la aplicación trabaja bajo un sistema que dispone de un SMBD relacional local, será este el primero que reciba las peticiones de datos, y tras comprobar que no puede atenderlas localmente, las transferirá al software encargado de la distribución remota.

Cada petición debe especificar el nombre de la base de datos a la que va dirigida y a partir de esta información, el componente cliente podrá determinar la ubicación del servidor. Para ello se dispone de una meta tabla donde se reflejan todos los posibles destinos de las peticiones que reciba. Toda incorporación de nuevas bases de datos a la red deberá quedar reflejada en los directorios de los clientes para ser reconocidas como parte de la base de datos distribuida. Un aspecto a tomar en cuenta es la integridad funcional de los datos. Si la operación es local y se produce algún error el propio sistema se lo comunica al gestor, en caso

de ser remota debe esperar la confirmación del servidor inmediato que la sirvió.

El problema de sincronización entre gestores independientes se resuelve mediante el protocolo de confirmación de dos fases. En la Fig. 3 se muestra un ejemplo del código de sincronización utilizada haciendo referencia a ligas remotas, incorporadas en instrucciones PL/SQL a través de procedimientos almacenados en Oracle 10g.

```
CREATE OR REPLACE PROCEDURE INVENTARIO.sincroniza_sucursal_servidor(
v_id_suc IN INTEGER, opInv IN INTEGER, opVenta IN INTEGER, opCxp IN INTEGER,
fecha_ini IN date, fecha_fin IN date)
AS
-- LOS parametros obligatorios son dblink, id_sucursal
-- opcionales son la tabla a sincronizar, una fecha en particular
-- se deberia borrar lo que hay en servidor pero k no esta en la sucursal
BEGIN
IF (opInv=1) THEN
insert into movto_inv@nacgen
(select * from movto_inv
where id_sucursal=v_id_suc
and (fecha between fecha_ini and fecha_fin)
and movto_inv.id_movto not in (select id_movto from movto_inv@nacgen where id_sucursal=v_id_suc
and (fecha between fecha_ini and fecha_fin)));
```

Fig. 3. Proceso de sincronización.

VI. Conclusiones

La selección de un sistema gestor de bases de datos de distribución gratuita como Oracle implementada en un servidor Linux robusto es una alternativa viable, ya que esto no representa el único costo al incurrir en la automatización de los procesos que requieran la integridad de los datos. La consistencia de la información en los nodos recae sobre el gestor local, el cual además brinda respuesta al usuario y al servidor remoto al mismo tiempo. Este sistema requiere ejecutarse sobre una plataforma estable con niveles de seguridad adecuados, al ser punto de acceso a la red de datos global [12].

La sincronización de datos a través de la implementación protocolo de dos fases, es una derivación de los SD [13] y de las BDD [8], y a pesar de la automatización utilizada se requiere de la supervisión humana para auditar y optimizar los procesos de validación en transmisión de datos, esta depende de factores externos al sistema que involucran a las comunicaciones, tecnología utilizada y fallos inherentes al diseño del sistema propuesto.

En esta etapa se considera la comunicación entre una sucursal y el servidor principal con la finalidad de mantenerlo actualizado a través de la implementación de servicios remotos RMI que permiten la comunicación horizontal entre los nodos, sin embargo no es recomendable habilitar este nivel de comunicación sin antes considerar las restricciones de seguridad como una alternativa confiable y la pertinencia de acceso a la información como parte de las políticas de la empresa.

VII. Referencias

- [1] Microsoft. <http://www.microsoft.com.mx/>
- [2] D. Reilly, M. Reilly. Java Network programming and Distributed Computing. 2002. Ed. Addison Wesley.
- [3] J. Gabillaud. Oracle 10g SQL, PL/SQL, SQL*Plus. Colección Recursos Informáticos. 2005. Ed. Ensi.
- [4] R. Goldberger. D, J. P. Incardona y L. Augusto. Linux en las PyMES. 1ª Edición. 2005. Ed. Grupo Norma.
- [5] D. Allen. Herramientas de Migración de Windows a Linux. 2005. Ed. Anaya Multimedia-Anaya Interactiva.
- [6] G. Coulouris, J. Dollimore y T. Kindberg. Sistemas Distribuidos: conceptos y diseño. 2005. Ed. Pearson Educación.
- [7] A. S. Tanenabum. Sistemas Distribuidos. Principios y Paradigmas. 2ª. Edición. 2007. Ed. Pearson Prentice Hall.
- [8] C. Gutiérrez. Diseño de un Sistema Gestor de Base de Datos Distribuida Basado en Oracle 9i. En: Revista de Ciencia, Tecnología y Medio Ambiente, Madrid, España, febrero del 2008. vol. VI.
- [9] R. Elmasri, S. Navathe. Fundamentals of Database Systems. 6ª Edición. 2010. Ed. Addison-Wesley.
- [10] J. M. Vara, *et al.* Diseño de base de datos objeto-relacionales con UML. 2005. Ed. Librería-Editorial Dykinson.

- [11] C. Date. An Introduction to Database Systems. 8ª Edición. 2004. Ed. Addison-Wesley.
- [12] M. T. Özsu, P. Valduriez. Principles of Distributed Database Systems. 3ª Edición. 2011. Ed. Springer.
- [13] F. J. Yaguez García, *et al.* Internet, TCP/IP y desarrollo de sistemas distribuidos. 2003. Ed. Fundación General de la UPM.

VIII. Autores

M. en C. Fidencio Servín Juárez obtuvo su título de Maestría en Ciencias con especialidad en Sistemas Computacionales por parte de la Universidad de las Américas-Puebla. Actualmente es profesor investigador en la Universidad de la Sierra Juárez con la especialidad en bases de datos y sistemas distribuidos.

M. en C. Leonor Adriana Cárdenas Robledo es profesora en la licenciatura en informática de la Universidad de la Sierra Juárez, realizó estudios de Maestría en Ciencias de la Computación con la especialidad en Ingeniería de software, en el Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET).

Lic. Daniel Santiago Paz realizó sus estudios de licenciatura en la Universidad de la Sierra Juárez en la especialidad de Informática, su trabajo de tesis versa sobre la temática de sistemas distribuidos y migración de datos heredados en entornos multiplataforma.

Sistema de Detección de Intrusiones basado en el Método de los Espacios de Entropía

*Pablo Velarde-Alvarado, **Alberto Martínez-Herrera, *Adalberto Iriarte-Solís

* Universidad Autónoma de Nayarit, Ciudad de la Cultura "Amado Nervo", Tepic, Nayarit, México. Tel. +52 (311) 211 88 21, email: {pvelarde, adalberto.iriarte}@uan.edu.mx

** Instituto Tecnológico y de Estudios Superiores de Monterrey, Ave. Eugenio Garza Sada 2501 Sur, Monterrey, Nuevo León, México. Tel. +52 (81) 8158 2269, email: a00798620@itesm.mx

Resumen

El Método de los Espacios de Entropía se basa en construir un espacio tridimensional del tráfico de red a nivel de flujo de paquetes. Cada punto del espacio es un valor tridimensional de la entropía de los rasgos libres para los conglomerados (*clusters*) de flujos observados durante una ranura de tráfico. La selección de rasgos para las nubes de puntos es realizada aplicando técnicas de Reconocimiento de Patrones (PR), como es el Análisis de Componentes Principales (PCA) y la Estimación de Densidad basada en Núcleos (KDE). El tráfico típico de la red es un modelo formado por una mixtura de Gaussianas (GM) y una Distribución Generalizada de Valor Extremo (GEV) que define el comportamiento de los rasgos seleccionados. Este modelo al ser integrado en un Sistema de Detección de Intrusiones Basado en Anomalías (A-NIDS) muestra ser efectivo en la detección de ataques reales llevados a cabo en una red de Área Local de un campus.

Palabras Clave: Sistema de Detección de Intrusiones, Reconocimiento de Patrones, Entropía, Análisis de Componentes Principales

I. Introducción

La detección e identificación oportuna de ataques y actividades maliciosas en computadoras y redes es un proceso fundamental de seguridad que se realiza de manera automática por un Sistema de Detección de Intrusiones a Nivel de Red o NIDS (*Network Intrusion Detection System*). Estos sistemas constituyen una segunda capa de seguridad al complementar la protección basada en políticas de filtraje de tráfico que brinda un *firewall*. Sin embargo, los *firewalls* frecuentemente no protegen contra ataques provenientes del interior de la red. Tampoco protegen conexiones que no pasen a través de él. De la misma forma ocurre cuando se trata de conexiones cifradas. Cuando hay que mantener un servicio abierto, entonces se requieren capacidades de Detección de Intrusiones. Desafortunadamente, suelen ocurrir demoras en la disponibilidad de actualizaciones para el NIDS que le permitan identificar nuevos ataques [1].

Concretamente, este tipo de desventaja está asociado a los NIDS basados en firmas o S-NIDS (*Signature-based NIDS*), tal como Snort [2]. Un S-NIDS emplea técnicas de comparación de patrones: poseen una base de datos con las firmas de ataques conocidos e intentan comparar estas firmas con el tráfico analizado, cuando se encuentra una similitud se activa una alarma. Para superar la desventaja de los S-NIDS se han desarrollado los NIDS basados en anomalías o A-NIDS (*Anomaly-based NIDS*) un ejemplo es PAYL, [3]. Un A-NIDS, durante la fase de entrenamiento requiere construir un modelo del comportamiento típico de determinados rasgos del tráfico. Posteriormente, el tráfico actual es comparado contra el modelo, donde una desviación significativa se marca como una anomalía. Un A-NIDS es capaz de detectar ataques recién liberados (*0-day attacks*) o ataques conocidos que han sido modificados. Actualmente, los S-NIDS siguen siendo más

fáciles de implementar y configurar que los A-NIDS.

Los ataques y las violaciones a la seguridad de la información parecen estar incrementándose en frecuencia e impacto. En este sentido, los S-NIDS podrían ver disminuida su capacidad de detección de ataques conforme aumenta su número y diversidad. Por ello, se espera que una mayor cantidad de A-NIDS sean desplegados para enfrentar las nuevas formas de ataque.

Para abordar las nuevas necesidades en el diseño de A-NIDS, diversos trabajos proponen métodos dirigidos a modelar el comportamiento típico de un conjunto seleccionado de rasgos del tráfico de una red. Tales modelos se resumen mediante un perfil de comportamiento basado en la entropía [4, 5, 6]. Los métodos basados en entropía están siendo propuestos como buenos candidatos para la nueva generación de A-NIDS. Una técnica de defensa basada en entropía brinda mayor sensibilidad en la detección de anomalías, esto se debe a que una actividad maliciosa cambia la aleatoriedad natural del tráfico de manera cuantificable.

En este artículo, es propuesto un A-NIDS que analiza el tráfico de red a nivel de flujo de paquetes. Un flujo de paquetes (o flujo de tráfico) es una secuencia unidireccional de paquetes con ciertas características comunes: dirección IP origen/destino, puerto origen/destino para TCP y UDP (tipo y código para ICMP ó 0 para otros protocolos), número de protocolo de nivel tres, ToS (*Type of Service*) e índice del interface de entrada, [7]. De hecho, existe libertad en la definición de un flujo; por lo tanto, para este trabajo consiste en una quintupla y un *inter-flow gap* (IFG) de 60 segundos. De esta forma, son agregados en flujos de tráfico todos aquellos paquetes de una ranura de tráfico que compartan los mismos valores de cabecera: dirección IP fuente, dirección IP destino, puerto fuente, puerto destino y protocolo. El retardo entre el primer y último paquete del flujo no debe exceder el IFG definido. Este enfoque de análisis del tráfico a nivel de flujo permite representar dicho tráfico como datos que constan de registros individuales de los flujos de tráfico. A partir de la información

generada por estos datos se propone un método para la construcción de los perfiles de comportamiento y las tareas de detección en el tráfico actual. Dicho sistema tiene una fase de entrenamiento donde los flujos son organizados en conglomerados y representados por valores tridimensionales de entropía que forman nubes de puntos. El análisis del comportamiento de las nubes de puntos es abordado mediante el Reconocimiento de Patrones. La selección y extracción de sus rasgos permitió definir un modelo para el comportamiento de la clase a la que pertenecen los conglomerados de flujos de tráfico no anómalos. De esta forma, es posible aplicar una técnica de clasificación supervisada para la identificación de las ranuras de tráfico anómalas y no anómalas.

II. Reconocimiento de Patrones

De manera general, el reconocimiento de patrones (*Pattern Recognition*, PR) se ocupa de los aspectos técnicos y matemáticos de la clasificación de diferentes objetos según sus rasgos o atributos (*features*). De acuerdo a [8], PR involucra los modelos matemáticos de objetos descritos por sus rasgos.

La terminología fundamental en este campo es la siguiente:

Patrón (*Pattern*). Es la representación física de los objetos. Pueden ser señales, imágenes o tablas de valores. Frecuentemente se refiere a los patrones como objetos, casos o muestras .

Rasgos (*features*). Los rasgos son mediciones, atributos o primitivas derivadas de las mediciones de los patrones y que pueden ser útiles para su caracterización .

Selección de rasgos (*feature selection*). Consiste en determinar cuál es el conjunto de rasgos más adecuado o relevante para describir a los patrones. La selección de rasgos relevantes, a partir del conjunto total de rasgos que describen a los patrones, se hace con dos motivos fundamentales: mejorar la clasificación y aumentar la velocidad de procesamiento.

Extracción de rasgos (*feature extraction*). Es una forma especial de reducción de dimensionalidad. Cuando la entrada de datos a un algoritmo es muy grande para ser procesada, y cuando se sospecha que es notoriamente redundante (es decir, muchos datos pero no mucha información), entonces, la entrada de datos debe ser transformada a la representación reducida del conjunto de rasgos (también llamado *vector de rasgos*).

Vector de rasgos (*feature vector*). Es la representación reducida de un conjunto de rasgos iniciales a un conjunto de rasgos útiles para la tarea de clasificación. Este conjunto de rasgos tiene una representación en forma vectorial, como se muestra en (1),

$$\mathbf{r} = r_1, r_2, \dots, r_d \in R \subset \mathbb{R}^d, \quad (1)$$

donde \mathbf{r} es el vector de rasgos y R es el dominio de dimensión d .

Espacio de rasgos (*feature space*). Es un espacio donde cada patrón es representado como un punto. Cada rasgo respecto al patrón, da una coordenada del punto en un eje del espacio. La dimensionalidad, d del espacio de rasgos es igual al número de rasgos usados, por ejemplo, si se usan dos rasgos, el espacio será un plano, con el primer rasgo en el eje x y el segundo rasgo en el eje y .

Clase (*class*). Son los estados de la naturaleza o categorías de objetos asociados con conceptos o prototipos.

Dado un patrón, su reconocimiento/clasificación puede consistir de una de las siguientes dos tareas

- 1) **Clasificación supervisada** (*supervised classification*). El patrón de entrada es identificado como miembro de una clase o una clase predefinida.
- 2) **Clasificación no supervisada** (*unsupervised classification*). El patrón se asigna a una clase hasta ahora desconocida.

El problema de reconocimiento se puede ubicar como una tarea de clasificación o categorización, donde las clases son definidas por el diseñador

del sistema (clasificación supervisada) o son aprendidas basándose en la similitud de patrones (clasificación no supervisada).

En este trabajo emplearemos el Reconocimiento estadístico de patrones (*statistical pattern recognition approach*). En este enfoque, cada patrón se denota mediante un conjunto de rasgos o mediciones y gráficamente se representa como un punto en un espacio de dimensión d . El éxito de este enfoque se basa en una selección adecuada de los rasgos, de forma tal que objetos que pertenezcan a clases diferentes ocupen regiones disjuntas (suficientemente separadas) en el espacio de rasgos. A partir de un conjunto de entrenamiento, para cada clase se tienen que establecer las regiones de decisión en el espacio de rasgos de forma que se distingan los patrones que pertenecen a clases diferentes. En el enfoque teórico de decisión estadística, las regiones de decisión se determinan mediante las distribuciones de probabilidad de los patrones que pertenecen a cada clase, la cual puede ser especificada o aprendida [9, 10].

III. Análisis de Componentes Principales

En muchas circunstancias, los datos pueden presentarse como una nube de puntos en un espacio de dimensión p . El Análisis de Componentes Principales (PCA), [11 - 13], es un método multi-variable que puede emplearse para extraer las características geométricas de tales puntos. PCA analiza un conjunto de datos (*dataset*) cuyas observaciones están descritas por una serie de variables (x_1, x_2, \dots, x_p) correlacionadas. PCA permite extraer la información más relevante del *dataset* y representarla como un conjunto de nuevas variables (y_1, y_2, \dots, y_p) , llamadas Componentes Principales que están descorrelacionadas entre sí y sus varianzas van decreciendo progresivamente, además de estar ordenadas por la cantidad de información que cada una retiene.

Cada y_j donde $j=1, 2, \dots, p$ es una combinación lineal de las (x_1, x_2, \dots, x_p) originales, es decir,

$$y_j = a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jp}x_p = \mathbf{a}_j^T \mathbf{x}, \quad (2)$$

donde $\mathbf{a}_j^T = [a_{1j}, a_{2j}, \dots, a_{pj}]$ es un vector renglón de coeficientes, $\mathbf{x} = [x_1, x_2, \dots, x_p]^T$ es un vector columna y T es la operación de transposición.

Una forma simple de maximizar la varianza podría ser aumentando el valor de los coeficientes de \mathbf{a}_j^T . Por ello, para mantener la ortogonalidad de la transformación la condición es que $|\mathbf{a}_j^T| = 1$. Es decir, $\mathbf{a}_j^T \mathbf{a}_j = \sum_{k=1}^p a_{kj}^2 = 1$. El primer componente se calcula eligiendo el vector \mathbf{a}_1 de modo que y_1 tenga la mayor varianza posible, sujeta a la condición de que $\mathbf{a}_1^T \mathbf{a}_1 = 1$. El segundo componente principal se calcula obteniendo \mathbf{a}_2 de modo que la variable obtenida, y_2 esté descorrelacionada con y_1 . De esta forma se obtienen (y_1, y_2, \dots, y_p) , descorrelacionadas entre sí, de manera que las variables aleatorias obtenidas vayan teniendo cada vez menor varianza.

IV. Entropía

La entropía se deriva de la idea de que una distribución de probabilidad representa incertidumbre, y de que algunas distribuciones tienen más incertidumbre que otras. Una medida de esta incertidumbre fue propuesta por J. C. Shannon en 1948, [14].

Para una distribución discreta de probabilidad $p(x_k)$ la entropía de Shannon, H de una variable aleatoria discreta X está dada por,

$$H(X) = -\sum_{k=1}^M p(x_k) \log_2 p(x_k), \quad (3)$$

donde M es la cardinalidad del alfabeto de X y x_k , $1 \leq k \leq M$ son los elementos del conjunto alfabeto. Una propiedad de esta función es $0 \leq H(X) \leq \log_2(M)$ en bits.

En [6] se propone el *estimador balanceado II* para la entropía de Shannon de un data set X , el cual está dado de la siguiente forma

$$\hat{H}^{bal-II}(X) = \frac{1}{N+2} \sum_{k=1}^M n_k + 1 \log \left(\frac{N+2}{n_k+1} \right), \quad (4)$$

donde n_k es la frecuencia de x_k y N es el tamaño del data set. Este estimador es una versión

mejorada de estimador balanceado propuesto en [15]. El nuevo estimador es computacionalmente más eficiente. Este tipo de estimadores buscan proveer un compromiso entre baja polarización (*low bias*) y pequeños errores estadísticos en series de datos pequeñas cuando se usa el estimador *naïve*.

V. El Método de los Espacios de Entropía

La construcción del espacio de entropía se realiza a nivel de flujo de tráfico, los espacios tridimensionales generados en condiciones de tráfico típico son utilizados para definir un perfil de comportamiento típico del tráfico de la red. El método comienza con la definición de la traza χ , la cual es dividida en m ranuras de tráfico no solapadas de duración máxima t_d segundos. En una ranura i se generan K_i flujos de paquetes, los flujos se definen bajo el concepto de 5-tupla y un *inter-flow gap* (IFG) de 60 segundos. Los cuatro campos del flujo identificados de la siguiente forma: $r=1$ para dirección IP fuente (*srcIP*), $r=2$ para dirección IP destino (*dstIP*), $r=3$ para puerto fuente (*srcPrt*) y $r=4$ para puerto destino (*dstPrt*). Posteriormente, en cada ranura i los flujos se agrupan (*clustering*) en función a uno de los cuatro campos definidos; cada campo define un *cluster key* o pivote. Ejemplificando (pero extendiendo para el resto de los campos), con $r=1$ se define el *pivote srcIP* y los *clusters* se forman con todos los flujos que tienen la misma dirección IP fuente. La cantidad de *clusters* formados dependerá de $|A_i^{r=1}|$, donde $A_i^{r=1}$ es el alfabeto de direcciones IP fuente que se observan en la ranura i . Entonces, cada *cluster* contiene los flujos que comparten la misma dirección IP fuente y con libertad de variación en el resto de campos del flujo, es decir $r=2, 3, 4$ los cuales reciben el nombre de *dimensiones libres*. La estimación de la entropía de las dimensiones libres en un *cluster* k se representa mediante el valor tridimensional $(\hat{H}_{srcPrt}, \hat{H}_{dstPrt}, \hat{H}_{dstIP})_k$, y se considera como un punto (*data-point*) en el espacio euclidiano. La cantidad de puntos de entropía en la ranura i , depende del número de

clusters formados, es decir, $|A_i^{r=1}|$. Los *data-points* de entropía en una ventana i son $(\hat{H}_{srcPrt}, \hat{H}_{dstPrt}, \hat{H}_{dstIP})_1, (\hat{H}_{srcPrt}, \hat{H}_{dstPrt}, \hat{H}_{dstIP})_2, \dots, (\hat{H}_{srcPrt}, \hat{H}_{dstPrt}, \hat{H}_{dstIP})_{|A_i^{r=1}|}$ la graficación en un *scatter plot* darán forma a datos de nubes de puntos (*Point Cloud Data*). Cuando aplicamos este procedimiento para el resto de los pivotes y para las m ranuras obtenemos los cuatro espacios de entropía de la traza χ .

En la Figura 1 se muestran los espacios de entropía para el pivote *srcIP* en tres trazas de tráfico, los ataques se llevaron a cabo en una red LAN de un campus, donde un segmento de la red se dejó vulnerable a los ataques de los gusanos Sasser y Blaster. La duración máxima de las ranuras de tráfico es de 60 segundos. En la Figura 1 (a) se muestra un patrón de *data-points* para una traza de ocho horas de tráfico típico del primer día de la semana; el patrón que forman las nubes de puntos se repite en las trazas de los cuatro siguientes días de la semana. Las Figuras 1 (b) y (c); muestran los patrones de 30 minutos de tráfico capturado durante la propagación del gusano de red Blaster y Sasser, respectivamente. Se aprecia una marcada diferencia con respecto al tráfico típico (benigno), principalmente porque los *data-points* se desplazaron hacia otras regiones del espacio. El alto volumen de flujos, generado por la propagación de los gusanos no se refleja como una mayor población de los puntos, por lo tanto, se infiere que existe una alta superposición de puntos. Los espacios de entropía son representados mediante el vector $\mathbf{X}^p \in \mathbb{R}^3$ mismo al que se aplica PCA con el fin de reducir dimensionalidad y generar un nuevo vector \mathbf{z} -

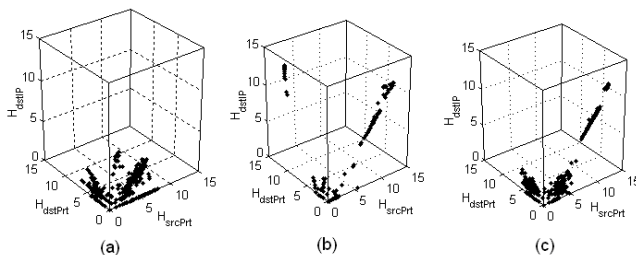


Fig. 1. Espacios de entropía para el pivote *srcIP*. (a) Tráfico estándar, (b) durante la propagación del gusano Blaster, (c) Gusano Sasser.

score $\mathbf{Z}^r \in \mathbb{R}^d, d \leq 3$.

El análisis de los *data-points* a nivel de ranuras del primer componente principal (PCA 1) se realizó mediante la estimación de densidad basada en núcleos (*Kernel Density Estimation, KDE*), usando un kernel gaussiano de 200 puntos y un ancho de banda h establecido mediante el criterio de Silverman (S1) $h = 1.06\sigma J^{1/5}$, donde J es el número de observaciones y σ es la desviación estándar de estas observaciones. Dicho análisis reveló que las ranuras de tráfico típico presentan bimodalidad gaussiana en su distribución de densidad. Los dos modos fueron identificados como *modo principal* y *modo lejano* como se ilustra en la Figura 2. El análisis de las trazas de los días de tráfico típico, mostró empíricamente que la media del modo lejano alternaba entre valores de 4.3 y -4.2 unidades del PCA 1, identificándose como modo lejano positivo y modo lejano negativo (este último con mayor frecuencia de ocurrencia). Se observó que durante el ataque del gusano Blaster, las ranuras que empezaron a contener el tráfico maligno, reflejaron una anomalía en el valor del modo lejano. Por ejemplo, las tres primeras ranuras registraron los valores de -9 , -11 y -13 unidades del primer componente principal, considerándose como anomalías dado que se alejaban del valor medio y del umbral típico señalado anteriormente, por lo tanto, esta anomalía es detectada de manera temprana. Un segundo rasgo que muestra sensibilidad para la detección de comportamiento anómalo en ranuras de tráfico es la desviación estándar del modo principal. Empíricamente se observó que la

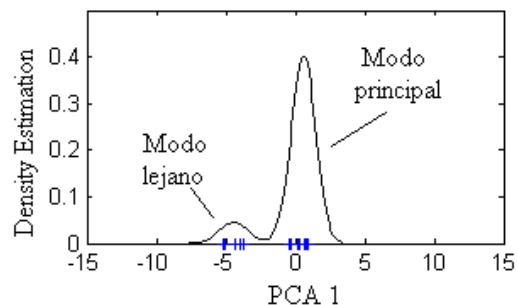


Fig. 2. La estimación de la densidad del PCA 1 presenta un comportamiento bimodal en una ranura i de tráfico.

desviación estándar del modo principal en condiciones normales es de 1.5 unidades. La propagación del gusano Sasser, originó una anomalía ya que la desviación estándar del modo principal en las ranuras experimentó un valor promedio de 0.4 unidades. En otro caso más, un ataque de escaneo que duró dos ranuras de tráfico, presentó valores anómalos en la desviación estándar del modo principal de 0.7 y 0.4 unidades.

KDE únicamente proporciona la forma de la densidad y no indica los parámetros de la misma. Sin embargo, se requiere un mecanismo para extraer cuantitativamente las dos propiedades de esta distribución que son usadas para la detección de intrusiones: la media del modo del lejano y la desviación estándar del modo principal. Esta tarea es llevada a cabo utilizando el modelo de mixtura de gaussianas (*Gaussian Mixture Model*, GMM), [16]. La implementación de GMM forma parte de *Statistics toolbox* de Matlab; *gmdistribution.fit* ajusta los datos a un modelo de mixturas gaussianas de K componentes mediante máxima verosimilitud usando el algoritmo EM (*Expectation-Maximization*) y devuelve los parámetros de la mixtura $\theta_k = [\pi_k, \mu_k, \sigma_k]_{k=1}^K$. Donde $K=3$ corresponde al número de modos de la densidad estimada mediante KDE del componente PCA 1. π_k son las proporciones de la mixtura; μ_k y σ_k son las medias y las desviaciones de las mixturas que está relacionada con la información requerida sobre los modos principales y lejanos.

Mediante el vector de mixturas θ_k , obtenido del primer componente principal y el pivote *srcIP*, se identifican los modos principal y lejano en una ranura i . El modo principal, corresponde a la mixtura k que tiene la mayor proporción π_k , es decir $\max(\pi_k)$. El modo lejano corresponde a la mixtura que presenta el mayor alejamiento, es decir $\max(|\mu_k|)$. La selección de rasgos para las tareas de clasificación de ranuras conduce a elegir a la media del modo lejano, denotada como r_1 y la desviación estándar del modo principal, denotada como r_2 . El comportamiento de estos dos rasgos, a través de las experimentaciones

realizadas, nos permite concluir que son adecuadas para separar en dos clases de tráfico: anómalo y no anómalo (o tráfico típico).

El análisis de las de tráfico típico mostró que el comportamiento para el rasgo r_1 seguía una distribución trimodal (dos modos lejanos y el principal) como se muestra en la Figura 3. El modelo que describe su comportamiento es una mixtura de gaussianas de tres componentes.

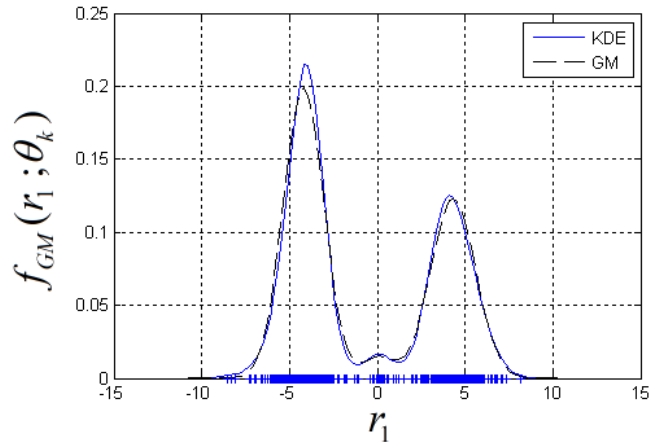


Fig. 3. Ajuste mediante mixtura de gaussianas del rasgo r_1 .

El modelo resultante del comportamiento del rasgo r_1 , es decir su perfil de comportamiento es:

$$f_{GM}(r_1; \theta_k) = \sum_{k=1}^3 \pi_k N(r_1 | \mu_k, \sigma_k^2), \quad (5)$$

$$= \sum_{k=1}^3 \frac{\pi_k}{\sqrt{2\pi\sigma_k}} e^{-\frac{r_1 - \mu_k}{2\sigma_k^2}}, \quad (6)$$

donde μ_k y σ_k representan la media y la desviación estándar del k -ésimo componente, π_k son las proporciones de mezclado, estos tres parámetros se representan mediante el vector de parámetros, $\theta_k = [\pi_k, \mu_k, \sigma_k]_{k=1}^3$. $N(r_1 | \mu_k, \sigma_k^2)$ representa una distribución gaussiana multivariable. Los valores de esta mixtura se muestran en la Tabla 1.

Para el segundo rasgo, r_2 , el análisis del conjunto de trazas con comportamiento típico mostró que su comportamiento seguía una distribución generalizada del valor extremo (*Generalized extreme value distribution*, GEV).

Parámetro	$k = 1$	$k = 2$	$k = 3$
Mezclado (π_k)	0.3946	0.5740	0.0314
Media (μ_k)	4.3239	-4.1988	0.0980
Varianza (σ_k^2)	1.6378	1.3208	0.7546

Tabla 1. Valores de los parámetros θ_k

El modelo resultante del comportamiento del rasgo r_2 , es decir su perfil de comportamiento es:

$$f(r_2; \mu, \sigma, \xi) = \frac{1}{\sigma} \left[1 + \xi * \left(\frac{r_2 - \mu}{\sigma} \right) \right]^{-1/\xi - 1} * \exp \left\{ - \left[1 + \xi * \left(\frac{r_2 - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (7)$$

para $1 + \xi * x_2 - \mu / \sigma > 0$, donde $\mu \in \mathbb{R}$ es el parámetro de localización, $\sigma > 0$ es el parámetro de escala y $\xi \in \mathbb{R}$ es el parámetro de forma. Los valores de ajuste de estos parámetros que definen el comportamiento de este rasgo son: $\xi = 0.2228$, $\sigma = 0.1221$ y $\mu = 0.9079$. El ajuste con estos valores se muestra en la Figura 4.

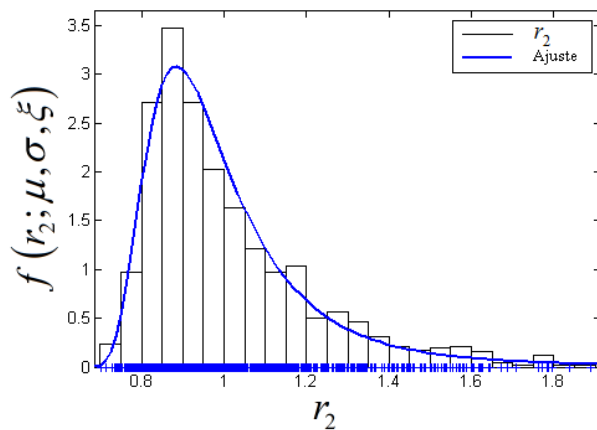


Fig. 4. Ajuste mediante distribución GEV del rasgo x_2 .

VI. Arquitectura del marco de referencia

En la Figura 5 se muestra el marco de referencia de un sistema de detección de intrusiones basado en anomalías, el cual emplea los perfiles de comportamiento basados en los espacios de entropía. Este sistema se puede dividir en dos

capas, la primera capa realiza funciones de entrenamiento de los rasgos seleccionados (r_1 y r_2) con el propósito de obtener el perfil de comportamiento descrito a través de los modelos probabilísticos (6) y (7). La segunda capa realiza las funciones de detección de intrusión, mediante el diagnóstico de los valores de los rasgos que presenta la ventana de tráfico bajo análisis contra el perfil de comportamiento del tráfico. A continuación se explican los bloques que componen el sistema.

En la capa de entrenamiento, el bloque de **pre-procesamiento** realiza el saneado de la traza para eliminar errores de captura o paquetes duplicados. También aquí se lleva a cabo el filtrado de paquetes de acuerdo al protocolo. **Generador de flujos**, en cada ranura de tráfico se generan flujos de paquetes de acuerdo a la definición de quintupla. **Clustering de flujos**, los flujos generados se agrupan de acuerdo a un pivote. La cantidad de *clusters* dependerá de la cardinalidad de los alfabetos de r . La **cuantificación de entropía** de las dimensiones libres en los clusters de flujos se realiza utilizando el estimador balanceado_II de la ecuación (4). Los *data-points* de entropía pasan a una etapa de **reducción de dimensionalidad** utilizando PCA. A partir del espacio transformado, se elige el primer componente principal del pivote srcIP y se realiza la selección de rasgos. La identificación de los rasgos se apoya en el uso del modelo GM. Identificados los rasgos útiles para la detección se procede al **perfilado** de los mismos.

En la capa de detección, el procedimiento es muy similar a excepción que no se llega a la generación del perfil; en su lugar, se extrae el valor de los rasgos para realizar el diagnóstico en función del nivel de desviación con respecto al modelo probabilístico del comportamiento típico. Un evento de red relacionado con una actividad maliciosa cambia de forma medible el comportamiento del tráfico; por ello, el sistema debe dar una respuesta oportuna alertando un cambio en la naturaleza de los flujos de tráfico observados en una ranura.

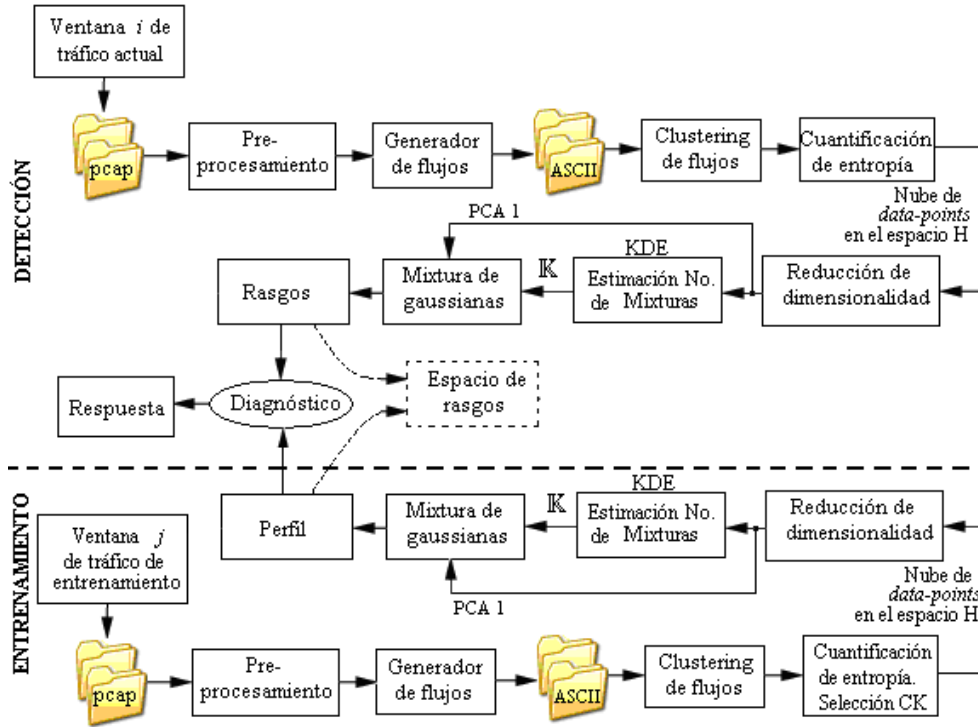


Fig. 5. Diagrama de bloques del marco de referencia para el A-NIDS

VII. Resultados

El espacio de entropía de una ranura de tráfico para el pivote srcIP, está formado por *data-points* tri-dimensionales que representan los múltiples valores de entropía de las dimensiones libres de los *cluster* de flujos. El número de puntos está determinado por la cantidad de direcciones IP fuente distintas que existen en la ranura. El primer componente principal simplifica el

análisis del comportamiento de estas dimensiones. Dos rasgos del primer componente principal logran captar oportunamente los cambios en el comportamiento del tráfico cuando ocurren los ataques. El espacio de rasgos permite observar la diferencia entre la clase de tráfico típico y la clase de tráfico anómalo. En la Figura 6 (a) se muestra el espacio de rasgos donde se distinguen dos clases de tráfico, el benigno y el originado por la propagación del gusano Blaster,

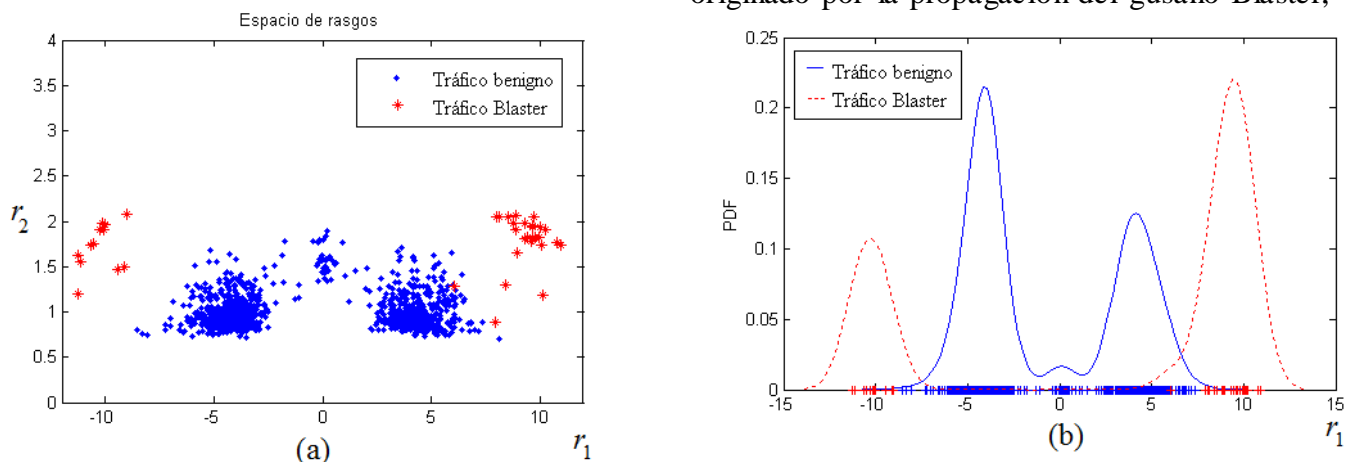


Fig. 6. (a) Espacio de rasgos con el tráfico benigno y tráfico anómalo por el gusano Blaster. (b) Distribución del rasgo r_1 para los tráficos benigno y por la propagación del gusano Blaster.

en el caso de tráfico benigno los puntos se agrupan en tres regiones; y el anómalo en dos regiones. La separación de clases permite aplicar una técnica de clasificación supervisada para identificar el tráfico normal de la red. Por otro lado, en la Figura 6 (b), se puede apreciar las distribuciones del rasgo r_1 para el tráfico típico y el anómalo originado por el gusano Blaster, observándose estadísticamente la diferencia de comportamientos. En el caso del gusano Sasser, las ranuras de tráfico durante su propagación se agrupan en una sola región, como lo muestra la Figura 7 (a), existe separación de clases, por lo tanto es posible distinguir un tráfico del otro. La Figura 7 (b), muestra la diferencia en términos de densidad del tráfico típico y anómalo del gusano Sasser cuando se observa el rasgo r_2 .

VIII. Conclusiones

En este artículo se presenta un sistema para detectar intrusiones que trabaja a nivel de flujo de paquetes. Se emplean técnicas estadísticas usadas en el Reconocimiento de Patrones. El análisis permite identificar dos rasgos que tienen las características necesarias para la identificación de ranuras anómalas asociadas con flujos de tráfico anómalo. La evaluación de los rasgos utilizando datasets de tráfico benigno y el capturado durante la propagación de ataques de gusanos y de un ataque de escaneo, mostraron la efectividad para la detección oportuna de ataques que generan variabilidad en los campos de los flujos. En base a los resultados, se propone un marco de referencia para aplicarlo en el diseño de un

Sistema de Detección de Intrusiones a Nivel de Red basado en Anomalías. Su funcionamiento se basa en la definición de los perfiles para los dos rasgos identificados. Las desviaciones del perfil son consideradas anomalías. El trabajo futuro está encaminado en contemplar nuevos escenarios de red con volúmenes de tráfico mayores, evaluar con otros tipos de ataques y mejorar la etapa de preprocesamiento incorporando un S-NIDS como Snort para eliminar ataques conocidos de forma que se dispongan de trazas que representen mejor el tráfico benigno.

IX. Referencias

- [1] T. V. Lillard. Digital forensics for network, Internet, and cloud computing: a forensic evidence guide for moving targets and data. 2010. Elsevier Inc. pp. 64
- [2] M. Roesch. "Snort - Lightweight Intrusion Detection for Networks". Proc. 13th USENIX Conference on System Administration, USENIX. 1999.
- [3] K. Wang & S. Stolfo. Anomalous Payload-Based Network Intrusion Detection. Recent Advances in Intrusion Detection, Springer Editors. 2004. pp. 203 – 222.
- [4] K. Xu K, Z. Zhang & S. Bhattacharyya. "Internet Traffic Behavior Profiling for Network Security Monitoring". Transactions on Networking, IEEE/ACM. 2008. Vol.16, No.3. pp. 1241 – 1252.
- [5] A. Wagner & B. Plattner. "Entropy Based

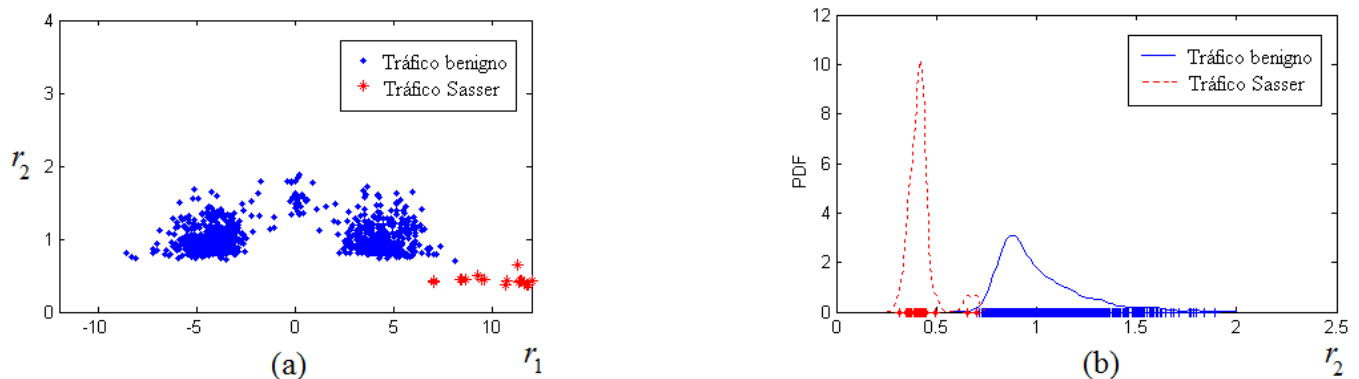


Fig. 7. (a) Espacio de rasgos con el tráfico benigno y tráfico anómalo por el gusano Sasser. (b) Distribución del rasgo r_2 para los tráficos benigno y por la propagación del gusano Sasser.

Worm and Anomaly Detection in Fast IP Networks" Proc. of the 14th IEEE International Workshop on Enabling Tech. 2005. pp. 172 – 177.

- [6] P. Velarde-Alvarado, C. Vargas-Rosales, D. Torres-Roman & A. Martinez-Herrera. "Detecting Anomalies in Network Traffic Using the Method of Remaining Elements". IEEE Communications Letters. 2009. Vol. 13. No. 6. pp. 462-464.
- [7] C. Malagón. NetFlow y su aplicación en seguridad. Ed. Red.es/RedIRIS. 2009. Boletín No. 87. pp. 33-42
- [8] J.P. Marques de Sa, Pattern Recognition: Concepts, Methods and Application, Springer, 2001.
- [9] L. Devroye, L. Györfi, G. Lugosi, A Probabilistic Theory of Pattern Recognition. Berlin: Springer-Verlag, 1996.
- [10] R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, New York: John Wiley & Sons, 1973.
- [11] R. A. Johnson, D. W. Wichern. Applied Multivariate Statistical Analysis. Prentice Hall, Englewood Cliffs, New Jersey, 1992
- [12] W. Härdle, Z. Hlavka. Multivariate Statistics: Exercises and Solutions. Springer Science+Business Media, LLC. 2007
- [13] C. M. Bishop. Pattern Recognition and Machine Learning. Springer Science+Business Media, LLC. 2006
- [14] J. C. Shannon. "A mathematical theory of communication". Bell System Technical Journal. 1948. Vol.27. pp. 379-423.
- [15] J. A. Bonachela, H. Hinrichsen & M. A. Muñoz. "Entropy Estimates of Small Data Sets". Journal of Physics A: Mathematical and Theoretical. 2008. No. 41
- [16] G. J. McLachlan and D. Peel. Finite Mixture Models. Wiley. 2000

X. Autores

Pablo Velarde-Alvarado actualmente se desempeña como profesor-investigador de tiempo completo en el Área de Ciencias Básicas e Ingenierías de la Universidad Autónoma de Nayarit. Recibió el título de ingeniero en electrónica por la Universidad Autónoma de Guadalajara en 1993. Los grados de maestría y doctorado en ciencias por parte del Centro de Investigación y Estudios Avanzados del IPN (CINVESTAV-IPN) en los años 2001 y 2009, respectivamente. Es miembro del Sistema Nacional de Investigadores Nivel Candidato, y cuenta con Reconocimiento de Perfil Deseable PROMEP. Sus principales líneas de investigación están relacionadas con el modelado del tráfico IP y el diseño de modelos estadísticos basados en entropía para sistemas de detección de intrusiones.

Alberto Martínez-Herrera está en su segundo año de Doctorado en Tecnologías de la Información y Comunicaciones en el Instituto Tecnológico y Estudios Superiores de Monterrey, (ITESM), Campus Monterrey. Es ingeniero en telecomunicaciones por parte de la Universidad Autónoma de Hidalgo. Sus principales líneas de investigación están enfocadas en la criptografía y en sistemas de seguridad de redes, principalmente sobre el protocolo DNS.

Adalberto Iriarte-Solís actualmente se desempeña como profesor-investigador de tiempo completo en el Área de Ciencias Económicas y Administrativas de la Universidad Autónoma de Nayarit, cuenta con Reconocimiento de Perfil Deseable PROMEP. Estudió un doctorado en Educación con especialidad en Tecnología Instruccional y Educación a Distancia en la Nova Southeastern University (NSU) en Miami, Florida. Sus líneas de investigación incluyen desarrollo de proyectos de investigación enfocados al mejoramiento de la infraestructura física y de seguridad en redes de voz y datos, así como el análisis y diagnóstico del uso de las herramientas tecnológicas actuales para la capacitación y formación de estudiantes y docentes.

Representación formal del conocimiento para apoyar la adopción de sistemas de información clínica

***Víctor H. Castillo, **Ana I. Martínez-García, *JRG Pulido, *José Luis Álvarez Flores, *Leonel Soriano Equigua, *Nicandro Farías Mendoza**

*Universidad de Colima, Km. 9 Carretera Colima-Coquimatlán, C.P. 28400, Coquimatlán, Colima, México.
victorc@ucol.mx.

**Centro de Investigación Científica y Educación Superior de Ensenada, Carretera Ensenada-Tijuana No. 3918, Zona Playitas, C.P. 22860, Ensenada, B. C. México. martinea@cicese.mx

Resumen

El problema de la adopción de los sistemas de información clínica (SIC) ha mostrado persistencia en entornos multiculturales. Ese problema tiene repercusiones sociales y económicas importantes. Se han establecido algunas propuestas para solventar esa problemática, las cuales se basan en el uso de conocimiento que disminuye la incertidumbre hacia innovaciones en esos sistemas. Sin embargo, no existe trabajo que reporte una estructura de representación formal de ese conocimiento, la cual ayude a caracterizarlo, entenderlo y aplicarlo. Este artículo expone y describe una ontología para apoyo a la adopción de los SIC, la cual ha sido evaluada mediante un enfoque cualitativo-heurístico. Los resultados de su evaluación la describen como una estructura de representación del conocimiento que sustenta una eventual mejora de los servicios de salud y una disminución en sus costos.

Palabras Clave: sistemas basados en conocimiento, ingeniería de conocimiento, adopción de tecnología, informática médica, administración de cambio.

I. Introducción

Varios países consideran que el sector del cuidado de la salud (SCS) debe ser uno de los aspectos centrales en su política pública [1, 2, 3]. Se estima que para el año 2050 el 21% de la población del planeta tendrá más de sesenta años. Ese hecho hará necesarios más servicios de salud, previendo esta situación, por ejemplo, el gobierno de China planea construir 3000 nuevos hospitales y el de los Estados Unidos de Norteamérica estima una inversión del 25% del producto interno bruto para 2025 en este sector [2]. Los Estados Unidos Mexicanos plantean también prioritario ese sector, como puede constatarse en su Plan Nacional de Desarrollo 2007-2012 [3].

Los estudios sobre el SCS identifican que hay aspectos a ser tratados en esta área, tales como: procesos complejos no coordinados, o información no disponible cuando es necesaria, aspectos que en ocasiones provocan que los

pacientes no reciban el cuidado adecuado en el momento requerido [4]. Una oportunidad, entre otras, para mejorar la calidad de los servicios de salud está relacionada con la adopción de sistemas de información clínica (SIC), ya que estos pueden favorecer la automatización de procesos susceptibles a fallas, el respaldo a toma de decisiones, o el apoyo para brindar información oportunamente al personal médico que la necesite. Sin embargo, incluso en países que se caracterizan por sus avances tecnológicos, como los Estados Unidos de Norteamérica, la adopción de SIC, y tecnología médica en general, es baja [5, 6], se reporta que la tasa de adopción de aplicaciones de registro médico electrónico en ese país proyecta, en un escenario optimista, un valor apenas cercano al 50% en 2011 [7]. Por situaciones como la anterior, el fenómeno de la adopción de sistemas de información [8, 9], y de SIC en particular [10, 11, 12], ha sido foco de atención en trabajos de investigación. En el

contexto del problema referido anteriormente, el término adopción se refiere al uso completo de las características de una innovación con el convencimiento de que ese es el mejor curso de acción disponible para la realización de una determinada actividad [13]. No obstante, considerando las ventajas de la tecnología de la información y comunicación (TIC) -relativas a reducción de costos y mejora de servicios cuando es usada para automatizar procesos- en la literatura se reportan pocos estudios que proponen enfoques automáticos, basados en TIC, para apoyar la adopción de SIC [14, 15]. Considerando que el proceso que lleva a la decisión de adoptar una tecnología innovadora en una organización es esencialmente una actividad de búsqueda y procesamiento de información en el que un individuo es motivado a reducir la incertidumbre sobre las ventajas y desventajas de tal innovación [13], algunas propuestas [14, 15] establecen que es posible que los enfoques para tratamiento de información y conocimiento puedan brindar apoyo al proceso de adopción de SIC. Sin embargo, tales propuestas no plantean las estructuras formales para representar el conocimiento que ayude a reducir la incertidumbre hacia innovaciones en tales SIC y, por lo tanto, a facilitar el proceso de adopción de los mismos.

Dentro de las herramientas para representación formal de conocimiento, las ontologías desempeñan una solución pertinente [16]. Una ontología es una representación formal de conocimiento que establece acuerdos acerca de asunciones compartidas y modelos del mundo que nos rodea [17]. En este artículo se define una ontología que modela el proceso de apoyo a la adopción de los SIC, y está organizado como se explica a continuación. En la segunda sección se describe un proceso de apoyo a la adopción de los SIC. La tercera sección detalla la ontología propuesta, posteriormente, en la cuarta se describe el proceso de validación de la misma para finalizar con la discusión del trabajo expuesto en este artículo en la quinta sección. Las

conclusiones respecto al mismo son expuestas en la sexta sección.

II. Un proceso de apoyo a la adopción de los sistemas de información clínica

Desde una perspectiva sistémica [18], un fenómeno puede embeberse en un modelo conceptual que eventualmente origina un modelo de flujo de información. Ese flujo, a su vez, puede contribuir a la elaboración de sistemas de información. Debido a eso pueden transformarse modelos conceptuales en sistemas de información, y viceversa. Así, con base en una arquitectura de software para apoyar el desarrollo de sistemas que facilitan la adopción de SIC [14] desarrollamos un proceso que apoya a los médicos en la adopción de tales sistemas. El proceso se basa en escenarios de interacción para definir tres elementos importantes: 1) los requerimientos de conocimiento necesarios para apoyar a los médicos en la adopción de SICs; 2) el tipo perfil de usuario del médico que requiere ese conocimiento; y 3) las circunstancias bajo las cuales le será otorgado tal conocimiento.

El proceso se conforma de cinco etapas. La primera ayuda al médico a hacerlo consciente de una innovación en el SIC, por ejemplo, de un cambio en la interfaz de usuario del sistema. En la segunda etapa el proceso establece que debe asistírsele al médico en la promoción de interacción social con el fin de que disminuya la incertidumbre hacia la innovación, lo cual se puede hacer proporcionándole automáticamente los contactos que apoyan esa interacción. La tercera etapa apoya al médico a que realice actividades de autoaprendizaje, lo cual se puede llevar a cabo proporcionándole los documentos que le ayuden a disminuir la incertidumbre hacia una innovación en el SIC. Por otra parte, en la cuarta etapa el usuario evalúa, con respecto a una situación específica, las ventajas y desventajas de una innovación en términos de una red social (a través de contactos) y/o autoaprendizaje (basados en los documentos que le han sido facilitados previamente). Finalmente, en la quinta etapa se

verifica si el médico tiene una decisión favorable hacia la adopción de la innovación en el SIC o una de rechazo. Adoptar significaría que aún ante un evento que causa incertidumbre en el uso del SIC, el usuario ha decidido utilizar todas sus características funcionales.

El diseño del proceso explicado previamente fue la base para la definición de una ontología para apoyo a la adopción de los SICs.

III. Una ontología para apoyar el proceso de adopción de los sistemas de información clínica

Las etapas definidas en el proceso para apoyo a la adopción de los SIC, descrito en la sección previa, pueden representarse mediante una ontología, la cual se ilustra en la Fig. 1.

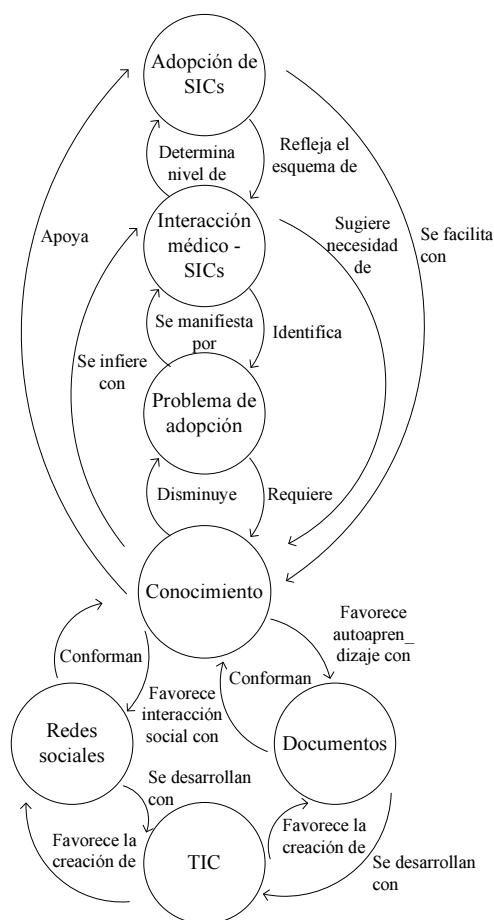


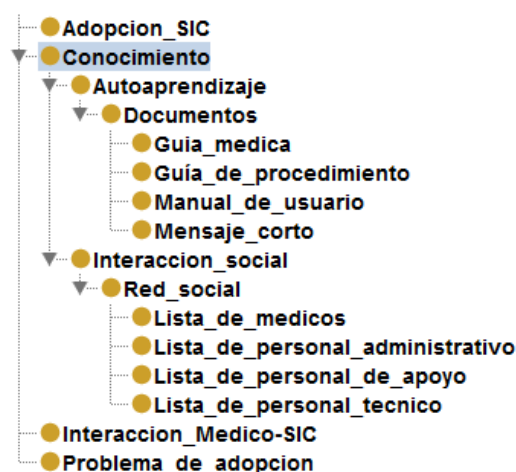
Fig. 1. Ontología para apoyar la adopción de los sistemas de información clínica.

La ontología describe que el problema de la adopción de SIC se manifiesta a través del análisis de la interacción médico-SIC. Ésta determina el nivel de adopción de tales sistemas. Asimismo, la adopción de los SIC se refleja por un esquema de interacción médico-SIC que identifica el problema de adopción de un médico. Tal problema requiere de conocimiento para ser solventado, el cual favorece tanto el autoaprendizaje, como la interacción social, aspectos fundamentales para disminuir la incertidumbre hacia una innovación. Ese conocimiento puede proporcionarse a través de documentos o con la disposición de redes sociales, elementos que pueden facilitarse con TIC. Así, como se aprecia en la ontología de la Fig. 1, ese conocimiento disminuye un problema de adopción de un SIC. Como se observa en esa figura, la adopción de SIC se facilita con conocimiento, el cual se infiere a través de la interacción médico-SIC.

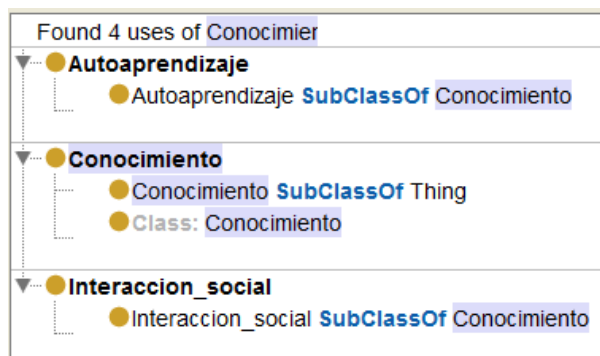
Debido a que la ontología diseñada sirve para formalizar conocimiento destinado a realizar una tarea particular –en este caso, para apoyar la adopción de los SIC–, ésta se clasifica como una ontología de aplicación. Para su diseño utilizamos la metodología Methontology [19], la cual establece seis fases que pueden implementarse en algún modelo de ciclo de vida: especificación, conceptualización, formalización, integración, implementación, y mantenimiento. Para el desarrollo de la ontología utilizamos un modelo de ciclo de vida incremental. Con base en Methontology, las fases de especificación, conceptualización y formalización se fundamentaron en el proceso para apoyo a la adopción descrito en la sección previa de este documento. Debido a que en la literatura no se reportan ontologías para el dominio de aplicación tratado en este artículo, la fase de integración no fue realizada en el proceso de diseño. Por otra parte, el proceso de implementación de la ontología se llevó a cabo con la plataforma Protégé [20]. La Fig. 2 ilustra la descripción de la jerarquía de clases mediante un árbol (ver Fig.

2a) y a través de una descripción formal (ver Fig. 2b). Como se observa en la Fig. 2a, la interacción social puede favorecerse mediante una lista de personal técnico o de médicos, mientras que el autoperendizaje se favorece a través de documentos como manuales y comunicados.

El enfoque de evaluación utilizado para valorar la ontología se centró en dos aspectos: 1) implementación de la ontología en una aplicación [21], esto es, a través de su implementación en Protégé; y 2) por las personas contempladas en el desarrollo de la misma, quienes intentan evaluar qué tan bien cumple con un conjunto predefinido de criterios, estándares, y requerimientos [22]. El primero de los aspectos de evaluación ayudó a la verificación de la arquitectura, mientras que el segundo apoyó la validación de la misma.



(a)



(b)

Fig. 2. Jerarquía de clases para la ontología diseñada (interfaz visual de Protégé).

Para la verificación, la plataforma de implementación Protégé ayudó a garantizar la exactitud de la ontología; mientras que para la validación, la opinión de ingenieros de conocimiento y médicos garantizó que la ontología pueda representar el dominio de aplicación en forma correcta.

IV. Resultados

La cantidad de conceptos en la ontología para apoyo a la adopción de los SICs, y su codificación, se muestra en la Tabla 1. En ésta puede observarse que se definieron 22 clases en total, de las cuales una pertenece a la identificación del problema de adopción, 12 a la descripción del tipo de conocimiento necesario para apoyar la adopción, cinco para identificar el tipo de interacción entre el médico y el SIC, y cuatro para describir el problema de adopción.

	Código	Número de clases
Adopción_SIC	ADxx	1
Conocimiento	COxx.xx	12
Interacción Medico-SIC	INxx	5
Problema de adopción	PRxx	4
-	-	22

Tabla 1. Principales clases y propiedades de la ontología.

El proceso de verificación se llevó a cabo con OWL Lint [20], el cual es un componente de extensión de Protégé que ayuda a llevar un control de calidad y depuración de una ontología. Este proceso no detectó alguna clase sin unir o redundancia de los conceptos definidos.

Por otra parte, el proceso de validación de la ontología se llevó a cabo mediante un enfoque cualitativo-heurístico que involucró a un experto en ingeniería de conocimiento y dos médicos con experiencia en el uso de sistemas de expediente médico electrónico y de tecnología de la información y comunicación. Para tal proceso primeramente se les mostró, al ingeniero y a los

médicos, una descripción del proceso para apoyo a la adopción de SIC. Después se les expuso una descripción de la ontología basada en ese proceso. Posteriormente se les mostraron cuatro escenarios de interacción a través de los cuales determinaron la utilidad de la ontología para representar el conocimiento que ayude a apoyar la adopción de un SIC. Un escenario de interacción describe la forma en que la tecnología de la información y comunicación ayuda a una persona a resolver un problema en un dominio de aplicación [23]. Al final de la explicación de los escenarios, el ingeniero y ambos médicos respondieron un cuestionario para evaluación de escenarios basado en el trabajo de Lewis [24]. Los participantes afirmaron consistentemente que la ontología describe el uso del conocimiento necesario para apoyar la adopción del sistema de expediente médico electrónico ilustrado en los escenarios de interacción.

V. Discusión

El apoyo a la adopción de los SIC representado en la ontología propuesta en este artículo se fundamenta en el uso de documentos y redes sociales. Estos dos artefactos, entregados al usuario de SIC de acuerdo a su entorno situacional, conforman conocimiento y han sido descritos como importantes para apoyar el proceso de adopción de los SIC [14, 15]. Esa clase de conocimiento se considera explícito [25], es decir, altamente estructurado, que puede administrarse y, por tanto, proporcionarse mediante TIC. El conocimiento explícito es fácil de transmitir y de adquirir [26].

Los resultados del presente trabajo presentan algunas consideraciones que deben tomarse en cuenta al extenderlos a otros contextos. La muestra por conveniencia, utilizada para elegir a los sujetos que evaluaron la ontología, tiene la desventaja de que no permite extender, con un nivel de confianza conocido, los resultados reportados en este artículo. Sin embargo, la selección de esa muestra tiene el beneficio de que los sujetos elegidos para la evaluación tienen un grado elevado de experiencia en modelos de

representación formal, por un parte, y en sistemas de información clínica, por otra. Eso, conjuntamente con el proceso de verificación realizado con la ayuda de una herramienta de software, da pertinencia a los resultados que exponemos en el presente trabajo.

VI. Conclusiones

En este artículo se expone una ontología para representar el conocimiento que apoye la adopción de los SIC. El fundamento conceptual de la misma se basa en un proceso para apoyo a la adopción de tales sistemas [14, 15]. Los resultados de evaluación muestran la exactitud de la ontología, así como un alto grado de aceptación por personal que se dedica a la ingeniería del conocimiento y a la utilización de dichos SIC. Consideramos que el diseño de la ontología para apoyo a la adopción de esos sistemas proporciona un precedente para el desarrollo de aplicaciones de software que apoyen automáticamente la adopción de los sistemas de información clínica, lo cual eventualmente redundará en la mejora de los servicios de salud y en la disminución de sus costos asociados.

VII. Referencias

- [1] B. Middleton, *et al.*, "Accelerating U.S. EHR Adoption: How to Get There From Here. Recommendations Based on the 2004 ACMI Retreat," *J Am Med Inform Assoc*, vol. 12, pp. 13-19, January 1, 2005.
- [2] E. Schwartz. (2005, September 12, 2005) Digital health and privacy. When Intel start worrying about your health, should yo be worried? *InfoWorld*. 8.
- [3] Presidencia de la República. (2007, 28 de agosto de 2007 en http://pnd.calderon.presidencia.gob.mx/pdf/PND_2007-2012.pdf). *Plan Nacional de Desarrollo 2007-2012, Estados Unidos Mexicanos. Eje rector 3: Igualdad de oportunidades, estrategia 5.3.* .

- [4] D. W. Bates, "The quality case for information technology in healthcare," *BMC Medical Informatics and Decision Making*, vol. 2, pp. doi:10.1186/1472-6947-2-7, 2002.
- [5] J. S. Ash and D. W. Bates, "Factors and Forces Affecting EHR System Adoption: Report of a 2004 ACMI Discussion," *J Am Med Inform Assoc*, vol. 12, pp. 8-12, January 1, 2005.
- [6] E. G. Poon, *et al.*, "Assessing the level of healthcare information technology adoption in the United States: a snapshot," *BMC Medical Informatics and Decision Making* vol. 6, pp. doi:10.1186/1472-6947-6-1, 2006.
- [7] E. W. Ford, *et al.*, "Resistance Is Futile: But It Is Slowing the Pace of EHR Adoption Nonetheless," *J Am Med Inform Assoc*, vol. 16, pp. 274-281, May 1, 2009.
- [8] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13(3) pp. 318-340, 1989.
- [9] V. Venkatesh, *et al.*, "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly*, vol. 27(3), pp. 425-478, September 2003.
- [10] K. Zheng, *et al.*, "Understanding technology adoption in clinical care: Clinician adoption behavior of a point-of-care reminder system," *International Journal of Medical Informatics*, vol. 74, pp. 535-543 August 2005.
- [11] D. E. Burke, *et al.*, "Exploring Hospitals' Adoption of Information Technology," *Journal of Medical Systems*, vol. 26(4), pp. 349-355, August 2002.
- [12] L. Poissant, *et al.*, "The Impact of Electronic Health Records on Time Efficiency of Physicians and Nurses: A Systematic Review," *J Am Med Inform Assoc*, vol. 12, pp. 505-516, September 1, 2005.
- [13] E. Rogers, *Diffusion of innovations*, 5th ed. New York: Free Press, 2003.
- [14] V. Castillo and A. I. Martínez, "A knowledge management architecture for supporting the adoption of clinical information systems," in *Eighth Mexican Int. Conf. on Computer Science (ENC 2007)*, Morelia, México, 2007.
- [15] V. Castillo, *et al.*, "A knowledge-based taxonomy of critical factors for adopting electronic health record systems by physicians: a systematic literature review," *BMC Medical Informatics and Decision Making*, vol. 10, p. 60, 2010.
- [16] M. Bosch, "Ontologies, different reasoning strategies, different logics, different kinds of knowledge representation: working together," *Knowledge Organization*, vol. 33, pp. 153-159, 2006.
- [17] T. R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing," *Int. J. Human-Computer Studies*, vol. 43, pp. 907-928, 1995.
- [18] P. Checkland and J. Scholes, *Soft systems methodology in action*, 1st ed ed. West Sussex, UK: Wiley, 1999.
- [19] M. Fernandez, *et al.*, "Methontology: from ontological art towards ontological engineering," in *AAA197 Spring Symposium Series on Ontological Engineering*, Standford, USA, 1997, pp. 33-40.
- [20] Stanford Center for Biomedical Informatics Research. (2011, Junio 24). *Protégé*. Available: <http://protege.stanford.edu/>
- [21] R. Porzel, Malaka, R., "A task-based approach for ontology evaluation," presented at the ECAI 2004 Workshop Ont. Learning and Population, 2004.
- [22] A. Lozano-Tello and A. Gómez-Pérez, "Ontometric: A method to choose the appropriate ontology," *J. Datab. Mgmt.*, vol. 15, pp. 1-18, 2004.
- [23] M. B. Rosson and J. M. Carroll, *Usability engineering: scenario-based development of human computer interaction* 1st ed. San Francisco, CA: Morgan Kaufmann, 2002.

- [24] J. R. Lewis, "IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use," *International Journal of Human-Computer Interaction*, vol. 7, pp. 57-78, 1995.
- [25] I. Nonaka and H. Takeuchi, *The knowledge-creating Company. How Japanese Companies Create the Dynamics of Innovation*. New York: Oxford University Press, 1995.
- [26] D. Leonard, "Knowledge transfer within organizations," in *Knowledge creation and management: new challenges for managers*, K. Ichijo and I. Nonaka, Eds., 1st ed New York, USA: Oxford University Press, 2007, pp. 57-68.

VIII. Autores

Víctor H. Castillo es Doctor en Ciencias de la Computación por el Centro de Investigación Científica y Educación Superior de Ensenada (CICESE), México. Tiene experiencia en modelado formal y desarrollo de software. Actualmente labora como profesor-investigador de tiempo completo para la Facultad de Ingeniería Mecánica y Eléctrica de la Universidad de Colima. Sus intereses de investigación se relacionan con los sistemas basados en conocimiento aplicados a la administración de cambio.

Dra. Ana I. Martínez-García realizó sus estudios de Licenciatura en Ciencias Computacionales en la Facultad de Ciencias de la Universidad Autónoma de Baja California. Obtuvo el grado de Maestría y Doctorado en Ciencias de la Computación de la Universidad de Manchester en Inglaterra. Actualmente es Investigadora titular del Departamento de Ciencias de la Computación del CICESE y miembro del Sistema Nacional de Investigadores. Sus áreas de interés incluyen: informática médica, la aplicación de la tecnología en la mejora de procesos, cómputo móvil e interacción hombre-computadora.

D. en C. JRG Pulido obtuvo su título de Doctor en Ciencias con especialidad en Computación en la Universidad de Nottingham, Inglaterra. Actualmente labora como profesor investigador de tiempo completo en la Facultad de Telemática de la Universidad de Colima.

M. en C. José Luis Álvarez Flores obtuvo su título de Maestría en Ciencias en la Universidad de Colima, México, donde labora actualmente como profesor investigador de tiempo completo en la Facultad de Ingeniería Mecánica y Eléctrica.

M. en C. Leonel Soriano Equigua obtuvo su título de Maestría en Ciencias con especialidad en Electrónica en CICESE, México. Actualmente labora como profesor investigador de tiempo completo en la Facultad de Ingeniería Mecánica y Eléctrica en la Universidad de Colima.

D. en C. Nicandro Farías Mendoza obtuvo su título de Doctor en Ciencias con especialidad en Computación en el Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Guadalajara, México. Actualmente labora como profesor investigador de tiempo completo en la Facultad de Ingeniería Mecánica y Eléctrica de la Universidad de Colima.

Sistemas de Rastreabilidad en el Piso de Fabricación Basados en SOA (ServiceOrientedArchitecture)

***Wendy Janett Guzmán González, **Cynthia Liliana Guzmán González, ***Juan Alberto Luna Sandoval**

*Universidad Autónoma de Chihuahua, Circuito # 1 nuevo campus universitario,6144429500,
wguzman@uach.mx

**Universidad Autónoma de Chihuahua, Circuito # 1 nuevo campus universitario, 6144429500,
cyguzman@uach.mx

***Instituto tecnológico de Jiménez, Ave. Tecnológico s/n, 6295423404, juanlunas@gmail.com

Resumen

Cada vez más las organizaciones dependen de su infraestructura de IT (Tecnologías de Información) para alcanzar sus objetivos. Pero en un entorno competitivo como el actual, aprovechar las oportunidades de negocio exige moverse con rapidez, con frecuencia IT no permiten estas respuestas rápidas ni disponen de la flexibilidad necesaria para competir de forma efectiva.

En el presente estudio se tiene como finalidad el desarrollo de una herramienta informática para el control de producción de empresas manufactureras aplicado al piso de fabricación con el objetivo de dotar de mayor agilidad a la organización, principalmente a la hora de adaptarlos adecuadamente a las nuevas necesidades.

Se exhibe también la adaptación de un sistema de rastreabilidad interna de productos a una Arquitectura Orientada a Servicios SOA (*ServiceOrientedArchitecture*) para el control de la lista de materiales BOM (*Bill Of Materials*) de cada producto, se crean de manera dinámica grafos para la representación y almacenamiento; esto es posible con el uso del control ActiveX AddFlow. Así mismo se expone la implementación de los servicios los cual se desarrolla con la herramienta WCF (*Windows CommunicationFoundation*) desarrollados bajo C# .NET 2010.

Palabras Clave: SOA, WCF, Servicio Web, lista de materiales

I. Introducción

En todo lugar donde se realice un proceso de transformación, llámese piso de fabricación, planta manufacturera, taller, se realizan actividades encaminadas a la transformación de piezas, partes o subproductos de un artículo para conformar un todo. A este todo denominado producto terminado, lo compone un cierto número de subensambles los cuales a su vez cada uno de ellos está constituido de elementos individuales únicos y básicos. Estos elementos individuales, son las partes esenciales que se deben de observar en su

recorrido de transformación desde que son parte de la materia prima, hasta que son partes o productos terminados.

Implantar un sistema de rastreabilidad que permita trazar y recuperar dichos registros para reconstruir el camino de la materia prima en la obtención o fabricación del producto desde su entrada hasta la entrega al distribuidores de gran ayuda, puesto que, es un enorme apoyo a las actividades encaminadas al control de calidad en el proceso de producción.

Por otra parte desarrollar el sistema informático utilizando una Arquitectura Orientada a Servicios permite alinear los recursos IT de forma más directa con los objetivos del negocio, ganando así un mayor grado de integración, proporcionando una inteligencia de negocio más precisa y más accesible con la cual se podrán adoptar nuevas decisiones, y ayudar a las empresas a optimizar sus procesos internos y flujos de información para mejorar la productividad individual, donde el resultado neto es un aumento notable en la agilidad de la organización.

II. Desarrollo

Una de las principales controversias que aparecen a la hora de renovar las aplicaciones en una empresa es la decisión de implantar una solución hecha a la medida o acudir a un paquete de software de aplicación. Aunque cada caso es distinto, en líneas generales la opción de paquete de software de aplicación tiene un menor TCO (*Total Cost of Ownership*), pero por otra parte se pierde la capacidad de diferenciación y adaptación.

SOA resuelve este problema permitiendo mejorar la capacidad de diferenciación de las soluciones a través de la configuración de los procesos y la libertad de acudir a cualquier tipo de servicio. De esta manera los paquetes de software de aplicación están empezando a evolucionar desde un modelo “cerrado” con mayor o menor grado de parametrización, a un modelo “abierto” donde se ofrece una librería de servicios y un conjunto de procesos base con un alto grado de parametrización mediante motores de reglas.

Trazabilidad

Se refiere a sistemas de registros destinados a seguir el flujo de productos o atributos de productos a través del proceso de producción y/o la cadena de suministros. [1]

Los sistemas de trazabilidad se caracterizan por su amplitud, penetración y precisión. La

amplitud describe la cantidad de información que el sistema registra. La penetración describe cuán lejos hacia arriba o hacia abajo en el proceso registra el sistema. La precisión refleja el grado de seguridad con el que el sistema puede localizar los movimientos o características del producto. [2]

Rastreabilidad

Según la norma ISO 22000 la define como: la capacidad de seguir el movimiento de un producto a través de etapas específicas de producción, procesado y distribución. Y aclara que el movimiento se puede realizar con el origen de los materiales, la historia del procesado o la distribución, pero recomienda que esté limitado a “un paso adelante y un paso atrás” en la cadena de suministro. [3]

Haciendo un bosquejo de los principales sistemas comerciales de rastreabilidad interna que actualmente se encuentran en el mercado se desarrollo el diagrama presentado en la Fig. 1 donde se describe las funciones mínimas que todo sistema de rastreabilidad debe de contener, lo anterior se obtuvo haciendo una selección de lo que los sistemas analizados tienen en común. Es importante señalar que el primer paso en un sistema de rastreabilidad es el mantenimiento de registros para cada proceso, desde la recepción de la materia prima hasta el embarque del producto terminado.

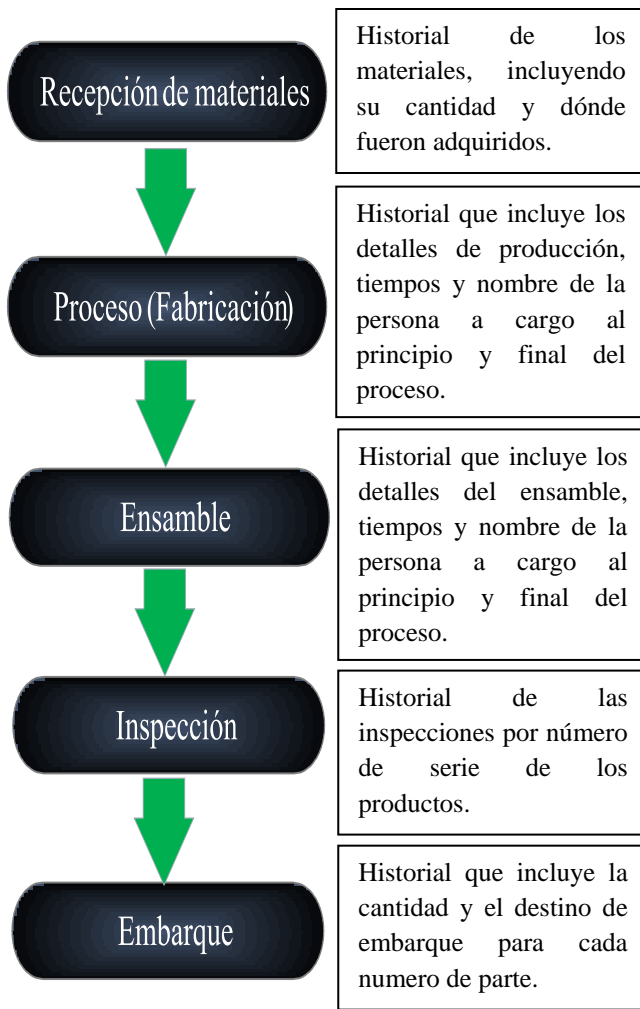


Fig. 1. Funciones de un sistema de rastreabilidad

BOM (bill of materials)

También conocida como estructura de cada producto final, es decir, el conocimiento exacto de los componentes y materiales que lo integran en clase y cantidad.[4]

Un producto puede estar formado por dos o más elementos, a su vez cada elemento está formado por otros subelementos convirtiéndolos en productos finales, subensables o subproducto.[5]

Los Diagramas pueden ser usados para proveer representación visual de cómo los productos finales están compuestos de ensambles y componentes. La Fig.2 muestra cómo el producto final P1 comprende los componentes C1 y C2 así como el ensamble A1. El ensamble

A1, consiste de los componentes C1 y C3. Los componentes C1 y C3 son externos, mientras que el componente C2 es producido del material M1. En la Fig.3 se explica el significado de cada uno de los símbolos usados en el grafo.

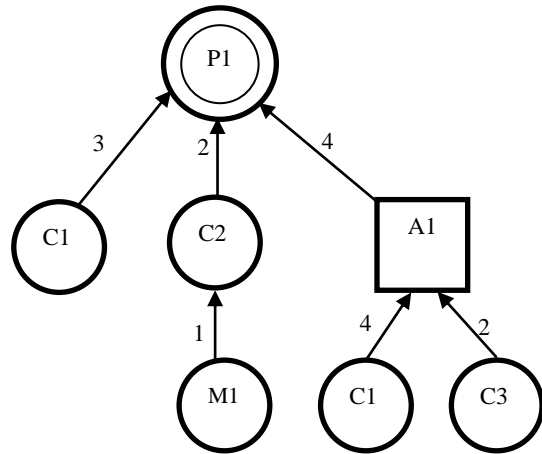


Fig.2. Representación grafica de un producto (grafo)

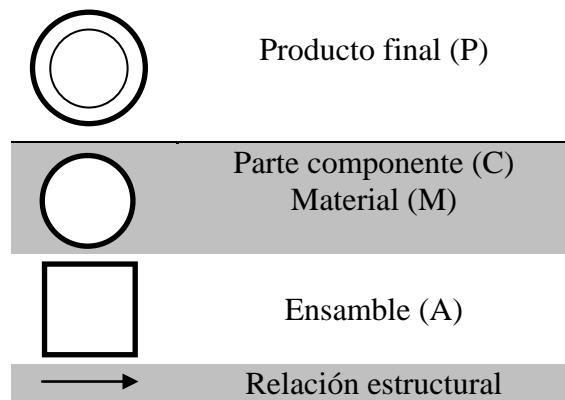


Fig.3. Simbología utilizada

Buscando la manera de almacenar en una base de datos la información de cada grafo, considerando la posibilidad de manejar productos que contengan un gran número de variantes y además contemplar que la base de datos sea independiente del producto a fabricar,

es decir, para cualquier producto, se obtuvo el diseño de tablas y sus relaciones presentado en la Fig. 4 de tal manera que pueda ser desplegado en pantalla el grafo como tal y editado cuantas veces sea necesario.

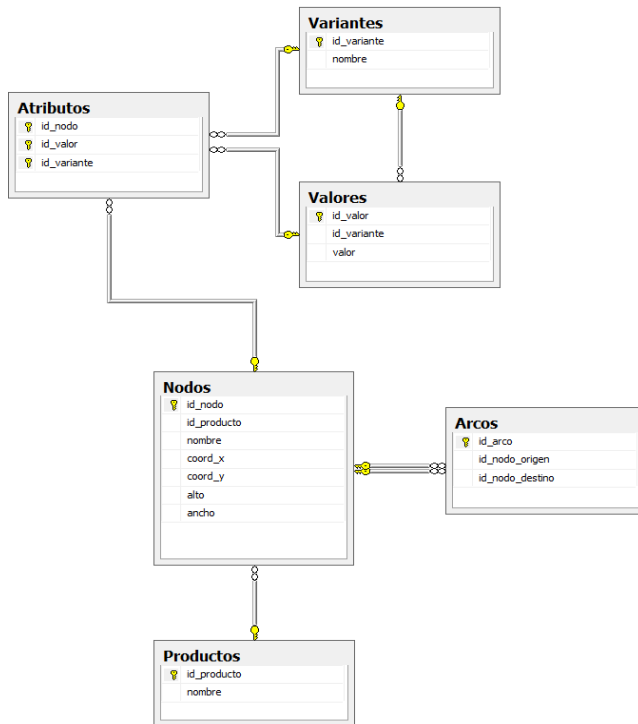


Fig.4.Modelo de datos

Donde la tabla Productos es la tabla servidora en la que se encuentra almacenada la información de cada grafo. En la tabla Arcos se tiene el control de cómo y cuáles nodos han sido conectados para cada uno de los grafos. En la tabla Nodos se asigna cada nodo conforme se van creando para cada grafo en particular. Finalmente la tabla Atributos contiene una relación con la tabla Variantes y a la tabla Valores, donde esta última especifica el valor que puede tomar una variante, esto es, porque se debe de considerar la existencia de varias familias de variantes, en donde cada una de ellas pueden tener diferentes cantidades de valores.

ActiveX AddFlow

Es un control de uso general para realizar diagramas. Su función consiste en desarrollar y desplegar datos o diagramas como de flujo de proceso, bases de datos, redes de comunicación, sistemas expertos, teorías de grafos, entre otros.[6]

Para fines de la presente aplicación una vez hecha la referencia y la inicialización del control, el objeto en el diagrama puede ser un *nodo* o un *enlace* (el *enlace* tiene el mismo significado que un arco o un link). Los diagramas se crean de forma interactiva sin alguna línea de código, además de que provee soporte para guardar, cargar, cortar, pegar e imprimir diagramas.[7]. En la Fig.5 se expone como inicializar el componente en C#.

```

AddFlow _diagrama;
DCProducto _prod;
private void InitDiagrama()
{
    _diagrama = new AddFlow();
    _diagrama.Parent = this;
    _diagrama.Dock = DockStyle.Fill;
    _diagrama.AutoScroll = false;
    _diagrama.AfterAddNode += new
    AddFlow.AfterAddNodeEventHandler(_diagrama_A
    fterAddNode);
}
  
```

Fig. 5. Inicialización del control ActiveX AddFlow

Servicio Web

El Servicio WWW, o simplemente Web, se puede definir como un amplio sistema multimedia de acceso a información heterogénea distribuida por toda la red en forma de documentos hipertextuales (hipertextos).[8] Surgió en 1990 en el CERN (Center Européen de Recherche Nucléaire)

SOA (ServiceOrientedArchitecture)

Es un estilo de Arquitectura de Software basado en la definición de servicios reutilizables, con interfaces públicas bien definidas, donde los proveedores consumidores de servicios interactúan de forma desacoplada para realizar los procesos de negocio.[9]

La documentación relacionada con esta nueva arquitectura es abundante, pero no concreta, es decir, que las metodologías para su implementación son propietarias. Empresas líderes como Software AG, TIBCO, Oracle, CBDI, Microsoft, cuentan con un framework completo con el soporte tecnológico asociado a la misma.[10]

Para el desarrollo del presente sistema se eligió WCF que pertenece al framework .NET 4.0 de Microsoft en Windows 7 para la creación de aplicaciones orientadas a servicios y aplicaciones distribuidas. Una de las principales consideraciones tomadas en cuenta para elegir WCF fue el modelo de programación unificado con el que cuenta, que es la unificación de numerosas capacidades que antes podíamos encontrar en distintas tecnologías.[11] Esto nos libera de tener que estar haciendo uso de más de una tecnología para cumplir satisfactoriamente los requerimientos

WCF (Windows CommunicationFoundation)

Es la tecnología de servicios Web de nueva generación de Microsoft, que facilitan la interconexión entre sistemas y aplicaciones dentro de la organización y a lo largo de la infraestructura geográficamente dispersas. Es el primer modelo de programación creado de principio a fin para facilitar el desarrollo de aplicaciones orientadas a servicios. WCF eleva los servicios Web a un nivel superior, al ofrecer al desarrollador un entorno de alta productividad para la creación de aplicaciones basadas en servicios Web seguros y fiables, capaces de interactuar entre plataformas.[12]

Hay varios puntos a considerar cuando se trabaja con WCF, en primer lugar se trabaja con servicios y EndPoints (puntos de acceso) que se comunican con clientes, independientemente de la plataforma. Un servicio es un programa que expone una colección de EndPoints. En la Fig. 6 se representa un servicio con una colección de EndPoints. Cada EndPoint es la puerta hacia el mundo exterior mediante una tecnología de transporte.

Un cliente es un programa que intercambia mensajes con uno o más EndPoints y un cliente puede exponer también un EndPoint para recibir mensajes de un servicio. Un EndPoint de un servicio está compuesto por una Address (dirección), un Binding (enlace) y un Contract (contrato).

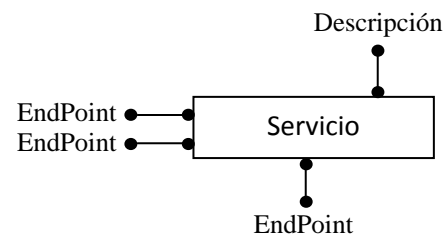


Fig. 6.EndPoints de un servicio

En la Fig. 7 se puede observar la composición de un servicio, donde cada extremo tiene un enlace asociado, un contrato de servicio y su implementación. Cada EndPoint tiene una dirección única que le distingue de los demás EndPoints para el mismo servicio, por lo cual, la unión de cada servicio también ha de ser único.

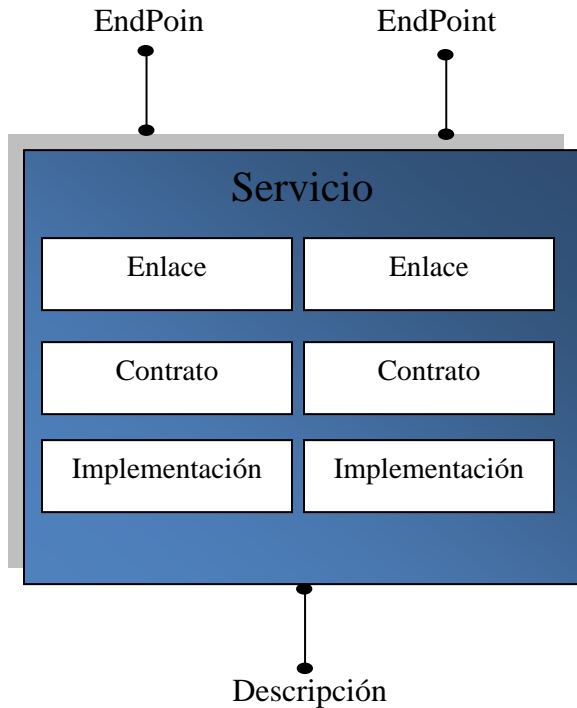


Fig. 7. Composición de un servicio WCF

Con el objetivo de emular el proceso se implementaron varios servicios, en cada uno de ellos se define un EndPoint, en el cual se debe especificar la dirección donde se encuentra el servicio, cómo se realiza la comunicación y el enlace que informa la manera en que se establece la comunicación.

Finalmente se definen las operaciones que realiza el servicio en el atributo [OperatioContract], dicho de otra manera, es para lo que fue creado el servicio. En la Fig. 8 se detalla la estructura de uno de los servicios de la aplicación, particularmente el requerido para crear la lista de materiales de cada producto. Los métodos que expone (contratos) son los CRUD (Create, Read, Update, Delete), el protocolo usado es el http definido en el Binding como WSHttBindingy lenguaje de comunicaciónhtml.

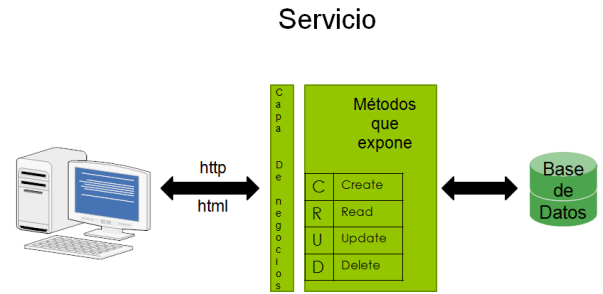


Fig.8. Estructura del servicio BOM

Continuando con el servicio definido previamente, es importante señalar como se hizo la implementación del control ActiveX AddFlow al servicio. En la Fig. 9 se especifica las capas del servicio utilizadas en la cual, se van creando objetos para cada una de ellas y así lograr la comunicación. Esto quiere decir, que a partir de que el usuario crea un grafo, se crea una entidad (arco o nodo), mediante código se transforma a objeto, y a través de los DCN [DataContract] definidos tanto en el cliente como en el servicio se logra la comunicación entre ambos, accedendo a la capa de negocios con un objeto y finalmente a la capa data Layer en donde se va almacenando de forma temporal la información de los movimientos realizados por el cliente. Una vez que el cliente da click en guardar la información contenida en la capa Data Layer se traslada a la base de datos en donde cada uno de los objetos nodo y arco tiene características propias, distintas a cualquier otro objeto.

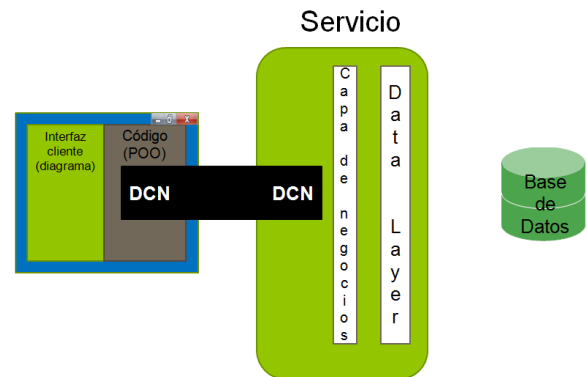


Fig. 9 Capas del servicio

Otro aspecto fundamental es el ambiente de alojamiento (host) en el cual el servicio se ejecuta. El host por excelencia para servicios bajo la plataforma de Windows es el ISS (*Internet InformationServices*). [13] El alojamiento en ISS añade beneficios, como el poder manejar puerta en marcha y la parada de un servicio de manera automática.(Solana 2009).

Los servicios de WCF pueden ser alojados en pocas versiones de IIS. Para alojar el sistema informático de la presente investigación de uso la versión IIS 7.0 en Windows 7 professional. En la Fig. 10 se observa la pantalla que demuestra que el IIS ha sido instalado correctamente en el sistema operativo.



Fig. 10. Internet InformationServices 7.0

Después de la instalación se debe proceder a hospedar el servicio en IIS, para lo cual se crea un nuevo sitio, indicando la ruta física, tipo de enlace, dirección IP y puerto. Lo anterior crea un archivo con extensión *.svc que al ejecutarlo queda hospedado el servicio. Ver Fig. 11



Fig. 11. Navegador mostrando el servicio en IIS

III. Resultados

El desarrollo de la presente investigación se llevó a cabo entre los meses de Agosto de 2010 a Junio de 2011. En primer término se trabajó con el control ActiveX AddFlow con el propósito de conocer su funcionamiento a fin de aplicarlo en el manejo y control de los grafos.

Continuando con un análisis detallado de los requerimientos mínimos de todo sistema de rastreabilidad considerando la manera en que trabajan e identificar posibles debilidades, lo anterior con el propósito de reunir la mayor cantidad de información útil y realizar un diseño de funcionalidades de un sistema genérico de rastreabilidad.

Posteriormente se realizó una búsqueda de información relacionado a SOA y WCF con la finalidad de enriquecer el marco de referencia y así conseguir el desarrollo de distintas aplicaciones de servicios y tomar la decisión de cuál es la forma que se adecua más al sistema informático. De igual manera se analizó cómo sería posible que el servicio fuera débilmente acoplado con la intención de que se cuente con un alto grado de flexibilidad, algo que en los servicios es indispensable, esto se logró en las especificaciones de los DCN, donde los clientes deben conocer la mínima cantidad información relacionada al servicio que invocan.

Con la definición previa de los requerimientos de todo sistema de rastreabilidad se desarrollaron los servicios necesarios, que permiten, en cualquier momento conocer información relacionada al proceso, producto y aspectos de calidad, así mismo mantener un monitoreo del piso de fabricación.

IV. Discusión

Durante el desarrollo de éste estudio surgieron ciertas limitaciones en cuanto a la disponibilidad de la información requerida, aunque la Arquitectura Orientada a Servicios cuenta con suficiente fundamento teórico, no es el caso para Windows Communication Foundation pues existe mucha información básica al respecto, careciendo aún de información especializada requerida para el desarrollo de aplicaciones como la mostrada en la investigación, en cuanto al desarrollo de las funcionalidades de rastreabilidad implementadas al sistema, se investigó ampliamente en sistemas comerciales con lo cual se adquirió gran cantidad de conocimientos relacionados al control de procesos, pues fue necesario involucrar bastantes consideraciones relacionados al ámbito manufacturero, a pesar éstas limitaciones identificadas se logró cumplir satisfactoriamente con el objetivo establecido.

V. Conclusiones

Es evidente que la idea fundamental de WCF es poder desarrollar aplicaciones SOA de manera simple, rápida y segura, planteando un modelo de desarrollo práctico en contra partida a otras implementaciones de SOA sobre java, como es el caso de IBM y Oracle.

Con la descripción de la aplicación presentada se pudo evidenciar que el desarrollo de aplicaciones basadas en la Arquitectura Orientada a Servicios utilizando el marco de trabajo de WCF se logra la integración de la flexibilidad con los requerimientos de negocio, produciendo sistemas débilmente acoplados que cubren perfectamente las necesidades de respuesta rápida a los cambios requeridos en cualquier momento de manera sencilla mediante la edición de los métodos del servicio afectados por dicho cambio. Logrando de ésta manera cumplir satisfactoriamente el objetivo planteado.

VI. Referencias

- [1] Pacheco, Miguel Merino. «¿Qué es la trazabilidad?» *Economía de la trazabilidad*, 2005: 12.
- [2] Escárcega, Mauro. 2009. <http://www.packaging.enfasis.com/notas/13999-presentan-tendencias-rastreabilidad-la-industria-farmaceutica> (último acceso: 28 de Noviembre de 2010).
- [3] ISO. *ISO 22000 standard procedures*. United States of America: bizmanualz, 2008.
- [4] Cuatrecasas, Luis. «Gestión competitiva de Stocks y procesos de producción.» 88. Barcelona, España: Talleres Graficos Vigor S.A., 2003.
- [5] Ramirez, Teo. *BOM (Bill of Materials - Estructura de Materiales)*. 27 de Julio de 2010. [http://jetaccess.net/groups/redbahn/wiki/ef9d5/6.2.7_BOM_\(Bill_of_Materials_-_Estructura_de_Materiales\).html](http://jetaccess.net/groups/redbahn/wiki/ef9d5/6.2.7_BOM_(Bill_of_Materials_-_Estructura_de_Materiales).html) (último acceso: 14 de Abril de 2011).
- [6] *Lassalle Technologies*. 2005. <http://lasalle.com>.
- [7] Trejo, Jaime. «Herramienta informática para la identificación de procesos asociados a la fabricación de muebles de mader.» Chihuahua, Chihuahua, Junio de 2007.
- [8] Ángel Cobo, Patricia Gómez, Daniel Perez. *PHP y MySQL Tecnologías para el desarrollo de aplicaciones Web*. España: Diaz de Santos, 2005.
- [9] Delgado, Andrea. *Desarrollo de aplicaciones con enfoque SOA*. 2006.

[10]Resnick, Steve. *Essential Windows Communication Foundation*. E.U.A.: Addison Wesley, 2008.

[11] Solana, Aroa. «Windows Communication Foundation.» 34. Madrid: Laurus, 2009.

[12]McMurtry, Craig. *Windows Communication Foundation*. USA: SAMS, 2008.

[13]Klein, S. *Windows Communication Foundation Concepts*. Indianapolis, USA: Wiley, 2007.

VII. Autores

Ing. Wendy Janett Guzmán González es ingeniera en Sistemas Computacionales, actualmente estudia la Maestría en Ciencias en Ingeniería Electrónica con especialidad en automática e informática industrial en el Instituto Tecnológico de Chihuahua.

M.S.M. Cynthia Liliana Guzmán González obtuvo su título de Maestro en Sistemas de Manufactura con especialidad en automática e informática industrial en el Instituto Tecnológico de Chihuahua.

M.S.M Juan Alberto Luna Sandoval obtuvo su título de Maestro en Sistemas de Manufactura con especialidad en diseño de productos y procesos en el Instituto Tecnológico de Chihuahua.

Buscador Web Móvil Semántico

***Miguel Félix Mata Rivera, **Raúl Otilio Hernández Carro, ***Sergio Alexis Zepeda Rosas**

*, **, ***Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas, Instituto Politécnico Nacional.

*, **, ***Avenida Instituto Politécnico Nacional No. 2580, Colonia Barrio la Laguna Ticomán, Delegación Gustavo A. Madero, CP. 07340 México D.F. *mmatar@ipn.mx, **rohernandezc88@gmail.com, ***szepedar0600@ipn.mx

Resumen

En la actualidad es común realizar búsquedas de productos y servicios en la Web para explorarlos, evaluarlos e incluso hacer compras o contratar algún servicio como el envío de paquetes de discos. Sin embargo, los buscadores actuales sólo realizan análisis sintáctico pasando por alto el significado de los términos y su contexto. Generar una herramienta que realice la función de filtrado y recuperación de resultados por su semántica, requiere utilizar diversas estrategias y algoritmos, de los cuales se utilizaron los referidos a la exploración de ontologías y las métricas de ponderación típicas en el área de recuperación de información con la finalidad de ofrecer al usuario una nueva alternativa de consulta, mediante la cual los resultados obtenidos posean un alto grado de relación con lo que realmente busca y de manera adicional, se muestren resultados que si bien no son concretamente el objeto de consulta puedan resultar interesantes dada la relación semántica que guarden entre ellos.

Por lo tanto se espera que los documentos recuperados sean de mayor relevancia, comparados con los que ofrece un buscador basado en *keywords* como lo es Google. Esto se indicará con las métricas de *recall* y *precision* y con tablas estadísticas.

Palabras clave: Búsqueda semántica, ontología, representación gráfica, recuperación de información.

I. Introducción

La búsqueda actual de información se realiza utilizando motores de búsqueda que recorren la Web, analizando páginas o documentos que contengan la palabra o término de la consulta y los ordena según su relevancia léxica o sintáctica, excluyendo posibles resultados que pueden ser relevantes para el usuario debido a que carecen de un procesamiento semántico que ofrezca resultados según el significado de los datos encontrados. Este último proceso se realiza normalmente utilizando una ontología.

La idea de la Web semántica es que los datos puedan ser utilizados y “comprendidos” por las computadoras sin necesidad de supervisión humana, de forma que los navegadores Web puedan ser diseñados para tratar la información

situada en las páginas Web de manera semiautomática.

Para poder explotar la Web semántica, se necesitan lenguajes de marcado apropiados que representen el conocimiento de las ontologías. La Web Semántica proporcionará un salto cualitativo sobre el potencial actual de la Web.

Las principales ventajas de esta innovación en Internet serán el desarrollo de aplicaciones con esquemas de datos comunes, fomento de las transacciones entre empresas por comercio electrónico y búsqueda de información con inferencias. Para poder lograr estos objetivos se necesita unificar los contenidos semánticos por medio de ontologías que formalicen este conocimiento de forma consensuada y reutilizable.

Es por ello que enriquecer la consulta con palabras clave y realizar una exploración basados en el contexto donde se ubican los términos de la consulta empleada permite que los resultados devueltos sean aquellos que realmente tengan un grado de relevancia mayor para el usuario comparado con los que entrega un buscador convencional.

II. Estado del arte

A lo largo de los últimos años diversos trabajos de investigación y proyectos se han dedicado a explorar las ventajas que posee la implementación de la Web semántica, así algunos de los trabajos relacionados con el sistema desarrollado como el Shopping Maps[1] que es un *mashup* semántico para modelos de datos heterogéneos en comercio electrónico. Se trata de un proyecto que realiza un *mashup* (aplicación Web híbrida) la cual entra a los sitios de eBay y Amazon para buscar productos de acuerdo con diversos criterios, además muestra en un mapa la información del producto, la ubicación de éste y otras ubicaciones donde es posible comprar y opiniones de otros clientes. Otro ejemplo de los intentos de implementación de algoritmos semánticos es Ubica-Negocio [2] apoyándose en la arquitectura de [8] y [9], básicamente es un asistente para establecer geográficamente un negocio por su semántica espacial y contexto de mercado. En el cual sugiere las ubicaciones más adecuadas para un negocio con base en su semántica espacial y el contexto de mercado a partir de los requerimientos del usuario. En un intento por conseguir darle un sentido de comprensión a la exploración de la Web NaturalFinder [3] es un complemento al integrarse con alguno de los buscadores actuales, es capaz de comprender el lenguaje natural. Swotti [4] es un buscador que rastrea las opiniones sobre productos, apoyándose en la Web semántica. Mientras que en la Universidad de San Carlos en Guatemala se desarrolló OntoGuate. [5] proyecto cuyo objetivo es la administración de una ontología de turismo y sus respectivas búsquedas. Una de las aproximaciones más recientes en el desarrollo de

la exploración de documentos en la web es el Buscador Web semántico de documentos [6] utilizando la teoría de [10] y [11], el trabajo es un sistema basado en semántica, el cual utiliza una ontología que define los conceptos y las relaciones del dominio afín. Finalmente el Sistema Multi-Agentes con un mapa auto-organizado. En [7] se utiliza la arquitectura de [12] y [13] con la teoría de [14] se desarrolló un sistema que realiza búsquedas procesando el significado de las palabras que conforman la consulta, a las cuales se le agrega una categoría para que esta búsqueda sea más congruente con el término de búsqueda original.

III. Metodología

La metodología considera un sistema capaz de gestionar consultas en el dominio de productos de música y películas por su semántica, esto para contextualizar el significado de la consulta, desplegando resultados que cumplan con las necesidades del usuario. Así, mediante el desarrollo de una aplicación Web y una aplicación en versión móvil, el usuario tendrá la posibilidad de realizar consultas a través de la Web o un teléfono celular con conectividad a Internet. El buscador semántico, conducido por ontologías procesa consultas por contexto y significado, con el fin de ofrecer al usuario resultados basados en significado y no únicamente por coincidencia de palabra, además de ponderar los resultados para su despliegue final.

El sistema hace uso de herramientas que los motores de búsqueda actuales no suelen utilizar, como es el procesamiento semántico y la ponderación de la información, partiendo del hecho de que la semántica nos permite procesar el significado en el contexto de la consulta y la ontología es el soporte para tomar una decisión basada en la experiencia de un dominio de conocimiento en particular.

Por lo tanto, el desarrollo de este sistema se enfocó en realizar una herramienta de búsqueda que recupere y pondere información por su semántica y contexto.

La consulta que ingrese el usuario tiene ciertas restricciones en cuanto a la longitud de los términos de búsqueda, tal que cumplan:

- No tener más de cinco palabras.
- Términos de no más de dos palabras.

La primera restricción se debe a que el procesamiento de exploración esta en función directa con el número de términos, lo que aumentaría el tiempo de exploración y de respuesta en el sistema.

La segunda restricción se considera para disminuir los procesos de desambiguación. La arquitectura general del sistema desarrollado se muestra en la Fig. 1.

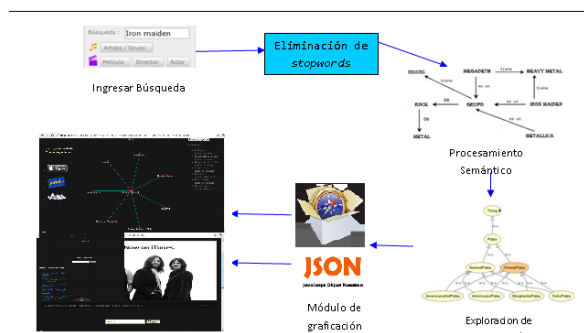


Fig. 1. Arquitectura del sistema

Como se observa en la Fig. 1, el sistema, recibe la consulta, y se procede a realizar el proceso de análisis sintáctico-semántico y la exploración de la ontología, para finalmente graficar los resultados. Cabe señalar que el módulo de graficación no es un mapeo de la ontología, sino que es un despliegue del conjunto de resultados con sus relaciones con otros conceptos o productos.

Asimismo se realizó una recopilación de 731 páginas Web relacionadas con los temas del dominio que se pretende abarcar en el mismo (un total de 149 artistas de música y películas), con la finalidad de almacenarlas y generar un *corpus* del cual se extraigan los resultados de las consultas solicitadas por el usuario. Se usó este método, puesto que cimentar nuestras consultas semánticas en los resultados de algún explorador

Web en tiempo real implicaba una mayor complejidad en la elaboración del sistema. Dicho *corpus* está almacenado en un directorio del servidor donde se ubica la aplicación Web con la finalidad de que pueda ser referenciado desde el mismo sistema. Para la selección de las páginas que guarden relación con la consulta ingresada por el usuario se diseñó una tabla en una base de datos que contiene la ruta de cada elemento existente en el *corpus*, así como los campos que nos permiten explorar parte del contenido de las páginas Web para poder decidir si éstas tienen relación con los criterios de la consulta enriquecida resultante de la exploración en la ontología. A continuación se mencionan las funciones que pueden realizarse tanto en la aplicación Web como en la aplicación móvil, describiendo a su vez las características de desempeño para cada aplicación.

IV. Funcionamiento de Aplicación Web

Es la encargada de fungir como interfaz entre el sistema y el usuario, ofreciendo las funciones de ingreso de términos de búsqueda, selección de resultados con base en parámetros determinados por la relaciones semánticas de la ontología. Así como el despliegue gráfico de los resultados obtenidos.

Este proceso de representación gráfica no es un mapeo o importación de la ontología, sino el procesamiento de los términos de la consulta, asociados a conceptos y relaciones semánticas. Adicionalmente, se cuenta con la funcionalidad de retroalimentación del sistema, la cual permitirá al usuario definir si los resultados entregados por el sistema en su primera interpretación son los que se esperaban. De no ser así, se retroalimentará la consulta con algún criterio semántico definido en la ontología, con el objeto de hacer aún más específico y útil el tratamiento de la información.

V. Funcionamiento de Aplicación móvil

La aplicación móvil del sistema es un intérprete de los resultados obtenidos como respuesta de

una consulta al sistema Web, desde un teléfono celular con conectividad a internet, la cual estará diseñada para detectar qué tipo de dispositivo ha entrado a la aplicación para ajustar la resolución de la pantalla, y realizar la consulta semántica, los metadatos entregados por el explorador a la aplicación móvil serán interpretados y graficados de tal forma que se desplieguen con el mismo formato que en la aplicación Web.

Con el desarrollo de este buscador semántico se persigue que la búsqueda realizada por el usuario sea más precisa y origine resultados relevantes con su mínima intervención.

La estructura de los géneros introducidos en la ontología se muestra en la Fig. 2.

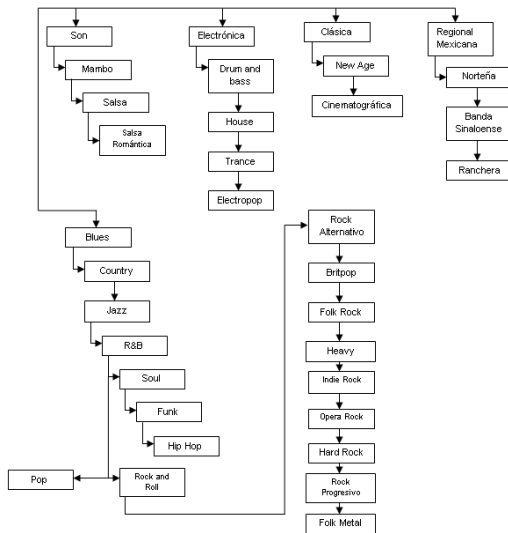


Fig. 2. Estructura conceptual de la ontología

La ontología se implementó en OWL (<http://www.w3.org/TR/owl-features/>) y se explora usando el lenguaje SPARQL (<http://www.w3.org/TR/rdf-sparql-query/>) y como se aprecia en la Fig. 2. A través de las relaciones semánticas existentes entre los conceptos es posible encontrar cuáles conceptos tienen relación con algún género de música así como su ubicación en cuanto a sus relaciones semánticas.

VI. Pruebas y resultados

Las pruebas se realizaron utilizando consultas en formato de texto libre. Por ejemplo, para la

consulta: “compay segundo” podemos apreciar en la Fig. 3 se muestran los resultados obtenidos para dicha consulta.

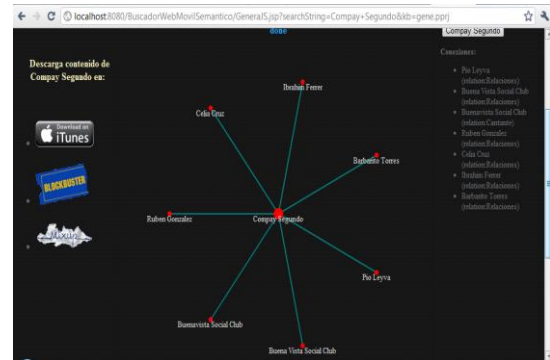


Fig. 3. Resultados de artistas relacionados con “compay segundo”

Como se observa en la Fig. 3 se muestran la graficación de los resultados para la búsqueda realizada, con la opción para retroalimentar, incluyendo enlaces a páginas que venden productos relacionados y las páginas relacionadas con la consulta.

Además como se aprecia en la fig. 3, se pueden mostrar apartados para describir su funcionamiento de una manera más óptima, dichas zonas corresponden a un módulo específico del sistema y que provee de una funcionalidad específica al usuario.

Estas zonas pueden dividirse en: zona de graficación, zona de enlaces externos, zona de retroalimentación y zona de páginas relacionadas con el artista. Ésta última zona se muestra en la Fig. 4, ya que es parte medular del sistema al ser el apartado donde se ven reflejados los enlaces que entrega al servidor después de haber sido procesada la consulta enriquecida traída desde la ontología.

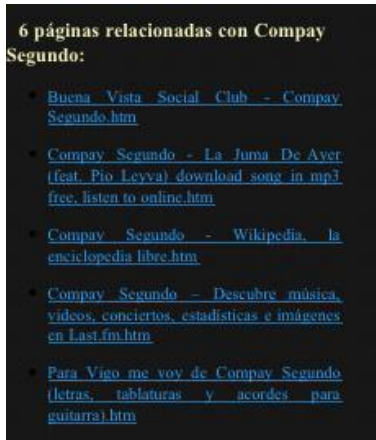


Fig. 4. Zona de enlaces a páginas relacionadas con el término de consulta

La interfaz gráfica en el móvil es prácticamente la misma que la pantalla principal de la interfaz Web, solo que ésta se ajusta a las dimensiones del *iPhone* (ver Fig. 5).

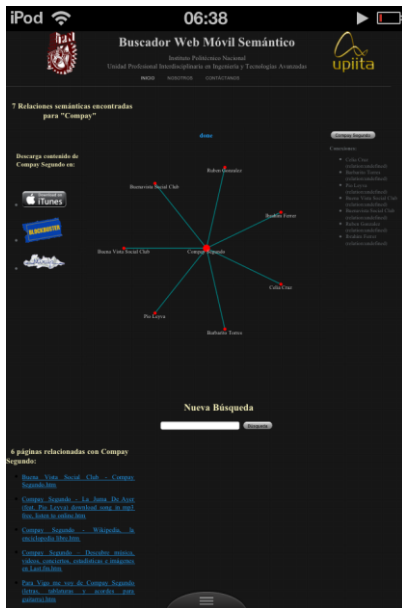


Fig. 5. Interfaz móvil del sistema

Para validar la funcionalidad del sistema en cuanto a la relevancia de los resultados entregados por el sistema a petición del usuario se realizaron pruebas obteniendo las medidas de *recall* y *precision*.

El *recall* es la probabilidad de que un documento relevante es recuperado en una búsqueda. La forma de calcularlo se muestra en la figura 6.

$$recall = \frac{|{\text{documentos relevantes}} \cap {\text{documentos recuperados}}|}{|{\text{documentos recuperados}}|} \quad (1)$$

Los resultados obtenidos se muestran en la tabla 1:

Artista	Recall
The Beatles	1
Compay Segundo	0.88
Iron Maiden	0.83
Queen	0.58
Mago de Oz	0.88
Guns N' Roses	1
Ibrahim Ferrer	1
LEd Zeppelin	0.75
Txus di Fellatio	0.6
Keane	0.6

Tabla 1. Valores de *recall* para diversos artistas

Presicion es la probabilidad que un documento (aleatoriamente seleccionado) recuperado sea relevante. La forma de calcularlo se muestra en la figura 6:

$$precision = \frac{|{\text{documentos relevantes}} \cap {\text{documentos recuperados}}|}{|{\text{documentos recuperados}}|} \quad (2)$$

Los resultados obtenidos se muestran en la tabla 2:

Artista	Precision
The Beatles	0.66
Compay Segundo	1
Iron Maiden	1
Queen	1
Mago de Oz	1
Guns N' Roses	0.66
Ibrahim Ferrer	0.66
LEd Zeppelin	1
Txus di Fellatio	1
Keane	1

Tabla 2. Valores de *precision* para diversos artistas

Una buena recuperación () significa que no se ha perdido nada, pero se puede tener una gran cantidad de resultados inútiles o irrelevantes (lo que implicaría la baja precisión).

Una alta precisión () significa que todo lo que obtuvo era un resultado relevante, pero no podría haber encontrado todos los elementos relevantes (lo que implicaría recuperación baja).

VII. Conclusiones

El sistema provee a los usuarios de una interfaz para realizar consultas sobre música y películas basada pro su significado, las principales ventajas y características, se enuncian a continuación:

El usuario posee la capacidad de seleccionar entre la entidad específica de su búsqueda ó alguna que esté directamente relacionada bajo los criterios de la ontología realizada.

Se despliega una interfaz atractiva con efectos visuales de animación provistos por el API *JIT* (Javascript InfoVis Toolkit).

La versión móvil permite a dispositivos que soporten el explorador *Safari* una interfaz exclusiva en cuestiones de diseño y funcionalidad, con lo cual vuelve al sistema portable y novedoso.

La construcción de la ontología, se baso en información de comercios establecidos y reconocidos a nivel mundial en el dominio de la música. La ontología es local dado que no deja de ser una agrupación categórica de los desarrolladores del sistema, por lo que se requiere su integracion con otros ontologias.

La ontología tiene un número reducido de conceptos, el cual debe extenderse de acuerdo a la metodologia de construccion.

La exploración al corpus está restringida a buscar de manera indexada en solo algunas etiquetas relevantes que la mayoría de las páginas web poseen para su fácil descripción y búsqueda por buscadores convencionales, sin embargo aun existe la posibilidad de robustecer la exploración a la páginas Web con la consulta enriquecida si se realizara una búsqueda aún más exhaustiva dentro del documento HTML.

Así se propone entonces el siguiente trabajo a futuro:

- Robustecer la cantidad de información recopilada tanto en la ontología como en el *corpus*, con la finalidad de hacer más profundo el análisis y por lo tanto, más específico.
- Realizar una exploración más compleja dentro de los documentos analizados para entregarlos como respuesta al usuario, así la precisión en la recuperación de resultados relevantes para el usuario se incrementa dado que el análisis se vuelve exhaustivo.
- Enriquecer de forma más compleja y con mayor sentido relacional las consultas, con términos que provean de un conocimiento mayor y un razonamiento

relacional basado en convenciones establecidas para lograr hacerlo universal.

Agradecimientos

Los autores de este artículo agradecen a la SIP-IPN, COFAA, EDD, y a la UPIITA por su apoyo .

VIII. Referencias

- [1] García, L. A. Shopping-MAPS: sistema de comercio electrónico basado en Mashup's con mapas geográficos integrados. Trabajo terminal en UPIITA-IPN 2010.
- [2] López, S. Méndez, R., Ubica-Negocio: asistente para establecer geográficamente un negocio por su semántica espacial y contexto de mercado. Trabajo terminal en UPIITA-IPN, 2010.
- [3] Buscador NaturalFinder <http://demos.bitext.com/DARPA/>, ultimo acceso: febrero de 2010.
- [4] buscador de opiniones Swotti , <http://www.swotti.com>, ultimo acceso: febrero de 2010.
- [5] Aldana, I., Chutan, A. L., Flores, H., Robles, E. OntoGuate, Universidad de San Carlos, 2008.
- [6] García, E., Buscador Web semántico de documentos. Trabajo terminal en UPIITA-IPN, 2010.
- [7] Ortíz, E. Sistema Multi-Agentes con un mapa auto-organizado. Trabajo terminal en UPIITA-IPN, 2010.
- [8] Horridge, M., Drummond, N., Jupp S., Moulton G., y Stevens R., *A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools Edition 1.2*. The University of Manchester, 2009.
- [9] Varelas, G., Voutsakis E., Raftopoulou P., Petrakis, G.M., Milios, E.E.

Semantic Similarity Methods in WordNet and their Application to Information Retrieval on the Web. Dept. of electronic and Comp. Engineering, Technical University of Crete, Faculty of Computing Science, Dalhousie University, 2005.

- [10] Zhang, L., Yu, Y., Zhou, J., Lin, C., Yang, Y. *An Enhanced Model for Searching in Semantic Portals*. APEX Data and Knowledge Management Lab, Dept. of Computer Science and Engineering Shanghai JiaoTong University, Dept. of Computer Science, HongKong University of Science and Technology, 2005.
- [11] Guha R., McCool, R., Miller, E. *Semantic Search*. IBM Research, Almaden, Knowledge System Lab, Standford, WEC/MIT. 72.
- [12] Koo, S.O., Lim, S.Y., Lee, S. J. *Building an Ontology based on Hub Words for Information Retrieval*. Language and Information Processing Laboratory, Kyungpook National University, 2003.
- [13] Pinheiro, W. A., Moura, A. M. *An Ontology Based-Approach for Semantic Search in Portals*. Military Institute of Engineering, 2004.
- [14] Raskin, R. *Semantic Web for Earth and Environmental Terminology (SWEET)*. Jet Propulsion Laboratory, 2003.

IX. Autores

Dr. Miguel Felix Mata Rivera, obtuvo su grado de doctor en ciencias de la computacion en el CIC-IPN, actualmente dirige el laboratorio de computo movil en la UPIITA-IPN. Sus areas de interes son los SIG, la semantica espacial y el computo ubicuo y movil.

Ing. ³Sergio Alexis Zepeda Rosas es Ingeniero en telemática egresado en junio de 2001 de la UPIITA- IPN

Ing. Raúl Otilio Hernández Carro es Ingeniero en telemática egresado en junio de 2011 de la UPIITA- IPN

Software para el Cálculo de los Parámetros de Resistencia del Suelo

* José Fredi Robles Vázquez, **Rubén Herrera Galicia, ***Héctor Hernández de León

Instituto Tecnológico de Tuxtla Gutiérrez, Carretera panamericana Km 1080, 6150380

*esmat@live.com.mx. ** h_galicia24@hotmail.com *** hectorhdeleon@gmail.com.

Resumen

En el presente trabajo se desarrolló un programa de computadora que resuelve un sistema de ecuaciones denominado relación de Mohr-Coulomb, que calcula los parámetros de resistencia del suelo. Para lograrlo se trazan tres círculos a partir de los esfuerzos principales (σ_1 y σ_3) aplicados a una muestra de suelo durante tres ensayos Triaxiales y se calcula la envolvente de falla de los círculos. La recta no siempre es tangente a los tres círculos por lo que es necesario definir cuál es la tangente de dos de los tres círculos que entregará los valores deseados, para ello el programa calcula el esfuerzo normal σ y el esfuerzo tangente τ y posteriormente se ajustan con el método de regresión lineal que dan como resultado a los diagramas pq. Para garantizar que ángulo de fricción interna y la cohesión del suelo son correctos se realiza un segundo método para el cálculo de la envolvente, en el que se obtiene el promedio de los tres ángulos y las tres cohesiones de cada ensayo. El software descrito demuestra sus cualidades de efectividad, rapidez de cálculos y facilidad de uso, frente a los programas comerciales que debido a su complejidad y costos elevados no son utilizados.

Palabras Clave: Cámara Triaxial, Circulo de Mohr, Envolvente de falla, Mecánica de Suelos.

I. Introducción

Morh presentó una teoría sobre la ruptura de los materiales. Esta teoría afirma que un material falla debido a una combinación crítica de esfuerzo normal y de esfuerzo cortante, y no solo por la presencia de un esfuerzo máximo normal o bien de un esfuerzo máximo cortante, así entonces, la relación funcional entre un esfuerzo normal y un esfuerzo cortante sobre un plano de falla, la definió como se presenta en la ecuación 1 [1, 2, 3, 7].

$$\tau_f = f(\sigma) \quad (1)$$

donde τ_f = esfuerzo cortante sobre el plano de falla.

σ = esfuerzo normal sobre el plano de falla

Y se definen como:

τ_f = fuerza cortante resistente/área de sección transversal.

σ = fuerza normal/área de la sección transversal del espécimen.

Para la mayoría de los problemas de mecánica de suelos, es suficiente aproximar el esfuerzo cortante sobre el plano de falla como una función lineal del esfuerzo normal, la cual se escribe como se presenta en la ecuación 2.

$$\tau_f = c + \sigma \tan \phi \quad (2)$$

donde c = cohesión

ϕ = ángulo de fricción interna

La gráfica del círculo de Mohr permite obtener el ángulo de fricción interna, y la cohesión del suelo a partir de su análisis geométrico como se aprecia en la Fig. 1, para poder determinar estas incógnitas es necesario extender la línea de envolvente de falla al lado negativo del eje de las equis (x), a partir de aquí se deduce la ecuación (3) y (4). Una vez definidas se obtienen de la ecuación (5 a la 9).

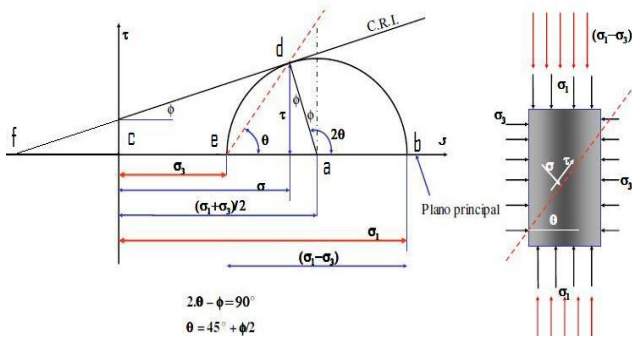


Fig. 1. Círculo de Mohr y envolvente de falla.

$$\overline{fa} = fO + Oa = c \cot \phi \frac{\sigma_1 + \sigma_3}{2} \quad (3)$$

$$\overline{ad} = \frac{\sigma_1 - \sigma_3}{2} \quad (4)$$

$$\sin \phi = \frac{\frac{\sigma_1 - \sigma_3}{2}}{c \cot \phi \frac{\sigma_1 + \sigma_3}{2}} \quad (5)$$

$$\sigma_1 = \sigma_3 \frac{(1 + \sin \phi)}{(1 - \sin \phi)} + 2c \frac{(\cos \phi)}{(1 - \sin \phi)} \quad (6)$$

$$\frac{(1 + \sin \phi)}{(1 - \sin \phi)} = \tan^2 \phi \left(45 + \frac{\phi}{2} \right) \quad (7)$$

$$\frac{\cos \phi}{(1 - \sin \phi)} = \tan \phi \left(45 + \frac{\phi}{2} \right) \quad (8)$$

$$\sigma_1 = \sigma_3 \tan^2 \phi \left(45 + \frac{\phi}{2} \right) + 2c \tan \phi \left(45 + \frac{\phi}{2} \right) \quad (9)$$

La ecuación (9), representa el criterio de falla de Mohr-Coulomb expresada en términos de los esfuerzos de falla.

El método de solución anexo a este trabajo es el de los Diagramas “pq”, que se obtienen a partir de los esfuerzos aplicados a la muestra de suelo, quedan definidos a partir del cálculo del esfuerzo de compresión denotado por la letra p (10), y el correspondiente al esfuerzo de corte máximo denotado por la letra q (11), creando una trayectoria de esfuerzos [9]. Para convertir a esta

trayectoria en una línea recta promedio, se utiliza el método de regresión lineal [4].

$$p = (\sigma_1 + \sigma_3)/2 \quad (10)$$

$$q = (\sigma_1 - \sigma_3)/2 \quad (11)$$

En fechas recientes el estudio más aproximado al que se aborda en este trabajo es el de diseño de un aparato biaxial [5], el que establece que es de mucha importancia, toda vez que, a través de él se puede observar el comportamiento del suelo y determinar las propiedades del mismo, cuando es sometido a un estado de deformación plana, ya que la mayoría de los problemas en ingeniería geotécnica se solucionan bajo este supuesto.

Derivado de la falta de información en el tema ha sido necesario desarrollar un programa que resuelve el problema de la obtención del ángulo de fricción interna y la cohesión de los suelos. Por lo que se propone la solución a la ecuación 9, mediante programación en el software Visual Basic de Microsoft. Que debe entregar resultados numéricos y gráficos.

La problemática que en este artículo se aborda está dirigida a los ensayos Triaxiales que son utilizados en la ingeniería civil para determinar los parámetros de resistencia de los suelos. Permiten predecir futuros asentamientos en las edificaciones y de esta manera establecer medidas de reforzamiento o mejoramiento del material. Existen tres tipos de ensayos Triaxiales, Consolidado no Drenado (CU), Consolidado Drenado (CD), y no Consolidado no Drenado (UU). El tipo de ensayo que se utiliza en el presente trabajo es el Consolidado no Drenado y la solución a este ensayo se presenta mediante los círculos de Mohr. El esfuerzo o tensión se define como una fuerza por unidad de área, con unidades en psi o MPa. Para conocer el comportamiento de un cuerpo sometido a esfuerzos se modela mediante un cubo, en la que las tensiones son aplicadas en las caras de dos formas. Los esfuerzos normales actúan de manera perpendicular (normal) a una de las caras y tiende a la tracción o compresión, donde la tensión cortante actúa tangente al plano (Se suele denotar por la letra griega *tau* τ), como se muestra en la

Fig. 2. Todo suelo sometido a esfuerzos de compresión tomará el comportamiento descrito.

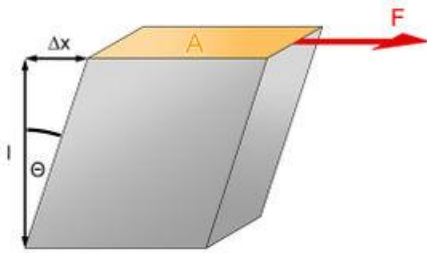


Fig. 2. Deformación del cubo sometido a un esfuerzo cortante.

II. Desarrollo

Se calculan los círculos de Mohr, para una prueba Triaxial, con los datos de entrada que son: El esfuerzo desviador máximo σ_1 y la presión de confinamiento aplicada a la cámara Triaxial σ_3 , en cada prueba [1, 2, 3, 6, 7, 8, 9], estos datos se obtienen a través de un sensor de presión para el que se utilizó la celda de carga de la compañía Logicbus con una capacidad de 3000 lbs. y un LVDT de la compañía Vishay con un corrimiento de 5 cm.

Para ello se resuelve la ecuación 1, que se obtiene a partir de los círculos de Mohr.

Existe un método de solución en las pruebas Triaxiales denominado diagramas “pq”, el cual consiste en obtener todos los puntos en los que se tocan la tangente a cada círculo con la envolvente de falla.

Uniéndolos todos los puntos generados para el diagrama “pq” se traza la línea que genera la envolvente de falla de los círculos de Mohr, y al ser una línea recta se puede calcular el ángulo de inclinación y todos sus componentes.

Si se aprecia para obtener el ángulo de fricción interna θ se necesitan dos ecuaciones con dos incógnitas, no obstante son tres círculos los que se dibujan, por lo tanto uno debe ser descartado para su análisis, ya que la recta se puede formar con solo dos puntos uno en cada círculo, usualmente el ingeniero experto en tensiones

toma como dato el que según su experiencia sea mejor, la diferencia con la solución propuesta es de que se utiliza el diagrama “pq”, se comparan los resultados con cada prueba de cada par de círculos generados y el que más se acerque al comparado se toma como resultado, ya que los círculos de Mohr son hasta la fecha una de las mejores representaciones de estado de corte de los materiales, como se podrá apreciar en los resultados obtenidos.

El Método de promedio se obtiene, al realizar un promedio de los tres ángulos de fricción interna calculados y de los tres valores de cohesión encontrados.

La solución en ambos métodos de inicio es la misma, primero se calcula el valor de los parámetros y posteriormente se aplican las soluciones antes mencionadas.

III. Resultados

Los Resultados se obtuvieron mediante el análisis del software propuesto de nombre **Analizador de suelos 2.0**.

El cual contiene dos métodos de solución:

El método por comparación con diagrama “pq” y el método por promedio de los datos de ángulo de fricción interna y cohesión. Ambos presentaron los resultados buscados y se muestran en las tablas (1, 2).

DATOS			
Ensayo	Presión lateral kg/cm^2	Esfuerzo Desviador kg/cm^2	
1	1	1.788	
2	2	2.848	
3	3	3.993	
RESULTADOS			
Ángulo de Fricción Interna	Cohesión kg/cm^2	Esfuerzo Normal kg/cm^2	Esfuerzo Tangencial kg/cm^2
20.2676°	0.2536	1.5843	0.8386
20.2676°	0.2536	2.9307	1.3358
20.2676°	0.2536	4.3049	1.8433

Tabla 1. Comparando con diagrama “pq”.

DATOS			
Ensayo	Presión Lateral kg/cm ²	Esfuerzo Desviador kg/cm ²	
1	1	1.788	
2	2	2.848	
3	3	3.993	
RESULTADOS			
Ángulo de Fricción Interna	Cohesión	Esfuerzo normal	Esfuerzo tangencial
20.8111°	0.2268	1.5764	0.826
20.8111°	0.2268	2.9181	1.3359
20.8111°	0.2268	4.2872	1.8563

Tabla 2. Resultados Promediando los datos.

Los resultados graficos se aprecian en las figs. (3, 4).

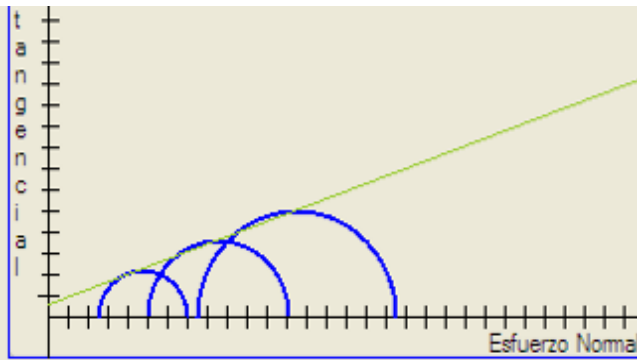


Fig. 3. Círculos de Mohr obtenidos con los resultados de la tabla 1.

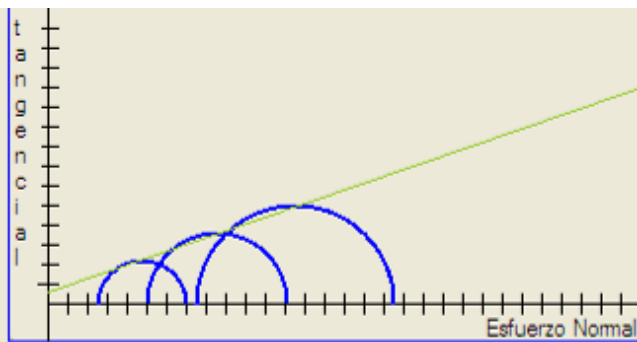
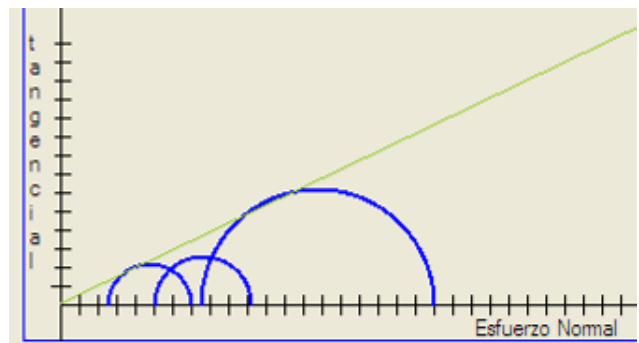
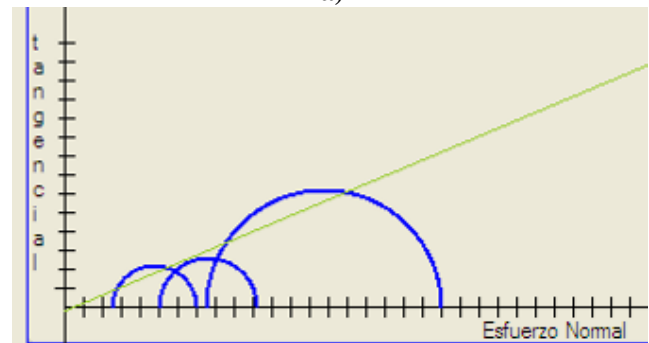


Fig. 4. Círculos de Mohr obtenidos con los resultados de la tabla 2.

Se puede Observar que ambos resultados son satisfactorios. Aun que para comprender las diferencias se propone al segundo ensayo de tal manera que se aprecie un círculo mas pequeño lo que representaría un problema en la muestra de suelo. Como se puede apreciar en la fig. 5. En este ejemplo ilustrativo se demuestra el funcionamiento del programa utilizando ambos métodos. Para el caso B) de la fig. 5 es muy evidente que a pesar de que se promedian los datos el valor del ángulo de fricción interna se mantiene.



a)



b)

Fig. 5. Ejemplo ilustrativo. a) Comparando con diagrama “pq”. b) Promediando los datos.

La interfaz gráfica del programa propuesto se puede apreciar en la Fig. 6, en la que es evidente que su uso no resulta complicado, toda vez que fue diseñada para la aplicación de un ensayo Triaxial.

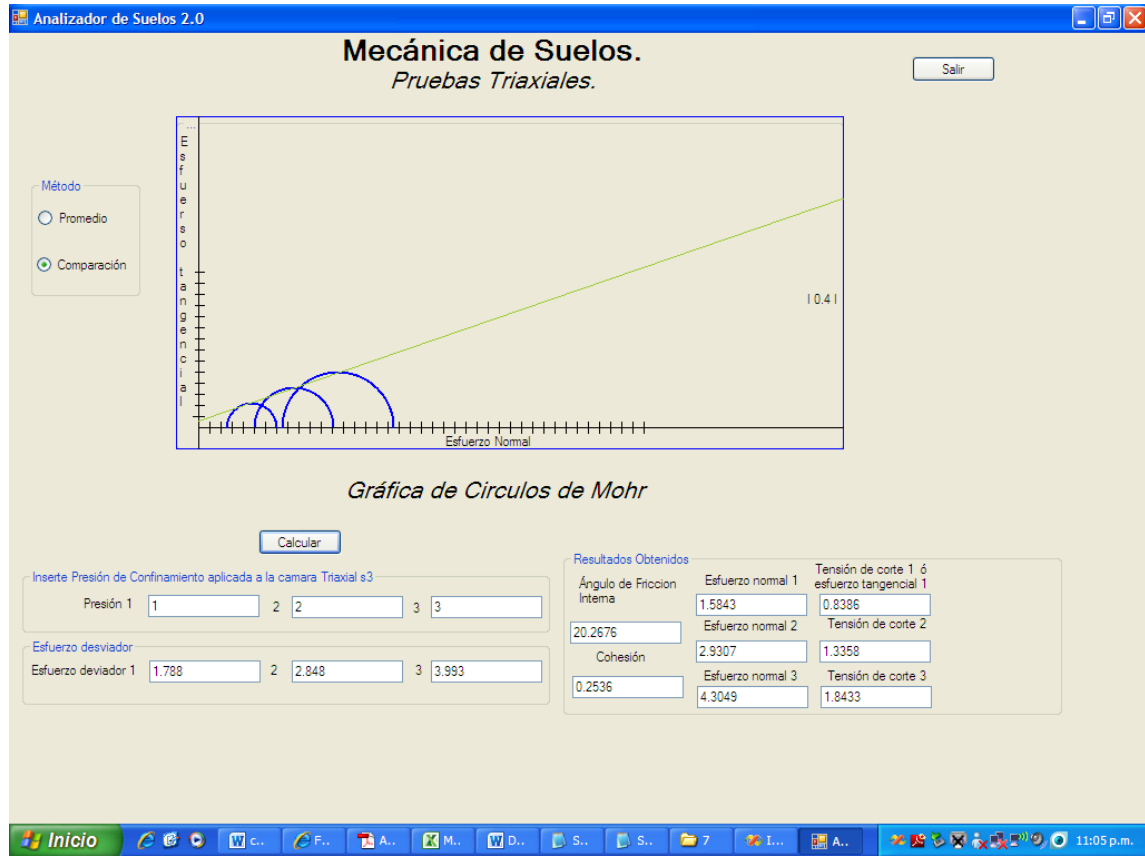


Fig. 6. Pantalla del programa Analizador de suelos 2.0.

IV. Discusión

Para poder concluir que el método que se desarrollo a lo largo de esta investigación es el correcto se debe comparar contra uno que se utilice actualmente, por lo que se realizaron pruebas y se compararon con los del laboratorio de mecánica de suelos de la Universidad Autónoma de Chiapas (UNACH), de las cuales surgieron los resultados mostrados en las tablas (3, 4, 5).

Para las siguientes pruebas se le llamará “Método 1”, al promedio de los ángulos y cohesión y como “Método 2”, al de comparación con los diagramas “pq”, y los resultados obtenidos por el laboratorio de mecánica de suelos de la Universidad Autónoma de Chiapas como “UNACH”.

Los resultados obtenidos al comparar con los que obtuvieron en laboratorio en pruebas reales demuestran que los métodos de solución aquí propuestos obtienen valores reales y correctos, y

que se pueden aplicar como instrumento de cálculo en los laboratorios de Mecánica de suelos, como herramienta de la Ingeniería civil o en el empleo de uso didáctico para la cátedra de geotecnia o suelos.

DATOS			
Ensayo	Presión lateral kg/cm ²	Esfuerzo Desviador kg/cm ²	
1	0.5	0.7	
2	1	1.13	
3	1.5	1.54	
RESULTADOS			
	Método 1	Método 2	UNACH
Ángulo	17.2028	16.9047	17
Cohesión	0.1057	0.1149	0.09

Tabla 3. Prueba 1.

DATOS			
Ensayo	Presión lateral kg/cm ²	Esfuerzo Desviador kg/cm ²	
1	0.5	1.03	
2	1	1.52	
3	1.5	1.96	
RESULTADOS			
	Método 1	Método 2	UNACH
Ángulo	18,4991	17.7916	18
Cohesión	0.2095	0.2334	0.22

Tabla 4. Prueba 2.

DATOS			
Ensayo	Presión lateral kg/cm ²	Esfuerzo Desviador kg/cm ²	
1	0.5	0.93	
2	1	1.13	
3	1.5	1.4	
RESULTADOS			
	Método 1	Método 2	UNACH
Ángulo	10.946	10.9693	11
Cohesión	0.2776	0.2866	0.3

Tabla 5. Prueba 3.

V. Conclusiones

El Software propuesto funciona apropiadamente para el cálculo de los parámetros de resistencia del suelo, sin embargo se puede incluir a la ecuación de relación Morh-Coulomb el parámetro de presión de poros para el cálculo de los parámetros efectivos.

Los programas comerciales no son amigables a los usuarios, la falta de uso de los mismos demuestra que además de costosos no tienen una interfaz familiar.

En los laboratorios de mecánica de suelos, se prefiere el uso de equipos antiguos por la familiaridad con dichos aparatos.

El programa propuesto demuestra su fácil aplicación además de que los sensores de desplazamiento y de presión son de fácil instalación en los equipos tradicionales.

VI. Referencias

- [1] Braja M. Das. Fundamentos de Ingeniería Geotécnica. Universidad de Sacramento. California.
- [2] C. R. Vallecilla Bahena, El círculo de Mohr. Fundamentos y Aplicaciones. Edición Primera. 2004. Universidad Santo Tomas. Colombia.
- [3] E. Juárez Badillo. A. Rico Rodríguez, Mecánica de Suelos, Tomo 1.
- [4] J. E. Freud, I. Miller, M. Miller. Estadística Matemática con Aplicaciones. Dirección electrónica de la página Web. Mes y año de la consulta. Sexta Edición. 2000. Prentice Hall. México.
- [5] W. Ruíz & A. Lizcano. Universidad de los Andes, Bogotá D.C., Colombia. Diseño de un Aparato Biaxial.
- [6] T. William Lambe Robert V. Whitman. Soil Engineering. Series in. John Wiley & Sons.
- [7] B. C. Punmia, Ashok Kumar Jain, Arun Kumar Jain. Soil Mechanics And Foundations. 16th Edition. Laxmi publications. 2005.
- [8] Ali Aisen. Soil Mechanics Basic Concepts and Engineering applications. Swets & Zeitlinger B. V., Lisse, The Netherlands.
- [9] Arvind V. Shroff, Dhananjay L. Shah. Soil Mechanics and Geotechnical engineering. A. A. Balkema & Zeitlinger Publishers. 2003. India.

VII. Autores

Ing. José Fredi Robles Vázquez, es ingeniero en electrónica y estudiante de la maestría en ciencias en Ingeniería Mecatrónica del Instituto Tecnológico de Tuxtla Gutiérrez.

Dr. Rubén Herrera Galicia, es profesor investigador del Instituto Tecnológico de Tuxtla Gutiérrez, obtuvo el título de Doctor en Ciencias Técnicas con especialidad en electrónica por la *Technical University of Warsaw*, Polonia.

Dr. Héctor Hernández de León es egresado del IPN como Ingeniero en Electrónica en 1982. Maestría en Ciencias en Ing. Electrónica por el CENIDET en 2001. Doctorado en Ciencias en Sistemas Automáticos por el INSA-Toulouse, Francia en 2006. Ha publicado diferentes artículos en el Iberamia y el Core de IEEE. Actualmente labora como docente e investigador en el Instituto Tecnológico de Tuxtla.

Simulación de Vehículos Reactivos con Ruedas Diferenciales en Webots

**Carlos Delgado-Mata, Hiroki Alejandro Sasamoto Cruz, Miguel Angel Carlón,
Ramiro Velázquez**

Universidad Panamericana campus Bonaterra, Aguascalientes, Ags. cdelgado@up.edu.mx

Resumen

En el presente artículo se muestra el desarrollo de vehículos reactivos con ruedas diferenciales en el entorno para simular robots Webots. El trabajo está inspirado en mecanismos de psicofísica, en particular de propiocepción e interocepción y en el mecanismo de klinokinesis que se encuentra en algunos animales. En el documento se explica cómo construir un robot en este simulador y se menciona que el código para determinar las conductas de los vehículos es Java. Adicionalmente, se describe el desarrollo de una aplicación en Java con conexión a base de datos MySQL, para almacenar y graficar trayectorias de los vehículos. Por otro lado, se muestran algunas trayectorias de los recorridos de los vehículos y finalmente se encuentran algunas similitudes y evidentemente las diferencias que yacen en estos vehículos.

Palabras Clave: Agresivo, Braitenberg, Miedo, Vehículos, Webots,

I. Introducción

Para el buen desempeño de entes artificiales, es importante que estos contengan mecanismos de retroalimentación que les permitan tener conocimiento de estados internos. Un área de conocimiento útil para conocer los estados internos de un ser vivo es la psicofísica, término acuñado por Fechner [1], que se encarga del estudio del estímulo físico y como es percibido [2]. Para conocer estos estados internos los humanos disponemos de mecanismos como propiocepción [3] e interocepción [4]:

Propiocepción que es el sentido que informa al organismo de la posición de los músculos, además ayuda a regular la dirección y rango de los movimientos y por tanto permite respuestas autonómicas.

Interocepción que es el sentido que permite establecer el estado de órganos internos, por ejemplo un dolor de estómago.

Estos sistemas son fundamentales tanto, para tomar decisiones, como para conservar su propia integridad y seguridad. Por otro lado, la

retroalimentación en sistemas artificiales está particularmente desarrollada en control e inteligencia artificial. Por ejemplo, en el cambio de dirección y velocidad autónomo de un misil, se utiliza como parámetro la posición geoespacial en cada instante en relación al objetivo que debe alcanzar. Por otro lado, en el caso de un sistema o programa computacional puede tener sus propios estados internos; por ejemplo, la cantidad de ciclos cumplidos en sentencias tipo *do...while* o *for*, o la cantidad de memoria disponible para una operación determinada.

Por otro lado, ha habido interés en desarrollar sistemas que simulen comportamientos psicológicos más complejos de lo que en realidad son. Un ejemplo relevante son los vehículos de Braitenberg, quién en su trabajo seminal [5] nombró algunos de sus robots con los términos de miedo, amor y agresión.

Otro comportamiento interesante que desarrollan algunos animales, desde bacterias y protozoarios, hasta mamíferos como las ovejas es klinokinesis,

donde la frecuencia de cambio de dirección se altera [6].

El principal objetivo de este trabajo es la de observar los comportamientos de algunos de los vehículos de Braitenberg y el comportamiento de klinokinesis, para este fin se utilizó el simulador de robots llamado Webots [7].

II. Desarrollo

El diseño de un vehículo de Braitenberg es sencillo; en la Fig. 1 se muestra la configuración de sensores y ruedas que se usarán más adelante: 2 ruedas y 2 sensores.

El robot desarrollado fue dotado de dos ruedas y dos sensores de luz. Para proporcionar la forma al modelo geométrico deseado al robot, se debe seleccionar el árbol de objetos que se encuentra en la sección de la interfaz gráfica *Scene tree*. Al seleccionar el nodo *children* y oprimir el botón *Add New*, se agrega un nuevo árbol al cual se le deben agregar en primera instancia un sólido. A su vez, el nuevo sólido tendrá un árbol y en éste habrá nuevamente un nodo con la leyenda *children*. Dentro de este nodo se selecciona aquél que dice *Transform* y es aquí donde se le incorporará un nodo *Shape* para formar el modelo geométrico correspondiente a las llantas.

Al mismo nivel del sólido que forma la llanta, se agregan los sensores como nuevos nodos. De manera semejante, se le dan la forma, color y ubicación que se requiera.

Sobra decir que se necesitan dos llantas y dos sensores, pero cabe recalcar que cada sólido o sensor que se utiliza puede ser guardado y ser utilizado como plantilla, para posteriormente ser utilizado en modelos geométricos similares. En esta evaluación, el modelo geométrico del vehículo (robot) quedó constituido como se muestra en la Fig. 2.

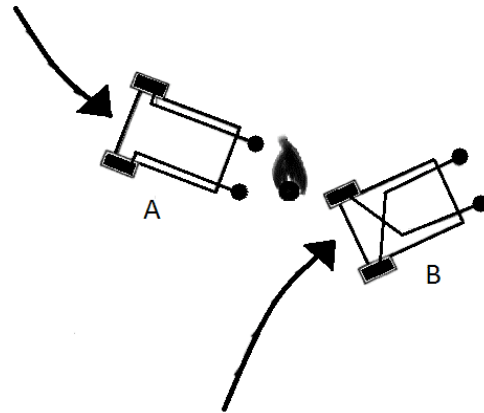


Fig. 1. Vehículo agresivo (A) y vehículo con miedo (B).

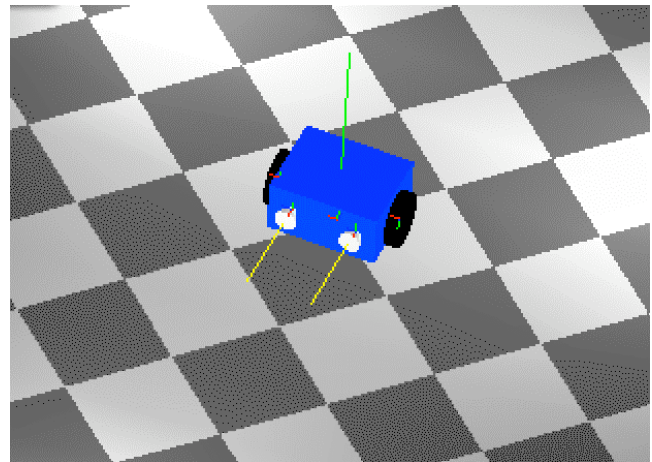


Fig. 2. Vehículo construido.

Una vez construido el robot, se prosigue con la programación del mismo. Usando las librerías, clases y métodos ya existentes y la posibilidad de crear nuevas funciones en el lenguaje Java, resulta sencillo desarrollar el comportamiento deseado.

Al llamar la clase *DifferentialWheels*, que ya existe en Webots, se puede usar gran variedad de objetos entre los cuales se encuentran los sensores de luz con los cuales se excitan los motores que mueven las llantas del robot. Para utilizarlos se declaran como *private LightSensor ls*.

En este proyecto sólo se usan dos sensores, por lo que una vez declarados falta asignarles el valor. Esto se hace dentro del constructor del programa. Para usar los valores que capturan los sensores se debe crear los objetos necesarios: `lsk=getLightSensor("ls");` y después habilitarlos con la instrucción `ls.enable(TIMESTEP)` dentro del mismo constructor.

Un problema que se puede encontrar al momento de construir un robot es el tamaño. En el propuesto en el presente trabajo la distancia entre sensores es muy pequeña, por lo que la diferencia entre las lecturas también lo es. Por lo tanto, una excitación de los motores usando una relación directa a la de los sensores, como Braitenberg lo propone, no es posible.

Una vez obtenidas las lecturas, se deben guardar a una variable para poder manipularlas: `double valor=ls.getValue();`

Para poder distinguir de manera rápida y sencilla las lecturas, a cada sensor se le asigna un multiplicador que depende de la relación entre ambos valores. Aquél que tenga una lectura mayor se multiplicará por .004 y aquél que tenga la menor lectura por .003. Estos valores fueron escogidos tal que la resultante se pudiera colocar como parámetro para la velocidad de los motores (véase Cod. 1).

```
if(Edd<Eii){
  Fiz=.8-Eii*.004;
  Fd=.8-Edd*.003;}
if(Edd>Eii){
  Fiz=.8-Eii*.003;
  Fd=.8-Edd*.004;}
if(Edd==Eii){
  Fiz=.8-Eii*.004;
  Fd=.8-Edd*.004;}
```

Cod. 1. Diferenciación de lecturas.

Para poder asignar una velocidad a los motores se usa la instrucción `setSpeed(Vi, Vd);` siendo V_i y

V_d los valores numéricos que se le quieran otorgar.

Para poder convertir el vehículo con ruedas diferenciales en un robot agresivo, el sensor derecho debe excitar a la llanta izquierda por lo que la instrucción quedaría así: `setSpeed(Vd,Vi);` de forma similar para designarle una conducta miedosa cada sensor excitará la llanta de su propio lado quedando la instrucción: `setSpeed(Vi,Vd);`.

En cuanto a los dos modelos restantes, la solución más práctica es la de restar al valor máximo que puedan alcanzar los motores el valor calculado a partir de la excitación de los sensores. Para este caso en concreto, el valor usado como máximo fue de .8.

Una vez creadas, tanto la parte del robot como su programación, hay que unirlos. En el árbol de características del robot diferencial hay una raíz que se llama *controller*, es ahí donde se indica la ubicación del controlador del robot; el código en Java.

Para el segundo experimento, se desarrollo un robot para simular el comportamiento de klinokinesis y además se desarrollo un entorno para graficar las trayectorias en un plano cartesiano donde se desplazan los robots. Para esto, se agregaron un componente para simular los giros aleatorios del robot de la misma manera que en [8] y otro componente para saber la ubicación de los robots.

III. Resultados

Los resultados del primer experimento se describen a continuación.

En la Fig. 3 se muestra el robot con miedo que exitosamente evadió la fuente que excitaba sus sensores. Se puede apreciar que mientras más cerca se encuentre de la fuente, menor será su radio de giro. Esta reacción se puede mejorar ajustando los valores por los cuales se multiplican los valores de los sensores.

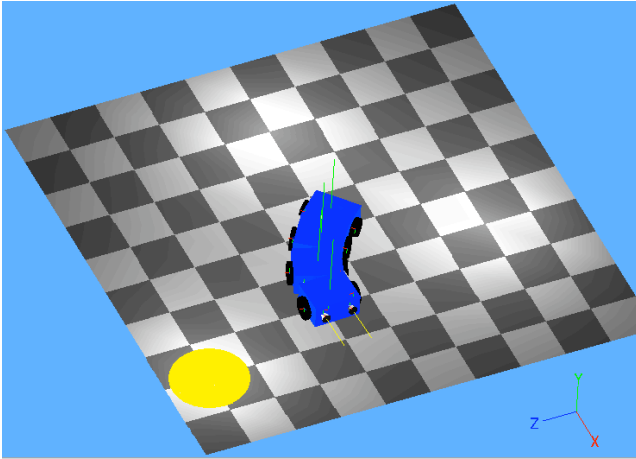


Fig. 3. Vehículo con miedo.

Por otra parte, en la Fig. 4 se puede apreciar el comportamiento agresivo que menciona Braitenberg, el de atacar la fuente que provoca la excitación, aquello que lo molesta. Este vehículo se detendrá hasta haber traspasado la fuente, debido a que ya no tendrá lectura en sus sensores dado que su fuente habrá quedado a sus espaldas.

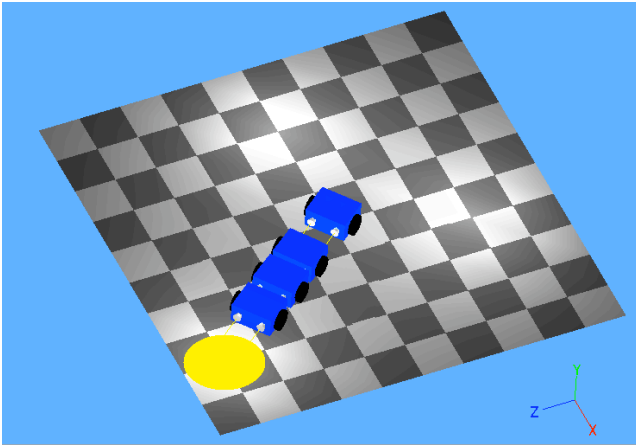


Fig. 4. Vehículo agresivo.

En la Fig. 5, el robot que ama se detiene a contemplar la fuente de excitación que admira. Mientras más se acerque más lento irá, hasta llegar un punto en el que probablemente se detenga por completo. La palabra probablemente cabe dado que sólo se detendrá si el robot sigue una línea recta donde ambos sensores obtengan la misma lectura.

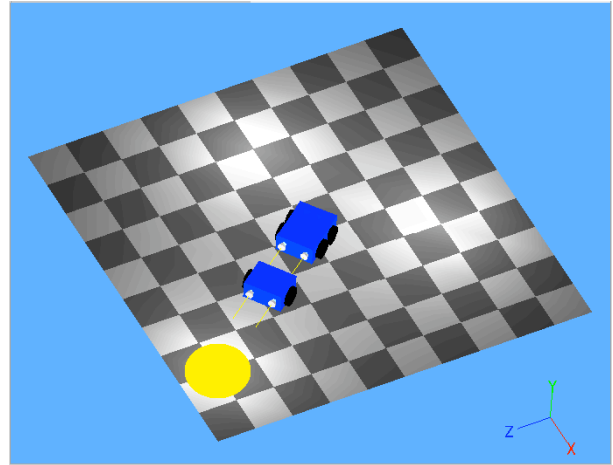


Fig. 5. Vehículo que ama.

El último ejemplo es el correspondiente al robot explorador (ver Fig. 6), el cual sería el análogo al agresivo, en cuanto a conexiones se refiere, pero un tanto similar al robot con miedo en cuanto a comportamiento se refiere, pues a pesar de tener diferentes nombres ambas conductas se alejan de la fuente de excitación más cercana. La diferencia yace en que cuando se encuentran en el vacío, el robot con miedo se quedará quieto mientras que el explorador ambulará hasta encontrar algo más que esquivar.

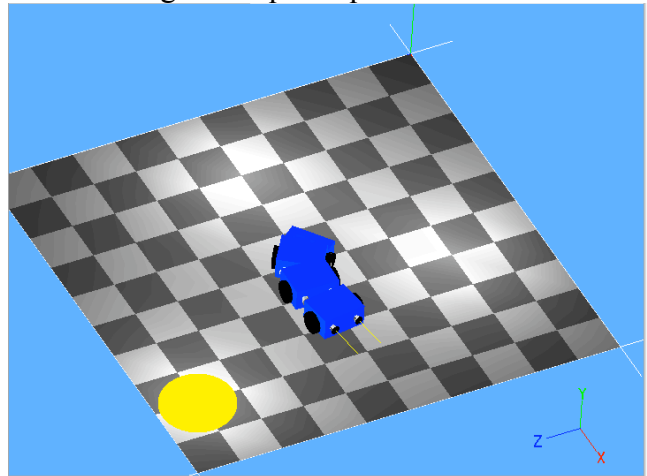


Fig. 6. Vehículo explorador.

Los resultados del segundo experimento (klinokinesis) se describen a continuación:

En el entorno se colocaron 5 robots, tal como se muestra en la Fig. 7

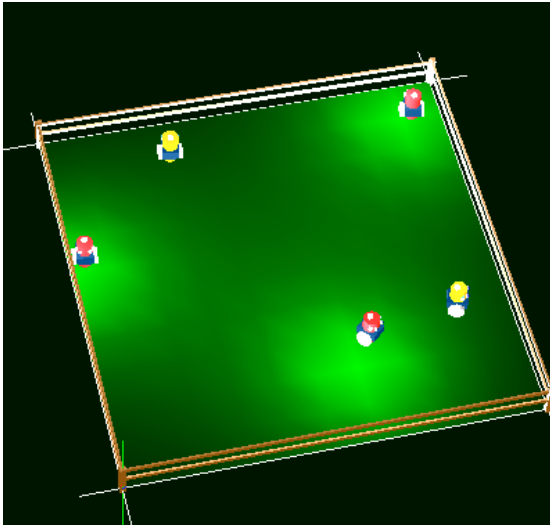


Fig. 7. Vehículos con comportamiento klinokinesis en webots.

Para graficar las posiciones de los robots de desarrollo una aplicación en java que recibe mensajes desde webots que son almacenados en una base de datos MySQL (ver Fig. 8), posteriormente se leen los datos almacenados para generar una gráfica con las trayectorias de los robots, como se muestra en la Fig. 9.

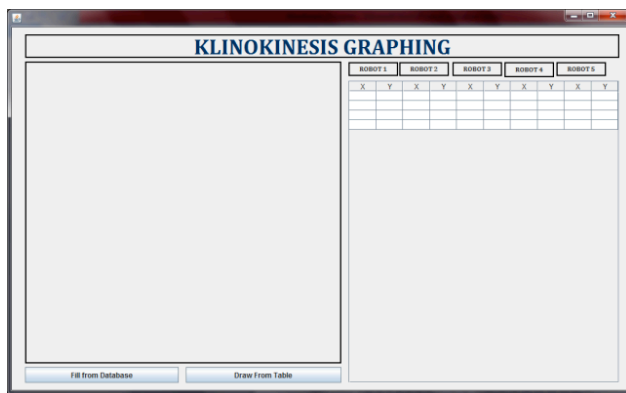


Fig. 8. Aplicación para graficar trayectorias.

Una vez leídas las coordenadas que fueron almacenadas en la base de datos MySQL, la aplicación de Java grafica las trayectorias que siguieron cada uno de los robots, como se muestra en la Fig. 10. Las trayectorias se representan con trazos de diferente color. Cabe mencionar, que para saber la dirección de la trayectoria se gráfico el punto de inicio con el trazo mas grueso y el punto final con el trazo mas delgado; esto es, el trazo se va adelgazando conforme el robot fue avanzando.

ROBOT 1		ROBOT 2		ROBOT 3		ROBOT 4		ROBOT 5	
X	Y	X	Y	X	Y	X	Y	X	Y
281	401	189	224	102	377	222	127	205	190
286	389	189	224	103	377	222	127	204	202
293	377	185	233	103	377	222	127	204	202
298	370	180	241	103	377	222	127	196	34
298	370	180	243	103	377	222	127	190	40
298	370	178	247	106	376	222	127	185	50
299	369	178	247	124	377	222	127	184	51
300	369	178	247	154	380	221	127	184	52
306	363	178	247	191	382	222	127	183	56
308	361	178	247	216	380	224	128	183	56
315	355	177	249	230	376	234	134	183	56
324	348	175	251	232	375	240	138	181	63
332	343	175	251	232	375	240	138	178	80
333	342	175	251	232	375	240	138	170	109
333	342	175	251	233	375	240	138	162	133
333	342	172	255	240	371	240	138	155	148
333	342	167	262	241	370	240	138	149	153
333	342	164	267	241	370	241	139	148	154
334	342	164	267	246	367	242	140	148	154
334	342	164	267	257	360	246	143	148	154
334	342	164	267	267	354	248	145	148	154
334	342	164	267	268	354	248	145	148	154
334	341	164	267	268	354	249	146	144	157
338	335	164	267	268	354	253	152	138	161
345	322	164	267	272	352	255	153	123	167
347	319	163	268	278	348	256	155	99	173
347	319	162	270	282	345	262	161	70	181
348	318	161	271	282	345	266	163	43	190
348	318	160	273	282	345	281	167	26	202
348	318	160	273	282	345	294	167	23	209

Fig. 9. Datos de trayectoria de robots, recuperados desde MySQL.

En la Fig. 10, donde se simula klinokinesis, cada uno de los robots tiene una trayectoria aleatoria. Sin estímulos que lo hagan tender hacia alguna dirección. No existe ninguna relación entre cualquiera de las trayectorias que se muestra en las Fig. 10. Es más, cada una de las simulaciones representadas con las tres imágenes se aprecia que las trayectorias seguidas por cada uno de los robots no se parece entre ellas por lo que argumentamos que esta simulación sencilla puede representar el comportamiento de klinokinesis.

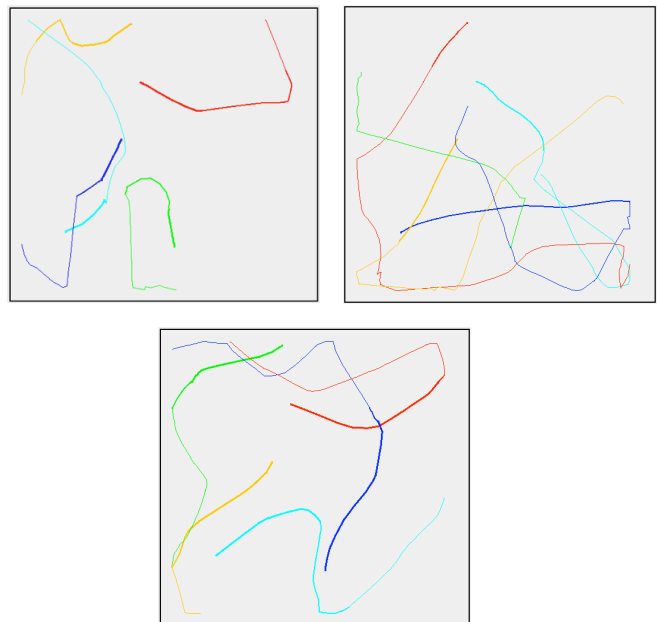


Fig. 10. Trayectorias generadas en tres simulaciones.

IV. Discusión

Al utilizar el entorno de simulación webots se puede percatar de la versatilidad que tiene al usarse, ya que se puede cambiar el entorno y observar cómo se comporta el robot bajo diferentes estímulos de manera muy sencilla: Esto es, se puede aumentar o disminuir el número de fuentes de excitación así como variar su ubicación. Por otro lado, también se puede desplazar el robot en el mundo creado y colocarlo como sea conveniente.

Ahora, analizando únicamente las trayectorias seguidas por las diferentes conductas, se puede apreciar que no basta con tan sólo cambiar de motor el valor del sensor; sino que, para cada conducta se tiene que calibrar de forma diferente.

Por ejemplo, el vehículo explorador, apenas detectó la luz empezó a cambiar su trayectoria. Para algunos esto no equivaldría a encontrar lo que se andaba buscando, quizá si se llegase a acercarse más a la fuente entonces se podría llamar un robot explorador. Aunque otro enfoque que se le otorga a esta conducta es aquella en la que se dice que a este robot le gusta permanecer más tiempo en aquellos lugares que le son gratos sin tener la intención de permanecer quieto.

V. Conclusiones

Se puede argumentar que los vehículos reactivos (aunque contengan un diseño sencillo) pueden simular conductas y actos tales como atacar, escapar, contemplar y vagar.

Sin importar los nombres con los cuales se les llamen a estos vehículos, el poder representarlos de manera virtual permite observar sus reacciones ante diferentes circunstancias sin la necesidad de crear un entorno físico. Por lo que se facilita la construcción de entornos para aquellos que sólo quieren estudiar las conductas de estos vehículos.

La importancia de este tema es más apreciada por personas que quieren profundizar en temas como inteligencia artificial orientada emocionalmente, pues es con este tipo de estudios que pueden

identificar los patrones a seguir para poder simular diferentes conductas.

La aplicación para trazar trayectoria se encontró de gran utilidad para analizar el comportamiento de los robots que simularon el mecanismo de klinokinesis.

También existe otra forma de aprovechar este conocimiento: intentar predecir el comportamiento y reacciones que podrían tener los seres vivos hacia cambios en su entorno.

VI. Referencias

- [1] G. T. Fechner, *Elemente der Psychophysik (Elements of Psychophysics)*, Germany, 1860.
- [2] G. Gescheider, *Psychophysics: the fundamentals (3rd ed.)*. Lawrence Erlbaum Associates, 1997
- [3] C.S. Sherrington, *On the proprioceptive system, especially in its reflex aspect*. *Brain* 29, p 467-85, 1907.
- [4] A.D. Craig, *How do you feel ? Interoception: the sense of the physiological condition of the body*. *Nat Rev Neurosci* 2002, 3: 655-666.
- [5] V. Braitenberg, *Vehicles: Experiments in Synthetic Psychology*. The MIT Press. Massachusetts. 1986.
- [6] K. Lorenz, *The foundations of Ethology*, Springer-Verlag, Heidelberg, Germany, 1981.
- [7] *Webots 6: Fast prototyping and simulation of mobile robots*, <http://www.cyberbotics.com/>, consultado 10 de Julio 2011.
- [8] C. Delgado-Mata, R.S. Aylett, *Fear and the Behavior of Virtual Flocking Animals*, In. *Proceedings of ECAL 2007, LNAI 4648*, Springer, Verlag, 2007.

Interfaz Gráfica y Sistema de Mando Inalámbrico para la Enseñanza de Niños con Problemas Neuromotores.

Mariana Sosa Terrazas, **José Sergio Orduña Díaz, ***José Luis Avendaño Juárez, *Gilberto Herrera Ruíz.**

*Facultad de Ingeniería Universidad Autónoma de Querétaro, cerro de las campanas s/n Col. Las Campanas Qro. Qro., (442)1921200 ext 6015, msosa24@alumnos.uaq.mx.

**Facultad de Ingeniería Universidad Autónoma de Querétaro, cerro de las campanas s/n Col. Las Campanas Qro. Qro., (442)1921200 ext 6015, sergio_orduna_diaz@hotmail.com

*** Facultad de Ingeniería Universidad Autónoma de Querétaro, cerro de las campanas s/n Col. Las Campanas Qro. Qro., (442)1921200 ext 6015, luis.avendano@uaq.mx.

****Facultad de Ingeniería Universidad Autónoma de Querétaro, cerro de las campanas s/n Col. Las Campanas Qro. Qro., (442)1921200 ext 6015, gherrera@uaq.mx.

Resumen

El presente trabajo se desarrolló mediante la vinculación de la Facultad de Ingeniería de la Universidad Autónoma de Querétaro (UAQ) y el Centro de Atención Múltiple (CAM) “Mis Primeros Pasos”, con el objetivo de brindar una herramienta tecnológica que apoye a la comunicación de los alumnos que tienen problemas neuromotores; dicha herramienta se divide en una interfaz gráfica y un sistema de mando inalámbrico de bajo costo. La interfaz gráfica sirve para estimular el desarrollo cognitivo del alumno y como un recurso más en la aplicación de sus conocimientos adquiridos en clase. Por otro lado, el sistema de mando inalámbrico permite comunicar al alumno con la interfaz, mediante el envío de mandos específicos con los cuales puede manipularla. El trabajo se encuentra en uso desde agosto del 2010 en el CAM.

Palabras Clave: Comunicación inalámbrica, interfaz gráfica, microcontroladores, módulo de comunicación inalámbrico, sistemas inalámbricos.

I. Introducción

El CAM “Mis Primeros Pasos” (ubicado en la 1ra. Privada de corregidora No. 22 col. San Pablo) con turno vespertino y clave: 22DML0011x, pertenece a la Unidad de Servicios para la Educación Básica en el Estado de Querétaro (USEBEQ). El CAM cuenta con una población de 84 alumnos, los cuales presentan capacidades distintas, como: debilidad visual, discapacidad auditiva, problemas neuromotores, discapacidad intelectual y algunos con discapacidad múltiple.

Por otro lado, en los últimos 30 años, el uso de las tecnologías de asistencia se ha vuelto una

herramienta muy importante para ayudar a que las personas que sufren alguna discapacidad tengan una mejor calidad de vida. El término de tecnología de asistencia incluye a cualquier dispositivo o software de apoyo, adaptabilidad y rehabilitación para aquellas personas con discapacidades. Este grupo de dispositivos y software permite a estas personas realizar tareas que antiguamente no podían cumplir o tenían gran dificultad para hacerlo [12].

Las tecnologías de asistencia proveen mejoras o cambios en los métodos de interacción con la tecnología necesaria para cumplir ciertas tareas. Estas tecnologías existen porque, en general, el resto de las tecnologías son pensadas y creadas

para personas sin discapacidades evidentes. Algunas de estas son: los programas de reconocimiento de voz, los sintetizadores de voz, los lectores de pantallas, texto predictivo, distintas configuraciones para el mouse, monitores especiales, etc. [11]

Por otro lado, en los últimos años se han desarrollado diferentes trabajos en el área de las tecnologías de asistencia utilizando herramientas tales como la computadora u otro tipo de tecnologías avanzadas [13].

Las tecnologías de asistencia pueden ayudar a las personas que presentan algún tipo de discapacidad, como pueden ser: traumatismo cerebral [14] y [15], accidentes cerebro-vascular [16] y [17], discapacidad de aprendizaje [18], [19] y [20], esclerosis múltiple [21]; y ha mostrado potencial en ayudar a algunas personas con demencia [22], autismo [23], [24] y [25], y retardo mental [26].

En el CAM en este último año se ha trabajado en la implementación de una herramienta tecnológica para ayudar a que los alumnos desarrollen nuevos conocimientos que les permitan tener una mejor calidad de vida. Dado que los alumnos de esta escuela tienen diferentes tipos y niveles de discapacidad se tiene una importante área de oportunidad para implementar diferentes proyectos que ayuden en la vida diaria de los alumnos.

El proyecto se adecuó a un alumno que cursaba el tercer grado de primaria especial, dicho alumno presentaba un gran avance con las metodologías que implementaba la maestra dentro del aula, se observó que tenía un gran interés por manipular diferentes objetos, también; se encontraba en un nivel silábico en el cual lograba reconocer una gran cantidad de letras y formaba algunas palabras, en su mayoría en mayúsculas; podía contar sílabas cerrando los ojos por cada separación, lograba hacer relaciones entre los objetos y las palabras que lo definían, así mismo la correspondencia entre la cantidad de objetos con el número correspondiente y la identificación de colores.

De acuerdo a las necesidades presentadas por el CAM el proyecto debía cumplir con los siguientes requisitos: una interfaz gráfica donde se visualice el abecedario, asimismo al seleccionar una letra esta debe desplegarse en la parte inferior de la pantalla para ir formando palabras, una pantalla donde se trabajen las figuras geométricas asociándole un nombre y un color, otra pantalla donde se realicen operaciones básicas (suma y resta) y finalmente la implementación del juego del ahorcado [7] y todo esto controlado por el alumno a través de un pedal y un botón, para direccionar el cursor en la pantalla y así seleccionar cualquiera de los elementos que aparezcan en esta.

El trabajo aquí presentado está dividido de la siguiente forma: en la sección dos se explica el diseño, la programación de la interfaz gráfica en una computadora de escritorio y el desarrollo del sistema de mando inalámbrico, en la siguiente sección de este trabajo se exponen los resultados obtenidos, en la cuarta sección se hace una breve discusión de los mismos y finalmente se presenta la sección de conclusiones.

II. Desarrollo

II.I Interfaz Gráfica.

En esta sección se explica a detalle la implementación de la interfaz gráfica, la cual se desarrolló bajo Leguaje C y su plataforma fue Turbo C [1] y [3]. La aplicación cuenta con varias pantallas las cuales son: Pantalla principal (Menú) de la cual parte el acceso a las demás pantallas, otra es un teclado virtual que contiene todas las funciones y caracteres de un teclado normal, además, otra es la parte donde se trabajan las figuras geométricas básicas con nombres y colores, una más en la cual se realizan operaciones básicas (suma y resta) y se diseñó también otra pantalla donde se implementó el juego del ahorcado, todas estas aplicaciones se crearon para apoyar el desarrollo cognitivo del alumno. Ya que la interfaz gráfica es para niños se trató de hacer visualmente agradable y no sólo funcional (ver Fig. 1).

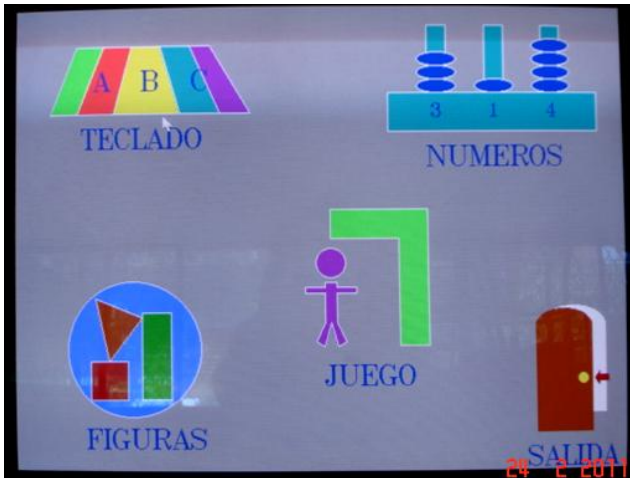


Fig. 1. Pantalla principal de la interfaz Gráfica (Menú).

El diagrama de flujo de la figura 2, presenta el funcionamiento general de la interfaz, al acceder al Menú el alumno mediante el sistema inalámbrico de mando puede seleccionar entre las cuatro opciones que se tienen, a partir de este momento el programa sólo funciona a través del sistema de mando o de un mouse externo. El alumno al entrar al Menú selecciona la opción que prefiere, entonces el programa verifica que la opción sea válida mediante una comparación con valores previamente establecidos, si la opción es válida el alumno accederá a la aplicación, de lo contrario saldrá de la misma.

El teclado virtual posee las mismas funciones de un teclado normal, ya que trabaja con todas las letras del abecedario, los números hindú-arábigos y caracteres especiales; contando con la opción de letras en mayúsculas o minúsculas, borra caracteres anteriores y saltar renglón.

Se consideró importante que cuando una letra fuese seleccionada esta resaltara de las demás por lo que se le agregó sonido y un cambio de color en la misma. Todos los caracteres que son seleccionados son colocados en la parte inferior de la pantalla (ver Fig. 3).

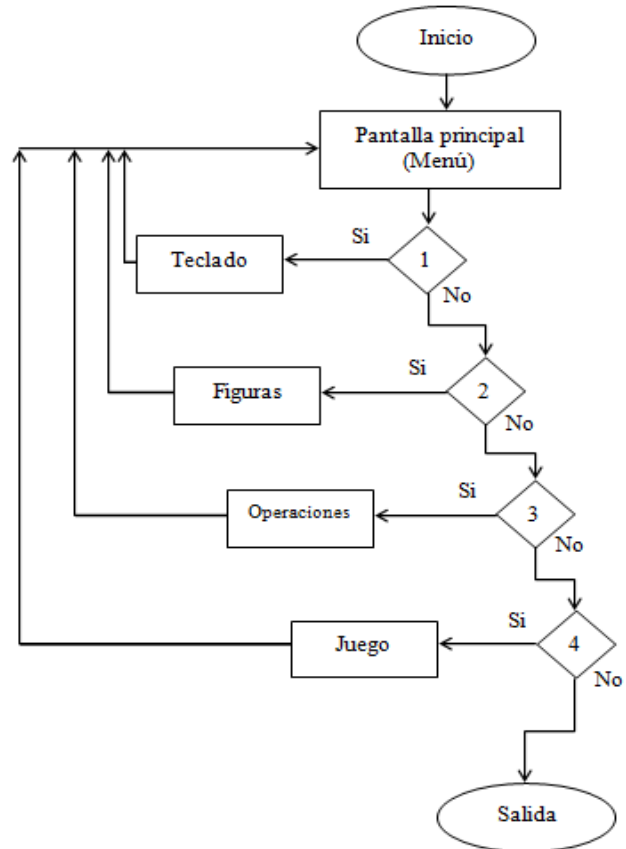


Fig. 2. Diagrama de flujo para el diseño de la Interfaz grafica.



Fig. 3. Teclado Virtual de la Interfaz Gráfica.

La pantalla con figuras geométricas básicas contiene al: círculo, cuadrado, triángulo y al rectángulo; la región interna de cada una de las figuras esta rellena de diferente color y en un recuadro tiene su nombre; una adecuación extra es que, en el momento que sea necesario, se puede ocultar el nombre y trabajar en la relación

de figuras, nombres y colores con el niño (ver Fig. 4).

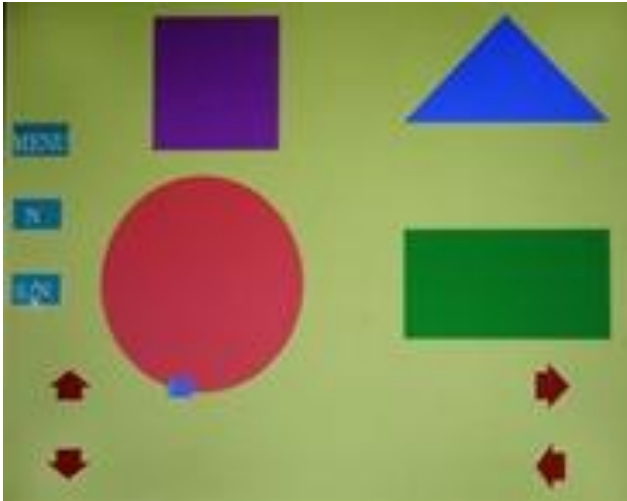


Fig. 4. Pantalla de Figuras Geométricas.

La aplicación de operaciones básicas realiza sólo sumas y restas de 2 dígitos y el resultado se presenta después de la operación. En esta pantalla sólo se muestran los números hindú-arábicos y los símbolos operacionales de la adición, sustracción e igualación; en el momento de ser introducido algún número por el alumno éste aparecerá en la parte central de la pantalla y de igual forma el resultado (ver Fig. 5).

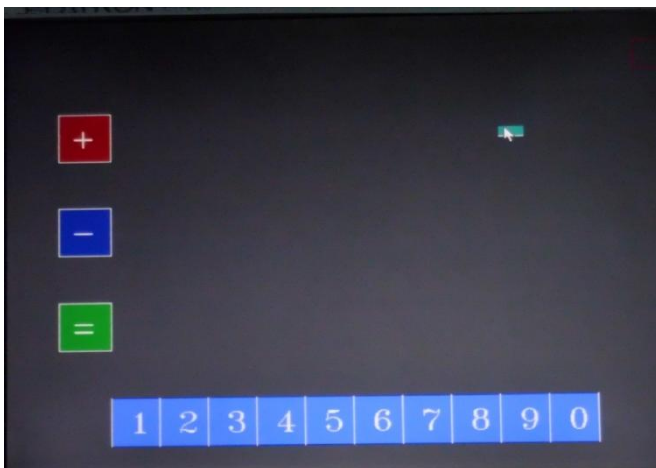


Fig. 5. Pantalla de Operaciones Básicas.

La implementación del juego del “Ahorcado” se realizó de tal forma que el alumno pueda comenzar a adivinar la palabra oculta ya sea letra por letra o por palabras enteras, mientras que en el costado de la pantalla se muestra la figura de

un “hombrecito” que se va dibujando cada que se genere un error. Cabe mencionar que la palabra a adivinar se programa al inicio del juego por el maestro y se pueden introducir palabras de hasta 40 letras (ver Fig. 6).



Fig. 6. Pantalla del Juego del Ahorcado.

II.II Sistema de Mando Inalámbrico.

El Sistema de Mando está compuesto por un Transmisor y un Receptor, cada uno construido para un propósito específico. El transmisor se encuentra montado en la silla de ruedas del niño, este tiene la encomienda de enviar las órdenes de mando para mover el cursor del mouse, dichas órdenes son codificadas para ser enviadas hacia el Receptor por medio de comunicación inalámbrica.

El Transmisor contiene cuatro botones correspondientes a los pedales que se montaron en la silla de ruedas del alumno (dos por pedal, izquierda-derecha y arriba-abajo) y un botón pulsador (tipo hongo) que su función es seleccionar un objeto en la interfaz gráfica. Todas las señales que se generan de los botones y del pulsador se procesan a través de un PIC (PIC16F876A)[5], las señales son acondicionadas y codificadas para ser enviadas a través de un dispositivo de transmisión inalámbrica (XBee-PRO® 802.15.4 OEM RF) [6]. El Transmisor utiliza una batería de alimentación de 9 volts (ver Fig. 8).

Por otro lado, el diagrama de flujo de la figura 7 describe el funcionamiento del Transmisor y este se puede describir de la siguiente manera: siempre está monitoreando si se presionó algún botón (pedales o el botón seleccionador tipo hongo), si lo hay verifica qué botón se está presionando y le asigna un carácter según sea el caso (“A” izquierda, “B” derecha, “C” arriba, “D” abajo, y “E” selección), este dato se envía a través del Módulo Maestro del Puerto Serial Síncrono (MSSP, Master Synchronous-Serial Port Module) [5] del PIC hacia el módulo de transmisión inalámbrica. En caso de que se genere un carácter erróneo que no corresponda a ninguno de los caracteres antes mencionados el sistema se mantiene en estado de espera de un carácter válido. El software utilizado para compilar fue el PIC C Compiler [8], y para grabar el programa en el PIC se usó el programador MasterProg [4], [9].

El Módulo Maestro del Puerto Serial Síncrono (MSSP Master Synchronous-Serial Port Module) [5], transmite por el puerto serial el carácter correspondiente del botón presionado, de esta forma se codifica cada botón y se sabrá con exactitud qué carácter se envió. El programa está preparado para que en caso de detectar un botón presionado y este no se encuentre en la lista, simplemente no hace nada y se regresa a preguntar si está presionado algún botón, esto es con el fin de que el programa no tenga una incoherencia y que el PIC pueda entrar en error o modo fallo o simplemente deje de operar. En la figura 9 se presenta la conexión eléctrica del módulo de transmisión inalámbrica y el PIC [2].

El Receptor está compuesto por una sección de decodificación, y dependiendo de la señal decodificada se moverá uno de los dos motores a pasos que se encuentran en la sección de Movimiento, cada uno de los motores tiene una función específica ya que uno gobierna el movimiento horizontal del cursor, mientras que el otro se encarga del movimiento en el eje vertical, esto se logra acoplando un codificador (Encoder)

a la flecha de cada motor, así cada que se active cualquiera de los motores este hará girar el codificador y un sensor fotoeléctrico (para monitorear la posición) detectará hacia donde fue el movimiento (dicho movimiento será el del cursor del mouse en la interfaz gráfica), cada que se detecta un movimiento, este es enviado a una tarjeta que sirve de interfaz y que trabaja bajo el protocolo PS2 [10], de esta forma el puntero del mouse se moverá en las direcciones ya establecidas (Arriba-Abajo, Izquierda-Derecha), lo anterior se ejemplifica en la figura 10.

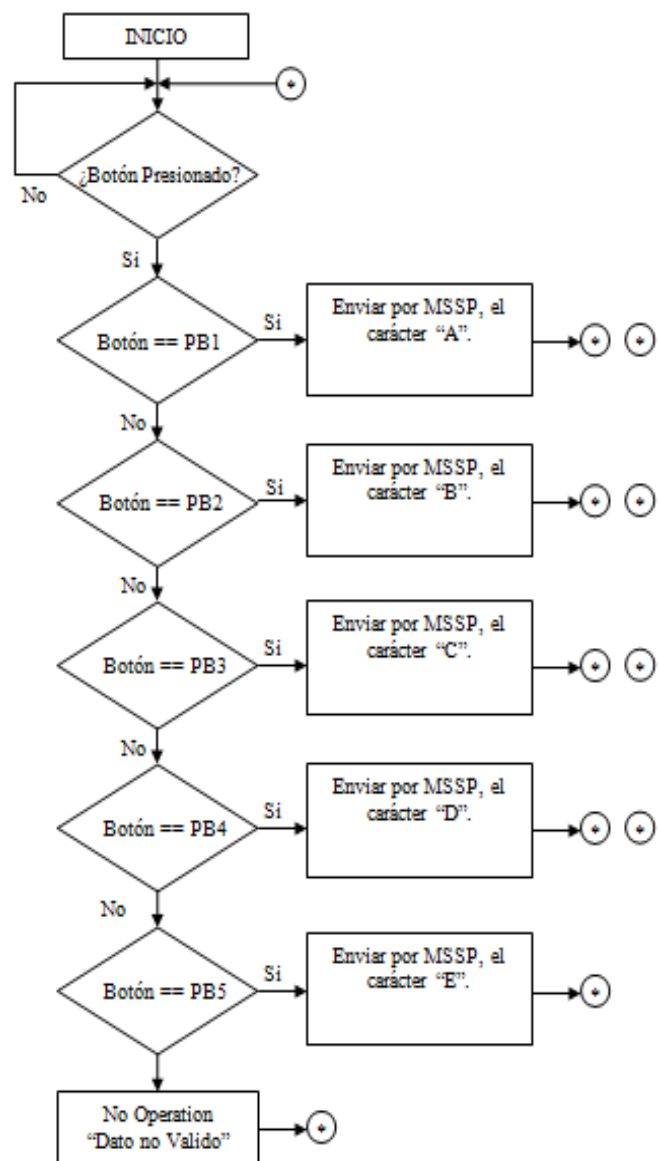


Fig. 7. Diagrama de flujo del Emisor.

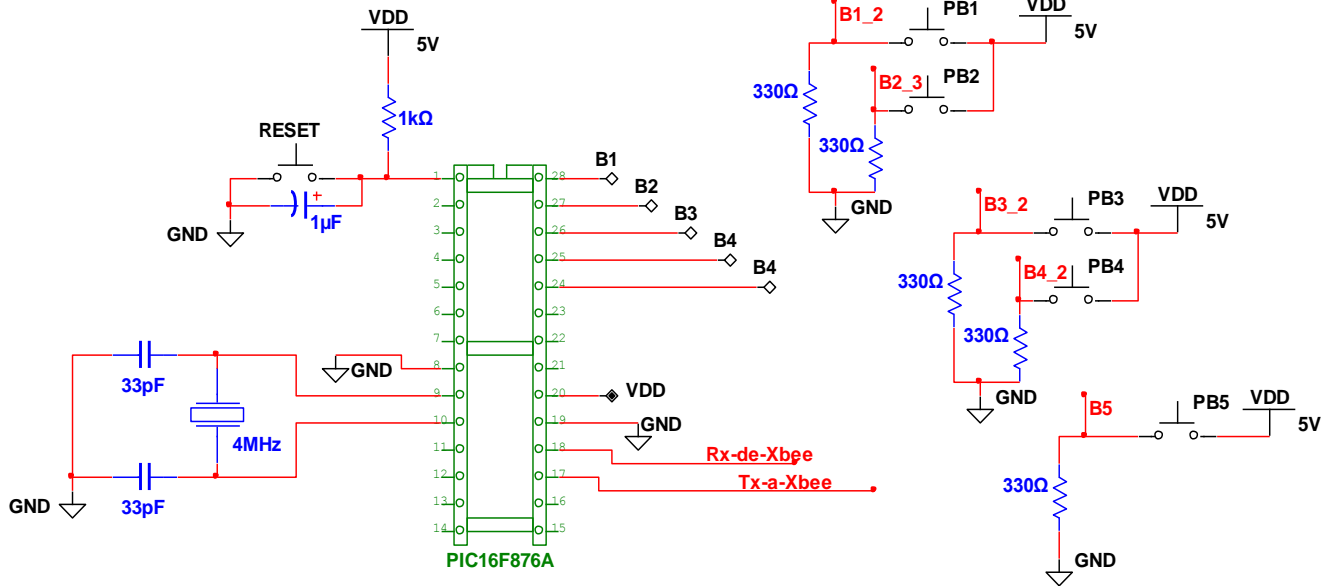


Fig. 8. Diagrama eléctrico del Emisor.

En el diagrama de flujo de la figura 11 muestra el monitoreo continuo de datos en el registro del puerto serial, si lo hay verifica que clave se está recibiendo y ejecuta la acción asignada a esa clave, en este caso son los caracteres “A”, “B”, “C”, “D”, y “E”. En caso de haber un dato y no corresponder a ninguna de las anteriores claves, el sistema lo marca como un carácter no valido, regresando al inicio del programa para esperar un nuevo dato en el registro del puerto serial, esto es con el fin de que el programa no tenga alguna incoherencia y esto provoque que el PIC entre en algún modo de fallo.

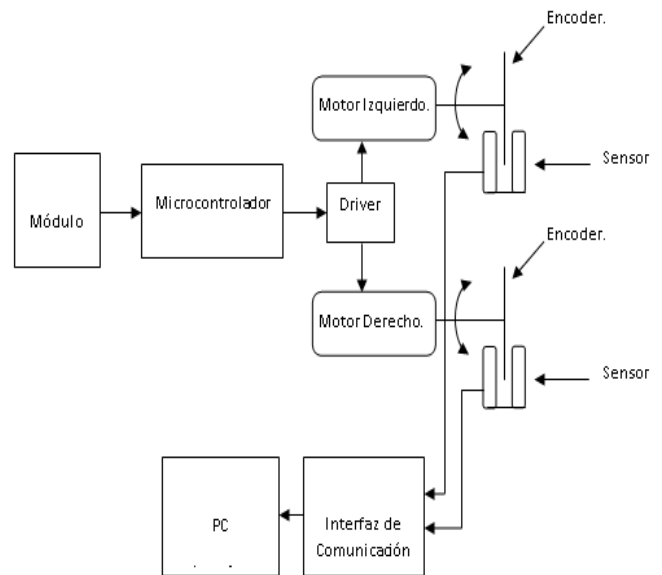


Fig. 10. Diagrama a bloques del Receptor.

Para terminar con esta sección se presentan las adecuaciones que se le hicieron a la silla de ruedas donde el niño pasa la mayor parte del tiempo y la usa como su transporte ya que sufre de parálisis cerebral y atetosis (movimientos involuntarios en todo el cuerpo).

En la figura 12 se presentan los pedales que se instalaron en la silla de ruedas y que sirvieron para que el niño pudiera desplazar el cursor

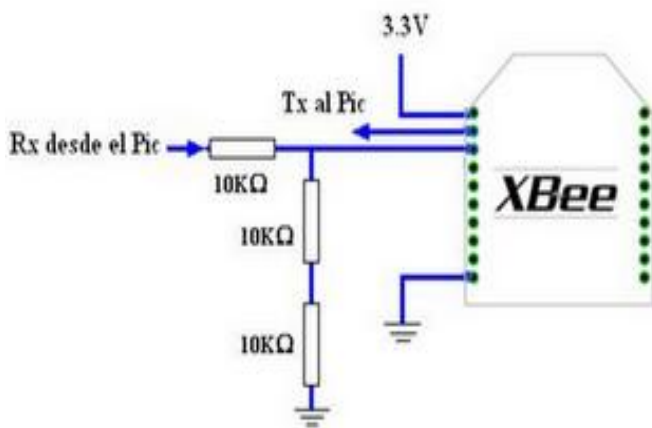


Fig. 9. Diagrama eléctrico del módulo del Emisor.

dentro de la interfaz gráfica y así seleccionar lo que el deseara. El niño esta acomodado de tal forma que con sus pies (punta y talón) pueda

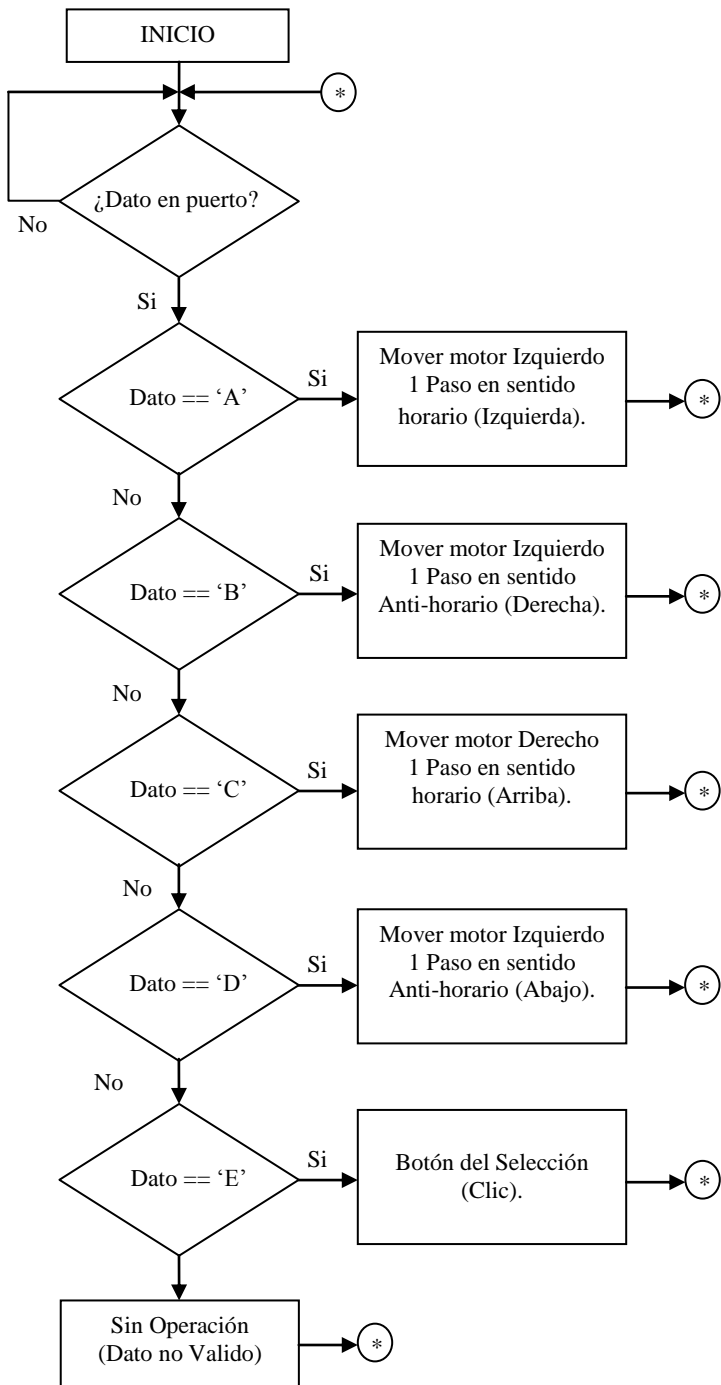


Fig. 11. Diagrama de flujo implementado en el Receptor.

realizar los movimientos de arriba-abajo, izquierda-derecha, del cursor del mouse. Cabe mencionar que por ser una silla acondicionada

para este propósito los pedales son móviles, es decir se pueden montar y desmontar, lo cual tiene como consecuencia que cuando el niño crezca sea necesario ajustarlos o cada que se siente en la silla se tenga que calibrar el sistema.

Por otro lado se colocó un botón para realizar la función del “clic” del mouse (seleccionar objetos). En la figura 13 se muestra dicho botón que el niño usa para seleccionar las diferentes opciones que se muestran en la interfaz gráfica.

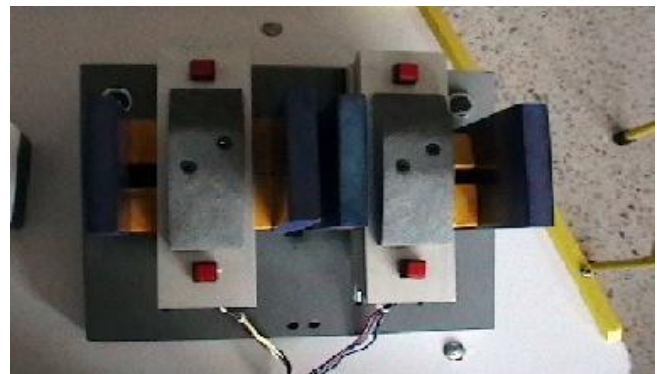


Fig. 12. Pedales instalados en la silla de ruedas.



Fig. 13. Botón para la selección en la interfaz gráfica.

III. Resultados

Para las pruebas finales del sistema fue necesario diseñar las etapas de alimentación de los diferentes módulos que lo componen. En las figuras 14 y 15 se presentan los módulos del

Emisor y del Receptor. Mientras que en la figura 16 se presenta el sistema ya terminado.

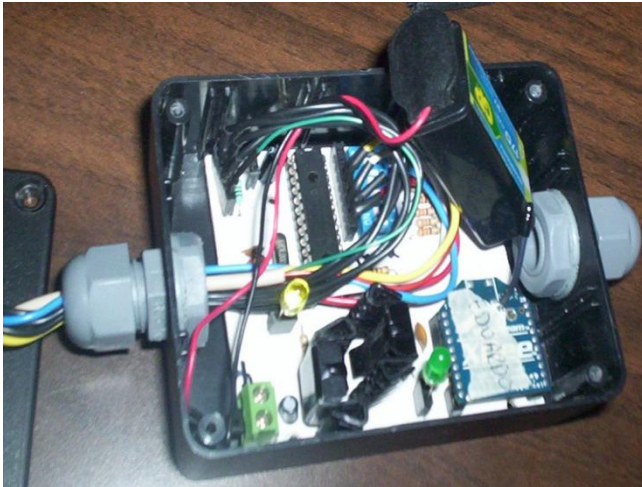


Fig. 14. Módulo del Emisor.

Junto con las adecuaciones en la silla de ruedas se hicieron varias pruebas con el sistema, el alumno probó y ayudo al equipo a calibrar el sistema para hacerle las adecuaciones finales, en la imagen de la figura 17 se presenta la silla del alumno mientras se le están haciendo los ajustes necesarios.



Fig. 15. Módulo del Receptor y fuente de alimentación.

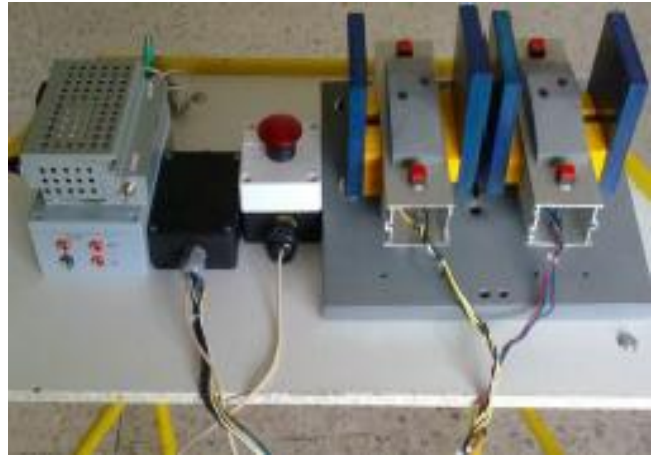


Fig. 16. Sistema de mando inalámbrico para la comunicación con la interfaz gráfica.

IV. Discusión

El proyecto se entregó en el mes de junio de 2010 y es usado por el alumno desde agosto de ese año. Es necesario comentar que los participantes del proyecto trabajaron varias sesiones con el niño para que se adecuara al sistema y que se realizaran las diferentes actividades implementadas en la interfaz gráfica.

Algo importante de mencionar sobre este proyecto es que está diseñado para el uso particular de un sólo niño. Se debe de trabajar para que el sistema sea genérico y que impacte en el aprendizaje de más alumnos y por lo tanto tengan una mejor calidad de vida y una mayor oportunidad de desarrollo.



Fig. 17. Adecuaciones a la silla de Ruedas.

V. Conclusiones

En el presente trabajo se ha explicado el diseño e implementación de una adecuación tecnológica para el desarrollo pedagógico de un niño con problemas neuromotores; el cual fue calibrado y diseñado bajo las sugerencias y características requeridas por la maestra del niño y el mismo niño, de acuerdo a sus condiciones e interacción con la herramienta. Dicho proyecto se realizó en dos partes: el diseño de una interfaz gráfica y un sistema de mando inalámbrico. La interfaz está compuesta por las siguientes pantallas: teclado virtual, figuras geométricas, juego del ahorcado y pantalla de operaciones. Todas ellas fueron programadas en Lenguaje C y su plataforma Turbo C. La interfaz gráfica es controlada por el niño mediante un sistema de mando inalámbrico, y así el alumno desde su silla de ruedas gobierna los movimientos del mouse que está en la interfaz, el alumno puede moverlo a la izquierda, a la derecha, arriba o abajo. Dicho sistema está compuesto por unos pedales, un sistema inalámbrico de transmisión-recepción, una comunicación serial del receptor a la PC donde está cargada la interfaz gráfica. Cabe mencionar que se usó un microcontrolador PIC como medio de control y sistema inalámbrico XBee-PRO® 802.15.4 OEM RF para la transmisión y recepción de las señales de mando. Otra de las aportaciones de este trabajo es el hecho de que el sistema está pensado para usar baterías de 9V, lo que lo hace un sistema portable y económico.

Así el resultado final de este proyecto es una adecuación una adecuación tecnológica de asistencia personalizada que brinda un resultado adecuado para el niño y sus maestras.

VI. Referencias

- [1] B. Ezzell, Programación de gráficos en Turbo C++. 1^{ra} Edición. 1993. Editorial Díaz de Santos. España.
- [2] E. García, Compilador C CCS y simulador PROTEUS para Microcontroladores PIC. 1^{ra} Edición. 2008. Alfaomega. México.
- [3] A. L. Joyanes Aguilar, Curso de programación con Turbo C++, 2da edición. 1996. McGraw Hill. España.
- [4] A. R. Pallás, Microcontroladores: fundamentos y aplicaciones con PIC, 2007, Marcombo, España.
- [5] Hoja de datos del microcontrolador PIC16F87X. <http://ww1.microchip.com/downloads/en/devicedoc/30292c.pdf>. abril del 2010.
- [6] Hoja de datos del módulo XBee / XBee-PRO OEM RF Modules. Digi International Inc. USA. 2008.
- [7] Juego del ahorcado. <http://www.misjuegos.com.mx/destreza/ahorcado-en-espanol/>. Mayo 2010.
- [8] Manual de usuario del PIC C http://www.ccsinfo.com/downloads/ccs_c_manual.pdf. Mayo 2010.
- [9] Dispositivos soportados por el Programador MasterProg <http://edutronika.webs.com/Publicacion/DISPOSITIVOS.pdf> Mayo 2010.
- [10] Protocolo PS2 <http://www.computer-engineering.org/ps2protocol/>. Mayo 2010.
- [11] Definición de Tecnología de Asistencia <http://www.alegsa.com.ar/Dic/tecnologia%20de%20asistencia.php>. Agosto 2011
- [12] LoPresti, Edmundo F., et al. (2004). Assistive technology for cognitive rehabilitation: State of the art, Technology for cognitive rehabilitation. En: Neuropsychological Rehabilitation, AT Sciences, Pittsburgh, PA, USA., Enero-Junio, No. 14 (1/2), pp. 5-39.
- [13] Bergam, M.M (1998). A proposed resolution of remediation-compensation controversy

controversy in brain injury rehabilitation. *Cognitive Technology*, 3(1), pp. 45-51.

[14] NIH (1), pp. 1-41 (United States National Institutes of Health) (1998). Rehabilitation of persons with traumatic brain injury. NIH Consensus Statement, 16

[15] Wilson, B. A., Evans, J. J., Emslie, H., & Malinek, V. (1997). Evaluation of NeuroPage: A new memory aid. *Journal of Neurology, Neurosurgery, and Psychiatry*, 63, pp. 113-115.

[16] Evans, J.J., Emslie, H., & Wilson, B. A. (1998). External cueing systems in the rehabilitation of executive impairments of action. *Journal of the International Neuropsychological Society*. 4(4), pp. 399-408.

[17] Wright, P., et al., (2001). Comparison of pocket-computer memory aids for people with brain injury. *Brain Injury*, 15(9), pp. 787-800.

[18] Higgins, E. L., & Raskind, M. H. (1995). Compensatory effectiveness of speech recognition on the written composition performance of postsecondary students with learning disabilities. *Learning Disability Quarterly*, 18 (Spring), pp. 159-174.

[19] Raskind, M. H., & Higgins, E. L. (1995). Effects of speech synthesis on the proofreading efficiency of postsecondary students with learning disabilities. *Learning Disability Quarterly*, 18 (Spring), pp. 141-158.

[20] MacArthur, C. A., Ferretti, R. P., Okolo, C. M., & Cavalier, A. R. (2001). Technology applications for students with literacy problems: A critical review. *Elementary School Journal*, 101 (3), pp. 273-301.

[21] Allen, D. N., Goldstein, G., Heyman, R. A., & Rondinelli, T. (1998). Teaching memory strategies to persons with multiple sclerosis. *Journal of Rehabilitation Research and Development*, 35(4), pp.405-410.

[22] Zanetti, O., et al., (2000). Utilizing and electronic memory aid with Alzheimer's disease patients. A study of feasibility. Sixth International Stockholm/Springfield Symposium on Advances in Alzheimer Therapy, Stockholm, Sweden.

[23] Strickland, D., Marcus, L., Mesibov, G. B., & Hogan, K. (1996). Brief report: Two case studies using virtual reality as a learning tool for autistic children. *Journal of Autism and Developmental Disorders*, 26 (6), pp. 651-660.

[24] Imamura, K. N., Wiess, T., & Parham, D. (1990). The effects of hug machine usage on behavioral organization of children with autism and autistic-like characteristics. *Sensory Integration Quarterly*, 27, pp. 1-5

[25] Werry, I., Dautenhahn, K., Ogden, B., & Harwin, W. (2001). Can social interaction skills be taught by social agent? The role of a robotic mediator in autism therapy. Proceedings CT2001, The Fourth International Conference on Cognitive Technology: Instruments of mind, Warwick, UK.

[26] LoPresti, E. F., Friedman, M. B., & Hages, D. (1997). Electronic vocational aid for people with cognitive disabilities. Proceedings of the Rehabilitation Engineering Society of North America (RESNA) (pp. 514-526). Arlington, VA: RESNA Press.

VII. Autores

C. Mariana Sosa Terrazas alumna del segundo año de la Licenciatura de Matemáticas Aplicadas en la Facultad de Ingeniería de la UAQ. Sus intereses son la programación y la optimización.

TSU. José Sergio Orduña Díaz obtuvo su título de TSU en el 2006 actualmente es alumno de la carrera de Ingeniería en Automatización en la Facultad de Ingeniería de la UAQ. Sus

principales interés son la programación de sistemas electrónicos.

M. en C. José Luis Avendaño Juárez alumno de la carrera de electrónica de la UAM-Azcapotzalco, obtuvo su título de Maestría en Ciencias en la Sección de Posgrado de la Facultad de Ingeniería de la UAQ. Sus principales intereses son la electrónica y la instrumentación.

Dr. Gilberto Herrera Ruiz Obtuvo su título de Doctorado en Ingeniería con especialidad en Automatización en 1992, profesor de tiempo completo de la Facultad de Ingeniería de la UAQ desde 1997. Sus principales intereses son la Automatización y los Sistemas Electrónicos.

Propuesta de Sistema de Conteo Electrónico de Personas que Utilizan el Servicio de Sanitarios de la Central Internacional de Autobuses de Culiacán

***Eduardo González Román, **Oscar Martínez López
Mario García Irigoyen, *Juan Manuel Mejía Camacho, *****Canek Portillo Jiménez**

Universidad Autónoma de Sinaloa (UAS). Facultad de Ingeniería. Ingeniería en Procesos Industriales (IPI).
Av. de la Américas y Blvd. Universitarios S/N, Culiacán, Sinaloa. (667) 7134043.

* lalo_erg_1587@hotmail.com

** martinez_lopez_89@hotmail.com

**** mejiacamacho@hotmail.com

***** canekportillo@hotmail.com

***IPI / Centro de Capacitación para el Trabajo Industrial # 32 (CECATI 32).

Fernando Cuén 1134, Ruiz Cortínez, Culiacán, Sinaloa. (667) 7454506. irigoyenmario@hotmail.com.

Resumen

Se realiza propuesta de un sistema de conteo electrónico de personas, para el control de la cantidad de usuarios de los sanitarios de la Central Internacional de Autobuses de Culiacán. Se ha fabricado un prototipo utilizando el microcontrolador PIC16F877A y otros elementos electrónicos y mecánicos de bajo costo, que sirven de apoyo al sistema mecánico actualmente instalado. Entre los datos que despliega el sistema a través de una pantalla LCD de 2 x 16 se tiene: el número de personas que ingresan a utilizar el servicio sanitario, el nombre del establecimiento donde será colocado el dispositivo, además de que permite realizar cortes de caja según se requiera.

Palabras Clave: Contador de personas, Microcontrolador PIC16F877A, Sensor IR, Sistema de conteo electrónico.

I. Introducción

La Central Internacional de Autobuses de Culiacán cuenta con servicios sanitarios, cuya utilización tiene un costo que debe pagar el usuario. Los responsables de la concesión actualmente han instalado un sistema mecánico de conteo del número de personas que utilizan el servicio (ver Fig. 1). Con esta información se realizan los cortes de caja requeridos con apoyo de una persona auxiliar.

A partir de lo anterior, surge un acercamiento entre los concesionarios de los servicios sanitarios de la terminal de autobuses y la Facultad de Ingeniería de la UAS, a través de la Licenciatura en Ingeniería en Procesos Industriales (IPI), donde se plantea el requerimiento de un apoyo electrónico al sistema

mecánico existente, con el objetivo de dar a su sistema mayor certidumbre y capacidad operativa.

Actualmente existen sistemas de conteo electrónico disponibles en el mercado [1, 2, 3], con diversos grados de complejidad de acuerdo a la aplicación. Se considera que dadas las características que se presentan en este escenario, tales sistemas no son adecuados o tienen tal capacidad, que serían subutilizados. Además hay que tomar en cuenta el costo elevado.

Por otro lado, en la Licenciatura de Ingeniería en Procesos Industriales los estudiantes llevan como parte de su plan de estudio, un curso de microcontroladores, donde se les pide como requisito la realización de un proyecto final. Es por ello que se aprovecha esta oportunidad y se

asigna este proyecto a un par de estudiantes de IPI, quienes con cierta guía de profesores, van desarrollándolo hasta realizar un prototipo como propuesta para la solución a la problemática particular planteada. Se desarrolla en sí un sistema electrónico de conteo sencillo y de bajo costo que complementaría al sistema mecánico existente. La principal contribución de este trabajo es el enfoque de aplicación real a lo visto en el curso del programa educativo. Esto influye en la motivación y creatividad de los estudiantes, contribuyendo a un mejor aprendizaje.



Fig. 1. Sistema mecánico de conteo actual

Las siguientes secciones del documento se distribuyen de la siguiente manera: en la sección II se presenta el diseño del sistema; el principio de operación lo tratamos en la sección III; en la sección IV se tiene lo relativo a la Programación del PIC; y los Resultados en la sección V. Por último se tienen las conclusiones.

II. Diseño del sistema

En esta sección se aborda lo referente al diseño del sistema de conteo. Se subdividió en dos partes: mecánica y electrónica.

La parte mecánica se refiere a la caja protectora (ver Figuras 2 y 3) compuesta de una tapadera con sistema de seguridad bajo llave y aloja en su interior la circuitería electrónica necesaria. Las dimensiones de la caja son de 25 cm x 18.8 cm x 8 cm, en su parte frontal interna tiene espacios para la visualización de una pantalla LCD y

botones auxiliares de reinicio e iluminación. La parte posterior cuenta con una entrada para la alimentación de energía eléctrica y para la conexión con el sensor infrarrojo (IR). La parte electrónica del sistema está formada por elementos que se alojan dentro de la caja y un par de componentes colocados fuera de ella, pero en conexión con los elementos del interior.

Dentro se coloca una placa donde se montan los componentes electrónicos. Por fuera de la caja se tienen al conjunto sensores IR y las superficies reflectoras.

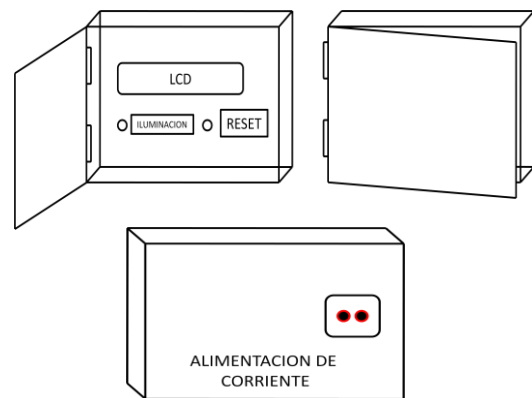


Fig. 2. Diseño inicial de la caja protectora. Parte frontal (interna y externa) y posterior



Fig. 3. Caja protectora abierta (arriba: acercamiento parte superior).

En la fabricación del prototipo se utilizan diversos materiales, herramientas y componentes electrónicos, los cuales presentamos en las

Tablas 1 y 2. Se utiliza el microcontrolador PIC16F877A [4, 5] por su versatilidad, y facilidad en la comunicación con la pantalla LCD 2x16. En la Fig. 5 se presenta el diagrama electrónico completo del subsistema.

Cantidad	Herramienta y materiales
1	Cautín tipo estación 45W
1	Rollo de soldadura 0.8mm 250 gr
1	Limpiador para soldadura
1	Programador p/pics
1	Caja de seguridad (25x18.8x8)
1	Moldura alfombra (22mmx0.91mts).
1	Canaleta 4pz. 1.10mts.

Tabla 1. Herramientas y materiales.

Cantidad	Componente
1	Microcontrolador 16F877A
4	Cable multiconductor calibre 22 (m).
4	Cable calibre 10 en mts.
2	Resistencias 100, 10k
2	Condensadores electrolíticos 33pf
1	Cristales de cuarzo 4Mhz.
2	Push button N.O
1	Fuente de poder ATX 400w
1	Potenciometro precisión multi. 10k
1	Placa fenólica 10X20
1	LCD 16X2 stn backlight
1	Sensor IR
1	Socket 40 pins 16F877A
1	Convertidor CD-CD 12-5v

Tabla 2. Componentes electrónicos

El circuito se simuló en el programa Proteus [6], para implementarlo posteriormente en tarjeta impresa. Proteus tiene la característica de simular circuitos electrónicos complejos integrando varios tipos de desarrollos con microcontroladores.

III. Principio de funcionamiento

El funcionamiento está basado en interrupciones. Mediante el uso de un sensor se envía un haz infrarrojo hacia una superficie reflejante.

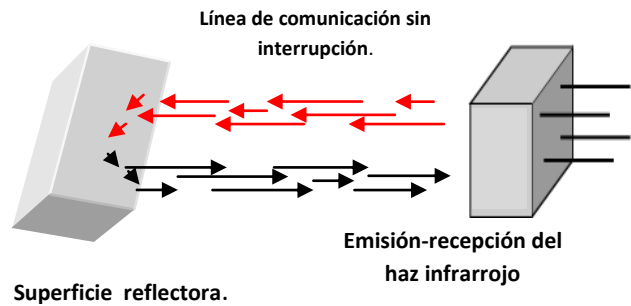


Fig. 4. Principio de funcionamiento (fase: en espera de interrupción).

Cuando se tiene el haz continuo reflejándose, el sensor se encuentra en la fase de espera para enviar la señal de interrupción al microcontrolador (ver Fig. 4).

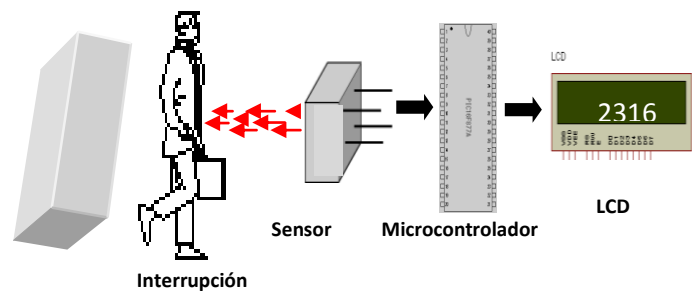


Fig.5. Principio de funcionamiento (fase: ocurre interrupción).

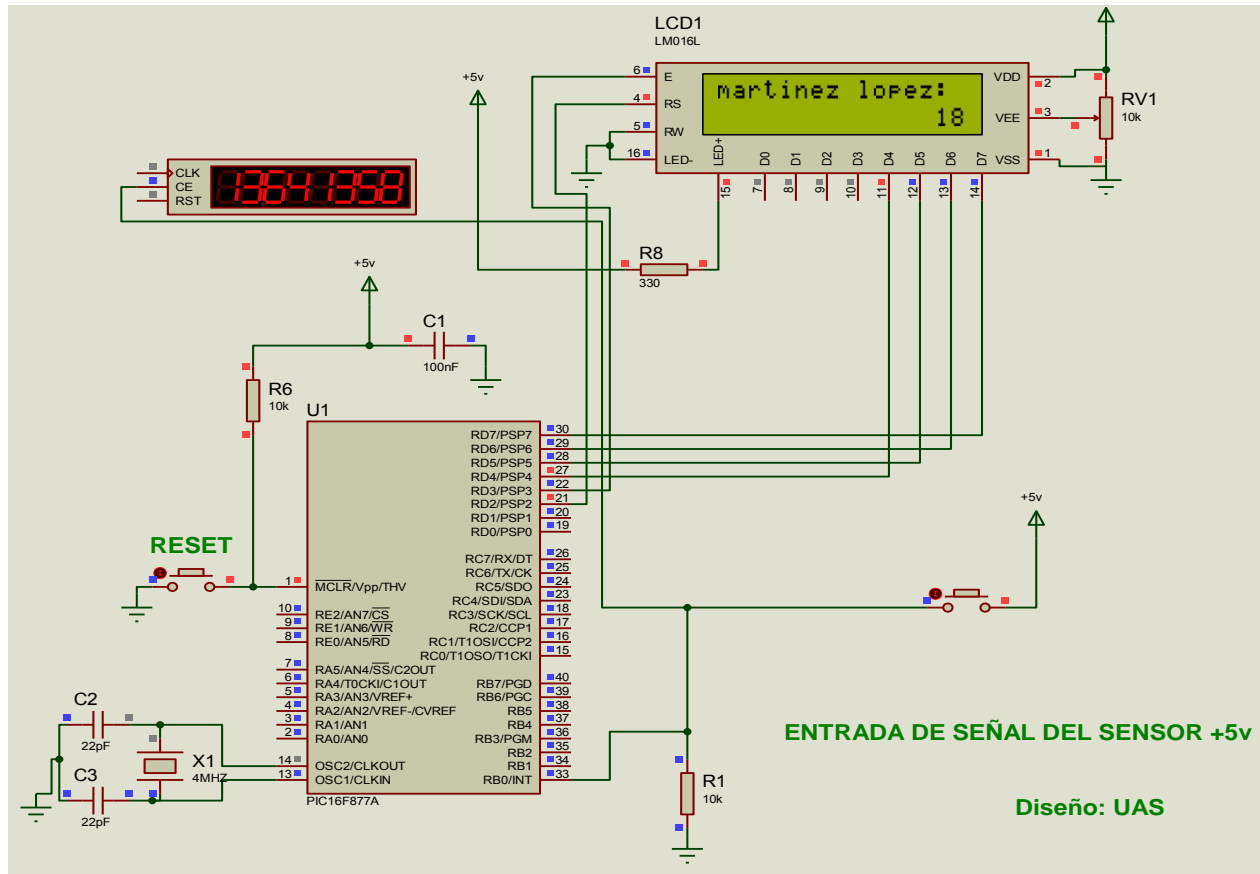


Fig. 6 Diagrama electrónico

Cuando ocurre una interrupción entre el haz y la superficie debido al paso de una persona (ver Fig. 5), la interrupción es detectada por el sensor que envía una señal al microcontrolador, para que éste a su vez después de un realizar cierto procesamiento, se comunice con la pantalla LCD indicándole que muestre la información y llevando un conteo en orden ascendente cada vez que se detecte una nueva interrupción.

Al microcontrolador se le ha establecido un límite de conteo de hasta 1,000,000, con un retardo que puede ser manipulado desde software para ajustarlo al tiempo promedio de cruce peatonal.

IV. Programación del PIC16F877A

El programa que se le introdujo al microcontrolador permite conocer el número de personas que han usado el servicio y devolver

este número a cero en cada corte de caja al presionar un botón.

Se utilizó como sistema de programación para el microcontrolador la plataforma de MikroC [7], que es un software que utiliza lenguaje de programación en C para microcontroladores, transformando el código a lenguaje ensamblador, que es el utilizado por éstos para abrir, cerrar o comunicar puertos con un fin específico.

Se utilizó el PortD como salida de comunicación con la pantalla LCD y el PortB.B0 como la entrada de la interrupción para el conteo de personas; la entrada MCLR se utiliza como el reset, que permite el reinicio del conteo.

El PortD se utiliza para la comunicación del PIC con la LCD, enviándole los datos de cuántas personas han ingresado hasta ese momento, para que ésta la despliegue en la pantalla. Por otro lado, por el PortB.B0 se detecta la interrupción. El sensor infrarrojo le manda una señal de 0 volts

si no se ha presentado ninguna interrupción en el haz de infrarrojo, y de 5 volts cuando una persona ha pasado a través del haz, indicándole al PIC que incremente el conteo. En la figura 7 se observa el diagrama de flujo del programa.

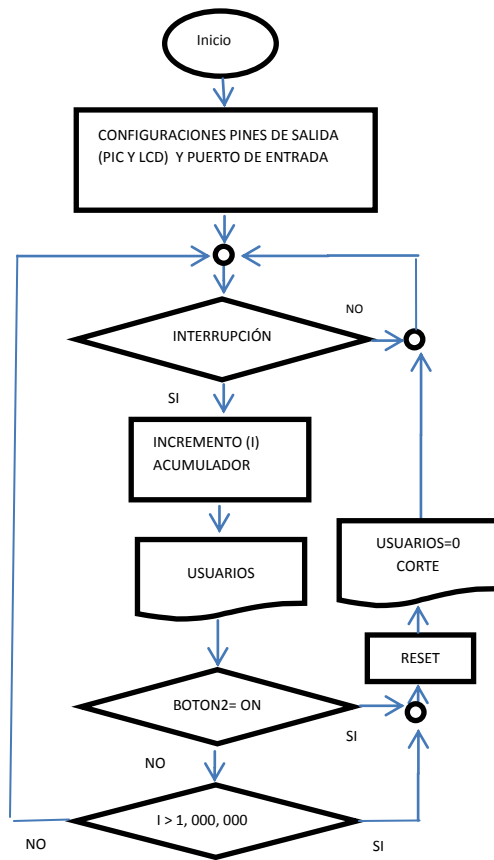


Fig.7. Diagrama de flujo

V. Resultados

En esta sección se presenta el prototipo. Se ha probado su funcionalidad en una exposición de proyectos internos de la carrera de IPI (ver Fig. 8).

Se simulan dos paredes a través de dos columnas delgadas de madera, donde se colocan un par de sensores IR y un par de superficies reflejantes. Para su implementación se propone el escenario que aparece en la Fig. 9, donde se muestra cómo

podría ir colocado el conjunto de sensores y superficies reflectoras.



Fig.8. Prototipo en funcionamiento

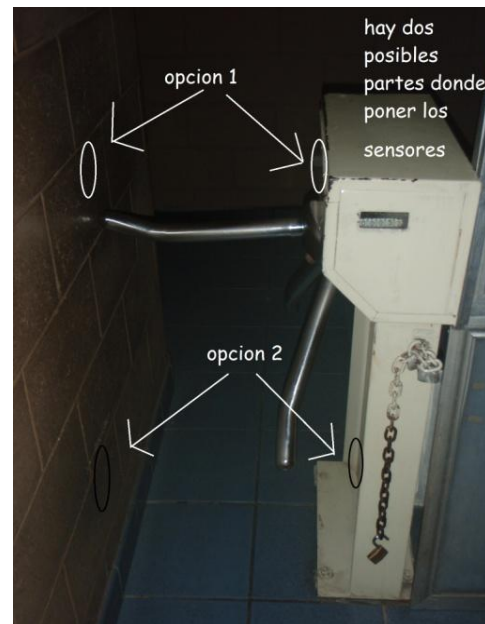


Fig.9. Opciones para la instalación de los sensores.

VI. Conclusiones

A partir de una problemática real se planteó una posible solución, desarrollándose un prototipo de un sistema electrónico de conteo, aplicando los conocimientos teóricos adquiridos en la asignatura de microcontroladores de la carrera de Ingeniería en Procesos Industriales.

El sistema de conteo se propone para auxiliar, como complemento y para dar mayor certidumbre en el control del número de personas que utilizan el servicio de los sanitarios en la central de autobuses.

El sistema es compacto, ligero y relativamente sencillo. Entre las posibles mejoras se tiene la inclusión de un dispositivo transmisor/receptor para que permita el monitoreo a distancia.

Otra aplicación posible es el control del número de pasajeros que utilizan el servicio de transporte urbano de la ciudad de Culiacán.

VII. Referencias

- [1] Contador comercial de Pogen. <http://www.pogen.mx>. Mayo del 2011.
- [2]. Contador comercial de Contarpersonas. <http://www.contarpersonas.com>. Mayo de 2011.
- [3]. Contador comercial de Tiendatec. <http://www.tiendatec.com.mx>. Mayo de 2011.
- [4]. Hojas de datos técnicos PIC16F877. Microship Technology Inc. Estados Unidos. 2001.
- [5]. J. M. Angulo Usategui. Microcontroladores DSPIC, diseño práctico de aplicaciones. 1era Ed. 2006. Paraninfo. Madrid. 768 pp.
- [6]. E. García Breijo. Compilador C CSS y Simulador PROTEUS. 1ra Ed. 2008. Marcombo. Barcelona. 344 pp.
- [7]. Guía de Referencia mikroC. MokrioElektronica. Bulgaria. 2009.

VIII. autores

Eduardo González Román es estudiante de la Licenciatura en Ingeniería en Procesos Industriales, en la Universidad Autónoma de Sinaloa.

Oscar Martínez López es estudiante de la Licenciatura en Ingeniería en Procesos Industriales, en la Universidad Autónoma de Sinaloa.

Mario García Irigoyen realizó estudios de Ingeniería Electrónica en la Universidad Autónoma de Baja California. Actualmente es profesor en la Licenciatura en Ingeniería en Procesos Industriales. Colabora también en Centro de Capacitación para el Trabajo Industrial # 32.

Juan Manuel Mejía Camacho realizó estudios de Maestría en Ingeniería en la UABC; cursó estudios de Ingeniería en electrónica en el Instituto Tecnológico de Culiacán. Es profesor en la Licenciatura en Ingeniería en Procesos Industriales.

Canek Portillo Jiménez realizó estudios de Maestría en Ciencias el CICESE; cursó estudios de Ingeniería en electrónica en el Instituto Tecnológico de Culiacán. Es profesor en la Licenciatura en Ingeniería en Procesos Industriales.

Control de una Red de Iluminación a través de un Sistema CAN Embebido

Raúl Alvarado Lara, José L. Hurtado Chávez, Elías Rodríguez Segura, José J. Díaz Carmona, Alfredo Padilla, Nimrod Vázquez

Instituto Tecnológico de Celaya. Departamento de Ingeniería Electrónica. Av. Tecnológico y A. G. Cubas S/N, Col. Alfredo V. Bonfil, Celaya, Gto. (461)6117575 ext. 384. elias@itc.mx

Resumen

El siguiente artículo describe el diseño e implementación de un Sistema Electrónico Embebido para el control de una red de iluminación basada en CAN. El sistema propuesto, tiene como finalidad el control de una red de iluminación, el cual incluye el controlador CAN, un balastro electrónico para lámpara fluorescente, un controlador para LED's de alta brillantez y un sensor de intensidad luminosa. Dentro de las funciones que desempeña dentro de los dispositivos de iluminación se encuentran: el encendido y el apagado de los dispositivos así como el control de intensidad luminosa de los mismos, además, cuenta con la función de detección de fallos para saber de manera rápida si algún dispositivo dentro de la red de iluminación ha fallado, haciendo mención al tipo de dispositivo e identificador del mismo. Por último se muestran resultados experimentales del sistema completo junto con la interfase de usuario.

Palabras Clave: Balastro Electrónico, CAN, LED de alta brillantez, Control embebido.

I. Introducción

En los últimos 15 años, la industria de la iluminación ha proveído de componentes para aplicaciones del control de la iluminación basados en sistemas analógicos. También ha definido un estándar para la comunicación digital entre los componentes individuales de un sistema de iluminación, el protocolo DALI (*Digital Addressable Lighting Interface*) [1]. Con la adopción del protocolo DALI por parte de la industria, se tiene la primera manifestación de la transición de analógico a digital en la comunicación entre los dispositivos de control, lográndose establecer una estructura de un sistema de control de iluminación genérico, el cual consta de un controlador, sensores y actuadores, como se muestra en la figura 1. El controlador por lo general es una computadora personal (PC) o un panel digital, en él se realiza las diferentes funciones de control y monitoreo del sistema.

En la literatura existen diseños reportados sobre balastos electrónicos digitales con control de la intensidad luminosa, tales como [2] y [3]. En

ambas referencias se propone un sistema de iluminación controlado digitalmente basado en el protocolo DALI por medio de una computadora personal.

Los sensores se utilizan para lograr un control automático del ambiente, entre los sensores más utilizados se tienen los de presencia y los de intensidad luminosa. El actuador por lo general es un balastro electrónico digital para el manejo de lámparas fluorescentes.

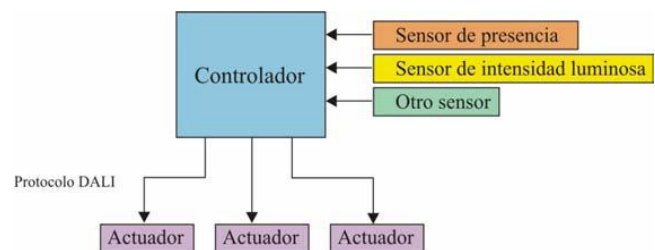


Fig. 1. Estructura de un sistema de control de iluminación genérico.

La comunicación serial digital resuelve el problema del cableado presente en el control analógico debido a que en el control digital todos

los dispositivos comparten un mismo cable, simplificando el cableado y la instalación, además de agregar una mayor funcionalidad al sistema de iluminación con la ayuda del software de la aplicación. Sin embargo, el protocolo DALI no fue creado como un protocolo de comunicación entre el controlador y los diferentes sensores del sistema de iluminación, esto es, DALI es un protocolo para la comunicación entre el controlador y el actuador [4]. Por lo que el transporte de la información de los distintos sensores del sistema de iluminación al controlador se realiza por medio de otro protocolo de comunicación mediante un medio físico distinto.

Debido a la falta de integración presente con el protocolo DALI en un sistema de control de iluminación se requiere de la introducción de otro protocolo de comunicación capaz de integrar los distintos actuadores y sensores del sistema. Respecto al actuador, éste no debería estar limitado al manejo de lámparas fluorescentes, sino también integrar otras tecnologías de iluminación, tales como leds de alta brillantes y lámparas de alta intensidad de descarga.

La selección del protocolo de comunicación dependerá de las prestaciones del protocolo tales como: capacidad de direccionamiento lógico, cantidad de datos, nivel de seguridad, robustez, entre otras características importantes.

En [5] se presenta una red Touch-Dim que es posible conectarla a un sistema DALI con el objetivo de tener un control accesible del sistema de iluminación desde cualquier punto de la red. En el trabajo presentado en [6] se propone un método de diseño de un sistema a control remoto para iluminación LED basado en CAN.

En este trabajo se propone la implementación de un sistema de iluminación basado en el protocolo de comunicación bus CAN. Se propone el diseño y la implementación balastro electrónico digital con control de la intensidad luminosa para operar una lámpara fluorescente de 28 watts, un controlador para operar un arreglo en serie de 6

leds de alta brillantez (HBLEDS), un sensor de intensidad luminosa, y por último el controlador CAN para establecer la comunicación con todos los dispositivos. La estructura del sistema de iluminación propuesto se presenta en la figura 2.

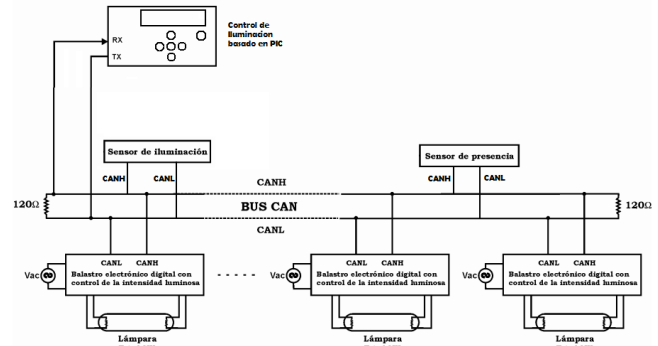


Fig. 2. Sistema de iluminación propuesto.

II. Desarrollo

El protocolo bus CAN es quien ofrece una mayor capacidad de direccionamiento lógico y permite la mayor longitud en el cable del bus. Con esta gran capacidad de direccionamiento lógico disponible se pueden asignar nodos como actuadores, cuyas funciones son las de recibir las instrucciones codificadas de las distintas funciones por parte del controlador a través del cable del bus. Estos nodos, los cuales estarán integrados a cada balastro, decodificarán dichas instrucciones y efectuarán la acción de control solicitada en el balastro. También es posible asignar nodos como sensores, los cuales podrán codificar y transmitir la información de su correspondiente variable de control (nivel de iluminación, presencia, etc.) por el mismo cable del bus al controlador. De esta manera se pretende integrar tanto actuadores como sensores al sistema de iluminación mediante un mismo protocolo de comunicación y un mismo medio físico. Los 8 bytes de datos que son posibles de transmitir en un mensaje mediante el protocolo bus CAN son suficientes para llegar a cabo la codificación de las distintas funciones del control en un sistema de iluminación, tales como: control de la intensidad luminosa, encendido y apagado de lámparas, etc.

Características del bus CAN

1. La comunicación está **basada en mensajes** y no en direcciones, todos los nodos CAN son capaces de transmitir y recibir datos y varios pueden acceder al bus de datos simultáneamente. Un nodo emisor envía el mensaje a todos los nodos de la red, cada nodo, según el identificador del mensaje, lo filtra y decide si debe procesarlo inmediatamente o descartarlo.

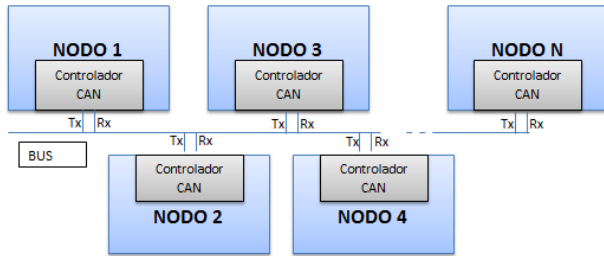


Fig. 3. Red CAN

Para visualizar mejor el funcionamiento de la comunicación CAN, se puede ver las conexiones ya mencionadas en la figura 3, con lo cual se puede lograr un mejor entendimiento de su funcionamiento.

2. **Flexibilidad:** Se pueden añadir nodos adicionales al bus sin necesidad de reprogramar todo el sistema para que se reconozcan los nuevos nodos

3. Gran **fiabilidad** en la transmisión:

- Detecta errores, los señala, envía mensaje de error y reenviará el mensaje corrupto una vez que el bus vuelva a estar activo.

- Distingue entre errores temporales y errores permanentes desconectando automáticamente los nodos defectuosos.

- Puede operar en ambientes con condiciones extremas de ruido e interferencias.

4. La velocidad de transmisión depende de la longitud de la red. Cada red puede alcanzar los 5000 metros de longitud. En la tabla 1 se puede apreciar las diferentes velocidades de transferencia que se pueden lograr, dependiendo de la longitud del bus.

Capa Física del bus CAN

La capa física es responsable de la transferencia de bits entre los distintos nodos que componen la red. Define aspectos como niveles de señal, codificación, sincronización y tiempos en que los bits se transfieren al bus. Como se muestra en la tabla 2, se especifican tres niveles de la capa física con distintas funciones cada uno. Los distintos niveles son:

- Señalamiento Físico: El Controlador CAN (En este caso, el implementado en el PIC18f258)
- Medio Físico Adjunto: Es la parte que se encarga de adecuar la señal (En este caso, el Transceiver MCP2551)
- Interfaz dependiente del Medio: (En este caso, el Transceiver MCP2551)

Tasa de Transferencia	Longitud del bus
1 Mbit/s	30 m
800 Kbit/s	50 m
500 Kbit/s	100 m
250 Kbit/s	250 m
125 Kbit/s	500 m
62,5 Kbit/s	1000 m
20 Kbit/s	2500 m
10 Kbit/s	5000 m

Tabla 1. Relación Tasa de Transferencia – Longitud del Bus

Especificación	Capa OSI		Implementación
Especificado por el Diseñador del sistema	Capa de Aplicación		
Estándar ISO 11898	Capa de Datos	Control Lógico	Controlador CAN Ejem. PCx82C200
		Control de Acceso Medio	
	Capa Física	Señalización Física	Transceiver CAN
		Medio Físico Adjunto	
		Interfaz dependiente del Medio	
	Medio de Transmisión		

Tabla 2. Especificación del BUS CAN Según ISO 11898

Los nodos conectados al bus interpretan dos niveles lógicos denominados:

Dominante: la tensión diferencial (CAN_H - CAN_L) es del orden de 2.0 V con CAN_H = 3.5V y CAN_L = 1.5V

Recesivo: la tensión diferencial (CAN_H - CAN_L) es del orden de 0V con CAN_H = CAN_L = 2.5V [5,6].

Hardware

En la figura 4 se muestra el diagrama de la interfaz del usuario o Controlador CAN. La interfaz del usuario está constituida por un teclado y una pantalla a través de los cuales es posible establecer el encendido, el apagado y los parámetros de intensidad luminosa de cualquiera de los nodos de la red. También permite la lectura de los sensores conectados a la red, y un sistema de control automático en el que por medio de las lecturas del sensor realiza el control de los dispositivos de iluminación.

El transistor 2N2222 se utiliza para el control del backlight de la LCD, ya que dentro del programa principal se implementó una subrutina para que este solo se encienda cuando el teclado sea activado, y se apaga después de 5 segundos de no detectar cambios en el mismo. Esto se hizo para reducir el consumo de energía del circuito.

Software

La figura 5 muestra como se despliega la información en el display LCD, se muestra la distribución de las funciones del programa creado para el control del sistema de iluminación, ordenadas de manera jerárquica, comienza en el menú principal, y va descendiendo en niveles hasta llegar a las funciones específicas que se pueden establecer.

El programa fue enteramente hecho en lenguaje C. se utilizó para la programación del mismo el IDE PCWHD de CCS que nos ofrecía ciertas ventajas como un lenguaje C orientado a microcontroladores, con lo que ciertas acciones del sistema eran programadas de una manera más adecuada que en un lenguaje C convencional.

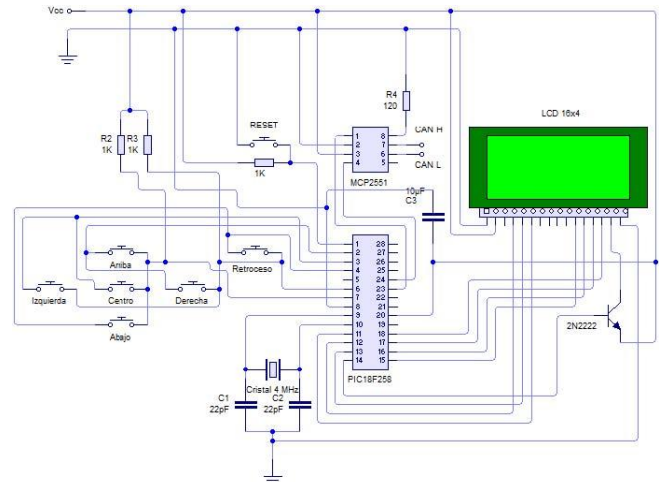


Fig. 4. Diagrama del controlador CAN.

El menú principal está organizado por 3 funciones: Acciones, que es el submenú donde se pueden ejecutar todas las funciones relativas al sistema de iluminación, Configuración de Dispositivos, es otro submenú, donde se lleva toda la organización de los dispositivos instalados en el sistema, así como la instalación de nuevos componentes al sistema. Y como último submenú se encuentra, Lista de Dispositivos, el cual tiene la función de mostrar al usuario todos los dispositivos que se encuentran actualmente instalados en el sistema.



Fig. 5. Funciones del programa como aparecen en la pantalla del controlador CAN

Al ingresar al submenú Acciones, se puede observar que hay 3 funciones que se pueden realizar desde el mismo: ON/OFF que es la función de Encendido o Apagado de un dispositivo tanto de forma individual, como de forma grupal. Al seleccionar la función ON/OFF, el programa pasa a una pantalla donde muestra al

usuario que tipo de acción puede realizar, encender o apagar alguno de los dispositivos dados de alta en el sistema.

Esta función puede ser aplicada tanto a Balastros con lámpara del tipo T5 de 28 watts como a un controlador con luz LED.

A continuación se encuentra la función de Intensidad, que tiene la capacidad de poder modificar la intensidad luminosa de tanto una lámpara como de un controlador de LEDS.

Dentro del sub-menú de Sensores, se encuentran 2 funciones de trabajo: La lectura de Sensor, y el control automático. La función de lectura de sensor permite al usuario probar los sensores instalados en el sistema, con lo cual al seleccionar esta función, el microcontrolador le hace un requerimiento de lectura al sensor, y en la pantalla se muestra el dato del sensor leído.

Esta función es útil, a la hora de saber si un sensor está trabajando de forma adecuada y está haciendo lecturas de la luz ambiental de forma correcta. Que dicho de otra manera nos ofrece un monitoreo de los sensores del sistema.

La función de Control Automático, es hacer un control de uno o varios dispositivos de iluminación instalados en el sistema mediante las lecturas proporcionadas por un sensor de iluminación deseado. Logrando con esto un control del sistema independiente de la interacción del usuario.

Al seleccionar esta función dentro del sub-menu de sensores, el sistema nos muestra como primer pantalla las instrucciones para el funcionamiento de la misma, que consiste en seleccionar primeramente el sensor del cual se tomaran las lecturas para controlar el o los dispositivos seleccionados, después de esto se seleccionaran el o los dispositivos a controlar.

Como paso final, se seleccionara el tiempo de muestreo del control, que se puede seleccionar desde 0.1 minutos (6 segundos) hasta 9.9 min (594 segundos), que es el tiempo que transcurrirá

entre lectura y lectura del sensor para hacer una actualización del nivel de intensidad actual de los dispositivos de iluminación seleccionados.

Dentro del sub-menú de Configuración de Dispositivos, se encuentra 3 opciones de configuración, que son: Agregar dispositivo, Borrar dispositivo, y Borrar Todos.

La función de Agregar Dispositivo, permite instalar un dispositivo nuevo en el sistema. En primera instancia se pregunta el identificador que se asignara al nuevo dispositivo, posteriormente pregunta el tipo de dispositivo (iluminación o sensor) y finalmente si se asigna a un grupo o no. Es importante mencionar que el controlador CAN tiene la máxima prioridad (ID=000). La asignación de identificadores esta abierta al usuario, ya que él es quien selecciona el identificador de acuerdo a la prioridad que desea en u cierto nodo.

La función de Borrar dispositivo, nos permite desinstalar algún dispositivo presente en el sistema, en caso de que esté haya tenido un fallo, o ya no sea necesario en el mismo. La última opción del sub-menú, que es BORRAR TODOS, realiza un borrado de toda la memoria EEPROM interna del microcontrolador con lo que el sistema sabe que todos los dispositivos han sido eliminados y esta listo para la instalación de nuevos dispositivos.

La función de Lista de dispositivos, dentro del menú principal, nos permite observar en la pantalla del sistema, los dispositivos instalados actualmente, y al seleccionar dispositivos específicos, nos permite saber algunos datos que nos podrían ser de utilidad a la hora de usar el sistema, estos son: Identificador, Intensidad (Dispositivos de iluminación), Ultima Lectura (Sensores), Grupo al que pertenecen.

Finalmente, se encuentra la función de Detección de Fallos del sistema. Cuando alguna de las lámparas del sistema esta en baja emisión o se encuentra rota, el microcontrolador dentro del

balastro electrónico manda un mensaje de interrupción avisando de la falla del dispositivo.

Cuando el sistema recibe la interrupción, pasa a la rutina de detección de fallos, y muestra en la pantalla un mensaje de advertencia, avisando el tipo de dispositivo que se encuentra fallando y el identificador del mismo. Cabe mencionar que solo se cubren fallas en los balastos electrónicos y no fallas de comunicación.

En la figura 6 se muestra el controlador CAN implementado.



Fig. 6. Controlador CAN implementado.

El balastro electrónico se compone principalmente de 3 etapas: etapa digital, etapa de corrección del factor de potencia y etapa del inversor resonante. La etapa digital es la encargada de la gestión y enlace de la información a través del cable del bus, será la responsable de recibir y transmitir la información, decodificarla y efectuar la acción solicitada.

La etapa de corrección del factor de potencia hará que el balastro demande de la línea eléctrica una corriente senoidal y en fase con el voltaje de entrada, lo cual permitirá que el balastro presente una baja distorsión armónica y tenga un alto factor de potencia. Con lo anterior, el balastro satisface la norma EN61000-3-2 Clase C, que

regula la cantidad de armónicos que un equipo puede inyectar a la línea.

La última etapa, inversor resonante, se encarga del manejo de la lámpara y se diseñó para que el balastro sea capaz de operar una lámpara fluorescente del tipo lineal T5 de 28 watts. Debe de operar a la lámpara en optimas condiciones del tal forma de no acortar la vida útil de la lámpara debido a un mal manejo, así como de disponer de un proporcionar un medio para poder efectuar el control de la intensidad luminosa. El diagrama del balastro electrónico y su implementación se muestra en las figuras 7 y 8.

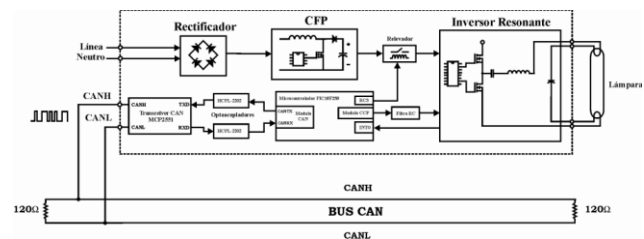


Fig. 7. Diagrama del Balastro Electrónico



Fig. 8. Balastro electrónico implementado.

La topología seleccionada para operar como controlador de leds de alta brillantez es un convertidor reductor controlado mediante el circuito HV9910 del fabricante Supertex.

Éste circuito permite la operación eficiente de leds de alta brillantez conmutando a frecuencia fija el mosfet del convertidor reductor. En la figura 9 se muestra la implementación del controlador.

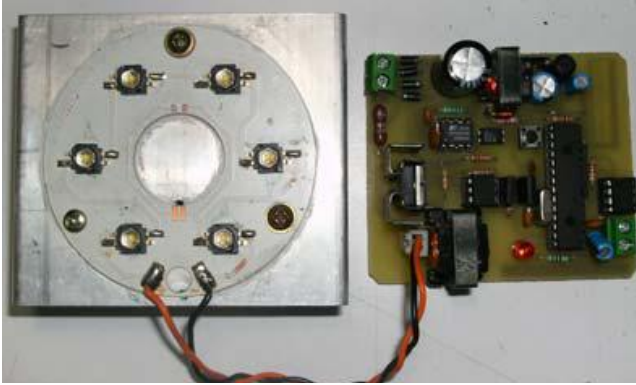


Fig. 9. Controlador para LED de alta brillantez.

III. Resultados

Para propósitos de prueba del sistema de iluminación basado en bus CAN se implementaron tres balastos electrónicos para operar una lámpara fluorescente cada una de ellos, un balastro electrónico para leds de alta brillantez, controlador CAN y un sensor de intensidad luminosa.

En la figura 11 se muestra el funcionamiento del sistema completo con el sensor de intensidad luminosa para una lámpara fluorescente T5 y un arreglo de 6 HBLED's.

Se puede apreciar que cuando el sensor de intensidad luminosa es obstruido (simulando la noche) ó recibe poca iluminación, el sistema cambia gradualmente la intensidad de los dispositivos de iluminación desde un estado de apagado hasta la intensidad máxima, mientras que cuando el sensor detecta que la iluminación del ambiente aumenta, hace la operación inversa.

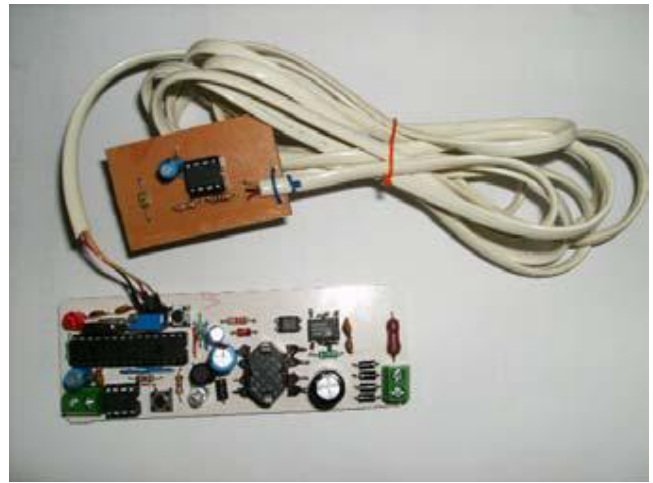


Fig. 10. Circuito para el sensor de intensidad luminosa.

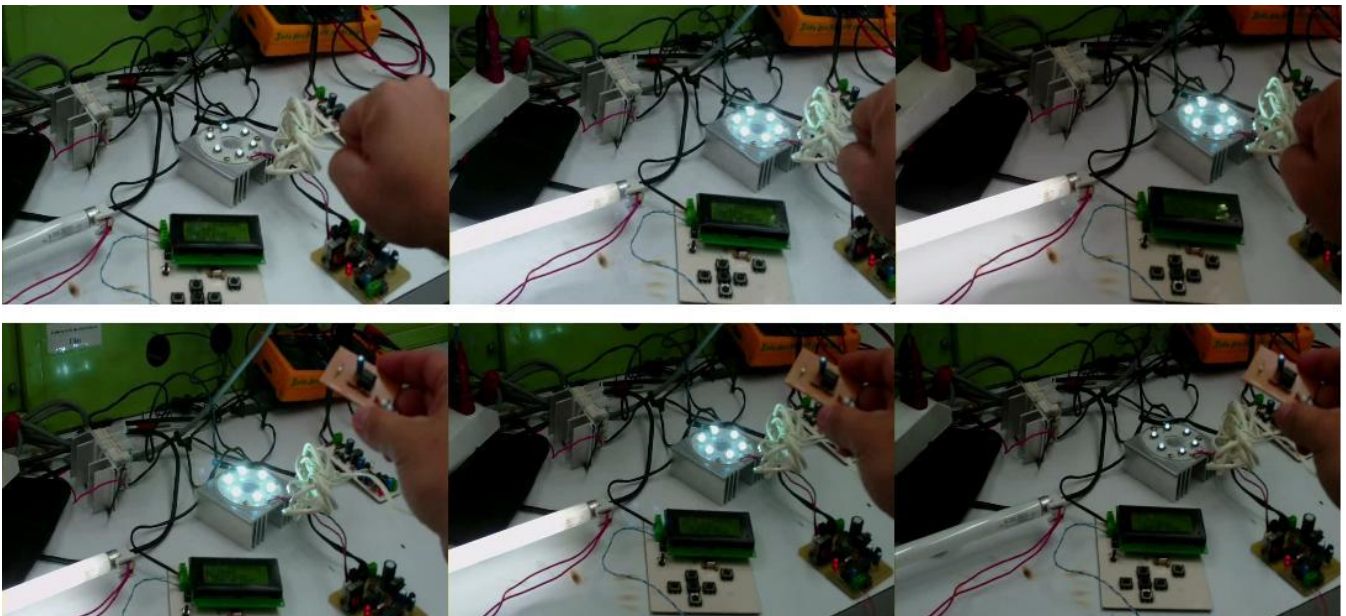


Fig. 11. Cambio de intensidad luminosa (Menor a Mayor, Imagen Superior), (Mayor a Menor, Imagen Inferior)

Para hacer pruebas de funcionamiento se escoge el modo manual y una vez que todo este operando correctamente, se pasa a modo automático, el cual es controlado por los sensores de intensidad luminosa y de presencia.

El sistema se mantendrá en este modo de operación hasta que el usuario cambie a modo manual para agregar dispositivos o hacer pruebas a los sensores. En caso de pérdida de energía el sistema no requiere de una batería de respaldo para guardar la configuración, ya que esta se queda grabada en memoria y cuando se reestablece la energía el sistema arranca con la configuración establecida de forma automática

IV. Conclusiones

El sistema de iluminación automatizado incluye funciones de variación de intensidad luminosa, control de apagado y encendido, y detección de fallo de lámpara. Con este sistema es posible controlar hasta un máximo de 2047 balastos o controladores de LED's, lo que lo hace un sistema integral. Al incorporarse éste sistema de iluminación se puede conseguir hasta un 70% de ahorro de energía.

V. Referencias

- [1] DALI Activity Group ZVEI-Division Luminaires, Manual, 2001, www.dali-ag.org.
- [2] C. Contenti and T. Ribarich, "Digitally addressable DALI Dimming Ballast" in Proc. Applied Power Electronics Conference and Exposition, Dallas Texas, 2002, pp. 936-942.
- [3] ST7DALI Evaluation Kit User Manual Release 1.0, 2004.
- [4] F. Rubinstein, S. Treado, and P. Pettler, "Standardizing Communication between Lighting Control Devices: a Role for IEEE P1451," in Proc. of Industry Applications Conference, 2003.
- [5] J. Juang, H. Wane, C. Lu, W. Chew, C. Hsu, "A Touch-Dim Network for the

Dimming Control of Lighting System", in Proc. International Conference on Computational Aspects of Social Network, 2010, Taiyuan, China.

- [6] Qing-Min Lai, "Remote Control System for Lighting Based On CAN Bus", in Proc. Second International Conference on Mechanic Automation and Control Engineering, 2011, Mongolia, China.
- [7] Application Note: A CAN Physical Layer discussion, Microchip Technology, 2002.
- [8] MCP2551 High-Speed CAN Transceiver. Microchip Technology, Data Sheet, 2003.

Paint con FPGA: Sistema para la Evaluación de las Interfaces VGA y PS/2 de la Tarjeta de Desarrollo Nexys2

Felipe Santiago Espinosa, **Josué Ayala Ramírez, ***Andrés Ramírez Macías, *Roberto Carlos Vázquez Carrasco**

Instituto de Electrónica y Mecatrónica
Universidad Tecnológica de la Mixteca
Carretera a Acatlima Km. 2.5, Huajuapán de León, Oaxaca, 69000, México
Tel. 953-53-202-14, correo electrónico: *fsantiag@mixteco.utm.mx, **joshu_10@hotmail.com,
max_mousetv@hotmail.com, *houc_darko89@hotmail.com

Resumen

En el presente documento se describe la implementación en un FPGA de un sistema que emula el funcionamiento de un programa para dibujar, similar al programa Paint de Windows. La implementación fue realizada en la tarjeta de desarrollo Nexys2, manufacturada por la empresa Digilent. La tarjeta incluye un FPGA Spartan3E-500 de la firma Xilinx, como elemento principal de procesamiento, además de otros dispositivos para el almacenamiento de datos y el manejo de periféricos. El sistema se desarrolló con la finalidad de evaluar dos de las interfaces de la tarjeta: VGA y PS/2. El diseño requirió de un planteamiento modular, empleando una metodología Top-Down, para posteriormente implementar todos los módulos codificando en VHDL. La interfaz PS/2 se utilizó para la entrada de información, por medio de un ratón. La interfaz VGA se manejó bajo un esquema de 640 x 480 píxeles, dedicando 8 bits por píxel. Por lo tanto, también se utilizó a la memoria RAM M45W8MW16, de la firma Micron, para alojar la información necesaria para el refresco de la pantalla. La memoria tiene una capacidad de 8 Mbyte x 16 bits y es parte de los circuitos integrados incluidos en la tarjeta Nexys 2.

Palabras clave: FPGA, Paint, PS/2, SRAM, VGA, VHDL.

I. Introducción

En la actualidad, el uso de FPGAs para la implementación de sistemas basados en hardware y software cada vez es más demandante. Los FPGAs son dispositivos que proporcionan alta velocidad de procesamiento, flexibilidad en su organización y sobre todo, concurrencia en la realización de tareas. Características que los hacen elegibles para algunas aplicaciones, ante microprocesadores o circuitos integrados a medida [1].

En el mercado se ofertan diferentes tarjetas de evaluación y desarrollo, las cuales, además de un FPGA, incluyen una diversidad de periféricos para aumentar su versatilidad. Un ejemplo es la tarjeta Nexys2, manufacturada y distribuida por la

empresa Digilent [2], en la figura 1 se muestra la vista superior de esta tarjeta.

Las características principales de la tarjeta son [3]:

- FPGA Spartan-3E de la firma Xilinx, con 500K compuertas.
- Memoria Micron M45W8MW16, PSDRAM de 16 MByte.
- Memoria Flash de Intel de 16 MByte.
- Oscilador de 50 MHz.
- 75 terminales de E/S directas del FPGA, disponibles en los conectores de expansión (conector Hirose FX2 con 43 señales y 4 conectores PMod, cada uno con 8 señales).

- Periféricos para evaluación que incluyen: 8 LEDs, 4 displays de 7 segmentos, 4 botones, 8 interruptores deslizantes.
- Puertos: PS/2, VGA y RS232 para aplicaciones. USB para programación y alimentación.

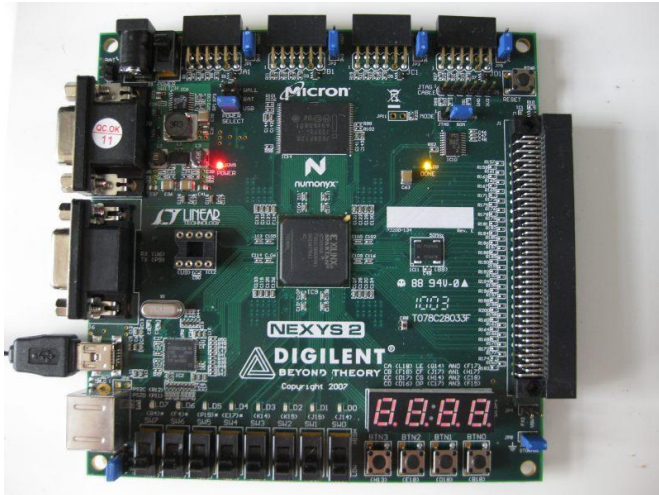


Fig. 1. Vista superior de la tarjeta Nexys2.

Con todas sus características, la tarjeta Nexys2 es ideal durante el proceso de aprendizaje de FPGAs, permitiendo realizar sistemas completos en un solo dispositivo.

Cualquier sistema electrónico requerirá de periféricos para entrada y salida de datos, la tarjeta Nexys2 cuenta con diferentes puertos para ello. En [1, 4, 5] se documentan algunos módulos para el

manejo de periféricos con un FPGA, empleando VHDL.

Específicamente, en [4] se detallan 2 controladores para VGA, uno del tipo gráfico y otro del tipo texto. El controlador del tipo gráfico es la base para un juego simple de ping-pong, ejemplo orientado para la tarjeta *Spartan-3 Starter*, codificado en VHDL.

Con el sistema Paint se busca la evaluación práctica del puerto PS/2 y VGA, en el primero se conecta un ratón y en el segundo un monitor, creando un sistema interesante y de fácil manejo, el cual sirve de base para el desarrollo de sistemas con requerimientos mayores de procesamiento.

II. Organización del Sistema Paint

El sistema se diseñó empleando la metodología Top-Down, que consiste en la división de un sistema complejo en módulos de menor jerarquía, los cuales a su vez se conformarán por otros módulos, hasta llegar a niveles tan simples, cuya implementación no requerirá de mucho esfuerzo [6, 7].

En la figura 2 se muestra la organización jerárquica del sistema, en donde se observan 4 módulos principales y 6 bloques en un segundo nivel. Todos los módulos se codificaron en VHDL, aunque para algunos sólo fue necesaria una adecuación a partir de las descripciones encontradas en [4]. En el diseño de sistemas basados en VHDL es común y conveniente la reutilización de módulos [8].

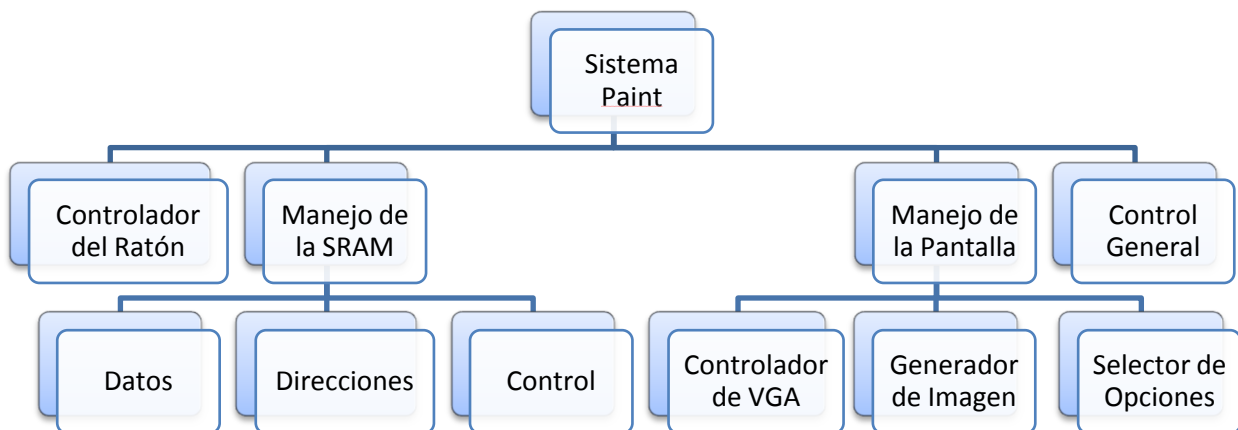


Fig. 2. Organización jerárquica del Sistema Paint.

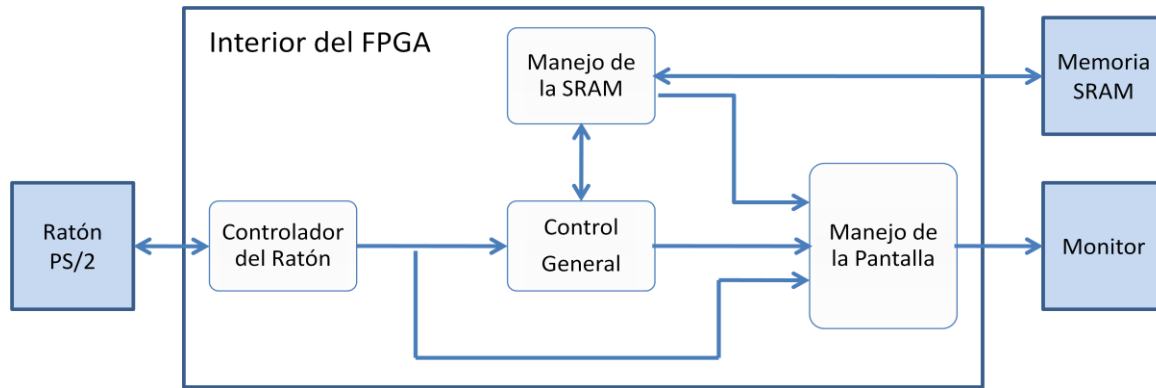


Fig. 3. Diagrama a bloques del Sistema Paint.

El módulo más importante es el Control General, puesto que se encarga de sincronizar y controlar a los módulos restantes (con excepción del Controlador del Ratón).

En la figura 3 se muestra la relación existente entre los módulos de mayor jerarquía.

Cuando el Controlador del Ratón establece la conexión con el dispositivo, envía señales de posición y activación de alguno de sus botones al Control General, el cuál gestiona la escritura o lectura de determinados pixeles en la SRAM, la dirección en SRAM se relaciona con la posición en la pantalla y el contenido con color del pixel.

El Manejo de la Pantalla implica la generación de las señales de sincronía para el monitor, la lectura de la SRAM para obtener el valor de los pixeles que ha escrito el usuario y la generación de pixeles con valores preestablecidos para crear la interfaz gráfica.

En las siguientes secciones se describen los módulos que integran al Sistema Paint, con sus correspondientes bloques.

III. Controlador del Ratón

Este módulo hace posible la interacción con el usuario. El módulo indica la posición de la pantalla en donde se encuentra el puntero del ratón y si se ha presionado alguno de sus botones, en la figura 4 se pueden ver sus entradas y salidas.

Las señales PS2_CLK y PS2_DATA son líneas bidireccionales propias de la interfaz PS/2, se

mantendrán en un nivel lógico alto mientras no haya actividad en el ratón.

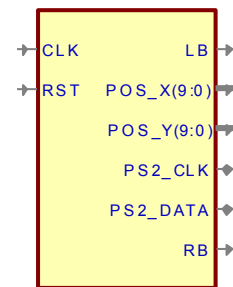


Fig. 4. Controlador del Ratón.

Cada vez que el ratón se mueve o si se presiona uno de sus botones, envía una secuencia de 3 palabras de 11 bits al FPGA. Los 11 bits incluyen: 1 bit de inicio (siempre es '0'), 8 bits de datos (primero el LSB), 1 bit de paridad par y 1 bit de paro (siempre es '1') [3].

El ratón genera la señal de reloj, la cual puede estar en un rango de 20 a 30 KHz. El FPGA va tomando la información en cada flanco de bajada. De los 24 bits de datos, 10 son empleados para indicar la ubicación en cada coordenada, X o Y, 2 indican las actividades en los botones y los otros 2 permanecen sin uso, en un ratón con 2 botones.

El Controlador del Ratón se basa en una máquina de estados por medio de la cual se sondearán las señales PS2_CLK y PS2_DATA, dejando en los vectores POS_X y POS_Y la posición del cursor y poniendo en alto a LB cuando se presione al botón izquierdo o a RB si se presiona el botón derecho.

IV. Manejo de la SRAM

Con este módulo se tendrá acceso a las localidades de la memoria SRAM externa, ya sea para guardar o extraer un dato. Se trata de una memoria Micron M45W8MW16, PSDRAM de 16 MByte, la cual puede trabajar en modo síncrono o asíncrono, en la figura 5 se ilustra el diagrama de tiempos para un acceso en modo asíncrono [9].

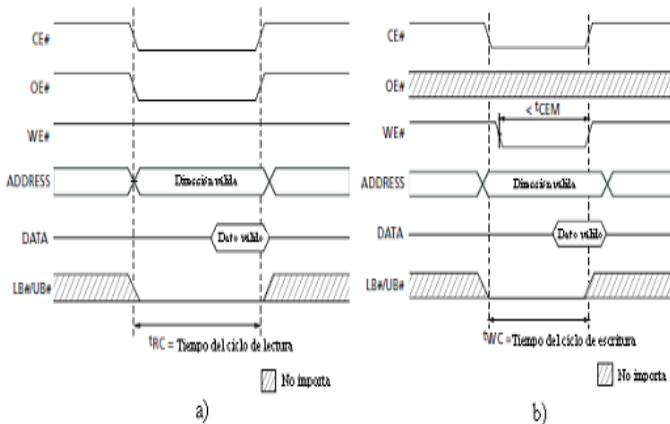


Fig. 5. Tiempos de acceso a la SRAM M45W8MW16 en modo asíncrono (a) Operación de lectura (b) Operación de escritura.

El tiempo mínimo requerido para una operación de escritura y lectura en el modo asíncrono es de 70nS. En el modo síncrono se puede utilizar una frecuencia hasta de 80Mhz.

En la tabla 1 se muestran las señales necesarias para el manejo de la SRAM.

Manejo de Datos

Aunque la memoria tiene un bus de datos de 16 bits, únicamente se emplearon datos de 8 bits, porque cada dato va a corresponder con el valor de un píxel.

Es necesario un bloque para el manejo de los datos porque el bus es bidireccional, por lo tanto, su objetivo básicamente consiste en la separación del bus bidireccional en 2 buses, uno de entrada y otro de salida.

En la figura 6 se muestran las señales de este bloque, con R y W se determina si se trata de una lectura o escritura. Mientras no exista una escritura el bus será puesto en alta impedancia, pudiendo realizar únicamente lecturas.

Señal	Descripción
ADDR (23:0)	Bus de direcciones de 24 bits.
DATA (15:0)	Bus de datos de 16 bits (bidireccional).
OE	Señal de habilitación de salida.
WE	Señal de habilitación de escritura.
MT-ADV	Dirección válida.
MT-CLK	Señal de reloj para SRAM en modo síncrono.
MT-UB	Señal de activación de byte superior para direccionamiento por byte.
MT-LB	Señal de activación de byte inferior para direccionamiento por byte.
MT-CE	Habilitación del chip.
MT-WAIT	Señal de espera para SRAM en modo síncrono.

Tabla 1. Señales necesarias para la SRAM.

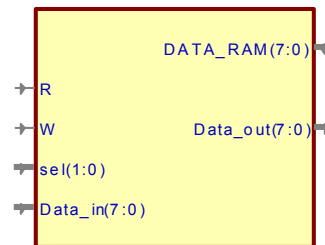


Fig. 6. Bloque para el manejo de datos de la SRAM.

Con el bus de selección se determina si se escribirá el dato o si la localidad será limpiada (escribiendo 0xFF), el bus es de 2 bits para hacerlo compatible con el manejo de direcciones.

Manejo de Direcciones

El manejo de direcciones es fundamental, dado que se podrán realizar diferentes acciones. En realidad este bloque se compone de 4 bloques de menor jerarquía, como se muestra en la figura 7, aunque todos ellos son muy simples.

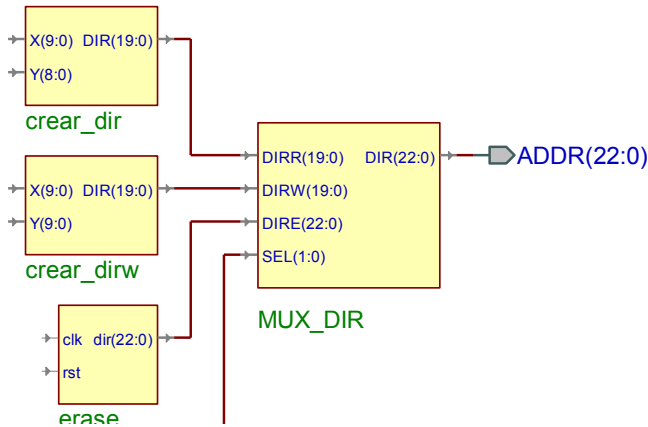


Fig. 7. Bloque para la generación de las direcciones de acceso a la SRAM.

La memoria normalmente será leída, en este caso la dirección se crea a partir de las coordenadas de los pixeles, durante el recorrido para su despliegue, proporcionadas por el Controlador de VGA.

La memoria también será escrita, cuando se mantenga presionado el botón izquierdo del ratón, la dirección se genera con las coordenadas proporcionadas por el Controlador del Ratón.

Finalmente, la memoria será completamente borrada (escribiendo 0xFF en todo su contenido), esta acción se realizará con la inicialización del sistema, con este contenido, en la pantalla se tendrá un color blanco. La selección de la dirección se realiza con el bus SEL, es cual es generado por el Control General, para el manejo del multiplexor de la salida del bloque.

Control de la SRAM

El bloque de control de la SRAM genera las señales necesarias para su manejo. En la figura 8 se muestran sus entradas y salidas, la SRAM normalmente se estará leyendo, a menos que se indique que se realizará una escritura.

Este es un bloque secuencial basado en una máquina de estados, dado que la señal de reloj de la tarjeta tiene una frecuencia de 50 MHz (periodo de 20 nS), con diferentes estados se cumplirá el tiempo mínimo de 70 nS, mientras se realiza la secuencia de habilitaciones correspondiente (ver figura 5).

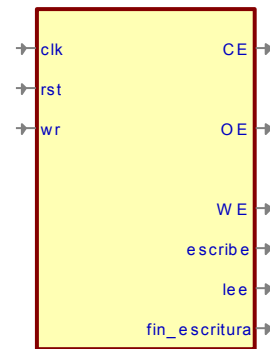


Fig. 8. Bloque para la generación de las señales de control para la SRAM.

V. Manejo de la Pantalla

Las señales necesarias para producir una imagen son [5]:

Señal de luminancia. Prácticamente es la señal de video, pero solo representa un punto de la imagen (pixel). Para desplegar la imagen completa se tiene que mostrar elemento por elemento, por lo cual se debe realizar una exploración, como se muestra en la figura 9.

La exploración se realiza con dos contadores, uno para el barrido horizontal y otro para el vertical. El contador horizontal recorre de izquierda a derecha, y el vertical de arriba a abajo. Una vez que se ha alcanzado el punto inferior derecho se continuará nuevamente con el punto superior izquierdo.

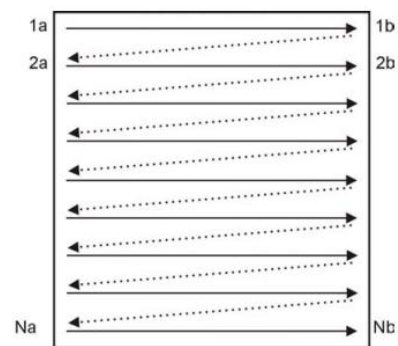


Fig. 9. Exploración de la pantalla.

Pulsos de sincronización. Se utilizan para controlar la trayectoria del haz de electrones. Hay pulsos de sincronización horizontal y vertical, ambos son activos en un nivel lógico bajo. El pulso horizontal marca el inicio y fin de una línea de

exploración, mientras que el pulso vertical indica el inicio y fin de una pantalla.

Pulsos de blanqueo. El objetivo de estos pulsos es poner la pantalla en negro, puesto que durante ese tiempo no se desplegará ningún punto de la imagen. La señal de blanqueo horizontal es independiente a la de blanqueo vertical.

En el sistema Paint, la escritura en la SRAM se realizará únicamente durante los pulsos de blanqueo, ya que cuando éstos no ocurran, se realizarán lecturas.

En la figura 10 se muestra la temporización en las señales empleadas para el manejo de la pantalla.

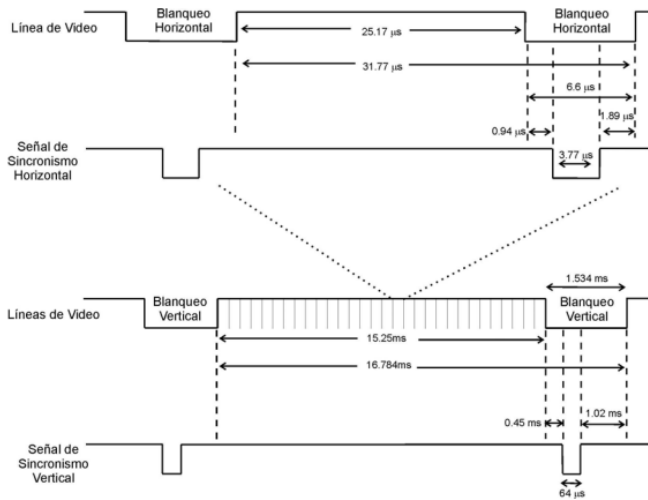


Fig. 10. Temporización en las señales de la pantalla.

El módulo para el manejo de la pantalla incluye 3 bloques: el Controlador de VGA, para generar las señales de sincronía, el Generador de Imagen, para determinar si en la pantalla se colocará un pixel obtenido de SRAM o un pixel con un valor predefinido, y el Selector de Opciones, para definir el color a mostrar en la pantalla, en el momento en que se presione el botón izquierdo del ratón.

Controlador de VGA

En la figura 11 se muestran la interfaz de este bloque, el cual se va a conectar directamente con el puerto VGA.

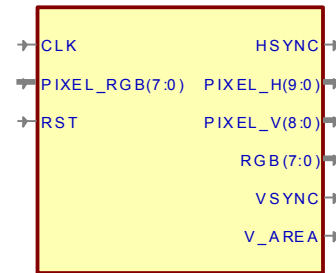


Fig. 11. Bloque Controlador de VGA.

En HSYNC y VSYNC se tendrán los pulsos de sincronía, de acuerdo con los tiempos mostrados en la figura 10, y en RGB se colocará el valor instantáneo del pixel.

El bloque se basa en contadores, los cuales se reflejarán en los vectores PIXEL_H y PIXEL_V, para que internamente se identifique el pixel al que se está haciendo referencia. La señal V_AREA permitirá detectar si un pulso de blanqueo está activo, para considerar una posible escritura en SRAM.

Generador de Imagen

Con este bloque se determina si en la pantalla se colocará un pixel obtenido de SRAM o un pixel con un valor predefinido. El bloque genera la información que corresponde con el puntero del ratón y con las opciones que se muestran en la pantalla, como la paleta de colores, la selección del lápiz o del borrador.

En la figura 12 se muestra la interfaz de este bloque, se observa que recibe las coordenadas del Controlador de VGA y del Controlador del Ratón, comparando ambas coordenadas se determina si en RGB_sal se colocará la información del cursor del ratón o se dejará la información recibida en RGB_ENT.

La entrada SEL define si se mostrará el lápiz o el borrador. La entrada Ver_color representa el color que actualmente está habilitado (el color con el que se pintará en la pantalla).

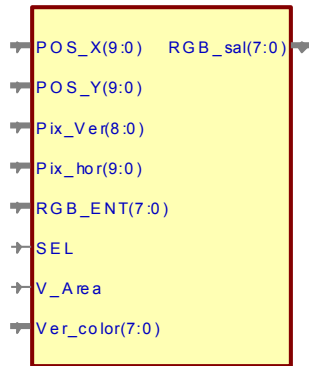


Fig. 12. Bloque Generador de Imagen.

Selector de Opciones

Como su nombre lo indica, este bloque tiene como objetivo almacenar el color con el que se pintará y limitar el área de las barras de herramientas y paletas (Figura 13). Las opciones mostradas en pantalla son: Color, Lápiz y Borrador.

El valor del color activo se almacenará en la SRAM cuando el usuario dé un clic izquierdo a una parte del área de dibujo.

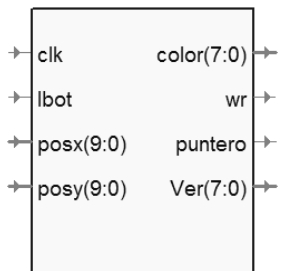


Fig.13 Bloque Selector de Opciones.

El bloque verifica la posición del ratón para activar la opción que corresponda con el área en donde el usuario haya dado un clic izquierdo.

VI. Control General

El Control General gestiona las acciones que deben tomar los distintos módulos, a partir de las decisiones tomadas por el usuario. También sincroniza el tiempo de escritura, ya que sólo se puede escribir en la SRAM cuando se presenta la etapa de blanqueo. En la figura 14 se muestra su interfaz.

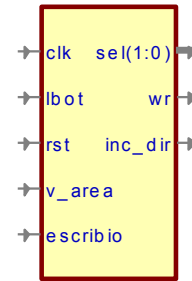


Fig. 14. Módulo del Control General.

Al alimentar el sistema, el Control General hace que las localidades de la memoria sean llenadas con el valor 0xFF, que corresponde al color blanco, con el fin de no contener “basura” en la memoria y mostrar la pantalla en blanco.

Posteriormente activa al control de la SRAM en modo de lectura y permanece en ese estado hasta que el usuario, al presionar el botón izquierdo del ratón, active la escritura en la SRAM.

Desde el Control General se determina si en la SRAM se realizará una lectura o escritura, y por lo tanto, también se define de donde proviene la dirección de acceso.

El comportamiento del Control General se esquematiza con la máquina de estados mostrada en la figura 15.

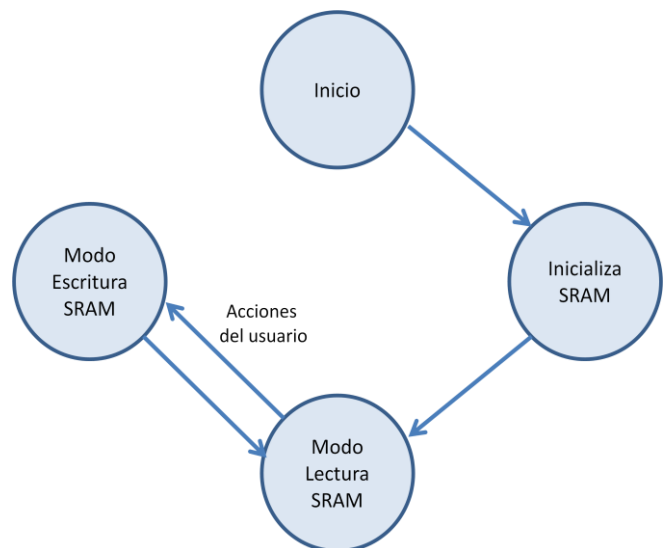


Fig.15. Máquina de estados del Control General.

VII. Resultados

Integrando todos los módulos se construye al sistema Paint, cuya interfaz es mostrada en la figura 16. Las señales en la parte superior corresponden a la memoria SRAM, aunque se agregan algunas para inhabilitar a la memoria Flash, dado que comparten algunos buses en la tarjeta Nexys2.

Las siguientes señales corresponden con el puerto VGA y por último se tienen las señales para el manejo del ratón, las salidas “lbot” y “rbot” hacen referencia a los botones, mostrando su estado en 2 LEDs de la tarjeta.

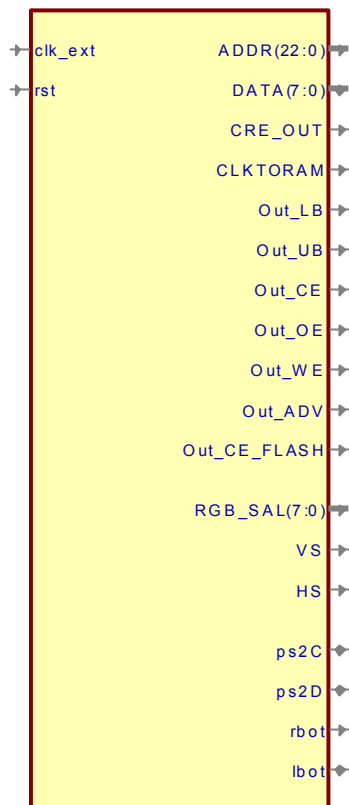


Fig. 16 Sistema Paint.

Se probó al sistema completo, obteniendo un buen desempeño al momento de mostrar los pixeles pintados en la pantalla. Las opciones colocadas en la barra de herramientas cumplen con los requerimientos mínimos de un sistema para realizar dibujos.

En la figura 17 se muestra la interfaz gráfica que genera el sistema. Se puede observar la barra de herramientas en la parte superior de la pantalla, la cual contiene una área de opciones de puntero (parte superior izquierda) y una paleta de selección de colores. El área de trabajo se ha limitado a la parte inferior de la barra de herramientas impidiendo que el usuario pinte sobre el área de la paleta y demás opciones.

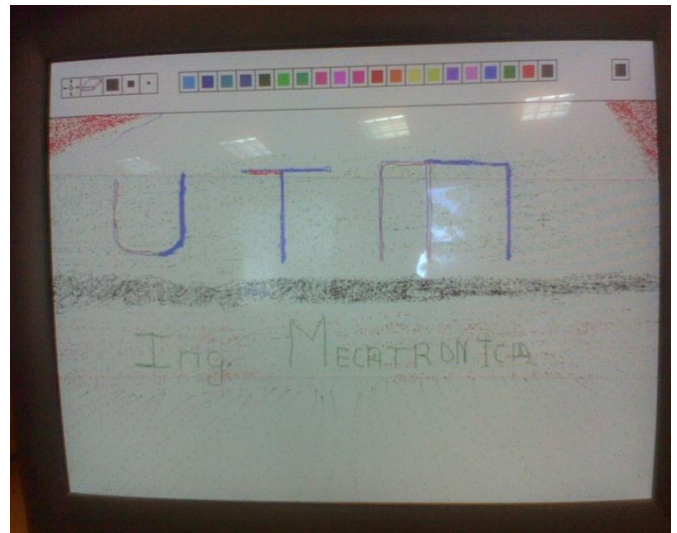


Fig. 17. Sistema Paint, puesto en marcha.

En la figura 18 se puede ver la conexión del monitor en el puerto VGA, y la conexión del ratón en el puerto PS/2. Ambos puertos en la tarjeta Nexys2.

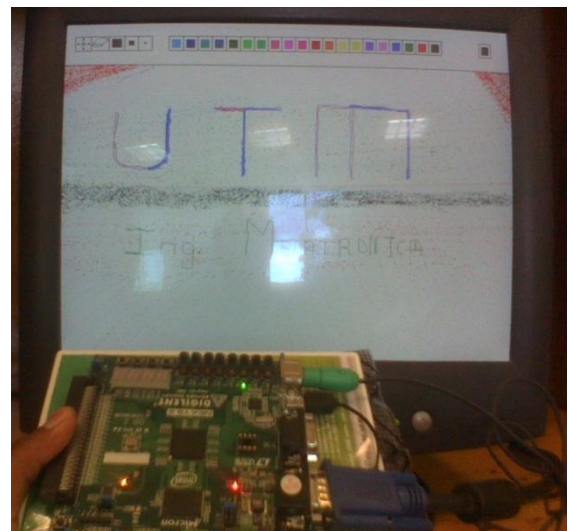


Fig. 18. Sistema implementado.

En la tabla 2 se muestra un resumen de los recursos ocupados por el sistema Paint en el FPGA Spartan3E, demostrando que son muy pocos, con respecto al total de recursos disponibles.

Recurso	Usado	Disponible	Uso
Slices	858	4, 656	18 %
Bloques I/O	56	232	24 %
Bloques RAM	0	20	0 %
GCLKs	4	24	16 %

Tabla 2. Recursos empleados por el sistema Paint.

VIII. Conclusiones

El sistema Paint permitió evaluar favorablemente los puertos VGA y PS/2 incluidos en la tarjeta Nexys2, se diseñó e implemento un sistema completo que hace uso de ellos.

Además, también fue necesario el uso de la memoria M45W8MW16, incluida en la tarjeta. En principio no se había considerado, sin embargo, con una resolución en el monitor de 640 x 480 y colores de 8 bits, implica un total de 307, 200 bytes, que agotan los recursos del FPGA.

El sistema funcionó de manera correcta, no obstante, cuando se encuentra presionado el botón izquierdo del ratón y éste se mueve rápidamente, no se alcanzan a pintar todos los puntos en la pantalla, esto puede deberse a que la duración del pulso de blanqueo no es suficiente para escribir en todas las direcciones relacionadas. El acceso a la memoria es asíncrono, por lo que una posible mejora consistiría en un acceso síncrono, el cual puede realizarse a una velocidad más alta, con ello, se podrán escribir más píxeles durante los pulsos de blanqueo.

En la tabla 2 se puede ver que aún quedan demasiados recursos para el desarrollo de aplicaciones, basadas en la tarjeta Nexys2, que requieran utilizar un ratón y un monitor, como elementos de entrada y salida. Por ejemplo, podría acondicionarse un brazo robótico, haciendo que el

brazo dibuje en alguna superficie lo que el usuario dibujó en la pantalla.

IX. Referencias

- [1] P. Wilson, "Design Recipes for FPGAs", 2007, Newnes of Elsevier.
- [2] Digilent Inc. - Digital Design Engineer's Source, <http://www.digilentinc.com/>, última visita: Junio de 2011.
- [3] Digilent Nexys2 Board Reference Manual, Doc: 502-107, Copyright Digilent, Inc., June 21, 2008.
- [4] P. Chu, "FPGA Prototyping by VHDL Examples, Xilinx Spartan-3Version", A John Wiley & Sons inc., Publication 2008
- [5] J. C. Tepozán Ríos, "Diseño de controladores para la tarjeta de desarrollo XSA-100". Tesis de Ingeniería, Universidad Tecnológica de la Mixteca, Mayo de 2004.
- [6] S. Brown, Z. Vranesic, "Fundamentos de Lógica Digital con Diseño VHDL". 2ª Ed., 2006, Editorial McGraw-Hill.
- [7] R. Romero Troncoso, "Sistemas Digitales con VHDL"; Legaria Ediciones.
- [8] F. Pardo, J. A. Boluda, "VHDL, Lenguaje para síntesis y modelado de circuitos", Ed. Alfaomega, 2000
- [9] M. A. Mendoza Manzano, "Procesamiento y Análisis Digital de Imágenes Mediante Dispositivos Lógicos Programables", Tesis de Ingeniería, Universidad Tecnológica de la Mixteca, Febrero de 2009.

X. Autores

M. C. Felipe Santiago Espinosa es Maestro en Ciencias con especialidad en Electrónica por parte del INAOE, Licenciado en Electrónica por parte de la BUAP, desde 1998 labora como Profesor – Investigador en la Universidad Tecnológica de la Mixteca, adscrito al instituto de Electrónica y Mecatrónica.

Josué Ayala Ramírez es estudiante de la Ingeniería en Mecatrónica, en la Universidad Tecnológica de la Mixteca.

Andrés Ramírez Macías es estudiante de la Ingeniería en Mecatrónica, en la Universidad Tecnológica de la Mixteca.

Roberto Carlos Vázquez Carrasco es estudiante de la Ingeniería en Mecatrónica, en la Universidad Tecnológica de la Mixteca.

Descripción de un Sistema de Adquisición Controlado por FPGA para Procesamiento de Pulsos Provenientes de PMTs

*Rubén Conde, *Humberto Salazar, *Oscar Martínez* y **L. Villaseñor **

* Facultad de Ciencias Físico Matemáticas, BUAP, Puebla, México, rconde@cfm.buap.mx

** Instituto de Física y Matemáticas, San Nicolás de Hidalgo, Morelia, México.

Resumen

Se presenta una versión de diseño e implementación de un sistema de digitalización de hasta 100MSPS para cuatro señales independientes provenientes de fotomultiplicadores (PMTs) basado en FPGA, sistema que será utilizado en el observatorio *Cosmic High Altitude Radiation Monitor* (CHARM) como instrumentación para la detección de la radiación Cherenkov producida por rayos cósmicos. Este sistema de adquisición de datos y procesamiento de pulsos es adaptado a una tarjeta de desarrollo Nexys2 de Digilent Inc, en su diseño se utilizando lenguaje de descripción hardware y también se utiliza un soft-processor para lograr una arquitectura con versatilidad en los procesos que requieren paralelismo y control en ellos. Se muestran resultados para las primeras pruebas realizadas en los PMTs, así como las primeras gráficas de carga y amplitud para la caracterización de los PMTs.

Palabras Clave: FPGA, Microblaze, CHARM.

I. Introducción

El observatorio híbrido CHARM, ubicado a 4300 m.a.s.l. equivalente a 430gr/cm^2 y localizado en (N $18^\circ 59.1$, W $97^\circ 18.76$) cerca de la Ciudad de Puebla en México, incluye detectores de radiación Cherenkov en agua y Chrenkov en aire. La meta principal de este observatorio será determinar la composición de la masa de la partícula primaria del rayo cósmico con energías entre 10^{16} y 10^{18} eV usando las ventajas que tienen las técnicas de detección empleadas y la ventaja que ofrece el observatorio por su ubicación geográfica.

Una de las tareas más comunes en los experimentos en superficie es el estudio de la composición primaria en la segunda rodilla del espectro de energía. Tomando ventaja de la altitud y de algunas ventajas encontradas en el volcán Pico de Orizaba cerca de la ciudad de Puebla, se tiene el Observatorio CHARM y en este se tienen este tipo de estudios.

Este observatorio usará las ventajas de la detección de radiación Cherenkov en agua como medida confiable del número de partículas secundarias debido al rayo cósmico primario y a

su buena determinación de dirección de arribo del rayo cósmico primario debido a la diferencias en tiempo de arribo.

Esto dará una estimación de la energía primaria y una estimación de la dirección de arribo, y utilizando la función de distribución lateral se podrá determinar la posición del *core* del chubasco. La energía total y la dirección de arribo de un Extensive Air Shower ó Chubasco Aereo Extenso (EAS) se determinará empleando el arreglo de detectores de radiación Cherenkov en Agua (WCD).

Como técnica complementaria de detección para obtener información relacionada con la composición de la partícula primaria, se propone un arreglo de detectores de radiación Cherenkov en aire (ACD) distribuidos en un arreglo más compacto que el WCD para determinar el perfil temporal de un EAS en su viaje por la superficie terrestre. La principal meta es determinar la altitud a la cual un EAS alcanza la máxima producción de partículas secundarias o X_{max} . Esta cantidad es directamente relacionada con la masa del primario.

Usando estas técnicas se espera determinar la tendencia de la masa del primario en la región de

la segunda rodilla del espectro de energía de la radiación del rayo cósmico.

La instrumentación correspondiente a cada tanque dentro de un arreglo EAS generalmente consiste en la detección de radiación Cherenkov en aire y en agua, sensores de parámetros ambientales, GPS, y comunicación inalámbrica entre otros. Con esto se tiene que la modularidad persistente en el arreglo, por lo que se decide desarrollar un elemento con estos elementos para extenderse posteriormente a todo el arreglo.

Dadas las simulaciones y los avances que se tienen en los prototipos existentes en la electrónica, buscamos la adecuación de los parámetros físicos para hacer una correcta instrumentación que nos permita el estudio correcto de los fenómenos de interés para este observatorio.

II. Desarrollo

El desarrollo de este sistema está basado en un FPGA Spartan3E (XC3S500E-FG320 4ns) que permite la flexibilidad de incrustar un procesador *microblaze* RISC de 32Bits y arquitectura Harvard creado por Xilinx[]. El uso de este procesador es totalmente personalizado por ser desarrollado con lógica programable, y requiere muy pocas conexiones y ningún componente adicional externo, al cual se le pueden instanciarse distintos recursos de hardware tales como puertos serie, timers, Ethernet, e IP de propósito específico y general. Transformándolo en un procesador configurado a la medida del desarrollo.

Se ha incorporado un procesador *microblaze* en la arquitectura implementada, y también se utiliza el lenguaje de descripción hardware (VHDL) para completar la implementación del proyecto, incrementando así su potencialidad final. Separando el proyecto en dos partes principales:

- Por un lado mantenemos los módulos de procesamiento del pulso, independientes entre sí. Con la idea de obtener el paralelismo en los módulos y buscar una

granularidad fina (a nivel de celdas básicas de las cuales se compone un FPGA, o grupos de CLB – Bloque Lógico Configurable) para mantener la eficiencia de máxima velocidad y mínima área de implementación. Con esto logramos extender los módulos de un canal a varios canales sin necesidad de rediseñarse. En este caso se usa la descripción de los módulos usando VHDL.

- La segunda parte es realizada por un procesador *microblaze* insertado, el cual tiene como idea realizar todas las tareas que tengan una naturaleza secuencial como el control e intercomunicación entre los módulos con el exterior del FPGA. Señalando que se usa el lenguaje C para la descripción de los algoritmos, lo que permite una mayor agilidad en la depuración de código y reducción en el tiempo de desarrollo del proyecto final.

La tarjeta desarrollada consta de dos bloques principales (Fig. 1) donde cada bloque contiene:

1. Dos entradas tipo lemo a través del conector EPL.00.250 [],
2. Un convertidor analógico a digital AD9216[6] dual de hasta 100MSPS,
3. Un reloj configurable para el ADC,
4. Dos bus datos de 10bits,
5. Dos amplificadores diferenciales,
6. Esta tarjeta consta de dos módulos como el listado del punto 1 al 5,
7. y conectividad para los dos bloques con la tarjeta de desarrollo Nexys2 a través del conector HIROSE FX2-100P.
8. Usa los recursos de la tarjeta Nexys2-500E para la implementación de la arquitectura y la comunicación con la PC.

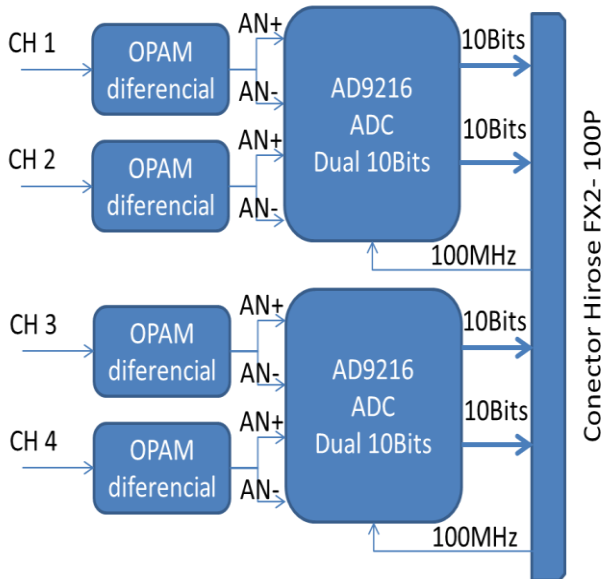


Fig. 1 Diagrama a bloques de la tarjeta de cuatro canales.

El diagrama eléctrico tiene un acoplamiento en DC, convirtiendo cada señal de entrada con referencia simple a señal de salida diferencial para ajustarse a la entrada del ADC dual, basado en el AD8138 de Analog Device[6].

Este circuito es diseñado para lograr un máximo desempeño del AD9216 estableciendo un menor ruido en la señal y baja impedancia. Ya que, la salida diferencial del AD8138 ayuda a balancear la entrada diferencial del AD9216.

Características del DAQ Hardware

La tarjeta hija de cuatro canales analógicos es basada en el ADC Dual AD9216-105MSPS de 10Bits de resolución de Analog Devices [], con un rango dinámico de conversión de -1 a +1Volts y ancho de banda de 320MHz a -3dB.

Esta tarjeta es diseñada para tener conectividad con la tarjeta Nexys2 de Digilent Inc a través de su conector Hirose FX2-100 y utiliza la comunicación RS232 de esta tarjeta para intercambiar con la PC.

La tarjeta Nexys2 es una tarjeta de desarrollo basada en el FPGA Spartan3E de Xilinx. Contiene un puerto USB2 de alta velocidad, 16 Mbytes de memoria RAM y ROM, y varios puertos de entrada salida digitales lo que lo hace

ideal para poderse extender a cualquier sistema digital, incluyendo un procesador empotrado pasado en el microblaze de 32Bits RISC de Xilinx. Un puerto de comunicaciones de baja velocidad como RS232.

PCB

El diseño del PCB es desarrollado utilizando el software de diseño Altium [6], y su fabricación (Fig. 2) fue hecha en nuestras instalaciones del laboratorio de Astropartículas en la FCFM-BUAP.

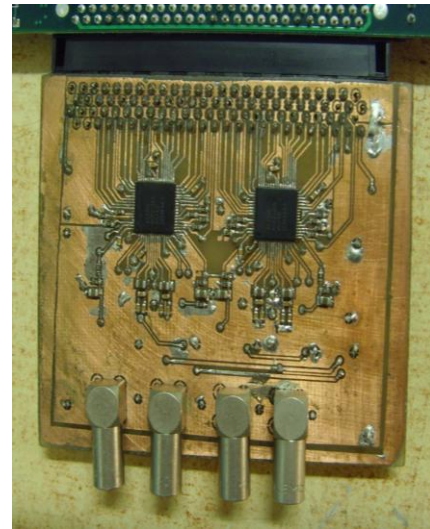


Fig. 2 Fabricación de la tarjeta digitalizadora de cuatro canales

III. Resultados

Mostraremos algunas primeras pruebas realizadas a la tarjeta entre estas es el histograma de ruido de la misma.

Para obtener el histograma se adquieren las señales en cada uno de los cuatro canales analógicos. El número de muestras adquirido a 100MSPS es:

Canal 1: 39 705 datos
 Canal 2: 43747 datos
 Canal 3: 62945 datos
 Canal 4: 54871 datos

Dado que los recursos en el FPGA son limitados, y para no detener la velocidad de conversión de 100MSPS se propone adquirirlos por bloques de 1024 datos, o lo que es lo mismo bloques de

$1024(10\text{ns}) = 1,024\text{useg}$, como lo muestra el diagrama de la Fig. 3.

Los datos en cada canal son tomados en forma independiente y secuencial, no se busca el paralelismo entre ellos. Si fuese el caso no habría cambios en el ruido entre los canales, ya que su conversión a digital está siempre presente activo.

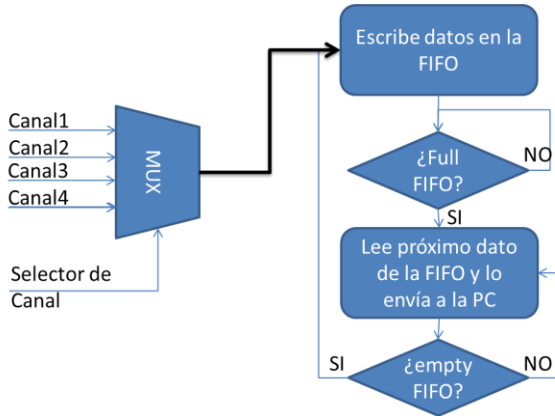


Fig. 3 Se elige un canal de captura, y se toman bloques de 1,024useg y posteriormente son enviados a la PC. Este tiempo de envío es la separación en tiempo que existe entre cada bloque de datos.

Antes de mostrar los valores de conversión a digital, la Tabla 1 muestra la equivalencia de conversión de salida con desplazamiento binario (*offset binary*), utilizada en configuración de polarización propuesta para AD9216.

Tabla 1 Código de salida

Código	(Vin+)-(Vin-)	Offset Binario
1023	>0.998 V	11 1111 1111
1023	+0.998 V	11 1111 1111
1022	+0.996 V	11 1111 1110
.	.	.
.	.	.
513	+0.002 V	10 0000 0001
512	+0.0V	10 0000 0000
511	-0.002 V	01 1111 1111
.	.	.
.	.	.
1	-0.998 V	00 0000 0001

0	-1.00 V	00 0000 0000
0	<-1.0 V	00 0000 0000

La Fig. 4 muestra parte de los datos obtenidos por el sistema de adquisición, también se puede observar la NO variación de la línea base de las señales.

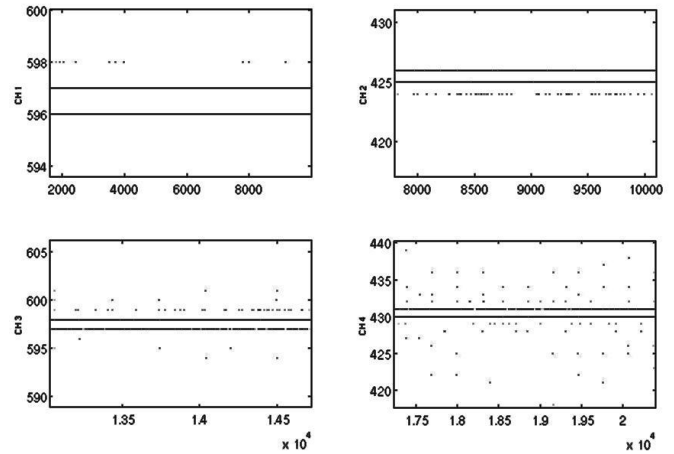


Fig. 4 Gráficas para la señal de entrada en la tarjeta donde las señales son tomadas de la señal de referencia (GND), es decir, la señal entrante es cortocircuitada a GND. De esta forma se tiene el ruido del sistema de nuestra tarjeta junto con la Nexys2.

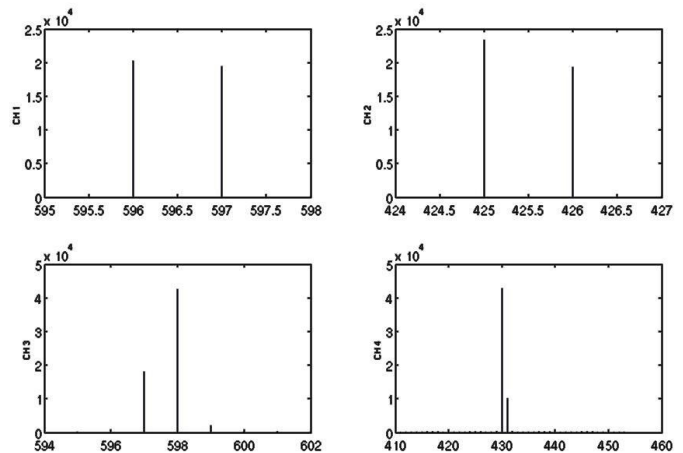


Fig. 5. Histogramas para cada canal. Ver Tabla 1

Adquisición libre y perfil de trazas

Manteniendo el mismo diagrama de flujo de la Fig. 3, se realiza la adquisición de los cuatro

canales en forma continua, y de la misma forma se realiza la adquisición de las trazas y son registradas con su respectivo valor de umbral a partir de la línea base que es tomada del promedio de 2×16 datos. La siguiente Fig. 6 muestran una corrida de la adquisición de la señal tomada de una señal de generador de funciones. El pulso tiene una frecuencia de 500KHz, y dado que el generador de funciones utilizado tiene una frecuencia máxima de 2MHz nos da un pulso a esta frecuencia no completamente definido pero nos sirve para mostrar el funcionamiento de los módulos implementados.

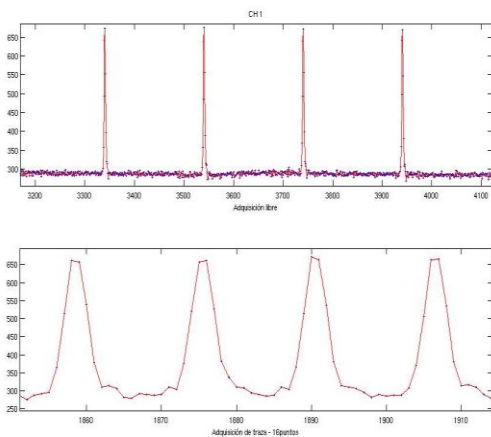


Fig. 6 Datos de uno de los cuatro canales, digitalizados a 100MSPS es decir cada 10ns, en su modo de adquisición libre y en la gráfica de abajo se muestra la toma de perfiles de traza de 16 puntos por cada una de ellas.

Histogramas de carga y amplitud máxima de Detector Cherenkov en Agua

La calibración de energía absoluta de un tanque es requerida para medir la energía Cherenkov depositada en cada detector por chubascos secundarios para reconstruir los parámetros del primario (en particular la energía primaria). La unidad de calibración adoptada es el *Vertical Equivalent Muon* (VEM) definida como la carga promedio colectada por un tanque que ha sido atravesado verticalmente en su centro por un simple muon de alta energía. Estos muones pueden ser fácilmente identificados con dos plásticos centelleadores colocados arriba y abajo del tanque, centrado en su eje [] y []

Muestro datos tomados de un WCD con nuestro sistema a una razón de muestreo de 10ns y comunicación serial de 115200 baud por segundo. La Fig. 7 y Fig. 8, muestran la distribución de carga y la distribución de amplitud. La carga está determinada la integración de 16 puntos que conforman el perfil de la traza registrada.

NOTA: los histogramas mostrados no tienen un mismo número de perfiles de traza, pero si tienen más de 100 mil perfiles de traza en cada registro de las señales.

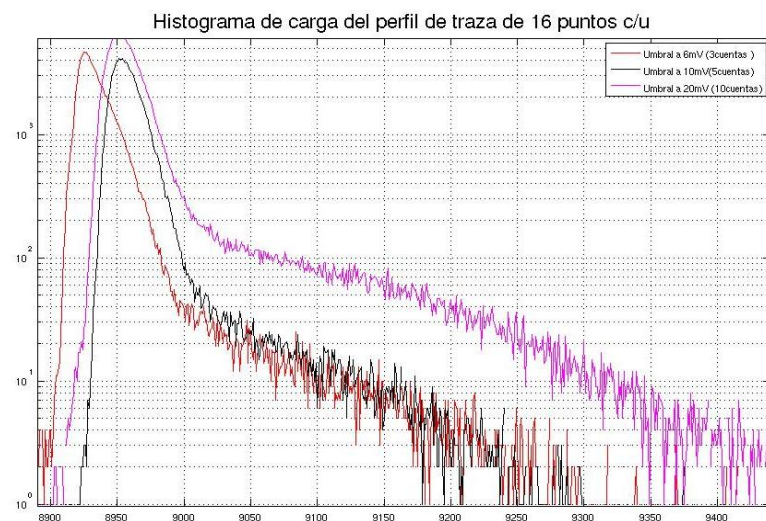


Fig. 7 Histograma de carga de señales detectadas por un WCD con nuestro sistema de digitalización propuesto para tres umbrales de disparo

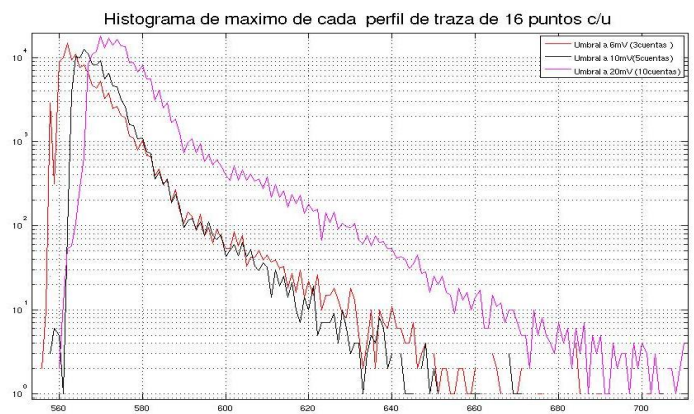


Fig. 8 Distribución de amplitud de señales detectadas por un WCD con nuestro sistema de

digitalización propuesto para tres umbrales de disparo.

IV. Discusión

Este sistema representa una importante tarea resuelta ya que sustituye a versiones anteriores de diseños propuestos para adquisición y procesamiento de pulsos provenientes de los fotomultiplicadores como medios de sensor de radiación de Cherenkov.

Se ha incorporado un procesador microblaze en el FPGA y aunado con diseños creados en VHDL se logran arquitecturas cada vez más complejas para realizar procesamiento de pulsos en cada canal de conversión de forma independiente.

V. Conclusiones

Se presentó el diseño, e implementación de un sistema de adquisición de datos para cuatro canales analógicos que tiene aplicación en el observatorio CHARM, aún bajo construcción en el Pico de Orizaba ubicado en el Parque Nacional Sierra Negra, como medio de digitalización para las señales provenientes de PMTs.

Se han mostrado resultados que ayudará a determinar la calibración del PMTs, obteniendo en su primera fase los histogramas de carga y máximo umbral en cada perfil de traza generados por el PMT.

VI. Referencias

- [1] <http://www.xilinx.com>
- [2] <http://mexico.newark.com/lemo/epl-00-250-ntn/rf-coaxial-epl-jack-r-a-50ohm-solder/dp/70C8070>
- [3] <http://www.digitlentinc.com>
- [4] <http://www.analog.com>
- [5] <http://www.altium.com>

[6] X. Bertou and For the Pierre Auger Collaboration, "Background radiation measurement with water Cherenkov detectors", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 639, Issue 1, 21 May 2011, Pages 73-76.

[7] A. Etchegoyen for The Pierre Auger Collaboration, "Muon-track studies in a water Cherenkov detector", Nuclear Instruments and Methods in Physics Research. April 2005 Pages 602 -612.

VII. Autores

M. en C. José Rubén Conde Sánchez estudiante de doctorado en FCFM-BUAP

Dr. Humberto Salazar Profesor investigador del FCFM-BUAP

Dr. Oscar Martínez Profesor investigador de la FCFM-BUAP

Dr. Luis Villaseñor Profesor investigador de Instituto de Física y matemáticas San Nicolás de Hidalgo Morelia Michoacán.