



**SEP**  
SECRETARÍA  
DE EDUCACIÓN  
PÚBLICA



TECNOLÓGICO  
NACIONAL DE MÉXICO



## Robótica y Computación. Investigación y Desarrollo.

Coordinadores de la Edición:  
Iliana Castro Liera  
Mario Cortés Larrinaga



ISBN: 978-607-97174-4-2



9 786079 817442

Esta página fue dejada en blanco intencionalmente.





## Robótica y Computación. Investigación y Desarrollo.

D. R. ©INSTITUTO TECNOLÓGICO DE LA PAZ  
Blvd. Forjadores de B.C.S. No. 4720  
Col. 8 de octubre  
C.P. 23080  
La Paz Baja California Sur, México

Primera edición mayo de 2019  
ISBN: 978-607-97174-4-2



Reservados todos los derechos. Ninguna parte de este libro puede ser reproducida, archivada o transmitida, en cualquier sistema –electrónico, mecánico, de fotorreproducción, de almacenamiento en memoria o cualquier otro– sin hacerse acreedor a las sanciones establecidas en las leyes, salvo con el permiso escrito del titular de los derechos de la obra.

Diseño de la portada: Iliana Castro Liera

Esta página fue dejada en blanco intencionalmente.

**Coordinador General**

Dr. Lorenzo Fidel Cota Verdugo - Instituto Tecnológico de La Paz

**Coordinador del Programa Técnico**

Dr. Marco Antonio Castro Liera - Instituto Tecnológico de La Paz

**Comité Revisor**

Dr. Francisco Herrera Fernández - UCLV Cuba

Dr. Alejandro Israel Barranco Gutiérrez - Instituto Tecnológico de Celaya

Dr. Víctor Adrián Santibáñez Dávila - Instituto Tecnológico de La Laguna

Dr. Ricardo Emmanuel Campa Cocom - Instituto Tecnológico de La Laguna

Dr. Alejandro Enrique Dzúl López - Instituto Tecnológico de La Laguna

M.C. Edmundo Javier Ollervides Vázquez - Instituto Tecnológico de La Laguna

Dr. Eduardo Javier Moreno Valenzuela - CITEDI Tijuana

Dr. Eusebio Bugarín Carlos - Instituto Tecnológico de Ensenada

Dra. Ana Yaveni Aguilar Bustos - Instituto Tecnológico de Ensenada

Dr. Jesús Alberto Sandoval Galarza - Instituto Tecnológico de La Paz

M.S.C. Iliana Castro Liera - Instituto Tecnológico de La Paz

Dr. Saúl Martínez Díaz - Instituto Tecnológico de La Paz

Dr. Israel Marcos Santillán Méndez - Instituto Tecnológico de La Paz

Dr. Marco Antonio Castro Liera - Instituto Tecnológico de La Paz

M.C. Jesús Antonio Castro - Instituto Tecnológico de La Paz

M.A.T.I. Luis Armando Cárdenas Florido - Instituto Tecnológico de La Paz

M.S.C. José Luís Gómez Torres - Instituto Tecnológico de La Paz

M.S.C. Diego García Molleda - Instituto Tecnológico de La Paz

M.S.C. César Higuera Verdugo - Instituto Tecnológico de La Paz

**Coordinadores de la edición**

M.S.C. Iliana Castro Liera - Instituto Tecnológico de La Paz

Dr. Mario Cortés Larrinaga - Instituto Tecnológico de La Paz

# Índice general

<b>Portada</b>	<b>I</b>
<b>Comités</b>	<b>v</b>
<b>Índice</b>	<b>VI</b>
Desarrollo de un sistema de control para un robot cartesiano de 3 grados de libertad mediante un sistema embebido para que funcione como impresora 3D . . . . .	1
Diseño y análisis cinemático de un robot paralelo para rehabilitación de Tobillo . . . . .	8
Experiencia de uso de un sistema de interacción humano-robot basado en dialogos gestuales . . . . .	15
Procesamiento de imágenes en hardware reconfigurable . . . . .	20
Ocultamiento de información en imágenes digitales mediante la esteganografía sin incrustación de datos . . . . .	24
Estimación de la matriz de cuantización de archivos previamente almacenados en formato JPEG . . . . .	29
Generación de imágenes HDR a partir de imágenes LDR . . . . .	34
Sintonización óptima fuera de línea del control disparado por eventos para un robot manipulador . . . . .	38
Diseño de un observador para un sistema hipercaótico utilizando programación genética . . . . .	45
Implementación de un sensorless control a un motor de inducción por medio de un observador de Luenberger adaptativo . . . . .	50
Diseño, construcción y control de un sistema carro-péndulo . . . . .	56
Control de movimiento para un robot con articulaciones flexibles basado en linealización por realimentación . . . . .	62
Control PID para la regulación de velocidad del péndulo con rueda inercial . . . . .	67
Control y simulación de un dispositivo de asistencia activo . . . . .	73
Identificación de ataques en redes de cómputo utilizando redes neuronales artificiales . . . . .	77
Identificación de huella plantar humana, utilizando descriptores de bajo costo computacional . . . . .	82
Reconocimiento invariante a rotación de texturas mediante la Transformada de Cúmulos Coordinados . . . . .	87
Detección y delimitación de grietas en estructuras de inmuebles utilizando procesamiento de imágenes . . . . .	93
Determinación de parámetros estadísticos de la pupila en el ojo humano . . . . .	98
Detección de raíz mediante técnicas de visión artificial . . . . .	104

Entrenamiento de una CNN para autoconducción de vehículo prototipo con Raspberry Pi . . . . .	110
VANT y fotogrametría digital para evaluar la morfodinámica costera . . . . .	115
Análisis comparativo de filtros para mediciones de sensores en aplicaciones de seguimiento de personas en interiores . . . . .	120
Sistema de generación de aplicaciones NFC para Android . . . . .	125
Formación de una red heterogénea de Robots mediante acoplamiento difusivo estático	130
Adaptación del vehículo submarino BlueROV2 para banco pruebas . . . . .	136
Identificación paramétrica de un quadrotor con validación experimental . . . . .	142
Diseño de un sistema de interacción natural con robots móviles utilizando gestos de la mano . . . . .	148
Fuzzy color description on Raspberry PI 3 . . . . .	154
Coficiente en FR basado en sistema difuso para ecuaciones diferenciales difusas .	158
Predicción de indicadores del Tecnológico Nacional de México con técnicas de Minería de Datos . . . . .	164
Identificación de patrones en antibiogramas mediante técnicas de aprendizaje automático . . . . .	169
Recoloring Ishihara plates with PSO algorithm and proposed equations . . . . .	174

# Desarrollo de un sistema de control para un robot cartesiano de 3 grados de libertad mediante un sistema embebido para que funcione como impresora 3D.

C. Jiménez, S. Vergara, M. A. Vargas, F. Reyes.

Facultad de Ciencias de la Electrónica opción en Automatización.

Benemérita Universidad Autónoma de Puebla.

Puebla, Pue. México.

caro.jimenez.martinez@hotmail.com

**Resumen**—En este trabajo se presenta el desarrollo mecánico, la cinemática directa y el modelo dinámico para un robot cartesiano de 3 grados de libertad (gdl). Se describe el desarrollo del sistema de control para dicho robot mediante un sistema embebido basado en una tarjeta FPGA (Field Programmable Gate Array) de la familia Cyclone V de Altera, el cual le permite controlar la posición de cada motor, al igual que le permite controlar la temperatura deseada de la cama caliente y del extrusor de material, de modo tal que el robot cartesiano de 3 gdl pueda funcionar como impresora 3D. Además, se presenta la caracterización del termistor que trae integrado tanto el extrusor de material como la cama caliente mediante el cual se puede determinar la temperatura actual del sistema para después cerrar el lazo de control de temperatura empleando el algoritmo de control tangente hiperbólica.

**Palabras clave**—FPGA, impresora 3D, control robot cartesiano.

**Abstract**—This paper presents the mechanical development, the direct kinematics and the dynamic model for a cartesian robot with 3 degrees of freedom (dof). The development of the control system for said robot is described by means of an embedded system based on an Altera Cyclone V FPGA (Field Programmable Gate Array), which allows it to control the position of each motor, as well as to control the temperature desired from the heatbed and the material extruder, so that the 3 dof cartesian robot can function as a 3D printer. In addition, the characterization of each thermistor is presented, which includes both the material extruder and the heatbed through which the current temperature of the system can be determined to then close the temperature control loop using a hyperbolic tangent control.

**Keywords**—FPGA, 3D printer, cartesian robot controller.

## I. INTRODUCCIÓN

La falta de desarrollo tecnológico en nuestro país es una de las principales causas de que muchas de las empresas mexicanas no puedan competir con empresas extranjeras, esto se debe a que mientras las empresas nacionales tienen que importar tecnología en muchos casos obsoleta, las empresas extranjeras usan tecnología de punta en sus procesos de producción. Por lo

anterior, es muy importante ganar experiencia en el desarrollo de máquinas que en un futuro puedan usarse en sistemas automatizados de producción.

Los robots manipuladores se encuentran mayormente en lo que es el área industrial ya que ofrecen ventajas considerables por el hecho de que pueden realizar tareas difíciles o peligrosas para el ser humano, así como también pueden realizar dichas tareas en menor tiempo y con mayor calidad y precisión. Dentro de los robots manipuladores se encuentra el robot cartesiano, el cual es uno de los dispositivos de mayor uso en las máquinas automatizadas gracias a que puede generar trayectorias complejas en un espacio tridimensional a partir de movimientos lineales independientes. De esta forma, pueden ser empleados para manipular objetos o en máquinas-herramienta de control numérico computarizado (CNC) como las cortadoras láser, perforadoras de placas, soldadura, centros de mecanizado o incluso como impresora 3D.

Las impresoras 3D se han vuelto muy populares hoy en día, gracias a las ventajas que ofrecen al poder realizar réplicas de modelos diseñados por ordenador convirtiéndolos en piezas 3D de distintos tamaños y formas según la aplicación para la cual se requieran. Debido a la versatilidad de este tipo de máquinas se les puede emplear casi en cualquier sector, por ejemplo, en los hospitales se ha empleado esta técnica de impresión 3D para la fabricación de órganos mediante materiales especiales, por otra parte, incluso se han desarrollado impresoras 3D para construir casas pequeñas de concreto.

Uno de los problemas más comunes que suelen tener las impresoras 3D de escritorio es que la posición de sus ejes se descalibra fácilmente debido a que funcionan mediante motores a pasos, es decir, en lazo abierto. Es por eso que se decidió realizar una impresora 3D que funcione mediante motores DC con encoders de cuadratura para determinar su posición actual (lazo cerrado), además de que se empleará una tarjeta FPGA para llevar a cabo el control, lo cual la vuelve única en su tipo, de modo que, mediante este sistema embebido se logren controlar los 3 gdl y la temperatura tanto de la cama caliente

como del extrusor de material, al igual que hacer monitoreo y diagnósticos del sistema en tiempo real.

## II. DIAGRAMA GENERAL DEL SISTEMA MECATRÓNICO

En la Fig. 1 se muestra el diagrama general del sistema mecatrónico, el cual se compone de un robot cartesiano de 3 gdl que puede ser manipulado por el usuario mediante comunicación wifi a través de una interfaz desarrollada en LabView donde se pueden modificar y visualizar los parámetros del sistema, como posiciones deseadas y ganancias de ajuste para los controladores.

Para que el robot se comunique con el sistema digital, se implementó una FPGA Cyclone V de la familia altera, la cual ejecutará los algoritmos de control para controlar la posición de los 3 gdl del robot cartesiano (x, y, z), así como también se encargará de mantener la temperatura deseada del extrusor de material y de la cama caliente donde se irá depositando el material extruido llevando a cabo la acción de impresora 3D.

En el caso de los 3 gdl, la FPGA se encargará de leer los encoders de cuadratura de cada servomotor para determinar su posición actual, aplicará la ley de control y mediante el generador de PWM (Pulse Width Modulation) mandará la señal a la etapa de potencia (puente H) de cada motor el cual proporcionará la energía necesaria a cada actuador para que se muevan a la posición deseada por el usuario.

En cuanto al extrusor de material y la cama caliente, estos traen integrado una resistencia eléctrica para calentarse y un termistor mediante el cual se determina la temperatura actual del dispositivo. Para que la FPGA pueda leer estas señales y procesarlas se debe hacer una conversión de la señal analógica a digital para lo cual se implementaron ADC0804 de 8 bits (uno para cada uno), se aplica la ley de control y mediante el generador de PWM se manda la señal de activación a la etapa de potencia (transistores de potencia), tanto del extrusor como de la cama caliente, proporcionando la energía necesaria para mantener cada actuador a la temperatura deseada por el usuario.

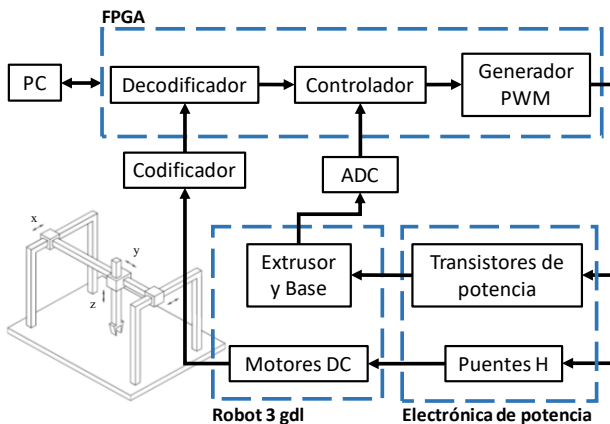


Fig. 1. Diagrama general del sistema mecatrónico.

## III. MECÁNICA DEL ROBOT

Se cuenta con un robot cartesiano de 3 gdl en el cual para poder realizar los movimientos prismáticos de este sistema con suma precisión se utilizaron rieles de alta precisión apoyados con carros lineales de precisión (Fig. 2a), esto con el fin de poder obtener el mejor desplazamiento lineal con mínima fricción en el sistema.

De igual manera, cuenta con tornillos embalados (Fig. 2b) para transformar el desplazamiento giratorio a lineal, ya que cuentan con la característica de reemplazar la fricción deslizante por una fricción de rodadura, con lo cual se incrementa la eficiencia del tornillo de 30% a un 90% ayudando a que el esfuerzo en los componentes motrices se vea disminuido significativamente y haciéndolos más durables.

También se le adaptaron interruptores de fin de carrera (Fig. 2c) a cada eje para evitar que el carro de alta precisión sobrepase los límites del área de trabajo y dañe el sistema mecánico.

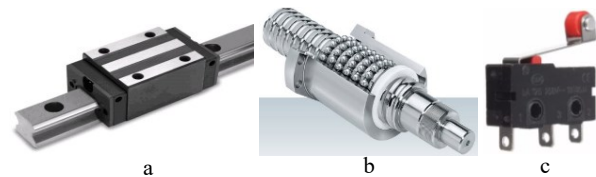


Fig. 2. Componentes mecánicos: a) Riel de alta precisión y carro de alta precisión, b) Tornillo embalado, c) Interruptor de fin de carrera.

Para los actuadores de los 3 gdl (ejes X, Y y Z) se utilizan motores reductores Z5D120-12 (Fig. 3a) que se alimentan a 12V, cuentan con la característica de proporcionar 3100 rpm (revoluciones por minuto) y generan un torque de 3-4 Nm [1]. Para poder ocupar los motores DC como servomotores tienen acoplado un encoder giratorio GTS06-VE-RA 1000 (Fig. 3b) de 1000 pulsos/revolución en la parte posterior de cada motor, mediante el cual se puede determinar la posición actual del actuador.

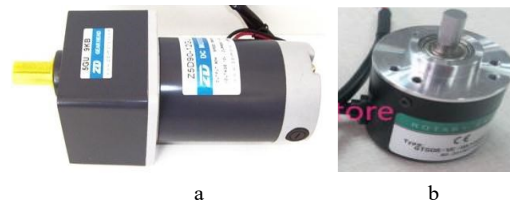


Fig. 3. Actuadores para los 3 gdl: a) Motor reductor, b) Encoder giratorio.

En cuanto a la herramienta de extrusión se implementó un extrusor de material Metal J-Head (Fig. 4a) el cual es alimentado con filamento de plástico (material a extruir, PLA) mediante un motor a pasos bipolar de cambio rápido de filamento, modelo EXTV22SMBU (Fig. 4b). Dicho extrusor cuenta con un termistor como sensor de temperatura de 100 K tipo NTC (Negative Temperature Coefficient), es decir, que al aumentar la temperatura su resistencia disminuye [2].

Finalmente, se le agregó una cama caliente MK3 (Fig. 4c) en la cual se irá depositando el filamento extruido para ir formando la pieza en 3D, esto con el fin de mantener el material caliente durante la impresión y así evitar pandeo en la pieza. Esta cama



tiene la característica de contar con fuente de alimentación doble, es decir, que se puede operar con 12V o 24V [3].

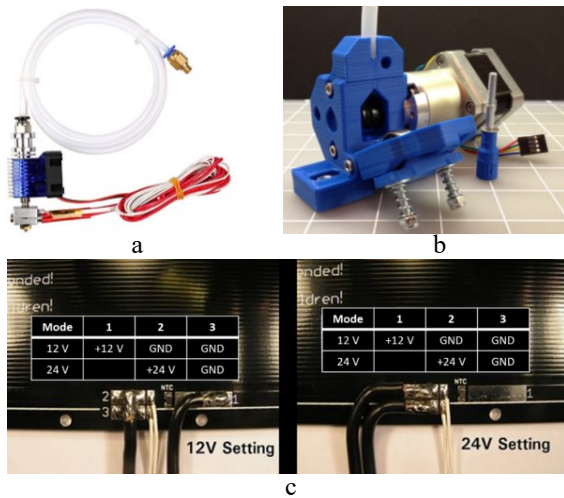


Fig. 4. Herramienta de extrusión: a) Extrusor de material, b) Alimentador de material, c) Cama caliente.

#### IV. DISEÑO DE LA CINEMÁTICA Y DINÁMICA

Es importante determinar el modelo dinámico para realizar un control más apropiado del robot y que además presente un mejor desempeño, de igual manera, es muy útil para analizar su comportamiento con diferentes algoritmos de control.

En la Fig. 5 se muestra la representación geométrica del robot cartesiano de 3 gdl, el cual permitirá visualizar el comportamiento del manipulador. Tanto el modelo de cinemática directa como el modelo dinámico se tomaron de [4].

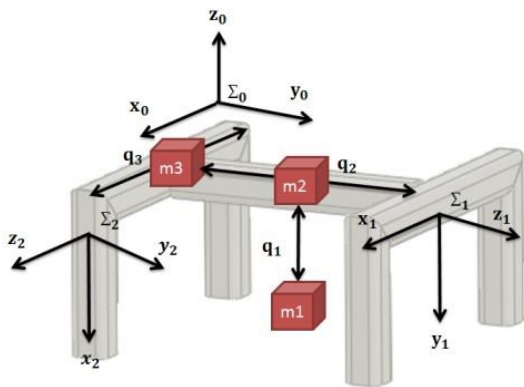


Fig. 5. Representación del robot cartesiano de 3 grados de libertad.

##### A. Cinemática

Para determinar la cinemática directa del robot cartesiano de 3 gdl se implementó la metodología de Denavit-Hartenberg [5], la cual ofrece un procedimiento sencillo para obtener la mínima representación de la orientación y la traslación de robots manipuladores y cuya estructura queda en representación de transformaciones homogéneas.

En la tabla I se presentan los parámetros de Denavit-Hartenberg.

TABLE I. PARÁMETROS DE DENAVIT-HARTENBERG.

Eslabón	$l_i$	$\alpha_i$	$d_i$	$\theta_i$
1	0	$-\pi/2$	$d_1$	0
2	0	$-\pi/2$	$d_2$	0
3	0	0	$d_3$	0

En donde, los cuatro parámetros Denavit-Hartenberg se definen de la siguiente manera:

- $l_i$  es la longitud del  $i$ -ésimo eslabón.
- $\alpha_i$  es el ángulo de torsión.
- $d_i$  es el desplazamiento de la articulación (offset).
- $\theta_i$  determina el ángulo de la articulación.

En la convención Denavit-Hartenberg cada transformación homogénea se representa mediante el producto de cuatro transformaciones básicas, como se muestra en (1).

$$H_{i-1}^i = H_{Rz_{i-1}}(\theta_i)H_{Tx_{i-1}}(d_i)H_{Tx_{i-1}}(l_i)H_{Rx_{i-1}}(\alpha_i) \quad (1)$$

La matriz de transformación homogénea para el robot cartesiano de 3 gdl queda como en (2).

$$H_0^3 = H_0^1 R_{z_0}^T H_1^2 H_2^3 = \begin{bmatrix} 0 & 0 & 1 & d_3 \\ 0 & -1 & 0 & d_2 \\ 1 & 0 & 0 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Por lo tanto, se tiene que el modelo de cinemática directa para el robot cartesiano de 3 gdl queda como se muestra en (3).

$$\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} d_3 \\ d_2 \\ d_1 \end{bmatrix} \quad (3)$$

##### B. Dinámica

El modelo dinámico cartesiano está dado por (4).

$$\tau = M(q)\ddot{q} + C(q,\dot{q})\dot{q} + g(q) + f_f(\dot{q}, f_e) \quad (4)$$

donde,

- $q \in \mathbb{R}^n$  es el vector de coordenadas generalizadas o posiciones articulares.
- $\dot{q} \in \mathbb{R}^n$  es el vector de velocidades articulares.
- $\ddot{q} \in \mathbb{R}^n$  es el vector de aceleraciones articulares.
- $M(q) \in \mathbb{R}^{n \times n}$  es la matriz de inercia, es simétrica y definida positiva.
- $C(q, \dot{q}) \in \mathbb{R}^{n \times n}$  es la matriz de fuerzas centrífugas y de Coriolis.
- $g(q) \in \mathbb{R}^n$  es el vector de fuerzas o pares gravitacionales.
- $f_f(\dot{q}, f_e) \in \mathbb{R}^n$  es el vector de pares de fricción que incluye la fricción viscosa, de Coulomb y estática.

Para este caso no se cuenta con la matriz de fuerzas centrípetas y de Coriolis ya que es igual a cero. Por lo tanto, el modelo dinámico para el robot de 3 gdl queda definido por (5).

$$\tau = \begin{bmatrix} m_1 + m_2 + m_3 & 0 & 0 \\ 0 & m_2 + m_3 & 0 \\ 0 & 0 & m_3 \end{bmatrix} \begin{bmatrix} \ddot{d}_1 \\ \ddot{d}_2 \\ \ddot{d}_3 \end{bmatrix} + g \begin{bmatrix} m_1 + m_2 + m_3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{bmatrix} \begin{bmatrix} \dot{d}_1 \\ \dot{d}_2 \\ \dot{d}_3 \end{bmatrix} + \begin{bmatrix} f_{c1} & 0 & 0 \\ 0 & f_{c2} & 0 \\ 0 & 0 & f_{c3} \end{bmatrix} \begin{bmatrix} \text{signo}(\dot{d}_1) \\ \text{signo}(\dot{d}_2) \\ \text{signo}(\dot{d}_3) \end{bmatrix} + \begin{bmatrix} f_{e1}[1 - |\text{signo}(\dot{d}_1)|] \\ f_{e2}[1 - |\text{signo}(\dot{d}_2)|] \\ f_{e3}[1 - |\text{signo}(\dot{d}_3)|] \end{bmatrix} \quad (5)$$

donde:

- $m_1, m_2, m_3$  son las masas de los servomotores.
- $b_1, b_2, b_3$  son los coeficientes de fricción viscosa.
- $f_{c1}, f_{c2}, f_{c3}$  representan los coeficientes de fricción de Coulomb.
- $f_{e1}, f_{e2}, f_{e3}$  son los coeficientes de fricción estática.

### V. ELECTRÓNICA DEL ROBOT

Para la etapa de potencia de cada motor se les diseñó un puente H en configuración Darlington que permita drenar 15A en modo saturación. Para la lectura de los encoders se hizo un divisor de voltaje, ya que los encoders manejan un voltaje de 5V en sus señales de salida y la FPGA solo maneja 3.3V, de esta manera, no se daña la FPGA y la lectura de las señales se logra hacer satisfactoriamente.

En la Fig. 6 se muestra el diagrama a bloques de la electrónica de control que se implementó tanto para el extrusor de material como para la cama caliente. Las etapas de potencia se diseñaron en base a transistores NPN en configuración Darlington para drenar 2.5A en modo saturación para el extrusor de material y 11.6A para la cama caliente.

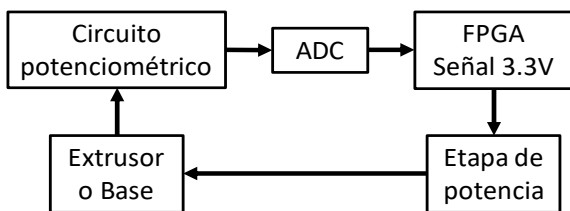


Fig. 6. Electrónica de control para el extrusor de material y la cama caliente.

Para determinar la temperatura actual del extrusor o de la cama caliente se hizo un circuito potenciométrico, formado por el termistor que trae integrado como sensor de temperatura y una resistencia colocados en serie a modo de divisor de voltaje, tomando el punto medio como medida de referencia. Después se caracterizó el termistor de ambos dispositivos variando el ciclo de trabajo de la señal PWM a una frecuencia de 100Hz y pasando 10 minutos de funcionamiento se registraba la temperatura actual del dispositivo (extrusor o base). Al graficar la respuesta de temperatura contra ciclo de trabajo, se puede observar que el comportamiento tanto del extrusor (Fig. 7) como de la cama caliente (Fig. 8) presenta una respuesta lineal.

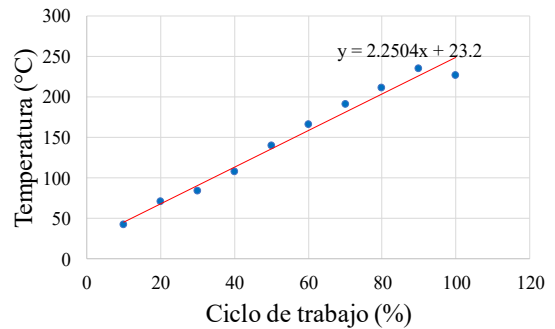


Fig. 7. Gráfica de linealidad del extrusor de material.

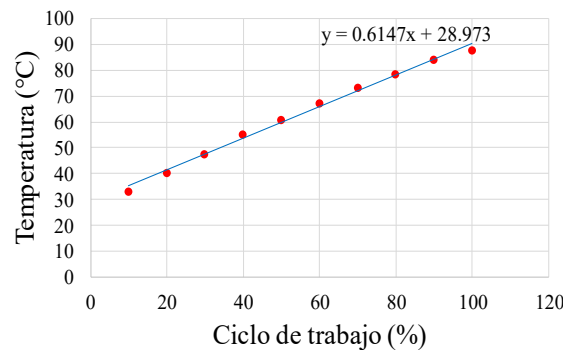


Fig. 8. Gráfica de linealidad de la cama caliente.

El circuito potenciométrico es alimentado por un voltaje que provoca una tensión de salida en el punto medio del divisor de voltaje, mediante el cual, al graficarlo con respecto a la temperatura se obtuvo la ecuación característica a la que trabaja el termistor y con la cual se puede determinar la temperatura a la que se encuentra el extrusor o la base.

La respuesta del termistor, tanto del extrusor de material como del de la cama caliente, Fig. 9 y Fig. 10 respectivamente, presentan un comportamiento no lineal al tener que su ecuación característica se adapta mejor con la curva de ajuste logarítmica.

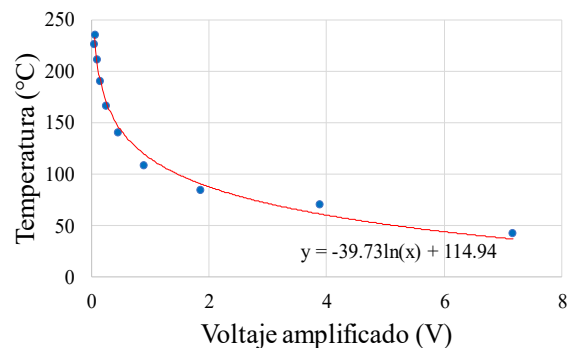


Fig. 9. Gráfica de respuesta logarítmica del termistor del extrusor de material.

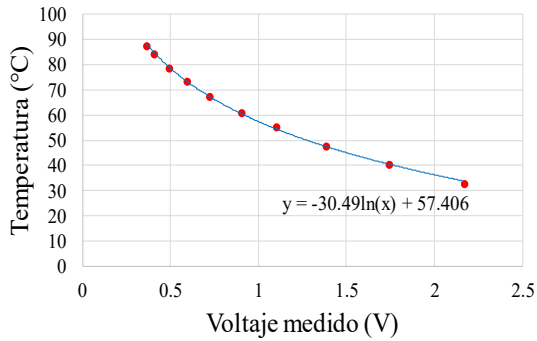


Fig. 10. Gráfica de respuesta logarítmica del termistor de la cama caliente.

Finalmente, para que la FPGA pueda interpretar la señal de los termistores, se utilizaron ADC0804 de 8 bits de resolución, cuyo circuito se diseñó con conexión en modo continuo [6].

## VI. IMPLEMENTACIÓN DEL CONTROL EN EL SISTEMA EMBEBIDO

### A. Firmware

Para llevar a cabo el control del robot se utilizó una tarjeta de desarrollo DE0-CV con FPGA de la familia Cyclone V de Altera, cuyo firmware necesario se desarrolló en el software Quartus II mediante lenguaje AHDL. Este firmware cuenta con protocolo de comunicación WIFI para poder tener comunicación entre una computadora y la tarjeta de adquisición de datos. Para realizar el control del robot se utilizó un microprocesador embebido en la FPGA desarrollado en la Maestría en Ciencias de la Electrónica opción en Automatización de la Facultad de Ciencias de la Electrónica de la Benemérita Universidad Autónoma de Puebla [7].

Este microprocesador se diseñó en base a arquitectura MIPS (Microprocessor without Interlocked Pipeline Stages), el cual es un procesador de 32 bits con arquitectura RISC (Reduced Instruction Set Computer) y cuenta con un banco de 32 registros de 32 bits. De igual manera, permite procesar las instrucciones recibidas en lenguaje ensamblador y realizar las correspondientes operaciones aritméticas, con el fin de llevar a cabo el control del robot. Además, a dicho firmware se le agregaron los bloques necesarios para procesar las señales de los encodes, de los ADC (Analog-Digital Converter) y también para generar las señales de control para cada actuador.

Para interpretar las señales A y B de los encoders, se le agregaron bloques para decodificar dichas señales, de igual manera, llevan a cabo el conteo de los pulsos de los encoders para poder determinar la posición actual de los actuadores que representan los 3 gdl y mediante dichos conteos se puede determinar el sentido de giro de cada motor.

Para interpretar y procesar las señales que se obtienen de los ADC y así determinar la temperatura actual de cada dispositivo (extrusor o base), se almacenan los 8 bits en un bus de 32 bits que posteriormente se convierten a punto flotante ya que el microprocesador solo puede leer datos en precisión simple.

Por otra parte, se le agregó el firmware necesario para realizar tanto los controles de posición como los controles de temperatura mediante PWM. En el caso de los 3 gdl, se manda la señal de PWM hacia los puentes H para activar los motores y según el conteo que se esté haciendo será el sentido de giro en cada motor. Para el extrusor de material y la cama caliente, la señal de PWM se manda hacia sus respectivas etapas de potencia y de esta manera, permitir que cada dispositivo empiece a calentarse.

### B. Software de control

#### 1) Control de posición.

Para realizar el control de posición de los 3 ejes del robot, se implementó un control cartesiano proporcional-derivativo (PD) más compensación de gravedad, como se muestra en la Fig. 11, ya que es el más conocido y empleado en el área de control automático. En este tipo de control, se tiene que la acción de control proporcional es directamente proporcional al error de posición y en cuanto a la acción de control derivativa, tiene el efecto de amortiguamiento o freno mecánico a través de la inyección de la velocidad articular, el cual tiene la finalidad de mejorar el desempeño del control proporcional [8].

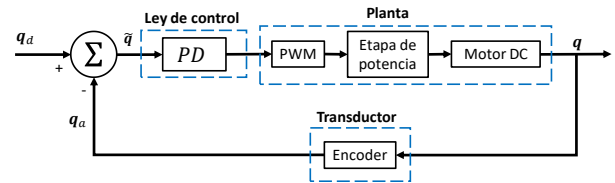


Fig. 11. Lazo cerrado para el control de posición.

Donde:

- $q_d$  es la posición cartesiana deseada (cm).
- $q_a$  es la posición cartesiana actual (cm).
- $\tilde{q}$  es el error de posición cartesiano ( $\tilde{q} = q_d - q_a$ ).

El control cartesiano PD está dado por (6):

$$\tau = K_p \tilde{q} - K_v \dot{q} + g(q) \quad (6)$$

donde,

- $K_p$  es la ganancia proporcional.
- $K_v$  es la ganancia derivativa.
- $\dot{q}$  es la velocidad de movimiento que se emplea para inyección de amortiguamiento.

La ecuación en lazo cerrado formada por el modelo dinámico cartesiano del robot de 3 gdl (5) y el algoritmo de control PD (6) está dada por (7):

$$\frac{d}{dt} \begin{bmatrix} \tilde{q} \\ \dot{\tilde{q}} \end{bmatrix} = \begin{bmatrix} \dot{\tilde{q}} \\ M^{-1}(q)[K_p \tilde{q} - K_v \dot{\tilde{q}} - C(q, \dot{q})\dot{\tilde{q}} - B\dot{\tilde{q}}] \end{bmatrix} \quad (7)$$

#### 2) Control de temperatura.

Para el control de temperatura, tanto del extrusor de material como de la cama caliente, se implementó un algoritmo de control tangente hiperbólico (*tanh*), ya que además de tener

sustento científico, una de las cualidades de este esquema de control es que permite operar al actuador (base o extrusor) en la parte lineal lejos de los límites de saturación.

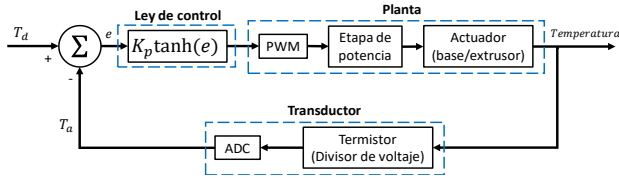


Fig. 12. Lazo cerrado para el control de temperatura.

Donde:

- $T_d$  es la temperatura deseada (°C).
- $T_a$  es la temperatura actual (°C).
- $e$  es el error de temperatura  $e = T_d - T_a$ .

Su algoritmo de control está dado por (8).

$$\tau = K_p \tanh(e) \tag{8}$$

donde  $K_p > 0$  y la función  $\tanh(e) = 0$  cuando  $e = 0$ .

### C. Software de usuario

#### 1) Compilador de posiciones deseadas.

El software necesario para que el robot cartesiano pueda realizar el proceso de impresión en 3D se describe en el diagrama de flujo de la Fig. 13.

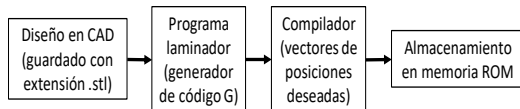


Fig. 13. Diagrama del software para impresión 3D.

El diseño que se desee imprimir en 3D se descarga de internet o se realiza mediante diseño asistido por computadora (CAD) y el archivo se guarda bajo la extensión *.stl*. Después, mediante un programa laminador, a partir del archivo *stl* se genera su código G el cual será introducido al compilador de posiciones deseadas (Fig. 14). Mediante la interfaz gráfica de LabView se diseña dicho compilador, en el cual se carga el archivo en código G “A” y cuyo código se puede visualizar en “B”. Después, hace una depuración para extraer las coordenadas de cada eje “C” y las va almacenando en vectores “D”. Finalmente, estos vectores se mandarán a RAM para llevar a cabo la acción de impresión en 3D.

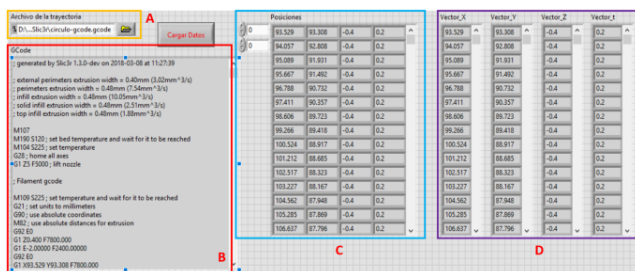


Fig. 14. Interfaz gráfica del compilador de posiciones deseadas.

#### 2) Interfaz de usuario.

Para llevar a cabo el control del robot, es decir, el control de posición de sus 3 gdl y los controles de temperatura del extrusor y la cama caliente, se utilizó una interfaz gráfica (Fig. 15) desarrollada en LabView [7], la cual mediante protocolo de comunicación wifi permite la comunicación entre la computadora y el microprocesador “A” embebido en la tarjeta de adquisición (FPGA), además trabaja a una velocidad de muestreo de 2.5ms.

En esta interfaz se carga un archivo que contiene el código “B” mediante el cual se realizará el control de todo el sistema, es decir, donde se cierran todos los lazos de control. Este código está basado en instrucciones en ensamblador para que el microprocesador las pueda interpretar y realizar las correspondientes operaciones aritméticas. En “C” se asignan las posiciones deseadas para cada gdl con sus respectivas ganancias proporcional y derivativa, así como también en “D” se asignan las temperaturas deseadas para el extrusor de material y la cama caliente con sus ganancias proporcionales. Finalmente, en “E” se ejecuta o se detiene el código de control.



Fig. 15. Interfaz de usuario para el control del robot.

## VII. RESULTADOS

En la Fig. 16 se presenta el robot cartesiano de 3 gdl que funcionará como impresora 3D, en el cual se muestra la estructura mecánica y la electrónica de potencia.

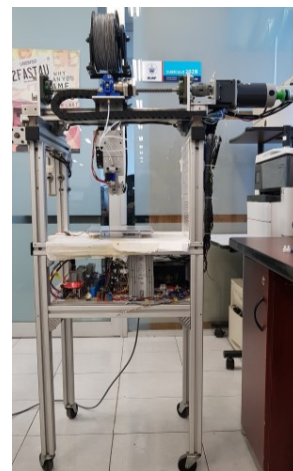


Fig. 16. Sistema mecatrónico para impresión en 3D.

Ya que el propósito de este robot es funcionar como impresora 3D, el método de impresión que llevará a cabo este robot será del tipo modelado por deposición fundida (FDM), el cual consiste en extruir en caliente un filamento de plástico, que para este caso se utilizará PLA. Este termoplástico presenta un rango de temperatura de impresión de 190-210 °C y para el caso de la cama caliente, se necesita que la cama se mantenga a temperaturas bajas, entre 30-50 °C.

A partir de estos rangos de temperatura, es decir, de 190-210 °C para el extrusor y de 30-50 °C para la cama caliente, se presentan las gráficas obtenidas al ejecutar los controles de temperatura para cada dispositivo. En la Fig. 17 se presenta la gráfica del comportamiento del extrusor al asignarle una temperatura deseada de 200 °C con una ganancia proporcional de 190.

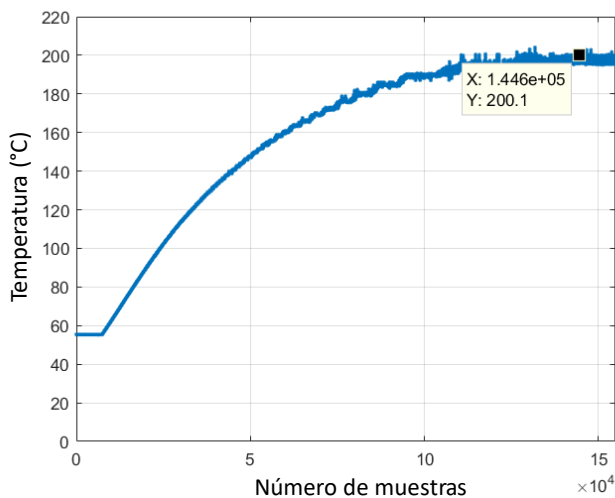


Fig. 17. Control de temperatura del extrusor de material a 200 °C.

En la Fig. 18 se muestra el comportamiento de la cama caliente al asignarle una temperatura deseada de 50 °C con una ganancia proporcional de 180.

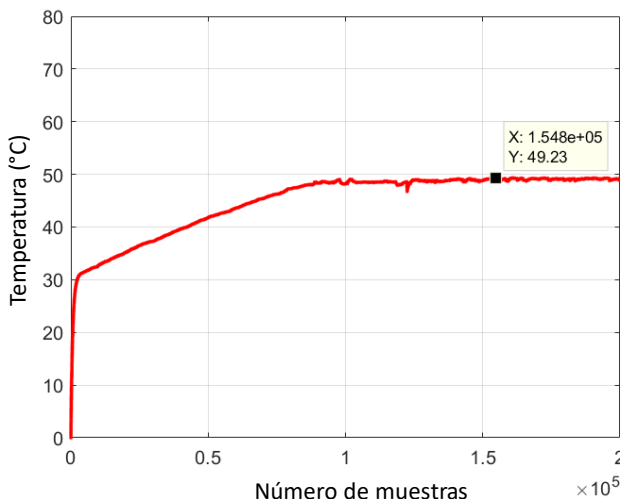


Fig. 18. Control de temperatura de la cama caliente a 50 °C.

## CONCLUSIONES

Realizar el control de temperatura, tanto para el extrusor de material como para la cama caliente, fue todo un reto ya que se trata de control no lineal, pero no porque la respuesta del control sea no lineal sino porque la naturaleza de los termistores es no lineal gracias a la respuesta logarítmica que presentan. Debido a esto, se tuvieron que hacer muchas pruebas y ajustes en el control de temperatura de cada planta, sobre todo para el control del extrusor, ya que el termistor presenta una respuesta en la que a mayor temperatura se tiene menor resolución, es por eso que para controlar a 200 °C se tuvo que amplificar el voltaje de salida para que el ADC pudiera leerlo con mayor facilidad, aún así se logró mantener ambos controles en el rango de temperatura que se necesita para cada uno (30-50 °C para la cama caliente, 190-210 °C para el extrusor de material).

La implementación de los circuitos de potencia, son de gran importancia ya que permiten drenar la corriente necesaria según la aplicación solicitada, en este caso, la etapa de potencia que se implementó para cada planta (base y extrusor) funciona de manera adecuada para el control de temperatura.

En cuanto al algoritmo de control empleado (*tanh*), al poner en funcionamiento cada sistema (extrusor o base), se logró mantener controlada la planta a la temperatura deseada teniendo un rango de error de  $\pm 1$  °C, lo cual no afecta al sistema ya que se encuentra dentro de los rangos de temperatura deseados.

Como trabajo a futuro se van a hacer pruebas para empezar a extruir el filamento y de esta manera poder concluir con la etapa de impresión en 3D.

## REFERENCIAS

- [1] Motores DC [online]. Disponible en: <https://longkang-motor.en.made-in-china.com/product/lozxPcJbOKYg/China-24V-DC-Gear-Motor-Z5D120-.html>
- [2] Extrusor Metal J-Head long distance [online]. Disponible en: [http://www.geeetech.com/wiki/index.php/Metal\\_J-head\\_long-distance](http://www.geeetech.com/wiki/index.php/Metal_J-head_long-distance)
- [3] Cama caliente PCB Heatbed [online]. Disponible en: [http://reprap.org/wiki/PCB\\_Heatbed](http://reprap.org/wiki/PCB_Heatbed)
- [4] A. López, S. Vergara, M. A. Vargas, F. Reyes, A. Palomino, "Design and construction of an acquisition system to control a Cartesian robot as a 3D printer," CIINDET, México, 2016.
- [5] Subir Kumar Saha., "Introducción a la Robótica," Mc Graw Hill, 1er ed., México, 2010, pp.100-106.
- [6] ADC0804 datasheet [online]. Disponible en: <https://www.intersil.com/content/dam/Intersil/documents/adc0/adc0803-04.pdf>.
- [7] G. Zepeda, S. Vergara, M. A. Vargas, F. Reyes, D. Pinto, D. Vilariño, "Diseño e implementación de un microprocesador RISC de 32 bits en una tarjeta con FPGA y comunicación inalámbrica," vol. 1, No. 1, Revista de Ciencias de la Ingeniería y Tecnología (RECIT), México, 2017.
- [8] F. R. Cortés, "Robótica, Control de robots manipuladores," Alfaomega Grupo Editor, 1st ed., México, 2011, pp.336-354.



# Diseño y Análisis Cinemático de un Robot Paralelo para Rehabilitación de Tobillo

Mauricio Arredondo Soto  
Departamento de Ingeniería Mecánica  
DICIS, Universidad de Guanajuato  
Salamanca GTO., México  
m.arredondosoto@ugto.mx

Marian G. Álvarez Pérez  
Departamento de Ingeniería Mecánica  
DICIS, Universidad de Guanajuato  
Salamanca GTO., México  
marian.alvarez.mec@gmail.com

Mario A. García Murillo  
Departamento de Ingeniería Mecánica  
DICIS, Universidad de Guanajuato  
Salamanca GTO., México  
garcia.mario@ugto.mx

J. Jesús Cervantes Sánchez  
Departamento de Ingeniería Mecánica  
DICIS, Universidad de Guanajuato  
Salamanca GTO., México  
jecer@ugto.mx

Agustín Vidal Lesso  
Departamento de Ingeniería Mecánica  
DICIS, Universidad de Guanajuato  
Salamanca GTO., México  
agustin.vidal@ugto.mx

**Resumen**—Los manipuladores paralelos han aparecido como una valiosa alternativa en la robótica para rehabilitación de tobillos, debido principalmente a la complejidad de esta articulación, así como su gran índice de lesiones. Sus ventajas en cuanto a movilidad, precisión y rigidez, con respecto a dispositivos del tipo órtesis, han hecho que hoy en día existan un número importante de trabajos relacionados. Sin embargo, la mayoría de las propuestas presentan restricciones de diseño que hacen poco viable su construcción.

Con el objetivo de solventar esta problemática, se propone un robot paralelo novedoso, de cuatro grados de libertad, con arquitectura 2-RPU + RRPS, capaz de llevar a cabo movimientos para rehabilitación pasiva que involucren la plantarflexión-dorsiflexión, inversión-eversión y aducción-abducción del tobillo. Así mismo, se analiza su comportamiento cinemático, en cuanto a posición, velocidad y aceleración, y por último se presenta un ejemplo numérico donde se contrastan resultados analíticos contra los del software especializado en simulación dinámica, MSC-Adams<sup>TM</sup>.

**Palabras clave**—Manipulador paralelo, rehabilitación de tobillo, cinemática, teoría de tornillos.

## I. INTRODUCCIÓN

El tobillo es una de las articulaciones más complejas del sistema motriz de los seres humanos, siendo el encargado de transmitir fuerzas y torques a la tierra durante la marcha [1]. No obstante, existen un sin número de factores que pueden conducir a la pérdida de algunas de sus funciones y/o a la reducción de su rango de movimiento.

La mayoría de las lesiones hacen que sea necesario someterse a un largo programa de rehabilitación para recuperar la fuerza de la articulación y mejorar su funcionamiento y movilidad. Muchos de estos ejercicios requieren de un alto grado de intervención por parte del terapeuta, al mismo tiempo que suelen implementarse para su realización, dispositivos bastante simples como: bandas elásticas, rodillos de espuma y tableros giratorios [2]. Estos aparatos presentan desventajas importantes como: a) no son interactivos, b) hay ejercicios que requieren de movimientos y posiciones muy específicos,

y estos dispositivos básicos no poseen la capacidad de llevarlos a cabo, c) no son capaces de dotar al terapeuta de información sobre el diagnóstico de la lesión, por lo que la evaluación de la tarea se deja sólo a la experiencia del terapeuta, al aspecto exterior de la articulación y a los comentarios del paciente.

Las desventajas de los sistemas convencionales han fomentado el creciente interés en el diseño y la creación de equipos automáticos para la terapia en cada una de sus modalidades.

Actualmente existen sólo unos pocos dispositivos automatizados disponibles comercialmente para la rehabilitación de tobillo, aunque esto puede deberse a varias razones, como la falta de un criterio de evaluación estandarizado que permita definir la eficiencia de un robot rehabilitador de tobillos [3], o la inexistencia de un único protocolo de rehabilitación que pueda considerarse el mejor [4], en lo que se refiere al diseño, la arquitectura supone un reto por sí misma. Así, la compleja biomecánica del tobillo dificulta la emulación realista de los movimientos [5], y la alta variabilidad en las dimensiones del tobillo en los pacientes [6] da lugar a la necesidad de un sistema que se adapte a distintos usuarios. Además, la mayoría de las arquitecturas propuestas se centran en la plantarflexión-dorsiflexión, tal que los movimientos de inversión-eversión y abducción-aducción, o bien no son tomados en cuenta, o dependen en gran medida de la ubicación del pie con respecto a los ejes de rotación de los robots. También, suelen presentar varias restricciones de alineación y un gran número de juntas costosas, lo que en términos generales, dificulta su posible fabricación.

En este trabajo de investigación, se propone una arquitectura novedosa sin grandes restricciones de diseño, capaz de realizar ejercicios de dorsiflexión-plantarflexión, inversión-eversión, abducción-aducción y sus movimientos combinados, haciendo posible el seguimiento de trayectorias complejas que lleven a una recuperación más completa del rango de movimiento.

## II. DESCRIPCIÓN DEL MANIPULADOR

El robot paralelo bajo estudio, mostrado en la figura 1, consiste en dos plataformas: una fija y una móvil, unidas a través de tres cadenas seriales, 2-RPU + RRPS.

La plataforma móvil tiene una configuración triangular con vértices ubicados en los puntos  $E_i$ , donde el subíndice  $i$  corresponde a la  $i$ -ésima extremidad, y se conecta a una plataforma fija definida por los puntos  $B_i$ . Junto con ellos, los puntos  $C_1, C_2, C_3$  y  $G_3$  representan la ubicación de los pares cinemáticos de todas las cadenas.

En las cadenas cinemáticas idénticas uno y dos, del tipo RPU, se actúa el par prismático, con origen en  $C_1$  y  $C_2$  respectivamente. En la tercera cadena, del tipo RRPS, los pares actuados son la segunda revoluta, ubicada en  $C_3$ , y el par prismático, que pasa por el punto  $G_3$ .

Con el fin de establecer la posición y orientación de la plataforma móvil con respecto al sistema de referencia coordenado  $XYZ$ , fijo en  $O$ , se ubica un segundo marco de referencia  $xyz$  sobre un punto  $P$  alojado en la plataforma móvil, como se observa en la figura 1.

Los cuatro grados de libertad del manipulador corresponden a movimiento plano general con respecto al vector unitario del eje  $Z$ ,  $\hat{k}$ , y una rotación en la dirección de  $y$ .

Así, a modo de garantizar dicha movilidad, los puntos  $C_1, C_2, E_1$  y  $E_2$  deben ser coplanares, además, es necesario que los ejes de revolución ubicados en  $B_1, B_2, B_3$  y el primer par de revoluta de cada junta universal ubicada en  $E_1, E_2$  sean perpendiculares a la base, mientras que la última revoluta de las juntas universales deben ser coplanares a los pares prismáticos de 1 y 2, y al mismo tiempo ser coaxiales, es decir, coincidir con la dirección de  $y$  del marco de referencia móvil. Por último, el eje de la revoluta actuada de la tercera extremidad debe ser paralela al plano de la base, así junto con la acción del par prismático de la cadena, esta extremidad controla la elevación del vértice  $E_3$ .

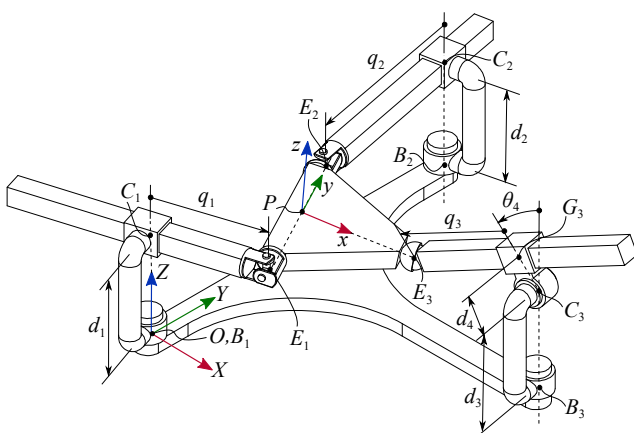


Fig. 1. Diagrama cinemático del Manipulador Paralelo 2-RPU+RRPS.

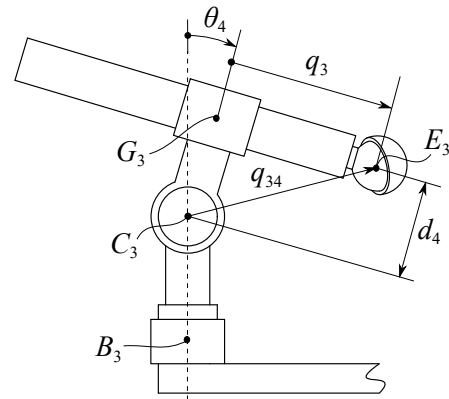


Fig. 2. Detalle de la cadena RRPS.

## III. ANÁLISIS DE POSICIÓN.

Inicialmente, de las dimensiones del manipulador, se pueden definir la posición de los puntos fijos  $B_i$  y  $C_i$ . Además se establecen las coordenadas generalizadas de los pares actuados  $\mathbf{q} = [q_1, q_2, q_3, \theta_4]^T$ . Las distancias entre los vértices de la plataforma móvil son también constantes y se definen a través de las longitudes  $e_{12}, e_{13}$  y  $e_{23}$ .

Las coordenadas  $q_1, q_2, q_3$ , por su parte, representan las distancias variables entre los puntos  $C_1-E_1$  y  $C_2-E_2$ , respectivamente para las cadenas 1 y 2, y la distancia entre  $G_3-E_3$  en la tercera extremidad. Mientras que  $\theta_4$  se mide con respecto al eje  $Z$  positivo del mismo sistema como se observa en la figura 2, y representa la rotación del par de revoluta actuado del manipulador.

Finalmente la posición del punto de estudio  $P$  con respecto a  $O$ , se describe a través del vector  $\mathbf{p} = [p_x, p_y, p_z]^T$ .

## A. Análisis directo de posición

El análisis directo de posición consiste en determinar los vectores de posición de los tres puntos  $E_1, E_2, E_3$ , ya que con esto se pueden obtener fácilmente el vector de posición de  $P$  y la orientación de la plataforma móvil, dadas las variables actuadas  $\mathbf{q} = [q_1, q_2, q_3, \theta_4]^T$ .

Los vectores de posición  $e_i$ , pueden definirse como:

$$\mathbf{e}_i = [X_i \ Y_i \ Z_i]^T, \quad i = 1, 2, 3.$$

Las figuras 1 y 2, sirven de apoyo para calcular de forma directa la componente en el eje  $Z$  de cada uno de los vértices de la plataforma móvil, como sigue:

$$Z_1 = d_1, \quad (1)$$

$$Z_2 = d_2 = d_1, \quad (2)$$

$$Z_3 = d_3 + d_4 \cos \theta_4 - q_3 \sin \theta_4. \quad (3)$$

De la figura 1 se observa que las longitudes de los pares prismáticos  $q_i$  están definidas como:

$$(\mathbf{c}_i - \mathbf{e}_i) \cdot (\mathbf{c}_i - \mathbf{e}_i) = q_i^2, \quad i = 1, 2. \quad (4)$$



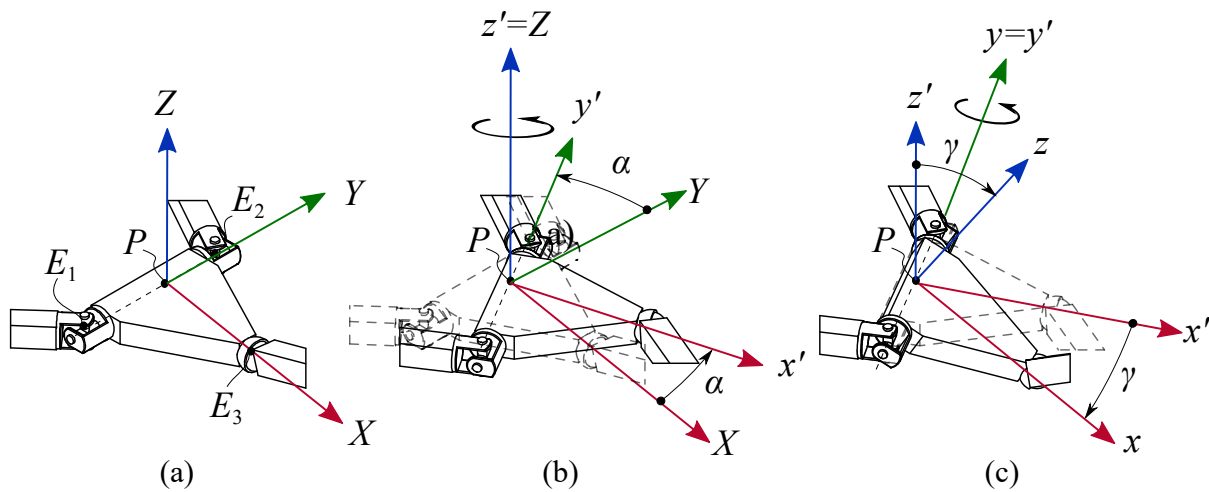


Fig. 3. Representación de los ángulos de Euler para el robot de rehabilitación en (a) posición inicial, y (b) primera rotación con respecto a  $Z$ , y (c) segunda rotación con respecto a  $y$ .

$$(\mathbf{e}_3 - \mathbf{e}_3) \cdot (\mathbf{e}_3 - \mathbf{e}_3) = q_{34}^2, \quad (5)$$

donde  $q_{34} = \sqrt{q_3^2 + d_4^2}$ . Además, las longitudes de la plataforma móvil  $e_{ij}$  entre los puntos  $E_1E_2E_3$  se pueden expresar como:

$$(\mathbf{e}_i - \mathbf{e}_j) \cdot (\mathbf{e}_i - \mathbf{e}_j) = e_{ij}^2, \quad (6)$$

$$(i, j) \in (1, 2), (2, 3), (1, 3)$$

Las ecuaciones (4), (5) y (6) forman un sistema de seis ecuaciones en las seis variables restantes,  $(X_i, Y_i, i = 1, 2, 3)$ , cuya solución determina completamente la pose de la plataforma móvil respecto a la fija. Entonces, el punto  $P$  se determina por el vector de posición:

$$\mathbf{p} = \frac{1}{2}(\mathbf{e}_1 + \mathbf{e}_2) \quad (7)$$

Finalmente, la pose de la plataforma móvil con respecto a la fija, puede sintetizarse en la matriz de transformación  $\mathbf{T}$ :

$$\mathbf{T} = \begin{bmatrix} \mathbf{R} & \mathbf{p} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \quad (8)$$

Donde  $\mathbf{R}$  es la matriz de rotación del sistema móvil con respecto al sistema fijo, definida como:

$$\mathbf{R} = [ \hat{\mathbf{u}}_x \quad \hat{\mathbf{u}}_y \quad \hat{\mathbf{u}}_z ], \quad (9)$$

y  $\hat{\mathbf{u}}_x, \hat{\mathbf{u}}_y, \hat{\mathbf{u}}_z$  son los vectores unitarios que definen la dirección y sentido de los ejes del marco de referencia  $xyz$  con respecto a  $XYZ$ .

El cálculo de estos vectores puede hacerse a través de las siguientes relaciones:

$$\hat{\mathbf{u}}_x = \frac{1}{h} (\mathbf{e}_3 - \mathbf{p}), \quad (10)$$

$$\hat{\mathbf{u}}_y = \hat{\mathbf{n}}, \quad (11)$$

$$\hat{\mathbf{u}}_z = \hat{\mathbf{u}}_x \times \hat{\mathbf{u}}_y, \quad (12)$$

donde  $\hat{\mathbf{n}}$  es el vector unitario que define la dirección y sentido del lado  $e_{12}$ ,

$$\hat{\mathbf{n}} = \frac{1}{e_{12}} (\mathbf{e}_2 - \mathbf{e}_1),$$

y la longitud  $h$  puede calcularse como:

$$h = \frac{1}{2} \sqrt{4e_{13}^2 - e_{12}^2}$$

### B. Descripción de la orientación de la plataforma móvil.

La matriz de rotación  $\mathbf{R}$  puede describirse descomponiendo la rotación total en varias rotaciones consecutivas. Para este mecanismo se utilizará la representación "ZYX" de los ángulos de Euler, la cual describe la matriz de rotación a través de tres rotaciones consecutivas con la forma:

$$\mathbf{R} = \text{rot}(\mathbf{z}, \alpha) \text{rot}(\mathbf{y}, \gamma) \text{rot}(\mathbf{x}, \psi). \quad (13)$$

Dado que el robot rehabilitador propuesto tiene únicamente dos grados de libertad rotacionales, y estas rotaciones coinciden con las dos primeras rotaciones en la descripción seleccionada, la tercera rotación  $\psi = 0$ . La figura 3 muestra las rotaciones de la plataforma móvil desde una orientación inicial, coincidente con el marco de referencia  $XYZ$  hasta su orientación final definida por el marco de referencia móvil  $xyz$ .

Así, la matriz de rotación  $\mathbf{R}$ , presentada en la ecuación (9), puede obtenerse en función de  $\alpha$  y  $\gamma$ , como sigue:

$$\mathbf{R} = \begin{bmatrix} \cos \alpha \cos \gamma & -\sin \alpha \cos \alpha \sin \gamma \\ \sin \alpha \cos \gamma & \cos \alpha \sin \alpha \sin \gamma \\ -\sin \gamma & 0 & \cos \gamma \end{bmatrix}, \quad (14)$$

### C. Análisis inverso de posición

Consiste en determinar los valores de las variables actuadas  $\mathbf{q}$  siendo conocida la pose de la plataforma móvil  $\mathbf{T}$ .

Los vectores de posición de los puntos  $E_1$ ,  $E_2$  y  $E_3$  se pueden calcular utilizando la expresión:

$$e_i = p + \mathbf{R}e_i^* \quad i = 1, 2, 3 \quad (15)$$

donde  $e_i^*$  son los vectores de posición de los puntos  $E_1$ ,  $E_2$  y  $E_3$  vistos desde el sistema de referencia móvil  $xyz$ .

Una vez conocidos los vectores de posición  $e_i$ ,  $i = 1, 2, 3$  los valores de las variables actuadas se pueden determinar fácilmente de las ecuaciones (3), (4) y (5).

#### IV. ANÁLISIS DE VELOCIDAD Y ACELERACIÓN

Los análisis de velocidad y aceleración del manipulador son resueltos mediante la teoría de tornillos. Los tornillos se modelan como se muestra en la figura 4. A menos que se indique lo contrario, estos están referidos al sistema fijo  $XYZ$ . El tornillo  ${}^a\mathcal{S}_i^b$  modela el movimiento de un cuerpo  $b$  con respecto a un cuerpo  $a$ , ambos pertenecientes a la cadena  $i$ . Nótese que el par esférico de la tercera pata, se descompone en tres pares de revoluta ( ${}^3\mathcal{S}_3^4, {}^4\mathcal{S}_3^5, {}^5\mathcal{S}_3^6$ ), cuyos ejes son ortogonales entre sí. Los tornillos asociados a los pares de revoluta se calculan como:

$$\mathcal{S} = \begin{bmatrix} \hat{s} \\ s_O \end{bmatrix} \quad (16)$$

donde  $\hat{s}$  es un vector unitario a lo largo del eje instantáneo del tornillo;  $s_O$  es el momento producido por el vector  $\hat{s}$  y el vector  $r_{O/P}$  el cual comienza en el punto  $P$  del eje instantáneo del tornillo y termina en el polo de referencia  $O$ .

Los tornillos para los pares prismáticos,  $q_i$ , del manipulador están dados como:

$$\mathcal{S} = \begin{bmatrix} \mathbf{0} \\ \hat{q}_i \end{bmatrix} \quad (17)$$

donde  $\hat{q}_i$ , es el vector unitario en la dirección del  $i$ -ésimo par prismático.

##### A. Análisis de Velocidad

El estado de velocidad de la plataforma móvil con respecto a la base,  $\mathbf{V}_O$ , puede ser representado como un tornillo infinitesimal,

$$\mathbf{V}_O = \begin{bmatrix} \boldsymbol{\omega} \\ v_O \end{bmatrix} = \sum_{m=0}^{n-1} m\omega_{m+1}^i m\mathcal{S}_i^{m+1} \quad (18)$$

donde  $\boldsymbol{\omega}$  y  $v_O$  son las velocidades angular y traslacional, respectivamente, de la plataforma móvil con respecto al polo de referencia  $O$ ; las  $m\omega_{m+1}^i$  representan las magnitudes de las velocidades de rotación y traslación del cuerpo  $b$  con respecto al  $a$  pertenecientes a la cadena  $i$ , y finalmente  $n$  es el número de pares cinemáticos de la cadena  $i$ .

La ecuación (18) puede ser escrita en forma matricial como:

$$\mathbf{V}_O = \mathbf{J}_i \boldsymbol{\Omega}_i \quad (19)$$

donde  $\mathbf{J}_i$ , es el Jacobiano de la  $i$ -ésima cadena cinemática y  $\boldsymbol{\Omega}_i$  es el vector de velocidades de los pares cinemáticos de dicha cadena.

El análisis directo de velocidad determina el estado de velocidad de la plataforma móvil con respecto a la base cuando

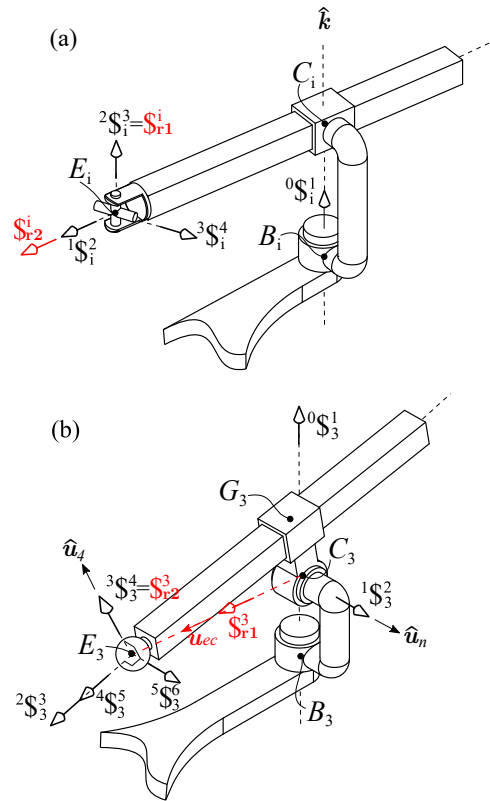


Fig. 4. Tornillos infinitesimales y recíprocos en (a) extremidades 1 o 2, y (b) extremidad 3.

las velocidades de los actuadores son conocidas. Un concepto útil para simplificar este análisis es el de reciprocidad. Dos tornillos  $\mathcal{S}_1$  y  $\mathcal{S}_2$  son recíprocos si la forma de Klein entre ellos es cero,  $\{\mathcal{S}_1; \mathcal{S}_2\} = 0$ .

Con esta definición puede demostrarse que el tornillo  ${}^2\mathcal{S}_1^3$  de las cadenas RPU el cual se renombrara,  $\mathcal{S}_{r1}^1$ , es recíproco a todos los tornillos asociados con los pares de estas cadenas. Entonces, aplicando la forma de Klein a ambos lados de la ecuación (18) para las cadenas 1 y 2, se obtiene:

$$\{\mathbf{V}_O; \mathcal{S}_{r1}^1\} = 0, \quad i = 1, 2. \quad (20)$$

Sin embargo, se busca un tornillo recíproco únicamente a los tornillos asociados a los pares cinemáticos pasivos. Un tornillo de revoluta en la dirección y sentido del par prismático lograría este objetivo, este tornillo se define como:

$$\mathcal{S}_{r2}^i = \begin{bmatrix} \hat{u}_i \\ e_i \times \hat{u}_i \end{bmatrix}, \quad i = 1, 2. \quad (21)$$

Aplicando la forma de Klein a ambos lados de la ecuación (18) para las cadenas 1 y 2, se obtiene:

$$\{\mathbf{V}_O; \mathcal{S}_{r2}^i\} = {}_1\omega_2^i = \dot{q}_1, \quad i = 1, 2. \quad (22)$$

En el caso de la tercera extremidad, RRPS, también deben determinarse dos tornillos recíprocos que permitan eliminar las velocidades de las juntas pasivas y mantener alguna de las velocidades de las juntas actuadas. Así, se define un tornillo

de revoluta,  $\mathbb{S}_{r1}^3$  con eje  $\hat{u}_{ec}$  y que pasa por el punto  $E_3$ , es decir:

$$\mathbb{S}_{r1}^3 = \begin{bmatrix} \hat{u}_{ec} \\ \mathbf{e}_3 \times \hat{u}_{ec} \end{bmatrix}, \quad (23)$$

Con lo que se obtiene:

$$\{\mathbf{V}_O; \mathbb{S}_{r1}^3\} = \{^2\mathbb{S}_3^3; \mathbb{S}_{r1}^3\} \dot{q}_3 = \nu_1 \dot{q}_3, \quad (24)$$

donde  $\nu_1$  representa un coeficiente determinado en función del producto escalar entre los vectores unitarios  $\hat{u}_{ec}$  y  $\hat{u}_3$ , es decir, el ángulo entre ambos vectores.

Con el fin de determinar la razón de velocidad en el par actuado restante,  $\theta_4$ , se evalúa un tornillo de revoluta,  $\mathbb{S}_{r2}^3 = {}^3\mathbb{S}_3^4$ , que a través de la forma de Klein con la ecuación (18) da como resultado:

$$\{\mathbf{V}_O; \mathbb{S}_{r2}^3\} = \{^1\mathbb{S}_2^3; \mathbb{S}_{r2}^3\} \dot{\theta}_4 = \nu_2 \dot{\theta}_4, \quad (25)$$

donde el coeficiente  $\nu_2$  es una relación con la distancia mínima entre los puntos  $E_3$  y  $C_3$ .

Agrupando las expresiones (20), (22), (24) y (25) en forma matricial, el estado de velocidad puede calcularse con la ecuación siguiente:

$$\mathbf{M}^T \Delta \mathbf{V}_O = \dot{\mathbf{Q}}, \quad (26)$$

donde  $\mathbf{M} = [{}^1\mathbb{S}_{r1}^1; {}^2\mathbb{S}_{r1}^2; {}^1\mathbb{S}_{r2}^2; {}^2\mathbb{S}_{r2}^3; {}^3\mathbb{S}_{r1}^3; {}^3\mathbb{S}_{r2}^3]$ ;  $\Delta = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}$  es un operador de polaridad definido por matrices  $3 \times 3$ , identidad  $\mathbf{I}$  y de ceros  $\mathbf{0}$ , y finalmente  $\dot{\mathbf{Q}} = [0, 0, \dot{q}_1, \dot{q}_2, \nu_1 \dot{q}_3, \nu_2 \dot{\theta}_4]^T$ .

La ecuación (26) puede manipularse para obtener una relación entrada salida entre las velocidades activas del manipulador,  $\dot{\mathbf{Q}}^a = [\dot{q}_1, \dot{q}_2, \dot{q}_3, \dot{\theta}_4]^T$ , y el estado de velocidad,  $\mathbf{V}_O$ , como sigue:

$$\mathbf{M}^T \Delta \mathbf{V}_O = \mathbf{B} \begin{bmatrix} 0 \\ 0 \\ \dot{\mathbf{Q}}^a \end{bmatrix} \quad (27)$$

Donde  $\mathbf{B} = \text{diag}(1, 1, 1, 1, \nu_1, \nu_2)$ .

Obviamente  $\det(\mathbf{M}^T) \neq 0$ , de otra forma el manipulador estaría en una posición singular. Entonces, el estado de velocidad puede ser calculado sin la necesidad de conocer las velocidades de las juntas pasivas, lo cual representa una ventaja significativa sobre otros modelos.

Habiendo obtenido el estado de velocidad, la velocidad traslacional del punto de estudio,  $\mathbf{v}_p$ , puede calcularse mediante cinemática clásica como sigue:

$$\mathbf{v}_p = \mathbf{v}_O + \boldsymbol{\omega} \times \mathbf{p} \quad (28)$$

Por otra parte, el análisis inverso de velocidad consiste en calcular las velocidades articulares,  ${}^m\omega_{m+1}^i$ , para un estado de velocidad determinado. Entonces, las velocidades de las juntas de cada extremidad del manipulador paralelo pueden determinarse mediante la ecuación (19).

## B. Análisis de Aceleración

En 1947 Brand [7] introdujo el concepto de estado de aceleración reducida,  $\mathbf{A}_O$ , de un cuerpo rígido:

$$\mathbf{A}_O = \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{a}_O - \boldsymbol{\omega} \times \mathbf{v}_O \end{bmatrix} \quad (29)$$

donde  $\boldsymbol{\alpha} = \dot{\boldsymbol{\omega}}$  y  $\mathbf{a}_O$  son la aceleración angular y lineal del efector final, respectivamente. Además, el estado de aceleración puede escribirse en términos de tornillos como sigue:

$$\mathbf{A}_O = \mathbf{J}_i \dot{\boldsymbol{\Omega}}_i + \mathcal{L}_i \quad (30)$$

donde  $\dot{\boldsymbol{\Omega}}_i$  el vector de aceleraciones de los pares cinemáticos de la  $i$ -ésima cadena,  ${}^m\omega_{m+1}^i$ . Además,  $\mathcal{L}_i$  es el tornillo de los productos de Lie, el cual está dado por:

$$\mathcal{L}_i = \sum_{j=0}^{n-2} \begin{bmatrix} j\omega_{j+1}^i & j\mathbb{S}_i^{j+1} \\ \sum_{k=j+1}^{n-1} k\omega_{k+1}^i & k\mathbb{S}_i^{k+1} \end{bmatrix} \quad (31)$$

donde  $n$  es el número de pares cinemáticos de la cadena  $i$ .

Realizando el producto de Klein con los tornillos recíprocos obtenidos en el análisis de velocidad a ambos lados de la ecuación (30) y agrupando matricialmente, se obtiene:

$$\mathbf{M}^T \Delta \mathbf{A}_O = \mathbf{B} \begin{bmatrix} 0 \\ 0 \\ \ddot{q}_1 \\ \ddot{q}_2 \\ \ddot{q}_3 \\ \ddot{\theta}_4 \end{bmatrix} + \begin{bmatrix} \{^1\mathbb{S}_{r1}^1; \mathcal{L}_1\} \\ \{^2\mathbb{S}_{r1}^2; \mathcal{L}_2\} \\ \{^1\mathbb{S}_{r2}^2; \mathcal{L}_1\} \\ \{^2\mathbb{S}_{r2}^3; \mathcal{L}_2\} \\ \{^3\mathbb{S}_{r1}^3; \mathcal{L}_3\} \\ \{^3\mathbb{S}_{r2}^3; \mathcal{L}_3\} \end{bmatrix} \quad (32)$$

Mas aun, la aceleración traslacional del punto  $P$  de la plataforma móvil está determinada como:

$$\mathbf{a}_P = \mathbf{a}_O + \boldsymbol{\alpha} \times \mathbf{p} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{p}) \quad (33)$$

El cálculo del estado de aceleración reducida,  $\mathbf{A}_O$ , usando la ecuación (32), no requiere conocer los valores de las aceleraciones de las juntas pasivas.

El análisis inverso de aceleración consiste en determinar las aceleraciones de las juntas del manipulador para un determinado,  $\mathbf{A}_O$ . Estas pueden encontrarse para cada extremidad despejando de la ecuación (30) como sigue:

$$\dot{\boldsymbol{\Omega}}_i = (\mathbf{J}_i)^{-1} (\mathbf{A}_O - \mathcal{L}_i) \quad (34)$$

## V. EJEMPLO NUMÉRICO

En esta sección se presenta como caso de estudio, un ejemplo numérico de los análisis de velocidad y aceleración directos del manipulador paralelo, en el cual se comparan las gráficas obtenidas mediante la programación de las ecuaciones obtenidas en la sección IV en un software de análisis matemático y las obtenidas mediante simulación en el software especializado MSC-Adams<sup>TM</sup>.

Los parámetros geométricos del robot se encuentran listados en la tabla I. Todas las dimensiones están dadas en milímetros ( $mm$ ).

Adicionalmente, en la tabla II se describe la posición inicial de los puntos  $E_1$ ,  $E_2$  y  $E_3$ .

TABLA I  
 PARÁMETROS GEOMÉTRICOS DEL MANIPULADOR

$\mathbf{b}_1 = (0, 0, 0)$
$\mathbf{b}_2 = (0, 600, 0)$
$\mathbf{b}_3 = (400, 300, 0)$
$\mathbf{c}_1 = (0, 0, 300)$
$\mathbf{c}_2 = (0, 600, 300)$
$\mathbf{c}_3 = (400, 300, 300)$
$e_{12} = e_{13} = e_{23} = 300; d_4 = 65$

 TABLA II  
 POSICIÓN INICIAL DE LOS PUNTOS  $E_i$ 

$\mathbf{e}_1 = (96.8447, 151.7270, 300)$
$\mathbf{e}_2 = (101.9903, 451.6829, 300)$
$\mathbf{e}_3 = (332.9459, 297.6988, 413.7912)$

Asumiendo que, para  $t = 0 - 5s$ , las variables actuadas están determinadas por las funciones:

$$q_1 = 180 + 20 \cos\left(\frac{\pi}{8}t^2\right), \quad q_2 = 180 + 20 \sin\left(\frac{\pi}{10}t\right)$$

$$q_3 = 115 + 25 \sin\left(\frac{\pi}{10}t^2\right), \quad \theta_4 = 330 + 15 \sin\left(\frac{\pi}{8}t\right)$$

donde la función  $\theta_4$  está dada en grados.

Finalmente, las gráficas de comparación se presentan en las figuras 5-8 y están estructuradas de la siguiente manera; a) gráfica obtenida mediante simulación dinámica, y a) gráfica obtenida mediante programación de las ecuaciones.

## VI. CONCLUSIONES

Las expresiones obtenidas en los análisis de velocidad y aceleración son lineales, sencillas y compactas mediante el uso de la forma de Klein del algebra de Lie. Además, una ventaja significativa de esta formulación es que no se necesita conocer las velocidades y aceleraciones de las juntas pasivas para resolver los análisis directos de velocidad y aceleración, respectivamente. Esto a su vez puede presentar una reducción importante en los tiempos de computo cuando se realiza el

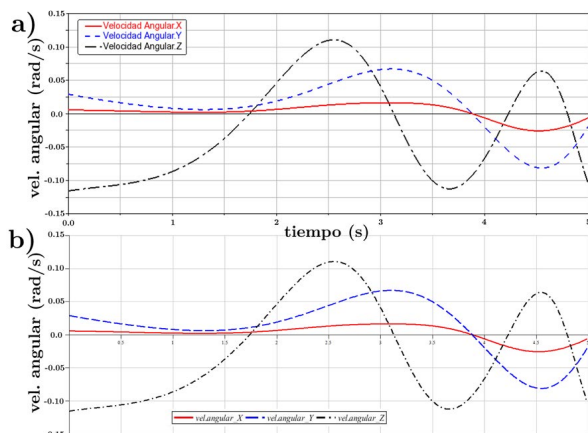


Fig. 5. Velocidad angular de la plataforma móvil

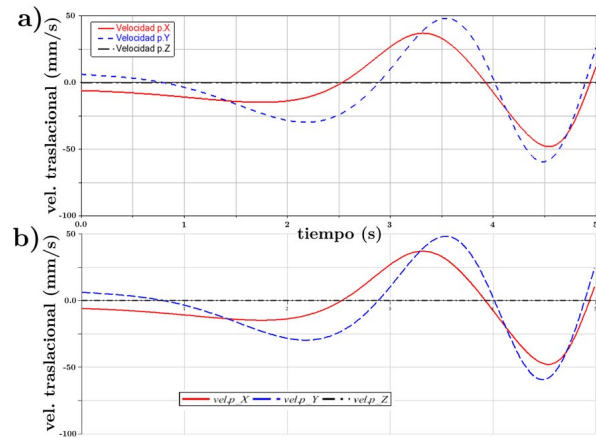
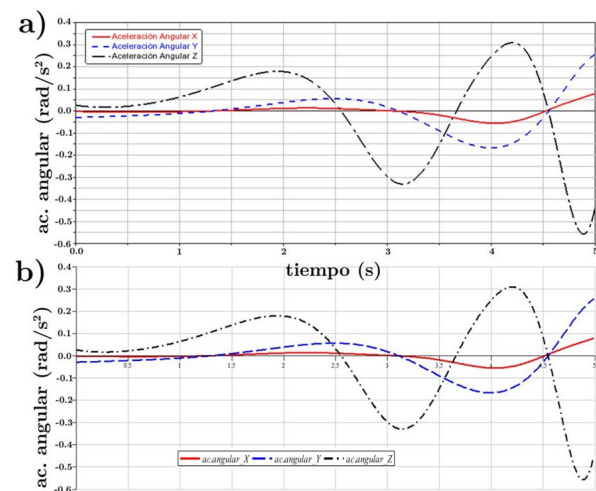
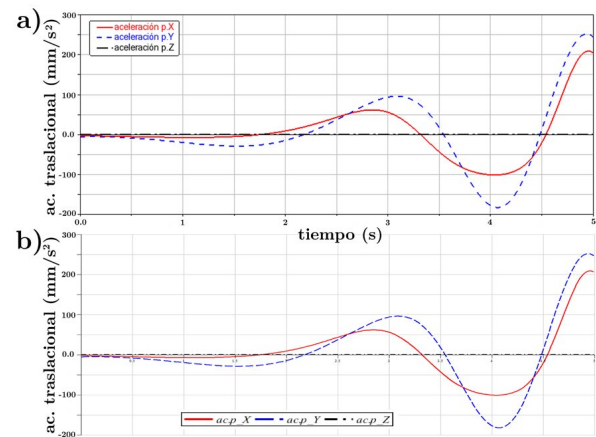

 Fig. 6. Velocidad traslacional del punto  $P$ 


Fig. 7. Aceleración Angular de la plataforma móvil


 Fig. 8. Aceleración traslacional del punto  $P$ .

control del robot y facilita la programación de trayectorias de

rehabilitación complejas.

Adicionalmente, los análisis de velocidad y aceleración dejan los cimientos para los análisis; cinético y generación de trayectorias, los cuales a su vez, servirían para incluir rehabilitación activa dentro de las características del manipulador.

#### REFERENCIAS

- [1] W. Jiacong, "Rehabilitation of ankle and foot," *Chin. J. Rehabil. Theory Pract.*, vol. 12, no. 14, pp. 1197–1198, 2008.
- [2] C. Syrseloudis and I. Emiris, "A parallel robot for ankle rehabilitation-evaluation and its design specifications," in *BioInformatics and BioEngineering, 2008. 8th IEEE International Conference, 2008*, pp. 1–6.
- [3] M. Zhang, T. Davies, and S. Xie, "Effectiveness of robot-assisted therapy on ankle rehabilitation: a systematic review," *Journal of Neuroengineering and Rehabilitation*, vol. 10, no. 1, pp. 1–16, 2013.
- [4] H. Krebs, J. Palazzolo, L. Dipietro, M. Ferraro, J. Krol, K. Rankeleiv, B. Volpe, and N. Hogan, "Rehabilitation robotics: Performance-based progressive robot-assisted therapy," *Autonomous Robots*, vol. 15, no. 1, pp. 7–20, 2003.
- [5] S. Carbes, S. Telfer, J. Woodburn, M. Oosterwaal, and J. Rasmussen, "A new multisegmental foot model and marker protocol for accurate simulation of the foot biomechanics during walking," in *Congress of the International Society of Biomechanics (ISB 2011), Brussels, Belgium, July, 2011*, pp. 3–7.
- [6] C. Barnett and J. Napier, "The axis of rotation at the ankle joint in man. its influence upon the form of the talus and the mobility of the fibula," *Journal of Anatomy*, vol. 86, no. 1, pp. 1–9, 1952.
- [7] L. Brand, *Vector and tensor analysis*. Wiley, 1947.

# Experiencia de uso de un sistema de interacción humano-robot basado en diálogos gestuales

Gustavo Castellón-Ramírez, Karina Reyes-Lío, Cristina Ramírez-Fernández, Eddie Clemente

Tecnológico Nacional de México/I.T. Ensenada

Maestría en Ciencias en Ingeniería Mecatrónica

Blvd. Tecnológico 150, Ex-ejido Chapultepec,

Ensenada, B.C., México 22780

Email: {gcastillon, sreyes, cramirez, eclemente}@ite.edu.mx

**Resumen**—Se presenta la evaluación de usabilidad y experiencia de uso de un sistema para la teleoperación de un robot móvil Xiao R Geek DS por medio de la interfaz natural Leap Motion para la coordinación vista-mano. La evaluación piloto con 32 sujetos, sugiere que el sistema es percibido como útil y usable, mientras proporciona experiencias de uso categorizadas como buenas. Además, la interacción con la interfaz natural es percibida como sencilla de usar y fácil de aprender. Por último, se generan sugerencias de diseño y algunos dominios de aplicación del sistema propuesto.

## I. INTRODUCCIÓN

La usabilidad de un sistema consiste en asegurar que es fácil de usar, eficiente en el uso y que genera una experiencia de usuario de disfrute [1]. En el diseño del sistema existen diversos estilos de interacción basados en la manipulación directa, el seguimiento de menús, el llenado de formas, la escritura de comandos, entre otros [2]. De acuerdo a Preece, Rogers y Sharp [3], la usabilidad y la experiencia de uso de un sistema, depende en gran medida del estilo de interacción y la retroalimentación proporcionada a las acciones de los usuarios. En particular, el estilo de interacción basado en una manipulación directa, permite a los usuarios disminuir la curva de aprendizaje y ampliar el espectro de uso de los sistemas [1].

En una manipulación directa el usuario puede utilizar interfaces llamadas “naturales” para interactuar directamente con el sistema [4], [5]. Actualmente, existen diversas interfaces naturales basadas en el reconocimiento de gestos o el reconocimiento de la voz, p. ej. [6], [7].

Una de las principales aplicaciones de las interfaces naturales es la manipulación a distancia de dispositivos robóticos que funcionan con los principios de la teleoperación [8], [9], [7]. Se ha visto que la teleoperación ayuda a completar tareas de manera satisfactoria en aplicaciones como la medicina [10], la búsqueda urbana y el rescate [11], el transporte [12], la agricultura inteligente [13], entre otros.

Particularmente, en las interfaces naturales basadas en gestos se efectúa la interpretación de movimientos que sirven como la entrada del sistema [14]. Cuando los gestos se centran en las manos, se utilizan sistemas de visión o sensores diversos como interfaces naturales [4]. Por ejemplo, Leap Motion es un sensor de gestos que permite la interpretación del movimiento

de las manos a través de imágenes [5]. El sensor Leap Motion se ha utilizado en la teleoperación de robots industriales para mejorar la robustez de las operaciones a distancia [8], [9]. Otra de las aplicaciones del sensor Leap Motion es el control de manos biónicas capaces de imitar, emular o simular los movimientos de una mano real [15]. También se ha utilizado en el control de drones [16], brazos robóticos [17], robots [18], entre otros. Sin embargo, a pesar de la funcionalidad que han mostrado estos sistemas de teleoperación por medio del sensor Leap Motion, hace falta evidencia de la percepción de los usuarios con respecto a la usabilidad y experiencia de uso generada.

La contribución de este trabajo es la evaluación de usabilidad y experiencia de uso de un sistema de teleoperación de un robot Xiao R Geek DS con una interfaz natural humano-robot para realizar tareas de coordinación vista-mano basadas en los gestos de la mano. El sistema hace uso de un enrutador que administra los mensajes entre el servidor y el sensor de gestos Leap Motion acorde a los movimientos del usuario. La evaluación piloto del sistema es realizada mediante la interfaz natural Leap Motion con la cual el usuario teleopera el robot móvil a través de un laberinto. Los resultados preliminares de la evaluación con 32 participantes, sugieren que el sistema tiene una percepción alta de usabilidad y buena experiencia de uso. Además, hay evidencia sobre la facilidad de uso y facilidad de aprendizaje de la manipulación directa usando el sensor Leap Motion. Por último, de acuerdo a los participantes, se generan sugerencias de diseño y algunos dominios de aplicación del sistema propuesto.

Este documento está organizado de la siguiente manera. En la Sección II se describen brevemente los componentes, las características técnicas y el funcionamiento del sistema propuesto. La evaluación de usabilidad y experiencia de uso se presenta en la Sección III. La Sección IV describe los resultados de la evaluación según las predicciones formuladas. Por último, en la Sección V se presentan las conclusiones y el trabajo futuro.

## II. DESCRIPCIÓN DEL SISTEMA

La función del sistema es la de brindar al usuario la capacidad de teleoperar un robot móvil mediante el uso de su mano dominante y una simple combinación de gestos.

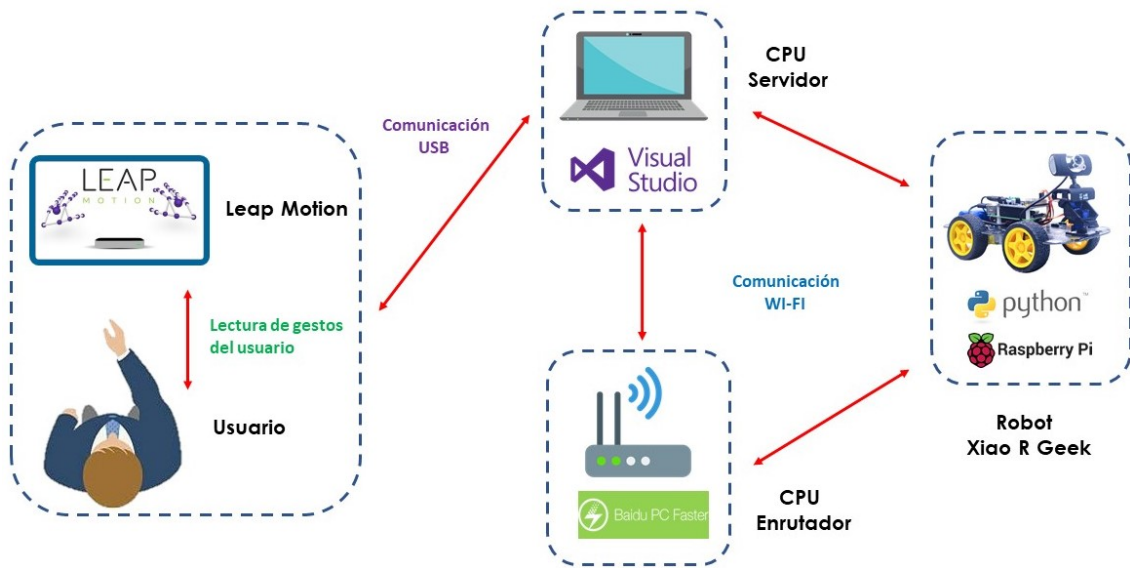


Figura 1: Arquitectura del sistema de interacción natural de teleoperación de un robot móvil.

Este sistema está compuesto por un enrutador, un servidor, un sensor de gestos Leap Motion y un robot Xiao R Geek DS Wireless WiFi Robot, en la Figura 1 se muestra la arquitectura del sistema.

La red “inalámbrica” es provista localmente por un enrutador instalado en una computadora portátil, a través de la aplicación *Baidu PC Faster*. La comunicación del sistema se realiza a través del protocolo TCP (Transmission Control Protocol), este protocolo garantiza que los datos sean transportados, y que el envío de información sea en el mismo orden que el de llegada. Para este tipo de protocolo se requieren dos tipos de usuarios, un servidor y uno o varios clientes. Para el desarrollo del sistema se estableció el servidor en un ordenador portátil, en una aplicación desarrollada en el lenguaje C Sharp; mientras que la aplicación para el cliente se desarrolló en Python. Lo anterior, debido a que el robot Xiao R Geek cuenta con el sistema operativo Raspbian que permite una interacción con los periféricos del robot a través de aplicaciones desarrolladas en Python3.

El servidor además de encargarse de la comunicación inalámbrica, se encarga de la interpretación de gestos en tiempo real realizados por la mano del usuario mediante el sensor Leap Motion. El sensor Leap Motion es un traductor de signos que cuenta con dos cámaras internas de una gran velocidad de captura [9]. Los gestos interpretados son mostrados en la Figura 2.

Además, el sensor Leap Motion cuenta con dos cámaras y tres LED infrarrojos que rastrean la luz infrarroja. El área de interacción de estas dos cámaras son dos pies sobre el dispositivo con un radio de dos pies a diferentes ángulos de detección. Con respecto al software Leap Motion aplica algoritmos avanzados a los datos sin procesar del sensor, un algoritmo de seguimiento interpreta los datos 3D e infiere las posiciones de las manos entre otros datos.

El robot es un Xiao R Geek DS FPGA tipo diferencial que cuenta con un sistema incrustado Raspberry Pi 3B, el cual utiliza un módulo BCM43438 para la comunicación WiFi LAN (Local Area Network). Además, el robot cuenta con una tarjeta diseñada con un driver L298P para motores DC de 12V y un reductor de voltaje SY8008. El robot también usa un paquete de baterías de litio 18650 con una placa de protección de alta corriente con especificaciones de 7.4V y 2200 mAh. Con este voltaje se alimentan 4 motores reductores DC de 6-12V del robot. Para un mejor control de los movimientos, los motores se mueven a menor velocidad durante los comandos de giro que son realizados sobre el eje del robot.

### III. EVALUACIÓN PILOTO

El estudio se condujo con estudiantes universitarios que participaron en una evaluación de usabilidad y experiencia de uso del sistema. En el estudio de usabilidad, participaron estudiantes de 18 a 27 años, con género, carrera y estatus

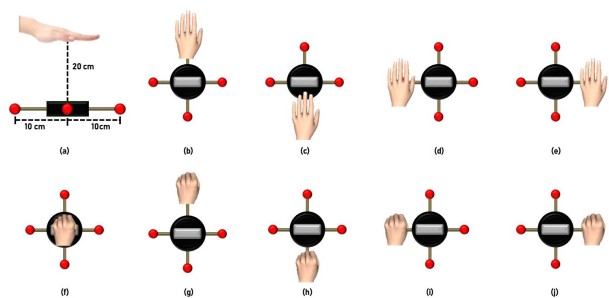


Figura 2: (a) parámetros de reconocimiento de gestos, (b) posición para avanzar, (c) posición para retroceder, (d) posición para girar a la izquierda, (e) posición para girar a la derecha, (f-j) gesto para detener el robot.



social indistinto. Los participantes firmaron un formato de consentimiento y estuvieron de acuerdo a ser video-grabados durante el uso de la interfaz natural para teleoperar el robot. La evaluación se llevó a cabo en un laboratorio de investigación de un instituto público local, equipado con dos computadoras, un robot Xiao R Geek, un sensor Leap Motion, y dos cámaras. El sistema se instaló en una computadora como servidor, mientras que la otra funcionó como enrutador. El robot Xiao R Geek fue configurado y calibrado conforme a los movimientos de la mano, detectados con el sensor de gestos Leap Motion (ver Figura 2). Las predicciones fueron: (1) los participantes se sentirán principalmente contentos al interactuar con el robot móvil, (2) la percepción de usabilidad del sistema para la teleoperación del robot móvil es alta, (3) la percepción de la interacción con la interfaz natural y el robot móvil es buena, y (4) todos los participantes lograrán terminar la tarea.

El procedimiento de la evaluación consistió en un mensaje de bienvenida y la respuesta de un cuestionario contestado en línea para conocer: i) datos demográficos, ii) experiencia en el uso de interfaces naturales (p. ej. Kinect, Wii motion, Leap Motion), y iii) estado de ánimo de los participantes. Entonces, a cada participante se le explicó el funcionamiento del sistema, las características del sensor Leap Motion, y cómo por medio de los gestos de la mano se teleopera el robot móvil. Se eligió este sensor su accesibilidad, precisión y capacidad para reconocer los gestos de la mano [19]. Los gestos programados para mover el robot móvil fueron la mano extendida hacia adelante, atrás, izquierda o derecha, y para detenerlo se usó la mano cerrada (ver Figura 2). La tarea consistió en usar los gestos de la mano para teleoperar el robot móvil desde un punto inicial a un punto final a través de un

laberinto. Esta tarea fue seleccionada porque permite conocer la coordinación vista-mano de los participantes (ver Figura 3). En la ejecución de la tarea todos los participantes de pie, usaron su mano dominante durante la interacción, teniendo como retroalimentación visual en tiempo real el movimiento del robot a través del laberinto (ver Figura 4).

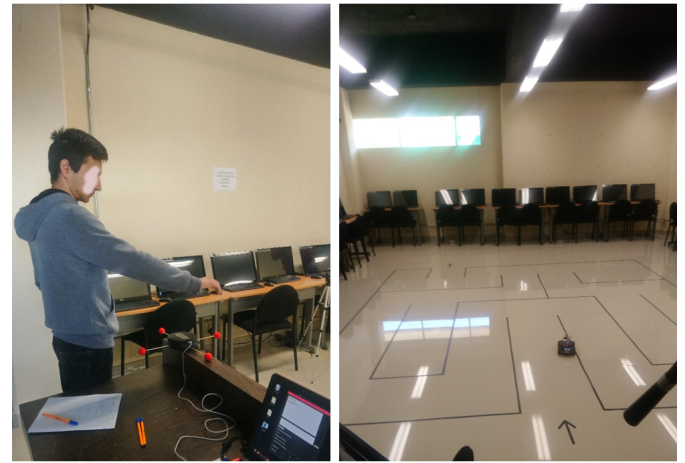


Figura 4: (a) participante con su mano usando el sensor Leap Motion, (b) movimiento del robot Xiao R Geek a través del laberinto.

La escala de usabilidad del sistema (SUS) [20] y la versión extendida del modelo de aceptación de la tecnología (TAM) [21] (5 puntos escala Likert) fueron aplicados al final del estudio. Durante el cuestionario de salida, los participantes indicaron su estado de ánimo al momento de interactuar con el robot y después de la evaluación, p. ej. si se encontraban molestos, nerviosos, aburridos, neutrales, relajados o contentos. También, en el cuestionario de salida se le preguntó a los participantes sobre los posibles usos, sugerencias o comentarios del sistema.

#### IV. RESULTADOS

Los participantes fueron 32 estudiantes universitarios (edad media  $\pm$  desviación estándar:  $21.5 \pm 1.97$ , 9 mujeres, 23 hombres). Con respecto al conocimiento de dispositivos de interacción natural, 29/32 conocen el Wii motion, 27/32 el kinect, y 2/32 el sensor de gestos Leap Motion. En el tiempo de uso de los dispositivos, el Wii motion es el que han usado por más tiempo (23/32: 10/32 más de un año, 1/32 entre 6 meses y 1 año, 4/32 de 1 a 3 meses, y 8/32 menos de 1 mes), seguido por el Kinect (25/32: 3/32 más de 1 año, 3/32 entre 6 meses y 1 año, 1/32 de 3 a 6 meses, 3/32 de 1 a 3 meses, y 15/32 menos de 1 mes), y ninguno de los participantes había usado el sensor de gestos Leap Motion.

##### IV-A. Experiencia de uso

La Figura 5 muestra el estado de ánimo promedio de los participantes obtenido en el cuestionario de entrada y el cuestionario de salida. Como se muestra en la Figura 5, antes

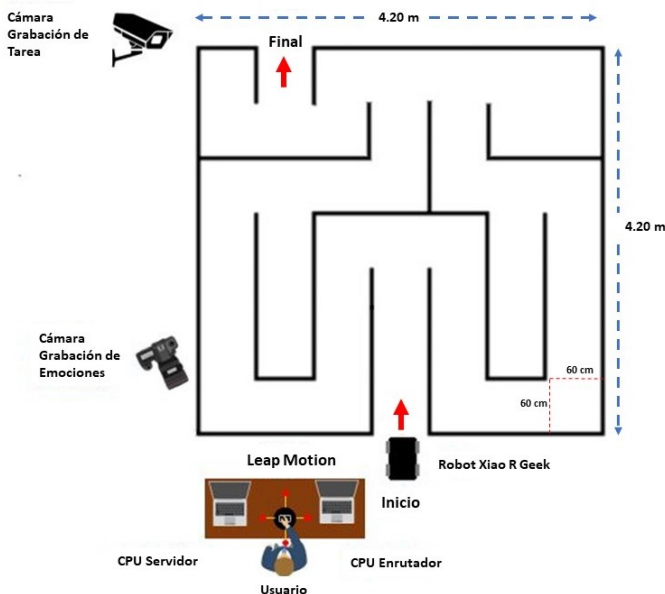


Figura 3: Escenario de la evaluación. Materiales: computadora enrutador, computadora servidor con el sistema, robot Xiao R Geek, sensor Leap Motion, y dos cámaras de grabación.

de la evaluación del sistema el 66 % de los participantes se sintió neutral y el 22 % relajado; durante la evaluación el 41 % se sintió nervioso, el 25 % neutral y el 22 % contento; mientras que al finalizar, el 19 % se sintió neutral, el 19 % relajado y el 59 % contento. Además, la Figura 6 muestra que el 75 % está totalmente de acuerdo de que el sistema es divertido (media 4.75/5).

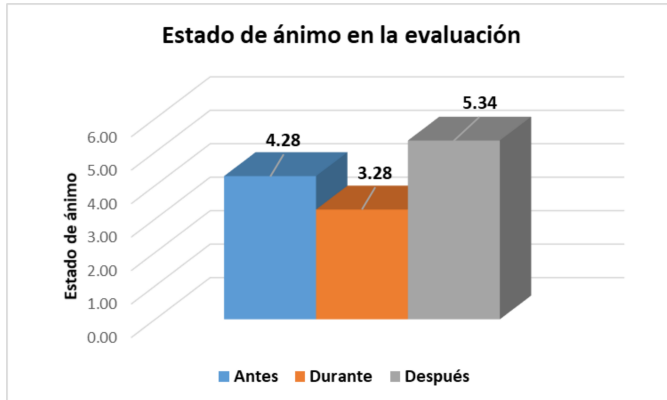


Figura 5: Promedio del estado de ánimo reportado por los participantes: antes (4.28/6), durante (3.28/6), después (5.34/6). Estado de ánimo: 1=molesto, 2=nervioso, 3=aburrido, 4=neutral, 5=relajado, 6=contento.

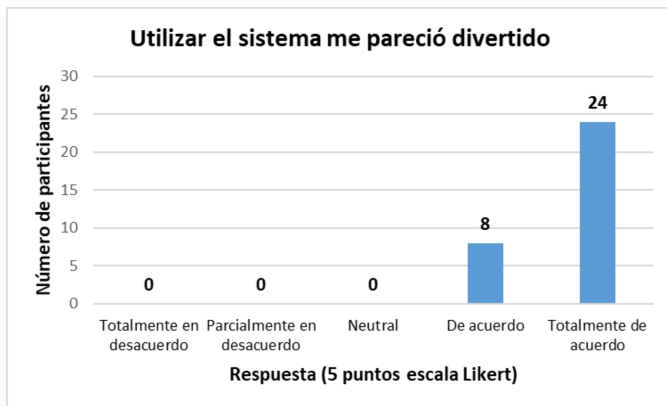


Figura 6: Experiencia de uso del sistema.

#### IV-B. Usabilidad

La intención de uso es alta, donde el 47 % está totalmente de acuerdo y el 34 % está de acuerdo en usar el sistema (media 4.16/5). De los cuales, 15/32 lo usaría de 1 a 2 días, 8/32 de 3 a 4 días, 6/32 por 5 días, y 3/32 de 6 a 7 días. Además, el 54 % lo usaría por 20 minutos o más. De acuerdo a las sugerencias de los participantes, el sistema tiene diversos dominios de aplicación, por ejemplo:

- para manipular/mover remotamente objetos, algunos comentarios son: “transportar objetos de un lugar a otro” (P2), “manejar muchos dispositivos a distancia, manejar a distancia objetos a otro planeta” (P4), “control de

herramientas en entornos peligrosos” (P6), “operaciones a largo alcance donde no entre el cuerpo humano” (P19), “alcanzar cosas que el robot pueda traer hacia mí sin necesidad de moverme” (P26), “posible coche para rescate” (P31);

- entretenimiento, algunos comentarios son: “para jugar o grabar cosas” (P7), “diversión” (P1, P8, P13, P16), “para jugar con él” (P24);
- manejo de objetos en la industria/hogar, algunos comentarios son: “facilitar el manejo de dispositivos” (P15), “controlar aparatos electrónicos” (P18);
- uso en dispositivos de asistencia para personas con capacidades limitadas, algunos comentarios son: “silla para discapacitados” (P24), “personas cuya movilidad sea limitada y a la vez esté en una silla de ruedas” (P30), “sillas de ruedas, pacientes con poca movilidad o sin extremidades” (P32);
- o recorrido de lugares remotos, un comentario es: “recorrer algún lugar a distancia” (P9).

Con respecto a la usabilidad, el 53 % está totalmente de acuerdo y el 25 % de acuerdo sobre la utilidad del sistema de control natural para manejar un robot de manera inalámbrica (media 4.31/5). Siendo fáciles de aprender los comandos para manejar el control del robot (media 4.28/5).

#### IV-C. Percepción de la interacción con la interfaz natural

Durante la interacción, los participantes están de acuerdo que el robot siguió sus órdenes fielmente en todo momento (media 3.56/5). Además, tuvieron buena experiencia al controlar el robot (media 3.88/5). Adicionalmente, la mayoría considera sencillo manejar el robot (18/32) (ver Figura 7), con el 66 % sintiéndose cómodo durante la interacción con el sistema. No obstante, se mostraron neutrales entre si es más sencillo interactuar de forma natural o usar un joystick (control remoto) (media 3/5).

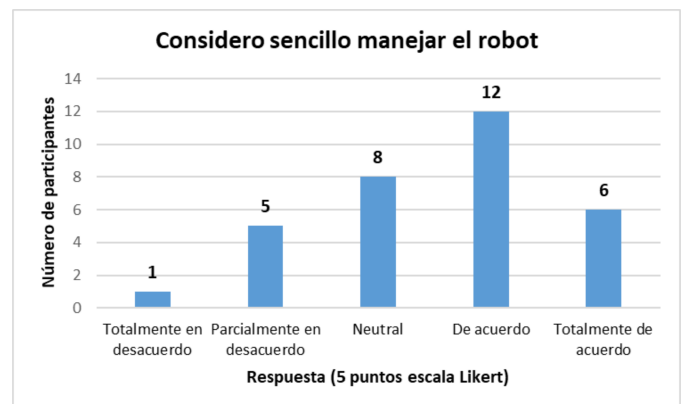


Figura 7: Percepción de la facilidad de uso en la interacción natural con el robot.

#### IV-D. Desempeño

Todos los participantes lograron terminar la tarea con un tiempo promedio de 1 min. 56 seg. y un promedio de 8.125

colisiones con las paredes del laberinto durante el recorrido de inicio a fin. Para mejorar el desempeño en la interacción humano-robot, los participantes realizaron diversas recomendaciones. En primer lugar, 10/32 de los participantes proponen mejorar la sensibilidad para reconocer los movimientos de la mano y sincronizarlos con la velocidad del robot, dado que de acuerdo a su percepción había un retardo. En segundo lugar, 6/32 de los participantes recomiendan modificar la velocidad de giro del robot, porque consideran que la velocidad era más alta con respecto a la velocidad de avanzar o retroceder. En tercer lugar, 5/32 de los participantes sugieren mejorar la estética del robot para identificar más fácilmente la parte delantera y trasera del mismo. Por último, 4/32 de los participantes recomiendan agregar otros gestos de la mano para realizar movimientos curvos, o modificar los comandos actuales para evitar el cansancio de la mano.

## V. CONCLUSIONES Y TRABAJO FUTURO

Se han presentado los resultados preliminares de una evaluación de usabilidad y experiencia de uso de un sistema para teleoperar un robot Xiao R Geek DS por medio del sensor Leap Motion. Los resultados principales son: (i) el sistema propuesto genera buenas experiencias de uso - el 75 % está totalmente de acuerdo de que el sistema es divertido (media 4.75/5); (ii) los sujetos reportaron que la aplicación es altamente usable (media 4.16/5) y el 100 % reportó usarla si la tuviera disponible; (iii) la interacción natural con el robot es sencilla de usar y fácil de aprender; y (iv) aún con la falta de experiencia en el uso interfaces naturales basadas en gestos de mano, todos los participantes lograron terminar la tarea asignada.

Además, los participantes consideran que el sistema puede usarse en diferentes dominios, como por ejemplo en la manipulación/control de objetos remotos, el entretenimiento, el manejo de dispositivos en la industria/hogar, la generación de dispositivos de asistencia para personas con capacidades limitadas, o el recorrido de lugares remotos.

Estos resultados proporcionan evidencia de la usabilidad y experiencia de uso de una interfaz natural basada en diálogos gestuales de la mano para teleoperar un robot móvil. Esta teleoperación se puede extender para aplicarlo en alguno de los dominios antes mencionados. Sin embargo, es necesario incluir las sugerencias al diseño actual, para tener un mejor establecimiento del alcance e impacto de estos resultados. Por último, como trabajo futuro, se tiene considerado usar el sistema en un dominio clínico en apoyo de la generación de dispositivos de asistencia. Asimismo, integrar mediciones fisiológicas para comparar los resultados de auto-reporte obtenidos en este trabajo.

## AGRADECIMIENTOS

Los autores desean agradecer el apoyo del Tecnológico Nacional de México y al Instituto Tecnológico de Ensenada a través del proyecto 6798.19-P.

## REFERENCIAS

- [1] A. Dix, J. Finlay, G. Abowd, and R. Beale, *Human-Computer Interaction 3rd. Ed.* Pearson/Prentice-Hall, 2004.
- [2] H. Vitense, J. Jacko, and V. Emery, "Multimodal feedback: an assessment of performance and mental workload," *Ergonomics*, vol. 46, no. 1-3, pp. 68-87, 2003.
- [3] J. Preece, Y. Rogers, and H. Sharp, *Interaction design: beyond human-computer interaction.* John Wiley & Sons, 2015.
- [4] B. Fang, F. Sun, H. Liu, and C. Liu, "3D human gesture capturing and recognition by the IMMU-based data glove," *Neurocomputing*, vol. 277, pp. 198-207, 2018.
- [5] A. Sharma, A. Yadav, S. Srivastava, and R. Gupta, "Analysis of movement and gesture recognition using Leap Motion Controller," *Procedia Computer Science*, vol. 132, no. Iccids, pp. 551-556, 2018.
- [6] M. Mortimer, B. Horan, and M. Seyedmahmoudian, "Building a relationship between robot characteristics and teleoperation user interfaces," *Sensors (Switzerland)*, vol. 17, no. 3, 2017.
- [7] D. Santiago, E. Slawiński, and V. Mut, "Delayed Trilateral Teleoperation of a Mobile Robot," *Mathematical Problems in Engineering*, vol. 2017, pp. 1-12, 2017.
- [8] C. Chen, L. Chen, X. Zhou, and W. Yan, "Controlling a robot using leap motion," *2017 2nd International Conference on Robotics and Automation Engineering, ICRAE 2017*, vol. 2017-December, pp. 48-51, 2018.
- [9] M. S. B. Piamba, M. S. B. Piamba, and O. A. V. Albán, "Interfaces naturales en la robótica: una revisión," *Scientia et Technica*, vol. 23, no. 1, pp. 112-118, 2018.
- [10] J. M. Gómez-de gabriel and W. Harwin, "Evaluation of sensor configurations for robotic surgical instruments," *Sensors (Switzerland)*, vol. 15, no. 10, pp. 27341-27358, 2015.
- [11] K. Nagatani, S. Kiribayashi, Y. Okada, K. Otake, K. Yoshida, S. Tadakoro, T. Nishimura, T. Yoshida, E. Koyanagi, M. Fukushima, and S. Kawatsuma, "Emergency Response to the Nuclear Accident at the Fukushima Daiichi Nuclear Power Plants using Mobile Rescue Robots," *Journal of Field Robotics*, vol. 7, no. PART 1, pp. 81-86, 2015.
- [12] D. Dolgov, S. Thrun, M. Montemerlo, and J. Diebel, "Path planning for autonomous vehicles in unknown semi-structured environments," *The International Journal of Robotics Research*, vol. 29, no. 5, pp. 485-501, 2010.
- [13] A. Barrientos, J. Colorado, J. del Cerro, A. Martinez, C. Rossi, D. Sanz, and J. Valente, "Aerial Remote Sensing in Agriculture: A Practical Approach to Area Coverage and Path Planning for Fleets of Mini Aerial Robots," *Journal of Field Robotics*, vol. 8, no. 5, pp. 667-689, 2011.
- [14] N. Mavridis, N. Giakoumidis, and E. Machado, "A Novel Evaluation Framework for Teleoperation and a Case Study on Natural Human-Arm-Imitation Through Motion Capture," *International Journal of Social Robotics*, vol. 4, pp. 5-18, 2012.
- [15] J. S. Artal-Sevil, J. L. Montanes, A. Acon, and J. A. Dominguez, "Control of a Bionic Hand using real-time gesture recognition techniques through Leap Motion Controller," *Proceedings of 2018 Technologies Applied to Electronics Teaching, TAAE 2018*, 2018.
- [16] A. Sarkar, K. A. Patel, R. K. Ram, and G. K. Kapoor, "Gesture control of drone using a motion controller," *2016 International Conference on Industrial Informatics and Computer Systems, CIICS 2016*, pp. 1-5, 2016.
- [17] Y. Pititeeraphab, P. Choitkunnan, N. Thongpance, K. Kullathum, and C. Pintavirooj, "Robot-arm control system using LEAP motion controller," *BME-HUST 2016 - 3rd International Conference on Biomedical Engineering*, pp. 109-112, 2016.
- [18] M. Kaya, B. Cubukcu, and U. Yuzgec, "Remote Vehicle Control with Leap Motion Touchless Sensor," in *3rd International Conference on Computer Science and Engineering*, 2018, pp. 292-297.
- [19] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the Leap Motion Controller," *Sensors (Switzerland)*, vol. 13, no. 5, pp. 6380-6393, 2013.
- [20] J. Brooke, "SUS - A quick and dirty usability scale," *Usability Evaluation in Industry*, pp. 189-194, 1996.
- [21] F. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319-340, 1989.

# Procesamiento de Imágenes en Hardware Reconfigurable

E. Morales, R. Herrera Charles, A. Calvillo Tellez  
 Instituto Politécnico Nacional-CITEDI  
 Tijuana, Baja California, México.  
 {emorales, robcharles, calvillo}@citedi.mx

J. R. Cárdenas-Valdez  
 TECNM-Instituto Tecnológico de Tijuana  
 Tijuana, Baja California, México.  
 jose.cardenas@tectijuana.edu.mx

**Resumen**—Los dispositivos de hardware reconfigurable tienen un papel preponderante en aplicaciones que requieren procesamiento en tiempo real como lo son: análisis de datos, automotrices, inteligencia artificial, telecomunicaciones, entre otras, además de que son ideales para aplicaciones de tipo cómputo concurrente, bajo consumo de potencia y rediseño de fabricación. En este trabajo se presenta un algoritmo detector de puntos de interés SURF que tiene un alto costo computacional, implementado en un dispositivo de hardware reconfigurable FPGA que se encuentra en la tarjeta de desarrollo ZYBO de Xilinx®, mostrando sus ventajas de rediseño y paralelismo aplicado al procesamiento de imágenes.

**Keywords**— *cómputo concurrente; FPGA; hardware reconfigurable; procesamiento; tiempo real.*

## I. INTRODUCCIÓN

El procesamiento de imágenes en tiempo real ha atraído la atención de la comunidad por el gran crecimiento que están presentando las aplicaciones en sistemas avanzados en asistencia de manejo (ADAS, Advanced Driver Assistance System), reconocimiento de rostros, localización y mapeo simultaneo (SLAM, Simultaneous Localization and Mapping), por mencionar algunos, por estas necesidades de desarrollo del procesamiento es crucial emplear dispositivos capaces de realizar el procesamiento en tiempo real, como lo son las unidades de procesamiento gráfico (GPU, Graphics Processing Unit), así como los dispositivos de hardware reconfigurable FPGA (Field Programmable Gate Array) y procesadores digitales de señales (DSP, Digital Signal Processor), que son los dispositivos que sobresalen en el procesamiento digital por realizar operaciones de manera concurrente [1], [2], [3].

Con el avance tecnológico en la fabricación de dispositivos de hardware reconfigurable además de sus ventajas de cómputo rápido, bajo consumo energético y costo de fabricación, una ventaja de vital importancia que presentan los dispositivos de hardware reconfigurable de la misma forma que su flexibilidad para modificar el diseño las veces necesarias, útil para el proceso previo a la fabricación [4], [5].

Existen algoritmos de procesamiento de imágenes que tienen un alto costo computacional como lo son los algoritmos descriptores, SIFT (Scale-invariant feature transform) [6], SURF (Speeded-Up Robust Features) [7], ORB [8], por mencionar algunos.

La aportación principal en este trabajo es la detección de puntos característicos, invariantes a la escala, rotación e iluminación, los algoritmos descriptores se dividen en dos

etapas. La primera es la detección, en ésta etapa se implementan algoritmos aplicados a bordes para identificar cambios de intensidad en la imagen y puntos o píxeles de interés, la segunda etapa consiste en la descripción de los puntos detectados por el algoritmo.

Este artículo se encuentra organizado de la siguiente manera: la sección II está dedicada a conceptos teóricos y modelado de procesamiento de imágenes, en la sección III, se presentan los diseños realizados, en la sección IV se muestran los resultados obtenidos como los recursos utilizados por la tarjeta de desarrollo. Finalmente, en la sección V se presentan las conclusiones obtenidas del desarrollo de este trabajo.

## II. MODELO DEL SISTEMA

### A. Hardware reconfigurable

Un hardware reconfigurable es un circuito que su principal característica es que sea programable para una aplicación en particular. También, representa la funcionalidad de un circuito que puede ser reprogramado según los requerimientos de las aplicaciones. Una de las principales características, es que el circuito es mayormente implementado de forma paralela que secuencial. Si una aplicación es implementada en un dispositivo FPGA, que se espera en tareas en diferentes componentes de hardware, esto implica que los FPGA trabajan de manera paralela, esta característica los hace ideales para aplicaciones de procesamiento de imágenes, principalmente en nivel bajo e intermedio de procesamiento como se muestra en la Fig. 1, en donde bajo nivel se trabaja directamente con los píxeles de la imagen, el nivel intermedio se trabaja con regiones y el nivel alto ya se hace la detección de objetos.

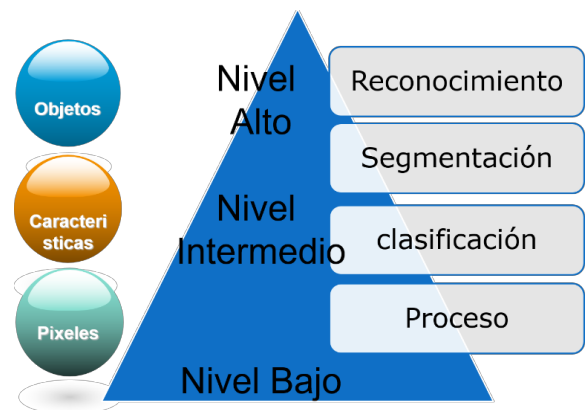


Fig. 1. Niveles de procesamiento de imágenes.

Identify applicable sponsor/s here. If no sponsors, delete this text box (sponsors).



**B. Procesamiento de imágenes**

El procesamiento de imágenes se define como una serie de operaciones matemáticas para obtener el resultado deseado, esto quiere decir, detectar algún evento característico, medir un objeto en específico en una imagen, clasificar uno o más objetos o describir una escena. El diagrama a bloques para reconocimiento de patrones se muestra en la Fig. 2 donde se obtiene la imagen por medio de un sensor y en la primera etapa se realiza el pre-proceso que normalmente se indica la supresión de ruido o alguna modificación necesaria para su manipulación, después se realiza una segmentación o localización que consiste en separar en secciones o descartar las partes de la imagen que no se requieren, en la extracción de rasgos se procesa la información para identificar qué es lo que se está observando, adicionalmente se realiza una comparación con una o varias referencias que se tengan previamente, se clasifica según el resultado obtenido en la comparación y por último se toma una decisión si es o no el objeto a reconocer.



Fig. 2. Procesamiento de Imágenes en tiempo real

**C. Procesamiento embebido y tiempo real**

Un procesamiento es embebido si en un componente o producto se introduce un sistema computacional, normalmente diseñado para realizar una tarea específica. El procesamiento de imagen embebido es de gran utilidad en las cámaras inteligentes, donde la cámara no solo captura la información del exterior, sino que también extrae la información requerida para la aplicación. Un procesamiento en tiempo real es cuando la respuesta a un evento ocurre dentro de un tiempo específico, en un sistema de procesamiento de imagen en tiempo real, el sistema captura, analiza, obtiene información y después utiliza la información para realizar una acción. Todo esto lo tiene que hacer en un lapso predeterminado, es importante mencionar que en algunos casos no se trabaja a la velocidad de cuadros por segundo (FPS, frame per second).

**D. Algoritmo SURF**

El algoritmo SURF [9-10] está constituido por dos etapas que a saber son la detección:

- Imagen integrada

- Matriz Hessiana
- Espacio de escalas
- Localización de puntos de interés

La segunda etapa que es el descriptor y realiza los siguiente:

- Asignar orientación
- Descripción basada en Haar wavelet [11]
- Relación entre dos imágenes

La imagen integrada consiste en realizar la suma de los valores de los píxeles en la imagen, que se expresa en la Ecuación ( 1),

$$I_{\Sigma(x)} = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j) \tag{1}$$

Donde  $I_{\Sigma}$  representa la imagen integrada,  $x$  el total de columnas de la imagen,  $y$  y el total de filas en la imagen.

La matriz Hessiana se conforma por las segundas derivadas parciales de una función como se muestra en la Ecuación ( 2) . Al calcular su determinante se encuentra el máximo, mínimo o punto de silla de una función.

$$H(I) = \begin{bmatrix} L_{xx}(I) & L_{xy}(I) \\ L_{xy}(I) & L_{yy}(I) \end{bmatrix} \tag{2}$$

Donde  $L_{xx}(I)$  representa la convolución de la función  $I$  con filtro gaussiano que representa la segunda derivada parcial como se expresa en la Ecuación ( 3) .

$$G(x, y) = \frac{(1)}{(2\pi\sigma^2)} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3}$$

En el dominio espacial se representa en las Figs. 3-4, para realizar el espacio de escalas se modifican de tamaño los kernels donde el espacio de escalas es invariante. La localización de puntos de interés se realiza comparando un píxel central con su vecindario de 26, permaneciendo si es máximo local o descartando si es menor a alguno de estos.

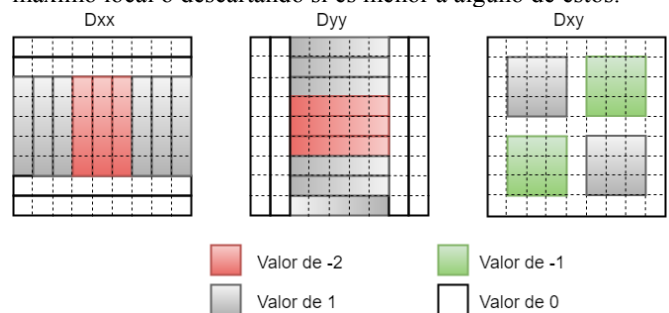


Fig. 3. Aproximación de segundas derivadas parciales por filtro de Gauss.

En la Fig. 4 se muestra el vecindario a comparar. donde el cuadro rojo representa el píxel central y los cuadros verdes el vecindario, si el píxel central es mayor a su vecindario, ese valor permanece y se realiza la localización del punto, en caso contrario se discrimina.

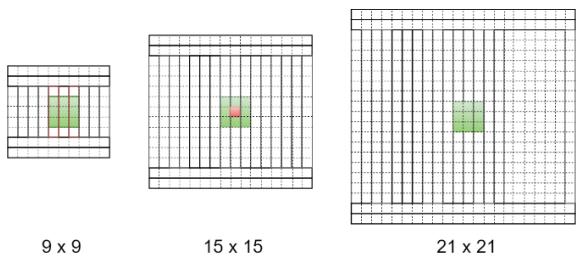


Fig. 4. Supresión de no máximos en un vecindario de 26 píxeles.

### III. DESARROLLO

En la plataforma de trabajo se realizó en el entorno de desarrollo integrado (IDE, Integrated Development Environment) Vivado de Xilinx® en lenguaje descriptor de hardware HDL, se utilizó una cámara ov7670 de omnivisión y un monitor con entrada VGA en una computadora portátil HP con sistema operativo Ubuntu 16.04. En la Fig. 5 se muestra el diagrama a bloques del circuito a implementar.

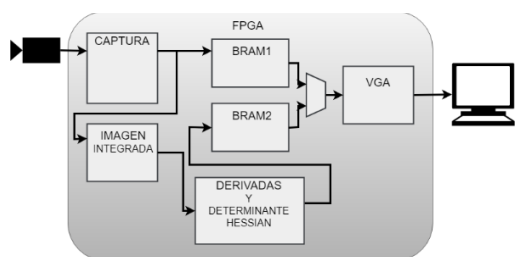


Fig. 5. Diagrama a bloques del circuito para el procesamiento en hardware reconfigurable.

Se obtienen los datos con una cámara ov7670, que envía la información de los píxeles de manera paralela, en la Fig. 6 se describe como son enviados los datos.

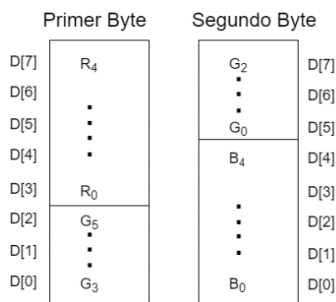


Fig. 6. Datos enviados desde la cámara ov7670.



Fig. 7. Circuito para calcular las derivadas parciales primer diseño.

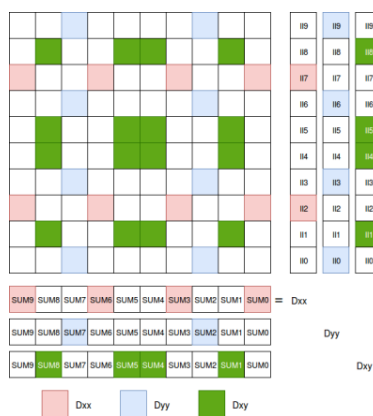


Fig. 8. Segundo diseño para calcular las segundas derivadas parciales.

### IV. RESULTADOS

Con los dos circuitos diseñados se obtuvieron los siguientes resultados en la Tabla I, se muestra los recursos utilizados para el diseño que se muestra en las Figs. 7-8.

TABLA I. RECURSOS UTILIZADOS DEL PRIMER DISEÑO

Elemento	Recurso Utilizado	Recurso Disponible	Porcentaje Utilizado
LUT	6588	17600	37.43%
LUTRAM	5344	6000	89.1%
FF	2894	35200	8.2%
BRAM	38	60	63.3%
DSP	7	80	8.8%
IO	39	100	39.0%
BUFG	5	32	15.6%
MMCM	1	2	50%

En la Tabla II se presentan los recursos obtenidos para el segundo diseño, que se muestra en la Fig. 8, donde se obtuvo una reducción considerable de la cantidad de LUTs en la descripción de hardware y el la cantidad de RAM utilizados por el segundo diseño para calcular las derivadas parciales del algoritmo detector de puntos de interés SURF, aunque en ambos casos la cantidad de recursos utilizados esta sobrada para aplicaciones de procesamiento de detección de bordes bajo el algoritmo de SURF.

TABLA II. RECURSOS UTILIZADOS DEL SEGUNDO DISEÑO

Elemento	Recurso Utilizado	Recurso Disponible	Porcentaje Utilizado
LUT	6201	17600	35.2%
LUTRAM	5168	6000	86.1%
FF	2430	35200	6.9%
BRAM	38	60	63.33%
IO	39	100	39.0%

Elemento	Recurso Utilizado	Recurso Disponible	Porcentaje Utilizado
BUFG	5	32	15.6%
MMCM	1	2	50%

V. CONCLUSIÓN

En este trabajo se desarrolló un sistema que reduce el número de registros de algoritmo detector de puntos de interés SURF donde una segunda propuesta reduce la cantidad de registros utilizados, sobre todo si se analiza la cantidad de LUTs utilizadas, como propuesta a trabajo futuro se propone utilizar HLS (High Language Synthesis) y comparar con los resultados obtenidos, con el fin de reducir la cantidad de recursos utilizados en la implementación en hardware. El sistema utilizado se basa en el cálculo de matrices Hessianas, escalamiento y localización de espacios de interés, además de asignar orientación y descripción de Haar wavelet, proceso implementado en un sistema de hardware reconfigurable FPGA. Cabe indicar que la reducción de recursos utilizados en la detección de puntos de interés permite identificar y procesar imágenes de mayor resolución.

REFERENCIAS

[1] J. Arróspide y L. Salgado, «Video based vehicle detection and tracking for driver assistance systems," *Securitas Vialis*, vol. 7, no.1-3, pp. 41--49, 2015.

[2] C. Kyrkou y T. Theocharides, "Accelerating object detection via a visual-feature-directed search cascade: algorithm and field programmable gate array implementation," *Journal of Electronic Imaging*, vol. 25, no. 4, pp. 0410131-04101314 2016.  
 A. HajiRassouliha, A. J. Taberner, M. P. Nash and P. M. Nielsen, "Suitability of recent hardware accelerators (DSPs, FPGAs, and GPUs) for computer vision and image processing algorithms," *Signal Processing: Image Communication*, vol. 68, pp. 101--119, 2018.

[4] B. Brousseau, "An Energy Efficient FPGA Hardware Architecture for the Acceleration of OpenCV Object Detection," University of Toronto: Tesis Doctoral, 2012.

[5] R. Cayssials, "sistemas embebidos en FPGA," Alfaomega Grupo Editor, 2014.

[6] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91-110, 2004.

[7] H. Bay, A. Ess, T. Tuytelaars y L. Van Gool, "Speeded-up robust features (SURF)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346--359, 2008.

[8] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," *International Conference on Computer Vision*, pp. 2564--2571, 2011.

[9] H. Bay et al., "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol 110, pp. 346-359, 2008.

[10] Bay, H., Tuytelaars, T. and Van Gool, L., "Surf: Speeded up robust features," In *European conference on computer vision*, vol. 3951, pp. 404-417, Graz, Austria, 2006.

[11] Gupta, D., and Choubey, S., "Discrete wavelet transform for image processing," *International Journal of Emerging Technology and Advanced Engineering*, vol. 4, no. 3, 598-602, 2015.



# Ocultamiento de Información en Imágenes Digitales Mediante la Esteganografía Sin Incrustación de Datos

Luis Alberto Sandoval Bravo, Clara Cruz Ramos, Rogelio Reyes Reyes

Instituto Politécnico Nacional

ESIME Unidad Culhuacán, Ingeniería en Computación

Av. Santa Ana No. 1000, Col. San Francisco Culhuacán, CP. 04430, Ciudad de México.

email: lsandovalb1500@alumno.ipn.mx, ccruzra@ipn.mx, rreyesre@ipn.mx

**Resumen**—Las técnicas actuales de esteganografía en imágenes digitales utilizan una imagen portadora e incrustan datos secretos en ella para formar la imagen stego, para lo cual se emplean métodos en el dominio espacial o en el dominio de la frecuencia; sin embargo, este tipo de métodos modifican directamente los bits de la imagen, alterando la intensidad de los píxeles y dejando huellas de modificación causadas por la incrustación de datos en la imagen portadora, lo que hace posible el esteganálisis exitoso. En este documento se propone un método de esteganografía en imágenes digitales sin incrustación de datos directamente en las imágenes, empleando una base de datos de imágenes indexada de acuerdo a su secuencia hash generada empleando un algoritmo hash robusto. Los resultados demuestran que el método propuesto es resistente a las técnicas comunes de esteganálisis para la obtención y/o destrucción de la información oculta.

**Keywords**—Esteganografía; esteganálisis; esteganografía sin incrustación de datos; algoritmo hash; imágenes digitales.

## I. INTRODUCCIÓN

Actualmente, con el avance de las tecnologías y telecomunicaciones aunado a la necesidad de estar constantemente comunicado y compartir información con otras personas como imágenes, videos o audios, muchas veces se requiere una comunicación secreta. La comunicación secreta nace de tener que enviar un mensaje con contenido confidencial a una persona. Una de las técnicas empleada comúnmente para el envío de información secreta es la esteganografía [1]. El principal objetivo de la esteganografía es poder ocultar la mayor cantidad de bits (mensaje) en una portadora sin levantar ninguna sospecha.

Existen diversos métodos de esteganografía para imágenes digitales que se han propuesto en la literatura [2-4]. Todos estos métodos utilizan una imagen portadora para ocultar la información secreta y generar la imagen stego. De manera general los métodos de esteganografía de imágenes se pueden dividir en dos grandes ramas: esteganografía con incrustación de datos y sin incrustación de datos [5].

Según la literatura [1], los métodos con incrustación de datos se pueden dividir en tres categorías: dominio espacial, dominio de frecuencia y métodos adaptativos. Los métodos comunes del dominio espacial incluyen la sustitución de LSB [2], la coincidencia de LSB [3], la paleta de colores [4] y los

métodos basados en histogramas [5].

Las modificaciones causadas por el mecanismo de incrustación en el dominio espacial en algunos casos no son detectadas fácilmente por el ojo humano, pero la información incrustada es sensible a los ataques comunes en imágenes. Para solucionar este problema, se han propuesto muchos métodos de esteganografía de dominio de frecuencia, como la tabla de cuantificación (QT) [6], la transformada discreta de Fourier (DFT) [7] y la incorporación de una transformación de wavelet discreta (DWT) [8]. La esteganografía adaptativa es un caso especial de los métodos de dominio espacial y dominio de frecuencia.

La ocultación de información sin incrustación de datos es un nuevo campo de investigación desafiante. De hecho, "sin cobertura" no quiere decir que no hay portador, pero en comparación con el ocultamiento de información convencional, la ocultación de información sin cobertura no requiere ninguna otra carga. Su característica distintiva es "sin incrustación", es decir, un operador no puede incrustar información secreta modificando el portador. El ocultamiento de información sin incrustación de datos no requiere modificación en las imágenes y podría resistir varias tecnologías de esteganálisis [9-12].

Uno de los métodos más comunes para ocultar los datos secretos "sin incrustación" es mediante la búsqueda de imágenes apropiadas que ya contengan la información. Estas imágenes se consideran imágenes stego, que se utilizan para la comunicación de datos secretos. Para un dato secreto dado, el problema principal reside en cómo encontrar estas imágenes originales, además de que, al realizar alguna modificación a la imagen como el cambio de escala, el cambio de luminancia o adición de ruido, los datos de la imagen también cambian, y por tanto la información se distorsiona o se pierde.

En este artículo se propone un método sin incrustación de datos, partiendo de una base de datos de imágenes y su secuencia hash para cada imagen que es generada por un algoritmo hash. Después, todas estas imágenes se indexan de acuerdo con sus secuencias hash para construir una estructura de índices. Para comunicar los datos secretos, el emisor primero transforma los datos secretos en una cadena de bits y los divide en varios segmentos con la misma longitud. Luego, para cada segmento, la imagen de la secuencia hash que es la misma que

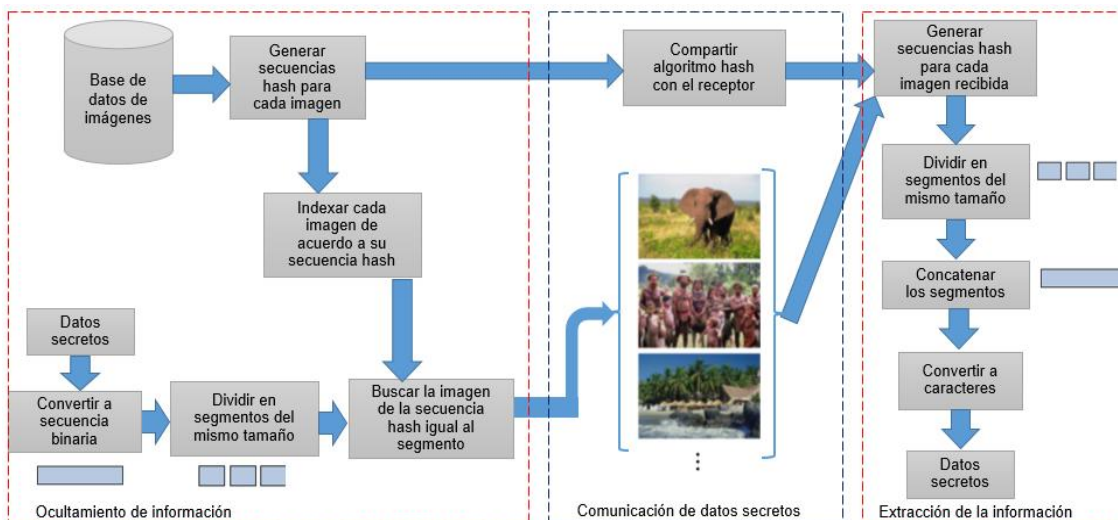


Fig. 1. Diagrama a bloques del método propuesto

el segmento se encuentra buscando en la estructura de los índices. Con esto, se obtienen una serie de imágenes asociadas a los segmentos, que se consideran imágenes estego y se transmiten al receptor. Por otro lado, el receptor recibe las imágenes en orden y empleando el mismo algoritmo hash utilizado en la codificación de los índices, el receptor puede concatenarlas para recuperar los datos secretos.

## II. MÉTODO PROPUESTO

El sistema propuesto cuenta con tres etapas: el ocultamiento de la información, la comunicación de datos secretos y la extracción de la información, véase la Fig. 1.

### A. Ocultamiento de la información.

En el diagrama a bloques del sistema propuesto para el ocultamiento de la información sin incrustación de datos mostrado en la figura 1, se puede observar una primera etapa en la cual se hace una recopilación de imágenes de internet para construir la base de datos con un mínimo de 256 imágenes, para asegurar que al menos una imagen tenga cada combinación con 8 bits (con 8 bits se tiene  $2^8=256$  combinaciones). Además, el tamaño de las imágenes deben ser múltiplos de 3 para poder hacer la división en bloques. Luego, se genera una secuencia hash para cada imagen en la base de datos, estas secuencias son generadas por el siguiente algoritmo hash:

Hay tres pasos principales en el algoritmo hash para una imagen dada:

- 1) Transformar la imagen a nivel de grises y dividirla en bloques no superpuestos de  $3 \times 3$  píxeles.
- 2) Calcular la intensidad promedio de cada bloque.
- 3) Finalmente, se comparan las intensidades de cada bloque en orden de zig-zag como se muestra la figura 2 para formar un vector binario de 8 bits, para esto, se compara si la intensidad del bloque es mayor o igual a la intensidad del siguiente bloque colocándose un "1" en el vector, de lo contrario, se coloca un "0".

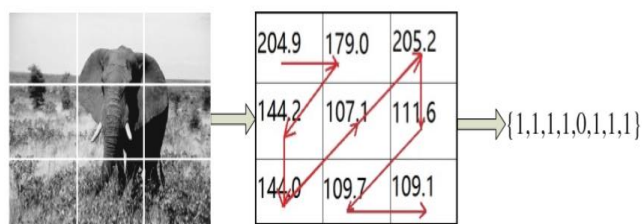


Fig. 2. Algoritmo hash empleado.

Para ocultar los datos secretos, primero se transforma la información en una cadena de bits y se dividen en varios segmentos con la misma longitud, en este caso en segmentos de 8 bits para que coincidan con el tamaño de las secuencias hash.

Luego, para cada segmento de bits, se busca la imagen con la secuencia hash que sea igual al segmento.

Al realizar la búsqueda de todas las imágenes en las cuales, sus secuencias hash coincidan con la consulta en la base de datos, tomará mucho tiempo. Para acelerar la búsqueda, primero se indexan todas las imágenes de la base de datos de acuerdo con su secuencia hash. Luego, se construye una estructura de índices, que contiene todas las posibles secuencias de hash de 8 bits como lo muestra la figura 3.

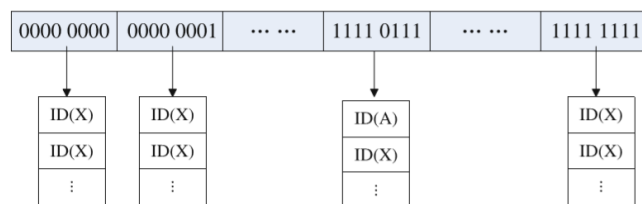


Fig. 3. Estructura de índices propuesta.

### B. Comunicación de datos secretos

Después de encontrar todas las imágenes que ya contienen los datos secretos, esas imágenes se transmiten al receptor una por una en orden. Para el receptor, todas las imágenes también se reciben en orden, ya que la transmisión y recepción de imágenes en orden asegura que los datos secretos puedan recuperarse sin desorden y ser entendidos. También el emisor

comparte el algoritmo de generación de secuencias hash con el receptor.

C. Extracción de la información

Una vez que el receptor obtiene todas esas imágenes con la información oculta, la secuencia hash de cada imagen recibida se genera con el mismo algoritmo hash utilizado por el emisor. Luego, de acuerdo con el orden de las imágenes recibidas, el receptor concatena todas sus secuencias hash de esas imágenes para recuperar los datos secretos. Debido al algoritmo, las secuencias hash de esas imágenes son insensibles a los ataques de imágenes mencionados como el cambio de escala, de luminancia o adición de ruido. Como resultado, los datos secretos recuperados serán los mismos que los datos secretos originales.

III. RESULTADOS EXPERIMENTALES

Para evaluar el sistema propuesto se utilizó una base de datos de imágenes (cifar-10), que se encuentra en internet, la cual contiene 10000 imágenes de 32x32 pixeles.

A. Robustez del algoritmo hash ante cambio de escala

Como se menciona anteriormente, se requiere que el tamaño de la imagen sea múltiplo de 3, por lo tanto, se realizaron pruebas con el algoritmo de generación de secuencias hash empleando diferentes imágenes, considerando las imágenes originales mostradas en la figura 4(a) y 4(b), a las cuales se les aplica un cambio de escala, de tal forma que tuvieran un tamaño múltiplo de 3, para poder generar las secuencias hash.

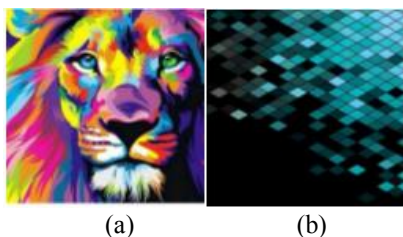


Fig. 4. Imágenes originales

Al realizar las diferentes pruebas, aplicando el algoritmo para la generación de secuencias hash en las imágenes cambiando su escala, se observó que para las diferentes escalas, siempre se obtenía la misma secuencia hash como se muestra en las figuras 5 y 6.

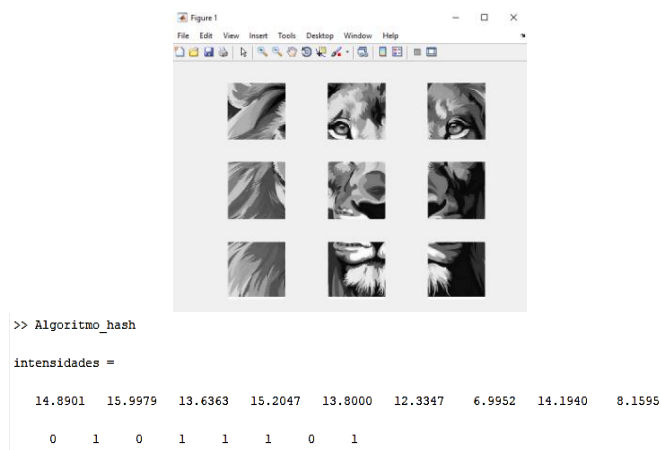


Fig. 5. Secuencia hash obtenida para la imagen 4(a)

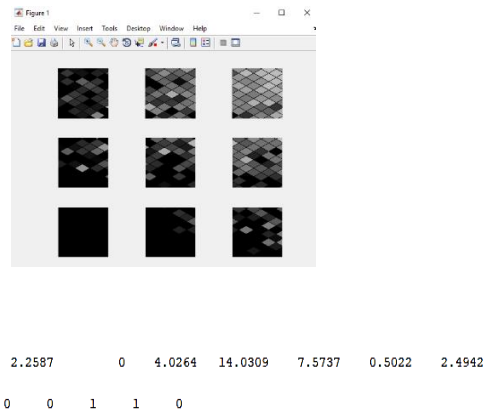


Fig. 6. Secuencia hash obtenida para la imagen 4(b).

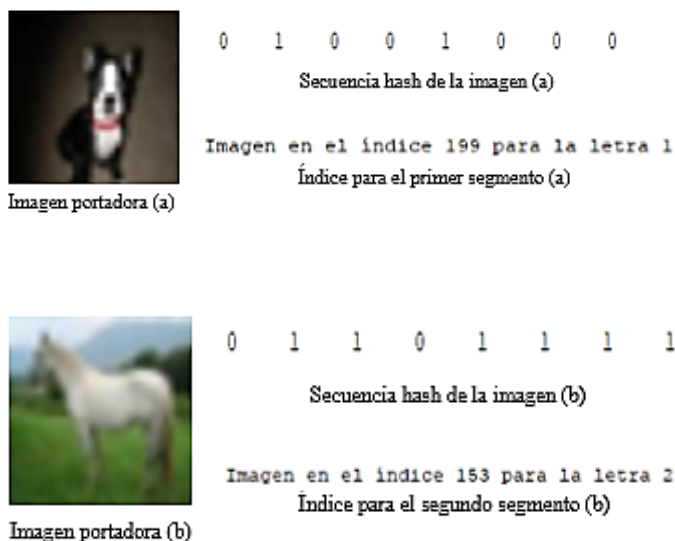
B. Ocultamiento de datos

Para evaluar el método propuesto, todas las imágenes de la base de datos se re-escalaron a un tamaño de 33x33 pixeles para poder implementar el algoritmo de generación de secuencias hash; los mensajes a ocultar son las palabras "Hola" y "silla".

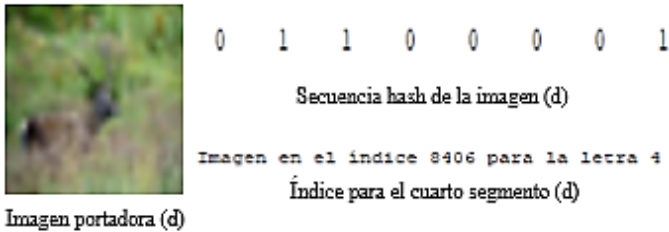
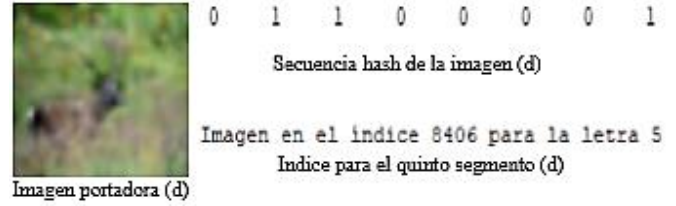
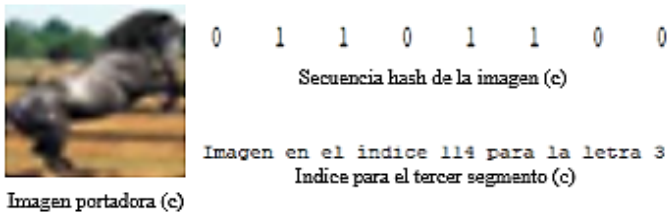
Al convertir los mensajes de texto en binario y dividirlos en segmentos de 8 bits, se represento cada segmento con el hash de una imagen cubierta, realizando la búsqueda en la estructura de índices, en donde coincidieran los 8 bits del segmento con los 8 bits de la secuencia hash para poder elegir alguna de las imágenes correspondientes a esa secuencia hash. A continuación, se muestra el texto en binario segmentado para el mensaje de texto a ocultar "Hola", así como los resultados de las imágenes y sus respectivos hash obtenidos:

H	0	1	0	0	1	0	0	0
o	0	1	1	0	1	1	1	1
l	0	1	1	0	1	1	0	0
a	0	1	1	0	0	0	0	1

A continuación, se muestran las imágenes correspondientes a las secuencias hash encontradas para el mensaje "Hola".







Cabe mencionarse que en el caso de que la letra del mensaje se repita, se puede duplicar la imagen correspondiente a la secuencia hash o se tiene la opción de elegir otra imagen de la estructura de índices que de igual forma, sea igual la secuencia hash.

C. Extracción de datos

En el lado del receptor, se reciben las imágenes en orden y con los pasos del algoritmo hash, se obtienen las secuencias hash de cada imagen, para posteriormente concatenar las secuencias hash tomando en consideración el orden en que fueron recibidas, para obtener posteriormente el mensaje oculto como se muestra en las figuras 7 y 8.

A continuación, se muestra el texto en binario segmentado para el mensaje de texto a ocultar “silla”, así como los resultados de las imágenes y sus respectivos hash obtenidos:

s	0	1	1	1	0	0	1	1
i	0	1	1	0	1	0	0	1
l	0	1	1	0	1	1	0	0
l	0	1	1	0	1	1	0	0
a	0	1	1	0	0	0	0	1

A continuación se muestran las imágenes correspondientes a las secuencias hash encontradas para el mensaje “silla”.

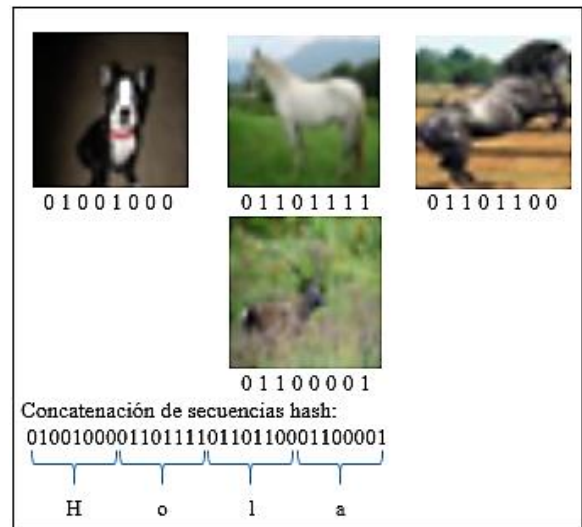
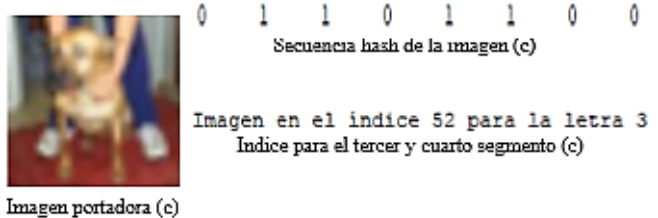
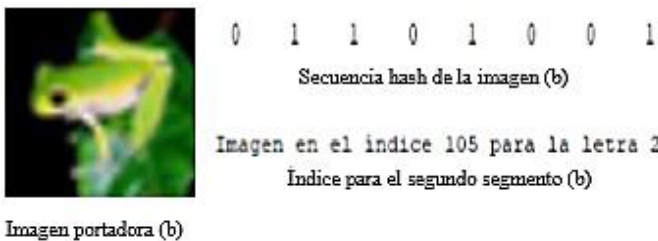
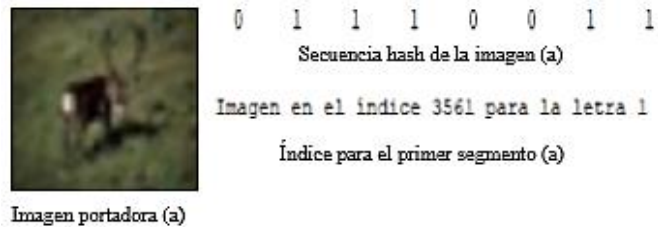


Fig. 7. Imágenes recibidas y extracción del mensaje “Hola”.

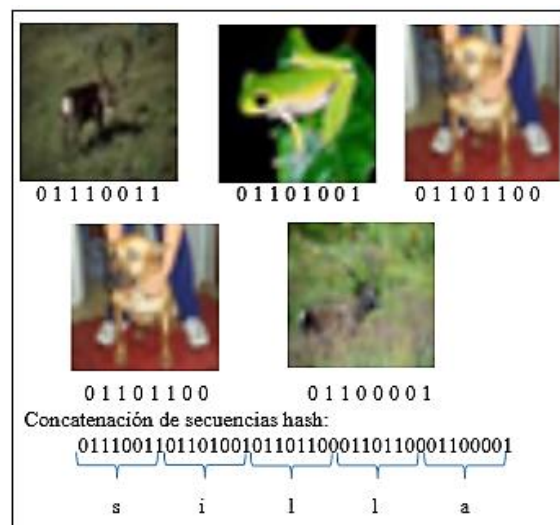


Fig. 8. Imágenes recibidas y extracción del mensaje “silla”.

#### IV. CONCLUSIONES

En este trabajo se propone un método de esteganografía sin incrustación de datos en imágenes digitales, utilizando una base de datos, la cual está construida con imágenes de internet, imágenes que son, con mucha frecuencia enviadas a otras personas, por lo tanto, reduce las sospechas de contener mensajes ocultos. De acuerdo a los resultados obtenidos, se demuestra que el sistema propuesto es robusto al ataque de reescalamiento de las imágenes portadoras, debido a que en todas las pruebas realizadas después de hacer estas modificaciones en la imagen, se pudo recuperar el mismo vector binario generado por el algoritmo. También el sistema propuesto emplea un método sencillo y rápido para indexar las secuencias hash, de tal forma que al generar la secuencia de imágenes que oculta el mensaje, esta se obtiene de manera inmediata.

#### AGRADECIMIENTOS

Agradecemos al Instituto Politécnico Nacional, al Consejo Nacional de Ciencia y Tecnología de México y a la Comisión de Operación y Fomento de Actividades Académicas (COFAA) del IPN por el apoyo otorgado para el desarrollo de este trabajo.

#### REFERENCIAS

- [1] Cheddad, A., Condell, J., Curran, K., Mc Kevitt, P.: Digital image steganography: survey and analysis of current methods. *Sig. Process.* 90(3), 727–752 (2010)
- [2] Wu, H.C., Wu, N., Tsai, C.S., Hwang, M.S.: Image steganographic scheme based on pixel-value differencing and LSB replacement methods. In: *Proceedings-Vision Image and Signal Processing*, vol. 152, pp. 611–615 (2005)
- [3] Mielikainen, J.: LSB matching revisited. *IEEE Sig. Process. Lett.* 13(5), 285–287 (2006)
- [4] Johnson, N.F., Jajodia, S.: Exploring steganography: seeing the unseen. *IEEE Comput.* 31(2), 26–34 (1998)
- [5] Li, Z., Chen, X., Pan, X., Zheng, X.: Lossless data hiding scheme based on adjacent pixel difference. In: *Proceedings of the International Conference on Computer Engineering and Technology*, pp. 588–592 (2009)
- [6] Li, X., Wang, J.: A steganographic method based upon JPEG and particle swarm optimization algorithm. *Inf. Sci.* 177(15), 3099–3109 (2007)
- [7] McKeon, R.T.: Strange Fourier steganography in movies. In: *Proceedings of the IEEE International Conference on Electro/information Technology (EIT)*, pp. 178–182 (2007)
- [8] Chen, W.Y.: Color image steganography scheme using set partitioning in hierarchical trees coding, digital Fourier transform and adaptive phase modulation. *Appl. Math. Comput.* 185(1), 432–448 (2007)
- [9] Ker, A.D.: Improved detection of LSB steganography in grayscale images. In: *Fridrich, J. (ed.) IH 2004. LNCS*, vol. 3200, pp. 97–115. Springer, Heidelberg (2004)
- [10] Ker, A.: Steganalysis of LSB matching in grayscale images. *IEEE Sig. Process. Lett.* 12(6), 441–444 (2005)
- [11] Xia, Z.H., Wang, X.H., Sun, X.M., Liu, Q.S., Xiong, N.X.: Steganalysis of LSB matching using differences between nonadjacent pixels. *Multimedia Tools Appl.* (2014). doi:10.1007/s11042-014-2381-8
- [12] Xia, Z.H., Wang, X.H., Sun, X.M., Wang, B.W.: Steganalysis of least significant bit matching using multi-order differences. *Secur. Commun. Netw.* 7(8), 1283–1291 (2014)

# Estimación de la Matriz de Cuantización de Archivos Previamente Almacenados en Formato JPEG

*Sergio Hernández Rojas*  
ESIME Unidad Culhuacán  
Instituto Politécnico Nacional  
Ciudad de México, México  
shernandezr1505@alumno.ipn.mx

*Rogelio Reyes Reyes*  
ESIME Unidad Culhuacán  
Instituto Politécnico Nacional  
Ciudad de México, México  
reyesre@ipn.mx

*Clara Cruz Ramos*  
ESIME Unidad Culhuacán  
Instituto Politécnico Nacional  
Ciudad de México, México  
ccruzra@ipn.mx

**Resumen**— En la actualidad, aproximadamente el 80% de las imágenes que se transmiten a través de la red son almacenadas utilizando el sistema de compresión JPEG, esto debido a la elevada capacidad de compresión sin aparentes cambios que posee. Este proceso consta de varias etapas, una de ellas es la estimación de la tasa de compresión y la calidad de la imagen resultante mediante el uso de la matriz de cuantización. Estas tablas son estándares ya definidos en relación con la cantidad de compresión deseada. En el presente artículo se desarrolla un algoritmo para la estimación de la matriz de cuantización de imágenes previamente almacenadas en formato JPEG, la cual puede ser utilizada para detectar cambios o alteraciones, llevar a cabo un esteganálisis de la imagen, así como en técnicas de marcado de agua.

**Palabras clave**— JPEG, Matriz de Cuantización, esteganografía, esteganálisis.

## I. INTRODUCCIÓN

El formato de Joint Photographic Experts Group (JPEG), es uno de los formatos más populares para gráficos web y el sistema de compresión más utilizado en imágenes digitales [1]. Hoy día el estándar JPEG es reconocido por la Organización Internacional de Estándares con el nombre de ISO 10918 [2,3].

El motivo por el cual este sistema de compresión es sumamente utilizado se debe a su elevada relación entre compresión y calidad, además de la gran capacidad que posee para representar una misma imagen con una cantidad considerablemente menor de datos, lo cual le permite al usuario obtener imágenes visualmente similares, pero porcentualmente muy distintas en cuanto al tamaño de la imagen se refiere.

El algoritmo de compresión JPEG explota dos defectos que el sistema visual humano posee para lograr eliminar información de una imagen, la cual el ojo no será capaz de detectar.

El primero de ellos se basa en que, no tenemos la misma capacidad para detectar las variaciones de crominancia que las variaciones de luminancia, en otras palabras, no somos tan sensibles a la variación de color en una imagen como lo somos a los cambios en la intensidad de luz. La segunda consiste en que, somos capaces de detectar ligeros cambios en el tono entre dos zonas de color adyacentes, pero cuando dicha diferencia de color es grande, ésta no tiene por qué ser codificada de forma precisa.

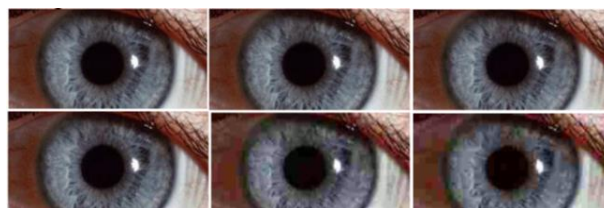


Figura 1. Imagen comprimida en JPEG con varias tasas de compresión y calidad.

Como se puede observar en la figura 1, el estándar JPEG otorga una gran gama de compresiones que pueden variar con factores de calidad desde 1 a 100. Con un grado de compresión de 1, la imagen es prácticamente la misma y obtenemos un archivo con un tamaño muy parecido a la imagen original. En cambio, si utilizamos un grado de compresión igual a 100, obtendremos un archivo de muy poco tamaño, pero con un gran número de distorsiones por cada bloque [3]. El método consta de seis pasos principales, los cuales serán brevemente explicados en un apartado posterior, haciendo mayor énfasis en la matriz de cuantización [4,5].

En el presente trabajo se desarrolla un método que nos permitirá determinar si una imagen BMP fue almacenada previamente en formato JPEG, de ser así, también se podrá estimar la tabla de cuantización con la cual fue comprimida dicha imagen, esto se realiza mediante un análisis estadístico al inspeccionar cuidadosamente todos los grupos de coeficientes DCT en cada uno de los bloques de  $8 \times 8$  píxeles de la imagen. Después, cada bloque es analizado para determinar si los valores de sus píxeles son valores truncados de la Transformada Inversa del Coseno (IDCT) dado un conjunto de coeficientes cuantificados mediante la matriz de cuantización extraída.

## II. DESARROLLO

La figura 2 muestra el diagrama a bloques general del estándar JPEG, los cuales serán descritos a continuación [1], [3-5].

### A. Cambio de espacio de color de RGB a YCbCr

El sistema visual humano es muy sensible a pequeños cambios en la luminancia (Y), pero no en la crominancia (CbCr), por ende, la parte de la crominancia puede perder gran cantidad de datos al ser comprimida sin que nosotros podamos notar algún

cambio aparente, cabe recalcar que este paso es opcional pero muy importante, ya que el resto del algoritmo trabaja de manera individual en cada componente de color.

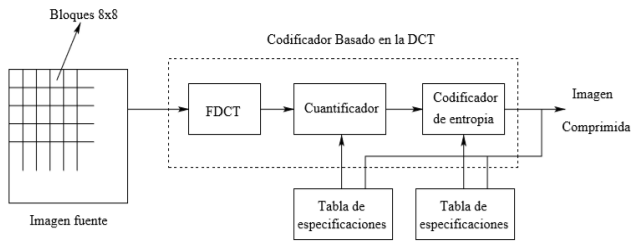


Figura 2. Esquema del codificador basado en la DCT [4].

Las ecuaciones matemáticas utilizadas para pasar del espacio de color RGB al espacio de YCbCr se muestran en (1).

$$\begin{aligned} Y &= (77/256) R + (150/256) G + (29/256) B \\ Cb &= -(44/256) R - (87/256) G + (131/256) B + 128 \\ Cr &= (131/256) R - (110/256) G - (21/256) B + 128 \end{aligned} \quad (1)$$

### B. Submuestreo

Mediante el submuestreo se puede reducir la resolución de algunos canales en los que se ha separado la imagen (luminancia y cromancia), con diferentes esquemas como, por ejemplo:

- 4:4:4 No hay submuestreo.
- 4:2:2 Reducimos a la mitad los canales de cromancia (Reducción en el eje X).
- 4:2:0 Reducimos a la cuarta los canales de cromancia (Reducción tanto en el eje X como en el Y).

Esto debido a los defectos que presenta el sistema visual humano con respecto a las variaciones de color y la densidad de la luz en una imagen.

### C. Transformada del Coseno Discreta (DCT)

El propósito básico de la DCT es convertir una señal de información (del dominio espacial) a información numérica en el dominio de la frecuencia (o dominio espectral), de tal manera que la información de la imagen exista en una forma cuantitativa para ser manipulada en la compresión mediante el siguiente procedimiento:

- Primero dividimos la imagen en bloques de 8x8 píxeles, que se procesan de manera independiente.
- Después cada bloque se transforma del dominio espacial al dominio de la frecuencia a través de la DCT, empleando (2).

$$F(u, v) = \frac{1}{4} C(u)C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos\left(\frac{(2x+1)u\pi}{16}\right) \cos\left(\frac{(2y+1)v\pi}{16}\right) \quad (2)$$

El resultado obtenido de esta transformación sirve para separar los componentes de baja y alta frecuencias tanto en la dirección horizontal como en la vertical. El bloque resultante (del mismo tamaño que el bloque original) se divide por la matriz de cuantificación, y cada entrada es redondeada. Los coeficientes de la matriz de cuantificación se diseñan a menudo para mantener ciertas frecuencias en la fuente para evitar pérdidas en la calidad de la imagen.

### D. Cuantización

Una vez efectuada la DCT se lleva a cabo la etapa de cuantización, en la cual se eliminan los coeficientes DCT con menor importancia. Esta transformación (con pérdidas), queda definida como la división de cada coeficiente de la matriz DCT de 8x8 entre el tamaño del escalón correspondiente, y el posterior redondeo al entero más próximo, como se muestra en (3).

$$F^Q(u, v) = \text{Round}\left(\frac{F(u, v)}{Q(u, v)}\right) \quad (3)$$

Si todos los pesos son 1, la transformación no tiene ningún efecto. En cambio, si los pesos aumentan al acrecentar la distancia respecto al origen, las frecuencias espaciales más altas se desechan con rapidez.

### E. Codificación de Huffman

Finalmente, en la etapa de codificación Huffman se convierten los valores de brillo de los píxeles de la imagen original en nuevos códigos de longitud variable, basado en su referencia de ocurrencia de la imagen. De esta manera, a los valores de brillo que ocurren más frecuentemente se les asignan los códigos más cortos y a los valores de brillo que ocurren con menos frecuencia se les asignan los códigos más largos. El resultado es que la imagen comprimida requerirá de menos bits para describir la imagen original.

## III. ALGORITMO PARA LA EXTRACCIÓN DE LA MATRIZ DE CUANTIZACIÓN

A continuación, se presenta el algoritmo para la obtención de la matriz Q de cuantización de una imagen previamente almacenada en JPEG, partiendo del modelo simplificado de dicha compresión.

Dada una imagen  $A$ , la cual es dividida en bloques no traslapados de 8x8 píxeles. Se calcula la transformada discreta del coseno (DCT) para cada bloque  $B_{orig}$  (con valores de píxeles enteros de 0 a 255), lo cual producirá 64 coeficientes de DCT.

Se denota el  $i$ -ésimo coeficiente de DCT del  $k$ -ésimo bloque como  $d_k(i)$ ,  $0 \leq i \leq 64$ ,  $k=1, \dots, T$ , donde  $T$  es el número total de bloques de 8x8 en la imagen  $A$ . Para cada bloque, cada uno de los 64 coeficientes de DCT es cuantizado y redondeado al



entero más cercano usando la matriz de cuantización  $Q$ , a este resultado redondeando se denota como  $D_k(i)$ .

$$D_k(i) = \text{Round} \left( \frac{d_k(i)}{Q(i)} \right) \quad (4)$$

Tomando como punto de partida esta ecuación, se analizan cuidadosamente todos los grupos de coeficientes DCT en cada uno de los bloques de  $8 \times 8$  que serán generados a partir de la imagen de la cual se pretende obtener la matriz de cuantización  $Q$ . Esto se lleva a cabo mediante el cálculo del error generado por la sumatoria de la relación que existe entre la resta del coeficiente DCT real menos el coeficiente DCT recomprimido de cada uno de los coeficientes de  $i$ ,  $d_1(i), \dots, d_T(i)$ .

$$E_i(q) = \frac{1}{T} \sum_{k=1}^T \left| d_k(i) - q * \text{Round} \left( \frac{d_k(i)}{q} \right) \right| \quad (5)$$

#### Descripción detallada del Algoritmo

Como entrada del algoritmo se tiene una imagen  $A$  con dimensiones de  $M \times N$ , a la cual se le aplicarán todos y cada uno de los siguientes pasos:

1. Convertir la imagen en formato RGB al formato YCbCr, empleando (1).
2. Dividir la imagen en bloques no traslapados de  $8 \times 8$  píxeles omitiendo las últimas filas o columnas si las dimensiones de la imagen no son múltiplos de 8.
3. Organizar todos los bloques obtenidos en forma de lista la cual tendrá dimensiones de  $8 \times M \times N$  (sin tomar en cuenta las filas y columnas excedentes).
4. Remover de la lista todos y cada uno de los bloques saturados (un bloque se encuentra saturado si posee al menos un píxel con valor de gris de 0 ó 255).
5. Calcule los coeficientes de DCT,  $d_k(i)$ ,  $0 \leq i \leq 64$ ,  $k=1, \dots, T$  para todos los bloques insaturados de  $8 \times 8$  píxeles, empleando (2).
6. Para cada coeficiente  $i$  se calculará y graficará el error  $E_i(q)$  empleando (5), como función del paso de cuantización  $q=1, 2, \dots, n$ . Donde el valor de  $n$  depende enteramente de la calidad con la cual fue comprimida la imagen a analizar, por lo cual no se recomienda tomar valores demasiado bajos, ya que esto conllevará a un truncamiento inadecuado de los posibles valores de la matriz de cuantización estimada. En esta etapa se podrá observar en la gráfica (véase figura 5) generada por el error calculado para cada uno de los diferentes pasos de cuantización que existen diversos mínimos locales, los cuales se emplean para determinar el paso correcto de cuantización para ese coeficiente de DCT.
7. Para cada coeficiente  $i$  se estimará el paso correcto de cuantización el cual será el mayor valor de  $q$  que genere

un mínimo local y que además sea múltiplo entero de todos aquellos valores de  $q$  anteriores que hayan generado una caída local.

8. En el paso anterior, existen algunos casos en los cuales la afirmación antes mencionada puede no ser verdadera, ya que, es posible que se generen algunos mínimos locales “fantasma” para valores de  $q$  más grandes que el paso correcto de cuantización, no obstante, esos mínimos son significativamente más grandes que el resto de los mínimos correctos. Para solucionar lo anterior, se asigna un umbral que logra filtrar satisfactoriamente estos mínimos “fantasma”. Este umbral dependerá de los coeficientes de DCT que hayan generado un mínimo local en cada uno de los 64 errores calculados, dado por la fórmula  $\mu_i + 3\sigma_i$ , donde  $\mu$  y  $\sigma$  son la media y la desviación estándar respectivamente.

Si el vector  $E_i(q)$  no posee ningún mínimo local a excepción de  $q=1$ , inspeccionaremos la diferencia que existe entre  $E_i(1)$  y  $E_i(2)$  para decidir si el paso correcto de cuantización será 1 ó 2. Para este tipo de casos existe una regla empírica obtenida a base de pruebas que arroja muy buenos resultados. Si  $E_i(1) < 0.6 * E_i(2)$ , se concluye que el paso de cuantización es igual a 1, de otra manera será 2.

#### IV. PRUEBAS Y RESULTADOS

Para la evaluación del algoritmo se realizaron diversas pruebas que involucran distintos formatos de imagen y la comprobación de las premisas presentadas en el algoritmo de la sección III.

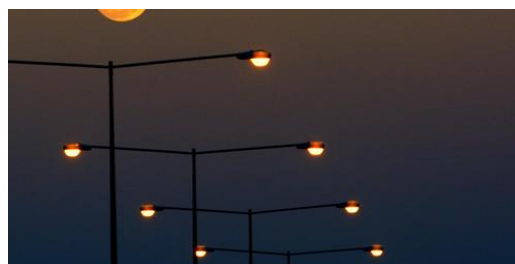


Figura 3. Imagen “calle” en formato BMP.



Figura 4. Imagen “calle” comprimida en JPEG con una calidad de 50%.

Para la primera prueba, la imagen mostrada en la figura 3 se comprimió mediante JPEG con un factor de calidad  $Q=50$ , posteriormente se almacenó en formato PNG y haciendo uso

del algoritmo propuesto se estimó la matriz de cuantización con la cual se realizó la compresión. En la tabla 1 se muestra la matriz de cuantización estándar empleada para  $Q=50$ .

Tabla 1. Matriz de Cuantización estándar para  $Q=50$ .

16	12	14	14	18	24	49	72
11	12	13	17	22	35	64	92
10	14	16	22	37	55	78	95
16	19	24	29	56	64	87	98
24	26	40	51	68	81	103	112
40	58	57	87	109	104	121	100
51	60	69	80	103	113	120	103
61	55	56	62	77	92	101	99

En la figura 5 se muestra la grafica del error  $E_1(q)$  para el coeficiente  $i=1$  (coeficiente DC) empleando (5), como función del *paso* de cuantización  $q=1,2, \dots, n$ , aplicado a la imagen mostrada en la figura 3 la cual se comprimió mediante JPEG con un factor de calidad  $Q=50$ . En la figura 5 se aprecian los distintos mínimos locales obtenidos para diferentes pasos de cuantización, también se puede observar que el paso correcto de cuantización es  $q=16$ .

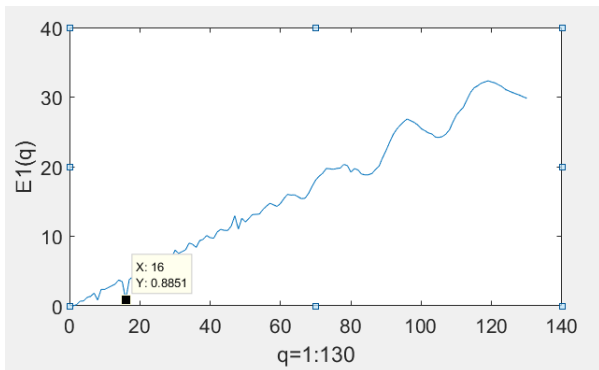


Figura 5. Gráfica del error  $E_1(q)$  para la imagen mostrada en la figura 4.

En las figuras 6, 7 y 8 se muestran las gráficas del error  $E_2(q)$ ,  $E_3(q)$ , y  $E_4(q)$  respectivamente para los coeficientes  $i=1, 2, 3$  y  $4$ , como función del *paso* de cuantización  $q=1,2, \dots, n$ , aplicados a la imagen mostrada en la figura 3 la cual se comprimió mediante JPEG con un factor de calidad  $Q=50$ . En las cuales se aprecian los distintos mínimos locales obtenidos para diferentes pasos de cuantización, también se puede observar que los pasos correctos de cuantización son  $q=12$ ,  $q=11$  y  $q=10$ .

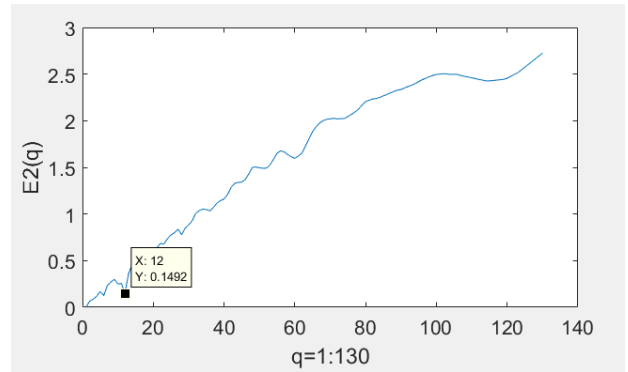


Figura 6. Gráfica del error  $E_2(q)$  para la imagen mostrada en la figura 4.

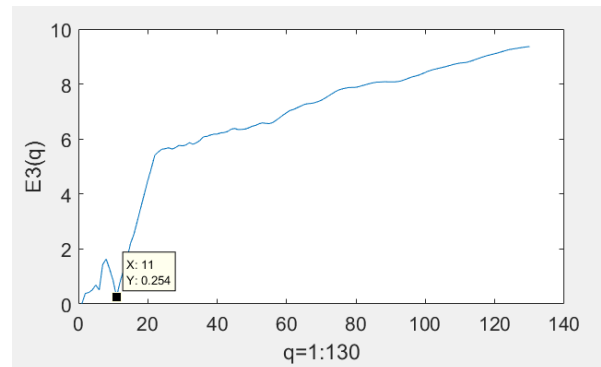


Figura 7. Gráfica del error  $E_3(q)$  para la imagen mostrada en la figura 4.

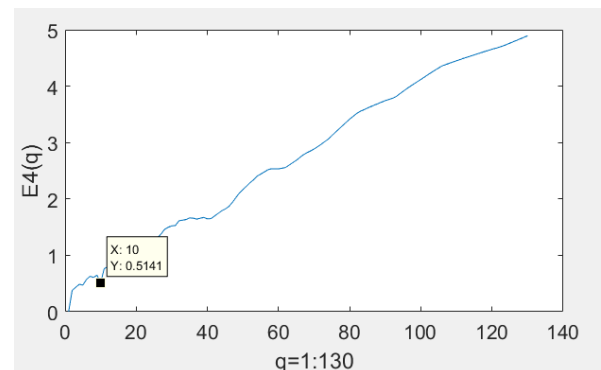


Figura 8. Gráfica del error  $E_4(q)$  para la imagen mostrada en la figura 4.

Tabla 2. Matriz de Cuantización estimada por el algoritmo propuesto para la imagen de la figura 4.

16	12	14	14	18	24	49	1
11	12	13	17	21	34	1	1
10	14	16	22	37	52	3	1
16	19	24	28	54	64	1	1
24	26	40	50	1	1	1	1
40	58	57	3	1	1	1	1
51	58	1	74	1	1	1	1
60	56	51	1	1	1	1	1

La segunda prueba consiste en aplicar el mismo proceso de estimación de la matriz de cuantización a una imagen BMP que no ha sido comprimida.



Figura 9. Imagen BMP sin compresión previa.

En la figura 10 se muestra la gráfica del error  $E_1(q)$  para el coeficiente  $i=1$  (coeficiente DC), como función del *paso* de cuantización  $q=1,2, \dots, n$ , aplicado a la imagen BMP sin compresión previa mostrada en la figura 9. En la figura 10 se aprecian que el único mínimo se obtiene en  $q=1$ .

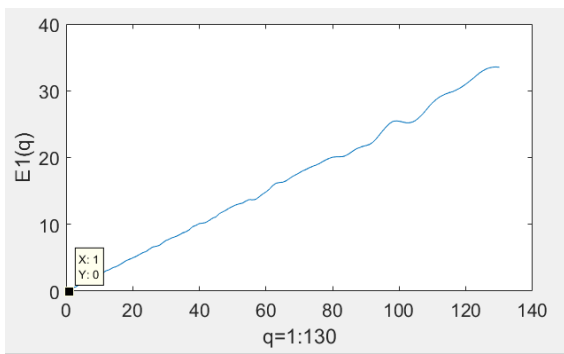


Figura 10. Gráfica del error  $E_{60}(q)$  para la imagen mostrada en la figura 9.

En la figura 11 se muestra la gráfica del error  $E_{60}(q)$  para el coeficiente  $i=60$ , aplicado a la imagen BMP sin compresión previa mostrada en la figura 9. donde se aprecia que el único mínimo se obtiene en  $q=1$ .

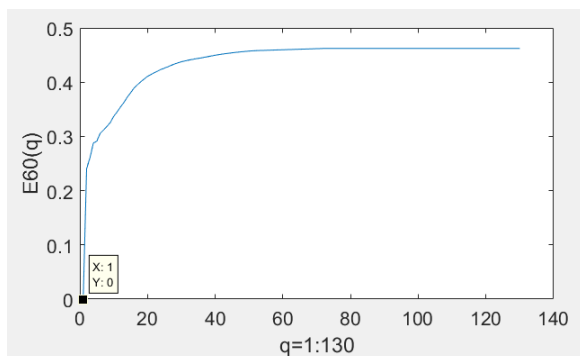


Figura 11. Gráfica del error  $E_{60}(q)$  para la imagen mostrada en la figura 9.

Tabla 2. Matriz de Cuantización estimada por el algoritmo propuesto para la imagen de la figura 7.

1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1

## V. CONCLUSIONES

En el presente artículo se propone y evalúa un algoritmo para estimar la matriz de cuantización de una imagen que haya sido previamente almacenada en formato JPEG. Los resultados demuestran que los valores obtenidos presentan pequeñas diferencias cuando se comparan con las matrices de cuantización estándar, esto se debe principalmente a que JPEG es un algoritmo de compresión con pérdidas, por lo cual, recuperar de manera exacta todos esos datos perdidos durante la compresión es una tarea prácticamente imposible para métodos similares de estimación. Cabe mencionar que, con la estimación de esta matriz de cuantización la cual es bastante cercana a la matriz de cuantización original, es posible detectar cambios o alteraciones realizadas a una imagen que haya sido previamente almacenada en este formato, también se puede llevar a cabo un estegoanálisis de la imagen para la detección de información oculta, y puede emplearse para mejorar la invisibilidad de marcas de agua.

## AGRADECIMIENTOS

Agradecemos al Instituto Politécnico Nacional, al Consejo Nacional de Ciencia y Tecnología de México, a la COFAA del IPN y a la Beca de Estimulo Institucional de Formación de Investigadores (BEIFI) del IPN por el apoyo otorgado para el desarrollo de este trabajo.

## REFERENCIAS

- [1] Charrier M., Cruz D.S. and Larsson M., "The next millennium compression standard for still images. Multimedia Computing and Systems", IEEE International Conference, 1999, pp. 131-132.
- [2] E. Molina Domínguez, "Sistema para compresión de imágenes en Matlab, a través de una serie de GUIs", Tesis de Licenciatura, Universidad de las Américas Puebla, Cholula, Puebla, México, 2004.
- [3] J. G. Ramírez Alfaro, Y. P. Orozco Zamora, C. Cruz Ramos, R. Reyes Reyes, "Algoritmo de extracción de la matriz de cuantización en archivos JPEG", ROC&C, Noviembre del 2012.
- [4] D. Salomon. (1938, June). Data. Compression (3rd edition). [online]. Available: [http://apachetechnology.in/ati/www/KC/dw/Saloman%20Data\\_Compression\\_Complete\\_Reference.pdf](http://apachetechnology.in/ati/www/KC/dw/Saloman%20Data_Compression_Complete_Reference.pdf) [Feb 21, 2019].
- [5] J. Fridrich, M. Goljan, and R. Du, "Steganalysis based on JPEG compatibility," SPIE Multimedia Systems and Applications IV, Denver, CO, August 20–24, 2001.

# Generación de Imágenes HDR a Partir de Imágenes LDR

*Maria T. Quiñonez Carbajal, Ricardo U. Osorno Ortiz, Clara Cruz Ramos,  
Rogelio Reyes Reyes, Volodymyr Ponomaryov*

Instituto Politécnico Nacional, ESIME Unidad Culhuacán, Ingeniería en Computación  
Tel. (55) 5729-6000 Ext. 73264 Fax (55) 56562058

Av. Santa Ana No. 1000, Col. San Francisco Culhuacán, CP. 04430, Ciudad de México.  
email: osournouriel@gmail.com, teresaquinonez96@gmail.com, ccruzra@ipn.mx, rreyesre@ipn.mx

**Resumen-** Actualmente se cuenta con monitores capaces de reproducir contenido en Alto Rango Dinámico (HDR) de luminancias, sin embargo, no son explotados al máximo, debido a que las cámaras fotográficas o de video con el que es capturado el contenido no captan todos los rangos de exposición. La fotografía en HDR permite tener imágenes más acordes a las visualizadas por el Sistema Visual Humano (HSV), ya que éste es capaz de ver hasta 20 pasos de iluminación en una escena con una adaptación visual extrema, mientras que las imágenes tradicionales de bajo rango dinámico (LDR) solo son capaces de capturar hasta 8 pasos de iluminación. Basado en esta problemática se propone un algoritmo computacional que permita generar imágenes HDR a partir de imágenes LDR basado en la técnica de mapa de bits de umbral mediano y la combinación de correlación cruzada normalizada ponderada para poder distinguir mejor los detalles y niveles de exposición en condiciones de abundante o escasa luz. Los resultados obtenidos demuestran una mayor visibilidad de detalle en regiones de la imagen en condiciones de luz extrema.

**Palabras clave—** Alto Rango Dinámico (HDR), Sistema Visual Humano (SVH), pasos de iluminación, adaptación visual.

## I. INTRODUCCIÓN

Actualmente las imágenes de alto rango dinámico nos permiten disfrutar de imágenes más vivas, más cautivadoras y más realistas, con las cuales podemos analizar y extrapolar los mejores detalles de una imagen aún en las sombras y altas luces como nunca antes, lo que significa que pueden manejar más detalles en las partes más brillantes y oscuras de una imagen, además de una amplia gama de colores, lo que permite inspirar a los profesionales y explotar al máximo los monitores HDR de última generación. En las últimas dos décadas, las imágenes HDR han revolucionado el campo de gráficos por computadora y otras áreas como la fotografía, procesamiento de imágenes, realidad virtual, efectos visuales y la industria de los videojuegos [1, 2].

Las imágenes HDR se pueden crear fundamentalmente de dos maneras. El primer método utiliza algoritmos de representación y otras técnicas de gráficos por computadora, el segundo método utiliza cámaras fotográficas convencionales (LDR) para generar imágenes HDR. Esto se puede lograr fotografiando una escena estática varias veces, donde para cada cuadro se varía el tiempo de exposición. Esto conduce a una

secuencia de imágenes, que se pueden combinar en una sola imagen HDR [3]. Este enfoque, generalmente, requiere que el sujeto permanezca inmóvil entre tomas, y la cámara también debe colocarse en un trípode. Esto limita el rango de fotografías que se pueden tomar. Afortunadamente, existen varias técnicas que alinean las imágenes, eliminan los fantasmas y reducen el efecto del destello de la lente, expandiendo así el rango de fotografías HDR que se pueden crear.

Al tomar exposiciones múltiples, cada imagen en la secuencia tendrá diferentes píxeles expuestos adecuadamente y otros píxeles subexpuestos o sobreexpuestos. Sin embargo, cada píxel se expondrá correctamente en una o más imágenes en la secuencia. Por lo tanto, es posible y deseable ignorar los píxeles muy oscuros y muy brillantes en los siguientes cálculos [4].

Yongqing Huo, Fan Yang y Chao Li, proponen una técnica la cual procesa imágenes en bajo rango dinámico para generar imágenes en alto rango dinámico [5]. El sistema, realiza un preprocesamiento de las áreas más luminosas de la imagen para evitar errores por medio del análisis de componentes principales (por sus siglas en inglés, PCA), y una expansión lineal basada en la luminancia de la imagen y la intensidad del monitor HDR. Los resultados presentan una alta robustez para imágenes con áreas muy luminosas, sin embargo el algoritmo presenta una alta complejidad computacional.

Andrey Vavilin y Kang-Hyun Jo generan una imagen HDR usando propiedades locales tales como la entropía, intensidad media, desviación de intensidad y nivel de detalles [6]. Esta técnica consiste de segmentar la imagen en regiones rectangulares clasificándolas como regiones de bajo rango dinámico o de alto rango dinámico, estas últimas siendo iteradas hasta que haya un valor determinado de píxeles en esa región. Por último se realiza la mezcla de exposición mediante una función gaussiana.

En este artículo se propone un algoritmo computacional que permita generar imágenes HDR a partir de imágenes LDR basado en la técnica de mapa de bits de umbral mediano y la combinación de correlación cruzada normalizada ponderada para poder distinguir mejor los detalles y niveles de exposición en condiciones de abundante o escasa luz.

## II. SISTEMA PROPUESTO

En esta sección se presenta la propuesta de solución para la generación de una imagen HDR a partir de un conjunto de imágenes LDR, cada una de ellas con diferente exposición. La Figura 1 muestra el diagrama a bloques general del sistema propuesto, el cual contiene dos bloques principales uno de *alineación* y otro de *mezcla de exposición*.

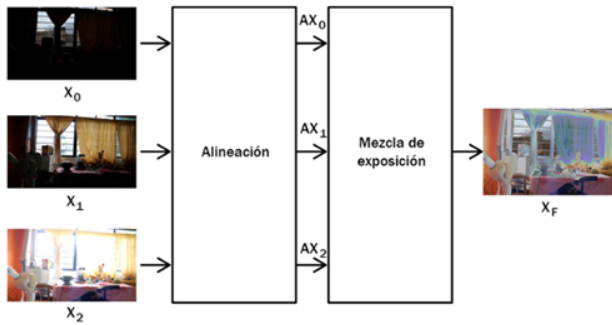


Fig. 1. Diagrama a bloques general del sistema propuesto.

## A. Bloque de alineación.

El bloque de *alineación* realiza un procesamiento de las imágenes de entrada  $X_0$ ,  $X_1$  y  $X_2$ , las cuales fueron capturadas con diferentes niveles de exposición (subexpuesta, normalmente expuesta y sobreexpuesta), obteniendo los desplazamientos de  $X_0$  y  $X_2$  con respecto a  $X_1$  representados por  $AX_0$  y  $AX_2$  respectivamente como se muestra en la Figura 2. El método de alineación utilizado está basado en el mapa de bits de umbral mediano [7].

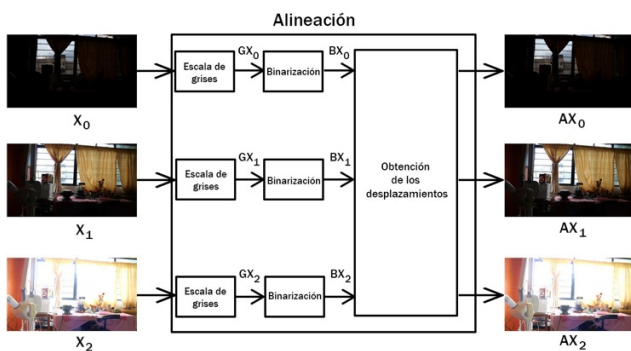


Fig. 2. Diagrama a bloques general de la alineación de las imágenes.

El sistema considera como entrada 3 imágenes en bajo rango dinámico  $X_0$ ,  $X_1$  y  $X_2$  las cuales son convertidas a imágenes a escala de grises ( $GX_0$ ,  $GX_1$  y  $GX_2$ ), posteriormente se calcula el MTB mediante una pirámide de imágenes de la siguiente manera:

Se aplica un proceso de binarización ( $BX_0$ ,  $BX_1$  y  $BX_2$ ). Una binarización convencional es aquella que convierte un pixel con 8 bits de información en un solo bit (0 ó 1) utilizando un valor de umbral fijo para toda la imagen.

Se crea una pirámide de imágenes con profundidad igual a 6 para cada exposición disminuyendo el muestreo en una proporción de 2 entre cada nivel. Se toman las dos primeras pirámides de imágenes (pirámide subexpuesta y normalmente expuesta), al nivel más pequeño de la imagen subexpuesta se le aplican nueve desplazamientos ((1,1), (1,0), (1, -1), (0,1), (0,0), (0, -1), (-1, 1), (-1,0), (-1, -1)) y se calculan las diferencias de cada uno con respecto al nivel más pequeño de la imagen normalmente expuesta. Se encuentra el desplazamiento con el menor error y este se duplica y se le aplica al siguiente nivel de la pirámide subexpuesta, una vez aplicado este desplazamiento se repite el proceso descrito anteriormente hasta llegar al nivel más grande de la pirámide, finalmente se obtiene la dirección de cambio final en las direcciones  $x$  e  $y$ . El proceso del párrafo anterior se repite entre la imagen (sobreexpuesta y normalmente expuesta) para obtener la dirección de cambio final entre ambas.

## B. Bloque de Mezcla de exposición.

El bloque de mezcla de exposición supone la mezcla de las tres imágenes alineadas para poder obtener una imagen HDR de salida  $X_F$ . Se realiza la mezcla de exposición para obtener una imagen con propiedades de rango dinámico extendido. Se toma como referencia la imagen normalmente expuesta, los pixeles subexpuestos que contenga esta imagen son combinados con los pixeles de la imagen sobreexpuesta y los pixeles sobreexpuestos son combinados con los pixeles de la imagen subexpuesta [7], la cual matemáticamente se expresa como (1).

$$X_F = \begin{cases} b(A_{X_{1ij}}) \cdot A_{X_{1ij}} + (1 - b(A_{X_{1ij}})) \cdot A_{X_{2ij}}, & \text{si } A_{X_{1ij}} < 127.5 \\ b(A_{X_{1ij}}) \cdot A_{X_{1ij}} + (1 - b(A_{X_{1ij}})) \cdot A_{X_{0ij}}, & \text{si } A_{X_{1ij}} > 127.5 \end{cases} \quad (1)$$

donde  $X_F$  es la imagen combinada,  $AX_0$ ,  $AX_2$  e  $AX_1$  son las imágenes subexpuesta, sobreexpuesta y normalmente expuesta alineadas,  $b$  es un coeficiente de mezcla, calculado por (2), cuyo valor está centrado en 127.5.

$$b(x) = \exp\left(-4 \frac{(x - 127.5)^2}{127.5^2}\right) \quad (2)$$

donde  $x$  es la posición del pixel.

## IV. PRUEBAS Y RESULTADOS

El sistema se desarrolló utilizando la plataforma de MathWorks Matlab® R2018a, en una computadora con procesador Intel (R) Pentium(R) CPU 2.80GHz, memoria RAM de 8.00GB y Sistema Operativo Windows 10 Home Single Language. Las imágenes digitales utilizadas para realizar las pruebas del algoritmo propuesto fueron tomadas con una cámara Canon T3i, con resolución de 5184x3456 pixeles en formato JPG, dichas imágenes presentan características variadas en luminosidad, textura y contraste.



A continuación, se muestran los resultados obtenidos del sistema propuesto con diferentes escenas. La Figura 3 muestra tres imágenes de prueba tomadas con diferentes niveles de exposición. La Figura 3(a) muestra la imagen normalmente expuesta de la escena “comedor”, la Figura 3(b) muestra una imagen sobreexpuesta de la misma escena y la Figura 3(c) la imagen subexpuesta de la misma escena, las cuales fueron capturadas con una diferencia de 2 pasos de iluminación (2EV) entre cada una de sus exposiciones y capturadas con la misma cámara, en donde se puede observar que en las tres imágenes existen detalles en la escena que no son perceptibles, lo que afecta la calidad y valor de la imagen.

La Figura 4 muestra la imagen HDR de la escena “comedor” resultante del sistema propuesto empleando las imágenes de la Figura 3, en la cual se pueden observar en los recuadros de color verde detalles de la escena difíciles de percibir en las imágenes normalmente expuestas.

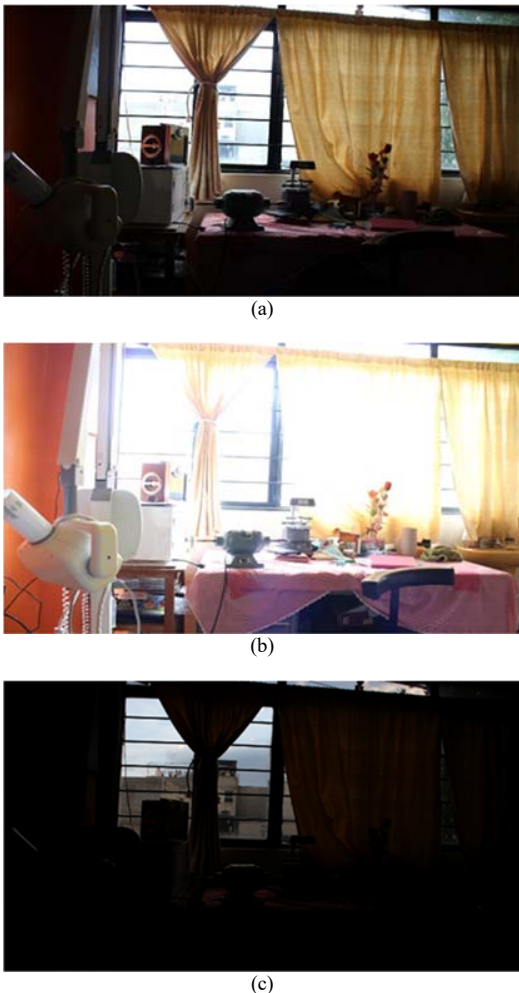


Fig. 3. Imágenes LDR de la escena “comedor”: (a) Imagen normalmente expuesta; (b) Imagen sobreexpuesta y (c) Imagen subexpuesta.

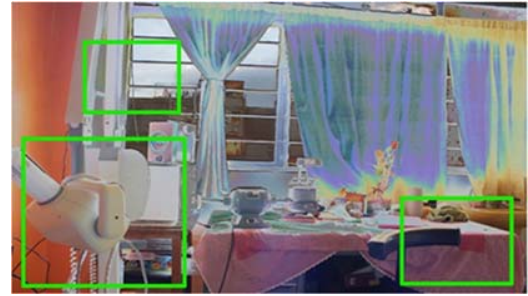


Fig. 4. Imagen HDR resultante de las imágenes LDR de la escena “comedor”.

La Figura 5 representa las imágenes de entrada para la escena “pared”. La figura 5(a) es la imagen normalmente expuesta, la figura 5(b) es la imagen sobreexpuesta y la figura 5(c) es la imagen subexpuesta, todas de la misma escena, mientras que la Figura 6 representa la imagen HDR resultante del sistema propuesto, en donde se pueden apreciar en los recuadros de color azul detalles mejorados a través del procesamiento, a pesar de que en las regiones sobreexpuestas hay píxeles que están fuera del rango de intensidad.

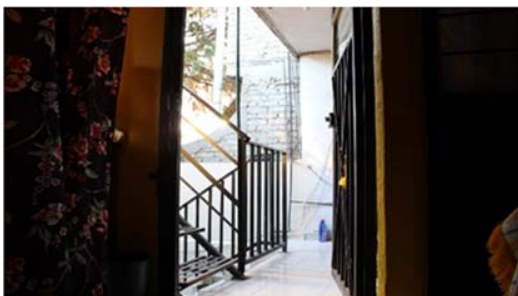


Fig. 5. Imágenes LDR de la escena “pared”: (a) Imagen normalmente expuesta; (b) Imagen sobreexpuesta y (c) Imagen subexpuesta.

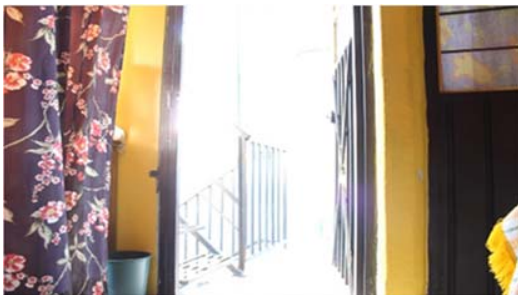


Fig. 6. Imagen HDR resultante de las imágenes LDR de la escena "pared".

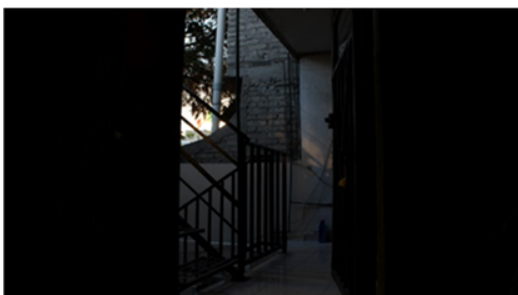
La Figura 7 representa las imágenes LDR capturadas a diferentes niveles de exposición de la escena "pasillo", en la que la imagen normalmente expuesta (Figura 7a) tiene detalles que no se pueden observar por la gran diferencia de luminosidad al momento de su captura.



(a)



(b)



(c)

Fig. 7. Imágenes LDR de la escena "pasillo": (a) Imagen normalmente expuesta; (b) Imagen sobreexpuesta y (c) Imagen subexpuesta.

La Figura 8 muestra la imagen HDR resultante del sistema propuesto de la escena "pasillo" en la cual, los recuadros de color púrpura muestran las regiones en las cuales se mejoraron los detalles, en el recuadro central se puede observar que aunque

hay pixeles fuera de rango, se alcanzan a distinguir texturas, mismas que no se aprecian en la imagen sobreexpuesta 7b.



Fig. 8. Imagen HDR resultante de las imágenes LDR de la escena "pasillo".

## V. CONCLUSIONES

El presente artículo propone y desarrolla un algoritmo que genera una imagen HDR a partir de tres imágenes LDR (normalmente expuesta, subexpuesta y sobreexpuesta), capturadas todas de la misma escena y con la misma cámara, pero con 2 pasos de iluminación de diferencia entre cada una de ellas. Las imágenes HDR resultantes del sistema propuesto amplían las propiedades del rango dinámico de las imágenes LDR, lo que permite percibir detalles en regiones de la imagen en donde el nivel de luminosidad es extremo y que en la captura convencional de una imagen LDR no se logran percibir.

## AGRADECIMIENTOS

Agradecemos al Instituto Politécnico Nacional, al Consejo Nacional de Ciencia y Tecnología de México, a la COFAA del IPN y a la Beca de Estimulo Institucional de Formación de Investigadores (BEIFI) del IPN por el apoyo otorgado para el desarrollo de este trabajo.

## REFERENCIAS

- [1] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec. High Dynamic Range Imaging Acquisition, Display and Image-Based Lighting. Second edition. USA: Morgan Kaufmann, 2010.
- [2] E. Cruz, "Aplicación de Súper-Resolución para el Mejoramiento de Algoritmos de Unión de Imágenes", tesis, Escuela Superior de Ingeniería Mecánica y Eléctrica Unidad Culhuacan, I.P.N., México, Ciudad de México, 2018.
- [3] F. Banterle, A. Artusi, K. Debattista, A. Chalmers. Advanced High Dynamic Range Imaging. Second edition. Boca Raton: Taylor & Francis, 2018.
- [4] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec. High Dynamic Range Imaging Acquisition, Display and Image-Based Lighting. Second edition. USA: Morgan Kaufmann, 2010.
- [5] Y. Hou, F. Yang, Chao Li "HDR image generation from LDR image with highlight removal", IEEE International Conference on Multimedia & Expo Workshops (ICMEW), China, 2015.
- [6] A. Vavilin, K. Jo, "Recursive HDR image generation from differently exposed images based on local image properties", International Conference on control, automation and systems, Seoul, Korea, October 2008.
- [7] A. Vavilin, K. Jo, "Fast HDR generation from multi-exposed multiple-view LDR images", 3rd. European Workshop on Visual Information Processing, Paris, France, 2011.



# Sintonización Óptima Fuera de Línea del Control Disparado por Eventos para un Robot Manipulador

Saul Enrique Benitez-Garcia,  
Miguel Gabriel Villarreal-Cervantes  
Laboratorio de Diseño Mecatrónico Óptimo,  
Departamento de Posgrado,  
CIDETEC, Instituto Politécnico Nacional,  
Av. Juan de Dios Bátiz s/n, 07700,  
Ciudad de México, México  
Email: sbenitezg1100@egresado.ipn.mx,  
mvillarrealc@ipn.mx

Efrén Mezura-Montes  
Centro de Investigación en Inteligencia Artificial,  
Universidad Veracruzana,  
Lomas del Estadio s/n, 91000  
Xalapa, Veracruz, México  
Email: emezura@uv.mx

**Resumen**—En años recientes ha surgido un fuerte interés de los investigadores por el Control Disparado por Eventos (CDE), que propone reducir el empleo periódico de los recursos de cómputo (uso del canal de comunicación y del procesador). Aún cuando este controlador ofrece disminuir el empleo de los recursos del sistema, requiere de una adecuada sintonización de sus parámetros para que el sistema realice de forma apropiada la tarea deseada. Por este motivo, en el presente trabajo se muestra la sintonización fuera de línea de un CDE con base en un enfoque de optimización, aplicado a un robot manipulador para el desarrollo de tareas de regulación. Para realizar la sintonización óptima del CDE, se detalla el modelo dinámico del robot manipulador, así como el diseño del CDE con base en una función de control de Lyapunov, necesarios para la formulación matemática del problema de optimización. El problema de optimización es resuelto por medio de una variante del algoritmo evolutivo de Evolución Diferencial, donde la solución obtenida es simulada y evaluada con base en cinco índices de rendimiento, que consideran el error en estado estacionario, el error transitorio y el número de actualizaciones de la estrategia de control.

**Palabras clave**—Control Disparado por Eventos, Sintonización óptima, Problema de optimización, Robot manipulador.

## I. INTRODUCCIÓN

Con la llegada de sistemas de cómputo de alto rendimiento a bajo costo y sistemas de medición de alta precisión se ha propiciado el crecimiento en el desarrollo y aplicación de las teorías de control avanzada y la robótica. Este avance se presenta en conjunto, ya que los robots presentan modelos no lineales, lo que ha servido como un buen caso de estudio para ejemplificar los conceptos generales del análisis y diseño de teorías de control avanzadas [1]. En consecuencia, los sistemas robóticos presentan una mejora significativa en su rendimiento. Particularmente, las teorías de control son empleadas en su mayoría en robots manipuladores, debido a su extenso campo de aplicación en el sector industrial [2], médico [3], de servicios [4], académico y científico [5], ello en comparación con los robots móviles [6]. La principal razón radica en su capacidad de manipular materiales, piezas, herramientas o dispositivos especializados.

Los sistemas de control actuales se implementan principalmente en sistemas digitales (control por computadora) [7], debido a que se caracterizan por ser sencillos de implementar, de modificar y además presentan un buen rechazo ante el ruido. Los sistemas de control por computadora pueden a su vez dividirse en dos: Control Disparado por Tiempo (CDT) y Control Disparado por Eventos (CDE). El primer enfoque consiste en realizar la medición de los parámetros del sistema y la actualización de la señal de control de manera periódica [8]; aunque este enfoque es el más empleado, la desventaja que posee, es la frecuencia de actualización de la señal de control, aún cuando el sistema ha alcanzado la posición de referencia. El segundo enfoque se basa en la ejecución de la estrategia de control mediante la activación de una función de evento, cuya activación se presenta cuando una restricción de rendimiento del sistema se viola.

El enfoque de CDE ofrece estabilidad y una disminución en el número de actualizaciones de la señal de control. En consecuencia, el uso del canal de comunicación y de la carga computacional disminuye y por consiguiente el consumo energético. En [9] se presenta una comparativa entre dos estrategias de control aplicadas a un robot móvil (3,0), por un lado un CDE y por otro lado el control par calculado. En dicho trabajo se obtiene una disminución del 23.73% en el número de actualizaciones de la señal del CDE en comparación con el requerido por el control par calculado (CTC). En [10] se propone el diseño de un controlador disparado por eventos basado en un control robusto. La propuesta de este diseño se valida mediante simulación en un robot de dos grados de libertad tipo SCARA, donde los resultados mostraron convergencia asintótica con o sin presencia de alguna perturbación. En [11] se propone estabilizar un péndulo invertido mediante una estrategia de control disparada por eventos y un mecanismo de activación obtenida mediante la metodología mostrada en [8], basada en un enfoque de estabilidad de Lyapunov. Los experimentos y análisis de resultados en tiempo real, muestran una reducción aproximada entre 98% y 50%, esto en comparación con el

esquema clásico que se presenta en el CDT.

Una de las principales problemáticas a las cuales los diseñadores de control se enfrentan posterior al diseño de sistema de control, se encuentra durante la implementación del esquema de control seleccionado, ya que durante esta etapa se requiere de un ajuste de los parámetros del controlador de tal forma que el sistema en lazo cerrado presente un comportamiento deseado. Esto ha motivado múltiples investigaciones referentes a la sintonización de controladores, que pueden ser clasificados como se muestra en [12]: *i)* métodos analíticos, donde las ganancias del controlador se obtienen por medio de un análisis de estabilidad del sistema en lazo cerrado; *ii)* métodos heurísticos, donde se considera la experiencia del diseñador para la sintonización manual del controlador; *iii)* métodos de optimización, donde se plantea el problema de sintonización como un problema de programación matemática y se resuelve mediante técnicas de optimización, para la obtención de parámetros fijos del controlador; y *iv)* métodos adaptativos, donde la sintonización se realiza en tiempo de ejecución, empleando un proceso de identificación en combinación con alguno de los métodos previos.

En este trabajo se presenta un fuerte interés por la sintonización de controladores mediante métodos de optimización fuera de línea, debido a que hoy en día las características de rendimiento deseadas en el diseño de control pueden ser tratadas como un problema de optimización como se muestra en [13], donde se considera el problema de regulación para un robot manipulador de dos grados de libertad y se propone realizar de forma sistemática el ajuste de las ganancias de un controlador PID. El ajuste de las ganancias se efectúa por medio de la técnica metaheurística PSO (del inglés Particle Swarm Optimización), con el que se obtuvieron resultados aceptables. En [14], el control PID del sistema robótico de dos grados de libertad se sintoniza proponiendo un problema de optimización multi-objetivo. El algoritmo evolutivo multi-objetivo NSGA-II se implementa para resolver el problema. Basándose en los resultados de la simulación, afirman que su enfoque es una alternativa viable. En [15] se presenta el ajuste de un control PID para un robot planar paralelo para el seguimiento de una trayectoria no lineal. El problema se plantea como un problema de optimización y se resuelve mediante el algoritmo de Evolución Diferencial. Las ganancias óptimas se aplicaron a un prototipo experimental y en simulación, donde los resultados obtenidos mostraron un seguimiento de trayectoria similar en el caso experimental y en simulación.

Debido al conflicto que existe entre el desempeño que ofrece el sistema de CDE para realizar una tarea y el número de activaciones de la función de evento, especialmente cuando fuerzas externas se encuentran presentes, es de especial interés obtener una sintonización adecuada del CDE de manera que se obtenga el mejor desempeño con el menor número de activaciones de evento durante el desarrollo de una tarea predefinida. Por tal motivo, en este trabajo de investigación se desarrolla con base en un enfoque de optimización la sintonización fuera de línea de un sistema de CDE aplicado en la simulación de un robot manipulador, con el propósito de

obtener un compromiso en la activación de eventos y el error en la tarea de regulación a varios puntos en el espacio.

En lo consecuente, en la Sección II se presenta el modelo matemático del robot manipulador empleado así como el diseño y análisis de estabilidad del CDE propuesto para este sistema. El planteamiento formal del problema de sintonización del CDE como un problema de optimización se define en la Sección III. Los resultados obtenidos al resolver el problema de optimización mediante una técnica de optimización se muestra en la Sección IV. Por último, en la Sección V se presentan las conclusiones del trabajo realizado.

## II. MODELADO Y CONTROL DE LA PLANTA

El modelo matemático del sistema a controlar es importante para el análisis y diseño del sistema de control. Por tal motivo, en esta sección se presenta el modelo matemático del robot manipulador utilizado en este trabajo, así como el diseño del control disparado por eventos para su estabilización.

### II-A. Modelo dinámico del robot manipulador

El diagrama esquemático del robot manipulador empleado en este trabajo se presenta en la Fig. 1, donde los parámetros del  $i$ -ésimo eslabón están dados por la distancia entre el eje de rotación al centro de masa  $l_{c_i}$ , el momento de inercia  $I_{z_i}$  con respecto al eje de movimiento  $z$ , la masa  $m_i$  y la longitud del eslabón  $l_i$ , con  $i = 1, 2, 3$ .

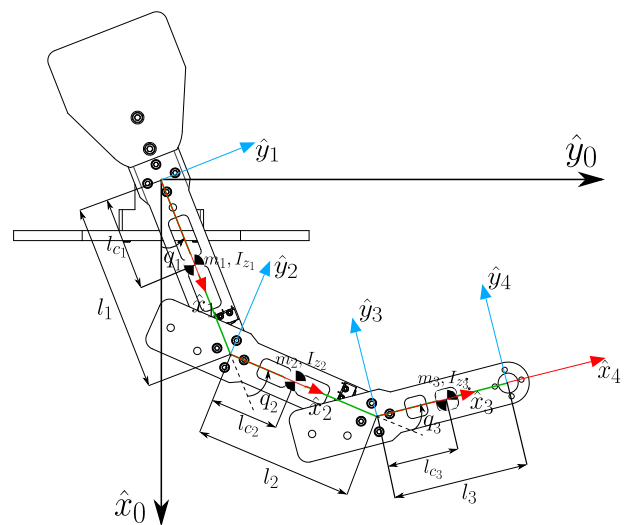


Figura 1. Diagrama esquemático del robot manipulador

El modelo dinámico del robot manipulador puede describirse en el espacio de estados como se muestra en (1), donde  $x = [x_1, x_2, x_3, x_4, x_5, x_6]^T = [q, \dot{q}] = [q_1, q_2, q_3, \dot{q}_1, \dot{q}_2, \dot{q}_3]^T \in \mathcal{R}^6$  es el vector de estados correspondiente a los desplazamientos  $q \in \mathcal{R}^3$  y velocidades angulares  $\dot{q} \in \mathcal{R}^3$ , expresados en el espacio articular de cada grado de libertad y  $\mathbf{0} \in \mathcal{R}^{3 \times 3}$  es una matriz de ceros. Para mayor información de los elementos que conforman a las funciones  $f(x)$  y  $g(x)$ , ver el Apéndice A.

$$\dot{x} = f(x) + g(x)u \quad (1)$$

donde:

$$f(x) = \begin{bmatrix} M^{-1}[-C[x_4, x_5, x_6]^T - G - F] \\ \mathbf{0} \\ M^{-1} \end{bmatrix} \in \mathcal{R}^6$$

$$g(x) = \begin{bmatrix} \mathbf{0} \\ M^{-1} \end{bmatrix} \in \mathcal{R}^{6 \times 3}$$

### II-B. Control Disparado por Eventos para el robot manipulador

En el presente trabajo, se propone emplear el enfoque de CDE mostrado en [8], donde la función de evento  $e(x, \bar{m})$  y la función de control  $k(\bar{m})$  se establecen en (2) y (3) respectivamente.

$$\bar{e}(x, \bar{m}) = -a(x) - b(x)k(\bar{m}) - \sigma \sqrt{a(x)^2 + \bar{\theta}(x)b(x)\Delta(x)b(x)^T} \in \mathcal{R} \quad (2)$$

$$k(\bar{m}) = -\Psi \Delta(x)b^T(x)\gamma(x) \in \mathcal{R}^n \quad (3)$$

con:

- $a(x) = \frac{\partial V}{\partial x} f(x)$  y  $b(x) = \frac{\partial V}{\partial x} g(x)$
- $\Delta(x) : \mathcal{X} \rightarrow \mathcal{R}^{p \times p}$  tal que  $\Delta(x) = \text{diag}(\delta_1(x), \delta_2(x), \dots, \delta_p(x))$  es una función suave y definida positiva en  $\mathcal{S} = \{x \in \mathcal{X} \mid \|b(x)\| \neq 0\}$
- $\bar{\theta}(x) := \omega(x)b(x)\Delta(x)b^T(x) - a(x)\omega(x)$  tal que  $\bar{\theta}(x) : \mathcal{X} \rightarrow \mathcal{R}$  es una función suave y  $\bar{\theta}(x)\|\Delta(x)\|$  se desvanece en el origen, asegurando la desigualdad  $\bar{\theta}(x)b(x)\Delta(x)b(x)^T + a(x)^2 > 0$  en  $\mathcal{S} - \{0\}$ .
- $\sigma$  es un parámetro de control ajustable en  $[0, 1]$ .
- $\Psi = \text{diag}(\psi_1, \psi_2, \dots, \psi_p)$  es una matriz diagonal de parámetros ajustables.
- $\gamma(x) \rightarrow \mathcal{R}$  es definido por:

$$\gamma(x) = \begin{cases} \frac{a(x) + \sqrt{a(x)^2 + \bar{\theta}(x)b(x)\Delta(x)b(x)^T}}{b(x)\Delta(x)b(x)^T} & \text{si } x \in \mathcal{S} \\ 0 & \text{si } x \notin \mathcal{S} \end{cases}$$

- $\bar{m} : \mathcal{X} \rightarrow \mathcal{R}^n$  es un parámetro de memoria, dado por:

$$\bar{m} = \begin{cases} x & \text{si } \bar{e}(x, \bar{m}) \leq 0, x \neq 0 \\ 0 & \text{en otra parte} \end{cases}$$

con  $x(0) = x_0$  y  $\bar{m}(0) = x(0)$

### II-C. Diseño de la Función de Control de Lyapunov y el Control Disparado por Eventos

Sea el cambio de variable definido como  $[e, \dot{\varphi}]^T$  y considere la Función de Lyapunov  $V : \mathcal{R}^n \rightarrow \mathcal{R}$ , dada por:

$$V(e, \dot{\varphi}) = \frac{1}{2} e^T P_1 e + \frac{1}{2} \dot{\varphi}^T P_2 \dot{\varphi} + e^T P_3 \dot{\varphi} \quad (4)$$

donde  $e = [\varphi - \bar{\varphi}]^T \in \mathcal{R}^3$  es el vector de error entre el vector de posición angular deseada  $\bar{\varphi}$  y el vector de posición angular actual  $\varphi = [x_1, x_2, x_3]^T$ , las matrices  $P_1 = \text{diag}[P_{11}, P_{12}, P_{13}]$ ,  $P_2 = \text{diag}[P_{21}, P_{22}, P_{23}]$  y  $P_3 = \text{diag}[P_{31}, P_{32}, P_{33}] \in \mathcal{R}^{3 \times 3}$  son matrices diagonales definidas

positivas y  $\dot{\varphi} = [x_4, x_5, x_6]^T$  es el vector de velocidad angular actual. Si

$$P_{1i} = \sqrt{\varrho_i \varrho_{i+3} + 2\sqrt{\varrho_i^3/\epsilon_i}} \quad (5)$$

$$P_{2i} = \sqrt{\varrho_{i+3}/\epsilon_i + \sqrt{4\varrho_i/\epsilon_i^3}} \quad (6)$$

$$P_{3i} = \sqrt{\varrho_i/\epsilon_i} \quad (7)$$

con  $\varrho_i, \varrho_{i+3} \in \mathcal{R}^+ \forall i = 1, 2, 3$ , entonces (4) es una FCL para el sistema dinámico (1) en el punto de equilibrio  $[e, \dot{\varphi}]^T = \mathbf{0} \in \mathcal{R}^6$ , considerando la estrategia de control dada por:

$$\mathbf{u} = \mathbf{M}(e + \bar{\varphi})(\bar{v}) + \mathbf{C}(e + \bar{\varphi}, \dot{\varphi})\dot{\varphi} + \mathbf{G}(e + \bar{\varphi}) \quad (8)$$

donde:

$$\bar{v} = -[\epsilon P_3 \quad \epsilon P_2] \begin{bmatrix} e \\ \dot{\varphi} \end{bmatrix} \quad (9)$$

y  $\epsilon = \text{diag}[\epsilon_1, \epsilon_2, \epsilon_3]$  es una matriz definida positiva de ganancias de retroalimentación.

Como la función de Lyapunov  $V$  (4) es una FCL para el robot manipulador dado por (1), es posible desarrollar la función de control de retroalimentación (3), que junto con la función de evento (2) estabilizan asintóticamente los estados  $[e, \dot{\varphi}]$  en el origen, donde:

$$a(x) = e^T P_1 \mathbf{I} \dot{\varphi} + \dot{\varphi}^T P_3 \mathbf{I} \dot{\varphi} \quad (10)$$

$$b(x) = e^T P_3 \mathbf{I} + \dot{\varphi}^T P_2 \mathbf{I} \quad (11)$$

con  $\mathbf{I} \in \mathcal{R}^{3 \times 3}$  como una matriz identidad.

## III. PLANTEAMIENTO DEL PROBLEMA DE OPTIMIZACIÓN

### III-A. Variables de diseño

El desempeño del CDE propuesto para este trabajo involucra los parámetros:  $\varrho_i, \varrho_{i+3}, \epsilon_i$ , con  $i = 1, 2, 3$ , necesarios para obtener los valores de las matrices  $P_1, P_2$  y  $P_3$  definidas en (5)-(7) respectivamente, de modo que se asegure que la relación de cambio de la FCL sea definida negativa, por lo tanto la función de Lyapunov (4) sea una FCL y consecuentemente los estados del sistema converjan asintóticamente al origen;  $\delta_i$  y  $\omega$  son parámetros a sintonizar considerados en la función de evento  $\bar{e}(x, \bar{m})$  definida en (2);  $\sigma$ , conocido como frecuencia de activación de la función de evento  $\bar{e}(x, \bar{m})$ ;  $\psi_i$  son parámetros que permiten un ajuste más preciso de la retroalimentación  $k(\bar{m})$  definida en (3). Por lo tanto, se establece como el siguiente vector de variables de diseño para el problema de optimización:

$$\bar{p} = [\varrho_i, \varrho_{i+3}, \epsilon_i, \delta_i(x), \sigma, \omega, \psi_i] \in \mathcal{R}^{17} \quad (12)$$

### III-B. Función de desempeño

La definición del problema de sintonización óptima del CDE se propone como un problema de diseño multi-objetivo, y se aborda como un problema de optimización mono-objetivo mediante una función de agregación. Para obtener esta función de agregación, se emplea el criterio de sumas ponderadas, el cual emplea parámetros constantes que indican la importancia de cada objetivo. En (13) se muestra la función de desempeño

$\bar{J}$  que se desea optimizar, donde  $\xi_1$  y  $\xi_2$  indican el grado de preferencia que se le proporciona a cada una de las funciones.

$$\bar{J} = \xi_1 \bar{J}_1(x) + \xi_2 \bar{J}_2(x) \quad (13)$$

El término  $J_1$  considera la precisión con la cual el sistema de control estabiliza al sistema robótico en una posición deseada (tarea de regulación). Tomando en cuenta que para la tarea de regulación lo más importante es disminuir el error cuando  $t \rightarrow \infty$  (estado estacionario), se considera la integral ponderada en el tiempo del error absoluto (ITAE del inglés Integral Time-weighted Absolute Error) mostrada en (14), donde  $\hat{e}_j = x_j - \bar{x}_j \forall j = 1, 2, 3$  es el error entre la posición actual  $x_j$  y la posición de referencia  $\bar{x}_j$ .

$$\bar{J}_1(x) = \sum_{j=1}^3 \int_0^{t_f} t |\hat{e}_j| dt \quad (14)$$

Por otro lado el término  $J_2$  mostrado en (15), toma en cuenta el número de actualizaciones de la señal de control, ya que el CDE ofrece disminuir el tráfico del canal de comunicación y el empleo del sistema de procesamiento.

$$\bar{J}_2(x) = \frac{1}{t_f} \int_0^{t_f} \eta(\bar{e}(t)) dt \quad (15)$$

donde:

$$\eta(\bar{e}(t)) = \begin{cases} 1 & \text{si } \bar{e} \leq 0 \\ 0 & \text{si } \bar{e} > 0 \end{cases}$$

### III-C. Restricciones de diseño

El problema de optimización queda sujeto a un conjunto de ecuaciones diferenciales que describen el modelo dinámico del robot manipulador. En (16) se muestra la restricción de igualdad  $h_1(\bar{p}, \bar{x}, t)$  que contempla el modelo dinámico del robot manipulador, considerando el reposo  $x(0) \in \mathcal{R}^n$  como el estado inicial del sistema.

$$h_1(\bar{p}, \bar{x}, t) : \frac{dx}{dt} - f(x) - g(x)u = 0 \quad (16)$$

donde:

$$u(\bar{p}, \bar{x}, t) = \begin{cases} u_{max} & \text{si } k(\bar{m}) > u_{max} \\ k(\bar{m}) & \\ u_{min} & \text{si } k(\bar{m}) < -u_{min} \end{cases} \quad (17)$$

Considerando el problema de regulación, se propone que el sistema de control estabilice al robot manipulador en  $n_p$  posiciones deseadas. Para tal propósito, en (18) se muestra la restricción de igualdad  $h_2(\bar{p}, \bar{x}, t)$  que contempla lo mencionado, donde  $t_p$  es el intervalo de tiempo en el cual el sistema de control deberá estabilizar al robot manipulador en cada posición y velocidad articular deseada y  $\rho_j \in \mathcal{R}^6 \forall j = 1, 2, 3, \dots, n_p$  es un vector de posiciones y velocidades angulares constantes en el espacio articular.

$$h_2(\bar{p}, \bar{x}, t) : \begin{cases} \bar{x} - \rho_1 = 0 & \text{si } t \in [0, t_p] \\ \bar{x} - \rho_2 = 0 & \text{si } t \in (t_p, 2t_p] \\ \vdots & \\ \bar{x} - \rho_{n_p} = 0 & \text{si } t \in ((n_p - 1)t_p, n_p t_p] \end{cases} \quad (18)$$

Por otro lado, con el objetivo de acotar el espacio de búsqueda de las soluciones, se consideran los límites inferior ( $\bar{p}_{min}$ ) y superior ( $\bar{p}_{max}$ ) de las variables de diseño para establecer la restricciones de límite  $g_1(\bar{p})$  y  $g_2(\bar{p})$ , como se muestra en (19) y (20) respectivamente.

$$g_1(\bar{p}) : \bar{p}_{min} - \bar{p} \leq 0 \quad (19)$$

$$g_2(\bar{p}) : \bar{p} - \bar{p}_{max} \leq 0 \quad (20)$$

### III-D. Formulación del problema de optimización

El problema de optimización consiste en encontrar el vector de diseño óptimo de parámetros de control  $\bar{p}^*$  que minimice la función objetivo (13), sujeto a la dinámica (16) del robot manipulador, a las posiciones y velocidades articulares de referencia (18) y a los límites en las variables de diseño (20) y (19). Por lo tanto, el problema formal de optimización se establece en (21):

$$\min_{\bar{p} \in \mathcal{R}^{17}} \bar{J}(\bar{p}) \quad (21)$$

Sujeto a:

$$h_1(\bar{p}, \bar{x}, t) = 0$$

$$h_2(\bar{p}, \bar{x}, t) = 0$$

$$g_1(\bar{p}) \leq 0$$

$$g_2(\bar{p}) \leq 0$$

## IV. RESULTADOS

### IV-A. Condiciones del proceso de optimización

El proceso de optimización es programado en MATLAB versión 9.1 sobre el sistema operativo Windows 10 de 64-bit, con un procesador Core i7 a 3.2 GHz con 16 GB de memoria RAM. El problema de sintonización óptima es resuelto con el algoritmo de Evolución Diferencial (ED) en su variante *Rand/1/Bin* (ver [16]), cuyas condiciones de operación fueron establecidas posterior a un proceso iterativo de prueba y error hasta obtener un desempeño adecuado, y quedan descritos a continuación:

- Número de individuos  $NP = 50$ .
- Número de generaciones máxima  $G_{max} = 1500$ .
- Factor de cruce  $CR = 0.63$ .
- Factor de escala  $F \in [0.3, 0.9]$  generado aleatoriamente en cada generación del algoritmo.
- Límites de las variables de diseño establecidos en la Tabla I.

Por otro lado, las condiciones de operación del problema de optimización se detallan a continuación:

- Índices de ponderación  $\xi_1 = 1000$  y  $\xi_2 = 0.01$ .
- Estado inicial del robot manipulador establecido en el reposo  $x(0) = [0, 0, 0, 0, 0, 0]^T$ .
- Número de posiciones y velocidades angulares deseadas de  $n_p = 4$ .
- Vectores de posiciones y velocidades deseadas  $\rho_j$  mostrados en la Tabla II, que representan puntos de equilibrio (PE) y puntos que no son de equilibrio (PNE) del sistema.
- Tiempo máximo de estabilización  $t_{pp} = 2s$ .

Tabla I  
 LÍMITES DE LAS VARIABLES DE DISEÑO

Parámetro	$\bar{p}_1$	$\bar{p}_2$	$\bar{p}_3$	$\bar{p}_4$	$\bar{p}_5$	$\bar{p}_6$	$\bar{p}_7$	$\bar{p}_8$	$\bar{p}_9$	$\bar{p}_{10}$	$\bar{p}_{11}$	$\bar{p}_{12}$	$\bar{p}_{13}$	$\bar{p}_{14}$	$\bar{p}_{15}$	$\bar{p}_{16}$	$\bar{p}_{17}$
$\bar{p}_{min}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$\bar{p}_{max}$	90E3	15E3	8E3	35	30	10	4	4	5	3	2	3	1	2	8	6	2

- Método de integración numérica de Euler con un tiempo de integración fijo de  $\Delta t = 1ms$ .

Tabla II

VECTORES DE POSICIÓN Y VELOCIDAD ANGULAR DESEADAS PARA EL PROCESO DE OPTIMIZACIÓN.

Posiciones y velocidades deseadas		Tipo de punto
Articular $\rho_j$	Cartesiana $\bar{\rho}_j$	
$\rho_1 = [\frac{\pi}{4}, \frac{\pi}{4}, \frac{\pi}{4}, 0, 0, 0]^T$	$\bar{\rho}_1 = [0.03, 0.36, \frac{3\pi}{4}, 0, 0, 0]^T$	PNE
$\rho_2 = [\pi, 0, 0, 0, 0, 0]^T$	$\bar{\rho}_2 = [-0.46, 0.00, \pi, 0, 0, 0]^T$	PE Inestable
$\rho_3 = [\frac{\pi}{2}, \frac{\pi}{4}, \frac{\pi}{3}, 0, 0, 0]^T$	$\bar{\rho}_3 = [-0.23, 0.25, \frac{13\pi}{12}, 0, 0, 0]^T$	PNE
$\rho_4 = [0, 0, 0, 0, 0, 0]^T$	$\bar{\rho}_4 = [0.46, 0, 0, 0, 0, 0]^T$	PE Estable

Tabla III

PARÁMETROS DINÁMICOS Y CINEMÁTICOS DEL ROBOT MANIPULADOR

Parámetro	Descripción	Valor	Unidades
$l_1$	Longitud de eslabón 1	0.18	$m$
$l_2$	Longitud de eslabón 2	0.15	$m$
$l_3$	Longitud de eslabón 3	0.13	$m$
$l_{c1}$	Longitud al centro de masa 1	-0.0679573760623	$m$
$l_{c2}$	Longitud al centro de masa 2	0.11864166275194	$m$
$l_{c3}$	Longitud al centro de masa 3	0.01042396720858	$m$
$I_{z1}$	Momento de inercia 1	0.07183176287911	$kgm^2$
$I_{z2}$	Momento de inercia 2	0.87694725642273	$kgm^2$
$I_{z3}$	Momento de inercia 3	0.00137475449407	$kgm^2$
$m_1$	Masa de eslabón 1	5.85264743808855	$kg$
$m_2$	Masa de eslabón 2	0.87694725642273	$kg$
$m_3$	Masa de eslabón 3	0.35093898923791	$kg$
$F_{v1}$	Fricción viscosa 1	1.16979065946303	$\frac{kg}{s}$
$F_{v2}$	Fricción viscosa 2	0.68395552317746	$\frac{kg}{s}$
$F_{v3}$	Fricción viscosa 3	0.80173520226598	$\frac{kg}{s}$

Es bien sabido que uno de los principales retos en la sintonización de controladores fuera de línea, es lograr un ajuste en simulación de los parámetros del controlador, tales que, proporcionen un adecuado comportamiento durante la etapa de experimentación en la planta real [17]. Por tal motivo, a menudo es necesario el empleo de esquemas que proporcionen al diseñador esta característica, tales como: Hardware-in-the-Loop (HiL), Model-in-the-Loop (MiL), Software-in-the-Loop (SiL) y Processor-in-the-Loop (PiL). El presente trabajo se acota al empleo del enfoque MiL, que consiste en incluir en el proceso de optimización un modelo matemático lo más aproximado posible a la planta real y para tal propósito, se incluye la fricción viscosa  $F_v = diag[F_{v1}, F_{v2}, F_{v3}] \in \mathcal{R}^{3 \times 3}$  al modelo matemático del robot manipulador (1). Los parámetros dinámicos y cinemáticos del robot manipulador obtenidos del proceso de identificación paramétrica se establecen en la Tabla III.

#### IV-B. Resultados del proceso de optimización

Una vez ejecutado el algoritmo de ED para resolver el problema de optimización configurado como se muestra en la subsección previa, se obtiene el vector de diseño mostrado en la Tabla IV. Este vector, se evalúa mediante experimentos en simulación, tomando en cuenta la siguiente configuración:

- Número de posiciones y velocidades deseadas  $n_p = 4$
- Vectores de posición y velocidad deseados  $\rho_j$  mostrados en la Tabla II.
- Tiempo máximo de estabilización  $t_{pp} = 5s$ .
- Método de integración numérica de Euler con un tiempo de integración fijo de  $\Delta t = 1ms$ .
- Estado de inicio del robot manipulador establecido en el reposo  $x(0) = [0, 0, 0, 0, 0, 0]^T$ .
- Parámetros dinámicos y cinemáticos del robot manipulador se muestran en la Tabla III.

En la Tabla V se muestran cinco índices de rendimiento: el máximo sobreimpulso (O), el tiempo de establecimiento (ST) al 5%, el error en estado estacionario (SSE), la magnitud del control aplicado ( $\|u\|$ ) y el porcentaje de eventos activados (ETP). En la Fig. 2 se aprecia el comportamiento del efector final en el sistema coordinado inercial  $\hat{x}_0 - \hat{y}_0$ , y en las Figs. 3a-3c se exponen las posiciones angulares del robot manipulador. El comportamiento de la señal de control correspondiente a cada grado de libertad del robot manipulador se encuentra en las Figs. 4a-4c. La función  $n(\bar{e})$  se ilustra en la Fig. 5, donde “1” indica el instante en el cual la función de evento se activa y en “0” cuando no se activa. De acuerdo con estos resultados, se describen a continuación las siguientes observaciones:

- El sistema no presenta sobreimpulsos, esto con base en el índice O y de las Figs. 3a-3c.
- De las Figs. 4a-4c se observa una mayor cantidad de variaciones cuando el sistema de control estabiliza al robot manipulador en los PNE, y previo a la estabilización en los PE. Este comportamiento se le atribuye a los efectos de la fuerza de gravedad, que afecta a los eslabones del robot manipulador ( $g \neq 0$ ) principalmente cuando el sistema busca posicionarse en los PNE. En contraste con los PE, la gravedad solo afecta al robot manipulador previo a alcanzar el estado estacionario, ya que una vez que el robot manipulador alcanza la referencia, la gravedad deja de afectar a los eslabones del sistema ( $g = 0$ ).
- La función de evento  $e(x, \bar{m})$  se activa (Fig. 5), visualmente, la mayor parte del tiempo de simulación, sin embargo el índice ETP indica que se reduce en un 15.3942% el número de actualizaciones de la señal de control.

- El índice ST sugiere que el tiempo de estabilización no excede 1s, y además el SSE indica que el error no sobrepasa el 5% de error una vez que el sistema alcanza el estado estacionario.

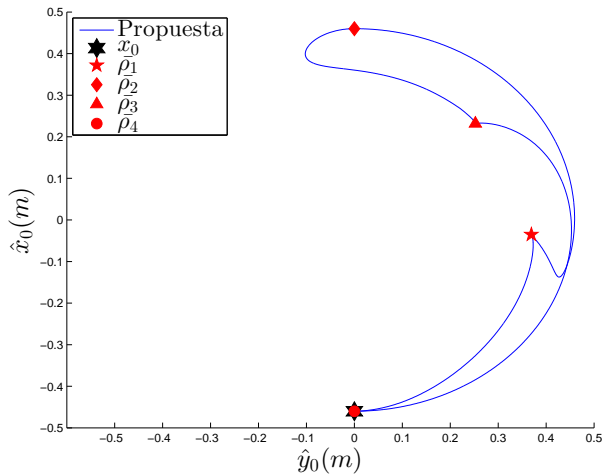


Figura 2. Comportamiento del efector final del robot manipulador en el espacio cartesiano  $\hat{x}_0 - \hat{y}_0$ .

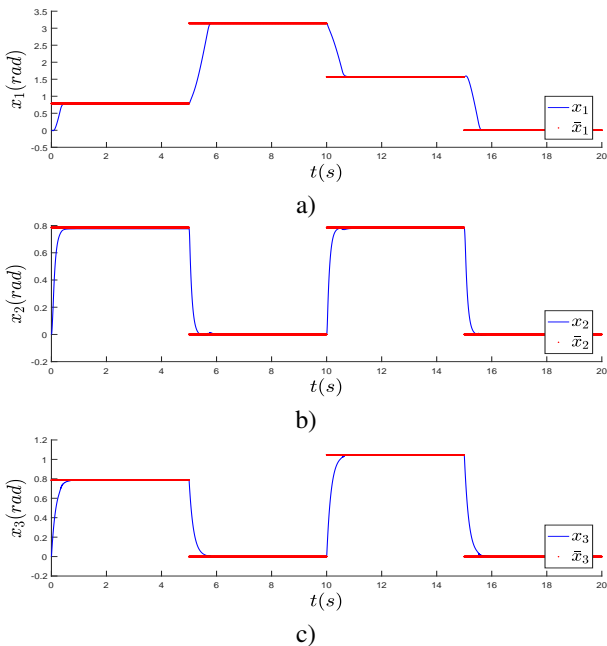


Figura 3. Comportamiento de cada grado de libertad del robot manipulador.

## V. CONCLUSIÓN

En el presente trabajo se desarrolló la sintonización fuera de línea de un CDE aplicado a un robot manipulador de tres grados de libertad, mediante un enfoque de optimización. El problema de optimización planteado contempla la minimización del error en estado estacionario para la tarea

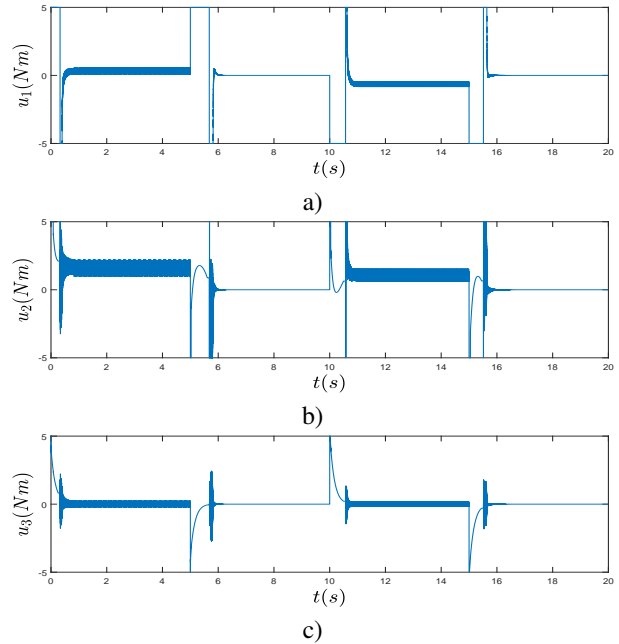


Figura 4. Comportamiento de la señal de control aplicada

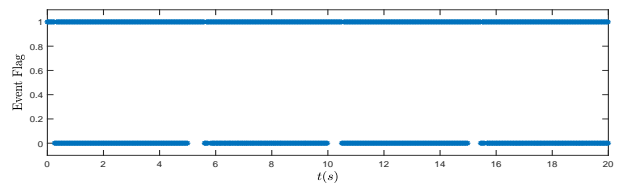


Figura 5. Función de activación de evento  $\eta(\bar{e})$ .

de regulación y el número de actualizaciones de la señal de control. La solución obtenida mediante el algoritmo de ED *Rand/1/Bin* se empleó para realizar pruebas en simulación, concluyéndose lo siguiente: 1) la sintonización del CDE con base en un enfoque de optimización indica que es una opción viable para obtener los parámetros del CDE, de manera que se presente un rendimiento aceptable en la tarea regulación ( $SSE = 0.0296rad(1.6959^\circ)$ ) y un decremento en el número de llamadas a la función de control (15.3952%), esto comparado con el número de actualizaciones requerido por un CDT; 2) la señal de control obtenida en simulación indica que existe una cantidad de variaciones mayor durante la estabilización de los PNE, ya que en estas posiciones los eslabones se ven afectados por la gravedad, por lo tanto, se puede comentar que puede existir alguna configuración de los parámetros del control, donde el sistema se encuentre en PE y no se requiera llamar a la función de control.

## AGRADECIMIENTOS

Los autores agradecen al Instituto Politécnico Nacional (IPN) y a la Secretaria de Investigación y Posgrado del IPN por el apoyo económico otorgado mediante el proyecto 20190239. El primer autor agradece al Consejo Nacional de Ciencia y

Tabla IV  
 VECTOR DE DISEÑO ÓPTIMO OBTENIDO

Parámetro	$\bar{p}_1$	$\bar{p}_2$	$\bar{p}_3$	$\bar{p}_4$	$\bar{p}_5$	$\bar{p}_6$	$\bar{p}_7$	$\bar{p}_8$	$\bar{p}_9$	$\bar{p}_{10}$	$\bar{p}_{11}$	$\bar{p}_{12}$	$\bar{p}_{13}$	$\bar{p}_{14}$	$\bar{p}_{15}$	$\bar{p}_{16}$	$\bar{p}_{17}$	$\bar{J}$	$\bar{J}_1$	$\bar{J}_2$
$\bar{p}_{min}$	57171.54	14907.13	7976.42	24.51	0.19	0.03	3.88	3.98	5.00	0.29	0.42	2.93	0.97	1.45	7.85	4.54	0.08	755.16	0.68	6784

 Tabla V  
 ÍNDICES DE RENDIMIENTO

O (5 %)	ST (s)(5 %)	SSE (rad)	$\ u\ $ (Nm)	ETP (%)
0.0000	0.4425	0.0296	1012.3132	84.6058

Tecnología (CONACYT), por haberle otorgado una beca para realizar sus estudios de maestría.

#### APÉNDICE

##### A. Detalles de la dinámica del robot manipulador

Las funciones  $f(x)$  y  $g(x)$  presentes en el modelo dinámico del robot manipulador expresado en (1), se presentan a continuación:

Los elementos de la matriz de inercias  $M \in \mathcal{R}^{3 \times 3}$  están dados por:

$$\begin{aligned}
 M_{11} &= I_{z_1} + I_{z_2} + I_{z_3} + l_1^2 m_2 + l_1^2 m_3 + l_2^2 m_3 + l_{c_1}^2 m_1 \\
 &\quad + l_{c_2}^2 m_2 + l_{c_3}^2 m_3 + 2l_1 l_2 m_3 \bar{\beta}_1 + 2l_1 l_{c_2} m_2 \bar{\beta}_1 \\
 &\quad + 2l_2 l_{c_3} m_3 \bar{\beta}_2 + 2l_1 l_{c_3} m_3 \bar{\beta}_3 \\
 M_{11} &= I_{z_1} + I_{z_2} + I_{z_3} + l_1^2 m_2 + l_1^2 m_3 + l_2^2 m_3 + l_{c_1}^2 m_1 \\
 &\quad + l_{c_2}^2 m_2 + l_{c_3}^2 m_3 + 2l_1 l_2 m_3 \bar{\beta}_1 + 2l_1 l_{c_2} m_2 \bar{\beta}_1 \\
 &\quad + 2l_2 l_{c_3} m_3 \bar{\beta}_2 + 2l_1 l_{c_3} m_3 \bar{\beta}_3 \\
 M_{12} &= I_{z_2} + I_{z_3} + l_2^2 m_3 + 2l_2 l_{c_3} m_3 \bar{\beta}_2 + l_1 l_2 m_3 \bar{\beta}_1 + l_{c_2}^2 m_2 \\
 &\quad + l_{c_3}^2 m_3 + l_1 l_{c_2} m_2 \bar{\beta}_1 + l_1 l_{c_3} m_3 \bar{\beta}_3 \\
 M_{13} &= I_{z_3} + l_{c_3}^2 m_3 + l_1 l_{c_3} m_3 \bar{\beta}_3 + l_2 l_{c_3} m_3 \bar{\beta}_2 \\
 M_{21} &= I_{z_2} + I_{z_3} + l_2^2 m_3 + 2l_2 l_{c_3} m_3 \bar{\beta}_2 + l_1 l_2 m_3 \bar{\beta}_1 \\
 &\quad + l_{c_2}^2 m_2 + l_{c_3}^2 m_3 + l_1 l_{c_2} m_2 \bar{\beta}_1 + l_1 l_{c_3} m_3 \bar{\beta}_3 \\
 M_{22} &= I_{z_2} + I_{z_3} + l_2^2 m_3 + 2l_2 l_{c_3} m_3 \bar{\beta}_2 + l_{c_2}^2 m_2 + l_{c_3}^2 m_3 \\
 M_{23} &= I_{z_3} + l_{c_3}^2 m_3 + l_2 l_{c_3} m_3 \bar{\beta}_2 \\
 M_{31} &= I_{z_3} + l_{c_3}^2 m_3 + l_1 l_{c_3} m_3 \bar{\beta}_3 + l_2 l_{c_3} m_3 \bar{\beta}_2 \\
 M_{32} &= I_{z_3} + l_{c_3}^2 m_3 + l_2 l_{c_3} m_3 \bar{\beta}_2 \\
 M_{33} &= I_{z_3} + l_{c_3}^2 m_3
 \end{aligned}$$

Los elementos de la matriz de fuerzas Centrígrafas y de Coriolis  $C \in \mathcal{R}^{3 \times 3}$  son:

$$\begin{aligned}
 C_{11} &= -x_5 l_1 l_2 m_3 \bar{\alpha}_2 - x_5 l_1 l_{c_2} m_2 \bar{\alpha}_2 - x_6 l_2 l_{c_3} m_3 \bar{\alpha}_3 \\
 &\quad - x_5 l_1 l_{c_3} m_3 \bar{\alpha}_5 - x_6 l_1 l_{c_3} m_3 \bar{\alpha}_5 \\
 C_{12} &= -(x_4 + x_5) l_1 l_2 m_3 \bar{\alpha}_2 - (x_4 + x_5) l_1 l_{c_2} m_2 \bar{\alpha}_2 \\
 &\quad - x_6 l_2 l_{c_3} m_3 \bar{\alpha}_3 - (x_4 + x_5 + x_6) l_1 l_{c_3} m_3 \bar{\alpha}_5 \\
 C_{13} &= -(x_4 + x_5 + x_6) (l_2 l_{c_3} m_3 \bar{\alpha}_3 + l_1 l_{c_3} m_3 \bar{\alpha}_5) \\
 C_{21} &= x_4 (l_1 l_{c_3} m_3 \bar{\alpha}_5 + l_1 l_2 m_3 \bar{\alpha}_2 + l_1 l_{c_2} m_2 \bar{\alpha}_2) \\
 &\quad - x_6 l_2 l_{c_3} m_3 \bar{\alpha}_3 \\
 C_{22} &= -x_6 l_2 l_{c_3} m_3 \bar{\alpha}_3 \\
 C_{23} &= -(x_4 + x_5 + x_6) l_2 l_{c_3} m_3 \bar{\alpha}_3 \\
 C_{31} &= (x_4 + x_5) l_2 l_{c_3} m_3 \bar{\alpha}_3 + x_4 l_1 l_{c_3} m_3 \bar{\alpha}_5 \\
 C_{32} &= (x_4 + x_5) l_2 l_{c_3} m_3 \bar{\alpha}_3 \\
 C_{33} &= 0
 \end{aligned}$$

Los elementos del vector de gravedad  $G \in \mathcal{R}^3$  son dados por:

$$G_{11} = gl_2 m_3 \bar{\alpha}_4 + gl_{c_2} m_2 \bar{\alpha}_4 + gl_1 m_2 \bar{\alpha}_1 + gl_1 m_3 \bar{\alpha}_1$$

$$+ gl_{c_1} m_1 \bar{\alpha}_1 + gl_{c_3} m_3 \bar{\alpha}_6$$

$$G_{21} = gl_2 m_3 \bar{\alpha}_4 + gl_{c_2} m_2 \bar{\alpha}_4 + gl_{c_3} m_3 \bar{\alpha}_6$$

$$G_{31} = gl_{c_3} m_3 \bar{\alpha}_6$$

Considerando:

$$\begin{aligned}
 \bar{\beta}_1 &= \cos(x_2), & \bar{\alpha}_1 &= \sin(x_1), & \bar{\alpha}_4 &= \sin(x_1 + x_2) \\
 \bar{\beta}_2 &= \cos(x_3), & \bar{\alpha}_2 &= \sin(x_2), & \bar{\alpha}_5 &= \sin(x_2 + x_3) \\
 \bar{\beta}_3 &= \cos(x_2 + x_3), & \bar{\alpha}_3 &= \sin(x_3), & \bar{\alpha}_6 &= \sin(x_1 + x_2 + x_3)
 \end{aligned}$$

#### REFERENCIAS

- [1] C. Canudas-De-Wit, B. Siciliano, and G. Bastin, *Theory of Robot Control*. London, England: Springer-Verlag, 1996.
- [2] E. D. Markus, J. T. Agee, and A. A. Jimohc, "Flat control of industrial robotic manipulators," *Robotics and Autonomous Systems*, vol. 87, pp. 226–236, 2017.
- [3] Y. Bouteraa, I. B. Abdallah, and J. Ghommam, "Task-space region-reaching control for medical robot manipulator," *Computers and Electrical Engineering*, p. 1–17, 2017.
- [4] I. M. D. Fonseca, L. C. S. Goes, N. Seito, M. K. da Silva Duarte, and E. J. de Oliveira, "Attitude dynamics and control of a spacecraft like a robotic manipulator when implementing on-orbit servicing," *Acta Astronautica*, vol. 137, pp. 490–497, 2017.
- [5] O. I. Borisov, V. S. Gromov, A. A. Pyrkin, A. A. Vedyakov, I. V. Petranovsky, A. A. Bobtsov, and V. I. Salikhov, "Manipulation tasks in robotics education," *IFAC-PapersOnLine*, pp. 22–27, 2016.
- [6] M. W. Spong, S. Hutchinson, and M. Vidyasagar, *Robot Dynamics and Control*. John Wiley and Sons Inc, 2005.
- [7] K. J. Aström and B. Wittenmark, *Computer-Controlled Systems: Theory and Design*. Beijing, China: Prentice Hall, 1997.
- [8] N. Marchand, S. Durand, and J. F. G. Castellanos, "A general formula for event-based stabilization of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 58, pp. 1332–1337, 2013.
- [9] M. G. Villarreal-Cervantes, J. F. Guerrero-Castellanos, S. Ramírez-Martínez, and J. P. Sánchez-Santana, "Stabilization of a (3,0) mobile robot by means of an event-triggered control," *ISA Transactions*, vol. 58, pp. 605–613, 2015.
- [10] N. S. Tripathy, I. N. Kar, and K. Paul, "An event-triggered based robust control of robot manipulator," *13th International Conference on Control, Automation, Robotics & Vision*, pp. 425–430, 2014.
- [11] S. Durand, J. F. Guerrero-Castellanos, N. Marchand, and W. F. Guerrero-Sánchez, "Event-based control of the inverted pendulum: Swing up and stabilization," *International Journal of Control, Automation and Systems*, pp. 1–10, 2013.
- [12] M. G. Villarreal-Cervantes, A. Rodríguez-Molina, C. García-Mendoza, O. Peñaloza-Mejía, and G. Sepúlveda-Cervantes, "Multi-objective on-line optimization approach for the dc motor controller tuning using differential evolution," *IEEE Access*, vol. 5, pp. 20 393–20 407, 2017.
- [13] O. Djaneye-Boundjou, X. Xu, and R. Ordóñez, "Automated particle swarm optimization based pid tuning for control of robotic arm," *Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS)*, vol. 39, p. 164–169, 2016.
- [14] H. V. H. Ayala and L. dos Santos Coelho, "Tuning of pid controller based on a multiobjective genetic algorithm applied to a robotic manipulator," *Expert Systems with Applications*, vol. 39, p. 8968–8974, 2012.
- [15] M. G. Villarreal-Cervantes and J. Alvarez-Gallegos, "Off-line PID control tuning for a planar parallel robot using de variants," *Robotics and Computer-Integrated Manufacturing*, vol. 64, pp. 444–454, 2016.
- [16] R. Storn and K. Price, "Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, pp. 341–359, 1997.
- [17] J. V. Carrau, G. Reynoso-Meza, S. García-Nieto, and X. Blasco, "Enhancing controller's tuning reliability with multi-objective optimisation: From model in the loop to hardware in the loop," *Engineering Applications of Artificial Intelligence*, vol. 64, p. 52–66, 2017.



# Diseño de un Observador para un Sistema Hipercaótico utilizando Programación Genética

Elia N. Martínez, Ramon Ramirez-Villalobos, Leonardo Trujillo, Paul J. Campos, Luis N. Coria

Departamento de Ingeniería Eléctrica y Electrónica

Tecnológico Nacional de México - Instituto Tecnológico de Tijuana

Tijuana, Baja California, México

Email: nictemtz94@gmail.com, {ramon.ramirez, leonardo.trujillo, paul.campos, luis.coria}@tectijuana.edu.mx

**Abstract**—Los sistemas hipercaóticos, por sus características, tienen aplicaciones en múltiples áreas de la ciencia e ingeniería. El estudio y desarrollo de este tipo de sistemas contribuye a resolver problemas diversos, siendo una de las aplicaciones más reportada la sincronización para encriptamiento y desencriptamiento de información. Para resolver el problema de sincronización, en este documento se presenta el diseño de un observador para un sistema hipercaótico tipo Lorenz. La metodología propuesta consiste en obtener un modelo equivalente del sistema hipercaótico, a partir de la serie de tiempo de las variables del sistema hipercaótico mediante el uso de un motor de modelado basado en programación genética. Posteriormente, se diseña un observador Lipschitz para el sistema equivalente obtenido. Simulaciones numéricas validan la convergencia del observador diseñado para el sistema obtenido mediante programación genética hacia las trayectorias del sistema hipercaótico tipo Lorenz.

## I. INTRODUCCIÓN

Un sistema caótico es un sistema que tiene una alta sensibilidad a condiciones iniciales, además presenta un comportamiento aparentemente aleatorio. Debido a lo anterior, en los últimos años se ha incrementado el estudio de los sistemas caóticos debido a las distintas aplicaciones que tienen en diversas áreas de la ingeniería. En 1963, Edward N. Lorenz desarrolló un modelo matemático en el cual estudiaba el comportamiento climático. Durante su estudio, se dio cuenta de que cualquier alteración por más mínima que fuera en las condiciones iniciales afectaría drásticamente las condiciones finales [1]. Posteriormente en 1979, Otto Rössler propuso el primer sistema hipercaótico. Dicho sistema, está conformado por un sistema de ecuaciones diferenciales de cuatro dimensiones [2].

Los sistemas hipercaóticos han sido intensivamente estudiados en la literatura. Comparados con la dinámica de los sistemas caóticos, los sistemas hipercaóticos tienen una dinámica más compleja. Debido a lo anterior, los sistemas hipercaóticos tienen potenciales aplicaciones en distintas ramas de la ciencia e ingeniería, por ejemplo: robótica móvil [3], sistemas de comunicaciones seguras [4], encriptamiento en sistemas biométricos [5], entre otras. En este contexto, en los últimos años, un reto interesante para la comunidad científica es la sincronización de sistemas hipercaóticos. Este problema puede ser resuelto mediante el diseño de observadores. Sin embargo, el diseño de observadores suele conllevar un análisis matemático complejo, y en ocasiones no se pueden diseñar

debido a la complejidad de los sistemas o por que no cumplen con las condiciones específicas requeridas por la metodología empleada.

Por otro lado, la programación genética es una metodología basada en los algoritmos evolutivos, la cual está inspirada en la evolución biológica para desarrollar automáticamente programas de cómputo. Es una técnica de aprendizaje automático utilizada para optimizar una población de acuerdo a una función de aptitud que evalúa la capacidad de cada programa para llevar a cabo una tarea específica [6]. De hecho, la programación genética puede ser entendida como una hiperheurística con un enfoque algorítmico para la síntesis automática de métodos heurísticos. En este contexto, la programación genética es un área de investigación que sido aplicada para resolver diversos problemas en diferentes áreas del conocimiento, entre ellas, generación de modelos dinámicos [7].

Basado en lo anterior, en años recientes se ha reportado estudios enfocados en la obtención de sistemas caóticos mediante el uso de programación genética. En [8], se intenta reproducir la dinámica del circuito de Chua a través de modelos autorregresivos no lineales con diferentes rasgos y ordenes. En [9] se utiliza programación genética en conjunto con diagramas de bifurcación para generar nuevos modelos de sistemas caóticos discretos. En [10], se utiliza programación genética multi-gen para evolucionar expresiones múltiples de las variables de estado del sistema de Lorenz. Posteriormente, en [11] reportan un sistema hipercaótico con órbitas heteroclinicas infinitas, mediante la misma metodología presentada en [10]. Finalmente, como continuación de la investigación realizada en [10], los mismos autores reportan en [12] otro conjunto de sistemas dinámicos cuyo comportamiento es similar al sistema de Lorenz.

La programación genética al ser un proceso automatizado para el análisis de datos ha ayudado a realizar de una manera más eficiente la resolución de problemas que conllevan un gran análisis matemático, donde con un acercamiento tradicional sería complicado resolverlo. Por ello, si se hace uso de la programación genética en el modelado de sistemas hipercaóticos se podría obtener un sistema equivalente cuya estructura cumpla con las condiciones de diseño y/o simplifique la complejidad. El objetivo del presente artículo es diseñar un observador para un sistema caótico, considerando un modelo equivalente mediante programación genética. La

metodología consiste en obtener un modelo equivalente de un sistema hipercaótico tipo Lorenz, mediante un motor de modelado basado en programación genética. Dicho modelo se obtiene a partir de una serie de tiempo de las variables del sistema hipercaótico, utilizando el programa Eureka de Nutonian [13]. Posteriormente, al modelo equivalente obtenido se le diseña un observador para sistemas tipo Lipschitz, al cual se le inyectan las salidas medibles del sistema hipercaótico. Cabe destacar que la metodología propuesta permite obtener el modelo equivalente de cualquier sistema hipercaótico, sin importar su complejidad.

Este trabajo de investigación se organiza como sigue. En la siguiente sección se presenta el modelo matemático del sistema hipercaótico tipo Lorenz considerado. En la tercera sección se muestra la metodología utilizada para obtener el modelo equivalente, mediante programación genética, así como, el diseño del observador. Posteriormente, se presentan los resultados obtenidos de la metodología propuesta. Finalmente, las conclusiones y trabajo futuro de esta investigación.

## II. SISTEMA HIPERCAÓTICO TIPO LORENZ

En 1963, Edward N. Lorenz desarrolló un modelo matemático para el movimiento de las masas de aire en la atmósfera. Como Lorenz estudiaba los patrones del tiempo, comenzó a darse cuenta que pequeñas variaciones en la condiciones iniciales de su modelo daban como resultado una diferencia en los patrones de comportamiento. Dicho modelo está descrito por la siguientes ecuaciones

$$\dot{x}_1 = \sigma(x_2 - x_1), \quad (1)$$

$$\dot{x}_2 = rx_1 - x_2 - x_1x_3, \quad (2)$$

$$\dot{x}_3 = -bx_3 + x_1x_2, \quad (3)$$

donde  $\sigma$ ,  $b$  y  $r$  son parámetros reales positivos.

Basados en el sistema de Lorenz, en la literatura se han reportado varios sistemas hipercaóticos. Un ejemplo de dichos sistemas hipercaóticos está definido por las siguientes ecuaciones [14]:

$$\dot{x}_1 = a(x_2 - x_1), \quad (4)$$

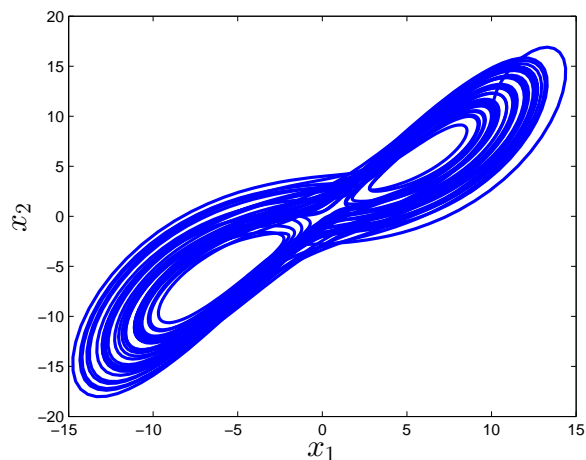
$$\dot{x}_2 = bx_1 - cx_2 + x_4 - x_1x_3, \quad (5)$$

$$\dot{x}_3 = -dx_3 + x_1x_2, \quad (6)$$

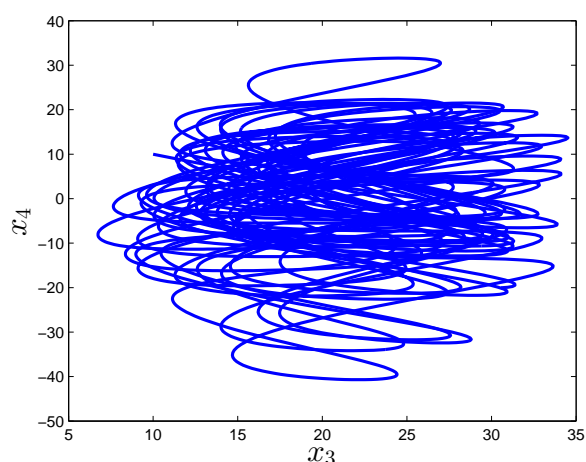
$$\dot{x}_4 = -kx_2 - rx_4, \quad (7)$$

donde  $a$ ,  $b$ ,  $c$ ,  $d$ ,  $k$  y  $r$  son parámetros reales positivos. El sistema (4)–(7) presenta un comportamiento hipercaótico para los valores de  $a = 12$ ,  $b = 23$ ,  $c = 1$ ,  $d = 2.1$ ,  $k = 6$  y  $r = 0.2$ .

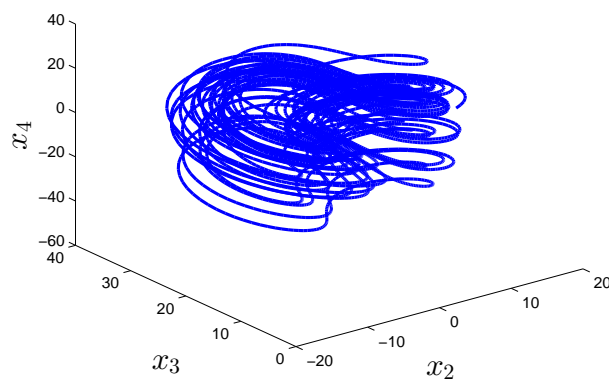
En la Fig. 1 se presentan los planos de fase de la dinámica hipercaótica del sistema (4)–(7). En la Figura 1(a) muestra la proyección  $(x_1, x_2)$ , la Figura 1(b) se muestra la proyección  $(x_3, x_4)$  y la Fig. 1(c) muestra la proyección  $(x_2, x_3, x_4)$  del atractor caótico.



(a) Plano de fase  $(x_1, x_2)$ .



(b) Plano de fase  $(x_3, x_4)$ .



(c) Atractor caótico  $(x_2, x_3, x_4)$ .

Fig. 1: Dinámica del sistema hipercaótico tipo Lorenz.

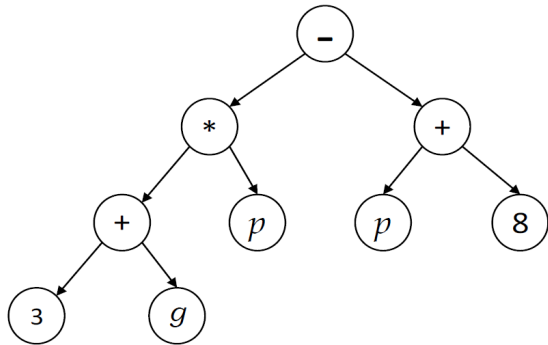


Fig. 2: Función representada mediante diagrama de árbol.

### III. METODOLOGÍA

#### A. Programación Genética

La programación genética (GP, por sus siglas en inglés) fue propuesta por Koza en 1992 [6], la cual está inspirada en selección natural y la genética. GP tiene como objetivo resolver problemas de manera automática partiendo de información *a-priori* del comportamiento deseado. En su forma más común, GP se puede entender como como algoritmo de aprendizaje supervisado que intenta construir una expresión válida sintácticamente usando un conjunto finito de funciones básicas y variables de entrada, guiada por un dominio dependiente de la función objetivo. Además, GP es una área de investigación que ha resuelto problemas diversos, incluyendo la generación de modelos dinámicos [7], [15]. En general, GP se caracteriza por dos rasgos principales, a diferencia de otras técnicas de computo evolutivo. En primer lugar, las soluciones evolucionadas representan expresiones sintácticas o programas válidos, que podrían ser utilizados como modelos. En segundo, la capacidad de GP para construir expresiones sintácticas directamente, sin asumir un modelo previamente, pueden permitir que preparar soluciones interpretables que sean capaces de resolver el problema, así como, proporcionar una visión del dominio del problema [15].

Comúnmente, las soluciones o modelos en GP se representan mediante diagramas de árbol. En la Fig. 2 se representa el modelo  $((3 + g) * p) - (8 + p)$ . Se puede observar que cada nodo del árbol tiene una función operador y cada hoja terminal tiene una variable o un valor constante; convirtiendo el árbol en un modelo matemático que puede ser resumido en una expresión matemática integral [16].

Dentro de este contexto, se ha comprobado con métodos numéricos que existen dificultades con respecto a la precisión y la convergencia de una solución. En tales circunstancias, se espera que GP sea una buena alternativa, dado que no solo encuentra una solución, sino que proporciona soluciones adicionales cuya estructura y calidad puede ser evaluada por el diseñador. Por lo tanto, dependiendo de la precisión y la simplicidad de la implementación, la elección de una solución depende de las especificaciones del usuario o los requisitos del problema [13].

En este contexto, GP puede ser de interés en el diseño de observadores. Puesto que, con GP se podría determinar un conjunto de sistemas dinámicos con comportamiento caótico compuesto por términos polinomiales, cruzados o de orden superior; que permitan satisfacer las condiciones de diseño de los observadores.

#### B. Modelo obtenido mediante Programación Genética

Para determinar un modelo equivalente, del sistema (4)–(7), se aplicó GP utilizando el Software Eureqa [13]. Primeramente, se recolectaron datos que emulaban la dinámica del sistema hipercaótico. Los datos fueron obtenidos mediante simulaciones numéricas en MATLAB/Simulink®, cuya cantidad corresponde a 40 segundos con un tiempo de muestreo de 0.001 segundos. Dichos datos fueron introducidos en una matriz de dimensiones  $m \times n$ , con  $m = 40000$  (cantidad de datos) y  $n = 4$  (dimensión del sistema hipercaótico). Posteriormente, una vez indicando el tipo de funciones y operaciones a considerar, los datos son cargados a Eureqa para generar los modelos se que correlacionan con el sistema (4)–(7). Es importante remarcar, que no es necesaria información *a-priori* del sistema. Además, como ya se mencionó, el algoritmo no solo genera una solución, si no que genera un conjunto de soluciones que son clasificadas en términos de su aptitud (error) y complejidad (tamaño dado por número de nodos). Finalmente, del conjunto de soluciones se selecciona el modelo que tenga mejor aptitud en cuanto a las consideraciones de diseño del observador.

#### C. Diseño del observador

Para el diseño del observador, considere un sistema dinámico de la forma

$$\dot{x} = Ax + Bu + f(x, u), \quad (8)$$

$$y = Cx, \quad (9)$$

donde  $x \in \mathbb{R}^n$  es el vector de estados,  $u \in \mathbb{R}^m$  es el vector de entradas de control y  $y \in \mathbb{R}^p$  es el vector de salida medible. Las matrices  $A$ ,  $B$ ,  $C$  son de dimensiones apropiadas.

Considere un observador de la forma

$$\dot{\hat{x}} = A\hat{x} + Bu + f(\hat{x}, u) + L(y - C\hat{x}), \quad (10)$$

donde  $\hat{x} \in \mathbb{R}^n$  es el vector de estados estimados. La matriz de ganancia del observador  $L$  es de dimensiones correspondientes.

El término no lineal  $f(x, u)$  en (8) es localmente Lipschitz en un dominio  $\mathcal{D}$  y uniformemente acotada en  $u \in \mathbb{R}^m$ , es decir, satisface

$$\|f(x, u) - f(\hat{x}, u)\| \leq \gamma \|x - \hat{x}\|, \quad x \in \mathcal{D}, \quad (11)$$

donde  $\gamma > 0$  es la constante de Lipschitz [17].

En el diseño del observador se hará uso del siguiente Teorema [18]

**Teorema 1.** Para la clase de sistemas y observador descritos de la forma en (8)–(9) y (10), respectivamente, si la matriz de ganancia del observador  $L$  es elegida tal que

$$\begin{bmatrix} (A - LC)^T P + P(A - LC) + \varepsilon \gamma^2 I_n & P \\ P & -\varepsilon I_n \end{bmatrix} < 0, \quad (12)$$

para alguna una matriz  $P$  simétrica y definida positiva, entonces esta elección de  $L$  conduce a estimaciones asintóticamente estables del observador (10) para el sistema (8).

#### IV. RESULTADOS

Primeramente, del conjunto de soluciones obtenidas mediante GP, se eligió el siguiente modelo

$$\dot{z}_1 = a_1 z_2 + b_1 z_4 - c_1 z_1, \quad (13)$$

$$\begin{aligned} \dot{z}_2 = & a_2 z_1 + b_2 z_4 + c_2 z_2 z_3 + d_2 z_1 z_3^2 \\ & - e_2 z_2 - g_2 z_1 z_3, \end{aligned} \quad (14)$$

$$\dot{z}_3 = a_3 z_1 z_2 - b_3 z_3 - c_3 z_3 z_2^2, \quad (15)$$

$$\dot{z}_4 = -a_4 z_1 - b_4 z_4 - c_4 z_2, \quad (16)$$

con los valores de parámetros que se muestran en la Tabla I.

TABLE I: Parametros del sistema

Parámetro	Valor
$a_1$	11.9973342817386
$b_1$	$3.7260882929797 \times 10^{-4}$
$c_1$	11.9965497255572
$a_2$	23.0343278659265
$b_2$	0.999994206103091
$c_2$	$0.8627809323498 \times 10^{-4}$
$d_2$	$6.87645909022622 \times 10^{-5}$
$e_2$	1.01387135386776
$g_2$	1.00333116916329
$a_3$	1.0012381644601
$b_3$	2.09982710879855
$c_3$	$5.81132216290191 \times 10^{-5}$
$a_4$	$1.71865610849075 \times 10^{-3}$
$b_4$	0.200121863576808
$c_4$	5.9984964056936

Posteriormente, el sistema (13)–(16) se reescribió de la forma (8) con  $z = [z_1, z_2, z_3, z_4]^T$  y

$$A = \begin{bmatrix} -c_1 & a_1 & 0 & b_1 \\ a_2 & -e_2 & 0 & b_2 \\ 0 & 0 & -b_3 & 0 \\ -a_4 & -c_2 & 0 & -b_4 \end{bmatrix}, \quad (17)$$

$$f(z) = \begin{bmatrix} 0 \\ c_2 z_2 z_3 + d_2 z_1 z_3^2 - f_2 z_1 z_3 \\ a_3 z_1 z_2 - c_3 z_3 z_2^2 \\ 0 \end{bmatrix}, \quad (18)$$

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (19)$$

En este caso la entrada de control es nula, es decir,  $u = 0$ . Considerando, la matriz de ganancia del observador  $L$  como

$$L = \begin{bmatrix} 0 & 0 \\ 250 & 0 \\ 0 & 100 \\ 0 & 0 \end{bmatrix}, \quad (20)$$

y los valores de parámetros  $\gamma = 1000$  y  $\varepsilon = \gamma^{-2}$ , existe una matriz  $P$  simétrica y definida positiva dada por

$$P = \begin{bmatrix} 0.3317 & 0 & 0 & 0 \\ 0 & 0.3316 & 0 & 0 \\ 0 & 0 & 0.3317 & 0 \\ 0 & 0 & 0 & 0.3317 \end{bmatrix} \times 10^{-5}, \quad (21)$$

entonces la desigualdad (12) del Teorema 1 se cumple.

Los resultados obtenidos se presentan a continuación. Las condiciones iniciales consideradas para el sistema (4)–(7) y del observador fueron elegidas como

$$x(0) = [x_1(0), x_2(0), x_3(0), x_4(0)]^T = 10 \in \mathbb{R}^4, \quad (22)$$

$$\hat{z}(0) = [\hat{z}_1(0), \hat{z}_2(0), \hat{z}_3(0), \hat{z}_4(0)]^T = 7 \in \mathbb{R}^4, \quad (23)$$

respectivamente. Las Fig. 3(a) y 3(b) muestran como converge el observador (línea punteada de color rojo) hacia los planos de fase  $(x_1, x_2)$  y  $(x_3, x_4)$  (línea solida de color azul), respectivamente, del sistema (4)–(7). Igualmente, en la Fig. 3(c) se muestra que las las trayectorias del observador (10) convergen al atractor caótico del sistema (4)–(7).

En la Fig. 4 se muestra la dinámica del error. Se puede observar que el error de estimación converge a cero de manera asintótica.

#### V. CONCLUSIONES

En este trabajo se diseñó un observador para un sistema hipercaótico, considerando un modelo equivalente obtenido mediante programación genética. Lo anterior se logró mediante el uso del motor de modelado Eureka de Nutonian. El conjunto de modelos se obtuvo a partir de la serie de tiempo de las variables de un sistema hipercaótico tipo Lorenz. A partir del conjunto de soluciones obtenidas, se eligió el conjunto de ecuaciones con mejor aptitud. Y basado en ésta selección, se diseñó un observador tipo Lipschitz para el sistema obtenido mediante GP. Al observador diseñado se le inyectan la salida medible del sistema hipercaótico para generar el error de estimación.

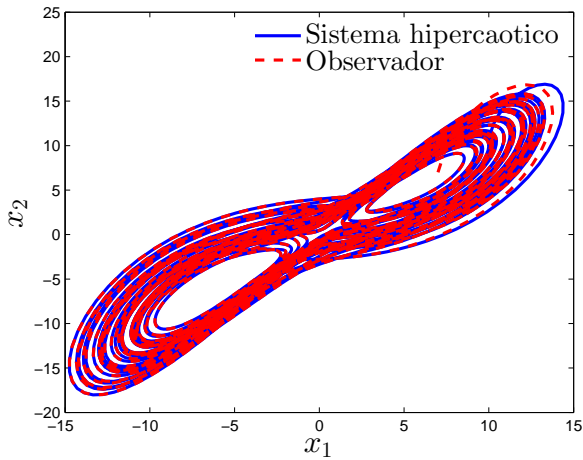
Los resultados obtenidos son validados mediante simulaciones numéricas. Investigaciones futuras considerarán sistemas hipercaóticos con terminos racionales y/o discontinuos.

#### AGRADECIMIENTOS

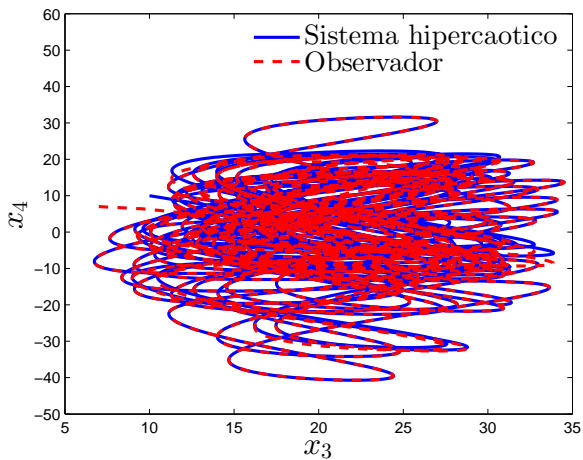
Este trabajo fue soportado por los proyectos 6006.19-P del Tecnológico Nacional de México, titulado “Observadores robustos para la solución de problemas en ingeniería” y el proyecto 6826.18-P del Tecnológico Nacional de México, titulado “Generación Automática de Sistemas Caóticos mediante Programación Genética”.

#### REFERENCES

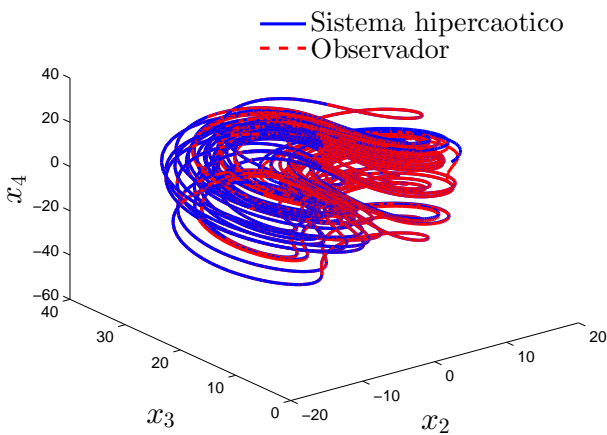
- [1] E. N. Lorenz. “Deterministic nonperiodic flow”. *Journal of the Atmosphere Sciences*, vol. 20(2), pp. 130–141, 1963.
- [2] O. E. RöSSLer. “An equation for hyperchaos”, *Physics Letters A*, vol. 71, pp. 155–157, 1979.
- [3] S. Vaidyanathan and A. Sambas and M. Mamat an W. M. Sanjaya, “A new three-dimensional chaotic system with a hidden attractor, circuit design and application in wireless mobile robot” *Archives of Control Sciences*, vol. 27(4), pp. 541–554, 2017.



(a) Plano de fase  $(x_1, x_2)$ .



(b) Plano de fase  $(x_3, x_4)$ .



(c) Atractor caótico  $(x_2, x_3, x_4)$ .

Fig. 3: Dinámica del sistema hipercaótico tipo Lorenz.

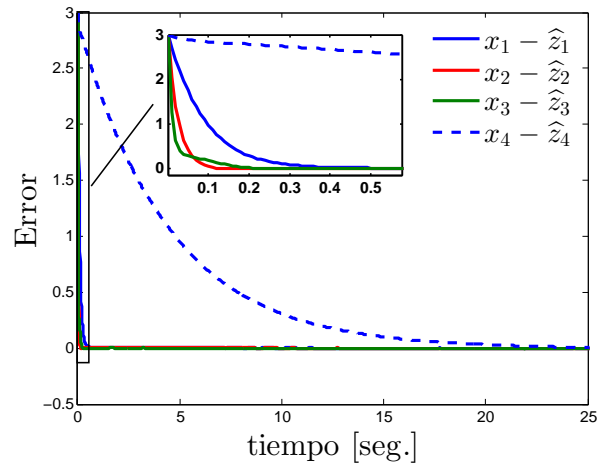


Fig. 4: Dinámica del error de estimación.

Based Realization of a Video Chaotic Secure Communication Scheme”, *International Journal of Bifurcation and Chaos*, vol. 28(13), pp. 1850160, 2018.

- [5] F. Abundiz-Pérez and C. Cruz-Hernández and M. A. Murillo-Escobar and R. M. López-Gutiérrez and A. Arellano-Delgado. “A Fingerprint Image Encryption Scheme Based on Hyperchaotic Rssler Map”, *Mathematical Problems in Engineering*, vol. 2016, Article ID 2670494, 15 pages, 2016.
- [6] J. R. Koza. *Genetic programming: on the programming of computers by means of natural selection*, MIT Press, 1992.
- [7] R. Poli and M. Graff and N. F. McPhee. “Free lunches for function and program induction”, *Proceeding of the tenth ACM SIGEVO forkshop of Foundations of genetic algorithms (FOGA)*, pp. 183–194, 2009.
- [8] K. Rodriguez and P. J. Flemming. “Evolution of mathematical models of chaotic system based on multiobjective genetic programming”, *Knowledge and Information Systems*, vol. 8(2), pp. 235–256, 2005.
- [9] I. Zelinka and S. Celikovskyy and H. Richter and G. Chen. *Evolutionary algorithms and chaotic systems*, Springer, 2010.
- [10] I. Pan and S. Das. “When Darwin meets Lorenz: Evolving new chaotic attractors through genetic programming”, *Chaos, Solitons and Fractals*, vol 76, pp. 141–155, 2015.
- [11] H. Wang and X. Li. “A novel hyperchaotic system with infinitely many heteroclinic orbits coined”, *Chaos, Solitons and Fractals*, vol 106, pp. 5–15, 2018.
- [12] I. Pan and S. Das. “Evolving chaos: Identifying new attractors of the generalized Lorenz family”, *Applied Mathematical Modelling*, vol 57, pp. 391–405, 2018.
- [13] M. Schmidt and H. Lipson, Eureka (Version 0.98 beta). Available from [www.nutonian.com](http://www.nutonian.com). (2014)
- [14] Y. Li and X. Liu and G. Chen and X. Liao. “A new hyperchaotic Lorenz type system: Generation, analysis, and implementation”, *International Journal of Circuit Theory and applications*, vol. 39(8), pp. 865–879, 2011.
- [15] K. Oluwaseun and J. A. A. Oyeboade. “Genetic programming: Principles, applications and opportunities for hydrological modelling”, *International Journal of Environmental and Ecological Engineering*, vol 8(6), pp. 330–336, 2014.
- [16] I. G. Tsoulos and I. E. Lagaris, “Solving Differential Equations with Genetic Programming”, *Genetic Programming and Evolvable Machines*, vol. 1, pp. 33-54, 2006.
- [17] H. J. Marquez. *Nonlinear Control Systems: Analysis and Design*, John Wiley & sons, 2003.
- [18] G. Panomchoeng and R. Rajamani. “Observer Design for Lipschitz Nonlinear Systems Using Riccati Equations”, *Proceeding of the 2010 American Control Conference (ACC)*, pp. 6060–6065, 2010.

[4] P. Chen and S. Yu and B. Chen and L. Xiao and J. Lü. “Design and SOPC-

# Implementación de un sensorless control a un motor de inducción por medio de un observador de Luenberger adaptativo

Aldrete Christian

Tecnológico Nacional de México

Intituto Tecnológico de Tijuana

Email: christian.aldrete@tectijuana.edu.mx

Coria Luis N.

Tecnológico Nacional de México

Intituto Tecnológico de Tijuana

Email: luis.coria@tectijuana.edu.mx

Ramirez-Villalobos Ramon

Tecnológico Nacional de México

Intituto Tecnológico de Tijuana

Email: ramon.ramirez@tectijuana.edu.mx

**Abstract**—En este trabajo es propuesto un sensorless control para un motor de inducción con un observador de Luenberger para estimar el flujo y velocidad del rotor mediante, las mediciones directas de corrientes en los devanados de estator del motor de inducción sin necesidad de ningún tipo de sensor que mida la posición del rotor. Por las características de ciertas aplicaciones, es común que el uso de un sensor dificulte la implementación física, con costos adicionales de mantenimiento, así como desgaste y problemas de ruido inherentes a los sensores. Para la validación de los resultados teóricos se emuló e implementó el observador propuesto en el esquema de Control de Campo Orientado integrado en la plataforma Technosoft DMCode-MS(IM).

## I. INTRODUCCIÓN

El motor de inducción es de las máquinas más utilizadas en la industria por su simplicidad, eficiencia, confiabilidad y bajo costo. El Control de Campo Orientado (FOC por sus siglas en inglés), es una técnica utilizada en aplicaciones con motores síncronos y asíncronos que exigen un control más robusto. Esta técnica desacopla la componente de corriente que causa el flujo en el rotor y la componente de corriente que causa el torque generado por el motor, esto se hace con el fin de tener mayor control de torque, posición y velocidad en los motores, que a comparación de otras técnicas típicas utilizadas en la industria, como el control por Volts/Hertz (V/Hz), hacen que a bajas velocidades el motor no sea capaz de generar el torque suficiente para operar el motor.

En el FOC se derivan dos métodos recurrentes, el FOC Indirecto y el FOC Directo [1], [2]. En el caso indirecto, es llamado así por la siguiente razón; un motor de inducción es categorizado como máquina asíncrona debido a que la velocidad del flujo y la velocidad del rotor nunca son iguales, la velocidad relativa del flujo con respecto a la del rotor es conocida como velocidad de deslizamiento, por ello se recurre al uso de un sensor, por ejemplo un encoder, es necesario conocer la posición del rotor para calcular la velocidad del flujo en el rotor y obtener las componentes de corriente que causan el flujo y el torque de la máquina. En el método directo el flujo del rotor es medido directamente, por ejemplo, se usan sensores de Hall equipados en un diseño especial para el motor, aquí no hay necesidad de medir la posición del rotor.

No obstante, ambos metodos usan sensores para el control, que en el caso de aplicaciones donde no se requieren de velocidades tan precisas pueden causar varios inconvenientes. El uso de sensores es invadir el diseño interno del motor resultando en un circuito más complejo, exponiéndose a ruidos y probables lecturas erróneas cuando se trabaja en ambientes hostiles, incluso algunas veces no es posible agregar un sensor al motor por las limitaciones de la aplicación. Para dar solución a este problema se proponen varias técnicas de control en la literatura para remover el sensor, comúnmente conocidas como control sensorless, los cuales, en la mayoría de los casos se basan en el método de FOC Directo, en el que se estima el flujo del rotor desde las terminales de alimentación del motor [3]–[5]. Las ventajas de remover el sensor impacta en el ahorro de costos de mantenimiento, reducción de espacio, mayor confiabilidad y eficiencia, todo lo anterior a cambio de algoritmos más elaborados, lo cual es posible gracias a los avances en el área de la electrónica de potencia y los microcontroladores. Durante los últimos años, el procesamiento de datos es posible a bajos costos con la consecuente implementación de controles más eficientes y sofisticados. Sin embargo, remover el sensor no es una tarea sencilla, el modelo matemático que describe el comportamiento de un motor de inducción es no lineal, multivariable y altamente acoplado. En el presente trabajo se propone un observador de Luenberger en el que se estima el flujo del rotor partiendo del análisis de la dinámica del error, estimando la velocidad por medio de un mecanismo de adaptación que será utilizado para lograr la convergencia del observador. Para validar los resultados se emulará y implementará en la plataforma Technosoft®DMCode-MS(IM).

El artículo se organiza de la siguiente manera. En la próxima sección se describe el modelo matemático del motor de inducción. Posteriormente, se presenta el diseño del observador de Luenberger y su matriz de ganancias. Así mismo, se muestra el cálculo para estimar la velocidad. Enseguida, se presentan los resultados obtenidos del observador de Luenberger adaptativo en lazo cerrado. En la sección inmediata se presentan las conclusiones de la investigación.



## II. MODELO MATEMÁTICO DEL MOTOR DE INDUCCIÓN

Las ecuaciones diferenciales que describen el motor de inducción están descritas en un marco de referencia estacionario  $(d,q)$  [6], considerando una velocidad angular  $\omega_g$ ,

$$\frac{d}{dt} i_{ds} = -\mu i_{ds} + \omega_g i_{qs} + \beta(\alpha\varphi_{dr} + \omega_r \varphi_{qr}) + \frac{u_{ds}}{\sigma}, \quad (1)$$

$$\frac{d}{dt} i_{qs} = -\mu i_{qs} + \omega_g i_{ds} + \beta(\alpha\varphi_{qr} + \omega_r \varphi_{dr}) + \frac{u_{qs}}{\sigma}, \quad (2)$$

$$\frac{d}{dt} \varphi_{dr} = -\alpha\varphi_{dr} + (\omega_g - \omega_r) \varphi_{qr} + \alpha L_m i_{ds}, \quad (3)$$

$$\frac{d}{dt} \varphi_{qr} = -\alpha\varphi_{qr} + (\omega_g - \omega_r) \varphi_{dr} + \alpha L_m i_{qs}, \quad (4)$$

$$\frac{d}{dt} \omega_r = \gamma(\varphi_{dr} i_{qs} - \varphi_{qr} i_{ds}) - \frac{T_l}{J}, \quad (5)$$

donde  $i_{ds}$ ,  $i_{qs}$ ,  $\varphi_{dr}$ ,  $\varphi_{qr}$ , representan las corrientes del estator y flujos del rotor respectivamente. Los parámetros del sistema se definen como  $\mu = R_s/\sigma + \alpha\beta L_m$ ,  $\sigma = \frac{L_s L_r - L_m^2}{L_r}$ ,  $\beta = \frac{L_m}{\sigma L_r}$ ,  $\alpha = \frac{R_r}{L_r}$  y  $\gamma = \frac{pL_m}{JL_r}$ . Las constantes  $R_s$  y  $R_r$  son las resistencias de estator y rotor,  $L_s$  y  $L_r$  son las inductancias de estator y rotor,  $L_m$  es la inductancia mutua,  $J$  es la inercia del motor,  $p$  es el número de pares de polos,  $T_l$  es la carga del motor y  $\omega_r$  es la velocidad angular del rotor.

El modelo del motor se trabajará en el marco de referencia estacionario, donde se define  $\omega_g = 0$ , también se adoptará el parámetro  $\omega_r$  como una constante y  $T_l = 0$ , por lo que las ecuaciones no lineales del modelo (1)-(5) se reducen a un sistema lineal, de la forma  $\dot{x} = Ax + Bu$ ,  $y = Cx$ , con

$$x = (i_{ds} \ i_{qs} \ \varphi_{dr} \ \varphi_{qr})^T, \quad (6)$$

$$u = (u_{ds} \ u_{qs} \ 0 \ 0)^T, \quad (7)$$

$$A = \begin{pmatrix} -\mu & 0 & \alpha\beta & \beta\omega_r \\ 0 & -\mu & -\beta\omega_r & \alpha\beta \\ \alpha L_m & 0 & -\alpha & -\omega_r \\ 0 & \alpha L_m & \omega_r & -\alpha \end{pmatrix}, \quad (8)$$

$$B = \begin{pmatrix} 1/\sigma & 0 \\ 0 & 1/\sigma \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad (9)$$

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \quad (10)$$

En este esquema tomar  $\omega_r$  como una constante es una consideración de diseño. Sin embargo, esta será contemplada como una variable  $\hat{\omega}_r(t)$ , más adelante se estimará y retroalimentará en la matriz de estados para cerrar el lazo de control.

## III. OBSERVADOR DE LUENBERGER

El observador que estima las corrientes de estator y flujo del rotor está diseñado como

$$\dot{\hat{x}} = \hat{A}\hat{x} + Bu + L(\hat{y} - y), \quad (11)$$

$$\hat{y} = C\hat{x}, \quad (12)$$

donde  $\hat{\cdot}$  significa los valores estimados y  $L$  es la matriz de ganancias de retroalimentación del observador.

Para verificar que el sistema es observable se realiza la prueba de observabilidad con las matrices  $A$  y  $C$ , definidas en (8) y (10), como sigue,

$$Ob = (C \ CA \ CA^2 \ CA^3)^T. \quad (13)$$

Utilizando un método numérico se calcula que el rango de la matriz  $Ob$  es 4. Consecuentemente, todas sus columnas o filas de la matriz  $Ob$  son linealmente independientes y por lo tanto el sistema es observable.

Para el diseño,  $L$  es propuesto en función de  $\hat{\omega}_r$  para que el observador sea estable en un rango específico de velocidades del rotor, vea [7].

$$L = \begin{pmatrix} l_1 & l_2 & l_3 & l_4 \\ -l_2 & l_1 & -l_4 & l_3 \end{pmatrix}^T, \quad (14)$$

donde

$$l_1 = (k-1)(-\mu - \alpha), \quad (15)$$

$$l_2 = (k-1)\hat{\omega}_r, \quad (16)$$

$$l_3 = (k^2 - 1)(\alpha L_m - \frac{\mu}{\beta}) - \frac{1}{\beta} l_1, \quad (17)$$

$$l_4 = -\frac{1}{\beta} l_2, \quad (18)$$

donde  $k > 0$  es una constante proporcional.

## IV. ESTIMACIÓN DE LA VELOCIDAD

La velocidad del rotor será estimada dentro de un esquema adaptativo analizando la dinámica del error del observador, en la Fig. 1 se muestra el esquema general para la estimación de la velocidad.

Considerado  $e = x - \hat{x}$ , la dinámica del error de estimación de las corrientes de estator y flujo del rotor a partir de (6), (8), (11) y (12) es descrito por la siguiente ecuación

$$\dot{e} = (A + LC)e - \Delta A \hat{x}, \quad (19)$$

donde

$$\Delta A = \hat{A} - A = \begin{pmatrix} 0 & 0 & 0 & \Delta\omega_r\beta \\ 0 & 0 & -\Delta\omega_r\beta & 0 \\ 0 & 0 & 0 & -\Delta\omega_r\beta \\ 0 & 0 & \Delta\omega_r\beta & 0 \end{pmatrix}, \quad (20)$$

$$\Delta\omega_r = \hat{\omega}_r - \omega_r, \quad (21)$$

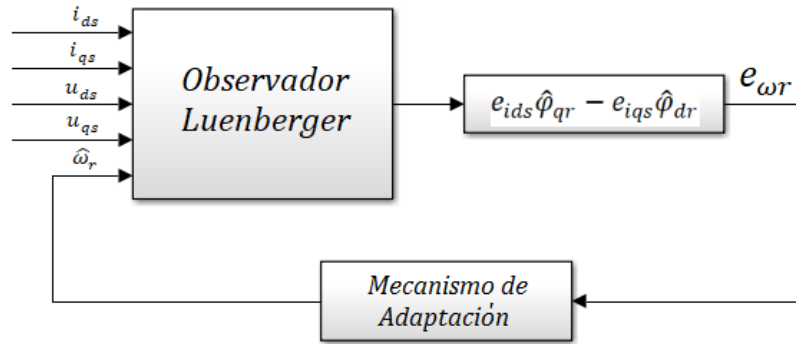


Fig. 1. Diagrama a bloques del observador de Luenberger adaptativo para estimar la velocidad del rotor.

donde  $\Delta A$  es la matriz de error causada por la variación de la velocidad  $\omega_r$ .

Para determinar la dinámica de error del observador se propone una función candidata a Lyapunov definida positiva

$$V = e^T e + \frac{(\hat{\omega}_r - \omega_r)^2}{\delta}, \quad (22)$$

donde  $\delta$  es una constante positiva. La derivada de (22) se resuelve como

$$\begin{aligned} \dot{V} = e^T ((A + LC)^T + (A + LC))e + 2 \frac{\Delta\omega_r}{\delta} \dot{\omega}_r \\ - 2\Delta\omega_r (e_{ids}\hat{\varphi}_{qr} - e_{iqs}\hat{\varphi}_{dr})\beta, \end{aligned} \quad (23)$$

donde  $e_{ids} = i_{ds} - \hat{i}_{ds}$  y  $e_{iqs} = i_{qs} - \hat{i}_{qs}$ . Con la matriz de ganancias  $L$  propuestas, la derivada  $\dot{V}$  satisface el criterio de estabilidad de Lyapunov, consulte [8]. Entonces con la finalidad de encontrar un esquema adaptativo para estimar la velocidad del rotor se procede a igualar el segundo y tercer término de (23) a cero,

$$-2\Delta\omega_r (e_{ids}\hat{\varphi}_{qr} - e_{iqs}\hat{\varphi}_{dr})\beta + 2 \frac{\Delta\omega_r}{\delta} \dot{\omega}_r = 0, \quad (24)$$

lo cual no modifica la estabilidad del sistema, de esta forma, (24) se desarrolla y se obtiene

$$\dot{\omega}_r = \int \delta (e_{ids}\hat{\varphi}_{qr} - e_{iqs}\hat{\varphi}_{dr})\beta, \quad (25)$$

por lo que la dinámica del error es

$$e_{\omega r} = e_{ids}\hat{\varphi}_{qr} - e_{iqs}\hat{\varphi}_{dr}. \quad (26)$$

Debido a que la velocidad del rotor puede cambiar rápidamente, dentro del mecanismo de adaptación de la Fig. 1, se propone un control PI con el fin de mejorar la respuesta de estimación de velocidad, dado por

$$\dot{\omega}_r = (K_p + \frac{K_i}{s})e_{\omega r}, \quad (27)$$

donde  $K_p$  y  $K_i$  son ganancias positivas arbitrarias y  $s$  es el operador de Laplace.

 TABLE I  
 ESPECIFICACIONES Y PARÁMETROS DEL MOTOR

Parámetro	Valor
Potencia nominal	0.37 kW
Voltaje nominal	400 V <sub>rms</sub>
Corriente nominal	0.95A
Velocidad nominal	3000 rpm
Número de polos	1
Inductancia del estator	0.4934 H
Inductancia del rotor	0.4934 H
Inductancia mutua	0.4867 H
Resistencia del estator	8.2 Ω
Resistencia del rotor	5.36667 Ω

## V. RESULTADOS

Para la validación de los resultados, el observador (11) y el controlador (27) se programan en el emulador de la plataforma Technosoft DMCode-MS(IM), posteriormente se realiza una implementación física. La programación es llevada a cabo mediante los procedimientos descritos en el manual de usuario de la plataforma [9]. Para el controlador del sistema se utilizan los controladores PI integrados en el software de la plataforma, los cuales están basados en el esquema FOC Indirecto mostrado en la Fig. 2. Los parámetros del motor están dados por la hoja técnica de especificaciones que se muestran en la Tabla I.

Las ganancias del controlador PI (27) usado para estimar la velocidad del rotor se eligieron como  $K_p = 500$  y  $K_i = 15000$ . La constante proporcional de (15)-(18) es  $k = 0.55$ .

Durante la toma de datos de los resultados, el encoder no cierra el lazo de control pero sus mediciones si son tomadas en cuenta para comprobar que las estimaciones de velocidad conciden con las reales. El resultado del control emulado basado en observador es mostrado en la Fig. 3, se observa que la velocidad estimada y medida por el encoder siguen a la referencia para velocidades cercanas de 900 rpm. En la Fig. 4 se alcanza una velocidad de referencia de 1800 rpm.

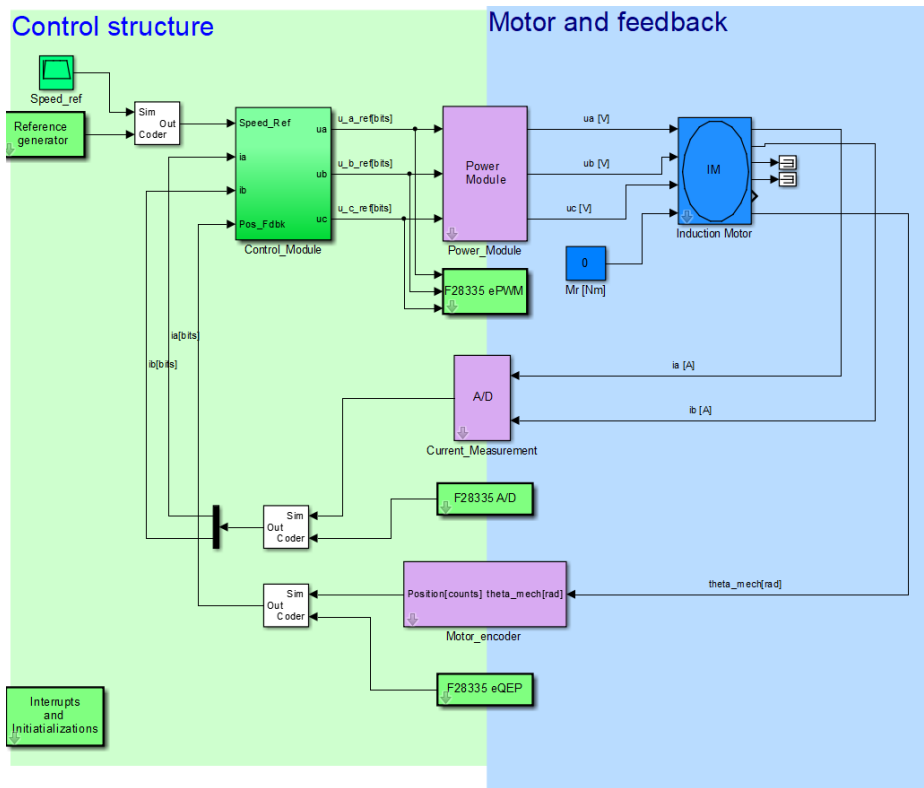


Fig. 2. Emulador de la plataforma Technosoft®DMCode-MS(IM).

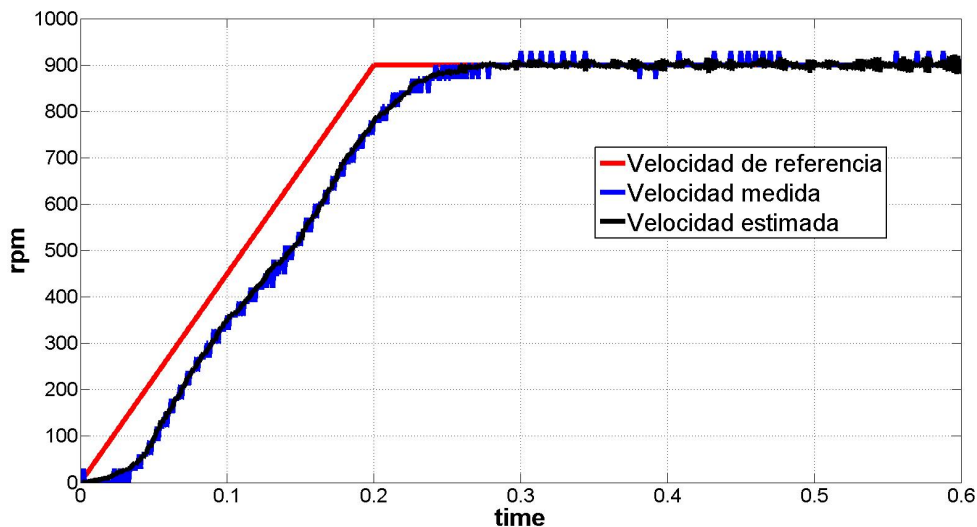


Fig. 3. Emulación: la velocidad estimada y medida del motor emulado alcanzan una velocidad deseada de 900 rpm.

Después de varias pruebas en el emulador de la plataforma, programación y ajustes en el código C para su compilación y descarga en el microprocesador, en la Fig. 5 se muestran los dispositivos usados para la implementación física del observador; una tarjeta como etapa de potencia, una tarjeta con DSP y un motor de inducción conectado en estrella a una

alimentación trifásica de 110V/60Hz.

Para realizar las pruebas físicas se consideraron las referencias usadas en las emulaciones. En la Fig. 6 muestra que tanto la velocidad estimada y real medida por el encoder alcanzan una velocidad de referencia de 900 rpm y en la Fig. 7 la velocidad real del motor alcanza la referencia en 1800 rpm.

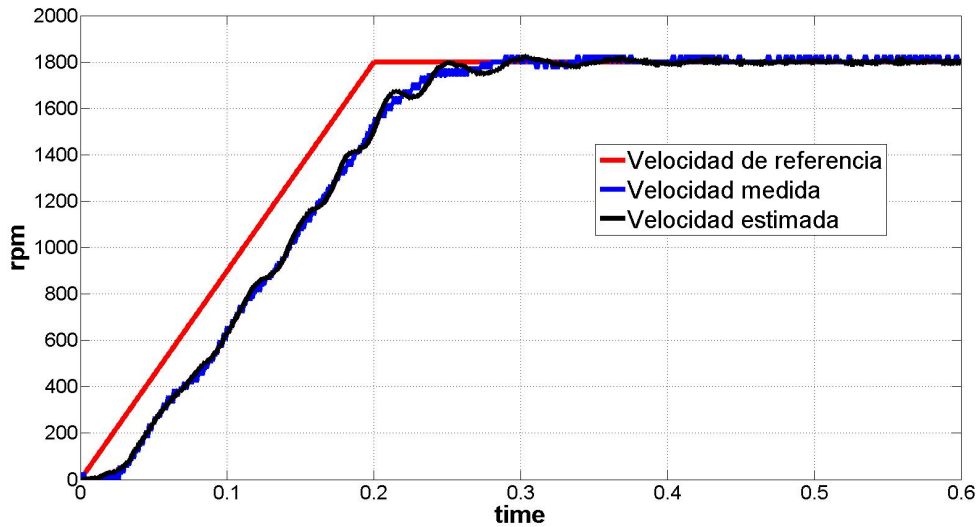


Fig. 4. Emulación: la velocidad estimada y medida del motor emulado alcanzan una velocidad deseada de 1800 rpm.

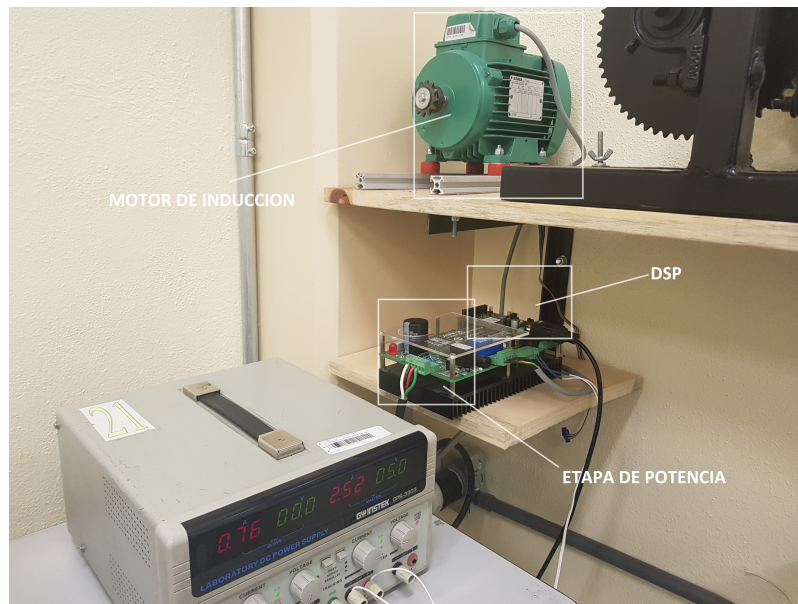


Fig. 5. Plataforma experimental Technosoft®DMCode-MS(IM).

## VI. CONCLUSIÓN

Se propuso un método de control sensorless para un motor de inducción dentro de un esquema de FOC Indirecto mediante un observador de Luenberger adaptativo para estimar la velocidad. El esquema sensorless fue puesto a prueba a nivel físico con la ayuda de la plataforma Technosoft®DMCode-MS(IM). Los resultados emulados fueron bastante satisfactorios para dar paso a la programación e implementación de la parte física. Los resultados finales muestran que la velocidad del motor y la velocidad estimada alcanzan la velocidad de referencia con éxito en estado estable, con ligeras variaciones en las rampas

de subida.

Se pretende en un trabajo futuro realizar pruebas bajo diferentes cargas en la parte física, también diseñar un observador que estime resistencia de estator y la constante de tiempo del rotor; ya que estos parámetros varían con la temperatura afectando las estimaciones del observador.

## AGRADECIMIENTOS

Este trabajo está soportado por el TecNM.

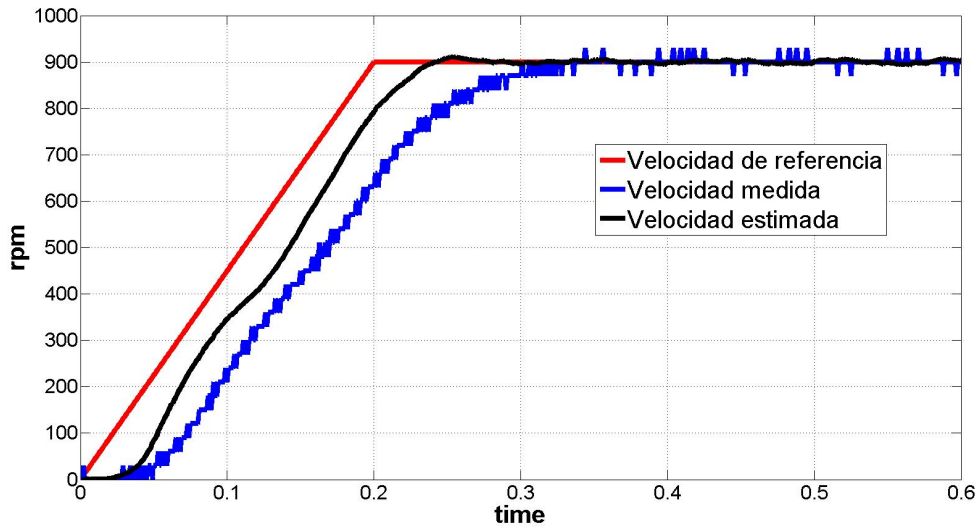


Fig. 6. Implementación: La velocidad estimada y medida real del motor alcanzan una velocidad deseada de 900 rpm.

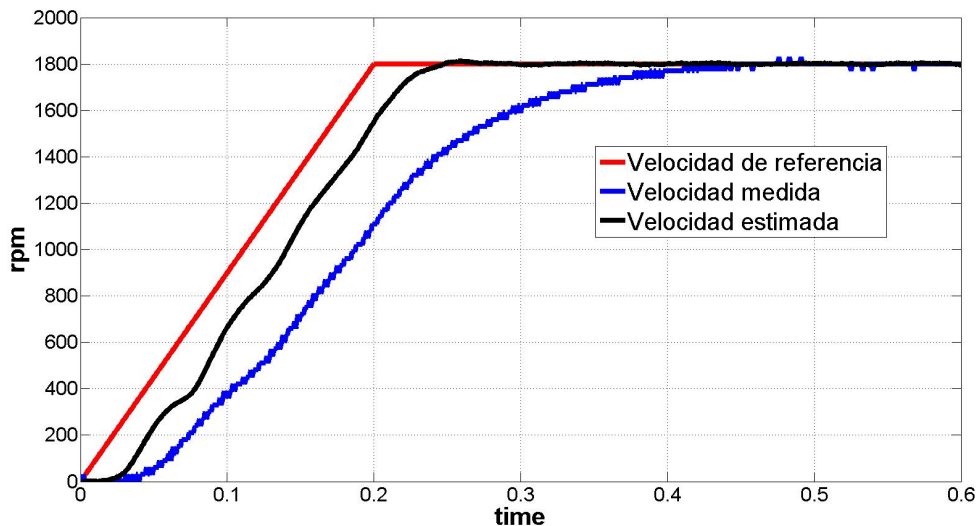


Fig. 7. Implementación: La velocidad estimada y medida real del motor alcanzan una velocidad deseada de 1800 rpm.

#### REFERENCES

- [1] M. Ahmad. (2010) High Performance AC Drives, Power Systems. springerlink.com.
- [2] Robyns, B., Francois, B., Degobert, P., & Hautier, P. (2012). Vector Control of Induction Machines. Longon: Springer.
- [3] Xu, D., Wang, B., Zhang G., Wang, G., & Yu, Y. (2018, Marzo). A review of Sensorless Control Methods for AC Motor Drive. CES Transactions on Electrical Machines and Systems, 2, pp. 104-115.
- [4] You, J., Wu, W., & Wang, Y. (2018). An Adaptive Luenberger Observer for Speed-Sensorless Estimation of Induction Machines. IEEE.
- [5] Zhang, Z., Zhengming, Z., Ting, L., Liqiang, Y., Wei, X., & Jianguo, Z. (2009, 06 November). A comparative study of Luenberger observer, sliding mode observer and extended Kalman filter for sensorless vector control of induction motor drives. IEEE, pp. 2466-2473.
- [6] Hamed, B., & Lassaad S. (2006). Speed Sensorless Indirect Stator Field Oriented Control of Induction Motor Based On Luenberger Observer. IEEE, pp. 2473-2478.
- [7] Kubota, H., Matsuse, K., & Nakano T. (1993). DSP-Based Speed Adaptative Flux Observer of Induction Motor. IEEE, pp. 344-348.
- [8] Saha, S., & Nayak, B. (2017). Sensorless vector control and selection of observer gain for speed control of indirect vector control induction motor drives.
- [9] Technosoft. (2013). DM-Code-S(IM) for MCK28335 plug-in for DMCD-Pro. Switzerland: Technosoft.

# Diseño, construcción y control de un sistema carro-péndulo

Iván Velázquez\*, Jesús Sandoval\*, Jerónimo Moyrón\*, César Higuera<sup>‡</sup>

Marco Castro\*

\* Tecnológico Nacional de México, Instituto Tecnológico de La Paz,  
Blvd. Forjadores, La Paz, B.C.S., 23080, MÉXICO.

E-mail: ivan\_180283@hotmail.com.mx, jsandoval@itlp.edu.mx, moyron\_jero@hotmail.com.mx, mcastro@itlp.edu.mx

<sup>‡</sup>Tecnológico Nacional de México, Instituto Tecnológico de Ensenada,  
Blvd. Tecnológico No. 150, Ex-ejido Chapultepec, Ensenada, MÉXICO  
E-mail: cesarh@itlp.edu.mx

**Resumen**—Se presenta una plataforma didáctica de un sistema subactuado que es popular en numerosos laboratorios de control automático: el carro-péndulo. La plataforma ha sido diseñada y construida en forma modular caracterizada por aprovechar el hardware-software de la marca Quanser<sup>®</sup> (un servo motor, una tarjeta de adquisición de datos y el software QUARC<sup>®</sup>) con que cuenta el laboratorio de control del Instituto Tecnológico de La Paz. La modularidad de la plataforma —que puede ser ensamblada y desarmada manualmente— permite reducir el espacio de almacenamiento. Resultados experimentales son presentados para validar el desempeño de un Regulador Cuadrático Lineal (LQR) diseñado para el control de posición del carro-péndulo propuesto.

**Index Terms**—Carro-péndulo, sistema mecánico subactuado, Quanser, LQR.

## I. Introducción

En las últimas décadas el carro-péndulo se ha convertido en un popular sistema del tipo pendular en control automático. A pesar de su simplicidad, el interés de su estudio se debe principalmente a sus características de no linealidad e inestabilidad, así como al desafío de controlarlo por ser un sistema subactuado —burdamente hablando, un sistema mecánico que tiene más grados de libertad que entradas de control—. Más aún, el sistema carro-péndulo como plataforma de validación de algoritmos de control resulta atractivo para estudiantes e investigadores que buscan comprobar sus resultados, ya sea de nuevos controladores o de aquellos establecidos en la literatura [1], [2], [3]. Ejemplos de controladores no lineales reportados para la regulación de un sistema carro-péndulo son mostrados en [2] y [3], donde ambas propuestas logran llevar el péndulo a su posición vertical superior, mientras el carro es llevado a una posición deseada [2] o únicamente al origen [3].

La plataforma didáctica propuesta en este trabajo es de fácil y rápido ensamblaje para experimentación y además es controlado con el software QUARC<sup>®</sup>, efectivo para simulación numérica y ejecución de control en tiempo real. Una característica de QUARC<sup>®</sup> es la posibilidad

de mostrar simultáneamente en una misma gráfica los resultados de simulación con los datos experimentales.

El resto del trabajo está organizado como sigue: En la sección II, se presenta el diseño del sistema carro-péndulo propuesto e información relevante sobre su construcción. El modelo matemático y el objetivo de control son mostrados en la sección III. La sección IV introduce el diseño de un control LQR, y resultados experimentales para ilustrar su desempeño son presentados en la sección V. Finalmente, damos algunas conclusiones en la sección VI.

## II. Diseño y construcción de un sistema carro-péndulo

El sistema carro-péndulo empleado comúnmente en la literatura de control para fines didácticos es un sistema mecánico subactuado que consiste de un carro moviéndose horizontalmente con un péndulo acoplado al carro, el cual gira libremente (Figura 1). El desplazamiento del

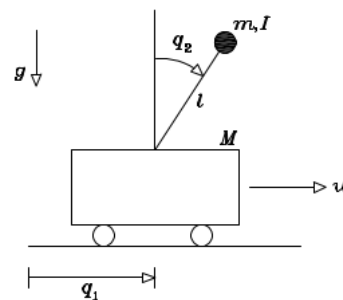


Figura 1. Diagrama esquemático de un típico sistema carro-péndulo.

carro es controlado con un actuador que aplica un par sobre él, mientras la ausencia de actuación en el péndulo determina la naturaleza subactuado del mecanismo [2]. En la Figura 1, las posiciones son  $q_1$  y  $q_2$ , mientras el resto de los parámetros del carro-péndulo: longitud, momento de inercia y masas están dados por  $l$ ,  $I$ ,  $m$  y  $M$ , respectivamente. La aceleración de la gravedad es  $g$ , mientras  $u$  es el par aplicado sobre el carro.



Enseguida, se describe el sistema carro-péndulo propuesto. Para el diseño se buscó que el equipo fuera modular tal que las dimensiones de cada una de las piezas fabricadas se ajustaron al equipo existente de la marca Quanser®. Un diagrama esquemático es mostrado en la Figura 2, donde las piezas en color azul (numeradas como 1 y 2) fueron construidas en una impresora 3D en material PLZ, después de ser diseñadas en el software Inventor 2018 de Autodesk®. El resto de las piezas corresponden al equipo existente de la marca Quanser®, las cuales están numeradas en la figura 2: base con servo motor en color negro –acoplado con la pieza 1–, péndulo numerado con 4 y base-encoder óptico numerado con 3.

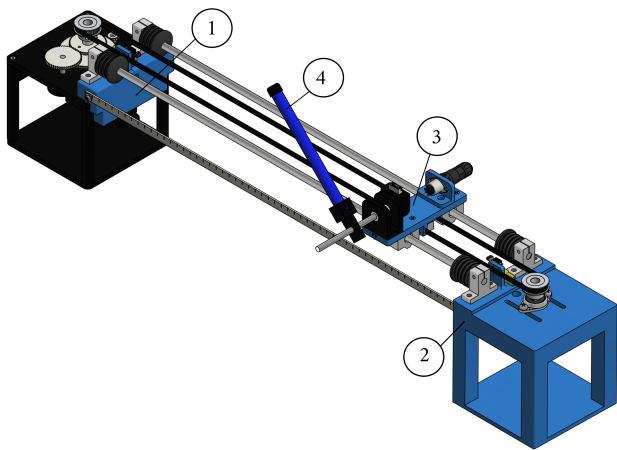


Figura 2. Esquema del sistema carro-péndulo propuesto.

Un acercamiento de la vista frontal del prototipo es mostrada en la figura 3, muestra en detalle el servo motor y el carro-péndulo desplazándose por un riel. Cabe mencionar que el desplazamiento traslacional del “carro” (base numerada con 3 en la figura 2) sobre el riel es producto de una fuerza  $T$  (en color rojo) producida por el acoplamiento del motor y una banda dentada que se mueve entre dos poleas. Esto es, el par del motor se convierte en una fuerza lineal debido al sistema de transmisión de la banda dentada y las poleas.

Los componentes principales del prototipo son:

1. Actuador: es parte de la unidad SRV02 de Quanser® (figura 4a). Para la medición de la posición angular del péndulo se emplea el módulo DBPEN-ROT de Quanser®, el cual ofrece una resolución de 4096 conteos por revolución en modo cuadratura. La adquisición de datos se realiza con la tarjeta Q8-USB v2 y para la etapa de potencia la unidad VoltPAQ-X1 Amplifier, ambos de Quanser® (figura 4b) [4].
2. Base fija: las guías o rieles están sujetas en el otro extremo, también por dos bases de 8 mm y a su vez están colocadas sobre una plataforma impresa, cuya función es la de contrabase. Se coloca un rodamiento

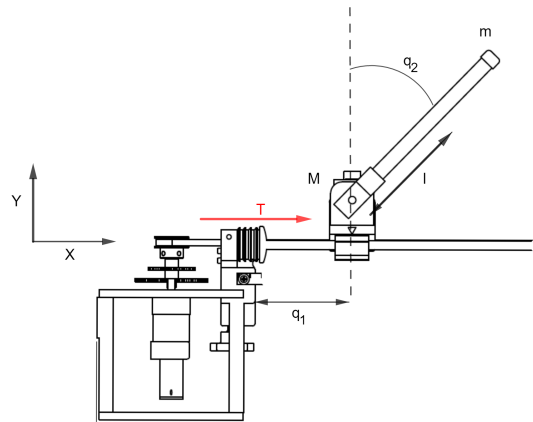


Figura 3. Acercamiento de la vista frontal del sistema carro-péndulo propuesto



(a) SRV02

(b) Q8-USB v2 y VoltPAQ-X1 Amplifier

Figura 4. Equipo Quanser®.

KFL180 para una polea dentada tipo GT2 y un interruptor límite como paro de emergencia del prototipo (figura 5).

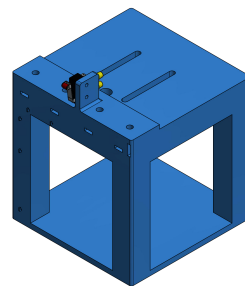


Figura 5. Base fija.

3. Carro (base móvil): consiste en una plataforma que se desplaza de manera horizontal mediante rodamientos lineales sobre dos guías de acero inoxidable. Para transmitir el movimiento entre el tren de engranes del servo motor y el carro, se utiliza un

sistema de poleas dentadas tipo GT2 y una correa dentada con un paso de 1 mm. También sobre el carro se encuentra montado el sensor para la posición angular del péndulo (figura 6).

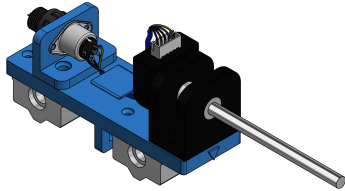


Figura 6. Carro (base móvil).

4. Péndulo: es parte del equipo Rotary Inverted Pendulum del mismo fabricante Quanser® y es mostrado en la figura 2 numerado con 4. Su longitud de 0,337 [m] y tiene un momento de inercia de 0,0012 [kg.m<sup>2</sup>] [4].
5. Soporte: el desplazamiento del carro se realiza sobre dos guías de 8 mm de acero inoxidable, las cuales están sujetas en uno de sus extremos a la unidad SRV02 mediante un soporte impreso y dos bases de piso de 8 mm. Este soporte puede ser manipulado manualmente por el usuario (para montaje y desmontaje) mediante una tuerca mariposa colocada en la parte inferior del soporte. Por motivo de seguridad del equipo y del usuario, se instaló un interruptor límite como paro de emergencia del prototipo.

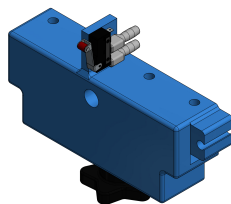


Figura 7. Soporte.

Finalmente, el prototipo final se muestra en la figura 8.

### III. Modelo matemático

#### III-A. Sistema mecánico: carro-péndulo

Las ecuaciones de movimiento sin considerar fricción, pueden ser escritas en forma compacta como:

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + g(q) = \tau \quad (1)$$

donde  $M(q)$  es la matriz de inercias,  $C(q)$  es la matriz de fuerzas centrífugas y de coriolis  $g(q)$  es el vector de pares

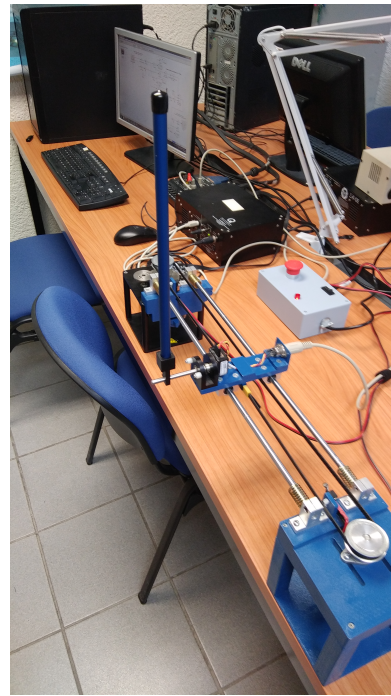


Figura 8. Sistema carro-péndulo propuesto.

gravitacionales y  $\tau$  es el vector de fuerzas externas. Para el modelo del carro-péndulo (1), se tiene que

$$M(q) = \begin{bmatrix} (M + m) & ml \cos(q_2) \\ ml \cos(q_2) & (ml^2 + I) \end{bmatrix}, \quad (2)$$

$$C(q, \dot{q}) = \begin{bmatrix} 0 & -ml \sin(q_2)\dot{q}_2 \\ 0 & 0 \end{bmatrix}, \quad (3)$$

$$g(q) = \begin{bmatrix} 0 \\ -mgl \sin(q_2) \end{bmatrix}, \quad (4)$$

$$\tau = \begin{bmatrix} F \\ 0 \end{bmatrix}, \quad (5)$$

donde  $F$  es la fuerza aplicada al carro para su desplazamiento traslacional. Por tanto, el modelo (1) con (2)-(5) puede ser rescrito explícitamente de la siguiente manera:

$$[M + m]\ddot{q}_1 + ml \cos(q_2)\ddot{q}_2 - ml \sin(q_2)\dot{q}_2^2 = F, \quad (6)$$

$$ml \cos(q_2)\ddot{q}_1 + [ml^2 + I]\ddot{q}_2 - mgl \sin(q_2) = 0. \quad (7)$$

#### III-B. Sistema eléctrico: motor de CD

El motor de CD incorporado en la unidad SRV02, es un motor de imanes permanentes del fabricante Faulhaber Coreless DC Motor, modelo 2338S006. Este es un motor de alta eficiencia que puede obtener una respuesta mucho más rápida que un motor de CD convencional [4]. Para obtener el modelo dinámico del motor se considera el esquema mostrado en la figura 9.

Se asume que la fuerza contraelectromotriz  $e_g$  es proporcional a la velocidad de rotación del rotor,  $\omega_m$ , y la constante de fuerza contraelectromotriz del motor  $k_m$  está dada por

$$e_g = k_m \omega_m \quad (8)$$

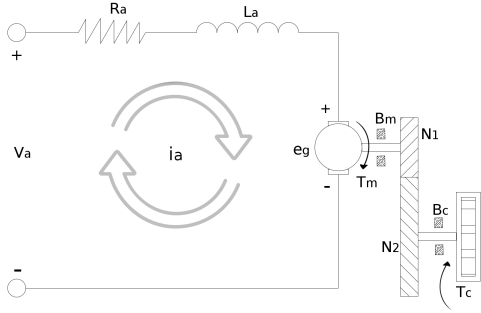


Figura 9. Circuito eléctrico del motor de CD y tren de engranes

Por la ley de voltajes de Kirchof puede obtenerse

$$V_a - R_a i_a - L_a \frac{di_a}{dt} - k_m \omega_m = 0 \quad (9)$$

donde  $V_a$  es el voltaje aplicado en las terminales del motor,  $i_a$  es la corriente de armadura,  $R_a$  y  $L_a$  son la resistencia de armadura e inductancia del motor, respectivamente.

Por otro lado, de acuerdo a la segunda ley de Newton,  $J\alpha = \tau$ , la inercia de carga incluye la inercia del tren de engranes y de cualquier carga externa conectada. La ecuación del eje del motor es expresado como

$$J_m \frac{d\omega_m}{dt} + B_m \omega_m + \tau_{ml} = \tau_m \quad (10)$$

donde  $J_m$  es el momento de inercia del motor,  $\tau_{ml}$  es el par de la carga reflejada sobre el eje del motor y  $B_m$  el coeficiente de fricción viscosa en el eje del motor. El par del motor expresado en termino de la corriente de armadura, la eficiencia del motor  $\eta_m$  y su constante de par  $k_t$ , se define como

$$\tau_m = \eta_m k_t i_a \quad (11)$$

Si se define el par en el eje de la carga  $\tau_l$ , por efecto de un par motor, como el producto de la eficiencia del tren de engranes  $\eta_g$ , la ganancia por la relación de engranes  $k_g$  y  $\tau_{ml}$

$$\tau_l = \eta_g k_g \tau_{ml} \quad (12)$$

se puede expresar  $\tau_{ml}$  de (12)

$$\tau_{ml} = \frac{\tau_l}{\eta_g k_g} \quad (13)$$

considerando la relación del tren engranes y la eficiencia del mismo, la velocidad del motor  $\omega_m$  puede definirse como

$$\omega_m = k_g \omega_l. \quad (14)$$

De (11) resolviendo para la corriente de armadura, y asumiendo que la inductancia del motor puede despreciarse ( $L_a \cong 0$ ), se tiene que

$$i_a = \frac{V_a - k_m \omega_m}{R_a} \quad (15)$$

sustituyendo (15) en (11)

$$\tau_m = \eta_m k_t \left( \frac{V_a - k_m \omega_m}{R_a} \right) \quad (16)$$

igualando (10) y (16), y sustituyendo (13),  $\tau_l$  queda

$$J_m \frac{d\omega_m}{dt} + B_m \omega_m + \frac{\tau_l}{\eta_g k_g} = \eta_m k_t \left( \frac{V_a - k_m \omega_m}{R_a} \right), \quad (17)$$

$$\tau_l = \frac{\eta_g k_g \eta_m k_t}{R_a} V_a - \left( \frac{\eta_g k_g \eta_m k_t k_m}{R_a} + \eta_g k_g B_m \right) \omega_m - \eta_g k_g J_m \frac{d\omega_m}{dt}. \quad (18)$$

### III-C. Modelo completo: carro-péndulo con motor de CD

La velocidad angular en el eje de la carga  $\omega_l$ , en este caso la polea dentada, expresada en términos de la velocidad del carro  $\dot{q}_1$  y el radio  $r$  de la polea, resulta

$$\omega_l = \frac{\dot{q}_1}{r} \quad (19)$$

para expresar la velocidad angular en el eje del motor en términos del desplazamiento lineal del carro, se sustituye (19) en (14)

$$\omega_m = k_g \frac{\dot{q}_1}{r} \quad (20)$$

De la definición de par, el par en el eje de carga es

$$\tau_l = Fr \quad (21)$$

donde la fuerza  $F$  es perpendicular a la dirección del radio  $r$  de la polea desde su centro, y es responsable del desplazamiento del carro. Si igualamos (21) con (18) y sustituimos (20) para obtener  $F$ , se obtiene

$$F = K_v V_a - B_{eq} \dot{q}_1 - J_{eq} \ddot{q}_1 \quad (22)$$

donde las constantes están definidas como

$$K_v = \frac{\eta_g k_g \eta_m k_t}{r R_a},$$

$$B_{eq} = \frac{\eta_g k_g^2 \eta_m k_t k_m}{r^2 R_a} + \frac{\eta_g k_g^2}{r^2} B_m,$$

$$J_{eq} = \frac{\eta_g k_g^2 J_m}{r^2},$$

por conveniencia y dado que el par que genera la fuerza para mover el carro es originado por el motor, éste se puede expresar en términos de un voltaje aplicado en las terminales del motor, igualando (22) y (6), resulta

$$[M+m+J_{eq}] \ddot{q}_1 + ml \cos(q_2) \ddot{q}_2 - ml \sin(q_2) \dot{q}_2^2 + B_{eq} \dot{q}_1 = K_v V_a \quad (23)$$

Las ecuaciones (7) y (23) representan la dinámica completa del sistema en términos del voltaje aplicado al motor.

Finalmente, el objetivo de control es llevar al péndulo a la posición vertical superior,  $q_2 = 0$ , con el carro en una posición deseada  $q_{d1}$ , iniciando desde una vecindad de esta configuración, donde la posición deseada del carro está acotada por los límites físicos del mecanismo. Formalmente, el objetivo de control puede ser establecido como

$$\lim_{t \rightarrow \infty} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} q_{d1} \\ 0 \end{bmatrix}. \quad (24)$$

#### IV. Diseño de un Regulador Cuadrático Lineal (LQR)

Puede verificarse que es posible obtener un modelo lineal alrededor del origen  $[q_1 \ q_2 \ \dot{q}_1 \ \dot{q}_2]^T = [0 \ 0 \ 0 \ 0]^T$  usando la serie de Taylor para aproximar el modelo no lineal (7) y (23) dado por [1]:

$$ml\ddot{q}_1 + [ml^2 + I]\ddot{q}_2 - mglq_2 = 0, \quad (25)$$

$$[M + m + J_{eq}]\ddot{q}_1 + ml\ddot{q}_2 + B_{eq}\dot{q}_1 = K_v V_a \quad (26)$$

Enseguida, se define el vector de estados

$$\mathbf{x} = [x_1 \ x_2 \ x_3 \ x_4]^T = [q_1 \ q_2 \ \dot{q}_1 \ \dot{q}_2]^T \quad (27)$$

con el fin de describir el modelo lineal (25) y (26) en su forma canónica:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u, \quad (28)$$

$$\mathbf{y} = \mathbf{C}\mathbf{x}. \quad (29)$$

Por tanto, luego de algunas manipulaciones algebraicas, el modelo (25) y (26) queda expresado de la siguiente manera

$$\frac{d}{dt} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} \dot{x}_3 \\ \dot{x}_4 \\ \frac{-I_{eq}B_{eq}x_4 - gl^2m^2x_2 + I_{eq}K_vV_a}{D} \\ \frac{mlBx_4 + ml(M+m+J_{eq})gx_2 - mlK_vV_a}{D} \end{bmatrix} \quad (30)$$

donde  $D = -m^2l^2 + (M + m + J_{eq})I_{eq}$  con  $I_{eq} = ml^2 + I$ . Note que la entrada de control de este sistema es el voltaje  $V_a$ . Considerando los valores numéricos de las tablas I y II en (30), se obtienen las matrices en (28) y (29) en forma numérica:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -0,2 & -37,57 & 0 \\ 0 & 46,2 & 173,5 & 0 \end{bmatrix}, \quad (31)$$

$$\mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1,4 \\ -6,5 \end{bmatrix}, \quad (32)$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (33)$$

Para el diseño del LQR se utilizó el software MATLAB®, con el fin de determinar la matriz de ganancias  $K$  de la ley de control:

$$V_a = -K\mathbf{x} \quad (34)$$

Tabla I  
Parámetros de la unidad SRV02

Símbolo	Descripción	Valor
$R_a$	Resistencia de armadura	2,6 $\Omega$
$k_t$	Constante de par del motor	$7,68 \times 10^{-3} \left[ \frac{\text{Nm}}{\text{A}} \right]$
$k_m$	Constante de fuerza contraelectromotriz	$7,68 \times 10^{-3} \left[ \frac{\text{V}}{\text{rad/s}} \right]$
$k_g$	Relación de transmisión total de alta velocidad	70
$\eta_m$	Eficiencia del motor	0.69
$\eta_g$	Eficiencia de caja de engranes	0.90
$J_m$	Momento de inercia del rotor	$3,9 \times 10^{-7} \text{ [Kg m}^2\text{]}$
$B_m$	Coefficiente de fricción viscosa	$1,5 \times 10^{-4} \left[ \frac{\text{Nm}}{\text{rad/s}} \right]$

Tabla II  
Parámetros del péndulo y carro

Símbolo	Descripción	Valor
$m$	Masa del péndulo	0,127 [kg]
$l$	Distancia del pivote al centro de masa	0,156 [m]
$M$	Masa del carro	0,258 [kg]

ya que el modelo (30) con (31)-(33) es controlable. La matriz  $K$  fue obtenida con el comando  $lqr$ , la cual está dada por

$$K = [-14,14 \quad -105,99 \quad -58,65 \quad -16,14] \quad (35)$$

donde se asignaron  $Q$  y  $R$  como

$$Q = \begin{bmatrix} 1000 & 0 & 0 & 0 \\ 0 & 1500 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 100 \end{bmatrix},$$

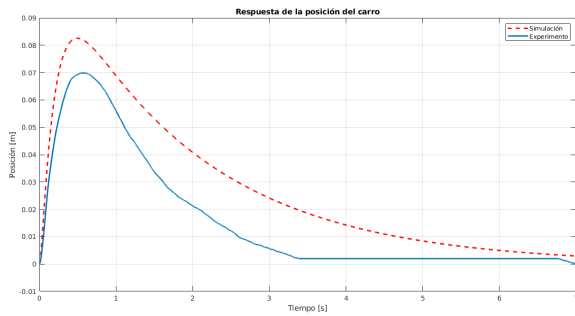
$$R = 1.$$

#### V. Resultados experimentales

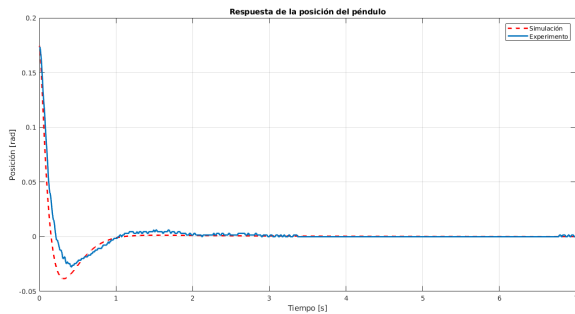
Se llevaron a cabo experimentos para validar el desempeño del LQR dado en (34) con (35). Para la ejecución del control se utilizó Matlab®-Simulink y Quarc® – una aplicación de Quanser®– instalado en un equipo de cómputo que aloja la tarjeta de adquisición de datos MultiQ-PCI®.

Las condiciones iniciales utilizadas para los experimentos fueron:  $[q_1(0) \ q_2(0) \ \dot{q}_1(0) \ \dot{q}_2(0)]^T = [0, 10\pi/180, 0, 0]^T$ . Las ganancias de (35) se modificaron a prueba y error a través de  $Q$  y  $R$  para mejorar la respuesta del experimento, hasta obtener la matriz:  $K = [-14 \quad -105 \quad -52 \quad -25]$ .

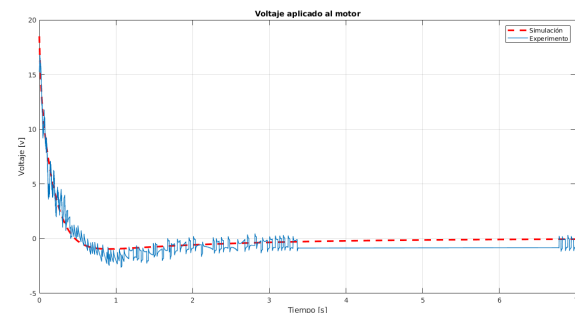
La línea punteada (rojo) en cada una de las gráficas mostradas en la figura (10) corresponde a la respuesta del sistema en simulación, mientras la línea continua (azul) representa los resultados del experimento. Puede notarse un error en el desplazamiento  $q_1$  (figura 10b), el cual está asociado a varios factores como: proceso



(a) Evolución temporal de la posición del carro  $q_1(t)$ .



(b) Evolución temporal de la posición del péndulo  $q_2(t)$ .



(c) Voltaje requerido al motor

Figura 10. Resultados de simulación y experimentales con  $q_{d1} = 0[m]$

de discretización del algoritmo de control, velocidades estimadas a través del método ordinario de Euler; así como a otros fenómenos no modelados. A pesar de lo anterior, el desempeño del controlador es aceptable dado que cumple el objetivo de control (24), inclusive ante una perturbación externa provocada por el usuario al golpear el péndulo momentáneamente.

## VI. Conclusiones

Se ha presentado una plataforma didáctica de un sistema carro-péndulo, y su correcta operación fue comprobada con la implementación de un control LQR que permitió cumplir el objetivo de control (24), al llevar al péndulo a su posición vertical superior con el carro en el origen. Como trabajo futuro se contempla validar algoritmos de control no lineal tanto para control de posición como para otros

objetivos de control que incluyen regulación de orbitas periódicas y regulación de velocidad constante (un control tipo seguimiento de trayectorias).

## AGRADECIMIENTOS

Este trabajo fue financiado parcialmente por el TecNM, contrato 5925.19-P

## Referencias

- [1] I. Siradjuddin, B. Setiawan, A. Fahmi, Z. Amalia, and E. Rohadi, "State space control using LQR method for a cart-inverted pendulum linearised model," *International Journal Mechanical Mechatronics Engineering*, vol. 17, no. 1, pp. 119–126, 2017.
- [2] J. Sandoval y R. Kelly, "Diseño de un nuevo IDA-PBC para la estabilización del sistema carro-péndulo," *Memorias del I Congreso Internacional de Robótica y Computación*, Los Cabos, Baja California Sur, ISBN: 978-607-95534-5-6, pp. 96–100, 2013.
- [3] R. Ortega, M. Spong, F. Gómez-Estern and G. Blankenstein, "Stabilization of a class of underactuated mechanical systems via interconnection and damping assignment," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1213–1233, 2002.
- [4] R. King, "User Manual," Inc. Quanser, may, 2001.

# Control de movimiento para un robot con articulaciones flexibles basado en linealización por realimentación

J. Moreno–Valenzuela

Instituto Politécnico Nacional–CITEDI  
Av. Instituto Politécnico Nacional 1310  
Col. Nueva Tijuana, Tijuana  
Baja California, México, 23435  
Email: moreno@citedi.mx

J. Montoya–Cháirez

Instituto Politécnico Nacional–CITEDI  
Av. Instituto Politécnico Nacional 1310  
Col. Nueva Tijuana, Tijuana  
Baja California, México, 23435  
Email: jamontoya@citedi.mx

V. Santibáñez

Tecnológico Nacional de México  
Instituto Tecnológico de La Laguna  
Blvd.Revolución y Cuauhtémoc SN  
Torreón, México, 27000  
Email: vsantiba@itlalaguna.edu.mx.

**Resumen**—El control en sistemas mecánicos subactuados ha sido una tendencia importante en los últimos años. Como ejemplo están el péndulo de rueda inercial y el giroscopio, que son dispositivos utilizados en sistemas aeroespaciales. También, los sistemas subactuados son excelentes sistemas para probar nuevas leyes de control. En este documento se estudia el caso de un robot con articulaciones flexibles, el cual, al presentar flexibilidad entre sus actuadores y articulaciones, es considerado un sistemas subactuado. Se estudia un controlador de seguimiento de trayectorias basado en la técnica de linealización por realimentación y se comprueban los resultados teóricos por medio de experimentos de tiempo–real. La complejidad del diseño del controlador reside en no sólo considerar las dinámicas de los elementos rígidos, sino también en considerar la flexibilidad en las articulaciones.

## I. INTRODUCCIÓN

Un robot con articulaciones flexibles, se define como un mecanismo con elementos flexibles entre sus articulaciones. El elemento flexible se puede introducir con el propósito de aumentar la seguridad de los humanos al interactuar con robots o se puede introducir de manera no intencionada por los actuadores al entregar la acción de control en las articulaciones. La idea de controlar este tipo de robots viene también de la necesidad de tener brazos robóticos más ligeros, en particular para aplicaciones espaciales [1].

Como ya se ha mencionado en varios trabajos, un robot con articulaciones flexibles se considera un sistema subactuado [2], [3]. La razón por la cual es un sistema subactuado, es por la flexibilidad presentada entre el actuador y el eslabón, lo cual hace imposible que la acción de control entregada por el actuador afecte directamente a el eslabón, como lo es con los robots manipuladores sin elementos elásticos [4].

Algunos trabajos realizados para esta clase de robots, se encuentran reportadas en [1]. El concepto de variedad integral se usó en [5] para aplicar la técnica de linealización por realimentación a un robot con articulaciones flexibles. En [6] se propuso un controlador PID difuso para un robot con articulaciones flexibles con incertidumbres en cargas variantes en el tiempo. En [7] se usó la metodología de moldeo de energía para desarrollar una clase de controladores para manipuladores

con articulaciones flexibles. En [8] se analizó un controlador adaptable basado en modos deslizantes, el cual es propuesto para un robot con una articulación flexible con incertidumbres no coincidentes. En [9] se presentó un control de impedancia para este tipo de manipuladores. En [10] se introdujo un controlador de seguimiento de trayectorias utilizando la técnica de linealización por realimentación transversa. Como último ejemplo, un control de posición por realimentación de estados con moldeo de energía fue aplicado en [11].

En este documento se desarrolla un controlador de movimiento para un robot con articulaciones flexibles de  $n$  grados de libertad. El controlador se deriva bajo la idea de usar la técnica de linealización por realimentación y el diseño de la función de salida inspirado por los diseños mostrados en [12]. Se pretende demostrar que las metodologías mostradas en [12] no están sujetas a sistemas de 2 grados de libertad solamente y que pueden ser aplicadas a sistemas subactuados de mayores grados de libertad. De hecho, el estudio realizado en este documento es una interpretación de linealización por realimentación en su versión no adaptable del controlador propuesto por Tian y Goldenberg [13].

En la siguiente sección, se presenta el modelo dinámico general de un robot con articulaciones flexibles de  $n$  grados de libertad y el planteamiento del objetivo de control. En la sección III se presenta el diseño del controlador. En la sección IV se presentan los resultados experimentales, y finalmente las conclusiones se presentan en la sección V.

## II. MODELO DINÁMICO Y PLANTEAMIENTO DEL PROBLEMA

### A. Modelo dinámico

El modelo de un robot con articulaciones flexibles en el plano horizontal es dado por [14], [15]

$$M(\mathbf{q})\ddot{\mathbf{q}} + C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + F_{vl}\dot{\mathbf{q}} + K(\mathbf{q} - \boldsymbol{\phi}) = \mathbf{0}_{n \times 1}, \quad (1)$$

$$J\ddot{\boldsymbol{\phi}} + F_{vm}\dot{\boldsymbol{\phi}} + K(\boldsymbol{\phi} - \mathbf{q}) = \boldsymbol{\tau}_m, \quad (2)$$



con  $M(\mathbf{q}) \in \mathbb{R}^{n \times n}$  como la matriz de inercia del robot,  $C(\mathbf{q}, \dot{\mathbf{q}}) \in \mathbb{R}^{n \times n}$  es la matriz de fuerzas centrífugas y de Coriolis,  $K \in \mathbb{R}^{n \times n}$  es la matriz asociada con la rigidez de los resortes,  $J$  es la matriz diagonal positiva definida de  $n \times n$  constante que contiene las inercias de los rotores y los engranes,  $F_{vl}$  y  $F_{vr} \in \mathbb{R}^{n \times n}$  son matrices diagonales constantes positivas definidas que contienen los coeficientes de fricción viscosa de las articulaciones, antes y después de los elementos flexibles respectivamente, y  $\boldsymbol{\tau}_m \in \mathbb{R}^n$  es el vector de par de entrada. Las coordenadas  $\mathbf{q}$  y  $\boldsymbol{\phi} \in \mathbb{R}^n$  se refieren a la posición del eslabón y la posición del rotor, respectivamente.

### B. Problema de control

En referencia al robot con articulaciones flexibles en (1)–(2), suponiendo que se conocen los parámetros del sistema y que los estados son medibles, el problema de control consiste en diseñar una ley de control  $\boldsymbol{\tau}_m$  tal que

$$\lim_{t \rightarrow \infty} \mathbf{e}_q(t) = \mathbf{0}_{n \times 1}, \quad (3)$$

donde

$$\mathbf{e}_q = \mathbf{q}_d - \mathbf{q}, \quad (4)$$

donde  $\mathbf{q}_d(t) \in \mathbb{R}^n$  es la posición deseada de los eslabones y se supone que es cuatro veces diferenciable con respecto al tiempo.

A continuación, se presentará la solución a este problema de control presentado en una forma constructiva, es decir, se obtendrá el sistema de lazo abierto en términos del error de posición y par, se propondrá la ecuación de salida, se propondrá el controlador usando la técnica de linealización por realimentación, y se discutirá la estabilidad de lazo cerrado, finalmente.

## III. DISEÑO DEL CONTROL BASADO EN LINEALIZACIÓN POR REALIMENTACIÓN

### A. Sistema a lazo abierto

Usando las ideas en [13], se define un cambio de variable

$$\boldsymbol{\tau} = K(\boldsymbol{\phi} - \mathbf{q}), \quad (5)$$

el cual se puede aplicar en la ecuación (2) para producir

$$JK^{-1}\ddot{\boldsymbol{\tau}} + F_{vm}K^{-1}\dot{\boldsymbol{\tau}} + \boldsymbol{\tau} = \mathbf{u}, \quad (6)$$

donde

$$\mathbf{u} = \boldsymbol{\tau}_m - JM(\mathbf{q})^{-1}[\boldsymbol{\tau} - C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} - F_{vl}\dot{\mathbf{q}}] - F_{vm}\dot{\mathbf{q}}. \quad (7)$$

Usando la definición del error de posición de los eslabones  $\mathbf{e}_q$  en (4) el error de par

$$\mathbf{e}_\tau = \boldsymbol{\tau}_d - \boldsymbol{\tau},$$

con  $\boldsymbol{\tau}_d$  como el par deseado. Entonces las ecuaciones (1) y (6) conforman el sistema a lazo abierto, el cual puede expresarse en la forma de espacio de estados, como

$$\frac{d}{dt}\mathbf{x} = \mathbf{f}(t, \mathbf{x}) + G(t, \mathbf{x})\mathbf{u}, \quad (8)$$

donde

$$\mathbf{x} = [\mathbf{e}_q^T \dot{\mathbf{e}}_q^T \mathbf{e}_\tau^T \dot{\mathbf{e}}_\tau^T]^T \in \mathbb{R}^{4n}, \quad (9)$$

$$\mathbf{f}(t, \mathbf{x}) = \begin{bmatrix} \ddot{\mathbf{q}}_d + M(\mathbf{q})^{-1}[C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + F_{vl}\dot{\mathbf{q}} - \boldsymbol{\tau}_d + \mathbf{e}_\tau] \\ \dot{\mathbf{e}}_q \\ \ddot{\boldsymbol{\tau}}_d + KJ^{-1}[F_{vm}\dot{\boldsymbol{\tau}}_d + \boldsymbol{\tau}_d - F_{vm}\dot{\mathbf{e}}_\tau - \mathbf{e}_\tau] \end{bmatrix}, \quad (10)$$

y

$$G(t, \mathbf{x}) = \begin{bmatrix} \mathbf{0}_{n \times n} \\ \mathbf{0}_{n \times n} \\ \mathbf{0}_{n \times n} \\ -KJ^{-1} \end{bmatrix}.$$

Nótese que un robot con articulaciones flexibles es un sistema subactuado de dimensión  $4n$  representado por los estados  $\mathbf{x}$  en (9). Por lo tanto, el objetivo de control en (3) no se puede cumplir con cualquier controlador [3].

Se hace la suposición de que los estados  $\mathbf{x} \in \mathbb{R}^{4n}$  son medibles. Bajo realimentación estática de  $\mathbf{u}(t, \mathbf{x})$  en (7), se puede obtener la señal auxiliar de control  $\boldsymbol{\tau}_m$ .

### B. Diseño de la función de salida y el control de linealización por realimentación

Inspirado en los controladores de linealización por realimentación desarrollados en [12], la función de salida consiste en una combinación lineal de los estados  $\mathbf{x}$ , como sigue:

$$\mathbf{y} = \alpha_1 \mathbf{e}_q + \alpha_2 \mathbf{e}_\tau + \beta_1 \dot{\mathbf{e}}_q + \beta_2 \dot{\mathbf{e}}_\tau,$$

donde  $\mathbf{y} \in \mathbb{R}^n$  y  $\alpha_1, \alpha_2, \beta_1$  y  $\beta_2$  son matrices constantes de dimensión  $n \times n$  a ser definidas.

A continuación, se implementará la técnica de linealización por realimentación para formular la entrada de control  $\mathbf{u}$  para el sistema (8). La derivada temporal de la función de salida  $\mathbf{y}$  es dada por

$$\begin{aligned} \dot{\mathbf{y}} &= \alpha_1 \dot{\mathbf{e}}_q + \alpha_2 \dot{\mathbf{e}}_\tau + \beta_1 \ddot{\mathbf{e}}_q + \beta_2 \ddot{\mathbf{e}}_\tau \\ &= \alpha_1 \dot{\mathbf{e}}_q + \alpha_2 \dot{\mathbf{e}}_\tau + \beta_1 \mathbf{f}_2 + \beta_2 \mathbf{f}_4 - \beta_2 KJ^{-1}\mathbf{u}, \end{aligned}$$

donde  $\mathbf{f}_2, \mathbf{f}_4 \in \mathbb{R}^n$  son el segundo y cuarto vector componente de  $\mathbf{f} \in \mathbb{R}^{4n}$  en (10), respectivamente. El grado relativo es  $m = n$ , dado que la salida tiene dimensión  $n$  y después de derivarla una vez con respecto al tiempo, la entrada de control aparece.

La ley de control se puede escribir como

$$\mathbf{u} = [\beta_2 KJ^{-1}]^{-1}[\alpha_1 \dot{\mathbf{e}}_q + \alpha_2 \dot{\mathbf{e}}_\tau + \beta_1 \mathbf{f}_2 + \beta_2 \mathbf{f}_4 + K_{po}\mathbf{y}], \quad (11)$$

donde  $K_{po}$  es una matriz simétrica definida positiva de dimensión  $n \times n$ , la cual convierte a la dinámica de salida en el siguiente sistema lineal

$$\dot{\mathbf{y}} = -K_{po}\mathbf{y},$$

el cual es globalmente exponencialmente estable, es decir,  $-K_{po}$  es una matriz Hurwitz.

El número de coordenadas independientes  $\eta_i(\mathbf{x})$  que expresan la dinámica interna es

$$i = 4n - m = 4n - n = 3n.$$

Estas coordenadas deben satisfacer

$$\frac{\partial \eta_i(\mathbf{x})}{\partial \mathbf{x}} G(t, \mathbf{x}) = 0_{1 \times 4n}.$$

Específicamente, se ha seleccionado

$$\boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_q \\ \dot{\mathbf{e}}_q \\ \mathbf{e}_\tau \end{bmatrix} \in \mathbb{R}^{3n}$$

como las coordenadas independientes que expresan la dinámica interna. Por lo tanto,

$$\mathbf{z} = \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \\ \mathbf{y} \end{bmatrix} = T\mathbf{x} = \begin{bmatrix} I_n & 0 & 0 & 0 \\ 0 & I_n & 0 & 0 \\ 0 & 0 & I_n & 0 \\ \alpha_1 & \beta_1 & \alpha_2 & \beta_2 \end{bmatrix} \begin{bmatrix} \mathbf{e}_q \\ \dot{\mathbf{e}}_q \\ \mathbf{e}_\tau \\ \dot{\mathbf{e}}_\tau \end{bmatrix} \quad (12)$$

es la transformación a obtener del sistema a lazo cerrado. Específicamente, la inversa de la transformación (12) es dada por

$$\mathbf{x} = T^{-1}\mathbf{z} = \begin{bmatrix} I_n & 0 & 0 & 0 \\ 0 & I_n & 0 & 0 \\ 0 & 0 & I_n & 0 \\ -\beta_2^{-1}\alpha_1 & -\beta_2^{-1}\beta_1 & -\beta_2^{-1}\alpha_2 & \beta_2^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \\ \mathbf{y} \end{bmatrix} \quad (13)$$

Las ecuaciones (12) y (13) son útiles en la obtención de la dinámica interna, la cual es dada por

$$\begin{aligned} \dot{\boldsymbol{\eta}}_1 &= \boldsymbol{\eta}_2, \\ \dot{\boldsymbol{\eta}}_2 &= \dot{\mathbf{q}}_d + M(\mathbf{q})^{-1}[C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + F_{vl}\dot{\mathbf{q}} - \boldsymbol{\tau}_d + \boldsymbol{\eta}_3], \\ \dot{\boldsymbol{\eta}}_3 &= -\beta_2^{-1}[\alpha_1\boldsymbol{\eta}_1 - \beta_1\boldsymbol{\eta}_2 - \alpha_2\boldsymbol{\eta}_3 + \mathbf{y}]. \end{aligned} \quad (14)$$

Nótese que (14) tiene la estructura de un manipulador de articulaciones rígidas, tomando en cuenta  $\boldsymbol{\tau}_d$  como la entrada de control. Para estabilizar la dinámica interna, se propone la siguiente señal de par deseado:

$$\boldsymbol{\tau}_d = M(\mathbf{q})[\ddot{\mathbf{q}}_d + K_d\dot{\boldsymbol{\eta}}_2 + K_p\boldsymbol{\eta}_1] + C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + F_{vl}\dot{\mathbf{q}}, \quad (16)$$

el cual tiene la estructura de un par calculado [4]. En [13], una estructura basada en pasividad para el par deseado  $\boldsymbol{\tau}_d$  fue usada. Sin embargo, en este documento se propone la señal en (16) como una alternativa. Cabe mencionar que la señal  $\boldsymbol{\tau}_d$  podría proponerse en general, como una señal que establezca la dinámica interna (14)–(15).

Sustituyendo (16) en (14), el sistema a lazo cerrado puede ser escrito como

$$\frac{d}{dt} \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \end{bmatrix} = \begin{bmatrix} 0_{n \times n} & I_n & 0_{n \times n} \\ -K_p & -K_d & M(\mathbf{q}_d - \mathbf{e}_q)^{-1} \\ -\beta_2^{-1}\alpha_1 & -\beta_2^{-1}\beta_1 & -\beta_2^{-1}\alpha_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \beta_2^{-1} \end{bmatrix} \mathbf{y}, \quad (17)$$

$$\frac{d}{dt} \mathbf{y} = -K_{po}\mathbf{y}, \quad (18)$$

Es posible demostrar que hay condiciones suficientes para que el sistema a lazo cerrado (17)–(18) sea exponencialmente estable, las cuales consisten en que la matriz

$$A(t) = \begin{bmatrix} 0_{n \times n} & I_n & 0_{n \times n} \\ -K_p & -K_d & M(\mathbf{q}_d - \mathbf{e}_q)^{-1} \\ -\beta_2^{-1}\alpha_1 & -\beta_2^{-1}\beta_1 & -\beta_2^{-1}\alpha_2 \end{bmatrix} \quad (19)$$

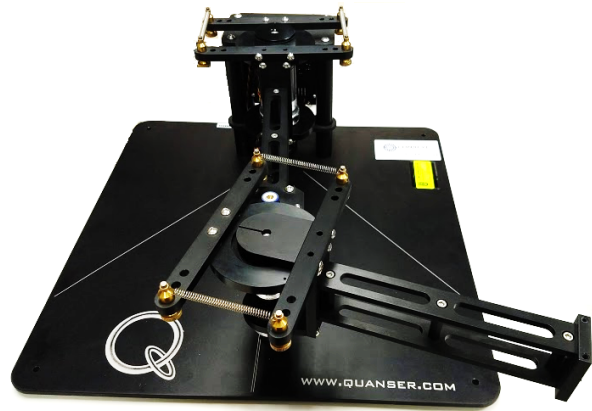


Fig. 1. Robot con articulaciones flexibles.

satisfaga la ecuación de Lyapunov

$$-\dot{P}(t) = P(t)A(t) + A(t)^T P(t) + Q(t), \quad (20)$$

con  $P(t)$  y  $Q(t)$  siendo matrices definidas positivas continuamente diferenciables simétricas y acotadas para todo  $t \geq 0$ , y que  $-K_{po}$  sea Hurwitz.

Las condiciones en la ganancias  $K_p$ ,  $K_d$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\beta_1$  y  $\beta_2$  para que  $A(t)$  cumpla (20) con alguna  $P(t)$  se puede obtener usando el análisis de estabilidad basado en Lyapunov.

De las ecuaciones (7) y (11), el controlador se puede resumir como sigue:

$$\begin{aligned} \boldsymbol{\tau}_m &= \mathbf{u} + JM(\mathbf{q})^{-1}[\boldsymbol{\tau} + C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + F_{vl}\dot{\mathbf{q}}] + F_{vm}\dot{\mathbf{q}}, \\ \mathbf{u} &= [\beta_2 KJ^{-1}]^{-1}[\alpha_1\dot{\mathbf{e}}_q + \alpha_2\dot{\mathbf{e}}_\tau \\ &\quad + \beta_1[\dot{\mathbf{q}}_d + M(\mathbf{q})^{-1}[C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + F_{vl}\dot{\mathbf{q}} - \boldsymbol{\tau}_d + \mathbf{e}_\tau]] \\ &\quad + \beta_2[\ddot{\boldsymbol{\tau}}_d + KJ^{-1}[F_{vm}\dot{\boldsymbol{\tau}}_d + \boldsymbol{\tau}_d - F_{vm}\dot{\mathbf{e}}_\tau - \mathbf{e}_\tau]] \\ &\quad + K_{po}\mathbf{y}]. \end{aligned} \quad (22)$$

Nótese que la primer y segunda derivada del par deseado, dada por  $\dot{\boldsymbol{\tau}}_d$  y  $\ddot{\boldsymbol{\tau}}_d$ , respectivamente, requieren la aceleración de los eslabones  $\dot{\mathbf{q}}$ , la cual puede obtenerse de la ecuación (1) dado que la matriz  $M(\mathbf{q})$  es invertible.

Una observación importante es que la implementación del controlador de linealización por realimentación (21)–(22), con  $\boldsymbol{\tau}_d$  en (16), requiere que la posición deseada de los eslabones  $\mathbf{q}_d(t)$  sea al menos cuatro veces diferenciable con respecto al tiempo.

## IV. EXPERIMENTOS

### A. Descripción de la plataforma

Para realizar experimentos se utilizó la plataforma mostrada en la figura 1, la cual consiste en un robot con articulaciones flexibles de 2 grados de libertad de Quanser. Los motores cuentan con una transmisión de 100 : 1 y de 50 : 1, respectivamente, y cuentan con codificadores ópticos de 4096 pulsos por revolución en cuadratura para cada motor y cada articulación. Los datos se obtienen mediante el dispositivo de adquisición de datos (DAQ) Q8-USB, el cual está optimizado para trabajar a tiempo real con Matlab Simulink®. Los motores se alimentan utilizando el amplificador de corriente de Quanser AMPAQ.

La trayectoria deseada es

$$\mathbf{q}_d = \begin{bmatrix} A \sin(\omega t) \\ -A \sin(\omega t) \end{bmatrix}, \quad (23)$$

TABLA I  
 LISTA DE PARÁMETROS DEL ROBOT CON ARTICULACIONES FLEXIBLES.

Símbolo	Valor
$\theta_1$	0.207184
$\theta_2$	0.017580
$\theta_3$	0.013163
$\theta_4$	0.216776
$\theta_5$	0.216776
$\theta_6$	0.006842
$\theta_7$	0.037675
$\theta_8$	0.002959
$\theta_9$	0.135564
$\theta_{10}$	9.358730
$\theta_{11}$	4.212811

donde  $A = \pi/9$  [rad] =  $20^\circ$  y  $\omega = 2\pi/T$  con  $T = 2$  [s]. Los elementos correspondientes al modelo dinámico en (1), para este caso son:

$$\begin{aligned}
 M(\mathbf{q}) &= \begin{bmatrix} \theta_2 + 2\theta_2 \cos(q_2) & \theta_3 + \theta_2 \cos(q_2) \\ \theta_3 + \theta_2 \cos(q_2) & \theta_3 \end{bmatrix}, \\
 C(\mathbf{q}, \dot{\mathbf{q}}) &= \begin{bmatrix} -\theta_2 \sin(q_2) \dot{q}_2 & -\theta_2 \sin(q_2) (\dot{q}_1 + \dot{q}_2) \\ \theta_2 \sin(q_2) \dot{q}_1 & 0 \end{bmatrix}, \\
 F_{vl} &= \begin{bmatrix} \theta_6 & 0 \\ 0 & \theta_7 \end{bmatrix}, \quad (24)
 \end{aligned}$$

y las matrices relacionadas con los motores mostradas en (2) para este caso son

$$\begin{aligned}
 J &= \begin{bmatrix} \theta_4 & 0 \\ 0 & \theta_5 \end{bmatrix}, \\
 F_{vm} &= \begin{bmatrix} \theta_8 & 0 \\ 0 & \theta_9 \end{bmatrix}, \\
 K &= \begin{bmatrix} \theta_{10} & 0 \\ 0 & \theta_{11} \end{bmatrix}.
 \end{aligned}$$

Los parámetros se encuentran en la tabla I donde fueron calculados en [15]. Las ganancias elegidas para la ley de control (21), (22) y para el par deseado (16), son

$$\begin{aligned}
 \alpha_1 &= \begin{bmatrix} 1500 & 0 \\ 0 & 4000 \end{bmatrix}, \quad K_{po} = \begin{bmatrix} 30 & 0 \\ 0 & 40 \end{bmatrix}, \\
 \alpha_2 &= \begin{bmatrix} 100 & 0 \\ 0 & 80 \end{bmatrix}, \quad K_p = \begin{bmatrix} 50 & 0 \\ 0 & 100 \end{bmatrix}, \\
 \beta_1 &= \begin{bmatrix} 50 & 0 \\ 0 & 50 \end{bmatrix}, \quad K_d = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix}, \\
 \beta_2 &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.
 \end{aligned}$$

Nótese que la posición deseada  $\phi_d$  se puede calcular indirectamente de la siguiente manera

$$\phi_d = K^- \tau_d + \mathbf{q}_d. \quad (25)$$

### B. Resultados experimentales

En la figura 2, se puede observar que tanto la posición de las articulaciones  $\mathbf{q}(t)$  como la posición de los motores  $\phi(t)$  siguen satisfactoriamente las trayectorias deseadas  $\mathbf{q}_d(t)$  y  $\phi_d(t)$  descritas en (23) y (25), respectivamente. La figura 3 muestra como el par generado por (5) sigue a la señal del par deseado  $\tau_d(t)$  mostrado en la ecuación (16). Finalmente, la Fig. 4 muestra la acción de control  $\tau_m(t)$  aplicada por los motores para cumplir el objetivo de control.

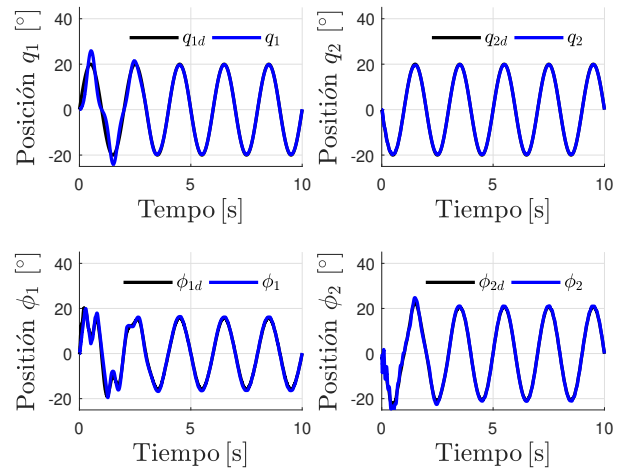


Fig. 2. Evolución temporal de las coordenadas  $q_{d1}(t)$ ,  $q_1(t)$ ,  $q_{d2}(t)$ ,  $q_2(t)$ ,  $\phi_{d1}(t)$ ,  $\phi_1(t)$ ,  $\phi_{d2}(t)$  y  $\phi_2(t)$ .

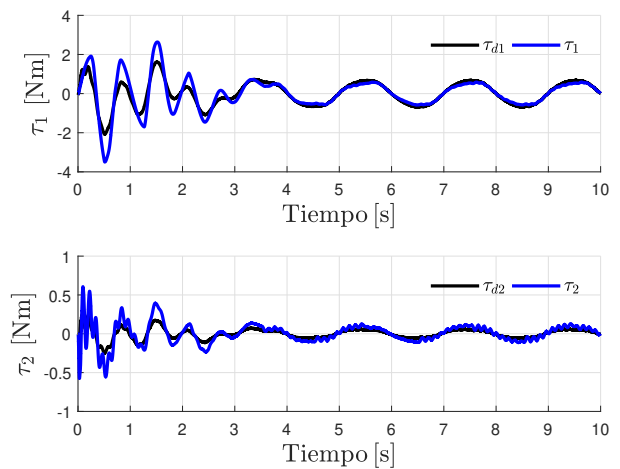


Fig. 3. Evolución temporal de las coordenadas  $\tau_{d1}$ ,  $\tau_1$ ,  $\tau_{d2}$  y  $\tau_2$ .

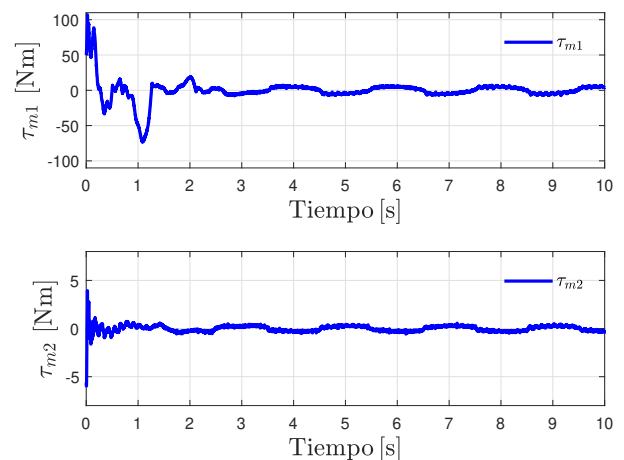


Fig. 4. Evolución temporal de las entradas de control  $\tau_{m1}$  y  $\tau_{m2}$ .

## V. CONCLUSIONES

En este documento se estudia un control de seguimiento de trayectorias aplicado a un robot con articulaciones flexibles. El controlador se basa en la técnica de linealización por realimentación, utilizando una función de salida generalizada, la cual permite controlar tanto las coordenadas subactuadas como las actuadas. Se demuestra que el método presentado en este documento funciona para cierta clase de manipuladores subactuados de más de dos grados de libertad. Como se observa, el método presenta un buen desempeño debido a que se tiene conocimiento del modelo dinámico, tanto de su estructura como de los parámetros.

## AGRADECIMIENTOS

Este proyecto es apoyado parcialmente por SIP-IPN, por los proyectos CONACYT A1-S-24762, CONACYT 134534 y Proyectos TecNM.

## REFERENCIAS

- [1] I. Fantoni y R. Lozano, *Non-linear control for underactuated mechanical systems*, Springer-Verlag, 2002.
- [2] R. Ortega, J. A. L. Perez, P. J. Nicklasson y H. Sira-Ramirez, *Passivity-based control of Euler-Lagrange systems: mechanical, electrical and electromechanical applications*, Springer Science & Business Media, 2013.
- [3] C. Ott, A. Albu-Schaffer, A. Kugi, y G. Hirzinger, "On the passivity-based impedance control of flexible joint robots", *IEEE Transactions on Robotics*, vol. 24, no. 2, pp. 416-429, 2008.
- [4] R. Kelly, V. Santibáñez, y A. Loría, *Control of robot manipulators in joint space*, Springer-Verlag, London, 2005.
- [5] M. W. Spong, K. Khorasani y P. V. Kokotovic, "An integral manifold approach to the feedback control of flexible joint robots", *IEEE Journal of robotics and Automation*, vol. RA-3, no. 4, pp. 291-300, 1987.
- [6] H. A. Malki, D. Misir, D. Feigenspan y G. Chen, "Fuzzy PID control of a flexible-joint robot arm with uncertainties from time-varying load", *IEEE Transactions on Control Systems Technology*, vol. 5, no. 3, pp. 371-378, 1997.
- [7] R. Kelly y V. Santibáñez, "Global Regulation of Elastic Joint Robots Based on Energy Shaping", *IEEE Transactions on automatic control*, vol. 43, no. 10, pp. 1451-1456, 1998.
- [8] A. C. Huang, y Y. C. Chen, "Adaptive sliding control for single-link flexible-joint robot with mismatched uncertainties", *IEEE Transactions on Control Systems Technology*, vol. 12, no. 5, pp. 770-775, 2004.
- [9] C. Ott, *Cartesian impedance control of redundant and flexible-joint robots*, Springer-Verlag, 2008.
- [10] B. Bischof, T. Glück, M. Böck y A. Kugl, "Path following control for elastic joint robots", *IFAC-PapersOnLine*, vol. 50, no. 1, pp 4806-4811, 2017.
- [11] W. Yin, L. Sun, M. Wang y J. Liu, "Nonlinear state feedback position control for flexible joint robot with energy shaping", *Robotics and Autonomous Systems*, vol. 99, pp 121-134, 2018.
- [12] J. Moreno-Valenzuela, y C. Aguilar-Avelar, *Motion Control of Underactuated Mechanical Systems*, Cham: Springer International Publishing AG, 2018
- [13] L. Tian y A. A. Goldenberg, "Robust adaptive control of flexible joint robots with joint torque feedback". *En memorias del 1995 IEEE International Conference on Robotics and Automation*, vol. 1, pp. 1229-1234, 1995.
- [14] C. Canudas de Wit, *Theory of robot control*, Springer-Verlag, London, 1996.
- [15] R. Miranda-Colorado y J. Moreno-Valenzuela, "Experimental parameter identification of flexible joint robot manipulators", *Robotica*, vol. 36, pp. 313-332, 2018.

# Control PID para la regulación de velocidad del péndulo con rueda inercial

Jerónimo Moyrón Durán  
Tecnológico Nacional de México  
Instituto Tecnológico de La Paz  
La Paz, B.C.S., México C.P. 23080  
Email: moyronmsc1702@itlp.edu.mx

Jesús Alberto Sandoval Galarza  
Tecnológico Nacional de México  
Instituto Tecnológico de La Paz  
La Paz, B.C.S., México C.P. 23080  
Email: jsandoval@itlp.edu.mx

**Resumen**—Recientemente en [1] fue presentado un controlador PID para la regulación de velocidad de una clase de sistemas mecánicos subactuados. El controlador propuesto permite regular solo la velocidad de las coordenadas actuadas del mecanismo, mientras las coordenadas subactuadas son llevadas a una posición deseada. En este trabajo se presenta dicho controlador aplicado a un mecanismo que pertenece a la clase de sistemas subactuados de [1]: el péndulo con rueda inercial. Simulaciones sobre el modelo de un péndulo con rueda inercial sin fricción son mostradas para validar el desempeño del regulador de velocidad propuesto.

**Palabras clave**—Sistemas mecánicos, control no lineal, regulación.

## I. INTRODUCCIÓN

Hace casi dos décadas fueron introducidos dos exitosos métodos de diseño de control para la regulación de una clase de sistemas mecánicos subactuados, los cuales se caracterizan por incorporar conceptos de energía e inyección de amortiguamiento [2], [3]. Aún cuando ambos métodos han permitido diseñar leyes de control para el control de posición de un amplio número de sistemas mecánicos subactuados de laboratorio, presentan el desafío de resolver un conjunto de ecuaciones diferenciales parciales. Recientes propuestas han logrado establecer condiciones específicas sobre la clase de sistemas mecánicos subactuados que se desea estabilizar usando variantes al método mostrado originalmente en [2], donde ya no es necesario resolver ecuaciones diferenciales parciales. Una de estas propuestas es mostrada en [4], la cual también utiliza conceptos de energía, pero con la característica que el control propuesto tiene una estructura del tipo PID. Cabe destacar que un paso importante en el diseño del controlador introducido en [4] es la incorporación de un lazo de control interno, el cual resulta de emplear la técnica de linealización parcial por realimentación.

Motivados por evitar la linealización parcial en el esquema de diseño presentado en [4], los autores en [1] presentan una propuesta alternativa en formulación hamiltoniana, la cual usa un cambio de coordenadas apropiado para el diseño de un control tipo PID. Además, en [1] se extiende el método de diseño propuesto (que originalmente aborda el problema de control de posición) para la regulación de velocidad constante en las coordenadas actuadas (articulaciones activas, esto es,

articulaciones dotadas con actuadores), mientras las coordenadas subactuadas (articulaciones pasivas, esto es, articulaciones que no cuentan con actuadores) son llevadas a una posición deseada.

En este trabajo se emplea el controlador de [1] para la regulación de velocidad de un sistema péndulo con rueda inercial sin fricción. Específicamente, se logra que la rueda gire a una velocidad constante deseada, mientras el péndulo es llevado a su configuración vertical invertida.

El resto del trabajo está organizado de la siguiente manera. En la sección II se presenta un breve resumen de la metodología introducida en [1] y su extensión para la regulación de velocidad de una clase de sistemas mecánicos subactuados con matriz de inercia constante, a la cual pertenece el péndulo con rueda inercial. La sección III muestra el diseño del controlador para el sistema mecánico propuesto en este trabajo. Resultados de simulación numérica son mostrados en la sección IV para ilustrar el desempeño del controlador propuesto. Finalmente, en la sección V se dan algunas conclusiones.

## II. REGULACIÓN DE VELOCIDAD DE UNA CLASE DE SISTEMAS MECÁNICOS SUBACTUADOS

En esta sección se presenta un breve resumen de la propuesta introducida en [1] para la regulación de velocidad de una clase de sistemas mecánicos subactuados con matriz de inercia constante<sup>1</sup>.

### II-A. Modelo dinámico (Planta a ser controlada)

Las ecuaciones de movimiento de la clase de sistemas mecánicos subactuados sin fricción viscosa considerada en este trabajo, pueden ser escritas como

$$\frac{d}{dt} \begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} 0_{n \times n} & I_{n \times n} \\ -I_{n \times n} & 0_{n \times n} \end{bmatrix} \begin{bmatrix} \nabla_q H(q, p) \\ \nabla_p H(q, p) \end{bmatrix} + \begin{bmatrix} 0 \\ G \end{bmatrix} u \quad (1)$$

donde la función de energía puede ser definida como

$$H(q, p) = \frac{1}{2} p^T M^{-1} p + V(q) \quad (2)$$

donde  $q, p \in \mathbb{R}^n$  son los vectores de posiciones y momentos generalizados, respectivamente,  $M = M^T > 0$  es la matriz

<sup>1</sup>El resultado presentado por [1] es más general y puede ser aplicado a sistemas mecánicos subactuados, cuya matriz de inercia depende de la coordenada subactuada (e.g. el sistema carro-péndulo).

de inercia constante,  $V(q)$  es la función de energía potencial, la cual se asume que es al menos una vez diferenciable y  $G \in \mathbb{R}^{n \times m}$  es la matriz de distribución de entradas de control, donde  $\text{rango}\{G\} = m$  y  $m < n$ . Se asume que la matriz  $G$  tiene la siguiente estructura:

$$G = \begin{bmatrix} 0_{s \times m} \\ I_{m \times m} \end{bmatrix} \quad (3)$$

donde  $s = n - m$ . Finalmente,  $u \in \mathbb{R}^m$  es el vector de entradas de control. Enseguida, los vectores de posiciones y momentos generalizados son separados de la siguiente manera:  $q = [q_u^T \ q_a^T]^T$ ,  $p = [p_u^T \ p_a^T]^T$  con  $q_u, p_u \in \mathbb{R}^s$  y  $q_a, p_a \in \mathbb{R}^m$ , así como también se asume que la matriz de inercia tiene la siguiente estructura:

$$M = \begin{bmatrix} m_{uu} & m_{au}^\top \\ m_{au} & m_{aa} \end{bmatrix} \quad (4)$$

donde  $m_{aa} \in \mathbb{R}^{m \times m}$ ,  $m_{au} \in \mathbb{R}^{m \times s}$  y  $m_{uu} \in \mathbb{R}^{s \times s}$ . Finalmente, se asume que la función de energía potencial puede dividirse en dos componentes:

$$V(q) = V_a(q_a) + V_u(q_u). \quad (5)$$

### II-B. Transformación de coordenadas

Un paso clave en la propuesta de [1] es el siguiente cambio de coordenadas global:

$$\begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} I_{n \times n} & 0_{n \times n} \\ 0_{n \times n} & T^\top \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix} \quad (6)$$

donde  $T \in \mathbb{R}^{n \times n}$  es una matriz de rango pleno que resulta de expresar la inversa de la matriz de inercia como sigue:

$$M^{-1} = TT^\top. \quad (7)$$

Al ser  $M$  una matriz constante se asegura la existencia de una matriz  $T$  de la forma

$$T = \begin{bmatrix} T_1 & 0_{s \times m} \\ T_2 & T_3 \end{bmatrix} \quad (8)$$

donde  $T_1 \in \mathbb{R}^{s \times s}$ ,  $T_2 \in \mathbb{R}^{m \times s}$  y  $T_3 \in \mathbb{R}^{m \times m}$ . Al derivar respecto al tiempo (6) y tomar en cuenta (1) se obtiene

$$\frac{d}{dt} \begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} 0_{n \times n} & I_{n \times n} \\ -T^\top & 0_{n \times n} \end{bmatrix} \begin{bmatrix} \nabla_q H(q, p) \\ \nabla_p H(q, p) \end{bmatrix} + \begin{bmatrix} 0 \\ T^\top G \end{bmatrix} u. \quad (9)$$

El sistema en malla abierta (9) puede ser reescrito equivalentemente en términos de las nuevas coordenadas definidas en (6) como sigue:

$$\frac{d}{dt} \begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} 0_{n \times n} & T \\ -T^\top & 0_{n \times n} \end{bmatrix} \begin{bmatrix} \nabla_q W(q, p) \\ \nabla_p W(q, p) \end{bmatrix} + \begin{bmatrix} 0 \\ T^\top G \end{bmatrix} u \quad (10)$$

donde se ha introducido una función escalar dada por

$$W(q, p) = \frac{1}{2} p^\top p + V(q) \quad (11)$$

tal que se cumplen las siguientes identidades:

$$\begin{aligned} \nabla_q W(q, p) &= \nabla_q H(q, p), \\ T \nabla_p W(q, p) &= \nabla_p H(q, p). \end{aligned}$$

### II-C. Modelo dinámico en términos de la función del error (vector de estado)

La extensión del control de posición al caso de regulación de velocidad constante mostrada en [1], se basa principalmente en la siguiente definición de función del error:

$$\tilde{q}(t) = \begin{bmatrix} \tilde{q}_u(t) \\ \tilde{q}_a(t) \end{bmatrix} := \begin{bmatrix} q_u(t) - q_u^* \\ q_a(t) - r t \end{bmatrix} \quad (12)$$

y

$$\tilde{p}(t) = \begin{bmatrix} \tilde{p}_u(t) \\ \tilde{p}_a(t) \end{bmatrix} := \begin{bmatrix} p_u(t) \\ p_a(t) - T_3^{-1} r \end{bmatrix} \quad (13)$$

donde  $q_u^* \in \mathbb{R}^s$  y  $r \in \mathbb{R}^m$  son vectores constantes. Usando (10), (12) y (13) se reescribe el modelo dinámico (9) en términos de las variables del error (12)-(13):

$$\frac{d}{dt} \begin{bmatrix} \tilde{q} \\ \tilde{p} \end{bmatrix} = \begin{bmatrix} 0_{n \times n} & T \\ -T^\top & 0_{n \times n} \end{bmatrix} \begin{bmatrix} \nabla_{\tilde{q}} \bar{W} \\ \nabla_{\tilde{p}} \bar{W} \end{bmatrix} + \begin{bmatrix} 0 \\ T^\top G \end{bmatrix} u \quad (14)$$

donde la función escalar  $\bar{W}$  es definida como

$$\bar{W}(\tilde{q} + q^*, \tilde{p}) = \frac{1}{2} \tilde{p}^\top \tilde{p} + V(\tilde{q} + q^*). \quad (15)$$

Enseguida, la siguiente entrada de control

$$u = \nabla_{\tilde{q}_a} V_a(\tilde{q}_a + q_a^*) + \mathbf{v}(\tilde{q}, \tilde{p}) \quad (16)$$

donde  $\mathbf{v}(\tilde{q}, \tilde{p})$  es una entrada de control auxiliar, permite reescribir el modelo dinámico (14) de manera alterna como sigue

$$\frac{d}{dt} \begin{bmatrix} \tilde{q} \\ \tilde{p} \end{bmatrix} = \begin{bmatrix} 0_{n \times n} & T \\ -T^\top & 0_{n \times n} \end{bmatrix} \begin{bmatrix} \nabla_{\tilde{q}} \bar{W}_d \\ \nabla_{\tilde{p}} \bar{W}_d \end{bmatrix} + \begin{bmatrix} 0 \\ T^\top G \end{bmatrix} \mathbf{v}(\tilde{q}, \tilde{p}) \quad (17)$$

con una nueva función escalar  $\bar{W}_d$  dada por

$$\bar{W}_d(\tilde{q}_u + q_u^*, \tilde{p}) = \frac{1}{2} \tilde{p}^\top \tilde{p} + V_u(\tilde{q}_u + q_u^*).$$

Finalmente, desarrollando (17) y luego de tomar en cuenta (8), (12) y (13), se obtiene un nuevo modelo dinámico en términos de la función del error (12)-(13) y la entrada de control auxiliar  $\mathbf{v}(\tilde{q}, \tilde{p})$ :

$$\frac{d}{dt} \begin{bmatrix} \tilde{q}_u \\ \tilde{q}_a \\ \tilde{p}_u \\ \tilde{p}_a \end{bmatrix} = \begin{bmatrix} T_1 \tilde{p}_u \\ T_2 \tilde{p}_u + T_3 \tilde{p}_a \\ -T_1^\top \nabla_{\tilde{q}_u} V_u(\tilde{q}_u + q_u^*) + T_2^\top \mathbf{v}(\tilde{q}, \tilde{p}) \\ T_3^\top \mathbf{v}(\tilde{q}, \tilde{p}) \end{bmatrix}. \quad (18)$$

El problema de control consiste en diseñar un control  $\mathbf{v}(\tilde{q}, \tilde{p})$  tal que el origen  $[\tilde{q}_u \ \tilde{q}_a \ \tilde{p}_u \ \tilde{p}_a]^\top = [0 \ 0 \ 0 \ 0]^\top$  de (18) sea un equilibrio asintóticamente estable, esto es, en forma compacta puede ser expresado como:

$$\lim_{t \rightarrow \infty} \begin{bmatrix} \tilde{q}(t) \\ \tilde{p}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{0}_n \\ \mathbf{0}_n \end{bmatrix}. \quad (19)$$

El cumplimiento de (19) significa que cada coordenada subactuada  $q_{u_i}(t)$  es llevada a una posición deseada  $q_{u_i}^*$ , mientras la coordenada actuada  $q_{a_i}(t)$  sigue una rampa  $r_i t$ , esto es, se desplaza a una velocidad constante  $r_i$ , de acuerdo a (12) y (13), respectivamente.



#### II-D. Control PID

Una solución al problema de control formulado en (19) es un control PID propuesto en [1], el cual está dado por:

$$\mathbf{v}(\tilde{q}, \tilde{\mathbf{p}}) = -k_e^{-1} \left[ K_P \bar{y}_d + K_I \int_0^t \bar{y}_d(s) ds + K_D \dot{\bar{y}}_d \right] \quad (20)$$

el cual es función de la variable:

$$\bar{y}_d = k_a \bar{y}_a + k_u \bar{y}_u \quad (21)$$

en lugar de la forma ortodoxa del control PID expresado en términos de la función del error (12)-(13), siendo

$$\begin{bmatrix} \bar{y}_u \\ \bar{y}_a \end{bmatrix} = \begin{bmatrix} T_2 \tilde{\mathbf{p}}_u \\ T_3 \tilde{\mathbf{p}}_a \end{bmatrix} = \begin{bmatrix} -m_{aa}^{-1} m_{au} \dot{q}_u \\ m_{aa}^{-1} m_{au} \dot{q}_u + \dot{q}_a \end{bmatrix} \quad (22)$$

Una estrategia conveniente que se hizo para reescribir la acción integral y derivativa de (20) en términos de  $\tilde{q}_a$  y  $\tilde{q}_u$ , fue acotar la clase de sistemas subactuados definido en (18) que cumplen las siguientes suposiciones:

**A1** Existe una función escalar  $V_N(\tilde{q}_u)$ , tal que

$$\dot{V}_N(\tilde{q}_u) = \bar{y}_u$$

o equivalentemente,

$$V_N(\tilde{q}_u) = -m_{aa}^{-1} m_{au} \tilde{q}_u. \quad (23)$$

**A2** Existen constantes  $k_e, k_a, k_u \in \mathbb{R}$  seleccionadas por el usuario que aseguran que la matriz  $\bar{K} \in \mathbb{R}^{m \times m}$  sea no singular ( $\det[\bar{K}] \neq 0$ ), la cual está definida por:

$$\bar{K} = k_e I_{m \times m} + k_a K_D T_3 T_3^\top + k_u K_D T_2 T_2^\top. \quad (24)$$

donde  $K_D \in \mathbb{R}^{m \times m}$  es una matriz definida positiva constante.

Tomando en cuenta (18), (21), (22) y las condiciones establecidas en las suposiciones **A1** y **A2**, nos permite describir (20) de la siguiente manera:

$$\mathbf{v}(\tilde{q}, \tilde{\mathbf{p}}) = -\bar{K}^{-1} [K_P \bar{y}_d + K_I [k_a \tilde{q}_a(t) + [k_u - k_a] \bar{V}_N(\tilde{q}_u)] + K_D k_u \bar{S}] \quad (25)$$

donde  $K_I \in \mathbb{R}^{m \times m}$  es una matriz definida positiva, y la matriz  $\bar{S} \in \mathbb{R}^m$  está dada por

$$\bar{S} = -T_2 T_1^\top \nabla_{\tilde{q}_u} V_u(\tilde{q}_u + q_u^*). \quad (26)$$

Note en (25) que la acción integral de (20), se ha sustituido por su equivalente en términos de  $\tilde{q}_a$  y  $\tilde{q}_u$ , esto es,

$$\begin{aligned} \int_0^t \bar{y}_d(s) ds &= \int_0^t [k_a \bar{y}_a(s) + k_u \bar{y}_u(s)] ds \\ &= \int_0^t [k_a (\dot{\tilde{q}}_a(s) - \bar{y}_u(s)) + k_u \bar{y}_u(s)] ds \\ &= k_a \tilde{q}_a(t) + (k_u - k_a) \bar{V}_N(\tilde{q}_u) \end{aligned} \quad (27)$$

#### II-E. Sistema en malla cerrada

Al sustituir la ley de control (25) en (18) se obtiene el sistema en malla cerrada

$$\frac{d}{dt} \begin{bmatrix} \tilde{q}_u \\ \tilde{q}_a \\ \tilde{\mathbf{p}}_u \\ \tilde{\mathbf{p}}_a \end{bmatrix} = \begin{bmatrix} T_1 \tilde{\mathbf{p}}_u \\ T_2 \tilde{\mathbf{p}}_u + T_3 \tilde{\mathbf{p}}_a \\ -T_1^\top \nabla_{\tilde{q}_u} V_u(\tilde{q}_u + q_u^*) - T_2^\top \bar{K}^{-1} [K_P \bar{y}_d + K_I [k_a \tilde{q}_a + (k_u - k_a) \bar{V}_N(\tilde{q}_u)] + K_D k_u \bar{S}] \\ -T_3^\top \bar{K}^{-1} [K_P \bar{y}_d + K_I [k_a \tilde{q}_a + (k_u - k_a) \bar{V}_N(\tilde{q}_u)] + K_D k_u \bar{S}] \end{bmatrix} \quad (28)$$

donde un conjunto de equilibrios de (28) está dado por

$$\mathcal{E} = \left\{ \begin{bmatrix} \tilde{q} \\ \tilde{\mathbf{p}} \end{bmatrix} \in \mathbb{R}^{2n} : \nabla_{\tilde{q}_u} V_u(\tilde{q}_u + q_u^*) = 0, \tilde{\mathbf{p}} = 0, \right. \\ \left. k_a \tilde{q}_a + (k_u - k_a) \bar{V}_N(\tilde{q}_u) = 0 \right\}. \quad (29)$$

Por tanto, para que el origen  $[\tilde{q}_u \ \tilde{q}_a \ \tilde{\mathbf{p}}_u \ \tilde{\mathbf{p}}_a]^\top = [0 \ 0 \ 0 \ 0]^\top$  sea un punto de equilibrio de (28) es necesario que

$$\nabla_{\tilde{q}_u} V_u(q_u^*) = 0.$$

#### II-F. Análisis de estabilidad

A continuación se presenta el análisis de estabilidad del punto de equilibrio en el origen  $[\tilde{q}_u \ \tilde{q}_a \ \tilde{\mathbf{p}}_u \ \tilde{\mathbf{p}}_a]^\top = [0 \ 0 \ 0 \ 0]^\top$ , para probar el cumplimiento del objetivo de control introducido en (19).

El análisis inicia con la función candidata de Lyapunov

$$H_d(\tilde{q} + q^*, \tilde{\mathbf{p}}) = \frac{1}{2} \tilde{\mathbf{p}}^\top M_d^{-1} \tilde{\mathbf{p}} + V_d(\tilde{q} + q^*) \quad (30)$$

donde  $\tilde{q}, \tilde{\mathbf{p}} \in \mathbb{R}^n$ , con  $M_d = M_d^\top > 0$  tal que

$$M_d^{-1} = \begin{bmatrix} \bar{A} & k_a k_u T_2^\top K_D T_3 \\ k_a k_u T_3^\top K_D T_2 & \bar{D} \end{bmatrix} \quad (31)$$

con

$$\bar{A} = k_u^2 T_2^\top K_D T_2 + k_e k_u I_s,$$

$$\bar{D} = k_e k_a I_m + k_a^2 T_3^\top K_D T_3.$$

Además, la función  $V_d : \mathbb{R}^n \rightarrow \mathbb{R}$  está dada por

$$\begin{aligned} \bar{V}_d(\tilde{q} + q^*) &= k_e k_u V_u(\tilde{q}_u + q_u^*) \\ &+ \frac{1}{2} \|k_a \tilde{q}_a + [k_u - k_a] \bar{V}_N(\tilde{q}_u + q_u^*)\|_{K_I}^2 \end{aligned} \quad (32)$$

la cual tiene un mínimo aislado en  $\tilde{q}^* = \mathbf{0}_n$ , esto es,

$$\tilde{q}^* = \arg \min V_d(\tilde{q} + q^*). \quad (33)$$

Algunas manipulaciones algebraicas permiten obtener la derivada temporal de  $H_d$  a lo largo de las trayectorias del sistema en malla cerrada (28):

$$\dot{H}_d \leq -\|\bar{y}_d\|^2. \quad (34)$$

Por tanto, de acuerdo a la teoría de Lyapunov puede concluirse que el origen es un equilibrio estable. Para probar que dicho equilibrio es asintóticamente estable, los autores en [1] utilizan argumentos de detectabilidad, esto es, concluyen que si  $\bar{y}_d$  es

una salida detectable del sistema en malla cerrada (28) entonces el equilibrio  $[\tilde{q}_u \ \tilde{q}_a \ \tilde{p}_u \ \tilde{p}_a]^T = [0 \ 0 \ 0 \ 0]^T$  es asintóticamente estable.

### III. REGULADOR DE VELOCIDAD PARA UN PÉNDULO CON RUEDA INERCIAL SIN FRICCIÓN

Siguiendo el procedimiento descrito en la sección anterior, a continuación se diseña un regulador de velocidad del modelo de un péndulo con rueda inercial sin fricción en las dos articulaciones, mostrado en la Figura 1.

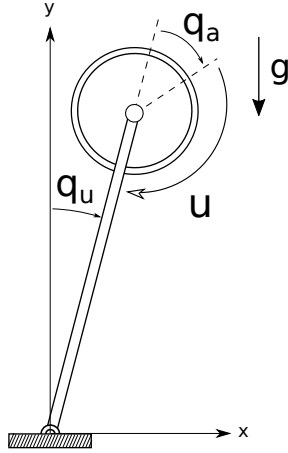


Figura 1. Péndulo con rueda inercial.

La dimensión del sistema es  $n = 2$ ,  $m = 1$  y  $s = 1$ . La matriz de inercia es

$$M = \begin{bmatrix} I_1 + I_2 & I_2 \\ I_2 & I_2 \end{bmatrix} \quad (35)$$

donde  $I_1$  e  $I_2$  son las inercia del péndulo y la rueda respectivamente. La función de energía potencial viene dada por

$$V(q) = V_u(q_u) \\ V_u(q_u) = a \cos(q_u) \quad (36)$$

donde  $a = (m_1 l_1 + m_2 l_2)g$ . Siendo  $m_1$  y  $m_2$  las masas del péndulo y la rueda, respectivamente,  $l_1$  y  $l_2$  las distancias a sus centros de masa y  $g$  la constante de aceleración de la gravedad. La matriz de distribución de entradas de control resulta

$$G = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (37)$$

y la matriz de transformación  $T$  definida en (8) es constante y está dada por

$$T = \begin{bmatrix} 1/\sqrt{I_1} & 0 \\ -1/\sqrt{I_1} & 1/\sqrt{I_2} \end{bmatrix}. \quad (38)$$

Por otro lado, las funciones del sistema en malla cerrada (28) son

$$M_d^{-1} = \begin{bmatrix} k_e k_u + K_D k_u^2 / I_1 & -k_a k_u K_D / \sqrt{I_1 I_2} \\ -k_a k_u K_D / \sqrt{I_1 I_2} & k_e k_a + k_a^2 K_D / I_2 \end{bmatrix} \quad (39)$$

mientras que

$$V_d(\tilde{q}) = k_u k_e a \cos(\tilde{q}_u) + \frac{1}{2} K_I [k_a \tilde{q}_a - (k_u - k_a) \tilde{q}_u]^2. \quad (40)$$

Aplicando el criterio de la segunda derivada

$$\nabla^2 V_d(0) = \begin{bmatrix} \frac{\partial^2 V_d}{\partial \tilde{q}_a^2} & \frac{\partial^2 V_d}{\partial \tilde{q}_a \partial \tilde{q}_u} \\ \frac{\partial^2 V_d}{\partial \tilde{q}_a \partial \tilde{q}_u} & K_I k_a^2 \end{bmatrix}$$

con

$$\frac{\partial^2 V_d}{\partial \tilde{q}_a^2} = -k_u k_e a + K_I (k_u - k_a)^2, \\ \frac{\partial^2 V_d}{\partial \tilde{q}_a \partial \tilde{q}_u} = -K_I k_a (k_u - k_a),$$

puede verificarse que  $V_d(\tilde{q})$  tiene un mínimo en cero ( $\tilde{q} = 0$ ) si  $k_u k_e < 0$ . Finalmente, la ley de control (16) queda

$$u = \mathbf{v}(\tilde{q}, \tilde{\mathbf{p}})$$

al ser  $V(q)$  independiente de  $q_a$  (por tanto,  $\nabla_{\tilde{q}_a} V_a(\tilde{q}_a + q_a^*) = 0$ ), donde:

$$\mathbf{v}(\tilde{q}, \tilde{\mathbf{p}}) = -\bar{K}^{-1} [K_P \bar{y}_d + K_I [k_a \tilde{q}_a - (k_u - k_a) \tilde{q}_u] \\ - \frac{a}{I_1} k_u K_D \sin(\tilde{q}_u)] \quad (41)$$

siendo

$$\bar{y}_d = \frac{k_u}{I_1} [p_a - p_u] + \frac{k_a}{I_2} [p_a - I_2 r], \quad (42)$$

$$\bar{K} = k_e + K_D \left[ \frac{k_a}{I_2} + \frac{k_u}{I_1} \right]. \quad (43)$$

### IV. SIMULACIONES

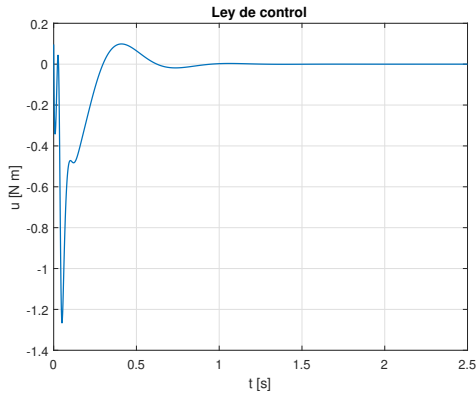
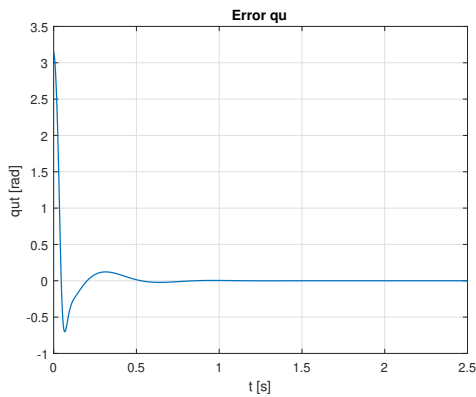
En esta sección se presentan los resultados numéricos de simulación para validar el desempeño del controlador (41). Los parámetros del modelo del péndulo con rueda inercial (Figura 1) de la marca Quanser y las ganancias del controlador (41) son mostradas en las Tablas I y II, respectivamente.

Tabla I  
PARÁMETROS DEL PÉNDULO CON RUEDA INERCIAL QUANSER

Parámetro	Descripción	Valor
$m_1$	Masa del péndulo	0.030 kg
$m_2$	Masa de la rueda	0.216 kg
$I_1$	Inercia del péndulo	$17.1 \times 10^{-5}$ kg m <sup>2</sup>
$I_2$	Inercia de la rueda	$2.5 \times 10^{-5}$ kg m <sup>2</sup>
$l_1$	Distancia al centro de masa del péndulo	0.058 m
$l_2$	Distancia al centro de masa de la rueda	0.120 m
$g$	Constante de aceleración de la gravedad	9.81 m/s <sup>2</sup>

Tabla II  
GANANCIAS DEL CONTROLADOR

Ganancia	Valor
$k_e$	1
$k_u$	-1
$k_a$	0.05
$K_I$	0.10
$K_D$	$3.42 \times 10^{-4}$
$K_P$	0.01


 Figura 2. Par requerido por la entrada de control  $u(t)$ .

 Figura 3. Evolución temporal del error en la articulación subactuada  $\tilde{q}_u(t)$ .

La condición inicial fue:

$$[q_u, q_a, p_u, p_a]^T = [\pi, 0, 0, 0]^T$$

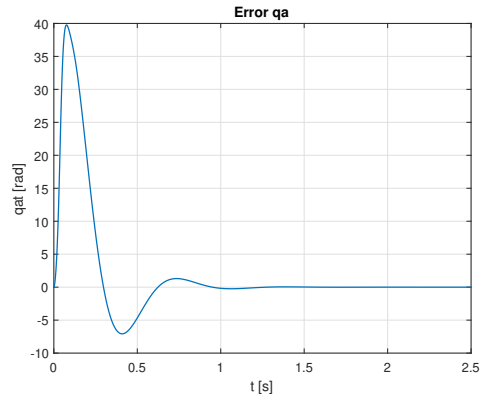
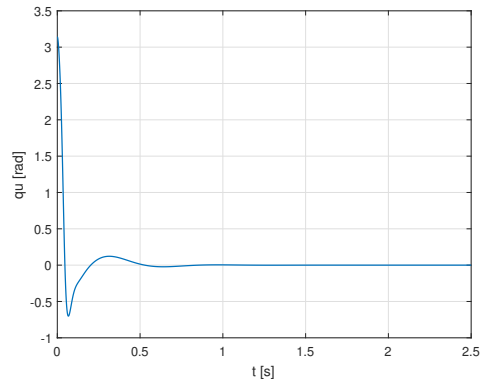
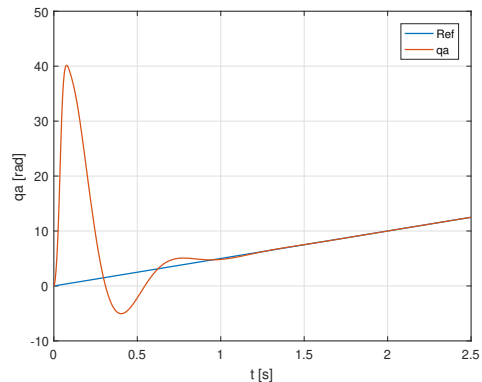
y la configuración deseada fue establecida para:

$$\begin{bmatrix} q_u^* \\ q_a^*(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 5t \end{bmatrix}.$$

En las Figuras 2, 3 y 4 se muestra la evolución temporal de la entrada de control y los errores de posición de cada articulación. Visualmente se aprecia cómo los errores de posición tienden a cero (ver figuras 3 y 4). Esto significa que se ha estabilizado el péndulo en su configuración invertida, mientras que la rueda sigue la referencia dada por la función rampa  $5t$ , lo cual significa de manera equivalente que la rueda gira a una velocidad constante de 5 [rad/s] (ver figuras 5 y 6).

## V. CONCLUSIONES

Se ha presentado un control PID para la regulación de velocidad un modelo de péndulo con rueda inercial sin fricción. El control PID corresponde al propuesto por [1], el cual aplicado al péndulo con rueda inercial cumple con las condiciones establecidas para el cumplimiento del objetivo de control (19).


 Figura 4. Evolución temporal del error en la articulación actuada  $\tilde{q}_a(t)$ .

 Figura 5. Evolución temporal de la articulación subactuada  $q_u(t)$ .

 Figura 6. Evolución temporal de la articulación actuada  $q_a(t)$ .

## AGRADECIMIENTOS

Este trabajo fue financiado parcialmente por el TecNM, contrato 5925.19-P.

## REFERENCIAS

- [1] J. G. Romero, R. Ortega, and A. Donaire, "Energy Shaping of Mechanical Systems via PID Control and Extension to Constant Speed Tracking," *IEEE Transactions on Automatic Control*, vol. 61, pp. 3551–3556, 2016.
- [2] R. Ortega, M. Spong, F. Gomez-Estern and G. Blankenstein, "Stabilization of a class of underactuated mechanical systems via interconnection and damping assignment," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1213–1233, 2002.

- [3] A. M. Bloch, N. Leonard and J. Marsden, "Controlled Lagrangian and the stabilization of mechanical systems I: The first matching theorem," *IEEE Transactions on Automatic Control*, vol. 45, no. 12, pp. 2253–2270, 2000.
- [4] A. Donaire, R. Mehra, R. Ortega, S. Satpute, J. G. Romero, F. Kazi and N. M. Singh, "Shaping the energy of mechanical systems without solving partial differential equations," *IEEE Transactions on Automatic Control*, vol. 61, no. 4, pp. 1051–1056, 2016.

# Control y simulación de un dispositivo de asistencia activo.

Antonio Martínez Chavelas  
Maestría en sistemas  
computacionales.  
Instituto Tecnológico de La Paz.  
La Paz, Baja California Sur.  
clefftony@gmail.com

Israel Santillán  
División de Estudios de Posgrado e  
Investigación  
Instituto Tecnológico de La Paz  
La Paz, Baja California Sur.  
israel.santillan@itlp.edu.mx

Jesús Sandoval Galarza  
División de Estudios de Posgrado e  
Investigación  
Instituto Tecnológico de La Paz  
La Paz, Baja California Sur.  
jsandoval@itlp.edu.mx

José Luis Gómez Torres  
División de Estudios de Posgrado e  
Investigación  
Instituto Tecnológico de La Paz  
La Paz, Baja California Sur.  
jgomez@itlp.edu.mx

**Abstract** — El presente trabajo consiste en el diseño, modelado dinámico, control y simulación de una prótesis activa de rodilla y tobillo. Los eslabones se proponen con dimensiones semejantes a las de un ser humano, esto con el fin de que pueda ser utilizado por un ser humano. El sistema cuenta con 2 grados de libertad, y se considera que el sistema sólo se mueve en el eje sagital. La marcha bípeda se dividirá en 4 etapas, cada etapa contará con su modelo dinámico y su ley de control. Al finalizar, se realizará una simulación por computadora para corroborar la viabilidad de la propuesta.

**Keywords** — *Prótesis, Robótica, Control, Dinámica.*

## I. INTRODUCCIÓN

Se le conoce como amputación al corte o separación de una extremidad del cuerpo mediante cirugía. Esta técnica se utiliza como una medida quirúrgica para controlar la propagación de una enfermedad en la extremidad afectada o para separar extremidades con heridas graves en las que no se puede salvar el miembro o donde los intentos de salvar el miembro han sido fallidos (comúnmente accidentes de tránsito). El resto de la extremidad residual después de la amputación aún pueden moverse y realizar esfuerzos.

La caminata realizada por personas con amputaciones en las extremidades inferiores es lenta, menos estable y requiere más energía, en comparación a la caminata realizada por una persona con las extremidades completas [1]. Además, los individuos con extremidades amputadas se tropiezan con mayor frecuencia y les resulta más difícil subir rampas, colinas y escaleras [2]. Éstas complicaciones se deben principalmente al uso de prótesis pasivas, las cuales no responden activamente a las perturbaciones ni contribuyen positivamente al realizar un trabajo, como lo hace un músculo natural. Actualmente existen prótesis activas modernas que mejoran significativamente la movilidad, ya que cuentan con actuadores (los cuales pueden ser actuadores rotacionales o lineales) que ayudan al usuario a realizar actividades sencillas.

El método más común para controlar estos sistemas consiste en dividir el ciclo de la marcha en diferentes etapas, cada una con su propia ley de control. Dichas leyes deben ser diseñadas para realizar una tarea en específico de forma eficiente. Así mismo, los parámetros arbitrarios de cada ley de control deben ser seleccionados cuidadosamente en función del objetivo a realizar y de las características del usuario, esto para realizar de forma eficiente el objetivo solicitado [3].

Esta metodología depende fuertemente en que el reconocimiento de la etapa actual durante la marcha sea preciso, ya que una ley de control que se active en el momento incorrecto puede ocasionar que el usuario tropiece [4].

En éste trabajo se propone un diseño de una prótesis de rodilla y tobillo para individuos con amputación transfemoral. Se propone que el diseño debe ser lo más simple posible, considerando materiales de fácil obtención y fácil fabricación. Así mismo, el sistema debe ser de fácil colocación y no invasivo para el usuario. La prótesis contará con actuadores del tipo rotacional (motores de CD) en la articulación de la rodilla y tobillo.

Se describe un modelo dinámico basándose en el método de Euler-Lagrange, el cual cuenta con la misma cantidad de objetivos de control como etapas en la que se dividió la marcha bípeda. Las trayectorias propuestas en cada etapa de la marcha deben imitar lo mejor posible la trayectoria desarrollada por un ser humano. Posteriormente, se realizará una simulación por computadora para corroborar la viabilidad de la propuesta.

## II. DISEÑO DE LA PRÓTESIS

La prótesis se diseña basándose en los movimientos que un ser humano es capaz de realizar. Los grados de libertad indican la cantidad de movimientos que pueden realizar cada junta, así como los ejes en los que podrán actuar las juntas mencionadas. El ser humano solo puede realizar un movimiento en la rodilla, por lo cual tiene un solo grado de libertad. El tobillo consta de tres grados de libertad [5], sin embargo, en el presente trabajo se utilizará sólo un grado de libertad. Por lo tanto, se concluye que el sistema tendrá 2 grados de libertad en total.

En [6] se recopilaron dimensiones físicas de diferentes personas, con diferentes edades y de diferentes países; también se muestra el promedio de los datos obtenidos. Dicha información es la que se utilizará para escoger la dimensión adecuada de la prótesis.

Descripción	Longitud [cm]
Altura de tibia	36.3
Altura de pie	11.5
Altura del piso a rodilla	47.8
Longitud de tibia	10
Longitud de pie	23
Anchura de tibia	5
Anchura de pie	9

Tabla 1: Dimensiones de la prótesis.

En la Tabla 1 se muestran las dimensiones de los elementos de la prótesis, mientras que en la figura 1 (Inspirado por [7]) se puede apreciar el sistema propuesto. Durante el diseño se consideraron las siguientes suposiciones:

**Suposiciones de Diseño (SD)**

- SD1.-** Los eslabones son de geometrías sencillas y no presentan deformaciones.
- SD2.-** La densidad de los eslabones es constante.
- SD3.-** El sistema cuenta con motores de corriente directa (CD) para realizar los movimientos.
- SD4.-** Los torques aplicados se desarrollan en las articulaciones de manera directa.
- SD5.-** El material propuesto es una aleación de aluminio.



Fig. 1. Propuesta de la prótesis.

III. ESTUDIO DE LA MARCHA

La marcha bípeda se puede dividir en dos fases principales: La fase de soporte y la fase de balanceo. Sin embargo en [8] se demostró que cada fase puede subdividirse en dos etapas, haciendo que el control de la prótesis sea más efectivo, como se muestra en la figura 2 [9].

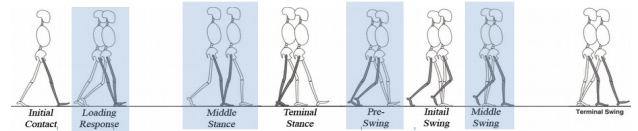


Fig. 2. Etapas de la marcha bípeda.

La etapa 1 comienza con la etapa de soporte, cuando el talón hace contacto con el suelo mientras que la rodilla se flexiona para absorber el impacto y el tobillo gira hasta quedar plano; la etapa 2 comienza cuando el talón empieza a levantarse del piso, esto quiere decir, la punta del pie funciona como pivote; la etapa 3 comienza cuando el pie deja de hacer contacto con el piso, es cuando empieza la fase de balanceo y el pie se desplaza hacia adelante; la etapa 4 comienza cuando el pie de balanceo empieza a descender y prepara el talón para hacer contacto con el piso, regresando así a la etapa 1.

El sistema se considera como una cadena cinemática abierta. Toda la trayectoria sucede en el plano sagital y, además, se considera que el sistema está alternando entre la fase de soporte y de balanceo constantemente. Para el desarrollo del modelo dinámico se consideraron las siguientes suposiciones:

**Suposiciones del Sistema (SS)**

- SS1.-** La prótesis cuenta con dos eslabones (pie y tibia) en una cadena cinemática abierta.
- SS2.-** El sistema es estrictamente planar, solamente se puede mover en el plano sagital.
- SS3.-** El sistema tiene pie por lo tanto tiene actuadores entre el sistema y el piso, a diferencia de otros trabajos ([8], [10]) en donde el sistema termina en punta y no existe actuador con respecto al suelo, dichos sistemas son subactuados.

**SS4.-** La fricción interna de los actuadores se desprecia.

Así mismo, aparte de los objetivos de control, se deben definir condiciones que el sistema debe cumplir durante la marcha, éstas condiciones son:

**Suposiciones de la Marcha (SM)**

- SM1.-** Durante la marcha, la prótesis alterna su función de balanceo a soporte de forma cíclica y continua, sin perturbaciones.
- SM2.-** La fuerza de reacción del piso al momento del impacto no es lo suficientemente grande para perturbar al sistema, por lo cuál se desprecia dicha fuerza.
- SM3.-** El pie no puede deslizarse por el suelo cuando



hacen contacto.

**SM4.-** La marcha se realiza de izquierda a derecha. La pierna de balanceo debe empezar detrás de la pierna de soporte, y el impacto debe suceder estrictamente delante de la pierna de soporte.

**SM5.-** El sistema se desplaza por una superficie plana sin inclinación o deformaciones.

Éstas suposiciones nos permiten limitar el alcance del modelo dinámico y entender cuál será el objetivo de control.

#### IV. MODELADO DINÁMICO

El modelado dinámico consiste en una ecuación que relaciona las fuerzas que se ejercen en el sistema con los efectos causados por la posición, velocidad y aceleración del sistema [11].

$$\tau = f(q, \dot{q}, \ddot{q}) \quad (1)$$

La ecuación (1) nos permite conocer los torques que se necesitan aplicar en las articulaciones para que los eslabones del bípodo desarrollen una posición, velocidad y aceleración deseada. El método a utilizar para desarrollar el modelo dinámico es el método de Euler-LaGrange, el cual se basan en el concepto de la energías.

A partir de las energías del robot se construye el lagrangiano  $L(q, \dot{q})$ , el cual se representa mediante la diferencia entre su energía cinética  $K$  y su energía potencial  $P$  [12]:

$$L(q, \dot{q}) = K(q, \dot{q}) - P(q) \quad (2)$$

Con esta ecuación de Lagrange (2), se puede llegar a una forma compacta del modelo dinámico del bípodo de  $n$  grados de libertad.

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = u \quad (3)$$

En donde:

$$M(q) = \frac{\delta}{\delta \dot{q}} \frac{\delta K(q, \dot{q})}{\delta \dot{q}} \quad (4)$$

$$C(q, \dot{q})\dot{q} = \dot{M}(q)\dot{q} - \frac{1}{2} \frac{\delta}{\delta q} [\dot{q}^T M(q) \dot{q}] \quad (5)$$

$$G(q) = \frac{\delta P(q)}{\delta q} \quad (6)$$

La ecuación (3) es el modelo dinámico compacto, (4) es la matriz de inercia/masa, (5) es el vector de fuerzas centrífugas y de Coriolis, (6) es el vector de gravedad y  $u$  es la acción de control.

Obtener la energía potencial consiste en sustituir los valores de masa de todos los eslabones, el vector de gravedad y los vectores del centro de masa (estos vectores nos indican donde se encuentran los centros de masa con respecto al origen) en la ecuación (7):

$$P(q) = \sum_{i=1}^n m_i g r_{ci} \quad (7)$$

Para obtener la energía cinética del sistema es necesario sumar la energía cinética lineal  $K_l$  y rotacional  $K_r$  del sistema (8).

$$K(q, \dot{q}) = K_l + K_r \quad (8)$$

En donde:

$$K_l = \sum_{i=1}^n \frac{1}{2} m_i v_i^2 \quad (9)$$

$$K_r = \sum_{i=1}^n \frac{1}{2} I_i \omega_i^2 \quad (10)$$

En donde  $m$  representa la masa de los eslabones, los vectores de velocidad lineal con respecto al origen se obtienen al derivar los vectores de posición del centro de masa con respecto al origen,  $\omega$  representan a los vectores asociados a las matrices anti-simétricas de velocidad angular,  $I$  representan a las matrices de momentos de inercia de las masas de los eslabones.

#### V. DISEÑO DEL CONTROLADOR

En el presente trabajo se utilizará un controlador basado en impedancias para generar los torques en las uniones. Éste acercamiento permite al usuario tener interacción con la prótesis al simular el comportamiento de la prótesis como si fuera una extremidad natural, al mismo tiempo que se genera un comportamiento estable y predecible [13]. En esencia, el comportamiento de la rodilla y del tobillo se obtiene mediante el cambio de objetivo de control, el cuál se realiza al cambiar de etapa de la marcha. El controlador actúa como si existiera un resorte y un disipador en la unión, de tal forma que se puede garantizar que la prótesis sea estable dentro de cada etapa.

Basándose en la definición del controlador por impedancias [14], el torque requerido en cada junta durante cada etapa de la marcha puede ser representado por una serie de funciones de impedancias pasivas. Basándose en los datos

obtenidos en [15], el torque en las uniones puede ser describirse mediante funciones que dependen de la posición y velocidad. Dicha función se describe en la ecuación (11)

$$u = k_1(q - q_d) + k_2(q - q_d)^3 + b\dot{q} \quad (11)$$

La ecuación (11) describe una función con una combinación de un término lineal de rigidez  $k_1$ , un término cúbico de rigidez  $k_2$  y un término lineal de disipación  $b$ , siendo  $q$  el vector de posiciones articulares del sistema,  $q_d$  el vector de posiciones que se desea que el sistema adopte,  $\dot{q}$  el vector de velocidades del sistema y  $u$  es el torque de la unión. Si se definen  $k_1$ ,  $k_2$  y  $b$  con valores positivos, entonces la junta convergerá a un equilibrio estable en  $q = q_d$  con  $\dot{q} = 0$  en cada una de las etapas de la marcha bípoda.

## VI. TRABAJO FUTURO

Realizar la simulación de cada etapa de la marcha bípoda con su respectiva ley de control aplicado a su respectivo modelo dinámico.

## REFERENCES

- [1] J. Johansson, D. Sherrill, P. Riley, P. Bonato, and H. Herr, "A clinical comparison of variable-damping and mechanically passive prosthetic knee devices," in *Amer. J. Phys. Med. Rehab.*, vol. 84, no. 8, pp. 563–575, 2005.
- [2] D. G. Smith, J. W. Michael, and J. H. Bowker, Eds., *Atlas of Amputations and Limb Deficiencies: Surgical, Prosthetic, and Rehabilitation Principles*. Rosemont, IL, USA: Amer. Acad. Orthopaedic Surgeons, 2004.
- [3] A. Simon, N. Fey, S. Finucane, R. Lipschutz, and L. Hargrove, "Strategies to reduce the configuration time for a powered knee and ankle prosthesis across multiple ambulation modes," presented at the IEEE Int. Conf. Rehab. Robotics, Seattle, WA, USA, 2013.
- [4] R. D. Gregg, T. Lenzi, L. J. Hargrove, and J. W. Sensinger, "Virtual constraint control of a powered prosthetic leg: From simulation to experiments with transfemoral amputees," *IEEE Trans. Robot.*, vol. 30, no. 6, pp. 1455–1471, 2014.
- [5] F. Hernández Stengele. "Diseño y construcción de prototipo neumático de prótesis de pierna humana". Universidad de las Américas Puebla, 2008. url: [http://catarina.udlap.mx/u\\_dl\\_a/tales/documentos/lep/hernandez\\_s\\_f/](http://catarina.udlap.mx/u_dl_a/tales/documentos/lep/hernandez_s_f/) (visitado 24-05-2016).
- [6] R.Á. Chaurand y col. *Dimensiones antropométricas de población latinoamericana: México, Cuba, Colombia, Chile*. Colección Modulo. Universidad de Guadalajara, Centro Universitario de Arte, Arquitectura y Diseño, División de Tecnología y Procesos, Departamento de Producción y Desarrollo, Centro de Investigaciones en Ergonomía, 2007. isbn: 978-970-27-1193-3.
- [7] F. Sup, H. Varol, and M. Goldfarb, "Upslope walking with a powered knee and ankle prosthesis: Initial results with an amputee subject," in *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 19, no. 1, pp. 71–78, 2011.
- [8] F. Sup, A. Bohara, and M. Goldfarb, "Design and control of a powered transfemoral prosthesis," in *Int. J. Robot. Res.*, vol. 27, no. 2, pp. 263–273, 2008.
- [9] Umberger, B. R. (2010). Stance and swing phase costs in human walking. *Journal of the Royal Society Interface*, 7(50), 1329–1340.
- [10] E. R. Westervelt, J. W. Grizzle, and D. E. Koditschek. Hybrid zero dynamics of planar biped walkers. *IEEE Transactions on Automatic Control*, 48(1):42–56, January 2003.
- [11] R. Kelly y V. Santibáñez. *Control de movimiento de robots manipuladores. Automática y Robótica*. Pearson Educación, 2003. isbn: 9788420538310.
- [12] Reza N Jazar. *Theory of applied robotics: kinematics, dynamics, and control*. Segunda edición. Springer Science & Business Media, 2010. isbn: 978-1-4419-1749-2.
- [13] Mochon, S. and McMahon, T. A. (1980). *Ballistic walking*. *Journal of Biomechanics*, 13(1): 49–57.
- [14] Hogan, N. (1985). *Impedance control: an approach to manipulation: Part 1—theory, Part 2—implementation, and Part 3—applications*. *ASME Journal of Dynamic Systems, Measurement and Control*, 107: 1–24.
- [15] Winter, D. A. (1991). *The Biomechanics and Motor Control of Human Gait: Normal, Elderly and Pathological*, 2nd edn. Waterloo, ON, University of Waterloo Press.

# Identificación de ataques en redes de cómputo utilizando redes neuronales artificiales

Javier Alberto Carmona Troyo  
Departamento de sistemas y Computación  
Instituto Tecnológico de La Paz  
La Paz B.C.S., México  
Email: javier.carmona@itlp.edu.mx

José Tadeo Rodríguez Solano  
Departamento de sistemas y Computación  
Instituto Tecnológico de La Paz  
La Paz B.C.S., México  
Email: tadeo@itlp.edu.mx

Alberto Ibarra Berber  
Departamento de sistemas y Computación  
Instituto Tecnológico de La Paz  
La Paz B.C.S., México  
Email: berber@itlp.edu.mx

Luis Armando Cárdenas Florido  
División de Estudios de Posgrado E Investigación  
Instituto Tecnológico de La Paz  
La Paz B.C.S., México  
Email: armando.cardenas@itlp.edu.mx

Armando Yuen Coria  
Departamento de sistemas y Computación  
Instituto Tecnológico de La Paz  
La Paz B.C.S., México  
Email: yuen@itlp.edu.mx

**Resumen**—El presente trabajo busca determinar un método rápido, económico y eficiente para la detección de intrusos en redes de área local. Esta forma de detección que se presenta está basada en técnicas de inteligencia artificial, específicamente en redes neuronales Artificiales.

La investigación se dividió en cuatro partes. Primero se analizaron diferentes contextos de problemáticas a las que nos enfrentamos en cuanto a seguridad informática. Posteriormente se identificaron los elementos y variables principales que interviene en el tráfico de redes, así como la aplicación de diferentes técnicas como las redes neuronales artificiales.

Posteriormente se presenta una propuesta de solución del problema y el planteamiento de los procedimientos desarrollados. Se concluye con los resultados obtenidos de las observaciones o pruebas realizadas.

**Keywords**—Network Security; Intruder Detection; patterns; artificial intelligence; fuzzy logic; RNA; IDS.

## I. INTRODUCCIÓN

Hoy en día, se visualiza la necesidad de desarrollar e implementar nuevos métodos y controles mas eficientes para mantener la información más segura y mas cuando esta es transmitida a través de las tecnologías de información. Esto es, debido al avance de la tecnología en el campo de las comunicaciones y la importancia de mantener la confidencialidad, la integridad y la disponibilidad de los datos, siendo estos los activos más importantes de cualquier organización. Estos controles de seguridad se plantean en redes de área local

interna o en las redes perimetrales que dan salida a dicha información. Al aplicar los controles de seguridad necesarios, la empresa, institución o compañía, se vería beneficiada ya que sus activos permanecerían seguros, evitando así, grandes pérdidas económicas o de imagen. Los piratas informáticos, al mismo tiempo en que van incorporando nuevas tecnologías, van adquiriendo más conocimiento acerca de las formas o métodos para llevar a cabo un vector de ataque sobre alguna vulnerabilidad que se tenga en las redes o aplicaciones de las empresas. Estos vectores de ataque que llevan a cabo los cibercriminales pueden ocasionar la pérdida de información, el secuestro digital o atentar contra la privacidad de los usuarios de la información o datos de la empresa o institución misma.

## II. MARCO TEÓRICO

### II-A. Seguridad de la información

La seguridad de la información engloba un conjunto de técnicas y medidas para controlar todos los datos que se manejan dentro de una institución y asegurar que no salgan de ese sistema establecido por la empresa. Principalmente este tipo de sistemas se basan en las nuevas tecnologías, por tanto la seguridad de la información resguardará los datos que están disponibles en dicho sistema y a los que solo tendrán acceso usuarios autorizados. Por otro lado, tampoco se podrán hacer modificaciones en la información a no ser que sea de la mano de los usuarios que tengan los permisos correspondientes. La seguridad de la información debe responder a tres cualidades principales:

- Crítica

- Valiosa
- Sensible

Por un lado debe ser crítica, ya que es una pieza fundamental para que la empresa pueda llevar a cabo sus operaciones sin asumir demasiados riesgos. También debe ser valiosa, puesto que los datos que se manejan son esenciales para el devenir del negocio y finalmente tiene que ser sensible, ya que al sistema solo podrán acceder las personas que estén debidamente autorizadas. Además, también hay que tener en cuenta que la seguridad de la información debe hacer frente a los riesgos, analizarlos, prevenirlos y encontrar soluciones rápidas para eliminarlos si se diera el caso [1].

#### II-B. Seguridad informática

La seguridad informática es un conjunto de normas, procedimientos y herramientas, que tienen como objetivo garantizar la disponibilidad, integridad, confidencialidad y buen uso de la información. La seguridad en un ambiente de red es la habilidad de identificar y eliminar vulnerabilidades. Estas normas incluyen horarios de funcionamiento, restricciones a ciertos lugares, autorizaciones, denegaciones, perfiles de usuario, planes de emergencia, protocolos y todo lo necesario que permita un buen nivel de seguridad informática minimizando el impacto en el desempeño de los trabajadores y de la organización en general y como principal contribuyente al uso de programas realizados por programadores [2].

#### II-C. Sistemas IDS (Sistema identificador de intrusos)

Como herramienta para protección de los intrusos en redes de área local o en infraestructuras de red, se crearon los IDS o sistemas identificadores de intrusos [3]. Estos sistemas, normalmente son colocados en la zona perimetral de un sistema de red. Estos sistemas detectan las posibles amenazas que pueden provocar un ataque a la infraestructura, detectando e identificando intrusos en la red.

Desde que el concepto de Detección de Intrusos fue propuesto por Anderson en 1980 [4]., muchas técnicas para su implementación han sido reportadas y estudiadas. Las técnicas para los sistemas de Detección de Intrusos (o intrusiones) IDS, se clasifican generalmente en dos categorías, la detección de mal uso y la detección de anomalías [5].

Los sistemas de detección de intrusos se pueden clasificar de diversas maneras, según diferentes parámetros que se apliquen; para esta investigación se tomó la clasificación propuesta por Iren Lorenzo Fonseca et al [6].

#### II-D. Inteligencia artificial

Elaine Rich [7] menciona que la inteligencia artificial, estudia como lograr que las máquinas realicen tareas que, por el momento, son realizadas mejor por los seres humanos.

Así mismo, Margaret Rouse [8] define a la inteligencia artificial como la simulación de procesos de inteligencia humana por parte de máquinas, especialmente sistemas informáticos. Estos procesos incluyen el aprendizaje (la adquisición

de información y reglas para el uso de la información), el razonamiento (usando las reglas para llegar a conclusiones aproximadas o definitivas) y la auto corrección [8].

#### II-E. Técnica de análisis de componentes principales

El análisis de componentes principales o PCA, es una técnica estadística que permite seleccionar información en un conjunto de  $n$  variables de interés en  $m$  nuevas variables independientes. En relación con las variables seleccionadas, cada una debe explicar una parte específica de la información y mediante combinación lineal, otorgan la posibilidad de resumir dicha información en pocos componentes que reducen la dimensión del problema [9].

### III. JUSTIFICACIÓN

Este tipo de investigación resulta significativa ya que, en el ámbito de las redes de área local, no se tienen los mecanismos de seguridad implementados en su máxima expresión, esto es debido al alto costo que se tiene para la implementación de Sistemas de Identificación de Intrusos o IDS (Intrusion Detection Systems) capaces de detectar en tiempo real una amenaza. Por ello, se propone un mecanismo utilizando técnicas de inteligencia artificial para la detección de intrusos, el cual resulte de bajo costo y pueda escalar hacia un mecanismo mucho más potente y con las mismas herramientas. Estas técnicas de inteligencia artificial están basadas en redes neuronal artificiales las cuales serán capaces de detectar un ataque al protocolo TCP como por ejemplo un mapeo de puertos NMAP.

Existen diversos IDS disponibles actualmente, algunos comerciales y otros de licencia libre. Uno de los más utilizados es SNORT, el cuál trabaja en base a reglas bien definidas. En este caso, la novedad será que, aplicando las técnicas propuestas de inteligencia artificial, permitirá no solo detectar un NMAP, si no también detecten cualquier invasión o ataque persistente a la capa de transporte.

En el caso de muchas instituciones federales de educación superior, no cuentan con mecanismos de seguridad apropiados para la detección y paro de ataques. Es por ello, que nace la idea de generar por medio de inteligencia artificial, un IDS capaz de detectar ataques a la infraestructura de red de área local y a su vez, sea funcional.

Este tipo de trabajos es importante, ya que no hay ningún sistema que pueda ser seguro en cuanto a ataques cibernéticos. Muchas veces tenemos que los sitios de red que manejamos tienen vulnerabilidades las cuales pueden ser aprovechadas por delinquentes cibernéticos.

México es uno de los principales países de Latinoamérica que frecuentemente es atacado en su infraestructura tecnológica, es por ello que debemos de hacer uso de la tecnológica de computación para poder detectar a tiempo a los posibles intrusos de nuestros sistemas para así poder poner un remedio

o solución a esto.

Esto impactará de sobre manera a las organizaciones o empresas ya que la seguridad de sus activos se podrá mantener adecuadamente, evitando grandes pérdidas de los mismos o pérdidas económicas las cuales podrán llevar a las instituciones a un estado de poca credibilidad.

#### IV. METODOLOGÍA

Para llevar a cabo el análisis de información que es transmitida a través de una red, se tomaron tramas de datos las cuales fueron capturadas durante una petición simple y otras capturadas durante un ataque NMAP para después obtener las características importantes de ellas, las cuales fueron procesadas por medio de redes neuronales artificiales. Esta técnica ayuda a determinar si se está llevando a cabo un ataque a la infraestructura de red.

El modelo utilizado para llevar a cabo el procesamiento de los datos obtenidos de las tramas de red, sigue los siguientes pasos:

- Obtención de datos: Consistió en tomar información captada por un sniffer de red, se utilizaron las herramientas Wireshark y Tcpdump.
- Datos a seleccionar: Se seleccionaron tramas TCP la cuales eran por el tipo de ataques, el punto de partida para siguientes pasos de análisis.
- Extracción de características deseables: Se obtuvieron de las tramas información relevante cuyas características representan a las de un ataque.
- Normalizar y reducir las dimensiones: Se llevó a cabo un preprocesamiento de información sobre el conjunto de datos el cual sirvió para alimentar un conjunto de datos.
- Clasificación: Se llevó a cabo una clasificación sobre un conjunto de datos tomados en el paso anterior.
- Conjunto de datos (DataSet): Para esta etapa, se tomó una parte de la base de datos para entrenar la red y otra parte para validar que esté funcionando correctamente.
- Clasificador: Dependiendo las reglas suministradas, se determina si existe un ataque o no.

#### V. ADQUISICIÓN DE DATOS

Para este trabajo fue necesario obtener un conjunto de datos sobre el cual se pudiera trabajar, y que permitiera entrenar el algoritmo de aprendizaje. Para esta tarea se utilizaron los programas de distribución gratuita Wireshark para la captura de tramas, VMware para la virtualización, NMAP para escanear puertos, hping3 (bajo plataforma Linux) como generador de ataques de denegación de servicio (DoS). Con estas dos últimas herramientas se simularon ataques en una red compuesta por máquinas virtuales Linux (Kali Linux 2.0), Windows XP SP2 y Windows 7.

Según el modelo propuesto, es necesario generar un conjunto de datos a partir de las señales de ataque capturados con Wireshark. Este conjunto de datos, fue exportado hacia archivos de arreglos en C los cuales posteriormente fueron procesados para ser analizados.

Los datos fueron tomados de una verificación de puertos con NMAP, de un ataque o denegación de servicios con ping y otro ataque con hping2, todos utilizando Kali Linux como máquina atacante. Una muestra de tramas de formato en "C" se muestra como sigue.

```

/* Frame (58 bytes) */ static const unsigned char pkt1[58] =
0x00, 0x0c, 0x29, 0x20, 0x54, 0xb8, 0x00, 0x0c, /* ..) T... */
0x29, 0x73, 0xa7, 0x4d, 0x08, 0x00, 0x45, 0x00, /* )s.M..E.
*/ 0x00, 0x2c, 0x89, 0xf2, 0x00, 0x00, 0x27, 0x06, /* .....? */
0x3b, 0x34, 0xc0, 0xa8, 0x26, 0xab, 0xc0, 0xa8, /* ;4.&... */
0x26, 0xaa, 0xed, 0x87, 0xd4, 0x38, 0x66, 0x4b, /* &...8fK
*/
0xf9, 0xbc, 0x00, 0x00, 0x00, 0x00, 0x60, 0x02, /* .....'.
*/ 0x04, 0x00, 0xa3, 0xb7, 0x00, 0x00, 0x02, 0x04, /* .....
*/ 0x05, 0xb4 /* .. */ ; /* Frame (60 bytes) */ static const
unsigned char pkt2[60] = 0x00, 0x0c, 0x29, 0x73, 0xa7,
0x4d, 0x00, 0x0c, /* ..)s.M.. */ 0x29, 0x20, 0x54, 0xb8, 0x08,
0x00, 0x45, 0x00, /* ) T...E. */ 0x00, 0x28, 0x1b, 0x93, 0x40,
0x00, 0x80, 0x06, /* (.@... */ 0x10, 0x97, 0xc0, 0xa8, 0x26,
0xaa, 0xc0, 0xa8, /* ....&... */ 0x26, 0xab, 0xd4, 0x38, 0xed,
0x87, 0x00, 0x00, /* &..8... */ 0x00, 0x00, 0x66, 0x4b, 0xf9,
0xbd, 0x50, 0x14, /* ..fK..P. */ 0x00, 0x00, 0xbf, 0x60, 0x00,
0x00, 0x00, 0x00, /* ...'.... */ 0x00, 0x00, 0x00, 0x00 /* .... */ ;

```

Los anteriores paquetes fueron ataques llevados a cabo desde una máquina con Kali Linux hacia una máquina virtual con sistema operativo Windows 7. Este ataque se llevó a cabo utilizando un escaneo de puertos con NMAP.

Estos datos fueron sujetos a un procesamiento para identificar patrones de ataque por medio del programa llamado "analyzer de Tramas", este programa fue realizado en lenguaje C# y lo que realiza es una selección de tramas TCP tomando todas aquellas cuyo valor del protocolo sea 06H.

Con la ayuda del programa analizador de tramas se convirtieron los valores hexadecimales de la captura en valores decimales para posteriormente pasarlos a un formato delimitado por comas. Cabe señalar que la adquisición y procesamiento de los datos adquiridos no se hace en tiempo real, sino que hay una serie de pasos para poder implementar la información como entrada de datos a la red neuronal artificial.

#### VI. NORMALIZACIÓN Y REDUCCIÓN DE DIMENSIONES

El proceso de normalización nos permite escalar los valores para que se establezcan en un rango específico y así facilitar

la manipulación de los datos. La normalización en este trabajo se llevó a cabo usando el proceso Z-Score de Matlab®, el cual usa la media y desviación estándar, donde el nuevo valor se halla de la siguiente forma:

$$V' = \frac{V - \bar{V}}{\sigma} \quad (1)$$

Donde:

$V'$  = Valor nuevo

$V$  = Valor actual

$\bar{V}$  = Promedio

$\sigma$  = Desviación estándar

Una vez normalizados cada uno de los conjuntos de datos se realizó el proceso de obtención de componentes principales, con el fin de depurar las características y así reducir las variables en cada conjunto de datos (de nuevo este proceso se realizó con el programa Matlab®). Un segmento de los datos normalizados se muestra en la figura 1:

1	2	3	4	5	6	7	8
-0.6320	0.1200	-0.7537	0.7367	2.9103	-2.4014	-0.0485	0.0812
-0.6320	0.1200	1.3511	-1.3687	-2.3700	2.9914	-0.0605	-0.1221
-0.6320	0.1200	1.3511	-1.3687	-2.3700	2.9914	-0.0591	-0.1221
-0.6320	0.1200	-0.7537	0.7367	2.9103	-2.4014	-0.0605	0.0810
-0.6313	0.1200	1.3508	-1.3532	-1.0249	2.1104	-0.0605	0.0812
-0.6320	0.1200	-0.7382	0.7365	2.0476	-1.0276	-0.0605	0.0812
-0.6313	0.1200	1.3509	-1.3532	3.7208	1.1737	-0.0605	0.0812
-0.6320	0.1200	-0.7382	0.7365	1.1305	3.8192	-0.0605	0.0812

Figura 1. Datos Normalizados

Al ser cada conjunto de datos de  $n \times 8$  ( $n$ , ya que cada conjunto difiere en la cantidad de tramas), el resultado de aplicar el proceso de componentes principales redujo cada uno de estos a una matriz de  $8 \times 8$ , donde cada columna representa un componente principal. Estos están ordenados en forma decreciente de importancia, de manera que el primer componente proporciona (o recoge) la mayor variabilidad posible de los datos.

Haciendo uso de la herramienta Neural Network de Matlab® se alimentó la red neuronal (figura 2) para clasificación de la forma que sigue: De 200 datos seleccionados de una trama de

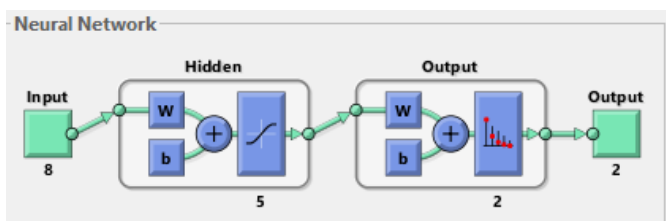


Figura 2. Red neuronal artificial generada con el toolbox de Matlab(R)

1500, se utilizó:

- 120 datos de entrenamiento
- 0 datos de prueba

En este caso, se utilizó para el entrenamiento, 80 datos correspondientes a tráfico normal y 40 datos para tráfico

tipo *flooding*. Para la prueba se utilizó, 60 datos de tráfico normal y 20 de tráfico tipo *flooding*. Cabe señalar que, para este trabajo, se tomaron datos de ataque tipo denegación de servicios (DoS) y diferentes configuraciones de la red neuronal.

## VII. ANÁLISIS DE RESULTADOS

Durante las pruebas realizadas con la red neuronal artificial se utilizó una red con aprendizaje supervisado. Se realizaron diferentes pruebas utilizando valores distintos en cada una de ellas con el objetivo de buscar resultados óptimos de la aplicación de la red neuronal.

Se armaron conjuntos de datos con los campos seleccionados a los cuales se les realizó un proceso de normalización y reducción de dimensionalidad por medio del procedimiento de análisis de componentes principales, dando como resultado conjuntos estandarizados en matrices de  $8 \times 8$  (8 filas por 8 columnas) para un total de 64 elementos por conjunto de datos que fueron usados como entradas al clasificador compuesto por una red neuronal con una capa oculta de neuronas variables y una salida de dos neuronas.

### VII-A. Caso 1

Para el primer caso de estudio se seleccionaron los siguientes valores (figura2):

- La función utilizada fue: *FeedForward Network*
- Número de entradas: 8
- Número de salidas: 2 (las cuales sirven para clasificar un ataque de un tráfico normal)
- Número de neuronas en la capa intermedia: 5

Resultado: En este caso, del 80 % de la muestra tomada como tráfico normal, se obtuvo el 63 % de clasificación, así como del 37 % de tráfico de DoS. En general, la clasificación en esta prueba fue del 65 % de clasificación correcta y 35 % de clasificación incorrecta.

### VII-B. Caso 2

En el segundo caso de estudio, se seleccionó lo siguiente (figura3):

- La función utilizada fue: *FeedForward Network*
- Número de entradas: 8
- Número de salidas: 2 (las cuales sirven para clasificar un ataque de un tráfico normal)
- Número de neuronas en la capa intermedia: 6

Resultado: En este caso, del 80 % de la muestra tomada como tráfico normal, se obtuvo el 81 % de clasificación para tráfico normal. Se obtuvo un 77 % de tráfico DoS con respecto al esperado. Se obtuvo en lo general una clasificación correcta del 55 % y 45 % de clasificación incorrecta.



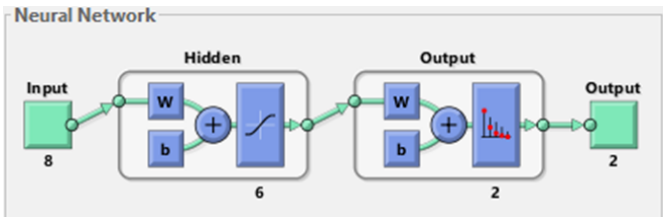


Figura 3. Red neuronal artificial generada con el toolbox de Matlab(R)

### VII-C. Caso 3

En el tercer caso, la red neuronal artificial con la que se trabajó fue la siguiente:

- La función utilizada fue: *FeedForward Network*
- Número de entradas: 8
- Número de salidas: 2 (las cuales sirven para clasificar un ataque de un tráfico normal)
- Número de neuronas en la capa intermedia: 10

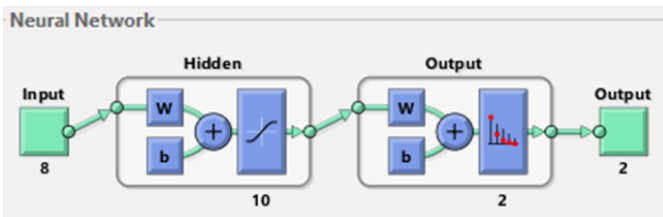


Figura 4. Red neuronal artificial generada con el toolbox de Matlab(R)

Resultado: Después del entrenamiento y dado los datos de prueba, del 80 % de la muestra tomada como tráfico normal, se obtuvo el 81 % de clasificación para datos de tráfico normal. Se obtuvo un 100 % de tráfico DoS con respecto al esperado. Se obtuvo en lo general una clasificación correcta del 95 % y 5 % de clasificación incorrecta.

### VII-D. Caso 4

Se realizó una última prueba con los mismos datos de entrada y salida, pero cambiando el número de neuronas en la capa intermedia, quedando la red como sigue (figura5):

- La función utilizada fue: *FeedForward Network*
- Número de entradas: 8
- Número de salidas: 2 (las cuales sirven para clasificar un ataque de un tráfico normal)
- Número de neuronas en la capa intermedia: 7

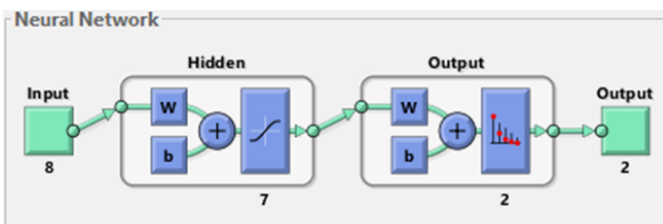


Figura 5. Red neuronal artificial generada con el toolbox de Matlab(R)

Resultado: El resultado obtenido en esta última prueba fue idéntico respecto al análisis con 10 neuronas, no se observó variación en los valores obtenidos.

## VIII. CONCLUSIONES

La aplicación herramientas basadas en redes neuronales para la identificación de ataques en redes de cómputo permite detectar peligros que no son considerados en los entornos comunes de operación.

Se analizaron varios casos de prueba con datos recolectados de una red real en operación y se logró obtener un modelo práctico y funcional que puede ser utilizado como apoyo en la identificación de tráfico de red no deseado. La aplicación de redes neuronales para la detección de ataques en las redes de cómputo puede apoyar en gran medida a los sistemas tradicionales dada su capacidad de aprendizaje y adaptación.

El modelo construido basado en redes neuronales artificiales, permite detectar paquetes peligrosos que forman parte de un ataque a la red de cómputo, a partir del conocimiento previamente adquirido con el análisis de anteriores paquetes, lo que no sucede con los sistemas no basados en redes neuronales.

El éxito de un sistema de redes neuronales artificiales para la detección de ataques en una red de cómputo se fundamenta en el adecuado entrenamiento y en la selección de un universo suficiente de paquetes de la red donde se utilice el sistema. Esto para que puedan detectar patrones o comportamiento que para otros sistemas basados en otras técnicas son descartados o no detectados.

## REFERENCIAS

- [1] B. S. OBS. (2017) Seguridad de la información. [Online]. Available: <https://www.obs-edu.com/int/blog-investigacion/sistemas/seguridad-de-la-informacion-un-conocimiento-imprescindible>
- [2] Icorp. (2018) Seguridad informática. [Online]. Available: <http://www.icorp.com.mx/soluciones/seguridad-informatica/>
- [3] K. Scarfone and P. Mell, "Guide to intrusion detection and prevention systems (idps)," National Institute of Standards and Technology, Tech. Rep., 2012.
- [4] J. P. Anderson, "Computer security threat monitoring and surveillance," *Technical Report*, James P. Anderson Company, 1980.
- [5] C. M. Bonilla, "Sistema neuronal de detección de intrusos," *Tendencias en ingeniería de software e inteligencia artificial*, 2008.
- [6] I. Lorenzo-Fonseca, F. Maciá-Pérez, F. J. Mora-Gimeno, R. Lau-Fernández, J. A. Gil-Martínez-Abarca, and D. Marcos-Jorquera, "Intrusion detection method using neural networks based on the reduction of characteristics," in *International Work-Conference on Artificial Neural Networks*. Springer, 2009, pp. 1296–1303.
- [7] E. Rich, K. Knight, P. A. González Calero, T. Bodega *et al.*, *Inteligencia artificial*, 1994.
- [8] R. Margareth. (2017) Inteligencia artificial o ia. [Online]. Available: <https://searchdatacenter.techtarget.com/es/definicion/Inteligencia-artificial-o-AI>
- [9] D. Peña, *Análisis de datos multivariantes*. McGraw-Hill España, 2013.

# *Identificación de huella plantar humana, utilizando descriptores de bajo costo computacional*

Marlene E. López-Jiménez, Raúl Santiago-Montero, David A. Gutierrez-Hernandez  
División de Estudios de Posgrado e Investigación  
Instituto Tecnológico de León; TecNM-León  
León, Gto. México

[M17240910@itleon.edu.mx](mailto:M17240910@itleon.edu.mx); [raul.santiago@itleon.edu.mx](mailto:raul.santiago@itleon.edu.mx)  
; [david.gutierrez@itleon.edu.mx](mailto:david.gutierrez@itleon.edu.mx)

Ignacio Hernandez-Bautista  
División de Estudios de Posgrado e Investigación  
Cátedra CONACYT/TecNM-León  
León, Gto. México  
[ihernandez@conacyt.mx](mailto:ihernandez@conacyt.mx)

**Abstract**—El presente artículo muestra una serie de métodos de extracción de características para el reconocimiento de objetos visuales, en este caso, clasificación de huella plantar. Después de mostrar el comportamiento de descriptores existentes (Histograma de gradientes Orientados), se proponen dos técnicas de extracción de características basadas en el concepto de compacidad, Factor E y Factor E Normalizado, como alternativas de bajo costo computacional. La base de datos empleada está conformada en su totalidad de 1,000 imágenes de huella plantar en niños de escuela primaria pública en México. Las propuestas de compacidad mostraron ser una propuesta eficaz y eficiente de descripción en regiones digitales, de manera que se obtuvo hasta un 99% de clasificación correcta de huella plantar

**Keywords**— *Arco Plantar Longitudinal Medial; Clasificación; Pie Plano; Pie Normal; Pie Cavo; Compacidad; Factor E; Factor E Normalizado; HOG.*

## I. INTRODUCCIÓN

Según la forma del arco del pie humano, se han identificado 3 morfologías distintas de pie: pie normal, pie plano y pie cavo. Los pies planos han sido considerados como factores que inciden en la aparición de padecimientos como; cansancio, lesiones en la zona del arco y tobillo, dolor de espalda, inestabilidad de la zona cervical superior, trastornos y degeneración de caderas, rodillas, entre otros [1,-3]. Uno de los temas más controvertido en el estudio del pie es la medición del Arco Plantar Longitudinal Medial (APLM) y su relación con los padecimientos descritos anteriormente. El estudio funcional del APLM es utilizado como medida útil, válida y de bajo costo para determinar el tipo de pie [3].

En este trabajo de investigación se muestran los resultados de clasificación, de huella plantar en niños de escuela primaria en México dentro del estado de Guanajuato, a través de procesamiento de imágenes digitales y técnicas de Inteligencia Artificial. Se propone una caracterización eficiente, de bajo costo computacional, basado en el concepto de compacidad. Se muestra como resultado que las propuestas tienen un alto valor discriminante, en comparación con descripciones como los Histogramas de Gradientes Orientados. Estos descriptores son fáciles de calcular e implementar, y sus

resultados para la identificación de la huella plantar en las 3 morfologías, mostraron alta eficacia.

## II. ANTECEDENTES

El diagnóstico de deformidades del pie en individuos requiere múltiples mediciones de la huella para calcular un conjunto de parámetros que determinan las características presentadas en el Arco Plantar Longitudinal Medial (APLM) [4]. Existen distintos métodos de evaluación del APLM, entre los principales se encuentran Índice de Arco Staheli (IAS), Ángulo de Clarke (AC), Índice de Chippaux-Smirak (ICS), Índice de Hernández Corvo (IHC) e Índice de Arco (IA) [3-5]. La discrepancia entre estos parámetros de la huella complica los intentos de clasificación racional de pies [1]. No obstante la medición de estos parámetros ha demostrado inconsistencia en el empleo de métodos tradicionales como el uso de Vernier o impresión de huella en tinta [6-9]. Existen varias propuestas en el diagnóstico asistido por computadora para la clasificación de tipo de huella plantar. En [4] se propone una aplicación de bajo costo computacional que permite identificar ejes característicos del pie y en el cual se muestra que algunas medidas establecidas por especialistas de manera manual tienen altos porcentajes de correlación al ser empleadas en el espacio digital. En [10] se propone la clasificación del pie plano mediante la combinación de indicadores múltiples con el proceso de red neuronal, para mejorar la precisión de la clasificación, sin embargo, no se realiza una propuesta de descripción distinta a las ya establecidas por médicos especialistas. Por otra parte, el diseño de plantillas de suela especial con sensores de presión para evaluar la distribución de presión plantar en diferentes partes claves, es un buen enfoque, pero es necesario fabricar diferentes tamaños de plantillas de suela para evaluar la distribución de la presión plantar en diferentes puntos clave [6,9]. Adicional a esto en [11] se mostró que este tipo de dispositivos genera un sesgo en la presión de huella plantar generada al ser subestimada la presión mostrada en la parte media del pie. Se han documentado otras técnicas para la reconstrucción y evaluación global de la huella plantar, en [12] se propone el uso de un algoritmo que realiza la detección de una parábola

en las áreas del arco y talón del pie, el método de elementos finitos es usado en múltiples investigaciones para la reconstrucción de huella [4,13]. Estas propuestas muestran buenos resultados pero el costo computacional es elevado en cuestión de tiempo y procesamiento. Para el diagnóstico masivo en la población es sumamente importante generar procesos de bajo costo computacional fáciles de implementar en hardware de bajo costo y rendimiento.

### III. MÉTODOS

En esta sección se presentará la serie de métodos realizada, para la clasificación de huella plantar en niños de escuela primaria. Se dará énfasis en la etapa de extracción de características, la cual es el punto principal de la investigación. Se realizó un estudio comparativo entre métodos conocidos de caracterización y propuestas basadas en el concepto de compacidad dando un enfoque a la eficacia de las mismas.

#### A. Base de datos

Para la realización de los experimentos de esta investigación, se generó una base de datos a partir de 50 niños con un rango de edad entre 7 y 9 años de escuela pública en México. A cada niño se le tomaron 10 muestras distintas

Las imágenes contenidas en la base de datos fueron adquiridas con el uso de un podómetro digital con una cámara web de 15 megapíxeles, en Fig. 2. a) se observa una muestra tomada de éstas imágenes.

#### B. Segmentación de la imagen

El proceso de segmentación y procesamiento de la imagen, se muestra en la Fig. 1. La imagen adquirida, originalmente en el espacio de color RGB, fue convertida a escala de grises. Posteriormente se aplicó un filtro anisotrópico para generar homogeneidad en las distintas áreas que presenta la imagen, adicionalmente el método de segmentación Mean Shift determina las distintas zonas de la imagen agrupadas por similitud de intensidad que se encuentra en la Fig. 2 b).

El método mean shift utiliza la composición de tonalidades en cada pixel aplicando el modelo de agrupamiento de k-medias [14]. Para la mayoría de las imágenes fueron establecidos 3 puntos de clusterización, para segmentar 3 partes distintas en las imágenes; el fondo, la piel del pie y la piel que ejerce mayor presión sobre la superficie

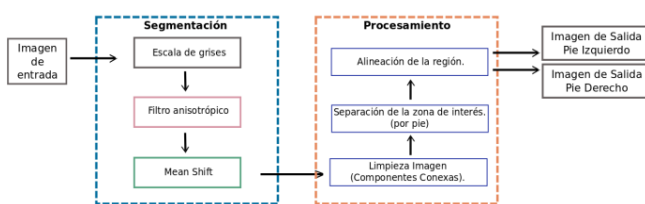


Fig. 1. Métodos utilizados para la segmentación y procesamiento de la imagen.

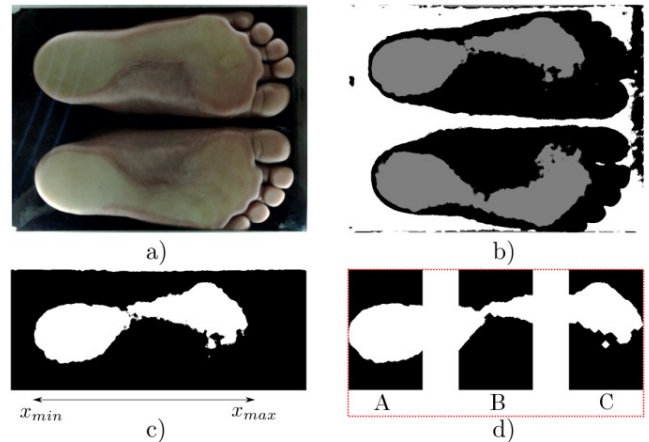


Fig. 2. Etapas realizadas para el análisis de la población; a) adquisición de la imagen; b) Segmentación de la imagen mediante mean shift; c) Imagen Binarizada y localización del eje mayor de la zona de interés; d) Separación de las zonas representativas del pie según el cálculo del IA.

plana del podoscopio, siendo, la última mencionada, la zona de interés.

#### C. Extracción de Características

Siendo utilizadas como base para la medición del Arco Plantar Longitudinal Medial, se definen las 3 partes principales del pie como Retropié, Mediopié y Antepié [1-4], las cuales pueden ser observadas en la Fig. 2 d) como A, B y C respectivamente. Partiendo de este supuesto la propuesta, en el presente artículo, de extracción de características basadas en compacidad, se enfocó a la descripción de estas 3 regiones. Para el desarrollo de la segmentación de la imagen como en esta sección se utilizó la herramienta MATLAB® [22].

##### 1) Modelo de Histograma de Gradientes Orientados

Se decidió utilizar los descriptores HOG, propuesto en [15], debido a su amplio uso en la caracterización de imágenes en problemas de clasificación como, el reconocimiento de gestos manuales, la clasificación de tipos de cáncer, detección de objetos, deformaciones en objetos, entre otros [15-19]. En el experimento generado para la extracción de características HOG se empleó la imagen binarizada de cada pie por separado, como se puede observar en la Fig. 2 c), analizando la huella completa y se realizó con un ventaneo de 64 x 64 píxeles.

##### 2) Modelo de descripción basado en Compacidad

Medir la compacidad de las formas planas se basa en la comparación entre la forma del objeto y la forma más compacta. Las medidas de compacidad tienen como objetivo ser adimensionales, simples de calcular, poseer un valor límite y este valor se correlaciona con la forma más compacta, en este caso, el cuadrado [20] y, tener un grado de discriminación entre una familia o clase de regiones similares [21].

#### a) Factor E

El factor E fue propuesto en [21], para asociar el grado de compacidad de una región digital con un valor cuantitativo. Este descriptor de forma utiliza el perímetro de contacto  $P_c$ , mismo que se expresa en unidades de longitud.

$$P_c = (Tn-P)/2 \quad (1)$$

En (1) se muestra la ecuación correspondiente al perímetro de contacto donde  $n$  representa el número de celdas regulares en  $T$  lados y un contorno con longitud  $P$ .

El factor E queda determinado como la razón entre las cardinalidades del perímetro de contacto del conjunto de celdas en la frontera de la región, para las zonas planas el factor E se determina por la relación expresada en (2).

$$\text{factor } E = (2P)/(4n-P) \quad (2)$$

#### b) Factor E Normalizado

Aunque el factor E cumple con los requerimientos de una medida de compacidad; este depende del número de celdas que componen a la región digital a evaluar. Ya que, cualquier región con un distinto número de celdas dará un resultado distinto de compacidad; da pauta a la propuesta de una medida que nos permita evaluar la similaridad de regiones con distinto número de celdas. El Factor e normalizado se encuentra expresado en (3). Siendo la relación obtenida entre el perímetro de la región y el perímetro de un cuadrado con área  $n$ , obtenido de la región a evaluar (4).

$$\text{Factor } E_{\text{normalizado}} = \text{Perímetro}_{\text{región}} / \text{Perímetro}_{\text{cuadrado}} \quad (3)$$

$$\text{factor } e_{\text{normalizado}} = \text{Perímetro}_{\text{región}} / 4(n)^{1/2} \quad (4)$$

Se realizó una sola etapa de extracción de características para el Factor E y Factor E normalizado, a partir de las cuales se generaron 4 propuestas distintas.

TABLA I. DESCRIPCIÓN DE EXPERIMENTOS REALIZADOS.

No. Experimento	Descriptor	Regiones Descritas
1	HOG	Huella total
2	FEN	A,B,C
3	FE	A,B,C
4	FEN/FE	A,B
5	FEN/FE	B

La Tabla I, indica el número de experimentos realizados, en total 5, donde se describe el tipo de descriptor utilizado y las regiones donde se aplica. En los experimentos 2 y 3 se generó un vector 3-dimensional, donde cada componente es un valor del descriptor por región. El experimento 4, son vectores 4-dimensionales; dos valores por

cada región. Finalmente, el experimento 5 es un vector 2-dimensional.

#### D. Clasificación

Para la etapa de Clasificación, en los tres tipos de morfología, se utilizó un conjunto de clasificadores, con el objetivo de medir el poder discriminante de nuestra propuesta. La implementación se hizo mediante plataforma JAVA y el paquete Weka® [23].

### IV. RESULTADOS

En esta sección se muestran los métodos utilizados para la clasificación de huella plantar, utilizando los descriptores propuestos a partir del concepto compacidad y comparándolo con el método de caracterización HOG. En la Tabla I se muestra el número de experimento y el descriptor o descriptores utilizados para la prueba de clasificación. En la Tabla II, un análisis del costo computacional de la extracción de características sustentado por los tiempos de procesamiento. Finalmente, la Tabla III, se describen los resultados de la eficacia de nuestra propuesta.

#### A. Extracción de características

La dimensión del vector resultante fue de 1188 características, las cuales fueron obtenidas de realizar el ventaneo con una matriz de 64x64 píxeles. En la Fig. 3. podemos visualizar el comportamiento de este descriptor.

En la Fig 2. d) podemos observar las distintas áreas establecidas para la evaluación de la huella en los experimentos 1-4 con descripción de compacidad.

TABLA II. TABLA COMPARATIVA DE PARÁMETROS DE PROCESAMIENTO POR CONJUNTO DE CARACTERÍSTICAS.

Descriptor	1. HOG	2. FEN	3. FE	4. FE/FEN	5. FE/FEN
T. P. (s)	0.40033	0.01733	0.0159	0.0228	0.0135
D	1188	3	3	4	2

En la Tabla II, Podemos observar el número de características obtenidas en cada experimento en la fila detonada por  $D$ . Así como el tiempo de procesamiento por experimento, obtenido a partir del promedio en procesamiento, de todos los individuos contenidos en la base de datos, en la fila con título " $T.P$ " medida en unidad de segundos, siendo así que los tiempos de procesamiento para los experimentos 2-5 representa tan solo el 11.84% del tiempo de procesamiento requerido en comparación de la extracción de HOG.

#### B. Clasificación

En la Tabla III, se muestran los resultados obtenidos tras utilizar distintos clasificadores bajo la técnica de validación cruzada en 10 folders definidos.

TABLA III. TABLA COMPARATIVA DE RESULTADOS DE CLASIFICACIÓN OBTENIDOS POR DESCRIPTOR.

Descriptor	1. HOG	2. FEN	3. FE	4. FE/FEN	5. FE/FEN
Naive Bayes	73.4%	71.1%	83.5%	85.2%	86.3%
MLP	--	76.9%	85.1%	86.8%	85.1%
Random Forest	<b>95.6%</b>	<b>84.3%</b>	<b>92%</b>	<b>92.2%</b>	<b>98.9%</b>
Random Tree	87.7%	77.9%	90.1%	90.2%	<b>99%</b>
1-NN	94.7%	82.4%	89.5%	90.6%	<b>99%</b>
3-NN	95.4%	83.7%	89.8%	92.1%	91.8%
Red Bayessiana	91%	74.7%	84.6%	85.5%	86.3%

De acuerdo a los resultados mostrados en la Tabla III, se puede apreciar la diferencia del rendimiento de las distintas propuestas de extracción de características. Los descriptors HOG muestran en su mayoría porcentajes sobre el 90%, mientras el descriptor FEN muestra porcentajes de clasificación entre los 70% y 80%, así mismo el descriptor FE

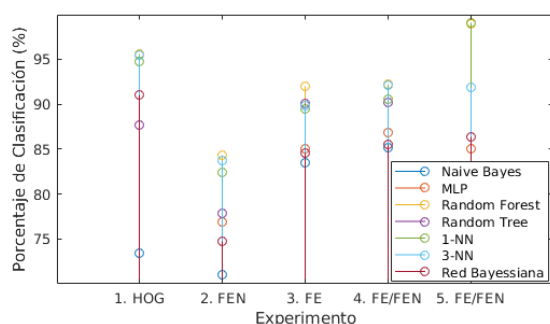


Fig. 4. Gráfica comparativa de porcentajes de clasificación agrupadas por experimento.

muestra mejores resultados en comparación del FEN. Por otra parte la combinación de ambas propuestas, Experimento 4 y 5, complementan una mejor descripción de la huella al tener porcentajes de clasificación mayores que incluso la descripción HOG, obteniendo en algunos casos, tan solo el 1% de clasificación errónea. En la Fig. 4 podemos observar una vista general de los porcentajes obtenidos por experimento, en la cual se puede observar que los mejores resultados fueron obtenidos por el experimento 1, 3 y 4.

Por otra parte, los árboles de decisión (Random Forest, Random Tree) mostraron los mejores resultados sobre los conjuntos de características obtenidos en la experimentación, superando a los métodos lineales como K Nearest Neighbours (1-NN, 3-NN), las redes neuronales (MLP) y los clasificadores probabilísticos, Naive Bayes y Redes Bayesianas, respectivamente. En el caso del Perceptrón Multicapa (MLP), la dimensión de los datos fue muy grande que la herramienta utilizada para las pruebas de clasificación no pudo desarrollar

el modelado de una red neuronal con 1118 características a evaluar.

## V. CONCLUSIONES Y TRABAJO FUTURO

La propuesta de este trabajo mostró que los descriptors con base en el concepto de compacidad, son una herramienta de bajo costo computacional y eficiente, para la clasificación de huella plantar en sus tres tipos de morfología. Como consecuencia, los tiempos de procesamiento disminuyeron considerablemente, tanto en la extracción de características, como durante el entrenamiento de clasificadores y el proceso de clasificación.

Comparado con el método HOG, nuestra propuesta reduce el espacio de características en más de 1,000 dimensiones. Esto impactó directamente en el entrenamiento de los clasificadores propuestos. Cabe mencionar que en el caso del Perceptrón Multicapa utilizar la propuesta HOG se vuelve prohibitivo.

Las propuestas de compacidad permiten generar modelos de clasificación más sencillos con altos niveles de clasificación correcta en este contexto. Se mostró que el Factor E y Factor E Normalizado son un conjunto sumamente poderoso a la hora de clasificar el tipo de huella logrando en conjunto hasta un 99% de clasificación.

Es importante hacer notar que la zona media del pie (*B*) se mostró como la más significativa para la clasificación, que la descripción de la huella plantar completa.

## AGRADECIMIENTOS

Los autores de este artículo agradecen a TecNM León-Celaya y a CONACYT por el apoyo económico brindado para la realización de este proyecto. Marlene Elizabeth López Jiménez con CVU 866095 agradece a CONACYT por la beca otorgada para llevar a cabo sus estudios de maestría.

## REFERENCIAS

- [1] Nikolaidou, M., Boudolos, K.: A footprint-based approach for the rational classification of foot types in young schoolchildren. En Rev. The Foot 16(2) (2006) 82–90.
- [2] Gutiérrez-Vilahú, L., Massó-Ortigosa, N., Rey-Abella, F., Costa-Tutusaus, L., Guerra-Balic, M.: Comparative study of plantar footprints in youth with down syndrome. International Medical Review on Down Syndrome 19(3) (2015) 36–42.
- [3] Razeghi, M., Batt, M.E.: Foot type classification: a critical review of current methods. En Rev. Gait & posture 15(3) (2002) 282–291
- [4] Ramírez, C. S. (2017). Análisis de dos métodos de evaluación de la huella plantar: índice de hernández corvo vs. arch index de cavanagh y rogers. Fisioterapia.
- [5] Onodera, A.N., Sacco, I. C. N. Morioka, E.H., Souza, P.S., de Sá, M.R., Amadio: What is the best method for child longitudinal plantar arch assessment and when does archmaturation occur? En Rev. The Foot 18(3) (2008) 142–149.
- [6] Maestre-Rendon, J., Rivera-Roman, T., Sierra-Hernandez, J., Cruz-Aceves, I., Contreras-Medina, L., Duarte-Galvan, C., & Fernandez-Jaramillo, A. (2017). Low Computational-Cost Footprint Deformities Diagnosis Sensor through Angles, Dimensions Analysis and Image Processing Techniques. Sensors, 17(11), 2700.
- [7] Lee, Y.-C.; Lin, G.; Wang, M.-J.J. Comparing 3D foot scanning with conventional measurement methods. J. Foot Ankle Res. 2014, 7, 44.

- [8] Mall, N.A.; Hardaker, W.M.; Nunley, J.A.; Queen, R.M. The reliability and reproducibility of foot typemeasurements using a mirrored foot photo box and digital photography compared to caliper measurements. *J. Biomech.* 2007, 40, 1171–1176.
- [9] Navarro, L.A.; García, D.O.; Villavicencio, E.A.; Torres, M.A.; Nakamura, O.K.; Huamani, R.; Yabar, L.F. Opto-electronic system for detection of flat foot by using estimat on techniques: Study and approach of design. In Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Buenos Aires, Argentina, 31 August–4 September 2010; pp. 5768–5771.
- [10] Aruntammanak, W., Aunhathaweep, Y., Wongseree, W., Leelasantham, A., y Kiattisin, S. (2013). Diagnose flat foot from foot print image based on neural network. En Biomedical engineering international conference (bmeicon), 2013 6th (pp. 1–5).
- [11] Urry, S. R., y Wearing, S. C. (2005). Arch indexes from ink footprints and pressure platforms are different. *The Foot*, 15 (2), 68–73.
- [12] Guerrero-Turrubiates, J.D.J.; Cruz-Aceves, I.; Ledesma, S.; Sierra-Hernandez, J.M.; Velasco, J.; Avina-Cervantes, J.G.; Avila-García, M.S.; Rostro-Gonzalez, H.; Rojas-Laguna, R. Fast parabola detection using estimation of distribution algorithms. *Comput. Math. Methods Med.* 2017, 2017, 6494390.
- [13] Morales-Orcajo, E.; Bayod, J.; de Las Casas, E.B. Computational foot modeling: Scope and applications. *Arch. Comput. Methods Eng.* 2016, 23, 389–416
- [14] Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. En *Rev. IEEE Transactions on pattern analysis and machine intelligence*, 24(5), 603-619.
- [15] Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. *IEEE Conf. Comput. Vis.*
- [16] Suleiman, A., & Sze, V. (2016). An energy-efficient hardware implementation of HOG-based object detection at 1080HD 60 fps with multi-scale support. *Journal of Signal Processing Systems*, 84(3), 325-337.
- [17] Bauer, S., Brunsmann, U., & Schlotterbeck-Macht, S. (2009, July). FPGA implementation of a HOG-based pedestrian recognition system. In *Proc. MPC-Workshop* (pp. 49-58).
- [18] Bakheet, S. (2017). An svm framework for malignant melanoma detection based on optimized hog features. *Computation*, 5(1), 4.
- [19] Korkmaz, S. A., Akçiçek, A., Binol, H., & Korkmaz, M. F. (2017, September). Recognition of the stomach cancer images with probabilistic HOG feature vector histograms by using HOG features. In *Intelligent Systems and Informatics (SISY), 2017 IEEE 15th International Symposium on* (pp. 000339-000342). IEEE.
- [20] Montero, R. S., y Bribiesca, E. (2009). State of the art of compactness and circularity measures. En *International mathematical forum* (Vol. 4, pp. 1305–1335).
- [21] Montero, R. S., y Bribiesca, E. (s.f.). Descriptores de regiones binarias en z2 y z3 basados en el concepto de compacidad. Universidad Autónoma de México. Unpublished.
- [22] Guide, MATLAB User's. "The mathworks." Inc., Natick, MA 5 1998, 333.
- [23] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1), 10-18.



# Reconocimiento invariante a rotación de texturas mediante la Transformada de Cúmulos Coordinados

Herrera Axel  
Centro de Investigaciones en Óptica A.C.  
León Gto.  
37100, México  
Email: axelherrerac@cio.mx

Cuevas Francisco J  
Centro de Investigaciones en Óptica A.C.  
León Gto.  
37100, México  
Email: fjcuevas@cio.mx

**Resumen**—Uno de los temas más abordados en aprendizaje automático está relacionado con el reconocimiento de patrones. Avances recientes en reconocimiento de patrones han aumentado la necesidad de mejorar los algoritmos para obtener características especiales de las texturas para poder realizar una mejor clasificación, la mayoría de estudios sobre análisis de texturas sólo se han llevado a cabo en un pequeño número de áreas por lo que aun sigue siendo limitado. Un desafío de esta técnica es reconocer texturas invariantes, específicamente a rotación. En el presente trabajo se propone un método invariante a rotación combinando la Transformada de Cúmulos Coordinados (CCR por sus siglas en inglés) y la transformada de Fourier, la CCR calcula el vector característico de la textura estudiada y la transformada de Fourier se utilizan y aprovechan las características invariantes a rotación que no tenía la CCR, dicho método ocupa imágenes binarizadas ya que este tipo de imágenes guarda mejor las propiedades estadísticas de la imagen. Para comprobar la eficiencia del método se realiza la clasificación de diferentes texturas reales que podemos tener en nuestro entorno.

## I. INTRODUCCIÓN

Desde sus inicios, el ser humano ha tenido la necesidad de plasmar lo observado para tener un registro de sus vivencias, sus momentos, de sus actividades cotidianas. Conforme ha avanzado la tecnología estos registros han podido llegar a digitalizarse y conservarse mucho más tiempo, gracias a esta apertura en el ámbito tecnológico se han desarrollado técnicas que ayudan a realizar el análisis estadístico de una imagen. Dentro de estos métodos estadísticos de análisis de imágenes digitales, recientemente se ha ido por el estudio de texturas. Dentro del ámbito industrial se ha vuelto una herramienta recurrente para los procesos que ocupan, algunas de estas industrias son: la médica, para mejoramiento de radiografías, el sensado remoto de imágenes, para poder realizar sensado desde otro lugar, la recuperación de imágenes basada en contenido, para ayudar a mejorar una imagen que este deteriorada o un objeto deteriorado que se le tomó una imagen y la segmentación documentada, para determinar los bordes de un objeto en específico. La clasificación de texturas es un procedimiento el cual obtiene información y características para clasificar algunos objetos (naturales o artificiales) métodos iniciales como el de Ojala et al. la obtuvieron los resultados que se esperaban para dicho análisis [1] pero, cuando se realiza el análisis de texturas en

diferentes direcciones las propiedades cambiarán de acuerdo a la dirección de la textura. En este artículo, se presenta una técnica de clasificación de texturas invariante a rotación. Para lograr esto se utilizan la transformada de Fourier combinada con la transformada de cúmulos coordinados (CCR por sus siglas en inglés). La CCR es el método que se propone para poder realizar el análisis estadístico de la imagen obteniendo un vector característico de la textura y la transformada de Fourier se utiliza porque nos ofrece características a rotación que otros autores aprovechan para realizar clasificación de texturas [2]. El artículo está desglosado de la siguiente manera: la sección 2 se explica el procedimiento de que realiza la CCR, sus propiedades y su método de binarización, en la sección 3 se muestra la Transformada de Fourier con sus características y propiedades y en la sección 4 se muestra como se integra estos dos métodos para realizar la invarianza a rotación para que la sección 5 se muestre los resultados y se explique el por qué se dan esos resultados.

## II. TRANSFORMADA DE CÚMULOS COORDINADOS

La Transformada de Cúmulos Coordinados (CCR) es un método no invasivo de clasificación de texturas propuesto por Kurmyshev-Cervantes [3] y Kurmyshev-Soto [4] el cual fue utilizado para que una imagen binarizada fuera caracterizada como un histograma de ocurrencia, el cual se convierte en descriptor de textura para realizar la clasificación. Para poder realizar esto se determina una ventana la cual recorrerá secuencialmente toda la imagen pixel a pixel dentro de la ventana se obtendrá un número en binario, ya que como la imagen es binaria nos dará cada pixel 0 o 1 (negro o blanco) este número en binario no lo podemos ocupar para el propósito que se necesita, por lo que se realiza una decodificación a decimal, esta decodificación consiste en hacer una revisión del número en binario del bit menos significativo (posición derecha) al bit más significativo (posición izquierda) de acuerdo a la posición que se encuentren los bits que valgan 1 se elevará 2 a la posición que se encuentre el bit 1 y se sumará con todos los demás bits iguales a 1 esto nos dará el valor en decimal. Este valor en decimal será el índice del histograma que se construirá de acuerdo a los números en decimal que la ventana detecta. Para entender matemáticamente el método lo explicaremos a continuación:



Dada una imagen binarizada de tamaño  $M \times N$ , es representada por la imagen binaria  $S^\alpha = \{s_{m,n}^\alpha\}$  tal que  $m = 1, 2, \dots, M$  y  $n = 1, 2, \dots, N$  son los índices de las filas y columnas de la imagen binarizada  $\alpha = 1, 2, \dots, 2^N$  es el índice de la mencionada imagen. Dicha imagen binarizada se caracteriza mediante el uso de una ventana rectangular de tamaño  $I \times J$  tal que  $I \leq N$  y  $J \leq M$ . Una vez establecido el tamaño de la ventana, esta recorre secuencialmente toda la imagen  $S^\alpha$ , moviéndola pixel a pixel y con la decodificación de binario a decimal que se realiza dentro de la ventana se obtiene un valor en decimal que incrementa unitariamente el valor decodificado del histograma en resumen, en la Figura 1 podemos ver un ejemplo del comportamiento de la ventana en la imagen binarizada.

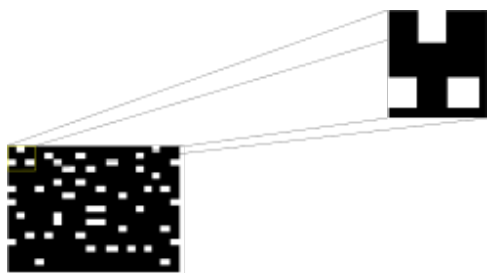


Figura 1. Representación de la CCR con su ventana de escaneo

El número de estados que puede tener el histograma dependerá del tamaño de la ventana, esto mismo es expresado matemáticamente:  $W = I \times J$  que es igual a  $2^w$ , el histograma de frecuencias se puede reducir más si se toman solo los valores diferentes a cero.

Por lo cual, la CCR calcula y relaciona a cada imagen binaria  $S^\alpha$  un histograma de frecuencias  $H_{(I,J)}^\alpha(b)$ , el cual representa el vector característico de la texturas escaneada por la ventana,  $\alpha$  representa el índice de la imagen,  $(I, J)$  representan el tamaño de la ventana de escaneo y por último la variable  $b = 2^W$  que representa la codificación del número en binario de acuerdo a la ventana de escaneo.

Con respecto a lo anterior, se puede determinar que el total de las frecuencias del histograma es igual al área total del mismo, esto es:

$$A = (M - I + 1) \times (N - J + 1) \quad (1)$$

Con este valor podemos normalizar el histograma  $H_{(I,J)}^\alpha(b)$ , el cual es considerado como una distribución en función de las frecuencias:

$$F_{(I,J)}^\alpha(b) = \frac{H_{(I,J)}^\alpha(b)}{A} \quad (2)$$

Es de vital importancia dejar claro como funciona la CCR para esto, un ejemplo claro de cómo se realiza la decodificación que necesita la CCR se puede representar por medio de una ventana de  $3 \times 3$ , cuyos códigos binarios y su decodificación se pueden observar en la Figura 2, donde el

número en decimal corresponde al índice del histograma de frecuencias.

Patron	Binario	Decimal
	101100001	353
	010101100	172
	010101101	173

Figura 2. Patrones binarios de acuerdo a la ventana de  $3 \times 3$ , su número en binario y su decodificación en decimal para el histograma de frecuencias de la CCR

Las propiedades de la CCR fueron establecidas en 2 teoremas fundamentales [5] los cuales son interpretados a continuación. El primer teorema establece la estructura de la CCR, específicamente en imágenes periódicas:

**Teorema 1.** Si un textón (celda primitiva) de una imagen binaria traslacional  $S^\alpha$  tiene tamaño  $\tau_1 \tau_2$  pixeles,  $\tau_1$  en una y  $\tau_2$  en otra dirección, entonces cualquier distribución de la CCR  $F_{(I,J)}^\alpha(b)$  toma no más que  $T = \tau_1 \tau_2$  valores diferentes de cero. Si la ventana de escaneo de la CCR tiene el tamaño igual o mayor al tamaño del textón,  $I \geq \tau_1$  y  $J \geq \tau_2$ , entonces  $F_{(I,J)}^\alpha(b)$  toma exactamente  $T = \tau_1 \tau_2$  valores diferentes de cero.

El siguiente teorema se enfoca en la relación del histograma con las funciones de auto correlación. Esto es que, el histograma de frecuencias no normalizado  $H_{(I,J)}^\alpha(b)$  contiene toda la información acerca de momentos de correlación de  $n$  puntos de la imagen  $S^\alpha$  si y solo si los vectores de separación entre  $n$  pixeles caben en la ventana de escaneo, eso es una función de distribución  $F_{(I,J)}^\alpha(b)$  la cual ofrece suficiente información de la función de probabilidad del  $n$ -ésimo-punto.

**Teorema 2.** Sea  $S^\alpha = \{s^\alpha(m, n)\}$  una matriz imagen binaria y  $F_{(I,J)}^\alpha(b)$  sea la función de distribución de la CCR. Si  $\max\{m_i\} \leq I$ ,  $\max\{n_i\} \leq J$  ( $i = 1, 2, \dots, k-1$ ) y  $k \leq I \times J$ , entonces ninguna función de auto correlación del  $k$ -ésimo orden para lograr una notación mas corta definiremos que  $\Delta = s^\alpha(m, n) s^\alpha(m + m_1, n + n_1)$ ,  $\alpha = S^\alpha(m + m_k, n + n_{k-1})$ ,  $\Omega = s^\alpha(m, n) s^\alpha(m + 1, n + n_1)$  y  $\xi = s^\alpha(m + m_k, n + n_{k-1})$ :

$$\langle \Delta \dots \alpha \rangle \equiv \lim_{M, N \rightarrow \infty} \sum_{m, n=1}^{M', N'} \Omega \dots \xi \quad (3)$$

puede ser reconstruida individualmente de  $F_{(I,J)}^\alpha(b)$  donde  $W = M \times N$  es el tamaño de la imagen, donde  $M' = M - \max\{m_i\}$ ,  $N' = N - \max\{n_i\}$ .

Este método a comparación de otros métodos propuestos para la clasificación de texturas citando algunos como Patrones Binario Locales (LBP) Filtro de Gabor, Momentos Invariantes de Hu, realizar más rápido el calculo, aun que el LBP es el más similar a la CCR las características que obtiene son diferentes (histograma diferente) usa una técnica específica de

binarización de imágenes con umbral local, que se define por el pixel central de un patrón local de la imagen. Después, el pixel central se excluye de cada patrón local. Por tanto, el sustento matemático del método de LBP es más limitado que el de la CCR ya que esta requiere la binarización como una etapa de preprocesado de imagen y, por tanto, puede usarse un umbral global o local adaptable. La clasificación basada en la CCR es más versátil, aunque las eficiencias terminales de clasificación por los dos métodos pueden ser muy cercanas en muchas aplicaciones.

La clasificación de patrones con la CCR conserva toda su información estructural para poder realizar un análisis más adecuado, por otro lado, la clasificación con la CCR puede ser orientado a bits, ya que lo leído en la ventana es un número en binario (0 o 1), lo que permitirá que el cálculo sea más rápido.

Cabe destacar que este método nos permite realizar una clasificación de texturas eficaz y de lo más adecuado, ya que la mayoría de los métodos estudiados se enfocan en realizar la clasificación de texturas por medio de imágenes en escala de grises, este método se basa en imágenes binarias ya que una imagen binaria guarda mejor su información estadística.

Sin embargo, uno de los problemas que se enfrenta en el escaneo de la imagen con la CCR es el tamaño de la ventana. Si es que se elige una ventana bastante grande, el histograma de la CCR será grande por lo que el computador tendrá problemas de administración de la información adecuadamente.

Una solución práctica para conectar los dos requisitos fundamentales (preservar información y capacidad de administración de la CCR) es reducir la escala de las imágenes que se van clasificando.

Este método se puede ocupar en la clasificación de texturas naturales, artificiales y computacionales, nos puede servir en determinar que tela estamos comprando esto se puede realizar con una app de celular que sea una herramienta auxiliar para el usuario esto se realiza mediante el enfoque la cámara en la tela y con una base de datos previamente cargada compara el vector característico obtenido con los de la base de datos y así determinar que tela se esta observando, se puede ocupar en el reconocimiento de corteza de árbol en un ecosistema estudiado, el mismo método se aplica que el de las telas pero con una base de datos diferente y se puede ocupar en la detección de fallas de mangueras industriales.

### II-A. Binarización

Como la CCR trabaja con imágenes binarizadas, es necesario determinar el tipo de binarización, por lo que para este proceso se elige la binarización por umbral. Esta viene declarada de la siguiente forma:

$$g(x, y) = \begin{cases} 0 & \text{si } f(x, y) \leq T \\ 1 & \text{si } f(x, y) > T \end{cases} \quad (4)$$

Dónde  $T$  es el umbral que va a dividir los dos modos dominantes del histograma de la imagen  $f(x, y)$  y  $g(x, y)$  es la imagen binarizada [6].

Para determinar el umbral adecuado se consideran la estadística de la imagen para permitir delimitar el objeto a estudiar, se logra mediante el método de Otsu.

El método de Otsu es un método estadístico el cual permite utilizar el histograma de la imagen en escala de grises, para poder determinar el umbral que separa los dos modos dominantes, el procedimiento es descrito a continuación:

Dada una imagen con niveles de intensidad  $D$  y asumiendo que el umbral buscado es  $T$ , las probabilidades acumuladas hasta  $T$  y desde  $T$  a  $D$  resultan ser:

$$w_1(t) = \sum_{z=1}^T P(z) \quad y \quad w_2(t) = \sum_{z=T+1}^D P(z) \quad (5)$$

Se obtiene las medias y varianzas asociadas:

$$\mu_1(t) = \sum_{z=1}^T zP(z) \quad y \quad \mu_2(t) = \sum_{z=T+1}^D zP(z) \quad (6)$$

$$\sigma_1^2(t) = \sum_{z=1}^T (z - \mu_1(t))^2 \frac{P(z)}{w_1(t)} \quad (7)$$

$$\sigma_2^2(t) = \sum_{z=T+1}^D (z - \mu_2(t))^2 \frac{P(z)}{w_2(t)}$$

Al final calcula la varianza ponderada:

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t) \quad (8)$$

Con estas fórmulas, permite elegir el mejor umbral  $T$  donde la varianza sea la mínima[7].

Hay que dejar muy en claro que al realizar el cálculo de la CCR si cambiamos de dirección la textura las características que se obtengan variaran dependiendo de la dirección que tome la textura ya que la CCR no es un método invariante por lo que se debe realizar un proceso extra para lograr la invarianza, para este experimento se toma la Transformada de Fourier.

### III. TRANSFORMADA DE FOURIER

Algunos de los autores han propuesto que la aplicación de la transformada de Fourier se debe realizar por medio de escala de grises y sacando el Espectro de la Transformada de Fourier en dos dimensiones [8], otra opción sería realizando una binarización del mismo y aplicando el mismo método, mientras que algunos autores mencionan que por medio de un algoritmo geométrico de corrección en la imagen se puede realizar el algoritmo antes mencionado obteniendo muy buenos resultados [9] en otras referencias, por medio de la transformada de Fourier directa con filtros nos permitirá realizar la detección de la textura en cuestión [10], [8] en este caso solo se ocupa la transformada y espectro de Fourier, En 2003, Li-Ma et al. demostraron que si la textura es periódica, entonces se puede inferir que el espectro de Fourier es periódico [11].

Para que pueda realizarse de una manera más sencilla y realizar un cálculo más eficiente se procede a ocupar la transformada de Fourier discreta de una dimensión. Esta, por definición es:

Sea una función  $f(x) \in \mathbb{R}$  se calcula la Transformada de Fourier, entonces la nueva función quedaría  $f(u) \in \mathbb{C}$  al que la

función de conversión de dicha transformada queda descrita de la siguiente manera:

$$f(u) = \frac{\sum_{l=0}^M \sum_{k=0}^N f(k) e^{-\frac{i2\pi kl}{M}}}{M} \quad (9)$$

Es de importancia mencionar que al obtener el espectro de Fourier, no estaría centrado, para evitar este problema se recurre a una corrección, de esta manera al reescribir la Transformada de Fourier queda de la siguiente manera:

$$f(u) = \frac{\sum_{l=0}^M \sum_{k=0}^N -1^l f(k) e^{-\frac{i2\pi kl}{M}}}{M} \quad (10)$$

Con dicha corrección obtendremos el espectro en el centro de la gráfica, la fórmula que describe adecuadamente el espectro es de la siguiente manera:

$$f(x) = \sqrt{\Re(f(u)) + \mathcal{C}(f(u))} \quad (11)$$

El espectro de Fourier en 2D por sus características definidas se establece que tiene propiedades invariantes a rotación y escala [12], [13], por lo que se puede inferir que estas mismas propiedades las contiene el espectro de Fourier en 1D.

#### IV. PROCEDIMIENTO DEL ALGORITMO

Retomando lo expuesto en las secciones anteriores, el algoritmo simple procede de la siguiente manera: Para poder realizar la CCR es necesario realizar la binarización adecuada de la imagen por que como bien se explicó, solo funciona con imágenes binarizadas entonces, para encontrar el umbral adecuado se procede a calcular el umbral con el método de Otsu, una vez obtenido el umbral se realiza la binarización por umbral a la imagen y es ahí donde la CCR realiza el cálculo pertinente y obtiene el vector característico. En este punto entra el procedimiento de mejora de la CCR.

El método propuesto es separado en 2 fases:

1. CCR
2. Espectro de Fourier

Cuando se carga una textura al algoritmo, el método calcula el histograma de la CCR posteriormente, ese mismo histograma se pasa al plano complejo con la transformada de Fourier, para poder obtener las propiedades invariantes de rotación, la conclusión del algoritmo es calcular el espectro de Fourier. Para ejemplificar el producto final se observa claramente en la figura 3:

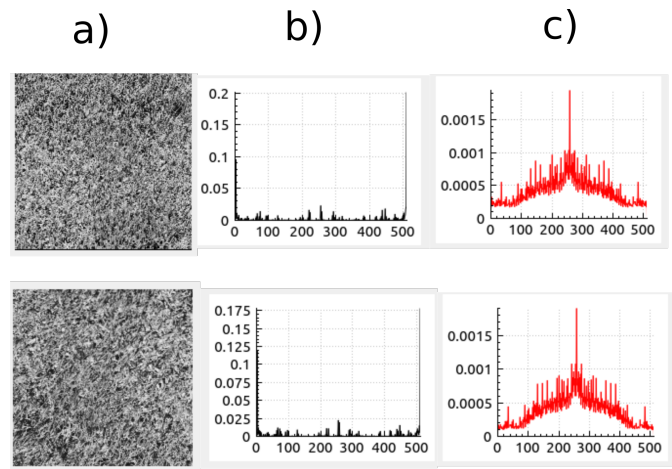


Figura 3. Columna a) Imagen original e imagen rotada a 135° Columna b) Histograma de la CCR normalizado de la imagen rotada y no rotada se puede observar como existe un cambio en el histograma de frecuencias al rotar la imagen Columna c) Histograma del Espectro de Fourier de una dimensión (EF1D) ahí se puede observar que no hay una variación visible o es mínima que no afecta la detección del patrón que es.

Como se puede observar en la figura 3, el vector característico de la textura tiene poca perturbación sin importar en que dirección se encuentre dicha textura, estas mismas perturbaciones que aparecen en el vector por el cambio de dirección que sufre la textura son vestigios de la CCR, pero con la ayuda de EF1D logra compensar esta variación por lo que, la Transformada de Cúmulos Coordinados combinada con el espectro de Fourier de una dimensión (CCR/EF1D) es un método que promete tener detecciones adecuadas.

#### V. EXPERIMENTO

La base de datos que es ocupada en el experimento son diez patrones de texturas tomados del álbum de Brodatz [14] que son mostrados a continuación:

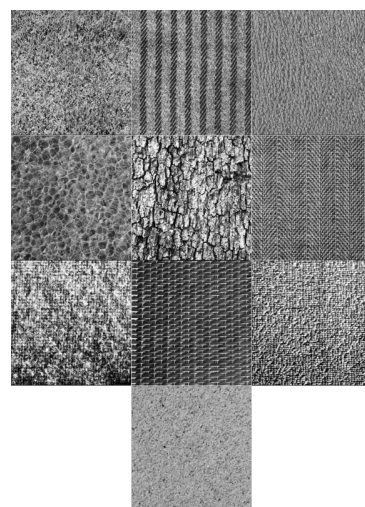


Figura 4. Base de datos de Texturas tomadas del álbum de Brodatz

Esta base de datos contiene diferentes texturas: pasto,tela,acrílico arrugado, burbujas sintéticas, madera, tejido de yute, lana, tejido de acero, tejido sintético y arena estas texturas se clasifican de acuerdo a la nomenclatura que Brodatz les puso (D9, D11, D24, D112, D12 H.E., D16, D19H.E., D65, D84 H.E. y D29). Para ellos se determinaron 7 direcciones pertinentes: 3 ángulos pequeño (5°,10°,15°) y 4 ángulo grandes (45°,135°,225°,315°) ya que se revisó que tan efectivo es el sistema por lo tanto, el total de imágenes que se ocuparan en el experimento son 70:

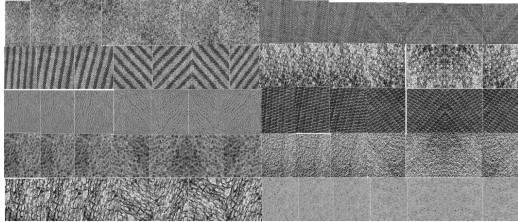


Figura 5. Conjunto de texturas rotadas a 5°,10°,15°,45°, 135°, 225° y 315° para mejorar el método invariante a rotación

La similitud de las texturas de la base de datos a uno de los patrones de texturas cargado al algoritmo se calculan utilizando la distancia de Hamming [15] de cada uno de los vectores calculados con la CCR/EF1D y se define de la siguiente manera:

$$d_{(I,J)}(S^\alpha, S^\beta) = \sum_{b=0}^{2^v-1} |H_{(I,J)}^\alpha(b) - H_{(I,J)}^\beta(b)| \quad (12)$$

Tal que el índice  $(I, J)$  indica el tamaño de la ventana de escaneo que es ocupada para la CCR/EF1D y  $N = I \times J$ . Dado un patrón de texturas  $S^P = [S_{mn}^P]$   $p = (1, 2, \dots, P)$  a clase correspondiente al patrón  $S^P$  es definido como una serie de imágenes  $S^\alpha$  tal que  $d_{(I,J)}(S^\alpha, S^\beta) = \min_n [d_{(I,J)}(S^\alpha, S^\beta)]$   $p = (1, 2, \dots, P)$ .

Se observa claramente que las reglas de distancia asignan a la imagen  $S^\alpha$  un patrón a la clase  $S^P$  si y solos si el histograma de la CCR/EF1D de  $S^P$  es el más parecido a  $S^\alpha$ .

En resumen, al comparar la base de datos con la imagen cargada al sistema, si alguno de ellos tiene el valor mas cercano a cero es el que corresponderá con el patrón adeudado.

La tabla 1 nos da un resumen de todas las texturas probadas y la figura 6 nos muestra la eficiencia de este método

Tabla I

TABLA DE RESULTADOS CON EL MÉTODO PROPUESTO, TOMANDO LA SIGUIENTE NOMENCLATURA: 1(D9), 2(D11), 3(D24), 4(D112), 5(D12 H.E.), 6(D16), 7(D19 H.E.), 8(D65), 9(D84 H.E.) Y 10(D29)

$\theta$	1	2	3	4	5	6	7	8	9	10
5	si	si	si	si	si	si	si	si	si	si
10	si	si	si	si	si	si	si	si	si	si
15	si	si	si	si	si	si	si	si	si	si
45	no	si	si	si	si	si	si	si	si	si
135	si	si	si	si	no	si	si	si	no	si
225	no	si	si	si	si	si	si	si	si	si
315	si	si	si	si	no	si	si	si	no	si

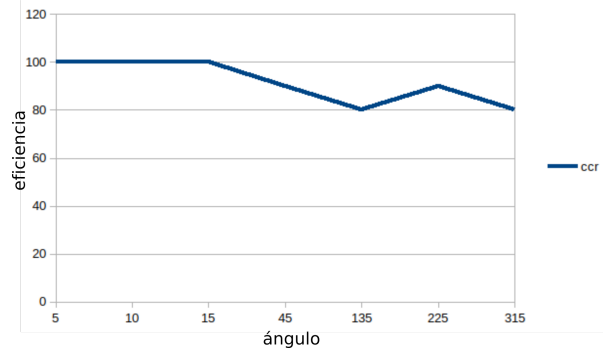


Figura 6. Eficiencia de la CCR/EF1D

Para obtener la eficiencia de las texturas se ocupa la siguiente fórmula:

$$R = \frac{N_k}{10^p} \times 100\% \quad (13)$$

Tal que,  $10^p$  corresponde al tamaño de la población, es decir, la cantidad de texturas que se encuentran en la base de datos,  $N_k$  corresponde al total de texturas con  $k$  ángulos que fueron detectadas correctamente. Esta fórmula es ocupada para cada ángulo, y el resultado es puesto en la tabla 2.

Tabla II

TABLA DE EFICIENCIA DE LOS ÁNGULOS

Ángulo	R
5°	100 %
10°	100 %
15°	100 %
45°	90 %
135°	80 %
225°	90 %
315°	80 %
Promedio	91.43 %

Los resultados de este estudio muestran que la gran mayoría de las texturas pudieron tener menos de 2 direcciones sin determinar correctamente, estos resultados son consistentes con estudios de otros métodos analizados y se sugiere que este método puede clasificarse perfectamente, sin embargo, este resultado no ha sido previamente descrito. Hay varias explicaciones posibles para estos resultados, al tener

imágenes que ya fueron procesadas previamente para realizar experimentos adecuadamente eliminan imperfecciones a la luz y por ende su detección puede llegar a resultar casi perfecta. Por lo tanto, se recomiendan estudios futuros sobre el tema actual.

## VI. CONCLUSIONES

Este documento ha dado cuenta de y las razones del uso generalizado de la CCR/EFID. En esta investigación, la meta era evaluar la textura en diferentes direcciones, uno de los descubrimientos más significativos que surgen de este estudio es que de acuerdo al ángulo y el tipo de textura que se presenta al algoritmo podrá o no clasificarla.

Los resultados de este estudio indican que la CCR compuesta con el Espectro de Fourier resulta una herramienta adecuada para mantener las características de las imágenes de texturas binarizadas. El estudio ha ido en cierta manera a mejorar nuestra comprensión de las texturas binarizadas, este no fue diseñado específicamente para evaluar el factor relacionado con la escala invariante. Además, se recomienda la investigación y experimentación en la CCR/EFID. Estos hallazgos sugieren varios cursos de acción para los nuevos métodos de rotación invariante.

## AGRADECIMIENTOS

Los autores agradecen al Centro de Investigaciones en Óptica y al CONACYT por el apoyo al reciente trabajo.

## REFERENCIAS

- [1] Mäenpää Topi, Ojala Timo, Pietikäinen Matti, and Soriano Maricor. Robust texture classification by subsets of local binary patterns. In *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, volume 3, pages 935–938. IEEE, 2000.
- [2] Li Wang and DC He. A new statistical approach for texture analysis. *Photogrammetric Engineering and Remote Sensing*, 56(1):61–66, 1990.
- [3] EV Kurmyshev and M Cervantes. A quasi-statistical approach to digital binary image representation. *Revista Mexicana de Física*, 42(1):104–116, 1996.
- [4] EV Kurmyshev and R Soto. Digital pattern recognition in the coordinated cluster representation. *Nordic Signal Processing Symposium, Espoo, Finland*, pages 463–466, 1996.
- [5] Raúl E Sánchez-Yáñez, Evgenii V Kurmyshev, and Francisco J Cuevas. A framework for texture classification using the coordinated clusters representation. *Pattern Recognition Letters*, 24(1-3):21–31, 2003.
- [6] Gonzalo Pajares, Alegre Enrique, and De la Escalera Arturo. *Conceptos y Metodos de vision por computadora*. Grupo de Vision del Comite Espanol de Automatica (CEA), 2016.
- [7] Liu Jianzhuang, Li Wenqing, and Tian Yupeng. Automatic thresholding of gray-level pictures using two-dimension otsu method. In *Circuits and Systems, 1991. Conference Proceedings, China., 1991 International Conference on*, pages 325–327. IEEE, 1991.
- [8] Takashi Matsuyama, Shu-Ichi Miura, and Makoto Nagao. Structural analysis of natural textures by fourier transformation. *Computer vision, graphics, and image processing*, 24(3):347–362, 1983.
- [9] João Batista Florindo and Odemir Martinez Bruno. Fractal descriptors based on fourier spectrum applied to texture analysis. *Physica A: statistical Mechanics and its Applications*, 391(20):4909–4922, 2012.
- [10] Robert Azencott, Jia-Ping Wang, and Laurent Younes. Texture classification using windowed fourier filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):148–153, 1997.
- [11] Li Ma, Tieniu Tan, Yunhong Wang, and Dexin Zhang. Personal identification based on iris texture analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (12):1519–1533, 2003.
- [12] Hamid A Jalab. Image retrieval system based on color layout descriptor and gabor filters. In *Open Systems (ICOS), 2011 IEEE Conference on*, pages 32–36. IEEE, 2011.
- [13] Kourosh Jafari-Khouzani and Hamid Soltanian-Zadeh. Radon transform orientation estimation for rotation invariant texture analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6):1004, 2005.
- [14] Phil Brodatz. *Textures: a photographic album for artists and designers*. Dover Pubns, 1966.
- [15] Herrera Axel and Cuevas Francisco J. Clasificación de texturas por el metodo de la transformada de cumulos coordinados. *XVII Congreso Nacional De Ingenieria Electromecanica Y De Sistemas (CNIES 2018)*, 2018.

# *Detección y Delimitación de Grietas en Estructuras de Inmuebles Utilizando Procesamiento de Imágenes*

María J. Hernandez Nava, Rocio Sandoval Salgado, Clara Cruz Ramos,  
Rogelio Reyes Reyes, Volodymyr Ponomaryov

Instituto Politécnico Nacional, ESIME Unidad Culhuacán, Ingeniería en Computación  
Av. Santa Ana No. 1000, Col. San Francisco Culhuacán, CP. 04430, Ciudad de México.

Tel. (55) 5729-6000 Ext. 73264 Fax (55) 56562058

email: mhernandezn1401@alumno.ipn.mx, rsandovals1500@alumno.ipn.mx, ccruzra@ipn.mx, rreyesre@ipn.mx

**Resumen** — *En los últimos años en México se han registrado una gran cantidad de sismos. Siendo el que se presentó el 19 de septiembre de 2017 con magnitud de 7.1° escala Richter el que causó más daños estructurales en muchos inmuebles de la Ciudad de México y zona conurbada, esto ocasionó que la revisión de los inmuebles dañados por personal calificado se efectuara de manera muy lenta, debido a que en México existe un número limitado de peritos para evaluar grietas; por tal motivo en el presente trabajo se propone desarrollar un sistema que auxilie al perito en la detección objetiva de grietas en un inmueble. Para la detección se emplea una conversión a escala de grises, posteriormente se aplica una umbralización para obtener una imagen binaria y la eliminación de regiones para reducir el ruido, finalmente mediante la esqueletonización se obtiene la delimitación y forma de la grieta. Los resultados obtenidos demuestran la correcta detección y delimitación de las grietas a partir de las fotografías de las mismas, así como también un valor promedio de 4.12 en la prueba subjetiva Mean Opinion Score (MOS).*

**Palabras clave** — *Detección de grietas, umbralización, esqueletonización, procesamiento de imágenes.*

## I. INTRODUCCIÓN

Un sismo es la vibración de la tierra producida por una rápida liberación de energía a causa del deslizamiento de la corteza terrestre a lo largo de una falla. La energía liberada se propaga en todas las direcciones desde su origen (foco o hipocentro) en forma de ondas [1]. En México, la mayoría de los sismos de gran magnitud ocurren por la subducción de la placa de Cocos por debajo de la placa de Norteamérica. La línea donde comienza la subducción se encuentra a pocos kilómetros fuera de las costas de Chiapas, Oaxaca, Guerrero y Michoacán. Por tanto, es en esa zona donde se localizan la mayoría de los epicentros [2]. En el año 2017 se registraron 26,364 sismos, y en el año 2018 se registraron 29,849 sismos en México. Siendo el del 19 de septiembre de 2017 con magnitud de 7.1° escala Richter el que causó más daños estructurales en la Ciudad de México [3].

Cuando ocurre un sismo, inmediatamente comienzan a suscitarse una serie de afectaciones de diferente magnitud a casas y edificios los cuales resultan afectados por grietas o incluso quedan inhabilitados. Para determinar si el inmueble es habitable o no, se requiere de una evaluación de los daños realizada por un perito, si los daños son muy graves se requiere de su demolición y reconstrucción. Las grietas son un factor

importante en esta evaluación ya que, dependiendo de su tipo, es el riesgo que representa para la estructura.

Se denomina grieta a la rotura que alcanza todo el espesor del elemento constructivo, dejándole inútil o debilitado para su posible función estructural. La grieta puede aparecer en cualquier elemento estructural o de cerramiento: pilares, vigas, muros, forjados, tabiques, etc. El método estándar para evaluar las grietas generalmente ha sido realizar una inspección visual y un diagnóstico acústico. Sin embargo, como pudo observarse en el sismo del pasado 19 de septiembre de 2017 el número de peritos para evaluar los inmuebles fueron insuficientes, además los resultados de la inspección al utilizar estas técnicas varían según la experiencia del evaluador, lo que hace que no haya suficiente objetividad [5]. Al tener un número limitado de peritos en México no se logra satisfacer la evaluación de todos los edificios afectados en casos de siniestros muy fuertes como el suscitado el pasado 19 de septiembre de 2017.

Debido a esta problemática se han propuesto técnicas y métodos enfocados en la detección y mejor visualización de las grietas empleando métodos automáticos o semiautomáticos de procesamiento de imágenes. Dentro de las propuestas se encuentran las realizadas por Ahmed Mahgoub Ahmed Talaba, Zhangan Huang, Fan Xia, Liu HaiMing [4] quienes proponen un método para la detección de grietas utilizando el método de Otsu y técnicas de filtrado múltiple, sin embargo, no detectan la forma del esqueleto de la imagen. Por otro lado, Lin Meng, Zhongkui Wang, Yoshiyuki Fujikawa y Shigeru Oyanagi [5] proponen un método de detección de grietas usando gradientes orientados, un método con resultados favorables para detectar el borde de la grieta y para detectar grietas en la superficie del concreto pero las imágenes resultantes presentan demasiado ruido en la detección.

Por lo tanto en este artículo se propone un método automático para detectar las grietas en una estructura de concreto mediante técnicas de procesamiento de imágenes basado en la esqueletonización para la eliminación de ruido, haciendo más efectiva la detección y delimitación de las grietas, cumpliendo con el objetivo de auxiliar a los peritos en la evaluación de la gravedad de la grieta presentada.

También se presenta una comparación del método propuesto con el método de Otsu y filtrado múltiple [4] mediante la aplicación de pruebas MOS [6].



II. SISTEMA PROPUESTO

La figura 1 muestra el diagrama a bloques del sistema propuesto para la detección de grietas utilizando procesamiento de imágenes.

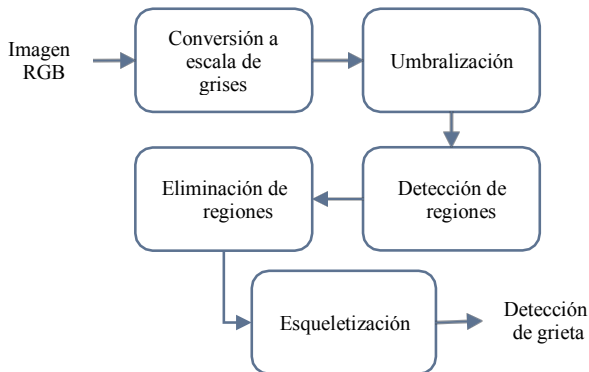


Fig. 1. Diagrama a bloques de detección de grietas

a) Se convierten los valores RGB de la fotografía capturada de la grieta en valores de escala de grises formando una suma ponderada de los componentes R, G y B mediante (1).

$$Y = 0.299R + 0.587G + 0.114B \quad (1)$$

b) Se realiza la umbralización de la imagen Y para obtener una imagen binaria I clasificando todos los píxeles en dos elementos: fondo y primer plano, esto se lleva a cabo mediante la evaluación de un umbral T de acuerdo a (2).

$$I = \begin{cases} 1, & Y(x, y) > T \\ 0, & \text{otro} \end{cases} \quad (2)$$

c) La detección de regiones consiste en la búsqueda de componentes conectados. Dos píxeles adyacentes en la imagen I definidos como p(x,y) y q(x,y), pertenecen al mismo componente conectado C si hay una secuencia de píxeles (p0, p1, ..., pn) de C cuando p0=p, pn = q, y pi es un vecino (adyacente al píxel) de pi-1 para n= 0,1,...,n [7]. La búsqueda de componentes conectados se realiza con los siguientes pasos:

- i. Para cada píxel p(x,y) con valor I=1 se establece una máscara de 3x3.
- ii. Se realiza una comparación de su valor con sus píxeles vecinos en 8 direcciones, como se ilustra en la figura 2.

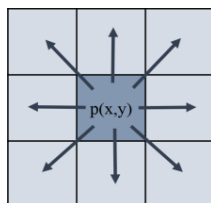


Fig. 2. Píxeles vecinos

Si se cumple (3) entonces son parte del mismo componente. Un ejemplo de esto se muestra en la figura 3(a), p y p0 pertenecen al mismo componente que llamaremos C, por otro lado, p1 no pertenece al componente conectado C.

$$Y_{p_0} = Y_p \quad (3)$$

- iii. Posteriormente se genera una matriz de etiquetas de regiones, y se asigna un identificador numérico único a cada conjunto de componentes conectados.

En la Figura 3(b) se muestra como a todos los píxeles pertenecientes al componente C se le asigna un valor de 1 en la matriz de etiquetas.

0	1	1	0	1
0	1	0	1	0
0	0	1	0	0
1	0	0	0	1
1	0	1	1	1

(a)

0	1	1	0	1
0	1	0	1	0
0	0	1	0	0
3	0	0	0	2
3	0	2	2	2

(b)

Fig. 3. a) Ejemplo de componentes b) Matriz de etiquetas

d) La eliminación de regiones tiene como finalidad eliminar texturas existentes en las regiones ajenas a las grietas, esto se lleva a cabo de la siguiente manera:

- i. De la matriz de etiquetas se obtiene el área (número de píxeles conectados) de cada región.
- ii. Para las regiones de área menor a N píxeles se cambia su valor a 0.

e) Finalmente se aplica el algoritmo de Esqueletonización Pavlidis a la imagen para obtener la forma de la grieta más visible, ya que se puede usar una representación esquelética de un objeto para describir su estructura [7], esto se lleva a cabo mediante la eliminación de píxeles del contorno del objeto que no afectan la conectividad del mismo.

- i. Se establece una máscara de 3x3.
- ii. Se comprueba si los píxeles contenidos en la máscara tienen similitud con plantillas que se muestran en la figura 4 [8]. Si hay similitud se considera esqueleto. Para los píxeles que no cumplen lo anterior se cambia su valor a 0.



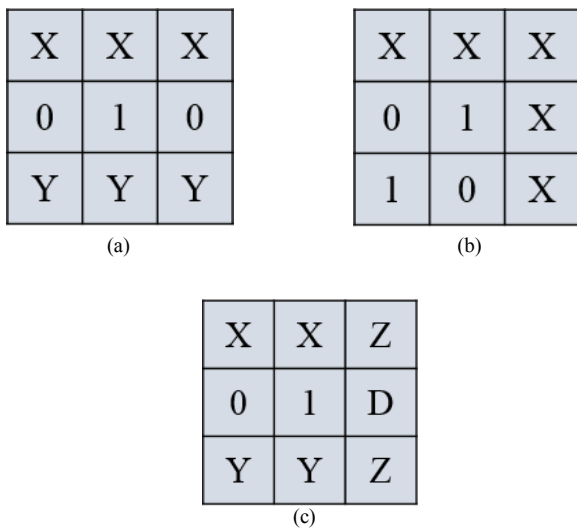


Fig. 4. Plantillas del algoritmo de eskeletonización Pavlidis

Para las plantillas de la figura 4(a) y 4(b) al menos un píxel de los grupos etiquetados como X o Y debe ser diferente de cero. Para la plantilla de la figura 4(c) al menos un píxel debe ser diferente de cero. Si ambos pixeles etiquetados con Z son diferentes de cero, entonces los etiquetados con X o Y pueden ser cualquier valor. De otra manera, al menos un miembro de cada par marcados con X y Y debe ser diferentes de cero, el píxel marcado con D es parte del esqueleto.

### III. PRUEBAS Y RESULTADOS

El sistema fue desarrollado en MatlabWorks Matlab R2018a (1.8.0\_152), en una computadora Acer aspire F5-573-57LP con procesador i5-6200U 6ª generación de procesadores Intel® Core 2.3 GHz con Turbo Boost hasta 2.8 GHz, 1TB en Disco Duro, Intel HD Graphics 520, DDR4-SDRAM, 16 GB de memoria interna, y sistema operativo Windows 10.

Para evaluar el sistema propuesto se utilizaron las imágenes mostradas en la figura 5, en formato JPG a color, las cuales presentan diferentes tipos de grieta, así como su ubicación en diferentes partes de la estructura de un inmueble. Las imágenes presentadas tienen una resolución de 700x420 píxeles, así como diferente luminosidad, textura y contraste.

Los resultados obtenidos para la grieta en el centro del muro de carga figura 5(a) se presentan en la figura 6.

La figura 6(a) muestra la imagen original a color antes de ser procesada, la figura 6(b) muestra la imagen original en escala de grises para su procesamiento más rápido y donde se aprecia que la fisura es más notable en la imagen grisácea. La figura 6(c) muestra la imagen binarizada para realizar un procesamiento más rápido, se usó un umbral  $T=80$ , la figura 6(d) muestra en color rojo donde se detectan áreas a eliminar, se consideraron áreas menores a un tamaño de  $N=30$  píxeles. La figura 6(d) muestra la imagen donde la grieta ha sido detectada, la figura 6(e) muestra finalmente el esqueleto de la grieta obtenida, donde se puede apreciar mejor la forma que tiene la grieta para poder auxiliar al perito a la hora de su visualización.

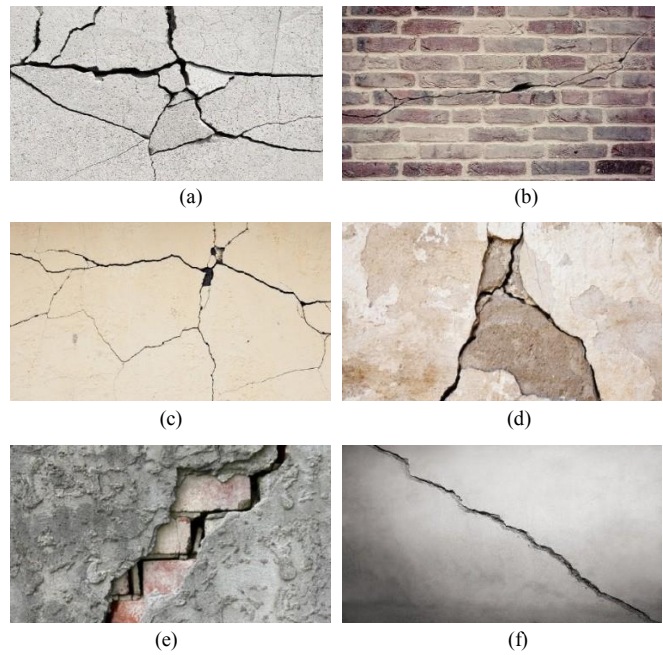


Fig. 5. Imágenes de grietas utilizadas (a) Grieta en el centro del muro de carga (b) Grieta superficial en muro divisorio hueco (c) Grieta en muro de contención (d) Grieta en muro divisorio (e) Grieta diagonal en muro de contención (f) Grieta diagonal en muro divisorio

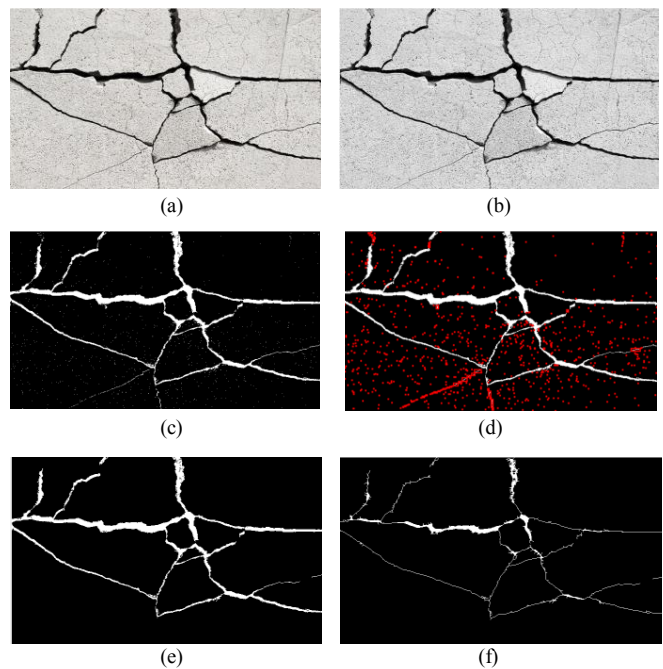


Fig. 6. Imagen procesada de grieta en el centro del muro de carga. a) Imagen original, b) Imagen en escala de grises, c) Imagen binaria, d) Áreas a eliminar, e) Grieta detectada y f) Esqueleto de la grieta.

Para la grieta en muro de contención (véase la figura 5(e)) los resultados obtenidos se muestran en la figura 7, en esta prueba se puede ver el beneficio de realizar una eliminación de regiones ya que al presentarse una grieta cerca de un muro con

acabado rustico se requiere la eliminación de la textura del mismo para realizar una detección objetiva.

En la figura 7(a) se muestra la imagen original, la imagen en escala de grises se muestra en la figura 7(b), la imagen binarizada con un umbral  $T=80$  en la figura 7(c), en la figura 7(d) se marcan en color rojo las áreas menores a un tamaño de  $N=30$  píxeles, al eliminarse estas regiones se obtiene la imagen de la figura 7(e) que es la imagen de la grieta que ya contiene menor ruido y finalmente obtenemos el esqueleto de la grieta figura 7(f).

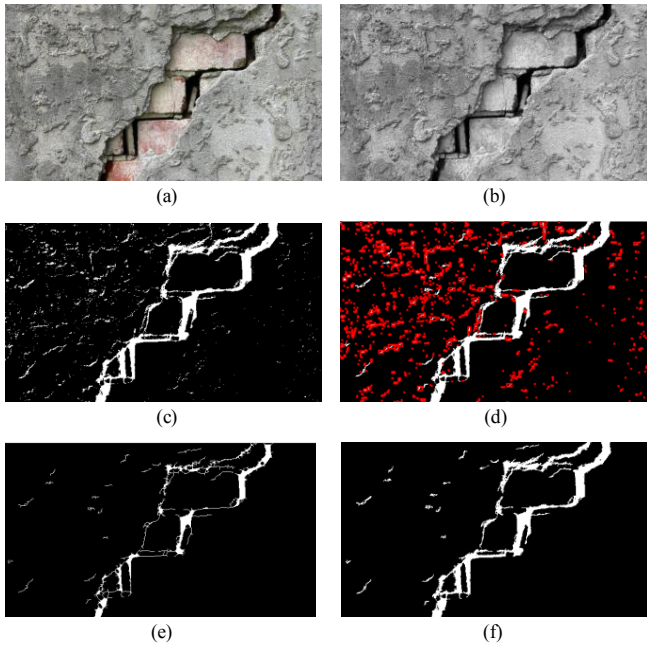


Fig. 7. Imagen procesada de grieta diagonal en muro de contención. a) Imagen original, b) Imagen en escala de grises, c) Imagen binaria, d) Áreas a eliminar, e) Grieta detectada y f) Esqueleto de la grieta.

Los resultados para la grieta superficial en muro divisorio hueco (véase la figura 5(b)) se muestran en la figura 8, a pesar de la textura del muro se detectó la grieta y se obtuvo su forma.



Fig. 8. Imagen procesada de grieta superficial en muro divisorio hueco a) Imagen original a color y b) Forma de la grieta al finalizar su procesamiento.

En el caso de la grieta en muro de contención (véase la figura 5(c)) los resultados se muestran en la figura 9, en la cual se puede observar que se detectó la forma de la grieta, pero se eliminaron algunas partes de ella debido a que eran de tamaño menor a 30 píxeles.

Para la imagen con la grieta diagonal en muro divisorio (véase la figura 5(d)) y la grieta en muro divisorio (véase la figura 5(f)), los resultados se muestran en la figura 10 y en la figura 11 respectivamente, en donde se puede observar que la detección de la grieta es favorable, la forma de la misma es clara y se eliminó el ruido en ambos resultados.

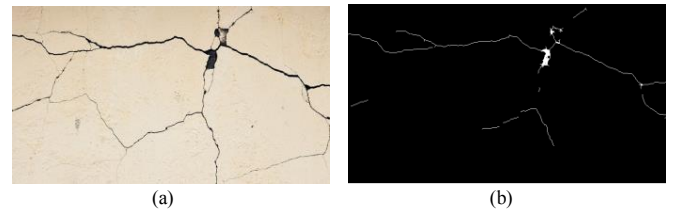


Fig. 9. Imagen procesada de grieta en muro de contención a) Imagen original a color y b) Forma y delimitación de la grieta detectada al finalizar su procesamiento.

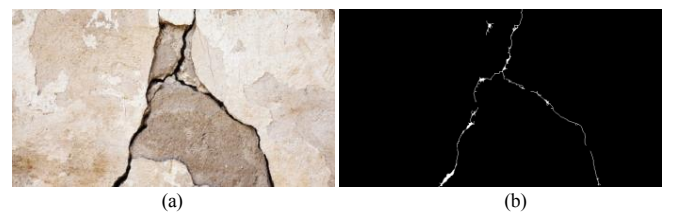


Fig. 10. Imagen procesada de grieta en muro divisorio a) Imagen original a color y b) Forma y delimitación de la grieta detectada al finalizar su procesamiento.



Fig. 11. Imagen procesada de grieta diagonal en muro divisorio a) Imagen original a color y b) Forma y delimitación de la grieta detectada al finalizar su procesamiento.

La correcta detección y delimitación de la grieta en la imagen es subjetiva, por lo tanto, para evaluar los resultados del sistema, se realizó una prueba MOS [8]. La prueba fue aplicada a 30 personas entre 20 y 50 años.

La prueba MOS se realizó mostrando la imagen original durante 5 segundos a las personas, después se les mostró la imagen de la grieta detectada y delimitada por el sistema propuesto. Los espectadores determinaron en qué medida era correcta la detección y delimitación de la grieta de acuerdo a su percepción.

Los criterios de evaluación para la prueba MOS fueron 5 (malo, pobre, justo, bueno y excelente), asignados a un valor entre 1 a 5, donde "1" es una nula detección y delimitación, y "5" representa una excelente detección y delimitación.

En la figura 12 se muestran los resultados obtenidos de la prueba MOS aplicada para la evaluación del sistema propuesto, representados en una gráfica de barras.

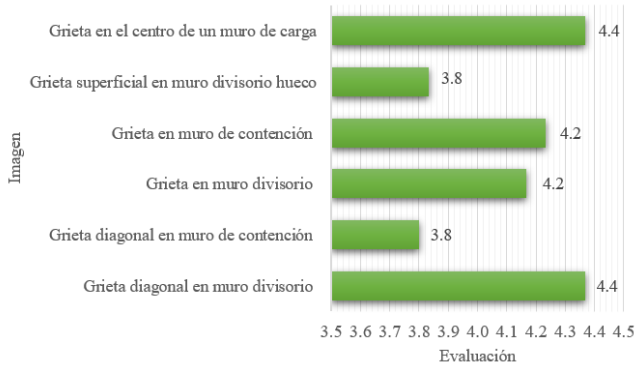


Fig. 12. Resultados de la prueba MOS aplicada al sistema propuesto

Tabla I. Resultados de la Prueba MOS

Imagen	Método de Otsu y filtrado múltiple [4]	Nuestro
Grieta en el centro de un muro de carga	4.1	4.4
Grieta superficial en muro divisorio hueco	1.5	3.8
Grieta diagonal en muro de contención	1.8	3.8
Grieta diagonal en muro divisorio	1.7	4.4

El promedio general obtenido en los puntajes de las pruebas MOS para las diferentes imágenes evaluadas fue de 4.12, este valor indica que independientemente de la variación de las características de las imágenes la detección y delimitación de la grieta en la imagen es considerada como buena para las personas a quienes se aplicó la prueba.

En la figura 13 se muestra una comparación entre los resultados obtenidos por el sistema propuesto en este artículo y el propuesto por Ahmed et. al. [4].

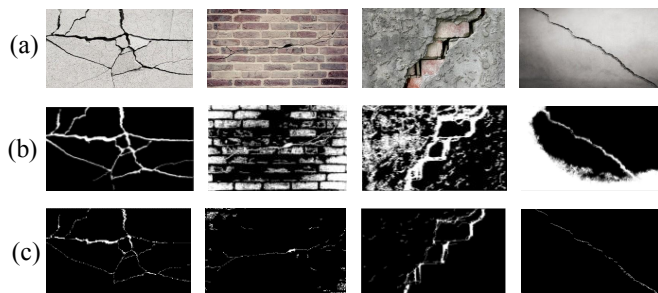


Fig. 13. Comparación de resultados entre el sistema propuesto y el propuesto en [4] a) Imágenes originales, b) Método propuesto en [4] y c) Sistema propuesto

Los resultados obtenidos en la figura 13 demuestran un mejor rendimiento para la detección y delimitación de la grieta en la imagen en el sistema propuesto comparado con los resultados del sistema propuesto en [4].

En la tabla 1 se muestran los resultados obtenidos de la prueba MOS del sistema propuesto y el sistema propuesto en [4].

#### IV. CONCLUSIONES

En este artículo, se propuso un método para la detección de grietas aplicando técnicas de procesamiento de imágenes. Las pruebas realizadas y presentadas demuestran la utilidad del método para detectar satisfactoriamente las grietas, con la eliminación de regiones y esqueletonización se redujo el ruido. Sin embargo para casos en que se presenten grietas con partes muy pequeñas o delgadas estas son eliminadas, sin embargo el sistema propuesto puede ajustarse modificando los valores de T y N para evitar que estas sean eliminadas. Como trabajo futuro se propone ampliar el enfoque para poder clasificar la grieta.

#### AGRADECIMIENTOS

Agradecemos al Instituto Politécnico Nacional, al Consejo Nacional de Ciencia y Tecnología de México, y a la Comisión de Operación y Fomento de Actividades Académicas (COFAA) del IPN por el apoyo otorgado para el desarrollo de este trabajo.

#### REFERENCIAS

- [1] Universidad de Costa Rica, “¿Qué es un sismo?”. Internet: [https://rsn.ucr.ac.cr/documentos/educativos/sismologia/3\\_669-que-es-un-sismo](https://rsn.ucr.ac.cr/documentos/educativos/sismologia/3_669-que-es-un-sismo) Marzo 10, 2015 [Enero 28, 2019].
- [2] Meli Piralla. Diseño estructural. México, LIMUSA noriega editores, 2001, pp. 407.
- [3] Servicio Sismológico Nacional, “Resumen mensual de sismicidad”. Internet: <http://www.ssn.unam.mx/sismicidad/resumen-mensual/> [Enero 10, 2019].
- [4] A. M. Ahmed, Z. Huang, F. Xi and L. Hai Ming, “Detection crack in image using Otsu method and multiple filtering in image processing techniques”, Optik, Volume 127, Issue 3, February 2016, pp. 1030-1033.
- [5] Lin Meng, Zhongkui Wang, Yoshiyuki Fujikawa and Shigeru Oyanagi, “Detecting Cracks on a Concrete Surface Using Histogram of Oriented Gradients”, 2015 International Conference on Advanced Mechatronic Systems (ICAMechS), Beijing, China, 22-24 Aug. 2015.
- [6] Streijl, Robert C., Stefan Winkler, and David S.Hands. "Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives." Multimedia Systems 22.2 (2016): 213-227.
- [7] Haralick, Robert M., and Linda G. Shapiro, Computer and Robot Vision, Volume I, Addison-Wesley, 1992, pp. 28- 48.
- [8] Gustavo Medina, Yessica Calderón, Gennadiy Burlak, “Recognition of characters through OCR (Optical Character Recognition)”, Programación Matemática y Software (2018) 10(1): 44-56. ISSN: 2007-3283.

# Determinación de Parámetros estadísticos de la pupila en el ojo humano

Rodolfo Romero Herrera  
ESCOM – IPN Departamento de ciencias e  
Ingenierías de la Computación  
Ciudad de México  
e-mail: rromeroh@ipn.mx

Saúl De La O Torres  
ESCOM – IPN Departamento de ciencias e  
Ingenierías de la Computación  
Ciudad de México  
e-mail: sdaleot@ipn.mx

Jose Elias Romero Martínez  
ESIME - IPN  
Ciudad de México  
e-mail: saile14398@gmail.com

**Abstract—** Se realizó un análisis de las demisiones de la pupila. Para la investigación se usó el método de Viola & Jones para detectar el iris; y mediante el uso de histogramas y su morfología se llevó cabo el seguimiento del ojo. Con técnicas estadísticas se obtuvo promedios, varianza y área de la pupila. El presente trabajo aporta mediante el software una aplicación económica y portable útil para estudios del ojo en cómputo afectivo (determinación de actividad del cerebro dependiente de la posición del ojo) o detección de enfermedades.

**Keywords—**Ojo; histograma; estadística; Viola & Jones;

## I. INTRODUCTION

Se puede considerar la localización de los ojos como un problema satisfactoriamente solucionado; sin embargo el reconocimiento de emociones humanas en los ojos es todavía un problema latente, que puede tener múltiples aplicaciones.

### A. Localización del Ojo Humano

Ubicar el ojo humano es un problema complejo. Sin embargo, la localización del ojo de manera no invasiva es un hecho en las interfaces hombre-computadora [1]; Los pupilómetros infrarrojos proporcionan una medida objetiva del tamaño de la pupila. Sin embargo no tienen un uso generalizado en la medicina [2]; Para localizar el ojo se puede usar filtros oculares, sin restricciones con respecto al fondo y el color de la [1]. Los métodos más frecuentes se basan en la postura frontal de la cara; a pesar de que el iris varía en las imágenes debido al cambio de posición de la cabeza, el color de los ojos, el tamaño, la forma, textura y el entorno ambiental [3][4][5]. Un método que comúnmente se propone es el de Viola y Jones [6] [7]; donde además se puede usar un detector de puntos de esquina y el algoritmo K-medias para agrupar los puntos vecinos y determinar las regiones candidatas a ojos [8]. En algunos casos, se usa una máquina de soporte vectorial (SVM) e histograma

de gradientes orientados (HOG) [2] [9]. Witama y Widodo operan con el clasificador HAAR para detectar la boca, los ojos y las cejas en la cara, y extraer las características con la wavelet de Gabor [10].

Zhang y Smith proponen debido a la falta de robustez, baja eficiencia, gran gasto computacional y fuerte dependencia de hardware; un algoritmo de reconocimiento de género con un enfoque de localización ocular modular, con reconocimiento de gestos en la mirada, lo que pone de manifiesto el efecto de los estados afectivos [11].

La detección de los ojos depende de una detección robusta de la cara; ya que se debe considerar cuando los ojos están total o parcialmente cerrados, con sombras, con gafas o con rotaciones [11] [12]. De cualquier forma es necesaria una base de datos de seguimiento ocular sobre imágenes de caras [13]. La detección de los ojos, lleva mucho tiempo recopilar datos de entrenamiento a gran escala y puede ser poco confiable [14].

## II. METODOLOGIA

Para obtener las características del ojo se realizan 4 pasos:

1. Localización del rostro.
2. Detección de los ojos dentro del rostro localizado.
3. Erosión y dilatación del ojo.
4. Medición de patrones estadísticos.

### A. Viola & Jones

Para la detección de rostros en una imagen se usó Viola-Jones. Este requiere de una vista frontal de la cara [15]. Las características del algoritmo de Viola-jones son:

**Robusto:** Con un porcentaje de verdaderos positivos de 99.9 % y de 3.33 % de falsos positivos [16].

**Detección:** El esquema de detección genérica con el que trabaja el algoritmo hace posible que pueda ser entrenado para la detección de otro tipo de objetos.



El algoritmo cuenta con 4 fases:

1. Selección de características tipo Haar.
2. Creación de una imagen integral.
3. Entrenamiento Adaboost con clasificadores en cascada.
4. Selección de características Haar

Una característica Haar se define sobre regiones rectangulares en una imagen en escala de grises, la cual se forma por un número finito de rectángulos; el valor escalar consiste en la suma de los píxeles de cada rectángulo; donde se aplica un cierto factor de peso. Ver la ecuación (1); donde  $\{r_1, \dots, r_N\}$  son aquellos rectángulos que forman la característica (ver Figura 2) y  $w_i$  el peso de cada uno. En la Figura 1 se muestran las características de tipo Haar de 2, 3 y 4 rectángulos.

$$Características_j = \sum_{i < N} W_i Suma\_rectangulos(r_i) \quad (1)$$

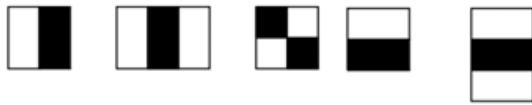


Figura 1. Rectángulos característicos del algoritmo.

Viola y Jones limitaron las características a 2,3 o 4 rectángulos. Los rectángulos deben ser adyacentes y del mismo tamaño. Las características se dividen en 3: De borde (2 rectángulos), De líneas (3 rectángulos) y de forma (4 rectángulos). Los pesos de los vectores son 1 y -1. El valor de un rasgo es el resultado de sumar los píxeles dentro rectángulo blanco y restar los píxeles dentro del rectángulo negro. Los rectángulos representan en general la región de los ojos; ya que, es más oscura que cualquier otra región de alrededor. Los rectángulos se pueden rotar 90° exceptuando la de 4, para obtener otras características; el peso de cada rectángulo también se puede invertir.

### B. La imagen Integral

La imagen integral es una representación de la imagen que puede ser calculada de manera rápida. El valor de la imagen integral  $ii(x, y)$  es igual píxel  $i(x, y)$  sumando con los píxeles de la imagen que se encuentren a la izquierda y arriba de la punto  $(x, y)$ . El cálculo de la imagen integral es iterativo y comienza en la posición (0,0). Ver ecuación (2):

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} (x', y') \quad (2)$$

La ecuación (2) hace el cálculo para un rectángulo de cualquier tamaño con 4 accesos a memoria. Para el cálculo de un rectángulo cuya esquina superior izquierda es  $(x, y)$  de tamaño  $(sx, sy)$  se accede al valor de la imagen integral de 4 esquinas como se observa en la Figura 2.

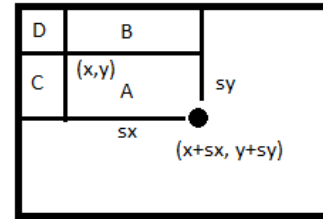


Figura 2. Cálculo de un rectángulo con la imagen integral.

Una desventaja con la imagen integral es que ocupa hasta 4 veces más memoria que la imagen original, ya que es la suma de los píxeles (ver Figura 3).

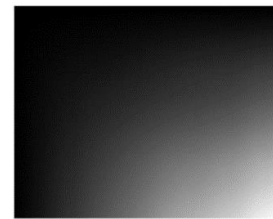


Figura 3. Representación en píxeles de la imagen integral.

### C. Entrenamiento Adaboost

En una ventana de 24x24 píxeles pueden existir hasta un total de 162.336 posibles características; Adaboosting selecciona las mejores características y construye un clasificador fuerte a partir de una combinación de clasificadores débiles. Cada clasificador presta atención donde los anteriores produjeron errores. Los clasificadores débiles deben tener como mínimo un 50% de precisión. Se utilizó el formato de color HSB para eliminar los efectos de iluminación o brillo intenso.

### D. Dilatación binaria

La dilatación ayuda a eliminar ruido. Se tiene una imagen A, y un elemento B, la dilatación se define con la ecuación (3) [8] [9]:

$$A \oplus B = \{d \in E^2: d = a + b \text{ para cada } a \in A \text{ y } b \in B\} \quad (3)$$

### E. Erosión binaria

La erosión de la imagen A con el elemento estructural B, combina dos conjuntos utilizando la sustracción de vectores. Ver ecuación (4):

$$A \ominus B = \{d \in E^2: d + b \in X \text{ para cada } b \in B\} \quad (4)$$

### F. Histograma

Para el seguimiento del ojo se toma una característica típica del mismo; por ejemplo, el color. Así, el color negro puede ser un patrón que permite identificar a la pupila. Esto se realiza

mediante un histograma que determina la variable con mayor magnitud y sus rangos.

G. Análisis Estadístico

Las ecuaciones (5), (6) y (7) se calcularon para encontrar el rango donde se localizan la mayoría de los datos, el cual está dado por la siguiente formula:

$$R = (\bar{X} - \sigma, \bar{X} + \sigma) \tag{5}$$

La media aritmética  $\bar{X}$  es el valor obtenido de sumar todos los datos de un evento y dividir entre el número total de datos. Dados los números  $a_1, a_2, \dots, a_n$ , la media aritmética es:

$$\bar{X} = \frac{\sum_{i=1}^n a_i}{n} = \frac{a_1 + a_2 + \dots + a_n}{n} \tag{6}$$

La **varianza** es la media aritmética de las desviaciones con respecto un punto medio, elevadas al cuadrado. Ecuación (2)

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} = \frac{\sum x_i^2}{n} - \mu^2 \tag{7}$$

La desviación estándar de un grupo de datos es una medida de cuánto se alejan los datos de su media.

$$\sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \tag{4}$$

H. Pruebas

Durante la realización de las pruebas se le pregunta al individuo que estado afectivo está sintiendo; y para analizar la respuesta se parte de la consideración de que al pensar es inevitable mover los ojos. Se sabe que determinadas partes del cerebro actúan para pensar, recordar, imaginar y mirar. La dirección de la mirada, en la mayoría de las veces, activa el cerebro como se muestra en la tabla 1 [17].

TABLA 1: EXPLICACIÓN DE LAS POSICIONES DE LA MIRADA

Posición de los Ojos	Significado
Ojos arriba a la izquierda	Construcción de imágenes en la mente. Uso de la imaginación para crear.
Ojos arriba a la derecha	La persona intenta visualizar la imagen que se encuentra guardada en su memoria
Ojos a la derecha	Ocurre cuando una persona intenta recordar o buscar en su memoria sonidos
Ojos a la izquierda	La persona intenta construir sonidos que conoce
Ojos para abajo y a la izquierda	Es cuando alguien se pone en contacto con un sentimiento. Por ejemplo si la persona está contenta o feliz.
Ojos para abajo y a la derecha	Es alguien que tiene un dialogo interno

Estas pautas básicas se pueden aplicar a personas diestras, pero a personas zurdas no del todo, debido al cruce de información que se lleva a cabo en el cerebro [17].

Registrado el diámetro de la pupila en cada frame del video, se grafican los resultados; son 20 imágenes por cada segundo de video. Así se obtienen una gráfica de barras como en la figura 4 y una gráfica continua como en la figura 5.

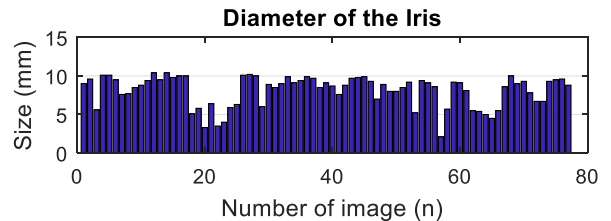


Figura 4 Grafica con barras.

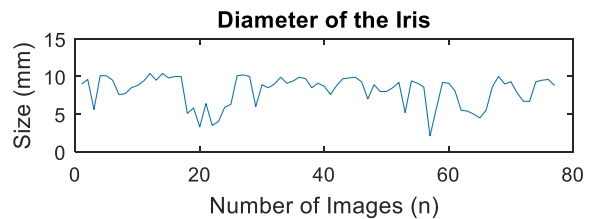


Figura 5 Grafica continua.

III. RESULTADOS

Se realizaron 100 pruebas con diferentes personas; para lo cual se graba en video al sujeto. Se obtienen 30 frames por segundo del video; los cuales son analizados uno por uno. La figura 6 muestra un frame de prueba.



Figura 6. Frame muestra

Se detecta el ojo de la persona, el cual es cortado. En la siguiente imagen de la figura 7 se puede observar dicho corte.

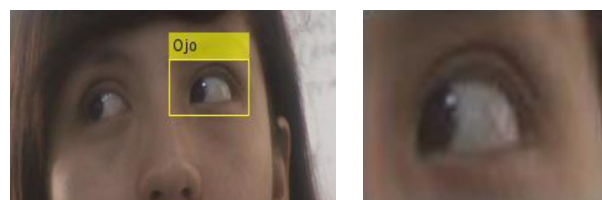


Figura 7. Detección del ojo

En la figura 8 se puede observar el frame binarizado. Las pestañas y zonas ajenas a la pupila se observa oscuras, como se ve en la figura 8 y 9; Sin embargo, estas otras superficies son anormales, mientras que la pupila es redonda. Se hace un análisis de distancias de cada píxel al fondo de la imagen. En las operaciones para aislar el iris se tiene un 1 si ese punto se considera parte de la pupila y 0 si no lo es. Así, los píxeles con mayor distancia al fondo son parte de la pupila.

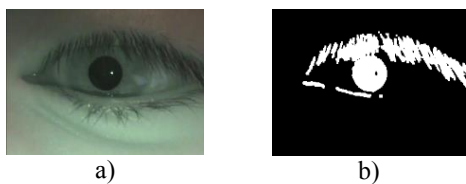


Figura 8 a) Imagen Original b) Imagen binarizada.

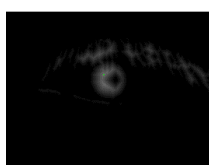


Figura 9. El color blanco representa la mayor distancia al píxel al fondo, el círculo tiene el mayor valor.

Entonces se dilata y erosiona para eliminar píxeles indeseables. Los orificios en la pupila se rellenan; y los elementos estructurantes evalúan si los píxeles tienen valor 1 o forman parte de la pupila. El proceso elimina huecos en la imagen. Ver figura 10 y 11.

	1	1		1	1	
	*	1		1	*	
	1	1		1	1	

Figura 10 Elementos Estructurantes.



Figura 11 a) Pupila sin rellenar b) Pupila rellena.

### A. Representación gráfica

En la figura 12 muestra los resultados obtenidos al medir el diámetro de una Iris. La figura 13, muestra el promedio obtenido por cada 20 frames y muestra cómo cambia el diámetro de la Iris. Es factible observar que hay diámetros muy reducidos lo puede ser una incongruencia debido a que el iris no cambia con tanta rapidez. Sin embargo, se verifico que estos casos corresponden a cuando la persona cerró el ojo, o la pupila se encuentra en alguna orilla o simplemente es un frame que no contiene un ojo. Sin embargo, afectan el análisis, por lo deben ser ignorados.

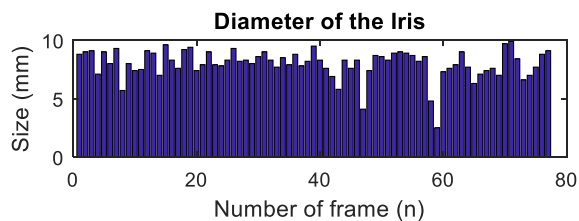


Figura 12 Grafica diámetro del Iris.

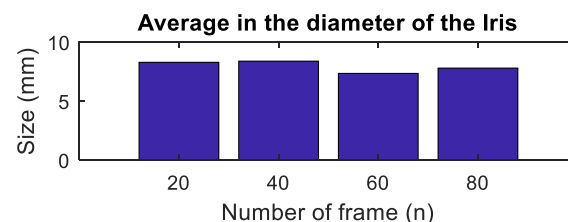


Figura 13 Promedio del diámetro

Es importante entonces también calcular la varianza y la desviación estándar. Se muestran en la tabla I encontrada después de analizar 7 videos de la misma persona (sujeto de prueba 1).

TABLE I. PARAMETROS ESTADÍSTICOS

	Video 1	Video 2	Video 3	Video 4	Video 5	Video 6	Video 7
Promedio	8.12	7.54	8	11.84	11.29	10.62	5.24
Valor Máximo	10.4	9.4	9.9	14.9	13.4	13	7.1
Valor Mínimo	2.1	3.8	2.5	5.3	6.7	6.5	3.6
Desviación estándar	1.95	1.2	1.23	1.85	1.34	1.52	0.64
Varianza	2.83	1.44	1.51	3.44	1.82	2.32	0.45
Moda	8.5	8.1	8.3	13.2	12.6	10.8	5.6
Mediana	8.9	7.85	8.2	12.1	11.5	10.8	5.4

Es importante localizar lo valor máximos y mínimos ya que pueden ser atípico y por lo tanto un problema, porque tienden a distorsionar los resultados del análisis de datos, en particular en las estadísticas descriptivas y las correlaciones. Se identifica un valor atípico cuando se alejan demasiado de la mediana. La mediana representa el valor intermedio entre los extremos de un conjunto de datos obtenido. El valor que aparece con mayor frecuencia en un conjunto de datos se calculó mediante el modo. La información provista por las medidas de tendencia central no es suficiente para caracterizar la distribución. El índice de variabilidad más conocido es la varianza.

### B. Aplicación a cómputo afectivo

Si a los sujetos de prueba se les toma un video experimentando diferentes emociones y se obtienen los rangos como los que se



muestran en la figura 14, se pueden comparar y delimitar las emociones por rango en cada individuo.

En el Sujeto de prueba 3, la emoción predominante es la alegría. En la gráfica de la figura 14, observamos que son varios los puntos que superan un tamaño de 5mm y es debido a que la persona emite una carcajada y el iris incrementa sus dimensiones.

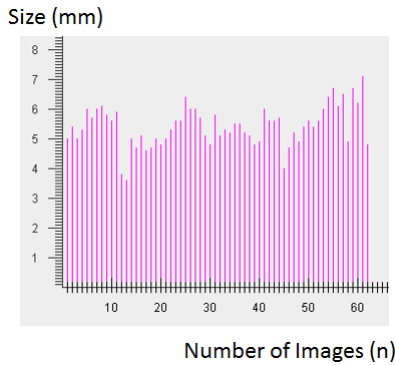


Figura 14 Grafica-1-sujeto-3.

### C. Aplicación Medición de diámetro para el estudio de enfermedades

La aplicación puede determinar el diámetro de la pupila si se tiene la distancia focal de la cámara y de esta manera apoyar al médico en el estudio de alguna enfermedad. Ver Figura 15.

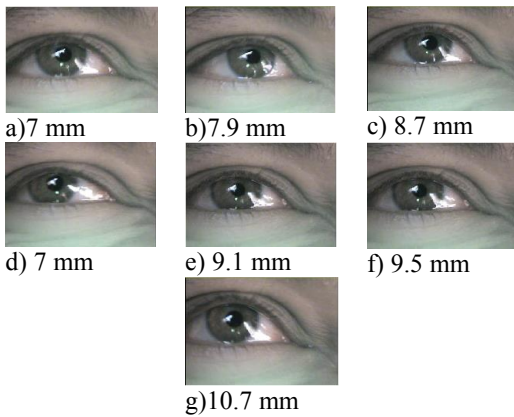


Figura 15. Imágenes de puntos altos.

### IV. CONCLUSIÓN

El iris registra cambios en el tamaño; es decir se genera una dilatación cuando es afectado por un agente externo o algún padecimiento. Por lo que es factible encontrar información que nos indique el estado de ánimo de una persona o alguna enfermedad cuando se mide algunas características del ojo, de ahí la importancia del software desarrollado.

Como trabajo a futuro se tiene contemplado observar el ojo y generar una base de datos con las imágenes tomadas.

### RECONOCIMIENTOS

Al IPN (Instituto Politécnico Nacional); ESIME (Escuela Superior de Ingeniería Mecánica y Eléctrica) y ESCOM (Escuela Superior de Cómputo); por el apoyo recibido para la realización del proyecto de investigación.

### REFERENCIAS

- [1] Ghazali, K.H., Jadin, M.S., Jie, M., Xiao, R. Novel automatic eye detection and tracking algorithm (2015) *Optics and Lasers in Engineering*, 67, pp. 49-56. Cited 9 times. DOI: 10.1016/j.optlaseng.2014.11.003 Document Type: Article
- [2] Larson, M.D., Behrends, M. Portable infrared pupillometry: A review (2015) *Anesthesia and Analgesia*, 120 (6), pp. 1242-1253. Cited 29 times. DOI: 10.1213/ANE.0000000000000314 Document Type: Review Source: Scopus
- [3] Kim, H., Jo, J., Toh, K.-A., Kim, J. Eye detection in a facial image under pose variation based on multi-scale iris shape feature (2017) *Image and Vision Computing*, 57, pp. 147-164. Cited 1 time. DOI: 10.1016/j.imavis.2016.10.003 Document Type: Article
- [4] A literature survey on robust and efficient eye localization in real-life scenarios (2013) *Pattern Recognition*, 46 (12), pp. 3157-3173. Cited 36 times. DOI: 10.1016/j.patcog.2013.05.009 Document Type: Article
- [5] Raghuvveera, T., Vidhushini, S., Swathi, M. Comparative study of CAMSHIFT and RANSAC methods for face and eye tracking in real-time video (2017) *International Journal of Intelligent Information Technologies*, 13 (2), pp. 63-75. Cited 1 time. DOI: 10.4018/IJIT.2017040104 Document Type: Article
- [6] Kong, W., Zhou, L., Wang, Y., Zhang, J., Liu, J., Gao, S. A system of driving fatigue detection based on machine vision and its application on smart device (2015) *Journal of Sensors*, 2015, art. no. 548602, . Cited 4 times. DOI: 10.1155/2015/548602 Document Type: Article
- [7] El Kaddouhi, S., Saaidi, A., Abarkan, M. Eye detection based on the Viola-Jones method and corners points (2017) *Multimedia Tools and Applications*, 76 (21), pp. 23077-23097. Cited 2 times. DOI: 10.1007/s11042-017-4415-5 Document Type: Article
- [8] E. Olmedo *et al.*, "A parallel approach for the training stage of the Viola-Jones face detection algorithm," *Intelligent Data Analysis*, vol. 21, (5), pp. 1097-1115, 2017.
- [9] Soetedjo, A., Somawirata, I.K. Implementation of eye detection using dual camera on the embedded system (2017) *International Journal of Innovative Computing, Information and Control*, 13 (2), pp. 397-409. Cited 1 time. Document Type: Article
- [10] Zhang, W., Smith, M.L., Smith, L.N., Farooq, A. Gender and gaze gesture recognition for human-computer interaction (2016) *Computer Vision and Image Understanding*, 149, pp. 32-50. Cited 11 times. DOI: 10.1016/j.cviu.2016.03.014 Document Type: Article.
- [11] Utama, P., Widodo, Ajie, H. A framework of human emotion recognition using extreme learning machine (2015) *Proceedings - 2014 International Conference on Advanced Informatics: Concept, Theory and Application, ICAICTA 2014*, art. no. 7005961, pp. 315-320. Cited 2 times. DOI: 10.1109/ICAICTA.2014.7005961 Document Type: Conference Paper
- [12] Liu, X., Graeser, A. Robust Face Detection with Eyes Occluded by the Shadow from Dazzling Avoidance System (2015) *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2015-October*, art. no. 7313472, pp. 2352-2357. Cited 3 times. DOI: 10.1109/ITSC.2015.380 Document Type: Conference Paper
- [13] Du, S., Liu, J., Liu, Y., Zhang, X., Xue, J. Precise glasses detection algorithm for face with in-plane rotation (2017) *Multimedia Systems*, 23 (3), pp. 293-302. Cited 2 times. DOI: 10.1007/s00530-015-0483-4 Document Type: Article
- [14] Xu, M., Ren, Y., Wang, Z. Learning to predict saliency on face images (2015) *Proceedings of the IEEE International Conference on Computer Vision, 2015 International Conference on Computer Vision, ICCV 2015*, art. no. 7410802, pp. 3907-3915. Cited 13 times. DOI: 10.1109/ICCV.2015.445 Document Type: Conference Paper
- [15] P. Viola *et al.*, "Robust real-time face detection," in 2001, . DOI: 10.1109/ICCV.2001.937709.

[16] Lerma, Joaquin Planells, Implementación del algoritmo de detección facial de Viola-Jone, 2009.

[17] Baró, T. La gran guía del lenguaje no verbal. Cómo aplicarlo en nuestras relaciones para lograr el éxito y la felicidad, Paidós, Barcelona, Buenos Aires, México, 2012.

# Detección de raíz mediante técnicas de visión artificial

López Guzmán,  
Guadalupe Aracely  
Instituto Tecnológico  
Nacional de México  
ITLP  
La Paz, B.C.S, México

Martínez Díaz,  
Saúl  
Instituto Tecnológico  
Nacional de México  
ITLP  
La Paz, B.C.S, México

Gutiérrez Jagüey,  
Joaquín  
Centro de Investigaciones  
Biológicas del Noroeste  
La Paz, B.C.S, México

Santillán Méndez,  
Israel Marcos  
Instituto Tecnológico  
Nacional de México  
ITLP  
La Paz, B.C.S, México

**Resumen**—Se presentan tres técnicas de segmentación para llevar a cabo la detección de raíz en una imagen capturada de la zona radicular de un cultivo. Las técnicas implementadas son segmentación por valor umbral global, segmentación por valor umbral por regiones y segmentación implementando el método de Wellner y haciendo uso de la imagen integral.

**Index Terms**—Sensores, Raíces, Procesamiento digital de imágenes, Segmentación de imágenes, Operadores Morfológicos.

## I. INTRODUCCIÓN

Se conoce como agricultura al conjunto de técnicas utilizadas en el cultivo de la tierra, considerada como una de las actividades económicas más importantes para distintas zonas a nivel mundial. A lo largo de la historia, ha ido evolucionando de tal manera que pueda responder a las distintas necesidades de optimización de calidad, producción, minimización del impacto ambiental y formas de cultivo, entre otros. Un ejemplo de esto es la Agricultura de Precisión (AP), que es la práctica agrícola centrada en la obtención de información detallada que permita adoptar soluciones adecuadas a cada cuestión particular [1]. La AP ayuda a mejorar distintos elementos clave entre los que destaca la supervisión de propiedades del suelo mediante la colocación de sensores.

Los sistemas de control automatizado de invernaderos se encargan de medir múltiples variables por medio de sensores desplegados en diferentes puntos dentro del invernadero. Los datos enviados por los sensores son almacenados, por lo general, en archivos de texto llano como registros operacionales; sin embargo, los agricultores y los investigadores carecen de herramientas que les permitan analizar y comprender estos registros [2].

Desde el punto de vista de la agricultura, el suelo constituye la principal reserva de agua para el crecimiento de las plantas cuyo objetivo consiste en mantener la vegetación en óptimas condiciones, asegurando una buena apariencia y un crecimiento adecuado. Es por esto que el conocimiento del contenido hídrico en el suelo es fundamental para determinar los momentos óptimos en que se deben realizar los riegos, así como la cantidad de agua requerida.

Existen distintas técnicas para determinar el contenido de agua en el suelo, entre las que destaca la estimación del

contenido volumétrico de agua en el suelo. Este contenido de agua se expresa como un porcentaje con base en el peso seco del suelo y el peso húmedo.

Los sensores en los sistemas de riego han cumplido un papel muy importante. Estos se clasifican de la siguiente manera: sensores de humedad, radiación solar, temperatura, entre otros.

En la actualidad existen sistemas de control de riego disponibles en el mercado y los cuales están compuestos por una red de sensores (WSN, *Wireless Sensor Network*) [3].

Existen distintas líneas de investigación cuyo objetivo es desarrollar e innovar tecnologías basadas en los principios de la AP. Un ejemplo de ello son los sistemas de apoyo a la toma de decisiones (DSS, *Decision Support System*) cuyo objetivo es la obtención, representación, almacenamiento y análisis de datos recolectados.

### I-A. Sensor de Irrigación basado en un dispositivo móvil

El sistema de irrigación basado en un dispositivo móvil o *Smartphone Irrigation Sensor* es un ejemplo de un DSS. El sensor de irrigación se configura como una unidad dentro del Sistema de Riego Automatizado (SRA), el cual consiste en dos componentes: Unidades de sensores inalámbricos (WSU, *Wireless Sensor Unit*) y unidad de información inalámbrica (WIU, *Wireless Information Unit*), enlazados mediante transceptores ZigBee que implementan una red de sensores inalámbrica (WSN, *Wireless Sensor Network*). El WIU posee un módulo GPRS para la transmisión de los datos hacia un servidor web [4].

El sensor está compuesto por un dispositivo móvil, control de iluminación basado en un LED de alta intensidad y un microcontrolador, los cuales están encapsulados en un contenedor hermético ubicado bajo tierra. El dispositivo móvil, mediante una aplicación, se encarga de llevar a cabo la captura de una imagen del suelo a través de un vidrio anti reflejante para estimar el riego requerido. El procesamiento de la imagen consiste en una transformación a escala de grises para posteriormente estimar el porcentaje de suelo húmedo relativo (RWS, *Relative Wet Soil*), el cual se basa en la diferenciación de píxeles de la imagen producida por diversos contenidos de agua en el suelo [5][6].

## II. DESCRIPCIÓN DEL PROBLEMA Y PROPUESTA DE SOLUCIÓN

El sensor de irrigación basado en un dispositivo móvil es perfectible en sus componentes físicos y lógicos para ampliar su uso y confiabilidad. A pesar de que las condiciones de funcionamiento son de tal manera que el sensor opere en un ambiente controlado, como son iluminación, suelo tamizado y uniforme, así como una barrera contra aparición de raíces, entre otras; las raíces siempre encuentran una manera de llegar a la zona de captura de la imagen. Esto trae consigo que la estimación del porcentaje de suelo húmedo relativo se vea alterada por la aparición de este nuevo elemento no contemplado en el algoritmo.

En la Figura 1a se observa la imagen obtenida en una etapa temprana del cultivo, mientras que en la Figura 1b se observa la misma región con la aparición de raíces.

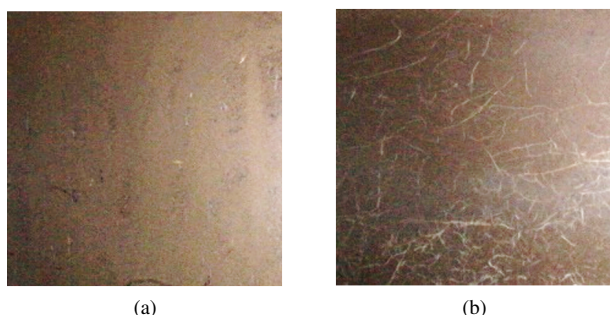


Fig. 1: (a) Imagen sin raíces e (b) imagen con raíces.

La problemática que se presenta en este trabajo es que una vez que se haya realizado la adquisición de la imagen, se debe evaluar y estimar el porcentaje de suelo húmedo. Sin embargo, a medida que la raíz crece, ésta busca absorber la mayor cantidad de nutrientes y agua, por lo que su presencia en la imagen capturada es inevitable, causando ruido en la estimación.

En este sentido, se propone la implementación de algoritmos que permitan estimar de forma robusta el área de la raíz mediante técnicas de segmentación, permitiendo así discriminar dicho elemento.

## III. FUNDAMENTOS

### III-A. Segmentación

Identificar un objeto particular dentro de una imagen es una tarea importante en algunas aplicaciones. La segmentación es el proceso de dividir una imagen en un conjunto de objetos y fondos [7]. El atributo más básico para la segmentación en una imagen monocromática es la amplitud de luminancia y los componentes de color para una imagen a color. Los bordes y la textura de la imagen también son atributos útiles para la segmentación [8].

Una iluminación poco uniforme es uno de los principales problemas asociados con el procesamiento digital de imágenes. Y es por eso que para abordar dicha problemática existen

distintas técnicas que permiten realizar el proceso de segmentación de una manera deseable.

### III-B. Método del valor umbral

El método de valor umbral es un método de segmentación que consiste en separar regiones de una imagen con base en la variación de la intensidad de los píxeles.

El método de valor umbral realiza una reducción de información en la que los únicos valores posibles son (0 y 1) y se consigue comparando su nivel de gris  $f(x, y)$  con cierto valor umbral  $t$  preestablecido, de tal forma que:

$$H(x, y) = \begin{cases} 0 & \text{si } f(x, y) < t \\ 1 & \text{si } f(x, y) \geq t \end{cases} \quad (1)$$

La selección de un umbral  $t$  es una tarea importante en el proceso de segmentación y su valor puede ser determinado de manera manual o puede ser el resultado de un método automatizado de selección de umbral óptimo.

### III-C. Umbral local

El umbral local es un método donde el valor umbral es calculado por regiones  $R_i$  y por lo tanto existirán distintos umbrales  $t_i$  para cada región, de tal forma que:

$$H(x, y) = \begin{cases} 0 & \text{si } f(x, y) < t_i \quad \forall (x, y) \in R_i \\ 1 & \text{si } f(x, y) \geq t_i \end{cases} \quad (2)$$

El número de regiones es considerado un factor importante para el tiempo de ejecución. Sin embargo, cuanto mayor sea este valor y mejor sea el método de selección de umbral, los resultados serán considerablemente mejores.

### III-D. Método de Otsu

El método de Otsu [9], es una estrategia de elección del valor umbral que implementa técnicas estadísticas para resolver el problema de selección del valor umbral óptimo.

El método de Otsu calcula el valor umbral óptimo de forma que la dispersión de cada región sea lo más pequeña posible y al mismo tiempo la dispersión entre regiones sea lo más grande posible, es decir, un umbral que produzca la mejor separación entre clases, la cual se encuentra definida por (3).

$$\sigma_\omega^2 = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t) \quad (3)$$

Si se supone que los píxeles de la imagen se agrupan en dos clases  $C_0$  y  $C_1$  (objeto y fondo de la imagen, o viceversa), mediante un umbral de nivel  $t$ , se tiene que la clase  $C_0$  se corresponde con los píxeles con niveles  $[0 \dots t]$  y la clase  $C_1$  se corresponde con los píxeles con niveles  $[t + 1, \dots, L - 1]$ . Por tanto, la probabilidad de que un píxel sea asignado a cada una de estas clases viene dada por:

$$\omega_0 = Pr(C_0) = \sum_{i=1}^k p_i = \omega(k) \quad (4)$$

$$\omega_1 = Pr(C_1) = \sum_{i=k+1}^L p_i = 1 - \omega(k) \quad (5)$$

Y los valores de las medias de las intensidades correspondientes a las clases  $C_0$  y  $C_1$  son:

$$\mu_0 = \sum_{i=1}^k iPr(i|C_0) = \sum_{i=1}^k \frac{iP_i}{\omega_0} = \frac{\mu(k)}{\omega(k)} \quad (6)$$

$$\mu_1 = \sum_{i=k+1}^L iPr(i|C_1) = \sum_{i=k+1}^L \frac{iP_i}{\omega_1} = \frac{\mu_T - \mu(k)}{1 - \omega(k)} \quad (7)$$

Donde (8) y (9) son la suma de las intensidades e intensidad promedio hasta el nivel  $k$  respectivamente de la imagen, y (10) es la intensidad media global de la imagen original.

$$\omega(k) = \sum_{i=1}^k p_i \quad (8)$$

$$\mu(k) = \sum_{i=1}^k ip_i \quad (9)$$

$$\mu_T = \mu(L) = \sum_{i=1}^k ip_i \quad (10)$$

### III-E. Imagen integral

Una imagen integral es una herramienta que puede ser implementada cuando se tenga una función de píxeles a números reales  $f(x, y)$  y se desea calcular la suma de esta función sobre una región rectangular de la imagen. Sin una imagen integral, la suma se puede calcular en tiempo lineal por rectángulo, calculando el valor de la función para cada píxel individualmente. Sin embargo, si se necesita calcular la suma en varias ventanas rectangulares superpuestas, es posible hacer uso de la imagen integral y lograr un número constante de operaciones por rectángulo con solo una cantidad lineal de preprocesamiento [9].

Para realizar el cálculo de la imagen integral, el valor de un píxel con coordenadas  $I(x, y)$  representa la suma de todos los píxeles arriba y a la izquierda de dicha posición y se calcula mediante (11).

$$I(x, y) = \sum_{i=0}^x \sum_{j=0}^y f(i, j) \quad (11)$$

En la Figura 2 se muestra la imagen integral donde la Figura 2a representa la imagen de entrada mientras que la Figura 2b representa la imagen resultante. El valor resaltado en la imagen integral es el resultado de realizar la suma consecutiva de los valores resaltados en la imagen de entrada.

3	8	2	4	3	11	13	17
1	7	3	0	4	19	24	28
4	2	7	6	8	25	37	47
1	8	5	2	9	34	51	63

(a)

(b)

Fig. 2: (a) Imagen de entrada e (b) imagen integral resultante.

Con la imagen integral, la suma de cada uno de los términos de  $f(x, y)$  que se encuentra en la región dada por los puntos  $(x_1, y_1)$  y  $(x_2, y_2)$  es almacenada en  $I(x, y)$  y puede ser calculada mediante (12).

$$I(x, y) = f(x, y) + I(x - 1, y) + I(x, y - 1) - I(x - 1, y - 1) \quad (12)$$

Asimismo, la suma del área de interés dada por los puntos  $(x_1, y_1)$  y  $(x_2, y_2)$  es estimada con (13).

$$\sum_{x=z_1}^{x_2} \sum_{y=y_1}^{y_2} f(x, y) = I(x_2, y_2) - I(x_2, y_2 - 1) - I(x_1 - 1, y_2) + I(x_1 - 1, y_1 - 1) \quad (13)$$

En la Figura 3 se observa que realizar la suma del área de interés resaltada en la Figura 3a es equivalente a realizar las operaciones de los elementos resaltados en la Figura 3b. Es decir, la suma de los elementos de la región de interés entre los puntos  $f(1, 1)$  y  $f(2, 2)$  de la Figura 3a es equivalente a realizar la suma de los elementos  $I(2, 2) - I(2, 0) - I(0, 2) + I(0, 0)$ .

3	8	2	4	3	11	13	17
1	7	3	0	4	19	24	28
4	2	7	6	8	25	37	47
1	8	5	2	9	34	51	63

(a)

(b)

Fig. 3: (a) Imagen de entrada e (b) imagen integral resultante.

### III-F. Método de Wellner

El método de Wellner [10] tiene como objetivo que cada píxel sea comparado contra el promedio de sus píxeles vecinos. Específicamente, un promedio aproximado de los últimos píxeles  $s$  observables al recorrer la imagen. Si el valor del píxel actual es  $t$  por ciento menor que el promedio, entonces se etiqueta como parte del objeto o *foreground*, de lo contrario es etiquetado como parte del fondo o *background*, de tal forma que:

$$T(n) = \begin{cases} 0 & \text{si } f(x, y) * c \leq f_s(n) * (k - t) \\ 1 & \text{de lo contrario} \end{cases} \quad (14)$$

Donde  $c$  representa el total de píxeles de la región de interés y  $f_s(n)$  es estimado como la suma de los  $N$  píxeles dentro del área de interés dada por los puntos  $(x_1, y_1)$  y  $(x_2, y_2)$ , obtenida mediante (15).

$$f_s(n) = \sum_{i=1}^w \sum_{j=1}^h f(x_2, y_2) - I(x_2, y_1 - 1) - I(x_1 - 1, y_2) + I(x_1 - 1, y_1 - 1) \quad (15)$$

siendo  $I$  la imagen integral,  $w$  y  $h$  el ancho y largo de la imagen respectivamente y los índices  $x_1$ ,  $x_2$ ,  $y_1$  y  $y_2$  calculados como:

$$\begin{aligned} x_1 &= i - \frac{s}{2} \\ y_1 &= j - \frac{s}{2} \\ x_2 &= i + \frac{s}{2} \\ y_2 &= j + \frac{s}{2} \end{aligned} \quad (16)$$

Para la Ecuación (14) y (16), Wellner propone utilizar un valor de  $s = \frac{1}{8}$  del ancho de la imagen y  $t = 0,15$ .

El método de Wellner realiza la comparación entre los píxeles más cercanos lo cual permite conservar las líneas que poseen fuerte contraste e ignorar aquellos elementos que posean cambios de gradiente suave.

#### IV. OPERADORES MORFOLÓGICOS

Se denomina mejora de la imagen al proceso de mejorar la calidad de una imagen para que sea más adecuada para algunas aplicaciones [7]. En esta fase se suele implementar el tratamiento morfológico.

El procesamiento de imágenes morfológicas es una colección de operaciones no lineales relacionadas con la forma o morfología de las características en una imagen.

##### IV-A. Erosión y dilatación

Las operaciones morfológicas básicas son erosión y dilatación. En la erosión se parte de la idea de probar sobre una imagen  $A$  una forma predefinida llamada elemento estructurante  $B$  (subconjunto del espacio o de la cuadrícula  $A$ ). La erosión está definida por (17).

$$A \ominus B = \{z \in E | B_z \subseteq A\} \quad (17)$$

Cuando los elementos del espacio sean menores que el elemento estructurante, serán eliminados. En la Figura 4 se puede observar la implementación de la función erosión en

una imagen con ruido (Fig. 4a) obteniendo como resultado una imagen más limpia (Fig. 4b).

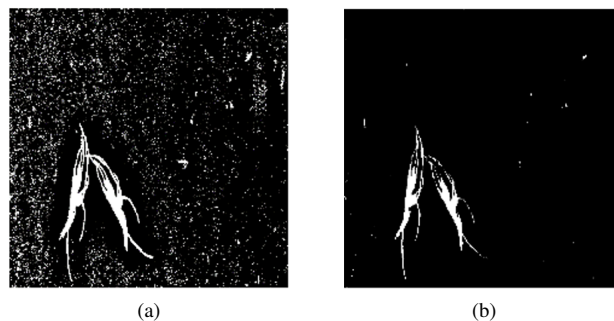


Fig. 4: (a) Imagen sin erosión e (b) imagen aplicando erosión con un elemento estructurante de línea.

El operador de dilatación usualmente implementa un elemento estructurante para expandir las formas contenidas en la imagen de entrada. En la Figura 5 se puede observar el resultado de aplicar el operador de dilatación a una imagen de entrada.

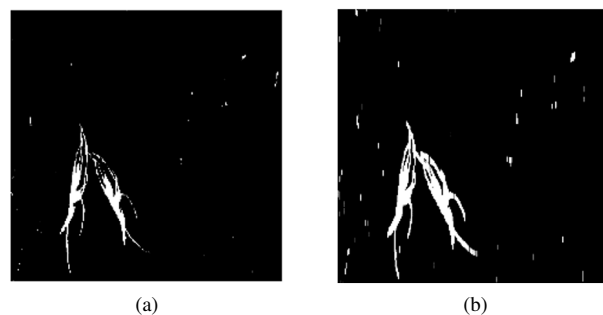


Fig. 5: (a) Imagen con erosión e (b) imagen aplicando dilatación con un elemento estructurante de línea.

La dilatación del elemento estructurante  $B$  sobre la imagen binaria  $A$  está definida por (18).

$$A \oplus B = \{z \in E | B_z \cap A \neq \emptyset\} \quad (18)$$

#### V. RESULTADOS EXPERIMENTALES

Para llevar a cabo el reconocimiento de la raíz en una imagen capturada, se probaron los algoritmos de segmentación implementando técnicas de selección de umbral global y umbral local. El desarrollo e implementación del algoritmo se llevó a cabo en el lenguaje C++ haciendo uso de la librería OpenCV.

Se realizaron pruebas con imágenes de dimensiones de 1100x1100 píxeles, obtenidas en un cultivo de calabacita tipo grey zucchini en un invernadero experimental implementando el sensor de irrigación basado en un dispositivo móvi. El tiempo total transcurrido entre la instalación del sensor de irrigación y la cosecha final de frutos duró aproximadamente 2 meses [6].



En la Figura 6 se puede observar las imágenes obtenidas por el sensor de irrigación en distintas etapas.

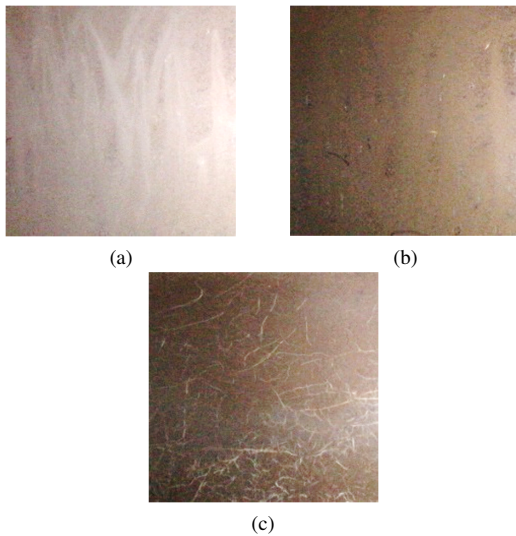


Fig. 6: (a) Imagen de zona radicular con suelo seco, (b) suelo húmedo y (c) suelo húmedo con aparición de raíz.

Para la segmentación mediante la técnica de valor umbral se implementaron las técnicas de segmentación con umbral global y segmentación con umbral local o por regiones. En la Figura 7 se observan el resultado de segmentar una imagen de entrada (Fig. 7a) con un valor umbral fijo a  $T_{global} = 125$  y la selección de un valor umbral implementando el método de Otsu (Fig. 7c).

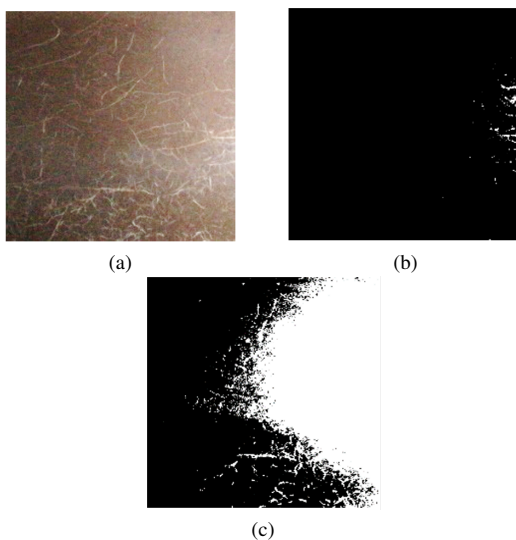


Fig. 7: (a) Imagen de entrada, (b) segmentación por valor umbral  $T_{global} = 125$  y (c) umbral global implementando el umbral óptimo por método de Otsu.

Asimismo, se realizó la segmentación de la imagen mediante la implementación de un valor umbral adaptativo por

regiones. En la Figura 8a-8c se observa el resultado de la segmentación implementado una región cuadrada de tamaño 71, 315 y 515 respectivamente.

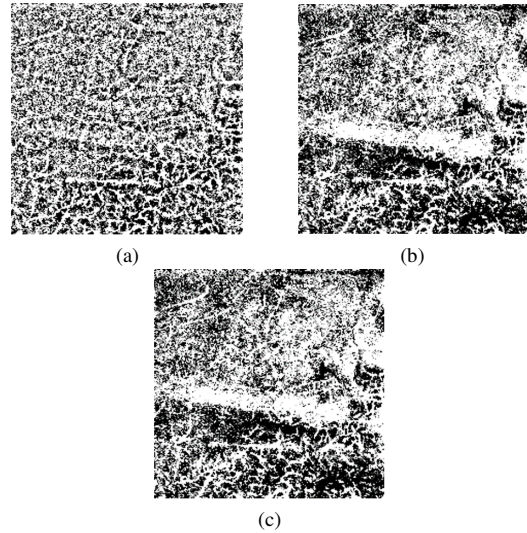


Fig. 8: Umbral global implementando el umbral óptimo por método de Otsu, (a) segmentación implementando umbral adaptativo con regiones de tamaño 71, (b) tamaño 315 y (c) tamaño 515.

Para la segmentación mediante la técnica del método de Wellner con la imagen integral se obtuvieron los resultados de la Figura 9a-9c utilizando distintos valores de  $k$ . Al finalizar la segmentación, se implementó el operador de erosión para eliminar regiones que no pertenecen al objeto.

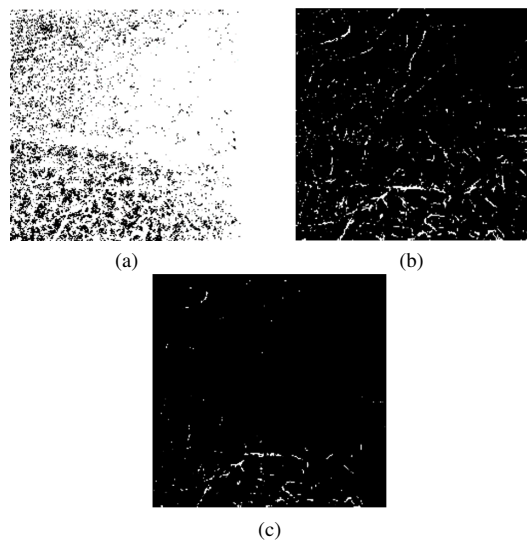


Fig. 9: Segmentación adaptativa con imagen integral y método de Wellner con un valor de (a)  $k = 1,0$ , (b)  $k = 1,25$  y (c)  $k = 1,35$ .

Para evaluar el desempeño de los métodos presentados, se probó un *DataSet* compuesto por un total de 152 imágenes



obtenidas en campo. En la Tabla I se presentan los porcentajes de precisión y tiempos de ejecución para cada método.

Método	Parámetros	Resultados
Umbral global	Umbral = 125	Porcentaje = 55.79 % Tiempo = 185 ms
	Umbral = Otsu	Porcentaje = 67.46 % Tiempo = 222 ms
Umbral local	kernel = 71	Porcentaje = 65.71 % Tiempo = 390 ms
	kernel = 315	Porcentaje = 64.79 % Tiempo = 459 ms
	kernel = 515	Porcentaje = 64.98 % Tiempo = 543 ms
Método de Wellner	$k = 1.0$	Porcentaje = 55.79 % Tiempo = 1068 ms
	$k = 1.25$	Porcentaje = 91.75 % Tiempo = 1112 ms
	$k = 1.35$	Porcentaje = 95.26 % Tiempo = 1109 ms

Tabla I: Porcentajes y tiempos de ejecución.

## VI. CONCLUSIONES

Se presentaron tres métodos de segmentación para detectar los objetos fondo y raíz. La segmentación con el método de valor umbral global arrojó resultados donde no es posible diferenciar los objetos contenidos en la imagen. Por otro lado, la segmentación con la selección de un valor umbral por regiones permite diferenciar la raíz, pero con una cantidad muy significativa de ruido. Los resultados experimentales muestran que al implementar la segmentación utilizando el método de Wellner y la imagen integral, se obtiene como resultado una imagen más clara en la cual es posible diferenciar de una mejor manera la raíz y el fondo. Asimismo, los porcentajes de precisión fueron considerablemente mejores respecto a los resultados de los otros métodos.

En conclusión, es posible observar que la imagen de entrada (Fig. 7a) presenta una iluminación que no es uniforme. Sin embargo, con base en la Tabla I, el método que presenta mayor problemática es el método por valor umbral local a pesar de que este es un método adaptativo. Por otro lado el método de Wellner presenta los mejores resultados con un porcentaje de 95.26 % con  $k = 1,35$ .

## REFERENCIAS

- [1] J. M. G. y. J. A. V. Manuel Pérez-Ruiz, Lucía Quebrajo Moya, "Introducción a la agricultura de precisión en el valle del Guadalquivir," *CAPE Agricultura de Precisión*, 2015.
- [2] C. A. Cobos, J. A. Timana, and R. V. Vallejo, "Sistema de soporte para la toma de decisiones para proceso de germinación y cultivo en invernaderos." *Revista de Biotecnología en el Sector Agropecuario*, vol. 6, no. 1, pp. 22–31, 2015.
- [3] U. D. of the Interior, B. of Reclamation, L. C. R. And, and T. S. Center, "Weather-and Soil Moisture-Based Landscape Irrigation Scheduling Devices Technical Review Report – 5th Edition," Tech. Rep., 2015.
- [4] J. Gutierrez, J. F. Villa-Medina, A. Nieto-Garibay, and M. A. Porta-Gandara, "Automated Irrigation System Using a Wireless Sensor Network and GPRS Module," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 1, pp. 166–176, jan 2014.
- [5] J. G. Jaguey, J. F. Villa-Medina, A. Lopez-Guzman, and M. A. Porta-Gandara, "Smartphone Irrigation Sensor," *IEEE Sensors Journal*, vol. 15, no. 9, pp. 5122–5127, sep 2015.
- [6] G. A. López Guzmán, "Desarrollo de un sensor de irrigación basado en un dispositivo móvil," Tesis de Licenciatura, Intituto Tecnológico de La Paz, 2014.
- [7] M. K. Pakhira, *Digital image processing and pattern recognition*. Prentice-Hall of India, 2011.
- [8] W. K. Pratt, *Digital Image Processing*. New York, USA: John Wiley & Sons, Inc., 2001.
- [9] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, jan 1979.
- [10] D. Bradley and G. Roth, "Adaptive Thresholding using the Integral Image," *Journal of Graphics Tools*, vol. 12, no. 2, pp. 13–21, jan 2007.

# Entrenamiento de una CNN para autoconducción de vehículo prototipo con Raspberry Pi

Jorge A. Armenta\*, José L. Gómez†, Iliana Castro ‡.  
 Tecnológico Nacional de México, I.T. La Paz, La Paz, B.C.S, México.  
 armentsistem@gmail.com\*, jgomez13016@itlp.edu.mx†, icaastro@itlp.edu.mx‡.

**Resumen**—En este trabajo se presentan los resultados de entrenar una Red Neuronal Convolutiva (CNN por sus siglas en inglés) de punto-a-punto para la autoconducción de un vehículo prototipo tipo Ackerman dentro de un circuito. Los datos obtenidos a partir del entrenamiento son utilizados para realizar simulaciones virtuales de autoconducción y, posteriormente implementarlas en el vehículo prototipo.

**Palabras clave**—Red neuronal convolutiva, CNN, autoconducción, vehículo autónomo.

## I. INTRODUCCIÓN

Las aplicaciones de detección vehicular son de gran interés por su utilización en acciones enfocadas a la autoconducción, logrando disminuir, principalmente, accidentes viales, los cuales son una de las diez principales causas de muerte según la Organización Mundial de Salud [1].

Para lograr la autoconducción se requiere dotar al vehículo con dispositivos y sensores que le permitan conocer el entorno en que se desplaza, tales como sensores de proximidad, radares y cámaras. Sin embargo, también se requiere poseer hardware potente que permita el procesamiento de los datos recibidos por los dispositivos para la toma de decisiones de autoconducción.

Otro componente clave que suele implementarse en los vehículos que se conducen por sí mismos es el uso de Redes Neuronales Convolutivas. Las CNN han revolucionado el reconocimiento de patrones, dejando de lado la definición a mano de características de interés usadas para realizar la clasificación [2].

En trabajos preliminares se han implementado las CNN punto-a-punto. En [3] se presenta el proyecto vehículo autónomo DARPA (DAVE), en el cual un vehículo es entrenado a partir de datos obtenidos mediante su conducción por un usuario. Los datos de entrenamiento incluyeron el video de dos cámaras y los comandos de direccionamiento del vehículo.

La empresa NVIDIA© creó un sistema robusto para la conducción en vías públicas con reconocimiento de características específicas como marcas de carril, barandillas u otros automóviles, evitando tener que condicionar las acciones de direccionamiento, basadas en la observación de estas características. Se demostró empíricamente que las CNN pueden aprender las tareas para mantenerse dentro de un carril y carretera, además, respetando la señalización de caminos o vías, planificación del camino y control detectadas [2].

En el presente trabajo se utiliza la arquitectura propuesta en [2], realizando simulaciones virtuales y en un prototipo a

escala, que tiene como objetivo ser usado como plataforma para evaluación de algoritmos de CNN.

## II. CNN Y SU ARQUITECTURA

Las CNN son un tipo de redes neuronales de múltiples capas, que están entrenadas con una versión del algoritmo de propagación hacia atrás, como casi todas las demás redes neuronales. Lo que diferencia a las CNN de las demás redes neuronales, es su arquitectura ya que están diseñadas para reconocer patrones a partir de datos de entrada con un preprocesamiento mínimo [4]. Se puede realizar el reconocimiento de patrones con una variabilidad extrema (como los caracteres escritos a mano), con robustez a las distorsiones y transformaciones geométricas simples [3], ver Figura 1.

Frecuentemente, las CNN se implementan para procesar imágenes, relacionando la entrada (imagen), directamente con una salida (imagen o vector) dependiendo las necesidades del problema. Entre las aplicaciones más utilizada se encuentran la detección y/o categorización de objetos dentro de una imagen, la clasificación de imágenes en general y la clasificación de una escena [5]. En esta última aplicación es posible generar una ponderación para conocer si está ocurriendo algo deseado o peligroso [3] [6] [7].

El enfoque de las CNN punto-a-punto (end-to-end), permite entrenar la red de convolución desde las imágenes capturadas hasta el ángulo de direccionamiento del vehículo [8]. Esto se logra utilizando núcleos de convolución para procesar la imagen por regiones y extrayendo las características de los patrones en diferentes capas [6]. Con la obtención de características en las imágenes, se correlacionan los datos de la imagen a un vector de salida que se utiliza para la toma de decisión del direccionamiento del vehículo en un entorno controlado.

La CNN utilizada en [2] consta de una capa de normalización de la imagen, cinco capas convolutivas para extraer las características de las imágenes y cuatro capas totalmente conectadas que funcionan como clasificador, en este caso clasifica la imagen asociándolo a un ángulo de giro, que es utilizado para el direccionamiento del vehículo. Se utilizan dos tipos de convoluciones, para las primeras tres capas el paso es de  $2 \times 2$  con kernel de  $5 \times 5$  y para las últimas dos capas se utiliza un kernel de  $3 \times 3$ .

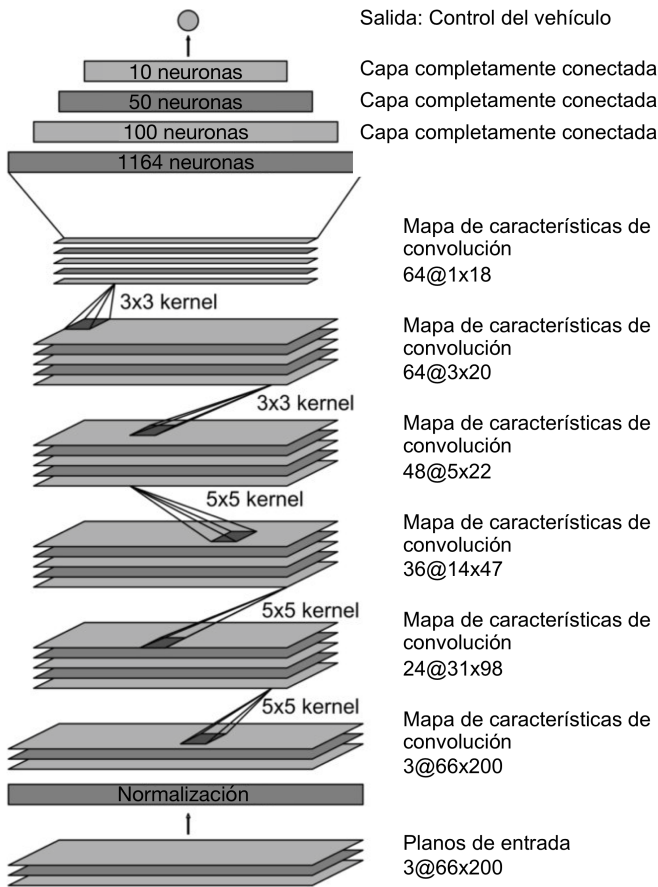


Figura 1. Arquitectura de una red neuronal convolucional.

### III. ENTRENAMIENTO DE LA CNN A PARTIR DE UN SIMULADOR DE AUTOCONDUCCIÓN

Para efectos de probar el desempeño de la CNN existe una amplia gama de plataformas de simulación. En [9] son comparadas de acuerdo a sus funcionalidades y capacidades, entre ellos se encuentran: Pelops, Prescan, Mathworks Simulink 3D Animations, RTMaps, Gazebo, SUMO, PTV Vissim, SimTram, unidad virtual de prueba Vires VTD, OpenDaVinci, CarSim y Udacity Self-driving Car Simulator.

En este trabajo se utiliza el simulador de Udacity©, que facilita una plataforma de entrenamiento para simular un vehículo virtual en un circuito [10].

El simulador tiene dos modalidades: modo entrenamiento y modo autónomo. El modo de entrenamiento permite que un usuario controle el vehículo sobre el circuito de forma supervisada, permitiendo recuperar en tiempo real datos como: perspectiva de tres cámaras montadas en el vehículo, ángulo del volante, velocidad del automóvil y activación del acelerador y del freno, así como almacenar los datos generados en una forma adecuada para su procesamiento. En el modo autónomo, se prueba el modelo de aprendizaje automático para evaluar la conducción sin salirse del camino.

#### III-A. Captura de datos

Los datos guardados para realizar el entrenamiento de la CNN son: imagen frontal, imagen lateral izquierda, imagen lateral derecha, ángulo de giro y rapidez.

Con los valores y archivos generados a partir del manejo del vehículo en el circuito, es posible generar múltiples entrenamientos con diferentes técnicas de procesamiento de datos con los cuales se obtiene la autoconducción del vehículo sin necesidad de ser operado por un usuario.



(a) Captura de imagen con traslación a la izquierda.



(b) Captura de imagen centrada.



(c) Captura de imagen con traslación a la derecha.

Figura 2. Imágenes del circuito tomadas desde la perspectiva del vehículo virtual, capturadas a partir del simulador Udacity's Self-Driving Car.

#### III-B. Entrenamiento de la CNN

Considerando la arquitectura descrita anteriormente se realizaron dos conjuntos de entrenamiento con los valores obtenidos del simulador Udacity's Self-Driving Car. Un conjunto de datos que contiene tres imágenes como entrada y otra que contiene una imagen como entrada.

Para llevar a cabo el entrenamiento se utilizaron las bibliotecas Keras y TensorFlow(desarrollada por google), ambas de código abierto. Estas bibliotecas permiten construir y entrenar redes neuronales convolucionales con una adecuada relación entre abstracción y eficiencia en el procesamiento [11]. El entrenamiento se realizó en una computadora con Sistema operativo Linux Mint 18.3 Sylvia con una tarjeta de video NVIDIA GeForce GTX TITAN X, en la cual se instaló TensorFlow versión 1.12.0, Python 3.5.2. Con cada conjunto de datos se efectuaron 10 épocas de entrenamiento.

La reducción de imágenes se realizó con el fin de probar cómo se comportaba el vehículo al entrenarlo con una sola imagen ya que el prototipo cuenta con una sola cámara. Se realizó la comparación de tiempos de entrenamiento y la rapidez de su implementación.

#### IV. PRUEBAS EN EL SIMULADOR

Una vez entrenada la red con los dos conjuntos de entrenamiento, se cargaron los pesos obtenidos en el simulador, obteniendo los datos que se presentan a continuación.

El simulador actúa como un servidor que envía como entrada a la CNN la imagen frontal del vehículo virtual y la red procesa la imagen y calcula el valor de ángulo de giro requerido para que el vehículo mantenga la trayectoria dentro del circuito, ver Figura 3. El criterio para evaluar el desempeño es el tiempo en que el vehículo completa el circuito de forma autónoma.

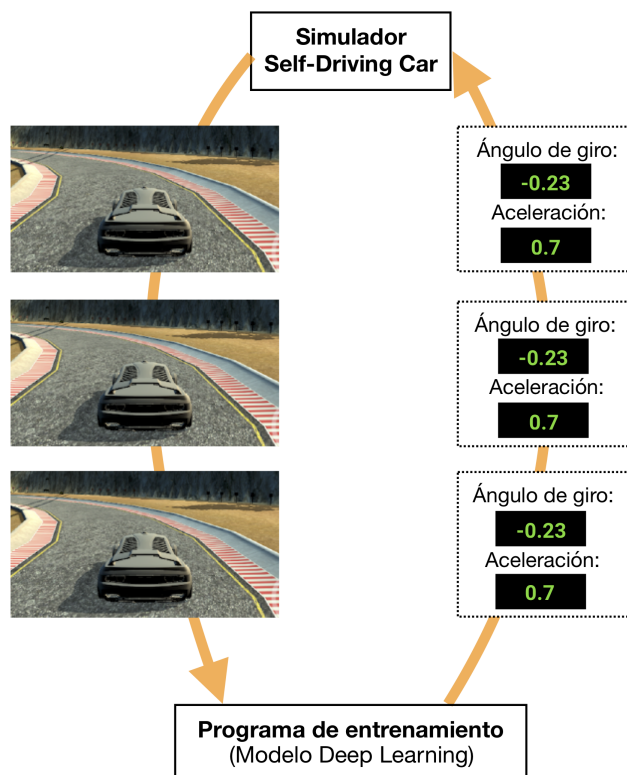


Figura 3. Metodología para la autoconducción a partir del entrenamiento de la CNN.

#### V. RESULTADOS

En las Tablas I y II se presentan los tiempos de simulación, con los dos conjuntos de datos propuestos, para tres y una imagen, mostrando las épocas 1 a la 10 en cada simulación.

Época	Simulación con tres imágenes	Terminación del circuito
1	1:55.15 s	Si
2	1:55.46 s	Si
3	1:54.13 s	Si
4	1:55.56 s	Si
5	1:54.16 s	Si
6	1:55.35 s	Si
7	1:55.53 s	Si
8	1:54.10 s	Si
9	1:55.63 s	Si
10	1:55.25 s	Si

Tabla I  
TIEMPOS DE EJECUCIÓN DE LA SIMULACIÓN DE AUTOCONDUCCIÓN EN LAS ÉPOCAS NÚMERO 1 HASTA LA 10, CON EL MODELO DE TRES IMÁGENES.

Época	Simulación con una imágenes	Terminación del circuito
1	00:32.55 s	No
2	1:55.11 s	Si
3	1:55.18 s	Si
4	1:55.26 s	Si
5	1:55.15 s	Si
6	1:55.68 s	Si
7	1:54.93 s	Si
8	1:55.64 s	Si
9	1:55.32 s	Si
10	1:55.97 s	Si

Tabla II  
TIEMPOS DE EJECUCIÓN DE LA SIMULACIÓN DE AUTOCONDUCCIÓN EN LAS ÉPOCAS NÚMERO 1 HASTA 10, CON EL MODELO DE UNA IMAGEN.

En la primera época de la simulación con el conjunto de datos de una imagen, el vehículo no logra completar el circuito, mientras que en la primera época del conjunto de tres imágenes sí efectuó el recorrido deseado completo. Para las siguientes épocas el vehículo concluye las trayectorias del circuito con tiempos de ejecución similares entre los dos modelos de entrenamiento.

#### VI. CAPTURA DE DATOS EN UN VEHÍCULO PROTOTIPO

Teniendo en cuenta que en las simulaciones es posible disminuir los datos de entrada, es decir pasar de tres imágenes a una sola, se propone la implementación en un vehículo prototipo que permite obtener los parámetros necesarios: imágenes, velocidad y ángulo de giro, para entrenar con ellos la CNN.

Para comprobar que los datos de entrenamiento son efectivos no sólo en las simulaciones, se realizará la implementación en un vehículo prototipo a escala, utilizando el *Smart Video Car Kit for Raspberry Pi* de SunFounder©, mostrado en la Figura 4. El vehículo cuenta con una cámara instalada con la cual se logra obtener las capturas de imágenes necesarias para calcular los ángulos de direccionamiento a partir de la CNN punto-a-punto entrenada con el conjunto de una imagen por cuadro.



Las características del prototipo y sus componentes son las siguientes [12]:

- Vehículo escala 1:12
- Tipo Ackerman
- Raspberry Pi 3 model B+ (1.4GHz 64-bit quad-core processor) con sistema operativo Raspbian Jessie
- 1 Micro Servo 9g SG90 TowerPro para el direccionamiento.
- 2 motores de tracción, Modelo: b48-1416LSD6



Figura 4. Vehículo prototipo *Smart Video Car Kit for Raspberry Pi* de SunFounder©.

La estructura del vehículo permite instalar otros componentes, sin embargo para no aumentar el consumo energético, se mantiene la implementación a partir de la obtención de datos monocolor.

Para la captura de datos desde el prototipo, se modificó el código utilizado por defecto en el mismo para su control. Se obtiene un conjunto de datos para entrenarse con la arquitectura descrita en [2] con una sola imagen de entrada, que se captura con la cámara instalada en el vehículo.

Se construyó una pista o circuito para manejar el vehículo prototipo y capturar los datos necesarios para su entrenamiento, ver Figura 5.

Se manejó el vehículo prototipo de forma supervisada por la pista realizando 4 vueltas, generando 745 imágenes y datos de conducción. Estos datos se procesaron y entrenaron de la misma manera que la versión de una sola imagen en el simulador. En la Figura 6 se muestra una imagen utilizada para el entrenamiento capturada desde la cámara del prototipo.

## VII. CONCLUSIONES

En las simulaciones, a partir de la época número cinco, el modelo que fue entrenado con una imagen de entrada, se comporta de manera similar al modelo con tres imágenes de entrada.

Se demostró que en un ambiente virtual es posible disminuir el número de datos para entrenar la CNN punto-a-punto, lo

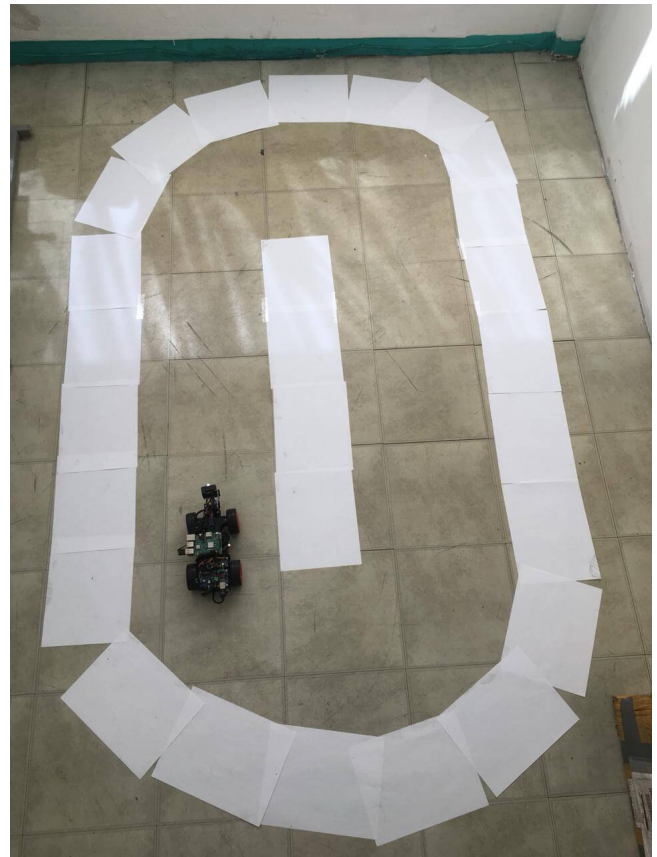


Figura 5. Circuito de entrenamiento para vehículo prototipo



Figura 6. Imagen del circuito tomada desde la perspectiva del vehículo prototipo.

que implica una reducción en el procesamiento y almacenamiento de los datos. La disminución de datos requeridos para el entrenamiento permite reducir la cantidad de dispositivos necesarios para la autoconducción del vehículo prototipo.

Se realizó el entrenamiento con datos de entrada prove-

nientes del prototipo. Como trabajo futuro se propone la implementación de los pesos obtenidos en el mismo.

#### REFERENCIAS

- [1] ORGANIZACIÓN MUNDIAL DE LA SALUD. Educación CTIM. Consultado el 16-Abril-2019. URL:[www.who.int/es/news-room/fact-sheets/detail/the-top-10-causes-of-death](http://www.who.int/es/news-room/fact-sheets/detail/the-top-10-causes-of-death)
- [2] BOJARSKI, M., DEL TESTA, D., DWORAKOWSKI, D., FIRNER, B., FLEPP, B., GOYAL, P., ZHANG, X., *et. al.* (2016). *End to end learning for self-driving cars*. arXiv preprint arXiv:1604.07316.
- [3] LECUN, Y., COSATTO, E., BEN, J., MULLER, U., FLEPP, B. (2004). *DAVE: Autonomous off-road vehicle control using end-to-end learning*. DARPA-IPTO Final Report.
- [4] LI, F. F., KARPATHY, A., JOHNSON, J. (2015). *CS231n: Convolutional neural networks for visual recognition*. University Lecture.
- [5] GOODFELLOW, I., BENGIO, Y., COURVILLE, A. (2016). *Deep learning*. MIT press.
- [6] KIM, J., CANNY, J. (2017). *Interpretable learning for self-driving cars by visualizing causal attention*. In Proceedings of the IEEE international conference on computer vision (pp. 2942-2950).
- [7] KRIZHEVSKY, A., SUTSKEVER, I., HINTON, G. E. (2012). *ImageNet classification with deep convolutional neural networks*. Advances in Neural Information Processing Systems (pp. 1097-1105).
- [8]
- [9] GRAZIOLI, F., KUSMENKO, E., ROTH, A., RUMPE, B., VON WENCKSTERN, M. (2017). *Simulation framework for executing component and connector models of self-driving vehicles*. In MODELS (Satellite Events) (pp. 109-115).
- [10] UDACITY. Public driving dataset. Consultado el 16-Abril-2019. URL:<https://www.udacity.com>.
- [11] ZOCCA, V., SPACAGNA, G., SLATER, D., ROELANTS, P. (2017). *Python Deep Learning*. Packt Publishing Ltd.
- [12] SUNFOUNDER (2012). Educación CTIM. Consultado el 16-Abril-2019. URL:<https://www.sunfounder.com/rpi-car.html>.



# VANT y fotogrametría digital para evaluar la morfodinámica costera

Germán Rojas-Silva, Guillermo Martínez-Flores, Enrique H. Nava-Sánchez  
Instituto Politécnico Nacional, Centro Interdisciplinario de Ciencias Marinas  
Departamento de Oceanología  
La Paz, B. C. S., México  
german.rojas@gmail.com, gmflores@ipn.mx, enava@ipn.mx

**Resumen**—Con la finalidad de estimar los cambios en la morfología costera en la barrera arenosa del estero San José, Baja California Sur, se realizaron cinco levantamientos topográficos durante un año a partir del 24/feb/2017, utilizando un VANT DJI Phantom 4 Pro y un GPS diferencial Leica Viva GS14. Las imágenes obtenidas fueron sometidas a un proceso fotogramétrico para generar los cinco Modelos Digitales del Terreno (MDT) con resolución espacial de 8 cm. Mediante la sustracción de MDT consecutivos se calculó la distribución espacio-temporal de los cambios volumétricos. Los mayores cambios se presentaron entre los levantamientos del 26/may y el 6/sep de 2017, con una erosión de 155,993 m<sup>3</sup> y acreción de 190,998 m<sup>3</sup>. Estos volúmenes representan los cambios que experimentó la barrera por la incidencia de la Tormenta Tropical Lidia el 1/sep de 2017. La descarga extraordinaria del arroyo San José, ocasionó una ruptura en la barrera con una longitud aproximada de 610 m y una ampliación en el delta del arroyo con una superficie de 10.83 ha. El RMSE en el eje z entre los MDT generados y los datos del GPS diferencial, oscilaron entre 0.62 y 6.58 cm.

**Palabras clave**—fotogrametría; Modelo digital del terreno; morfología costera, playas.

## I. INTRODUCCIÓN

El estudio de la morfología costera requiere de mediciones topográficas continuas de las playas, para evaluar cambios en la distribución del volumen de sedimento. Este monitoreo de la morfodinámica costera es de suma importancia, especialmente en los últimos años, en virtud de que la tasa de erosión de las playas se ha acelerado a nivel mundial a causa del impacto antrópico directo, así como por el aumento en el nivel del mar relacionado con el cambio climático [1-6].

La toma de decisiones para el manejo de costas por lo general se basa en estudios que investigan la dinámica costera con diversas técnicas de monitoreo de playas, ya sea de manera directa (mediciones con GPS o perfiles topográficos) o indirecta (LiDAR o fotogrametría aérea). Por lo tanto, las técnicas para el monitoreo en la playa se basan en observaciones *in situ* o en datos obtenidos mediante percepción remota [7, 8].

Los productos generados por levantamientos topográficos precisos y de alta resolución espacial son cada vez más convenientes para comprender los cambios y procesos morfológicos de las playas en escalas de días, meses o años, así como en el monitoreo de eventos episódicos como las tormentas.

Existen varias técnicas para monitorear la morfología y evolución de playas a partir de la generación de Modelos Digitales del Terreno (MDT). Estos incluyen sistemas ARGUS de videograbación, sensores LiDAR aerotransportados, exploración láser terrestre (TLS), monitoreo topográfico tradicional usando un GPS diferencial en Tiempo Real (RTK-DGPS), estación total, y técnicas fotogramétricas. Cada una de estas técnicas tiene sus ventajas y limitaciones en términos de cobertura espacial y temporal, precisión, experiencia operativa, necesidades y costos de software [9, 10]. Entre las más populares destacan las mediciones con GPS topográficos de alta precisión [7, 8], las estaciones totales o los escáneres láser terrestres [8, 11, 12]. Estas técnicas son relativamente económicas y tienen una precisión de 2 a 6 cm en la vertical [8, 13]. Sin embargo, requieren grandes esfuerzos para la cartografía de grandes áreas, y no proporcionan ortofotografías útiles para detectar elementos específicos del terreno [8].

Recientemente se han empezado a utilizar nuevas técnicas para la generación de MDT con VANT (vehículos aéreos no tripulados, también conocidos como drones) en geomorfología y estudios ecológicos [8, 14-17]. Estos se establecen como una alternativa a la topografía tradicional, especialmente cuando se trata de cartografiar con alta resolución espacial y temporal, además de ser una alternativa de bajo costo a la fotogrametría aérea clásica [18- 20]. La captura de datos solo requiere un dron con un sistema de control y cámara digital estándar comercial [21].

## II. DATOS Y MÉTODOS

Se realizaron cinco levantamientos topográficos a lo largo de un año, iniciando el 24 de febrero de 2017, utilizando un VANT DJI Phantom 4 Pro y un equipo GNSS diferencial Leica GS14. Las imágenes digitales obtenidas fueron procesadas con la técnica de fotogrametría digital (Fig. 1) con el software Agisoft Photoscan. Los procedimientos se desarrollaron en tres etapas: levantamiento topográfico, procesamiento de fotografías aéreas y cálculo de volúmenes. A continuación, se describen estas etapas.

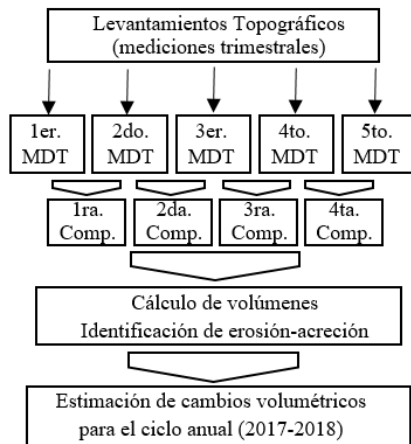


Fig. 1. Esquema conceptual. Levantamientos topográficos, comparación entre MDT (diferencia), cálculo de volúmenes y estimación de cambios.

### A. Levantamiento topográfico

a) Cobertura del levantamiento topográfico. Se estableció la cobertura del polígono de la región de interés (barrera arenosa del Estero San José), que comprende aproximadamente 1,800 m de playa entre el Hotel Holiday Inn y la Marina Puerto Los Cabos, con una superficie de 20 ha.

b) Bancos de nivel. Antes de realizar los vuelos se establecieron cuatro mojeneras (monumentos de concreto de 1.20 m de largo y 7.5 cm de diámetro), que fueron utilizados como bancos de nivel distribuidos cada 400 m a lo largo del área de estudio. Los bancos de nivel fueron georreferenciados con un equipo GPS diferencial Leica Viva GS14, asignándoles coordenadas UTM ligadas a la red geodésica del INEGI. Estos bancos de nivel sirvieron como referencia y enlace entre los levantamientos topográficos realizados.

c) Altura del vuelo. La altura del vuelo está en función de la precisión deseada para el levantamiento topográfico y a su vez, la precisión del levantamiento depende de la resolución de la imagen. Para definir la altura del vuelo según la resolución deseada o Distancia de Muestreo sobre el Terreno (GSD, por sus siglas en inglés), que en este caso es de 2 cm, se utilizó la ecuación (1):

$$GSD = \frac{H}{f} * \frac{d}{p} \quad (1)$$

Donde  $H$  es la altura del vuelo (m);  $f$  es la longitud focal de la cámara (8.8 mm);  $d$ , la dimensión del sensor (CMOS) de la cámara (13.20 × 8.80 mm);  $p$ , la resolución de la cámara (5472 × 3648 pixeles).

Los parámetros  $f$ ,  $d$  y  $p$  son valores específicos para la cámara del VANT DJI Phantom 4 Pro. Puesto que se desea una resolución espacial de 2 cm aproximadamente, se despeja la variable altura ( $H$ ) en la ecuación (1):

$$H = \frac{GSD * f * p}{d} = \frac{0.02 * 8.80 * 5472}{13.20} = 72.96 \quad (2)$$

En virtud de que la aplicación utilizada para volar el VANT (Pix4Dcapture), no tiene la opción de volar de manera autónoma a la altura calculada, se estableció una altura de 80 m, lo que resulta en una resolución espacial (GSD) de 2.19 cm.

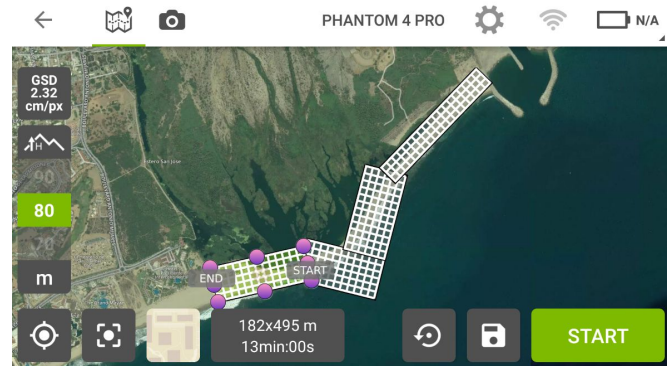


Fig. 2. Cobertura de los planes de vuelo generados en la aplicación móvil Pix4Dcapture.

d) Plan de vuelo. Una vez definido el polígono a medir y la altura del vuelo, se estableció el plan de vuelo. Este se generó con la aplicación para dispositivo móvil Pix4Dcapture (Fig. 2), en la cual se especifica el tipo de trayectoria y porcentaje de traslape entre fotografías, con lo cual la aplicación estima el tiempo de vuelo. El traslape de la fotografía es el empalme que presentan dos fotografías consecutivas. En estas zonas de traslape es en donde el software calcula las elevaciones del terreno. Se definieron 4 planes de vuelo con una cobertura de 6 ha y una duración de 18 min para cada levantamiento topográfico. El tipo de trayectoria elegido fue de doble malla (recorrido cuadrulado con dos trayectorias, una en sentido longitudinal y otra en sentido transversal a la línea de costa), especificando un traslape del 80% en las fotografías en ambas trayectorias, según las especificaciones recomendadas por [22].

e) Puntos de control terrestre. Los puntos de control terrestre (PCT) tienen la función de ubicar, orientar y escalar los productos generados como resultado del procesamiento de las imágenes (nube de puntos y MDT). Estos puntos se ubicaron con marcas de triplay de madera de 6 mm de espesor y de 50 × 60 cm por lado. En estas marcas se dibujaron cuatro cuadrantes y se pintaron de negro los cuadrantes opuestos con la finalidad de identificarlos fácilmente en las fotografías aéreas. Se distribuyeron las marcas sobre la barrera y posteriormente fueron georreferenciadas con un equipo GPS diferencial Leica Viva GS14 obteniendo su ubicación en coordenadas UTM ( $x$ ,  $y$ ,  $z$ ). Para el primer levantamiento se colocaron 81 PCT, para el segundo 55, en el tercero 64, para el cuarto y quinto levantamiento se colocaron 15. Los puntos de control obtenidos con este procedimiento fueron utilizados en la etapa de gabinete para la georreferenciación y procesamiento de las fotografías aéreas.

f) Vuelo y captura de fotografías. Una vez georreferenciados los puntos de control se procedió a realizar el vuelo vinculando la aplicación Pix4Dcapture con el VANT para seguir la trayectoria previamente establecida en el plan de vuelo. Para realizar los vuelos se consideró que la velocidad del viento fuera menor a 20 km/h según lo recomendado por [9], así como una marea baja para obtener fotografías con una mayor superficie de la playa.

### B. Procesamiento de imágenes

Esta etapa consistió en el procesamiento de las imágenes captadas con el VANT y los puntos de control obtenidos en campo utilizando el software Agisoft Photoscan, para generar los MDT. El flujo de trabajo fotogramétrico implementado en este software está basado en el protocolo de la estructura del movimiento (SfM por sus siglas en inglés), el cual ha sido descrito por [10, 23, 24]. El flujo de trabajo y la reconstrucción de MDT se describen brevemente a continuación.

a) Alineación de las imágenes. En esta etapa se identificaron los puntos de control en las fotografías, asignándoles las coordenadas correspondientes (las obtenidas en campo con el GPS diferencial). Agisoft Photoscan utiliza un algoritmo de seguimiento de rasgos (*feature tracking*) para identificar, igualar y supervisar la ubicación de los rasgos únicos [9, 25].

b) Generación de nubes de puntos. Con base en la posición estimada de cada imagen, el software Agisoft Photoscan calcula la profundidad por correlación entre los traslapes de cada par de imágenes y luego combina todas estas en una sola nube de puntos [25]. En esta etapa también se llevó a cabo el proceso de enmascarado de fotografías, que consiste en editar manualmente cada fotografía para eliminar las regiones que contienen superficie del agua (de mar y del estero), así como de las personas que transitaban por la playa durante el levantamiento. La nube de puntos resultante se generó con una densidad promedio (resolución del levantamiento topográfico) de 150 puntos/m<sup>2</sup>.

La nube de puntos generada fue clasificada por rangos de alturas y por la pendiente del terreno, eliminando puntos correspondientes a vegetación, así como los encontrados en los límites del levantamiento y las elevaciones erróneas provocadas por el efecto del oleaje.

c) Creación de malla y textura. A partir de la triangulación de la nube de puntos, el software Agisoft Photoscan reconstruye una malla poligonal 3D que representa la superficie de terreno. Esta malla se editó manualmente eliminando puntos aislados que generan distorsión en el MDT (principalmente en la cara de la playa por el efecto del oleaje).

Una vez que la malla fue editada, se procedió a generar la textura del MDT (Fig. 3). En esta etapa el software Agisoft Photoscan asigna la textura y colores de las fotografías aéreas al MDT.

Finalmente, los MDT generados fueron sometidos a un análisis estadístico para evaluar la precisión de los levantamientos topográficos.

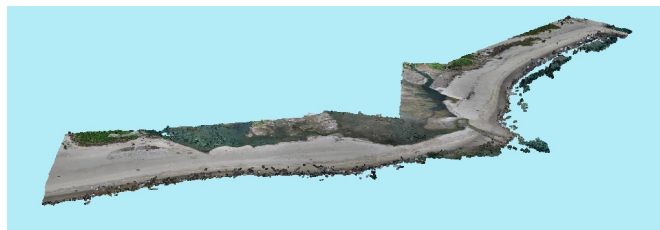


Fig. 3. Nube de puntos generada con el software Agisoft Photoscan (levantamiento del 15/feb/2018).

El análisis estadístico consistió en comparar las coordenadas generadas con el MDT en los puntos de control y en puntos arbitrarios de perfiles aleatorios georreferenciados con el equipo GNSS diferencial. El análisis estadístico (basado en [26]), emplea como estimador la raíz del error cuadrático medio (RMSE).

### C. Cálculo de volúmenes y zonas de erosión-acreción

Los MDT generados con el proceso de fotogrametría aérea para cada levantamiento topográfico se compararon entre sí, con la finalidad de evaluar los cambios morfológicos que presentó la barrera en los periodos de medición. Para la interpretación de estas variaciones se identificaron las zonas de cambio (erosión/acreción) calculando el volumen del sedimento entre levantamientos topográficos consecutivos (diferencia entre dos MDT).

Por ejemplo, para estimar los cambios entre los levantamientos de febrero y mayo, se restó el MDT de febrero al de mayo. La resta aritmética de elevaciones pixel a pixel se realizó con ArcGIS. Los valores positivos se interpretan como erosión, si es negativo como acreción y si es cero como estable. Posteriormente, se generaron una serie de perfiles en los que se representan con mayor detalle los cambios en la morfología de la playa (avance o retroceso de la línea de costa en el tiempo transcurrido entre los levantamientos).

## III. RESULTADOS Y DISCUSIÓN

La barrera arenosa presentó modificaciones considerables en su morfología por la incidencia de la tormenta tropical Lidia. La descarga del arroyo San José aportó al litoral un volumen de sedimento mayor a 200,000 m<sup>3</sup>. Los cambios más notables en la morfología de la barrera ante este tipo de eventos meteorológicos con lluvias extraordinarias se dan por el aporte de sedimento del arroyo cuando hay ruptura en la barrera. Esto genera una acreción súbita de sedimentos, que excede la capacidad del oleaje de tormenta para redistribuirlo en el corto periodo (meses). El sedimento disponible por la descarga se distribuye a lo largo de la barrera. Con las mediciones realizadas se puede deducir que la dirección del transporte es al oeste, donde se encuentran las mayores variaciones volumétricas, mientras que al este, las variaciones son menores.

El MDT del 24/feb/2017 (Fig. 4) representa la topografía inicial de la barrera arenosa, se levantaron 1800 m de longitud de playa en los cuales se observa que la playa tiene un ancho de entre 100 a 130 m en las partes más amplias y de entre 30 a 50 m en la parte estrecha, presentando una pendiente en la cara de la playa de entre 11 y 14°. Para El MDT del segundo levantamiento (26/may/17) la playa conserva los rasgos descritos del primer levantamiento, presentando estabilidad en la trasplaya y un avance de hasta 50 cm de línea de costa al centro y sureste de la barrera, así como un retroceso de hasta 1.5 m en los 400 m adyacentes a la escollera de la marina. En el MDT del levantamiento del 06/sep/2017, los cambios en la morfología de la barrera arenosa son más importantes (ruptura de 610 m a lo largo de la barrera, formación de un delta con una superficie de 10.83 ha con avance de ~250 m hacia el mar) que fueron generados por la descarga del arroyo San José, inducida por la tormenta tropical Lidia el 01/sep/2017.

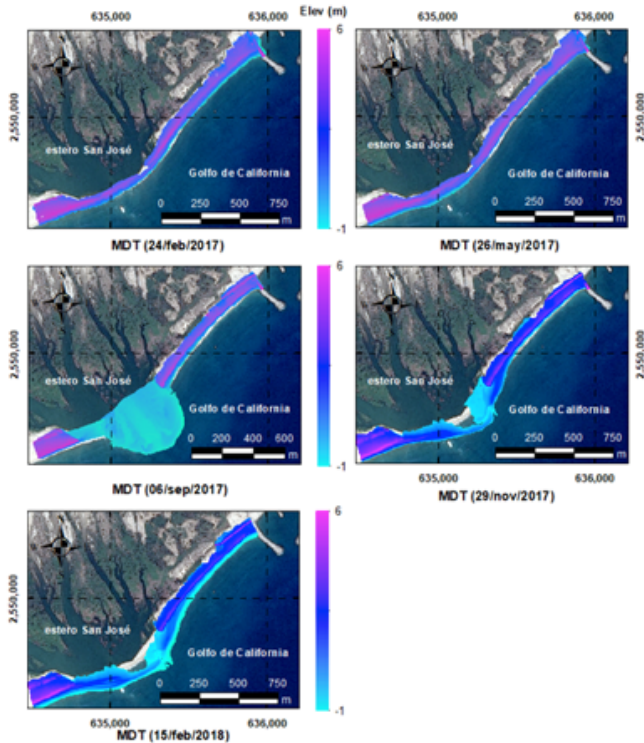


Fig. 4. Modelos digitales de elevación (MDT) resultado de la restitución fotogramétrica para cada uno de los levantamientos topográficos. Las elevaciones (en metros) están referidas al nivel medio del mar.

Para el levantamiento del 29/nov/2017 (Fig. 4), se observa que el delta disminuye 5 ha aproximadamente, incrementando su elevación de 50 cm a un rango de 1.5 a 3 msnm con respecto a la medición anterior, observándose que el sedimento se distribuye hacia el noreste de la barrera. En la última medición (15/feb/2018), se observa que en la zona suroeste al delta del arroyo se presentaron cambios más evidentes en la morfología de la barrera, con un avance de hasta 25 m y un crecimiento de una hectárea aproximadamente, con respecto a la medición del 29/nov/2017.

En la Figura 5 se representan los cambios morfológicos entre MDT consecutivos, siendo el más importante el que se dio entre mayo y septiembre de 2017. El oleaje de tormenta provocó un retroceso a lo largo de la línea de costa del área de estudio, que alcanzó aproximadamente 20 m. Para este periodo, los volúmenes de acreción y erosión estimados fueron de 155,993 m<sup>3</sup> y 190,998 m<sup>3</sup> respectivamente, con una acreción neta de 35,005 m<sup>3</sup>. Los mayores volúmenes se localizan en la zona deltaica de la barrera, donde la tasa de acreción media fue de 128 a 491 m<sup>3</sup>/m. Esta acreción se presentó en la zona de rompiente, frente a la ruptura de la barrera, por efecto del sedimento descargado por el arroyo, así como por el removido de la barrera por su ruptura, formando el delta antes mencionado.

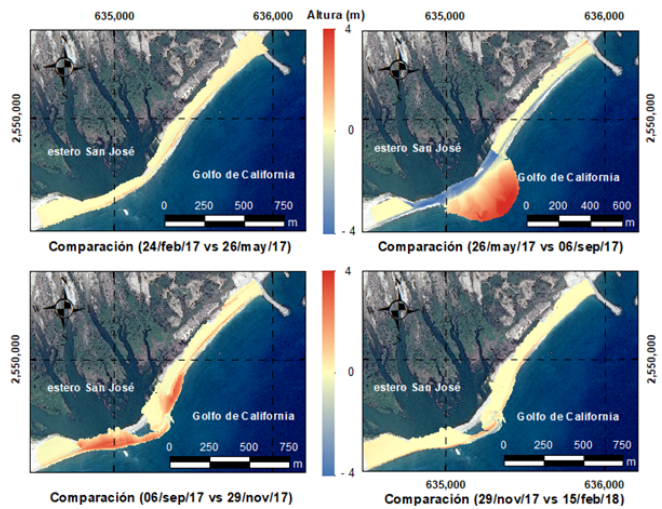


Fig. 5. Cambios morfológicos en la barrera arenosa del Estero San José en las fechas indicadas. La escala de colores indica pérdida o acumulación del sedimento en metros por pixel.

En la tabla 1 se muestran los errores obtenidos durante el procesamiento fotogramétrico para cada levantamiento topográfico.

Tabla 1. Densidad de puntos en los MDT, así como errores entre los MDT generados y las mediciones con el GPS diferencial (RTK).

Fecha	24-feb-17	26-may-17	06-sep-17	29-nov-17	15-feb-18
# fotografías	2201	1608	1950	1069	893
Densidad	156 pts/m <sup>2</sup>	168 pts/m <sup>2</sup>	163 pts/m <sup>2</sup>	122 pts/m <sup>2</sup>	127 pts/m <sup>2</sup>
<b>(a) Error (PC)</b>					
# (PC)	81	55	64	15	15
error en X (cm)	3.6	3.06	1.92	0.47	0.65
error en Y (cm)	4.4	2.15	1.38	0.23	0.43
error en Z (cm)	3.32	0.91	0.78	0.33	0.93
RMSE (cm)	<b>6.58</b>	<b>3.85</b>	<b>2.49</b>	<b>0.62</b>	<b>0.9</b>
<b>(b) Error (PA)</b>					
# Puntos aleatorios	60	131	-	-	-
RMSE (cm)	<b>13.26</b>	<b>5.1</b>	-	-	-

La Figura 6 muestra las zonas y magnitudes de erosión y acreción a lo largo de la barrera en volumen por metro lineal de playa (m<sup>3</sup>/m) entre los levantamientos realizados.

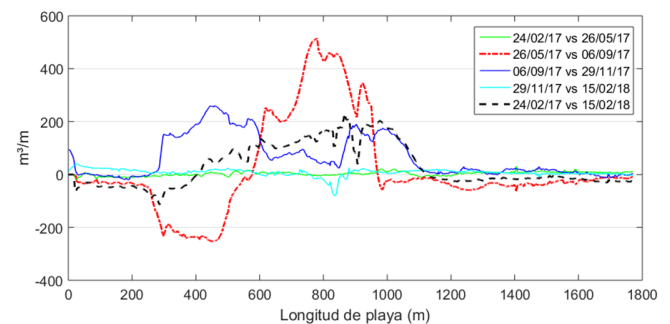


Fig. 6. Curvas de la variación de las tasas de erosión y acreción trimestral y anual (m<sup>3</sup>/m) a lo largo de la barrera arenosa.



IV. CONCLUSIONES

La metodología seguida para realizar los levantamientos topográficos con VANT y fotogrametría digital es eficiente. Conforme se desarrolló la investigación se realizaron diversas pruebas y se concluyó que aumentar la cantidad de puntos de control no mejora la precisión del levantamiento topográfico. Se obtuvo una mejor precisión cuando se disminuyó la cantidad de puntos de control, logrando reducir la cantidad de 81 a 15 puntos para una superficie de 30 ha, facilitando el trabajo de campo y de gabinete. En el procedimiento se observó que las fronteras de los MDT generados con el proceso fotogramétrico presentan distorsión, por lo que se recomienda hacer una cobertura de mayor tamaño que envuelva a la superficie de interés para que estas distorsiones queden fuera de la región de interés.

AGRADECIMIENTOS

Los autores agradecen al Centro Interdisciplinario de Ciencias Marinas del Instituto Politécnico Nacional por el apoyo para la realización de este trabajo (a través del Proyecto SIP-20172040 “Erosión y recuperación de playas impactadas por los huracanes Juliette y Odile en la región de Los Cabos, Baja California Sur”), así como a los profesores Dr. Marco Antonio Castro Liera y MC. Liliana Castro Liera del Instituto Tecnológico de La Paz, por las facilidades otorgadas para utilizar el equipo de cómputo del Departamento de Sistemas y Computación.

REFERENCIAS

[1] Fox, W.T., & Davis, R.A., 1978. Seasonal variation in beach erosion and sedimentation on the Oregon coast. Geological Society of America Bulletin, 89, 1541–1549.

[2] Birkemeier, W.A., Nicholls, R.J., Lee, G., 1999. Storms, storm groups and nearshore morphologic change. In: Kraus, N.C., McDougal, W. (Eds.), Coastal Sediments '99, vol. 2. ASCE, Long Island, U.S.A.;1109–1122.

[3] Anfuso, G., Benavente, J. and Gracia, J., 2001. Morphodynamic responses of nourished beaches in SW Spain, Journal of Coastal Conservation, 7, 71-80.

[4] Masselink, G., Pattiaratchi, C.B., 2001. Seasonal changes in beach morphology along the sheltered coastline of Perth, Western Australia. Marine Geology, 172, 243–263.

[5] Swales, A., 2002. Geostatistical estimation of short-term changes in beach morphology and sand budget. Journal of Coastal Research, 18, 338–351.

[6] Jeong-Min Lee, Jun-Yong Park, Jin-Yong Choi, 2013. Evaluation of Sub-aerial Topographic Surveying Techniques Using Total Station and RTK-GPS for applications in Macro-tidal Sand Beach Environment., Proceedings 12th International Coastal Symposium (Plymouth, England), Journal of Coastal Research, Special Issue No. 65; 535- 540.

[7] Morton, R.A., Leach, M.P.; Paine, J.G., and Cardoza, M.A., 1993. Monitoring beach changes using GPS surveying techniques. Journal of Coastal Research, 9, 702–720.

[8] Casella, E., Rovere, A., Pedroncini, A. et al., 2016. Drones as tools for monitoring beach topography changes in the Ligurian Sea (NW Mediterranean). Geo-Marine Letters, 36, (2).

[9] Brunier G., Fleury J., J. Anthony, Gardel A. Dussouillez P., 2016. Close-range airborne Structure-from-Motion Photogrammetry fo high-resolution beach morphometric surveys: Examples from an embayed rotating beach. Journal of Geomorphology 261; 76–88.

[10] James, M.R., Ilic, S., Ruzic, I., 2013. Measuring 3D coastal change with a digital camera. Coastal Dynamics. 893–904.

[11] Saye SE, Van der Wal D, Pye K, Blott SJ., 2005. Beach–dune morphological relationships and erosion/accretion: an investigation at five sites in England and Wales using LIDAR data. Geomorphology 72(1):128–155.

[12] Theuerkauf E.J., Rodriguez A.B., 2012. Impacts of transect location and variations in along-beach morphology on measuring volume change. J Coast Res 28(3):707–718.

[13] Saye SE, Van der Wal D, Pye K, Blott SJ., 2005. Beach–dune morphological relationships and erosion/accretion: an investigation at five sites in England and Wales using LIDAR data. Geomorphology 72(1):128–155.

[14] Floreano D., Wood R.J., 2015. Science, technology and the future of small autonomous drones. Nature 521(7553): 460–466.

[15] Colomina, I., Molina P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: a review. ISPRS J Photogramm Remote Sens 92: 79–97.

[16] Anderson K, Gaston K.J., 2013. Lightweight unmanned aerial vehicles will revolutionize spatial ecology. Front. Ecol. Environ. 11; 138–146. doi:10.1890/120150.

[17] Everaerts J., 2008. The use of unmanned aerial vehicles (UAVs) for remote sensing and mapping. Int Arch Photogramm Remote Sens Spat Inf Sci 37:1187–1192.

[18] Uysal M., Toprak A.S., Polat N., 2015. DEM generation with UAV Photogrammetry and accuracy analysis in Sahitler hill. Measurement 73 539–543.

[19] Eisenbeiss H., 2009. UAV Photogrammetry, Diss. ETH No. 18515, Institute of Geodesy and Photogrammetry, ETH Zurich, 2009, pp. 235.

[20] Colomina, I., M. Blázquez, P. Molina, M.E. Parés, M. Wis., 2008. Towards a new paradigm for high-resolution low-cost photogrammetry and remote sensing, IAPRS&SIS; 1201–1206.

[21] Haala, N., Cramer, M., Weimer, F., Trittler, M., 2011. Performance test on UAV-based photogrammetric data collection, ISPRS ICWG I/V UAV-g Conference, Zurich, Switzerland.

[22] Turner I.L., Harley M.D., Drummond C.D., 2016. UAVs for coastal surveying. Coastal Engineering, 114,19-24.

[23] Javernick, L., Brasington, J., Caruso, B., 2014. Modelling the topography of shallow braided rivers using structure-from-motion photogrammetry. Geomorphology 213, 166–182.

[24] Gonçalves, J.A., Henriques, R., 2015. UAV photogrammetry for topographic monitoring of coastal areas. ISPRS J. Photogramm. Remote Sens. 104, 101–111.

[25] Agisoft, 2015. Agisoft Photoscan User Manual: Professional Edition. Version 1.1. ([http://www.agisoft.com/pdf/photoscan-pro\\_1\\_1\\_en.pdf](http://www.agisoft.com/pdf/photoscan-pro_1_1_en.pdf)).

[26] Höhle, J., Höhle, M., 2009. Accuracy assessment of digital elevation models by means of robust statistical methods. ISPRS Journal of Photogrammetry and Remote Sensing 64 (5), 398–406.

# Análisis comparativo de filtros para mediciones de sensores en aplicaciones de seguimiento de personas en interiores

Héctor E. Villarreal Rodríguez, Arnoldo Díaz-Ramírez, Verónica Quintero Rosas, Mario A. Camarillo Ramos, Mayra L. Lizárraga Camberos

Departamento de Sistemas y Computación  
Tecnológico Nacional de México/IT de Mexicali

hevr93@gmail.com, {adiaz,veronicaquintero, mario.camarillo}@itmexicali.edu.mx, mayral.lizarraga@gmail.com

**Resumen**—El paradigma de Internet de las Cosas ha propiciado el uso de sensores diversos para monitorizar entornos en muchas áreas, como el hogar, vehículos, aviones, fábricas, por mencionar algunos. Las mediciones recabadas por los sensores están sujetas a errores por diferentes causas, principalmente por fuentes externas. Se le denomina *ruido* a esta perturbación en la medición. Para reducir o eliminar el efecto del ruido en la información recolectada es común el uso de filtros. Sin embargo, debido a que existe una enorme cantidad de filtros disponibles, es importante elegir el más adecuado dependiendo del contexto de la aplicación y la naturaleza de los datos.

Una de las aplicaciones más interesantes del Internet de las Cosas es la detección y registro de la ruta que sigue una persona en el interior de un edificio. Las lámparas o los artículos electrodomésticos, entre otros, pueden generar ruido que altere las mediciones de los sensores. En este artículo se presentan los resultados de un análisis comparativo de filtros para mediciones de sensores en aplicaciones de seguimiento de personas en interiores. La motivación de este trabajo es su aplicación en un componente que detecte los patrones de deambulación de personas afectadas por demencia, con la finalidad de tener un diagnóstico más certero del avance de su enfermedad. Los resultados obtenidos permitieron obtener datos de los sensores con una muy buena tasa de precisión.

## I. INTRODUCCIÓN

En años recientes se ha generado un gran interés por el paradigma del Internet (IoT) de las cosas [1]. En las aplicaciones IoT es común el uso de sistemas embebidos, equipados con sensores, que son colocados en cosas o personas para que de manera ubicua monitoricen y detecten eventos de interés. Además, utilizando tecnologías de comunicación inalámbrica, es posible que usuarios o sistemas informáticos reciban la información recolectada y puedan actuar en consecuencia. La irrupción del IoT ha motivado el uso de una enorme cantidad de sensores que recaban información de manera continua y que es enviada de manera periódica a una estación base, en donde es almacenada para que los usuarios u otros sistemas puedan consultarla. Las áreas de aplicación del IoT son muy amplias, incluyendo la salud, la agroindustria, ciudades inteligentes, ecología y manufactura, por incluir algunos. De hecho, el IoT es una de las tecnologías habilitadores de la Industria 4.0.

La fuente primaria de los datos que recolectan los sensores son mediciones de parámetros de fenómenos o sistemas físicos. Por ejemplo, la temperatura, la humedad o la presión

barométrica. Desafortunadamente, los datos recabados por los sensores están sujetos a diferentes fuentes de error [2]. Es común clasificar a las fuentes de error de las mediciones de los sensores en dos tipos: errores sistemáticos y errores aleatorios. Los primeros, conocidos también como de sesgo (*bias*), son causados por cambios en condiciones externas, como la humedad o temperatura, o por defectos en los mismos sensores (fallos, falta de calibración, etc). Por otra parte, los errores aleatorios, también conocidos como ruido (*noise*), pueden ser causados por interferencia de elementos externos, como la luz emitida por lámparas fluorescentes o campos electromagnéticos. Los primeros pueden ser corregidos por calibración o sustitución de los sensores que fallan. Sin embargo, las causas de los errores aleatorios difícilmente pueden ser controladas o evitadas, ya que en muchas ocasiones las fuentes de los errores son externas.

Los sensores muestran variaciones significativas en la precisión y exactitud de sus mediciones, así como tolerancia a ruido externo. Estas variaciones dependen en gran medida al tipo, costo y aplicación en la que se utiliza el sensor. Por ejemplo, se ha descubierto que las mediciones de los sensores de temperatura son seriamente afectadas por la radiación solar [3], por lo que es necesario aplicar técnicas de filtrado para reducir el error [4]. Otro ejemplo es el del sistema de satélites de posicionamiento global o GPS, cuyas mediciones pueden tener errores de varios metros [5].

Uno de los temas de mayor interés de aplicaciones IoT es el de la detección y seguimiento de personas en interiores. Se han detectado diversos problemas que pueden abordarse con técnicas de detección y seguimiento de personas en interiores, tales como seguridad [6], rescate de personas [7], salud [8], entre otras [9]. En [10] se clasifican las técnicas de localización de personas en interiores en *activos* y *pasivos*. En las técnicas activas, las personas llevan algún dispositivo adherido en su cuerpo, como pueden ser etiquetas RFID o sensores. En contraste, las técnicas pasivas no requieren que las personas lleven algún dispositivo, por lo que la localización se lleva a cabo a través del uso de técnicas tales como mediciones de la fuerza de señales de WiFi (*Received Signal Strength Indicator* o *RSSI*) o de procesamiento de video. Otra manera de clasificar a las técnicas de localización y seguimiento de personas es en: 1) técnicas que requieren infraestructura y



2) técnicas que no requieren infraestructura. En las primeras debe instalarse algún tipo de infraestructura previa en el sitio en el que se aplicará la técnica de localización. Por ejemplo, pueden utilizarse transmisores inalámbricos y medir la fuerza de la señal recibida para estimar la ubicación del sujeto [11]. Por otra parte, las técnicas sin infraestructura no requieren conocimiento previo del lugar ni equipo pre-instalado [12]. La técnica elegida depende en gran medida del contexto de la aplicación. Por ejemplo, en el caso de un sistema de seguridad aplicado a un edificio, es posible instalar infraestructura previa de tal manera que se utilice alguna técnica pasiva, que no sea del conocimiento del sujeto. Un ejemplo es la instalación de video cámaras y el uso de procesamiento de imágenes. Pero existen aplicaciones en las que no es posible conocer o tener acceso previo al sitio en el que se utilizará el sistema de localización y seguimiento. Por ejemplo, el rescate de personas en edificios siniestrados.

Recientemente se desarrolló un método de tipo activo para localización de personas en interiores, sin necesidad de infraestructura previa [13]. Este método es capaz de detectar y registrar, con una gran precisión, la trayectoria que sigue una persona mientras se desplaza en el interior de un edificio. El dispositivo, adherido al sujeto, está equipado con sensores que miden la aceleración y la dirección de su movimiento al desplazarse, además de la intensidad campo magnético terrestre. Para este fin se utilizan el magnetómetro, el giroscopio y el acelerómetro. Los sensores, a pesar de contar con una gran precisión en la actualidad, en ocasiones registran mediciones erróneas debido al ruido externo. Para tratar dichos errores se utiliza un filtro pasa baja [14] que reduce significativamente el efecto del ruido externo en las mediciones de los sensores que generan información analógica, como en el caso de los sensores utilizados en [13].

Con base en el trabajo de *Lizárraga et al.* [13], se determinó desarrollar un componente IoT para detectar los patrones de deambulación de personas afectadas con demencia, ya que se ha descubierto que el patrón de deambulación es un síntoma del grado de avance de la enfermedad [15]. Sin embargo, se descubrió que el filtro pasa baja era capaz de eliminar el ruido de los sensores si la distancia recorrida por la persona es mayor a veinte metros. Pero una persona afectada con demencia usualmente recorre una distancia menor en cada episodio de deambulación, por lo que fue necesario aplicar un filtro adicional. En este artículo se presenta un estudio comparativo de diversos filtros de señales digitales, utilizados en mediciones recabadas por un magnetómetro, un giroscopio y un acelerómetro, en aplicaciones para detectar las rutas seguidas por personas al desplazarse de un lugar a otro. Con los resultados del estudio comparativo fue posible diseñar un método con una baja tasa de error.

El artículo está organizado de la siguiente manera. La Sección 2 presenta brevemente los filtros digitales utilizados en el estudio comparativo. En la Sección 3 se muestran los resultados del estudio. Finalmente, se presentan las conclusiones y el trabajo futuro.

## II. FILTROS DE SEÑALES DIGITALES

Un filtro es un dispositivo o proceso que remueve componentes o características no deseada en una señal [16]. Originalmente los filtros eran diseñados utilizando inductores, capacitores y transformadores, y terminados por resistores representando la carga y resistencia interna de la fuente. Con la aparición de las computadoras, los filtros para señales digitales se comenzaron a implementar como procesos utilizando técnicas de integración e interpolación, y resolviendo ecuaciones diferenciales. Los filtros digitales tienen la ventaja de contar con una alta precisión y confiabilidad, además de ser más económicos.

Para eliminar características no deseadas de las mediciones de los sensores magnetómetro, giroscopio y acelerómetro, utilizados para registrar la ruta de una persona al caminar trayectos menores de 20 metros, se evaluaron los filtros que se describen a continuación.

### II-A. Filtro Elíptico

El filtro elíptico también se conoce como filtro Cauer y, a veces, incluso como filtro Zolotarev. El filtro se utiliza en muchas aplicaciones de RF (radio frecuencia) donde se requiere una transición muy rápida entre las frecuencias de banda de paso y de banda de parada. El filtro elíptico produce la transición más rápida que cualquier tipo de filtro, pero también muestra una ondulación de ganancia en las denominadas pasa banda y rechaza banda.

Los niveles de ondulación en la pasa banda y en la rechaza banda, se pueden ajustar de forma independiente durante el diseño. A medida que la onda en la pasa banda se acerca a cero, el filtro se convierte en un filtro de tipo I de Chebyshev, y cuando la onda en la rechaza banda se aproxima a cero, se convierte en un filtro de tipo II de Chebyshev. Si la ondulación en la en la pasa banda y en la rechaza banda se convierte en cero, entonces el filtro se transforma en un filtro Butterworth [17].

### II-B. Filtro Butterworth

El filtro Butterworth es un tipo de filtro de procesamiento de señal diseñado para tener una respuesta de frecuencia lo más plana posible en la pasa banda. También se conoce como un filtro de magnitud máxima plana. Fue descrito por primera vez en 1930 por el ingeniero y físico británico Stephen Butterworth en su artículo titulado "*On the Theory of Filter Amplifiers*".

La topología de filtro Butterworth se usa ampliamente en muchas aplicaciones de RF y de filtro general. Una de las características clave del filtro Butterworth es que tiene lo que se denomina una respuesta máxima plana dentro de su pasa banda. De hecho, el filtro Butterworth a menudo se considera como una buena forma de filtro que es adecuado para muchas aplicaciones, aunque no proporciona el corte más preciso [18].

### II-C. Filtro Chebyshev

Los filtros Chebyshev son filtros analógicos o digitales que tienen una caída más pronunciada y más ondulación de la pasa banda (tipo I) u ondulación de la rechaza banda (tipo

II) que los filtros Butterworth. Los filtros Chebyshev tienen la propiedad de que minimizan el error entre las características del filtro, pero con ondulaciones en la pasa banda. Este tipo de filtro lleva el nombre de Pafnuty Chebyshev porque sus características matemáticas se derivan de los polinomios de Chebyshev. Los filtros de tipo I de Chebyshev se denominan normalmente como "*filtros de Chebyshev*", los de tipo II se denominan generalmente como "*filtros de Chebyshev inversos*" [19], [20].

#### II-D. Filtro Savitzky-Golay

Los filtros de suavizado Savitzky-Golay se utilizan normalmente para *suavizar* una señal ruidosa cuyo intervalo de frecuencia (sin ruido) es grande. También se denominan filtros polinomiales de suavizado digital o filtros de suavizado de mínimos cuadrados. Los filtros Savitzky-Golay funcionan mejor en algunas aplicaciones que los filtros FIR estándar de promedio, que tienden a filtrar el contenido de alta frecuencia junto con el ruido. Los filtros Savitzky-Golay son más efectivos para preservar los componentes de señal de alta frecuencia, pero tienen menos éxito para rechazar el ruido.

Los filtros Savitzky-Golay son óptimos en el sentido de que minimizan el error de mínimos cuadrados al ajustar un polinomio a cuadros de datos ruidosos [21].

#### II-E. Suavizado con modelo polinomial de primer y segundo grado (*lowess* y *loess*)

Los nombres "*lowess*" y "*loess*" se derivan del término "*diagrama de dispersión ponderado localmente suave*", ya que ambos métodos utilizan la regresión lineal ponderada localmente para suavizar los datos.

El proceso de suavizado se considera local porque, al igual que el método del promedio móvil, cada valor suavizado está determinado por los puntos de datos vecinos definidos dentro del intervalo. El proceso está ponderado porque se define una función de peso de regresión para los puntos de datos contenidos en el intervalo. Además de la función de peso de regresión, puede utilizar una función de peso robusto, que hace que el proceso sea resistente a los valores atípicos. Finalmente, los métodos se diferencian por el modelo utilizado en la regresión: *lowess* usa un polinomio lineal, mientras que *loess* usa un polinomio cuadrático [19].

#### II-F. Versión robusta de *lowess* y *loess* que asigna menor peso a los valores atípicos en la regresión (*rloess* y *rloess*)

Si los datos contienen valores atípicos, los valores suavizados pueden distorsionarse y no reflejar el comportamiento de la mayor parte de los puntos de datos vecinos. Para resolver este problema, se puede suavizar los datos utilizando un procedimiento robusto que no está influenciado por una pequeña fracción de valores atípicos. El método asigna un peso cero a los datos fuera de las medias de las desviaciones absolutas [22].

#### II-G. Filtro de movimiento promedio

El filtro de movimiento promedio es un simple filtro FIR de paso bajo (respuesta de impulso finito), comúnmente utilizado para suavizar una serie de datos obtenidos por algún medio (e.g., recolectados por un sensor). Recibe  $N$  muestras de entrada y obtiene el promedio de esas  $N$ -muestras, produciendo un solo valor de salida. Es una estructura muy simple de LPF (filtro de paso bajo), que es útil para los científicos e ingenieros para filtrar el ruido no deseado de los datos recabados.

A medida que aumenta la longitud del filtro (el parámetro  $N$ ), la *suavidad* de la salida aumenta, mientras que las transiciones bruscas en los datos se hacen cada vez más pronunciadas. Esto implica que este filtro tiene una excelente respuesta en el dominio del tiempo, pero una respuesta de frecuencia pobre [23], [24]. El filtro MA realiza tres funciones importantes:

1. Recibe  $N$  puntos de entrada, calcula el promedio de esos puntos  $N$  y produce un solo punto de salida
2. Debido al cómputo/cálculos involucrados, el filtro introduce un tiempo definido de retraso
3. El filtro actúa como un filtro de paso bajo, con una respuesta de dominio de frecuencia deficiente y una respuesta de dominio de tiempo buena.

### III. ANÁLISIS COMPARATIVO DE FILTROS DE SEÑALES DIGITALES

Para la evaluación del desempeño de los filtros descritos en la sección anterior, se decidió utilizar un dispositivo similar al utilizado durante las pruebas realizadas por [13]. Dicho dispositivo es un dispositivo móvil, modelo «Samsung Galaxy s7». Al utilizar un dispositivo idéntico en ambos casos se tuvo la ventaja de poder replicar los experimentos realizados en su trabajo previo, además de llevar a cabo pruebas en ambientes distintos.

Este dispositivo cuenta con varios sensores con los que se recolectaron los datos de las trayectorias de las personas. Por ejemplo, para poder recolectar información de la aceleración del sujeto se utilizó el sensor «K6DS3TR». Para obtener el ángulo de inclinación del sujeto con respecto al eje magnético terrestre, se utilizó el giroscopio «K6DS3TR». Para obtener la intensidad del campo magnético terrestre en la localización actual del dispositivo, se utilizó el sensor «YAS537».

Se llevaron a cabo 30 experimentos en los que una persona llevó a cabo diferentes desplazamientos. Diez de ellos fueron en línea recta y diez fueron cíclicos; esto es, la persona hizo desplazamientos semi-circulares, regresando a un lugar cercano al inicio. Seis experimentos fueron desplazamientos rítmicos, en los que la persona hizo desplazamiento repetitivos, de un punto a otro y de regreso, dentro de una área delimitada. Finalmente, cuatro de ellos fueron desplazamientos aleatorios, a diferentes puntos, sin una área delimitada ni repeticiones.

La información recabada con los sensores fue procesada utilizando las técnicas descritas en [25]. Con las mediciones realizadas por los sensores del dispositivo, se obtuvieron valores del ángulo y la distancia del movimiento de la persona. Utilizando estos datos como coordenadas polares y aplicando un par de funciones matemáticas, se obtienen coordenadas

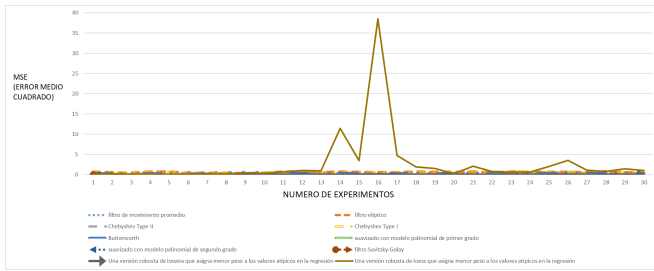


Figura 1. Comparativa entre los diferentes resultados de los suavizados/filtrados

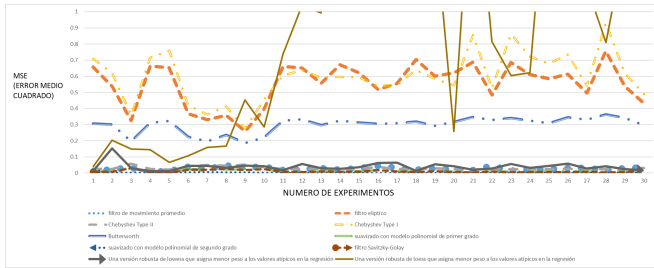


Figura 2. Filtros con mejor desempeño

rectangulares del sujeto, pudiendo identificar las posiciones del mismo mientras se desplaza.

Para comparar los valores generados por los filtros con respecto a los valores reales se utilizó el el error cuadrático promedio o MSE. El MSE es una medida global del error total entre dos vectores de datos. En ese caso, uno de los vectores contiene los valores reales del desplazamiento de la persona, mientras que el otro contiene los valores de las mediciones de los sensores después de haber sido procesado por el filtro. Entre menor sea el MSE, mas preciso es el filtro. De manera formal el MSE puede escribirse como

$$MSE = \frac{1}{N} \sum_{x=1}^N [Y_{clean}(x) - Y_{proc}(x)]^2$$

El Cuadro I muestra el promedio del MSE entre los diferentes tipos de filtros utilizados, aplicados datos obtenidos por las mediciones de los sensores.. Como se puede apreciar, el método de filtrado que mejor desempeño tuvo fue el de suavizado con modelo polinomial de primer grado o *lowess*, mientras que el menos preciso es la versión robusta del suavizado con un modelo polinomial de segundo grado o *rloess*.

La Fig. 1 muestra la comparativa por experimento, donde se puede apreciar la diferencia entre el desempeño de los filtros. La Fig. 2 muestra los mismos resultados pero permite apreciar el comportamiento de los filtros con mejor desempeño. Puede observarse que el que proporcionó mejores resultados es el de suavizado con modelo polinomial de primer grado o *lowess*. Cabe señalar que en el componente propuesto en [25] se utilizó el filtro de movimiento promedio, que a pesar de no ser el de mejor desempeño proporcionó muy buenos resultados. La razón de su elección fue el tener una menos complejidad de implementación con respecto a los mas eficientes, lo que lo

hace atractivo para aplicaciones que deben operar en tiempo real.

#### IV. CONCLUSIONES

La proliferación del uso de sensores propiciado por la irrupción del Internet de las Cosas ha motivado el uso de técnicas de filtrado para reducir o eliminar el ruido en los datos obtenidos por los sensores. El ruido es causado por elementos externos tales como lámparas, campos magnéticos u otros dispositivos. A la fecha se han propuesto una enorme cantidad de filtros, por lo que la elección del mas conveniente no es tarea trivial, y mas porque su elección está en gran medida determinada por el contexto de la aplicación.

Recientemente se propuso el desarrollo de un componente, con base en el paradigma IoT, para detectar los patrones de deambulaje en interiores de edificios, de personas afectadas por algún tipo de demencia, como la enfermedad del Alzheimer. Los desplazamientos de estas personas suelen ser de distancias pequeñas y continuos, por lo que fue necesario llevar a cabo un estudio comparativo para determinar el mejor filtro para mediciones obtenidas por un magnetómetro, un acelerómetro y un giroscopio, sensores utilizados en el desarrollo del componente. En este artículo se presentaron los resultados de dicho estudio comparativo, lo que permitió determinar cuáles fueron los filtros con mejor desempeño para el componente IoT desarrollado.

#### REFERENCIAS

- [1] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787 – 2805, 2010.
- [2] E. Elnahrawy and B. Nath, "Cleaning and querying noisy sensors," in *Proceedings of the 2nd ACM International Conference on Wireless Sensor Networks and Applications*, pp. 78–87, 2003.
- [3] X. Sun, S. Yan, B. Wang, L. Xia, Q. Liu, and H. Zhang, "Air temperature error correction based on solar radiation in an economical meteorological wireless sensor network," *Sensors*, vol. 15, no. 8, pp. 18114–18139, 2015.
- [4] B. Wang, X. Gu, M. L., and S. Yan, "Temperature error correction based on bp neural network in meteorological wireless sensor network," in *Proceedings of International Conference on Cloud Computing and Security*, 2016.
- [5] P. Ranacher, R. Brunauer, W. Trutschnig, S. V. der Spek, and S. Reich, "Why gps makes distances bigger than they are," *International Journal of Geographical Information Science*, vol. 30, no. 2, pp. 316–333, 2016.
- [6] V. A. Petrushin, G. Wei, and A. V. Gershman, "Multiple-camera people localization in an indoor environment," *Knowledge and Information Systems*, vol. 10, pp. 229–241, Aug 2006.
- [7] S. Zorn, R. Rose, A. Goetz, and R. Weigel, "A novel technique for mobile phone localization for search and rescue applications," in *Proceedings of IEEE International Conference on Indoor Positioning and Indoor Navigation*, pp. 1–4, 2010.
- [8] A. Redondi, M. Tagliasacchi, M. Cesana, L. Borsani, P. Tarrío, and F. Salice, "Laura — localization and ubiquitous monitoring of patients for health care support," in *Proceedings of the IEEE 21st International Symposium on Personal, Indoor and Mobile Radio Communications Workshops*, pp. 218–222, 2010.
- [9] A. Yassin, Y. Nasser, M. Awad, A. Al-Dubai, R. Liu, C. Yuen, R. Raulafs, and E. Aboutanios, "Recent advances in indoor localization: A survey on theoretical approaches and applications," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 1327–1346, 2017.
- [10] G. Deak, K. Curran, and J. Condell, "A survey of active and passive indoor localisation systems," *Computer Communications*, vol. 35, no. 16, pp. 1939 – 1954, 2012.
- [11] X. Luo, W. J. O'Brien, and C. L. Julien, "Comparative evaluation of received signal-strength index (rss) based indoor localization techniques for construction jobsites," *Advanced Engineering Informatics*, vol. 25, no. 2, pp. 355 – 363, 2011. Information mining and retrieval in design.

Procesos de Suavizados/Filtro	MSE
suavizado con modelo polinomial de segundo grado	0.001155739
filtro Savitzky-Golay	0.010586667
suavizado con modelo polinomial de primer grado	0.01239
filtro de movimiento promedio	0.02676
Chebyshev Type II	0.02969
Butterworth	0.298976667
Una versión robusta de lowess que asigna menor peso a los valores atípicos en la regresión	0.039403333
Filtro elíptico	0.553396667
Chebyshev Type I	0.59737
Una versión robusta de loess que asigna menor peso a los valores atípicos en la regresión	2.683873333

Cuadro I

MSE PROMEDIO DE CADA MÉTODO UTILIZADO PARA SUAVIZAR/FILTRAR LOS DATOS DE LOS EXPERIMENTOS

- [12] J. Shang, F. Gu, X. Hu, and A. Kealy, "Apfiloc: An infrastructure-free indoor localization method fusing smartphone inertial sensors, landmarks and map information," *Sensors*, vol. 15, no. 10, pp. 27251–27272, 2015.
- [13] M. L. L. Camberos, A. Diaz-Ramirez, J. Marquez-Barja, and V. Quintero, "Algoritmo de trazado de rutas en interiores sin infraestructura," in *In proceedings of International Conference on Computer Networks Applications*, pp. 64–69, Tecnológico Nacional de Mexico, 2017.
- [14] D. Ahn, J. . Park, C. . Kim, J. Kim, Y. Qian, and T. Itoh, "A design of the low-pass filter using the novel microstrip defected ground structure," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, pp. 86–93, Jan 2001.
- [15] D. Martino-Saltzman, B. B. Blasch, R. D. Morris, and L. W. McNeal, "Travel behavior of nursing home residents perceived as wanderers and nonwanderers," *The Gerontologist*, vol. 31, pp. 666–672, oct 1991.
- [16] B. A. Sheno, *Introduction to Digital Signal Processing and Filter Design*. Wiley Interscience, 2006.
- [17] M. D. Lutovac, D. V. Tomic, and B. L. Evans, *Filter Design for Signal Processing Using MATLAB and Mathematica*. Prentice Hall, 2000.
- [18] S. Butterworth, *On the Theory of Filter Amplifiers*, 1930.
- [19] S. Winder, *Analog and Digital Filter Design (EDN Series for Design Engineers)*. Newnes, 2002.
- [20] R. W. Daniels, *Approximation Methods for Electronic Filter Design: With Applications to Passive, Active and Digital Networks*. McGraw-Hill, 1974.
- [21] S. Orfanidis, *Introduction to Signal Processing*. Prentice Hall, 1995.
- [22] I. The MathWorks, *Filtering and Smoothing Data*. The MathWorks, Inc., 2019.
- [23] H. Chen and S. Chen, "A moving average based filtering system with its application to real-time QRS detection," in *Computers in Cardiology, 2003*, IEEE, 2003.
- [24] P. A. V. Oppenheim, *Lecture 12: Filtering*. MIT OpenCourseWare, Mar. 2019.
- [25] H. Villarreal, M. Lizarraga, A. Diaz-Ramirez, V. Quintero Rosas, and J. P. Garcia-Vázquez, "Componente del internet de las cosas para detectar patrones de deambulaje en pacientes con demencia," *Research in Computer Science*, vol. 148, no. 2, 2019.

# Sistema de generación de aplicaciones NFC para Android

José María Martínez Merlo  
Instituto Tecnológico de Orizaba

Departamento de estudios de posgrado e investigación  
Orizaba, México  
Email: iscmztjose@gmail.com

Ignacio López Martínez

Instituto Tecnológico de Orizaba  
Departamento de estudios de posgrado e investigación  
Orizaba, México  
Email: ilopez@ito-depi.edu.mx

Beatriz Alejandra Olivares Zepahua

Instituto Tecnológico de Orizaba  
Departamento de estudios de posgrado e investigación  
Orizaba, México  
Email: bolivares@ito-depi.edu.mx

Celia Romero Torres

Instituto Tecnológico de Orizaba  
Departamento de estudios de posgrado e investigación  
Orizaba, México  
Email: cromerotorres@hotmail.com

Luis Ángel Reyes Hernández

Instituto Tecnológico de Orizaba  
Departamento de estudios de posgrado e investigación  
Orizaba, México  
Email: l\_rh01@hotmail.com

**Resumen**—La problemática con las tecnologías emergentes, en especial NFC, es que caen en desuso, son sustituidas por otras tecnologías o simplemente desaparecen porque nunca llegaron a ser utilizadas como prometían, puesto que para usarlas se necesita forzosamente de la interacción de programadores que plasmen las ideas de los usuarios en aplicaciones funcionales. La eliminación de la figura del programador o bien la posibilidad de que los propios usuarios diseñen y construyan sus aplicaciones es una tendencia que cada vez toma más auge; un sistema generador de aplicaciones de NFC permitirá a los usuarios recurrir a la tecnología con las particularidades que cada consumidor necesite. El generador de aplicaciones de NFC establecerá una gama de usos generales y definirá una serie de acciones específicas que el usuario quiera personalizar permitiendo entonces la construcción de aplicaciones NFC específicas para usos particulares, las cuales pueden ser modificadas y actualizadas por los propios usuarios sin la necesidad de la intervención de un programador.

**Palabras clave**—NFC, generador de aplicaciones, desarrollo android.

## I. INTRODUCCIÓN

Actualmente los dispositivos móviles tienen la capacidad de utilizar una diversidad de aplicaciones, ya sea para juegos, utilidades o actividades empresariales, crear aplicaciones a medida para usos particulares de la industria y usuarios generales es cada vez más importante y a la vez difícil por la complejidad de cada una de estas. //

Entre las aplicaciones para fines particulares hay algunas que hacen uso de tecnologías emergentes con las que cuentan los dispositivos tales como NFC (Near Field Communication, Comunicación de Campo Cercano) que es una tecnología inalámbrica que proporciona comunicación entre dos dispositivos móviles que cuenten con tecnología NFC de ondas

de radio de corto alcance; utilizando la inducción de campo magnético para este fin [1].

NFC permite la comunicación entre dispositivos próximos, con un bajo nivel de poder de operación, está basado en la tecnología de RFID (Radio Frequency Identification, identificación por radiofrecuencias) y provee comunicación sin contacto entre dispositivos en un rango no mayor a 4 cm, opera en una frecuencia de 13.56 MHz y tiene una tasa de transferencia de hasta 424kb/s, el rango corto elimina cualquier interferencia, por lo que es una tecnología confiable [2].

La popularidad de NFC se incrementó en los últimos años, gracias a esto la industria de teléfonos inteligentes produce más dispositivos con estas funcionalidades y se integran sensores NFC en cada vez más dispositivos [3].

Un dispositivo NFC puede comunicarse también con NFC tags (Etiquetas NFC), las NFC tags contienen una pequeña cantidad de memoria y una antena, pueden ser configuradas como sea requerido, estas son dispositivos NFC pasivos, son generalmente de un tamaño pequeño (un par de pueden integrarse en objetos cotidianos [4].

Un generador de aplicaciones Android es una herramienta que permite facilitar el proceso de creación de aplicaciones. Todas las aplicaciones generadas por dichas herramientas siguen una misma estructura, por lo que se cuenta con código escrito en lenguaje de programación que contiene la estructura base de las aplicaciones y mecanismos necesarios para la creación de estas, creando un archivo XML (Extensible Markup Language, lenguaje de marcado extensible) como recurso para proveer la función esperada por el usuario [5].

Tipo	Características	Protocolo	Almacenamiento
Tipo 1 y 2	bajo costo y de aplicaciones únicas	ISO (International Organization for Standardization, organización internacional para estandarización)/IEC (International Electrotechnical Commission, Comisión Internacional Electrotécnica) 14443A	96 y 48 bytes respectivamente y son expandibles a 2 Kbytes.
Tipo 3	Especializadas y de uso en mercados asiáticos	Estándar industrial japonés X 6319-4.	Hasta 1 Mbyte.
Tipo 4	Utilizadas para aplicaciones con alta seguridad	Protocolo de intercambio de datos ISO-DEP.	32 Kbytes.
Tipo 5	Típicamente utilizadas industrialmente	ISO/IEC 15693.	8 Kbytes

Tabla I  
TIPOS DE TAGS NFC

## II. TRABAJOS RELACIONADOS

Después de la revisión de la literatura, no se encontró una herramienta como la que se propone en este documento; sin embargo existen muchos acercamientos que abonaron a la construcción y diseño del presente desarrollo; esta sección se divide en trabajos relacionados con NFC y aplicaciones móviles.

Alan y Birant [6] propusieron un sistema de cálculo de tarifas basadas en un servidor sin tener que obtener una tarjeta por cada ciudad, la creación del sistema propuesto se aplicó en cinco ciudades de Turquía con el estándar NFC (Near Field Communication, Comunicación de campo cercano) con el que recopilaban los datos para crear las reglas de asociación para el sistema propuesto utilizando el algoritmo FP-Growth como una técnica de minería de datos.

Tatiraju et al. [2] presentaron una alternativa de solución utilizando tarjetas de presentación compatibles con NFC donde los contactos comerciales son transferidos al teléfono inteligente del usuario, además de una aplicación de Android que permite al usuario navegar por todos los contactos y almacenar o recuperar estos de la nube. El modelo propuesto es una mejora sobre el existente sistema de almacenamiento de tarjetas de presentación usando NFC, ya que implementa estrategias de seguridad al utilizar encriptación, por lo que se evita el espionaje por parte de terceros.

Sharma et al. [1] definieron características de NFC como: tiene dos modos de uso activo y pasivo en ambos, cada etiqueta produce su propio campo electromagnético alternadamente para realizar el cambio de información, es compatible con RFID (Radio Frequency Identification, identificador de frecuencia de radio), entre otras. Además, compararon las ventajas y desventajas de NFC con tecnologías similares.

Los usos potenciales de NFC deben ser apoyados por bancos, comerciantes y compañías tecnológicas para proveer una plataforma segura a los usuarios de esta tecnología.

Las compras hasta hace poco se relacionaban con dinero en efectivo o tarjetas de crédito, con el auge de los teléfonos móviles nuevas formas de pago han surgido, por ejemplo, los Mobile payment, pagos por bienes, servicios y facturas con un dispositivo móvil como teléfono móvil, teléfono inteligente o asistente digital personal, aprovechando las tecnologías inalámbricas y otras tecnologías de comunicación, estos han crecido con la introducción de varios métodos de pago tales como el WAP (Wireless Application Protocol, Protocolo de Aplicación Inalámbrica), USSD (Unstructured Supplementary Service Data, Datos de Servicio Suplementario No Estructurado), SMS (Short Message Service, servicios de mensajes cortos), entre otros. Si bien cada método proporciona ventajas, estos carecen de la popularidad y gran uso de los métodos tradicionales de pago (efectivo, crédito y banca en línea).

En [7] indicaron las implicaciones teóricas útiles para la industria de desarrollo, mercados y manufactura, así como la aceptación de la NFC MCC (NFC Mobile credit card, Tarjeta de crédito móvil NFC) en el área de los ATM y mercados emergentes. Un tema difícil de abordar para los análisis estáticos (análisis que se realiza sin ejecutar el programa) es la reflexión (capacidad que tiene un programa para modificar su estructura normalmente a tiempo de ejecución), puesto que estos no la manejan adecuadamente, los marcos de trabajo de Android que hacen uso de análisis estático tienen gran cantidad de problemas debido a que generalmente eligen crear modelos en lugar de analizar el código dentro del marco de trabajo. Debido a que los modelos cuidadosamente diseñados son importantes se desea que sean reutilizables.

Blackshear et al. [8] presentaron una herramienta para el modelado de aplicaciones Android de propósito general para los análisis estáticos llamada DROIDEL, la cual genera automáticamente apéndices específicos de cada aplicación que obtienen el comportamiento reflexivo de una aplicación en particular.

La sustitución de llamadas reflexivas y la generación de código para simular el comportamiento se realiza a nivel de código de Java, la salida de DROIDEL es una aplicación Java con un solo punto de entrada que puede procesarse por cualquier plataforma de análisis existente.

En otro documento se menciona que la experiencia del visitante es la parte más importante de una visita al museo, esta comienza antes que llegar y continúa mucho tiempo después de la visita. Para mantener a los visitantes entretenidos, se experimentan con nuevos canales para lograr mayor participación de los visitantes, la intención principal es que la experiencia física de visitar un museo o un sitio histórico esté relacionada con los medios digitales.

En [9], se presentó el diseño, implementación, uso y evaluación de un Tangible Data Souvenir (recuerdo de datos tangible) para la interacción con exposiciones de un museo. Un Souvenir (recuerdo) se definió como la materialización de cómo se experimenta una visita por cada persona, y se



crea dinámicamente con una base de datos grabada durante la visita.

Los TDS se crean con información de las exposiciones en forma de cartas además, de contar con réplicas de piezas con sensores NFC que reproducen audios y videos dentro de lectores disponibles en los recorridos. Los TDS actúan además como la puerta de acceso a más contenido en línea.

NFC también permite una comunicación bidireccional de baja velocidad de datos entre dispositivos en un rango cercano, generalmente unos centímetros, la principal ventaja de NFC es eliminar la necesidad de realizar ajustes de configuración de red al utilizar alternativas como Bluetooth o WiFi. Esto se debe a su propiedad inherente de asociación por proximidad física; si dos dispositivos se pueden comunicar usando NFC, entonces implica que deben estar ubicados conjuntamente. En muchos casos, NFC se usa para iniciar y configurar automáticamente un canal de comunicación de alta velocidad de datos, como WiFi o Bluetooth. En [10] Nandakumar et al. describieron que la adopción de NFC se ve obstaculizada por los bajos niveles de penetración del hardware NFC debido a que la mayoría de los teléfonos no están habilitados para NFC por lo que las aplicaciones que usan esta tecnología seguirán estando muy limitadas.

Teniendo como base los trabajos relacionados antes mencionados se aprecia que hay muchas aplicaciones de la tecnología NFC de las cuales se realizó un análisis y se propone una arquitectura con una estructura que conforma el uso de NFC, desarrollo Android y generación de aplicaciones.

### III. ARQUITECTURA

El desarrollo del sistema fue realizado con la metodología UWE [11] puesto que fue la que mejor se adapta a las necesidades de un sistema basado en web, en la parte del SGBD se optó por MySQL al ser un gestor de código abierto y licencia libre además que este es popular para desarrollo de aplicaciones web, NetBeans es el IDE seleccionado gracias a la compatibilidad integrada con el desarrollo web utilizando JSF; debido a que el soporte para aplicaciones Android es superior en Android Studio, ya que este es el IDE recomendado en el mismo SDK de Android, se opta por este para realizar el desarrollo de la aplicación base que después se ocupará por el sistema para hacer aplicaciones personalizadas.

El sistema está basado en la web con el cual se pueden construir aplicaciones que hagan uso de la tecnología NFC en Android, donde el usuario personaliza la aplicación partiendo de una estructura base, para después incluir las personalizaciones utilizando lenguaje XML y por último generar un archivo APK que será entregado al usuario.

El generador de aplicaciones NFC sigue el patrón MVC mostrado en la Figura 1, donde las funciones de interacción con el usuario estarán en la vista, el controlador basado en Java tendrá las funciones de RMI [12] (Remote Method Invocation, Método de invocación remoto), las cuales traducirán los elementos con los que interactúa la capa de vista traduciéndose a XML para después interpretarse a código java dentro de la consola, además de contar con el acceso a la base de datos

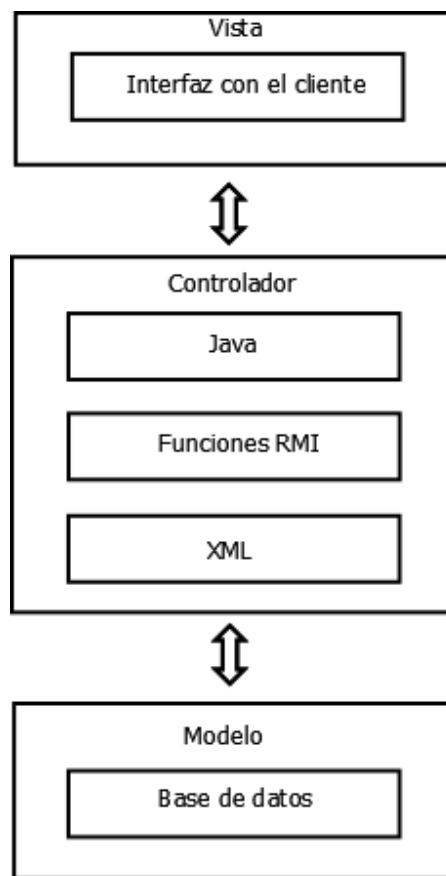


Figura 1. Diagrama MVC

para la autenticación del usuario y despliegue de cambios realizados por este anteriormente.

El proceso de construcción de las aplicaciones Android se muestra en la Figura 2, donde el generador invoca a las funciones de construcción por medio de RMI.

Debido a que el sistema utiliza JavaServerFaces, en el diagrama de casos de clases mostrado en la Figura 3 se tienen las clases principales de la aplicación, en donde la clase aplicación contendrá los atributos de las aplicaciones Android y la clase generador hará uso de los métodos RMI.

El generador hace uso de un código base de Android Es implementado con las clases NdefMessage y NdefRecord utilizando Java, provee los métodos para obtener y modificar los datos dentro de un Tag NFC que puede ser del tipo 1 al tipo 4, para el generador se utilizan etiquetas correspondientes al tipo 4 utilizando los estándares: MifareClassic y MifareUltralight.

La lectura del formato NDEF [13](NFC Data Exchange Format, Formato de intercambio de información NFC) se maneja con un sistema de despachamiento de tags que analiza el tag descubierto, clasifica los datos e inicia una aplicación que reciba estos datos categorizados, al tener una aplicación que maneje tags NFC se declara un filtro de intención y se solicita el manejo de datos.

El desarrollo de la aplicación base se realiza con el IDE

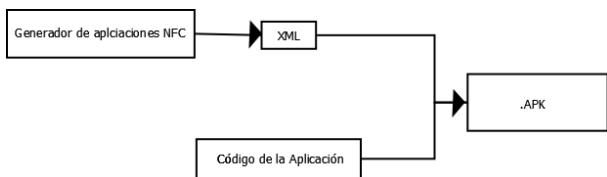


Figura 2. Proceso de construcción de la app



Figura 3. Diagrama de casos de clases

Android Studio, dicho IDE nos permite crear fácilmente dicha aplicación, para hacer uso de NFC dentro de una aplicación debe modificarse el manifiesto de la aplicación utilizando ; uses - permission android:name= "android.permission.NFC" /;, la API nivel 9 es la mínima versión soportada [14]. La generación de aplicaciones es posible en la línea de comando de Windows utilizando el contenedor de Gradlew. Dentro del generador se invoca la instrucción de uso del contenedor de Gradlew directamente en una clase java [15].

El código para la creación de la aplicación se muestra en la Figura 5, la aplicación puede utilizarse en cualquier dispositivo Android que integre NFC y Android 2.1 Eclair de la misma forma el usuario podrá realizar cambios a la aplicación así le convenga desde la pantalla de edición de el sistema.

La interacción con el sistema se lleva a cabo como se muestra en la Figura 6, Las funciones dentro del sistema que el usuario podrá llevar a cabo se muestran en la Figura 7.

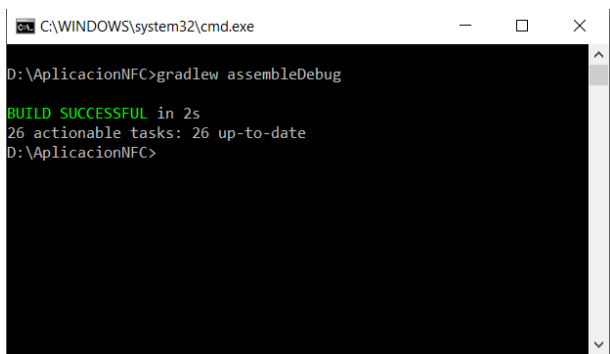


Figura 4. Generacion por medio de Gradlew

IV. CONCLUSIONES

El desarrollo de aplicaciones móviles es se ha convertido en parte importante de la computación, ya que la mayoría de los dispositivos móviles se ocupan día con día para distintas actividades en las que podría integrarse tecnología NFC.

```
public class Generador {
    public static void main(String[] args) {
        System.out.println("Se va a generar la apk");
        ejecutarComando("cmd /c cd D:\\AplicacionNFC assembleDebug");
    }
    public static void ejecutarComando(String comando){
        try {
            String comando="cmd /c cd D:\\AplicacionNFC gradlew assembleDebug";
            Process process = Runtime.getRuntime().exec(comando);
            InputStream inputStream = process.getInputStream();
            BufferedInputStream bufferedInputStream = new BufferedInputStream(inputStream);
            System.out.println("Aplicacion generada");
        } catch (IOException e) {
            System.out.println("Excepción: no se pudo generar el APK");
        }
    }
}
```

Figura 5. Código de generacion desde una clase java

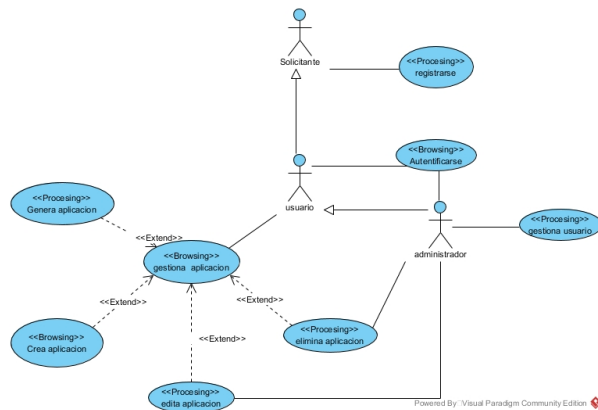


Figura 6. Diagrama de casos de uso

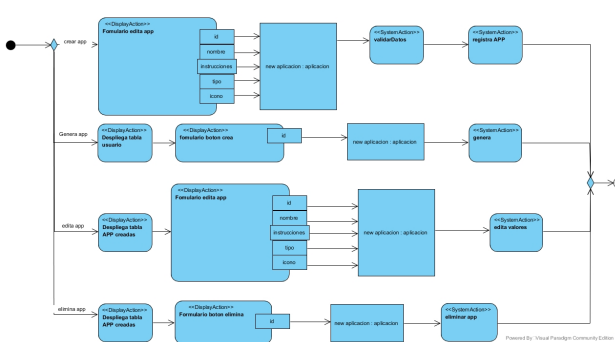


Figura 7. Diagrama de funciones de usuario

El desarrollo del sistema generador de aplicaciones Android NFC realizado en un entorno web, permitirá que el uso de las aplicaciones NFC sea más frecuente y se consolide como una tecnología de uso diario.

El desarrollo mencionado en este documento se encuentra en proceso, con casos de estudio, ya concluyó la fase de prototipo (se encuentra en la fase de pruebas piloto con casos de estudio), se pretende publicarlo en la web para pruebas de efectividad de servicio y medida de usabilidad.

El estado del arte sirvió para determinar;

- Patrón de diseño.
- Metodología.
- Tecnología java de desarrollo.

En este documento se explicó a grandes rasgos la estructura principal del sistema, cabe mencionar que el correcto fun-

cionamiento de las aplicaciones generadas con la herramienta dependerá de que se ocupen los Tags NFC adecuados y dado que el SDK de android es actualizado frecuentemente no se puede asegurar la compatibilidad completa en los nuevos dispositivos.

Los usuarios que hagan uso de esta herramienta tendrán una aplicación funcional la cual es compatible con los dispositivos Android que hagan uso de NFC, así como la posibilidad de crear más aplicaciones conforma a las necesidades que tengan en un futuro de forma sencilla y sin necesidad de recurrir a un programador.

#### REFERENCES

- [1] S. VIBHOR, P. GUSAIN, AND K. PRASHANT, *Near Field Communication*, Adv. Intell. Syst. Res., vol. 2013, no. Cac2s, pp. 342–345, 2013
- [2] R. TATIRAJU, R. GUPTA, S. SALUNKE, AND R. YEOLEKAR, *NFiD: An NFC based system for Digital Business Cards*, pp. 1510–1512, 2017
- [3] A. MESCHTSCHERJAKOV, C. GSCHWENDTNER, M. TSCHELIGI, AND P. SUNDSTRÖM, *Co-designing for NFC and ATMs: An Inspirational Bits Approach*, in Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services, 2013, pp. 422–427
- [4] U. JAMBUSARIA, N. KATWALA, AND D. MISTRY, *Secure smartphone unlocking using NFC*, Procedia Comput. Sci., vol. 45, no. C, pp. 465–469, 2015
- [5] L. BANDA AND I. LÓPEZ MARTÍNEZ, *Proceso para la Generación Automática de Aplicaciones de Realidad Aumentada para Dispositivos Móviles*, in Congreso Internacional de Investigación e Innovación en Ingeniería de Software, CONISOFT'13, 2013
- [6] U. D. ALAN AND D. BIRANT, *Server-Based Intelligent Public Transportation System with NFC*, IEEE Intell. Transp. Syst. Mag., vol. 10, no. 1, pp. 30–46, 2018
- [7] G. W. H. TAN, K. B. OOI, S. C. CHONG, AND T. S. HEW, *NFC mobile credit card: The next frontier of mobile payment?*, Telemat. Informatics, vol. 31, no. 2, pp. 292–307, 2014
- [8] S. BLACKSHEAR, A. GENDREAU, AND B.-Y. E. CHANG, *Droidel: A General Approach to Android Framework Modeling*, in Proceedings of the 4th ACM SIGPLAN International Workshop on State Of the Art in Program Analysis, 2015, pp. 19–25
- [9] D. PETRELLI, M. T. MARSHALL, S. O'BRIEN, P. MCENTAGGART, AND I. GWILT, *Tangible data souvenirs as a bridge between a physical museum visit and online digital experience*, Pers. Ubiquitous Comput., vol. 21, no. 2, pp. 281–295, 2017
- [10] R. NANDAKUMAR, K. K. CHINTALAPUDI, V. PADMANABHAN, AND R. VENKATESAN, *Dhwani: Secure peer-to-peer acoustic NFC*, Annu. Conf. ACM Spec. Interes. Gr. Data Commun. Appl. Technol. Archit. Protoc. Comput. Commun. ACM SIGCOMM 2013, vol. 43, no. 4, pp. 63–74, 2013
- [11] N. KOCH, A. KNAPP, G. ZHANG, AND H. BAUMEISTER, *Uml-Based Web Engineering*, in Web Engineering: Modelling and Implementing Web Applications, G. Rossi, O. Pastor, D. Schwabe, and L. Olsina, Eds. London: Springer London, 2008, pp. 157–191
- [12] ORACLE, *Getting Started Using Java™ RMI*, 2018. [Online]. Available: <https://docs.oracle.com/javase/7/docs/technotes/guides/rmi/hello/hello-world.html#start>
- [13] GOOGLE, *Ndef*, [Online]. Available: [https:// developer.android.com/reference/android/nfc/tech/Ndef](https://developer.android.com/reference/android/nfc/tech/Ndef). [Accessed: 11-Dec-2018]
- [14] ANDROID DEVELOPERS, *NFC BASICS*, 2018. [Online]. Available: <https:// developer.android.com/guide/topics/connectivity/nfc/nfcjava>
- [15] ANDROID DEVELOPERS, *Compilar Aplicacion desde línea de comando*, 2018. Available: <https:// developer.android.com/studio/build/building-cmdline?hl=es-419>

# Formación de una Red Heterogénea de Robots mediante Acoplamiento Difusivo Estático

Durazo-Acevedo Salvador, Aguilar-Bustos Ana Yaveni, Bugarin Carlos Eusebio  
División de Estudios de Posgrado e Investigación  
Tecnológico Nacional de México/I.T. Ensenada  
Ensenada, Baja California, México  
(correo-e: *sdurazo; aguilar; ebugarin@ite.edu.mx*)

**Resumen---**En este trabajo se describe un procedimiento para la formulación de redes heterogéneas de robots. El procedimiento se desarrolla en base al acoplamiento difusivo estático. La efectividad de la formulación se ejemplifica para una red con cinco robots que contiene dos tipos distintos de robots móviles terrestres a base de ruedas. Para demostrar la aplicabilidad del procedimiento a redes de robots se propone un objetivo de formación. El cumplimiento de dicho objetivo de formación se demuestra mediante la teoría de estabilidad en el sentido de Lyapunov y se presenta la simulación para la formación con y sin desplazamiento en formación.

**Palabras clave--**Redes de robots; acoplamiento difusivo; formulación; control de robots.

## I. INTRODUCCIÓN

En el presente trabajo se expone un procedimiento para formular una red con grupos de robots que estructural y funcionalmente son distintos. El procedimiento se apoya en la teoría de grafos y una selección muy particular de estrategia para el acoplamiento entre robots. Se describe además la estrategia de acoplamiento difusivo estático, la cual es una evolución natural de la formulación sistemática de red.

Una red de robots (RR) es heterogénea (RHeR) si los robots que la integran exhiben variantes dinámicas y funcionales [3, 6, 12, 17]. Estas variantes pueden corresponder a diferencias estructurales muy notorias como la construcción de sus articulaciones, distintas dimensiones de vectores de variables de estado, distinta dimensión de variables de salida, su forma de movilidad o bien pueden ser diferencias más sutiles como el alcance de comunicación o el ángulo de apertura de sus sensores. En un sentido estricto, todas las redes de robots son heterogéneas, ya que, bajo igualdad de condiciones en sus señales de entrada cada robot podría comportarse de formas distintas en virtud de sus diferencias sutiles.

La formación consiste en posicionar y orientar espacialmente en  $\mathbb{R}^2$  ó en  $\mathbb{R}^3$ , con cierta distribución deseada, a los robots móviles que forman parte de una red [9]. Dicha formación comúnmente es una etapa preliminar para el desarrollo de tareas colaborativas más complejas. Como ejemplos: (i) La formación para *seguimiento de líder* y la formación para *cobertura espacial*.

Particularmente se han propuesto y experimentado, en diversas redes, algunas leyes de control con el propósito de forzar una respuesta deseada denominada estado de sincronía [10, 12]. Para una red de robots, la sincronía se refiere a la habilidad de coordinar sus movimientos o su postura de manera inducida mediante la inserción de una señal de referencia. También se han propuesto y experimentado algunas leyes de control para lograr el consenso [13, 19]. Para una red de robots, el consenso se refiere a la habilidad de coordinar su postura y movimientos como resultado del acoplamiento mismo de red. En la actualidad el problema para sincronizar o consensar en una RHeR persiste en dos sentidos: el primero por su complejidad por la dinámica propia de los robots, por la complejidad en la estructura de la red y por su dificultad de coordinar un número elevado de robots en la red [16], y el segundo, en virtud de no contar con un procedimiento sistemático para formular la red y proponer leyes específicas de control según el objetivo de la red de robots.

Para resolver el problema de formación de una RR, tanto homogéneas como heterogéneas, se han utilizado distintos procedimientos y estrategias; por ejemplo: en [1] se utiliza el consenso y se valida experimentalmente en una red homogénea, en [2] se presenta una estrategia de control distribuido para lograr una formación rígida de red homogénea de robots en dos dimensiones, en [8 y 9] se aplica una ley de control no continua para estabilización de formación de una red homogénea mediante consenso, en [8] se incluye además la evasión de colisión entre robots, en [14] se propone la formación de una red heterogénea de robots utilizando un procedimiento de unificación de modelos dinámicos entre los robots y en [15] se utiliza la teoría de redes difusas neuronales para estabilizar una red heterogénea de robots omnidireccionales.

En la sección a continuación se aborda una propuesta de procedimiento sistemático para la formulación de ley de control para estas RHeRs; en la sección 3 se ejemplifica la formulación particular para una red con dos modelos distintos de robots móviles y las restricciones necesarias para implementar dicha formulación; en la sección 4 se incluye un objetivo de formación y una ley de control; también se describe la metodología para alcanzar el objetivo propuesto y se presentan simulaciones particulares para validar tanto la ley de control como el procedimiento de formulación.

## II. FORMULACIÓN DE UNA RED DE ROBOTS

Considérese el modelo general de un robot expresado por la ecuación diferencial [4, 11]

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}, \mathbf{u}), \\ \mathbf{y} &= \mathbf{h}(t, \mathbf{x}, \mathbf{u}),\end{aligned}\quad (1)$$

donde  $\mathbf{x} \in \mathbb{R}^r$  es el vector de variables de estado,  $\mathbf{y} \in \mathbb{R}^s$  es el vector de variables de salida,  $t$  el tiempo continuo y  $\mathbf{u} \in \mathbb{R}^p$  es el vector de variables de entrada. Tómese  $\mathbf{f}: \mathbb{R} \times \mathbb{R}^r \times \mathbb{R}^p \rightarrow \mathbb{R}^r$  y  $\mathbf{h}: \mathbb{R} \times \mathbb{R}^r \times \mathbb{R}^p \rightarrow \mathbb{R}^s$  como funciones continuas.

Considere un robot  $R_i$  que pertenece a una red de  $N$  robots, y considere un vector de entrada  $\boldsymbol{\vartheta}_i \in \mathbb{R}^{v_i}$  que representa a la señal externa para el control, de manera que su modelo es

$$\begin{aligned}\dot{\mathbf{x}}_i &= \mathbf{f}_i(t, \mathbf{x}_i, \mathbf{u}_i), \\ \mathbf{y}_i &= \mathbf{h}_i(t, \mathbf{x}_i, \mathbf{u}_i), \\ \mathbf{u}_i &= \mathbf{Y}_i(t, \mathbf{y}, \boldsymbol{\vartheta}_i),\end{aligned}\quad (2)$$

donde ahora  $\mathbf{f}_i: \mathbb{R} \times \mathbb{R}^{r_i} \times \mathbb{R}^{p_i} \rightarrow \mathbb{R}^{r_i}$ ,  $\mathbf{h}_i: \mathbb{R} \times \mathbb{R}^{r_i} \times \mathbb{R}^{p_i} \rightarrow \mathbb{R}^{s_i}$  y la función  $\mathbf{Y}_i: \mathbb{R} \times \mathbb{R}^s \times \mathbb{R}^{v_i} \rightarrow \mathbb{R}^{p_i}$  es un mapeo de  $\mathbf{y}$  y  $\boldsymbol{\vartheta}_i$  hacia el vector de excitación  $\mathbf{u}_i$ .

Aplicando la teoría de grafos [4, 16], se denota  $\mathcal{G} = (R, \mathcal{E})$  al conjunto de robots y el conjunto de conexiones entre ellos, donde  $R = \{R_i: i = 1, 2, \dots, N\}$  y  $\mathcal{E} = \{\mathcal{E}_{ij}(\mathbf{y}_j): i, j = 1, 2, \dots, N\}$  es el conjunto de  $N \times N$  conexiones dirigidas de la red. Una conexión dirigida  $\mathcal{E}_{ij}(\mathbf{y}_j) \in \mathbb{R}^{p_i}$  es una función que corresponde a un par de robots  $(R_i, R_j)$ , siendo la salida  $\mathbf{y}_j \in \mathbb{R}^{s_j}$  (salida del robot  $R_j$ ) el vector de información que se transfiere al robot  $R_i$  mediante una transformación  $\mathcal{E}_{ij}: \mathbb{R}^{s_j} \rightarrow \mathbb{R}^{p_i}$ .

Para la construcción de la función  $\mathbf{Y}_i$  se puede proponer que cada  $\mathcal{E}_i(\mathbf{y})$  [5, 12, 16, 18] se exprese como

$$\mathcal{E}_i(\mathbf{y}) = c \sum_{j=1}^N \mathcal{E}_{ij}(\mathbf{y}_j), \quad (3)$$

donde  $c > 0$  es la fuerza global de acoplamiento y  $\mathcal{E}_{ij}(\mathbf{y}_j) = g_{ij} \boldsymbol{\Gamma}_{ij} \mathbf{y}_j$ , con  $g_{ij}$  elemento de  $\mathbf{G} \in \mathbb{R}^{N \times N}$ , que tiene la particularidad de que cada  $g_{ii} = -\sum_{j=1, j \neq i}^N g_{ij}$ , por lo que la suma de los valores de sus elementos por renglón es nula;  $\boldsymbol{\Gamma}_{ij} \in \mathbb{R}^{p_i \times s_j}$  es la matriz de configuración interna con elementos  $\gamma_{ij,kl} \geq 0$  con  $i, j = 1, 2, \dots, N$ ,  $k = 1, 2, \dots, p_i$  y  $l = 1, 2, \dots, s_j$  que expresan las proporciones de contribución de las variables de salida  $\mathbf{y}_j$  hacia  $\mathbf{u}_i$  y en consecuencia, define las proporciones entre las variables de salida del robot  $R_j$  hacia las variables de entrada del robot  $R_i$ .

La generación de  $\mathbf{u}_i$  con acoplamiento difusivo estático y considerando  $\boldsymbol{\vartheta}_i$  en forma aditiva es [7]

$$\mathbf{u}_i = c \sum_{j=1}^N g_{ij} \boldsymbol{\Gamma}_{ij} \mathbf{y}_j + \boldsymbol{\vartheta}_i, \quad i = 1, 2, \dots, N. \quad (4)$$

## III. FORMULACIÓN DE UNA RHeR

Para la formulación de la RHeR en este trabajo se utilizan los dos tipos distintos de robots mostrados en la Figura 1.

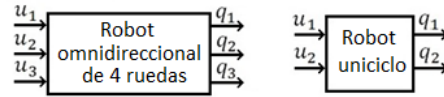


Figura 1. Representación simplificada para los robots: a) un robot omnidireccional, b) un robot tipo unicycle.

El modelo individual para ambos tipos de robots como parte de la red es

$$\begin{aligned}\dot{\mathbf{q}}_i &= \mathbf{u}_i, \\ \mathbf{y}_i &= \mathbf{q}_i, \\ \mathbf{u}_i &= \mathbf{Y}_i(t, \mathbf{q}, \boldsymbol{\vartheta}_i), \quad i = 1, 2, \dots, N.\end{aligned}\quad (5)$$

Utilizando el acoplamiento difusivo estático de acuerdo a (4) y para  $i = 1, 2, \dots, N$

$$\mathbf{u}_i = c \sum_{j=1}^N g_{ij} \boldsymbol{\Gamma}_{ij} \mathbf{q}_j + \boldsymbol{\vartheta}_i, \quad (6)$$

donde se ha sustituido  $\mathbf{y}_j = \mathbf{q}_j$ .

Para definir cada  $\boldsymbol{\Gamma}_{ij}$  de una red heterogénea se consideran las siguientes restricciones:

*Restricción 1.* La topología de red corresponde a una con raíz única y en la cual se inserta un líder virtual como referencia para la formación.

*Restricción 2.* El vector de variables de salida del líder virtual es homólogo con el vector de mayor orden de variables de salida de los robots; es decir, que corresponden a la misma condición de los robots omnidireccionales.

*Restricción 3.* Para cada robot  $R_i$ , existe un líder particular (referencia individual) y el cual sólo podrá ser un robot (o el líder virtual) con igual o mayor orden en su vector de variables de salida que el orden  $s_i$  del robot  $R_i$ .

*Restricción 4.* Suponiendo que existe la conexión  $\mathcal{E}_{ij}$ , el robot  $R_i$  tiene disponible la información de todas las salidas del robot  $R_j$ .

Dadas tales restricciones,  $\boldsymbol{\Gamma}_{ij}$  se propone con las siguientes connotaciones:

$\boldsymbol{\Gamma}_{ii} = \mathbf{I}_{s_i}$ , matriz identidad de orden  $s_i$ ;  $\boldsymbol{\Gamma}_{ij} = \mathbf{I}_{s_i}$  si  $s_i = s_j$ ;  $\boldsymbol{\Gamma}_{ij} = [\mathbf{I}_{s_i} \ \mathbf{0}]$  si  $i > j$  y si  $s_j > s_i$  con la matriz nula  $\mathbf{0} \in \mathbb{R}^{s_i \times (s_j - s_i)}$  y  $\boldsymbol{\Gamma}_{ij} = [\mathbf{I}_{s_j} \ \mathbf{0}]^T$  si  $j > i$  y si  $s_i > s_j$  con la matriz nula  $\mathbf{0} \in \mathbb{R}^{(s_i - s_j) \times s_j}$ .

Para ejemplificar la formulación de una RHeR considérese la red con acoplamiento difusivo estático, direccionada, no ponderada, construida con dos robots omnidireccionales y tres robots tipo unicycle como se ilustra en la Figura 2.

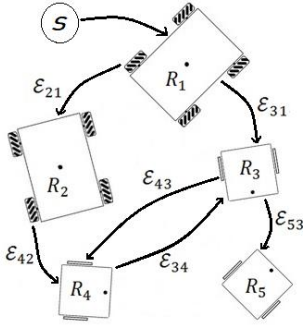


Figura 2. Red heterogénea de  $N = 5$  robots, direccionada y no ponderada, con  $R_1$  y  $R_2$  robots omnidireccionales y  $R_3, R_4,$  y  $R_5$  robots tipo unicycle.

La matriz de configuración externa resulta

$$\mathbf{G} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -2 & 1 & 0 \\ 0 & 1 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix}. \quad (7)$$

Cada  $\mathbf{\Gamma}_{ij}$  se define de la siguiente manera

$$\begin{aligned} \mathbf{\Gamma}_{ij} &= \mathbf{I}_3, \quad i, j = 1, 2. \\ \mathbf{\Gamma}_{ij} &= \mathbf{I}_2 \text{ para } i, j = 3, 4, 5. \\ \mathbf{\Gamma}_{ij} &= [\mathbf{I}_2 \ \mathbf{0}] \text{ para } i = 3, 4, 5 \text{ y } j = 1, 2 \text{ con } \mathbf{0} \in \mathbb{R}^{2 \times 1}. \\ \mathbf{\Gamma}_{ij} &= [\mathbf{I}_2 \ \mathbf{0}]^T \text{ para } i = 1, 2 \text{ y } j = 3, 4, 5 \text{ con } \mathbf{0} \in \mathbb{R}^{1 \times 2}. \end{aligned}$$

#### IV. FORMACIÓN

Considérese para la RHeR el objetivo de formación

$$\lim_{t \rightarrow \infty} \|\mathbf{e}_i(t)\| = 0, \quad i = 1, 2, \dots, N, \quad (8)$$

donde  $\mathbf{e}_i = \mathbf{q}_i - (\mathbf{\Gamma}_{ik}\mathbf{s}_i + \boldsymbol{\varphi}_i)$  es el error que se expresa con la diferencia entre los estados  $\mathbf{q}_i$  del robot  $R_i$  y un vector  $\mathbf{\Gamma}_{ik}\mathbf{s}_i + \boldsymbol{\varphi}_i$  fijado respecto a los estados de un líder particular  $\mathbf{s}_i$  y un desplazamiento deseado  $\boldsymbol{\varphi}_i$ .  $\mathbf{\Gamma}_{ik}$  representa un acoplamiento dimensional entre las variables de salida  $\mathbf{q}_k$  del líder particular para el robot  $R_i$  y las variables de salida  $\mathbf{q}_i$  del robot  $R_i$  bajo el contexto de las restricciones 1, 2, 3 y 4.

Supóngase que es posible seleccionar convenientemente un líder particular  $\mathbf{s}_i$  para cada robot de la red, donde al menos un robot posee un líder general con variables de salida  $\mathbf{q}_0 \in \mathbb{R}^{\max(s_i)}$ . Con la excepción de los robots cuyo líder particular es el líder general, para los demás robots se sugiere como líder particular a algún otro robot vecino con el cual se obtenga la trayectoria más corta hacia el líder general. Cada vector de postura final deseada es  $\boldsymbol{\varphi}_i = [\varphi_{i1} \ \varphi_{i2} \ \dots \ \varphi_{is_i}]^T$ . Los vectores de postura deseada se especifican respecto a la postura de su líder particular  $\mathbf{s}_i$ . La derivada de  $\mathbf{e}_i$  es  $\dot{\mathbf{e}}_i = \dot{\mathbf{q}}_i - \mathbf{\Gamma}_{ik}\dot{\mathbf{s}}_i$  o bien, sustituyendo  $\dot{\mathbf{q}}_i$  de acuerdo a (5) se tiene

$$\dot{\mathbf{e}}_i = \mathbf{u}_i - \mathbf{\Gamma}_{ik}\dot{\mathbf{s}}_i. \quad (9)$$

##### A. Ley de control propuesta

Para cada robot  $R_i$ , la ley de control propuesta para el objetivo de formación (8) es

$$\mathbf{u}_i = -cd_i\mathbf{\Gamma}_{ii}\mathbf{e}_i - c\sum_{j=1}^N g_{ij}\mathbf{\Gamma}_{ij}(\mathbf{s}_j + \boldsymbol{\varphi}_j) + \mathbf{\Gamma}_{ik}\dot{\mathbf{s}}_i, \quad (10)$$

$$i = 1, 2, \dots, N$$

donde  $d_i \geq 0$  es una ganancia para el controlador en el robot  $R_i$  ( $\mathbf{D} = \text{diag}\{d_i\}$ ). Luego de sustituir (10) en (6) resulta

$$\mathbf{u}_i = c\sum_{j=1}^N g_{ij}\mathbf{\Gamma}_{ij}\mathbf{q}_j - cd_i\mathbf{\Gamma}_{ii}\mathbf{e}_i - c\sum_{j=1}^N g_{ij}\mathbf{\Gamma}_{ij}(\mathbf{s}_j + \boldsymbol{\varphi}_j) + \mathbf{\Gamma}_{ik}\dot{\mathbf{s}}_i, \quad i = 1, 2, \dots, N, \quad (11)$$

y dado que  $\mathbf{q}_j = \mathbf{e}_j + \mathbf{s}_j + \boldsymbol{\varphi}_j$ , la dinámica del error  $\mathbf{e}_i$  resulta

$$\dot{\mathbf{e}}_i = c\sum_{j=1}^N g_{ij}\mathbf{\Gamma}_{ij}\mathbf{e}_j - cd_i\mathbf{\Gamma}_{ii}\mathbf{e}_i, \quad i = 1, 2, \dots, N. \quad (12)$$

Mediante el desarrollo de  $\mathbf{e}_i$  para  $i = 1, 2, \dots, N$  se tiene

$$\begin{bmatrix} \dot{\mathbf{e}}_1 \\ \dot{\mathbf{e}}_2 \\ \vdots \\ \dot{\mathbf{e}}_N \end{bmatrix} = c \begin{bmatrix} -\delta_1\mathbf{\Gamma}_{11} & g_{12}\mathbf{\Gamma}_{12} & \dots & g_{1N}\mathbf{\Gamma}_{1N} \\ g_{21}\mathbf{\Gamma}_{21} & -\delta_2\mathbf{\Gamma}_{22} & \dots & g_{2N}\mathbf{\Gamma}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ g_{N1}\mathbf{\Gamma}_{N1} & g_{N2}\mathbf{\Gamma}_{N2} & \dots & -\delta_N\mathbf{\Gamma}_{NN} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_N \end{bmatrix}, \quad (13)$$

donde cada  $\delta_i = d_i - g_{ii}$ . De manera que la dinámica del error en lazo cerrado es

$$\dot{\mathbf{e}} = -\mathbf{F}\mathbf{e}, \quad (14)$$

##### B. Equilibrio y estabilidad

Considérese el punto de equilibrio  $\mathbf{e} = \mathbf{0}$  de (14) y considérese la siguiente función candidata de Lyapunov

$$\mathbf{V} = \frac{1}{2}\mathbf{e}^T\mathbf{e} > 0 \quad \forall \mathbf{e} \neq \mathbf{0}, \quad (15)$$

por lo que la derivada temporal de  $\mathbf{V}$  es

$$\dot{\mathbf{V}} = \mathbf{e}^T\dot{\mathbf{e}}. \quad (16)$$

Tomando  $\dot{\mathbf{e}}$  de (14) y sustituyendo en (16) se tiene

$$\dot{\mathbf{V}} = -\mathbf{e}^T\mathbf{F}\mathbf{e}, \quad (17)$$

con



$$\mathbf{F} = c \begin{bmatrix} \delta_1 \mathbf{I}_3 & \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{O} \\ -\mathbf{I}_3 & \delta_2 \mathbf{I}_3 & \mathbf{O} & \mathbf{O} & \mathbf{O} \\ -\Gamma_{31} & \mathbf{O} & \delta_3 \mathbf{I}_2 & -\Gamma_{34} & \mathbf{O} \\ \mathbf{O} & -\Gamma_{42} & -\Gamma_{43} & \delta_4 \mathbf{I}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & -\Gamma_{53} & \mathbf{O} & \delta_5 \mathbf{I}_2 \end{bmatrix}, \quad (18)$$

donde  $\mathbf{O}$  es una matriz rectangular nula de dimensión conveniente de acuerdo a  $g_{ij}\Gamma_{ij} = \mathbf{O}$  si  $g_{ij} = 0$ .

Note que  $\dot{\mathbf{V}} < 0 \forall \mathbf{e} \neq \mathbf{0}$  si  $\mathbf{F} > 0$ ; por lo tanto, si  $\mathbf{F} > 0$ , entonces  $\mathbf{e} = \mathbf{0}$  es un punto de equilibrio asintóticamente estable en forma global y entonces se demuestra el cumplimiento del objetivo de formación (8).

### C. Simulación de formación para una RHeR con $N = 5$ robots

Considérese la formulación de la RHeR de la Figura 2 y la relación de los líderes particulares de cada robot conforme lo señala la Tabla I.

Tabla I. Selección de líderes particulares y descripción de acoplamientos.

$R_i$	$R_k$	$\Gamma_{ik}\mathbf{q}_k$ ( $\mathbf{s}_i = \mathbf{q}_k$ )	$\Gamma_{ik}\dot{\mathbf{q}}_k$
$R_1$	$R_0^a$	$\Gamma_{11}\mathbf{q}_0 = \Gamma_{11}[q_{01} \ q_{02} \ q_{03}]^T$	$\Gamma_{11}\dot{\mathbf{q}}_0$
$R_2$	$R_1$	$\Gamma_{21}\mathbf{q}_1 = \Gamma_{21}[q_{11} \ q_{12} \ q_{13}]^T$	$\Gamma_{21}\dot{\mathbf{q}}_1$
$R_3$	$R_1$	$\Gamma_{31}\mathbf{q}_1 = \Gamma_{31}[q_{11} \ q_{12} \ q_{13}]^T$	$\Gamma_{31}\dot{\mathbf{q}}_1$
$R_4$	$R_2$	$\Gamma_{42}\mathbf{q}_2 = \Gamma_{42}[q_{21} \ q_{22} \ q_{23}]^T$	$\Gamma_{42}\dot{\mathbf{q}}_2$
$R_5$	$R_3$	$\Gamma_{53}\mathbf{q}_3 = \Gamma_{53}[q_{31} \ q_{32}]^T$	$\Gamma_{53}\dot{\mathbf{q}}_3$

<sup>a</sup> Líder general virtual.

Nótese que el rango $\{\mathbf{F}\} = 9$  ya que el renglón 1 de  $\mathbf{G}$  es nulo. Se propone la matriz  $\mathbf{D} = \text{diag}\{d_1, 0, \dots, 0\}$  debido a que  $R_1$  tiene trayectorias direccionadas hacia el resto de los robots. Para que el rango $\{\mathbf{F}\}$  sea completo resulta suficiente  $d_1 \neq 0$  (lo que se comprueba con el procedimiento de eliminación de Gauss) y entonces  $\mathbf{e} = \mathbf{0}$  es un punto de equilibrio único de (14).

Si se propone  $d_1 > 0$  y  $d_2 = d_3 = d_4 = d_5 = 0$ , esto se sustituye en (18) y con  $\mathbf{A}_s = [\mathbf{F} + \mathbf{F}^T]/2$  se tiene

$$\mathbf{A}_s = \frac{c}{2} \begin{bmatrix} 2d_1 \mathbf{I}_3 & -\mathbf{I}_3 & -\Gamma_{13} & \mathbf{O} & \mathbf{O} \\ -\mathbf{I}_3 & 2\mathbf{I}_3 & \mathbf{O} & -\Gamma_{24} & \mathbf{O} \\ -\Gamma_{31} & \mathbf{O} & 4\mathbf{I}_2 & -2\Gamma_{34} & -\Gamma_{35} \\ \mathbf{O} & -\Gamma_{42} & -2\Gamma_{43} & 4\mathbf{I}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & -\Gamma_{53} & \mathbf{O} & 2\mathbf{I}_2 \end{bmatrix}, \quad (19)$$

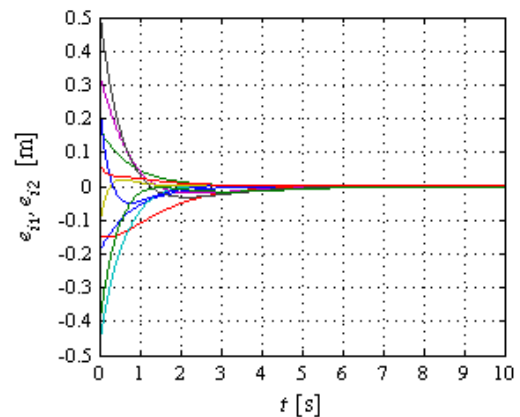
El análisis de acuerdo al Teorema de Silvester [11], para que  $\mathbf{F} > 0$  es suficiente que  $d_1 > 21/33$  y se cumple con que  $\mathbf{e} = \mathbf{0}$  es el único punto de equilibrio y el objetivo de formación (8) es satisfecho de forma global.

Las condiciones para la simulación en este planteamiento se fijaron como lo muestra la Tabla II. Se utilizaron constantes  $c = 1.25$  y  $d_1 = 1$ . La dinámica del líder virtual  $\dot{\mathbf{s}} = \mathbf{0}$  y con condición inicial  $\mathbf{s}(0) = [0 \ 0 \ 0]^T \text{m}$ .

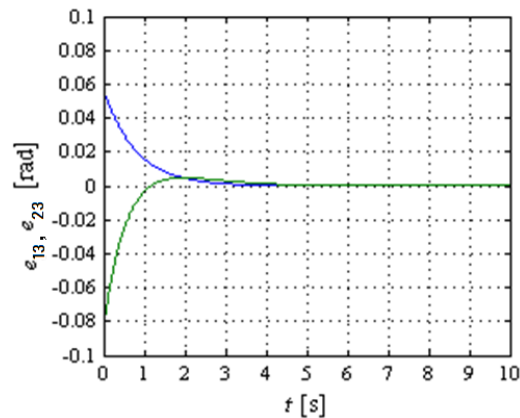
Tabla II. Condiciones de simulación para la red heterogénea con topología mostrada en la Figura 2 y correspondencia mostrada en la Tabla I.

$R_i$	$q_{ix}(0)$ [m]	$q_{iy}(0)$ [m]	$q_{i\theta}(0)$ [rad]	$\varphi_{ix}$ [m]	$\varphi_{iy}$ [m]	$\varphi_{i\theta}$ [rad]
$R_1$	0.1955	0.1706	0.0568	0.0	0.0	0.0
$R_2$	-0.0390	-0.0020	-0.0783	0.3	0.3	0.0
$R_3$	-0.1620	0.3492	-3.0882	-0.3	0.3	---
$R_4$	-0.1135	-0.3512	3.0847	-0.6	-0.6	---
$R_5$	0.0261	-0.1846	-0.0542	0.6	-0.6	---

La Figura 3 muestra la evolución de los errores respecto al tiempo. La Figura 3(a) muestra diez trazos correspondientes a errores de posición. La Figura 3(b) muestra sólo dos trazos correspondientes a los errores de orientación de los robots omnidireccionales, ya que para los robots tipo unicyclo no se consideró esta variable como variable de salida. Se aprecia que para  $t \approx 6$  s esta respuesta cumple con el objetivo de control en forma práctica aceptable.



(a)



(b)

Figura 3. Evolución de los errores en simulación: (a) errores de posición, (b) errores de orientación para  $R_1$  y  $R_2$ .

La Figura 4 muestra la evolución de las trayectorias para la formación.

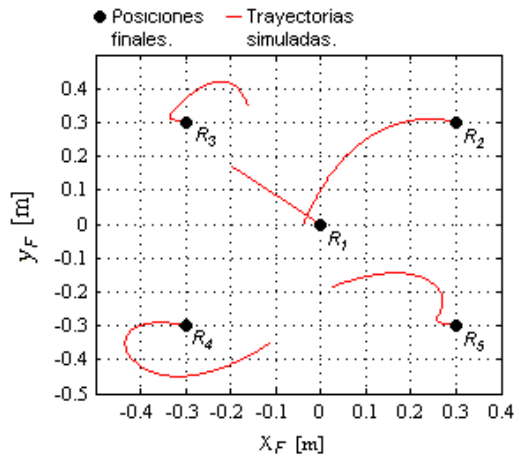


Figura 4. Evolución en simulación de las trayectorias para la formación de la red heterogénea.

#### D. Simulación con movimiento circular

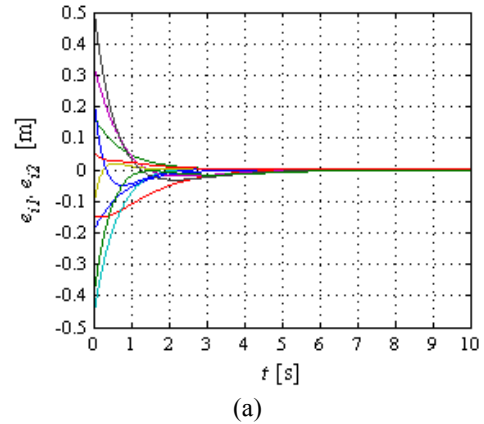
A continuación se presenta una simulación para la formación cuyo movimiento describe una circunferencia. Se utilizan las mismas condiciones descritas en la Tabla II. La constante global de acoplamiento es  $c = 1.25$  y la ganancia  $d_1 = 1$ . La dinámica del líder virtual fue  $\dot{s} = [0.15\sin(s_1) \ 0.15\cos(s_2) \ 0]^T$  y una condición inicial  $s(0) = [0 \ 0 \ 0]^T$  m.

La Figura 5 muestra la evolución de los errores mediante la simulación. La Figura 5(a) muestra los trazos correspondientes a los errores de posición. La Figura 5(b) muestra los trazos correspondientes a los errores de orientación de los robots omnidireccionales. Se aprecia que para  $t \approx 6$  s esta respuesta cumple con el objetivo de control en forma práctica aceptable y los robots siguen la trayectoria del líder virtual (ahora coincidente con el robot líder  $R_1$ ).

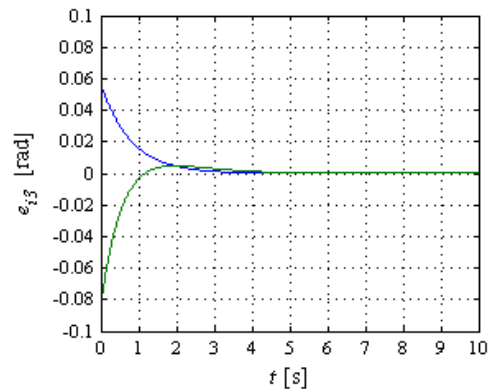
La Figura 6 muestra la evolución de las trayectorias para la formación y desplazamiento circular. La Figura 6 muestra 30 s de simulación, realizando 4.8 revoluciones en seguimiento del líder.

## V. CONCLUSIONES

Se ha presentado un procedimiento para formular una red heterogénea de robots. El procedimiento tiene su fundamento en la teoría de grafos y se propone la estrategia de acoplamiento difusivo estático. El acoplamiento propuesto se ejemplifica para una red de cinco robots que contiene dos tipos distintos de robots móviles terrestres a base de ruedas. Se ha desarrollado también la descripción para un objetivo de formación y el desplazamiento en formación para la red de robots. El cumplimiento del objetivo del control se demuestra mediante la teoría de estabilidad en el sentido de Lyapunov y se efectúan simulaciones acordes a la metodología planteada. Se describe el procedimiento para la formulación de red heterogénea y se resalta la ventaja de que la formulación es independiente de la misión particular de la red de robots e independiente del objetivo de control.



(a)



(b)

Figura 5. Evolución de los errores en simulación con un líder virtual en desplazamiento circular: (a) errores de posición, (b) errores de orientación de los robots  $R_1$  y  $R_2$ .

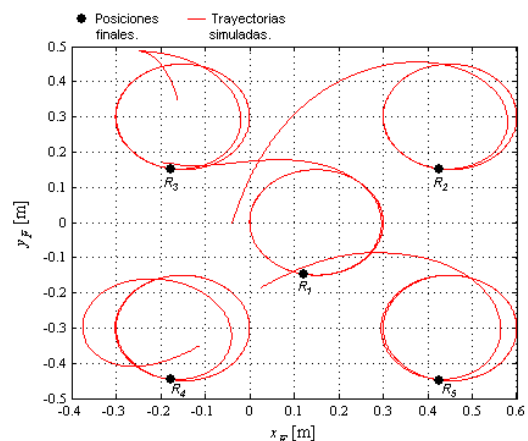


Figura 6. Evolución en 30 s de simulación de las trayectorias para la formación de la red heterogénea con líder con movimiento circular de radio 0.15 m.

Es importante mencionar algunas restricciones y que constituyen retos que se dejan aún vigentes en este trabajo: la estructura del modelo de los robots es de integrador simple y equiparable para ambos robots; la topología de la red requiere un líder en la raíz; la dificultad para el cómputo de la ganancia mínima crece exponencialmente en función del número de

robots que pertenecen a la red y; el procedimiento se aplica para una red con topología estática.

#### AGRADECIMIENTOS

Se agradece el apoyo financiero del Tecnológico Nacional de México (Proyecto 6719.18-P) y del Instituto Tecnológico de Ensenada.

#### REFERENCIAS

[1] Antonelli, G., F. Arrichiello, F. Caccavale, A. Marino, Decentralized Centroid and Formation Control for Multi-Robot Systems, IEEE International Conference on Robotics and Automation (ICRA), (2013) 3511-3516.

[2] Aranda, M., G. López-Nicolás, C. Sagüés, M. M. Zavlanos, Distributed Formation Stabilization Using Relative Position Measurements in Local Coordinates, IEEE Transactions on Automatic Control. DOI 10.1109/TAC.2016.2527719.

[3] Bezzo N., J. P. Hecker, K. Stolleis, M. E. Moses, R. Fierro, Exploiting Heterogeneous Robotic System in Cooperative Missions, arXiv:1509.00948v1 [cs.RO]. (Sep 2015).

[4] Bullo, F., J. Cortés, S. Martínez, Distributed Control of Robotic Networks: a Mathematical Approach to Motion Coordination Algorithms, Princeton University Press. (2009).

[5] Chen, G., Pinning Control and Synchronization on Complex Dynamical Networks, International Journal of Control, Automation, and Systems, 12(2) (2014) 221-230.

[6] Cruz, P., R. Fierro, W. Lu, S. Ferrari, Maintaining Robust Connectivity in Heterogeneous Robotic Networks, Proc. Of SPIE 8741, 87410N. (2013). Doi: 10.1117/12.2016236.

[7] Durazo, S., Yaveni, A., Bugarín, E., Control Law for a Robot Network Formation Using Static Diffusive Coupling. Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería (2017). DOI: 10.23967/j.rimni.2017.7.002.

[8] Fuady, S., A. Ravi, R. Bambang, Distributed Formation Control of Multi-Robot System with Obstacle Avoidance, International Conference on Robotics, Biomimetics, Intelligent Computational Systems (ROBIONETICS) (2013) 94-98.

[9] García, H., B. Jayawardhana, M. Cao, Distributed Rotational and Translational Maneuvering of Rigid Formations and Their Applications, IEEE Transactions on Robotics, 32(3) (2016) 684-697.

[10] Hu, C., J. Yu, H. Jiang, Z. Teng, Pinning Synchronization of Weighted Complex Networks with Variable Delays and Adaptive Coupling Weights, Springer Nonlinear Dyn, 67 (2012) 1373-1385.

[11] Kelly, R., V. Santibañez, Control de Movimiento de Robots Manipuladores, Ed. Pearson Education. Madrid, España. (2003).

[12] Seyboth, G. S., D. V. Dimarogonas, K. H. Johansson, F. Allgöwer, Static diffusive couplings in heterogeneous linear networks, IFAC Proceedings Volumes, 45(26) (2012) 258-263.

[13] Song, Q., F. Liu, J. Cao, W. Yu, M-Matrix Strategies for Pinning-Controlled Leader-Following Consensus in Multiagent Systems With Nonlinear Dynamics, IEEE Transactions on Cybernetics, 43(6). (Dec 2013). 1688-1697.

[14] Sun, Z., Anderson, B., Formation Feasibility on Coordination of Networked Heterogeneous System with Drift Terms, IEEE 55<sup>th</sup> Conference on Decision and Control (CDC) (Dec 2016), 3462-3467.

[15] Tsai, C-C., Chen, Y-S., Tai, F-C., Intelligent Adaptive Distributed Consensus Formation Control for Uncertain Networked Heterogeneous Swedish-Wheeled Omnidirectional Multi-Robots, SICE Annual Conference (Sep 2016), 154-159.

[16] Wieland, P., From Static to Dynamic Couplings in Consensus and Synchronization Among Identical and Non-Identical Systems, Ph. D. Thesis, Institute for Systems Theory and Automatic Control, University of Stuttgart (2010).

[17] Wu, H-L., Tsai, C-C., Tai, F-C., Adaptive Nonsingular Terminal Sliding-Mode Formation Control Using ORFWNN for Uncertain Networked Heterogeneous Mecanum-Wheeled Omnidirectional Robots, IEEE

International Conference on Systems, Man, and Cybernetics (SMC), (Oct 2017), 3317-3322.

[18] Yu, W., G. Chen, J. Lü, J. Kurts, Synchronization Via Pinning Control on General Complex Networks, SIAM J. Control Optimization. 51(2) (2013) 1395-1416.

[19] Zhong, W-S., G-P. Liu, C. Thoma, Global Bounded Consensus of Multiagent Systems With Nonidentical Nodes and Time Delays, IEEE Transactions on Systems, MAN, and Cybernetics—Part B: Cybernetics, 42(5). (Oct 2012). 1480-1488.

# Adaptación del vehículo submarino BlueROV2 para banco de pruebas

Abigayl Chargoy\*, Jesús Sandoval\*, César Higuera†, Eusebio Bugarin†.

\*Tecnológico Nacional de México, I.T. de La Paz, abiggailc@gmail.com, jsandoval@itlp.edu.mx.

†Tecnológico Nacional de México, I.T. de Ensenada, cesarh@itlp.edu.mx, ebugarin@ite.edu.mx.

**Resumen**—Se presenta una nueva configuración para el vehículo comercial BlueROV2 (2017) en la cual se realizaron adaptaciones en el software y hardware, a partir de la reconexión de los componentes electrónicos propios del vehículo, con lo cual es posible usarlo como plataforma experimental para la validación de algoritmos de control. Se implementó un algoritmo de visión monocular para conocer la orientación del vehículo incluyendo la cámara como sensor de posición, además de la lectura de los sensores y manipulación de los actuadores, donde el procesamiento de datos se realiza de manera interna en el submarino. Se realizaron experimentos con el vehículo a través de comunicación remota para corroborar la correcta lectura de los sensores y operación de los propulsores bajo el agua con la configuración propuesta.

**Palabras clave**—Vehículo submarino, BlueROV2, Pixhawk, visión monocular.

## I. INTRODUCCIÓN

El ambiente marino es uno de los entornos que ofrece numerosas actividades para el uso de robots llamados Vehículos Autónomos Subacuáticos (AUV) y los Vehículos Operados Remotamente (ROV), tales como: monitoreo de áreas protegidas, recopilación de datos, exploración submarina y conteo de especies acuáticas entre otras.

Una de las técnicas para conocer el desplazamiento y localización de un vehículo es el uso de cámaras como principal sensor, las cuales permiten recolectar datos a partir de imágenes capturadas y procesarlas para tomar decisiones sobre el control del vehículo, dado que el sistema de posicionamiento global (GPS) sólo puede ser utilizado en la superficie.

La configuración de fábrica del submarino BlueROV2 de la compañía *BlueRobotics*©, mostrado en la Figura 1, requiere de un programa externo, como el *QGroundControl*, el cual está limitado para procesamiento de imágenes e implementación de algoritmos de control.

Varios autores se han enfocado en estudiar nuevas aplicaciones y técnicas de control para el vehículo BlueROV2. Por ejemplo, en [9] y [10] los autores realizan modificaciones al submarino agregando cámaras, para la utilización de técnicas de SLAM, redes neuronales y con librerías de apoyo como TensorFlow y ROS, logran resolver retos de odometría, reconocimiento y detección de objetos mediante visión estereoscópica. En [4], [5], [6], [7] y [8] y [12] los autores estudian las técnicas para controlar vehículos submarinos, en estos casos es necesario conocer la configuración del submarino, el modelo matemático y la comunicación para hacer uso de software para su simulación e implementación.

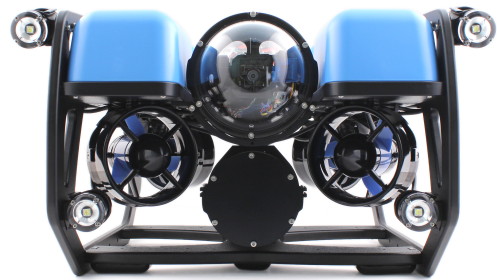


Figura 1. Vehículo submarino BlueROV2 [2].

El objetivo de establecer una nueva configuración en el vehículo BlueROV2 es poder implementar algoritmos de control incluyendo algoritmos de visión artificial monocular, sin necesidad de agregar más componentes electrónicos. Para utilizar el vehículo submarino como una plataforma de validación de diferentes algoritmos para el control de un vehículo submarino, es necesario realizar modificaciones, tanto en el hardware como en el software, que permitan manipular el vehículo sin el uso de programas y librerías predeterminadas para su operación remota. En este trabajo se presenta una nueva configuración con la implementación de un algoritmo de control desarrollado en C++ que permite el control de los actuadores, la lectura de los sensores y el procesamiento de imágenes.

## II. CONFIGURACIÓN DEL BLUEROV 2

La configuración propuesta por la compañía *BlueRobotics*© para operar el vehículo a través del cable de comunicación que permite la navegación, la transmisión de imágenes y la recolección de datos son los siguientes [2]:

- *Raspberry Pi 3*: Computadora de bajo consumo que comunica la cámara y la unidad inercial a la computadora del operador. Utiliza el sistema precompilado ejecutando el software *Companion* y Firmware ArduSub [1].
- *Unidad inercial Pixhawk v1*: Componente empleado para control de vehículos móviles [11]. Está dotado de sensores y algoritmos internos empleados para el localizar y desplazar el submarino.
- *Cámara*: Módulo de cámara marca Raspberry Pi v2. Se emplea para desplegar las imágenes en la computadora del operador.

- *Fathom-X Tether*: Permite la comunicación a través del BlueROV2 y la computadora externa [2].
- *Controlador de velocidad electrónico (ESC)*: Permite activar los actuadores a partir de la señal recibida del Pixhawk.
- *Actuadores*: Los propulsores empleados son el modelo T200 de BlueRobotics©. Cada actuador se controla mediante un ESC. En la Figura 2 se muestra la posición de los propulsores y su orientación. Las propelas de color verde tienen diseño de giro en sentido horario mientras que en las azules se tiene sentido antihorario. Los esquemas de conexión y de comunicación entre el operador y el vehículo submarino se presentan en las Figuras 3 y 4.

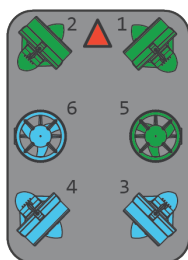


Figura 2. Configuración vectorizada de los propulsores. Los propulsores 1-4 son utilizados para desplazamiento hacia adelante/atrás y movimiento rotacional. Los propulsores 5-6 desplazan el vehículo arriba/abajo. El triángulo color rojo indica el frente del vehículo.

Para la operación del submarino se emplea el programa *QGroundControl*, el cual requiere de un control (joystick) *Xbox* de la compañía *Windows*© que es configurado para activar el servomotor de la cámara, actuadores, luces y modo de operación del vehículo. El programa *QGroundControl* permite visualizar las imágenes capturadas por la cámara y los datos obtenidos de los sensores del vehículo. La comunicación remota al submarino mediante los dispositivos *Fathom X* que están conectados entre sí a partir de un cable umbilical al computador externo con la computadora *Raspberry Pi* del vehículo. La *Raspberry Pi* es la responsable de transmitir los datos de la cámara, de los sensores del *Pixhawk* y actuadores del *BlueROV2*. La unidad inercial *Pixhawk* controla los propulsores y las lámparas del submarino.

La configuración mostrada en la Figura 4 tiene la limitación de implementar algoritmos de control y de visión artificial, la cual requiere de instalación de librerías y extensiones adicionales. Con el fin de remover esta limitación se plantea una nueva configuración en el vehículo submarino *BlueROV2*.

### III. NUEVA CONFIGURACIÓN DEL BLUEROV2

Para utilizar el vehículo submarino como banco de pruebas de algoritmos de control e incluyendo visión artificial se realizó la reconexión de componentes electrónicos. Para lograr los objetivos propuestos se realizaron cambios en las conexiones mostradas en la Figura 4 resultando la reconexión y la comunicación mostradas en la Figura 5 y 6, descrita a continuación.

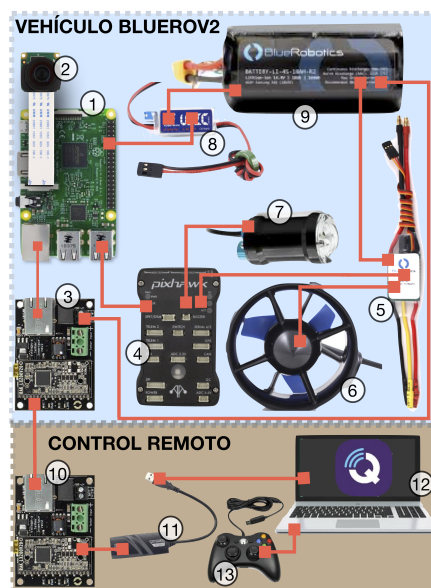


Figura 3. Esquema de conexión de los componentes para la operación del submarino. Los componentes numerados son: (1) *Raspberry Pi 3*. (2) *Raspi-Camera*. (3) y (10) *Fathom X*. (4) *Pixhawk v1*. (5) *Control de velocidades (ESC ×6)*. (6) *Propulsores (× 6)*. (7) *Lámparas (×4)*. (8) *Regulador 5V*. (9) *Batería LiPo 14.8V 4Cell*. (11) *Conector Ethernet-USB*. (12) *Computadora portátil con el programa QGround Control*. (13) *Control alámbrico de Xbox*.

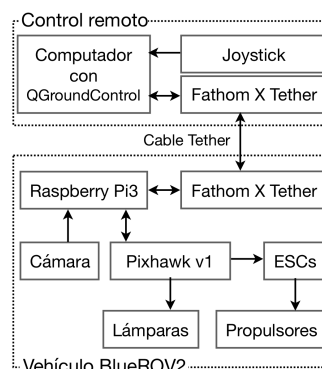


Figura 4. Esquema de comunicación del vehículo submarino *BlueROV2*.

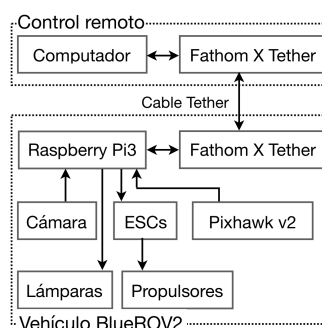


Figura 5. Propuesta de comunicación del vehículo submarino *BlueROV2*.

#### III-A. *Pixhawk 2*

Se actualizó el *Pixhawk 1* por la versión *Pixhawk 2*. La función de la unidad inercial es únicamente leer los sensores, dejando de lado el control del sistema de propulsión.



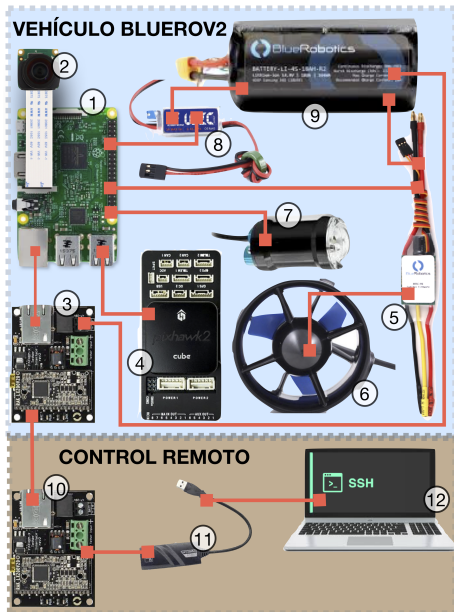


Figura 6. Propuesta de conexión de los componentes para la operación del submarino. Los componentes numerados son: (1) Raspberry Pi 3. (2) Rspi-Camera. (3) y (10) Fathom X. (4) Pixhawk v2. (5) Control de velocidades (ESC  $\times$  6). (6) Propulsores ( $\times$  6). (7) Lámparas ( $\times$  4). (8) Regulador 5V. (9) Batería LiPo 14.8V 4Cell. (11) Conector Ethernet-USB. (12) Computadora portátil.

### III-B. Sistema de propulsión

Se realizó una nueva configuración el sentido de giro de los propulsores para usarlos por pares, como se muestra en la Figura 7. Los ESC son conectados directamente en la Raspberry Pi para su control. La posición de las propelas no fue modificada manteniendo la configuración vectorizada propuesta por BlueRobotics©.

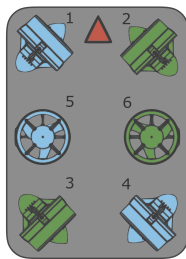


Figura 7. Modificación en la configuración vectorizada de los propulsores. Los propulsores de color verde cuentan con propelas con sentido horario y los azules con sentido antihorario.

### III-C. Raspberry Pi

La Raspberry Pi dispone de entradas y salidas de propósito general (GPIO) que pueden ser programados como PWM y se pueden usar para controlar los motores del vehículo. Los pulsos para activar los propulsores son generados por la Raspberry Pi y conectados a los ESC. La asignación de pines del GPIO a los propulsores y las lámparas se muestra en la

Tabla I. En la Figura 7 se describen las posiciones de los seis propulsores.

Propulsor	1	2	3	4	5	6	Lámparas
Pin GPIO	5	6	13	19	20	21	26

Tabla I

ASIGNACIÓN DE PROPULSORES DEL SUBMARINO Y PINES DEL GPIO DE LA RASPBERRY PI.

## IV. CONFIGURACIÓN DEL SOFTWARE

La computadora del vehículo tiene instalado el sistema operativo Rasbian GNU/Linux versión Debian 8 (Jessie), versión del kernel 4.14.81-v7+. Para el algoritmo de lectura de sensores y actuación de los propulsores se requiere la instalación de librerías OpenCV 3.0, PiGPIO y MAVLink.

Se desarrolló un programa para su ejecución en la Raspberry Pi al cual se puede ingresar de forma remota mediante protocolo Secure SHell (SSH). Este programa obtiene los datos de la cámara y del Pixhawk, y es usado para controlar tanto los actuadores como ejecutar los algoritmos de control, tanto en malla abierta como en malla cerrada, usando como retroalimentación los datos obtenidos de la lectura de los sensores y el procesamiento de imágenes.

### IV-A. Comunicación

La comunicación entre el submarino y el usuario utiliza el protocolo SSH a través del cable umbilical del vehículo. Este enlace permite la implementación de algoritmos de control, la visualización, la recopilación y la obtención de datos de los sensores para su análisis.

### IV-B. Visión monocular

La integración de la cámara como sensor permite obtener la orientación y distancia tridimensional de una imagen patrón. El algoritmo de visión monocular propuesto en [3], se presenta como pseudocódigo en el Algoritmo 1, permite calcular la orientación y distancia de separación de la imagen patrón con respecto a la cámara del submarino, a partir de la obtención de las matrices de rotación y translación calculadas mediante los píxeles que conforman a las imágenes capturadas.

En la Figura 8 se muestran resultados de la ejecución del algoritmo de visión por computadora para el cálculo de la postura de la imagen patrón (tablero de ajedrez). En la imagen capturada se observan los ejes dibujados sobre el tablero de ajedrez detectado con la orientación calculada.

### IV-C. Lectura de sensores

El algoritmo para la lectura de los sensores del Pixhawk mediante la librería MAVLink se describe en el Algoritmo 2.

### IV-D. Control del sistema de propulsión

Para el control del sistema de propulsión se ejecuta el algoritmo, propuesto en [3], que realiza la activación de los actuadores así como el aumento y la disminución de la potencia, que pueden ser utilizados por una ley de control implementada por el usuario. El pseudocódigo se muestra en el Algoritmo 3.





Figura 8. Resultados del algoritmo de visión para detección de postura de un patrón de tablero de ajedrez colocado a una separación de 30cm a la cámara del submarino.

#### Algoritmo 1 Distancia y orientación de un patrón reconocido

```

1: procedure          OBTENCIÓN          DE          POSTU-
   RA(MatrixIntrinseca, CoficienteDistorsion)
2:    $M \leftarrow MatrixIntrinseca$ ;
3:    $D \leftarrow CoficienteDistorsion$ ;
4:   repeat
5:     Read imageCamera
6:     if (FindChessboard(imageCamera)=true) then
7:       imageGray           $\leftarrow$  ConvertRGBtoGRAY
       imagenCamera
8:       PointsCorners  $\leftarrow$  CornerSubPix(imageGray)
9:       imagePointsChessboard           $\leftarrow$ 
       PointsCorners
10:      (rvec, tvec)           $\leftarrow$ 
       SolvePnP(imagePointsChessboard,  $M$ ,  $D$ )
11:       $rMat \leftarrow$  Rodrigues(rvec)
12:      angulosEuler[roll, pitch, yaw]           $\leftarrow$ 
       rotationMatrixToEulerAngles( $rMat$ );
13:       $DistPixels \leftarrow PointsCorners[inicio_x] -$ 
        $PointsCorners[fin_x]$ 
14:       $DistPixelsNoRotation \leftarrow (0.210 * \cos(roll) * \sin(yaw))$ ;
15:       $DistMetrosNoRotation \leftarrow$ 
       ( $DistPixelsNoRotation * (578 / DistPixels)$ )
16:      write angulosEuler
17:      write DistMetrosNoRotation
18:     end if
19:   until exit
20: end procedure
    
```

## V. EXPERIMENTOS REALIZADOS

En los experimentos realizados, el algoritmo fue implementado y ejecutado directamente en la computadora Raspberry Pi mediante la protocolo SSH, los cuales se realizaron en la fosa de clavados del gimnasio de usos múltiples del Gobierno

#### Algoritmo 2 Lectura de la unidad inercial

```

1: procedure C-UART-INTERFACE(puerto)
2:    $puertoPX4 \leftarrow puerto$ 
3:   while  $puerto = puertoPX4$  do
4:      $orientacion(x, y, z) \leftarrow giroscopio(x, y, z)$ ;
5:      $velocidad(x, y, z) \leftarrow acelerometro(x, y, z)$ ;
6:     write(orientacion( $x, y, z$ ));
7:   end while
8: end procedure
    
```

#### Algoritmo 3 Control de propulsores

```

1: procedure CONTROL(Eje, Ganancia, Tiempo)
2:   repeat
3:      $potencia \leftarrow (stop \pm Ganancia * 4)$ 
4:     if  $Eje == "x"$  then
5:       Activar propulsores para desplazamiento en  $x$ 
       con potencia durante Tiempo segundos.
6:     else if  $Eje == "y"$  then
7:       Activar propulsores para desplazamiento en  $y$ 
       con potencia durante Tiempo segundos.
8:     else if  $Eje == "z"$  then
9:       Activar propulsores para desplazamiento en  $z$ 
       con potencia durante Tiempo segundos.
10:    else if  $Eje == "roll"$  then
11:      Activar propulsores para rotación en roll con
       potencia durante Tiempo segundos.
12:    else if  $Eje == "yaw"$  then
13:      Activar propulsores para rotación en yaw con
       potencia durante Tiempo segundos.
14:    else if  $Eje == "l"$  then
15:      Activar lamparas con potencia.
16:    end if
17:  until  $Eje == "q"$ 
18: end procedure
    
```

del Estado de B.C.S., la cual tiene una profundidad máxima de 5m. Se usó una computadora portátil conectada con un cable de red tether a través de acceso remoto SSH al vehículo, que sirve como extensión del teclado y monitor para realizar las configuraciones, la implementación y ejecución de los algoritmos.

#### V-A. Cálculo de distancia mediante visión monocular

En los experimentos de visión artificial se sumergió el submarino a una profundidad de 20cm, colocando la imagen patrón a la misma profundidad y a una distancia 60cm, ver Figura 9. La captura de las imágenes se realizó cada segundo para evitar la saturación de la memoria de la Raspberry Pi. Se realizaron pruebas con el algoritmo de visión artificial en el laboratorio del Instituto Tecnológico de La Paz (Figura 8) observando que el error varía desde  $\pm 0.02m$  a  $\pm 0.11m$  de acuerdo las distancias reales de separación entre la cámara y la imagen patrón.

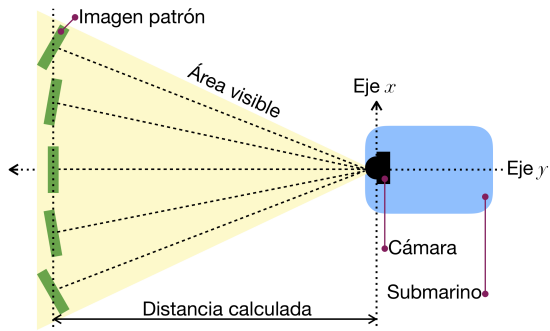


Figura 9. Esquema descriptivo de la posición de la cámara del BlueROV2 y desplazamiento de la imagen patrón para la realización de experimentos.

La variación en el error en la medición de la distancia entre la cámara y la imagen patrón es resultado por la distorsión generada por la cámara, donde la distorsión es significativa en los límites de las imágenes, es decir, aumenta al no estar centrada la imagen patrón en el espacio de visualización de la cámara.

La Figura 10 muestra una captura de la cámara del submarino durante las pruebas en la fosa de clavados, la cual es procesada para detectar y calcular la distancia de separación entre la cámara y la imagen patrón. Los ejes son dibujados con la orientación calculada a partir del algoritmo implementado. Con el tablero colocado a una distancia de 0.6m, los resultados de la distancia calculada obtenidos por el programa mostraban un rango de error de  $\pm 0.018m$  a  $\pm 0.12m$ , cabe destacar que los valores obtenidos bajo la superficie del agua y fuera de ella son similares.

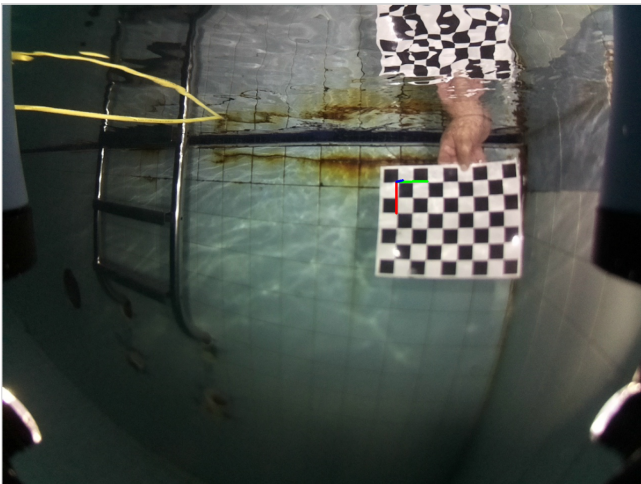
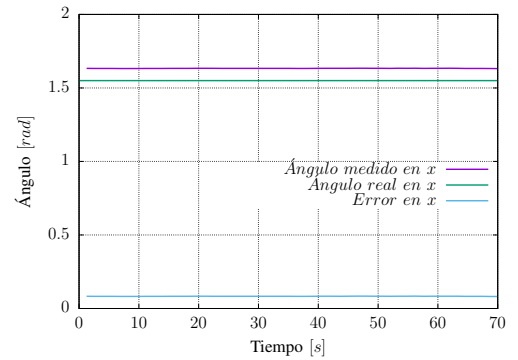


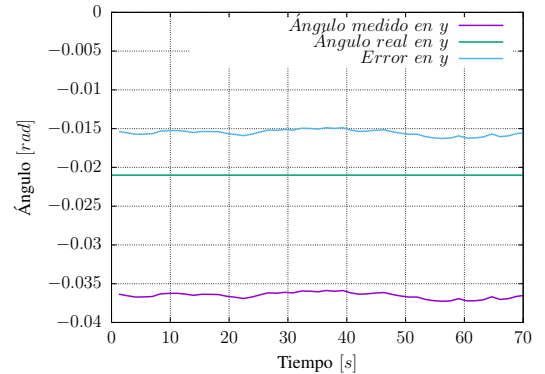
Figura 10. Detección de la distancia y orientación del tablero de ajedrez bajo el agua respecto al vehículo submarino.

### V-B. Lectura de sensores

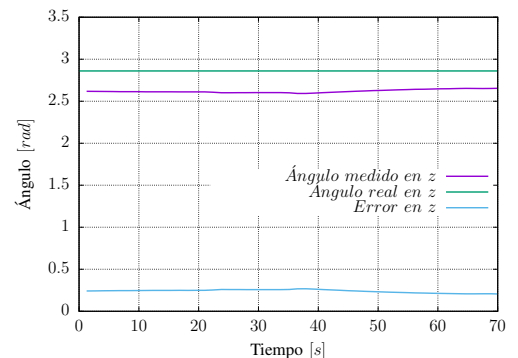
Se realizó la evaluación de la lectura de los sensores del componente Pixhawk 2 manteniendo el vehículo estático. En la Figura 11 se muestra los valores de orientación y velocidad con un equipo inercial externo al submarino y los valores medidos por la unidad inercial con respecto a los ejes  $x$ ,  $y$  y  $z$  del vehículo.



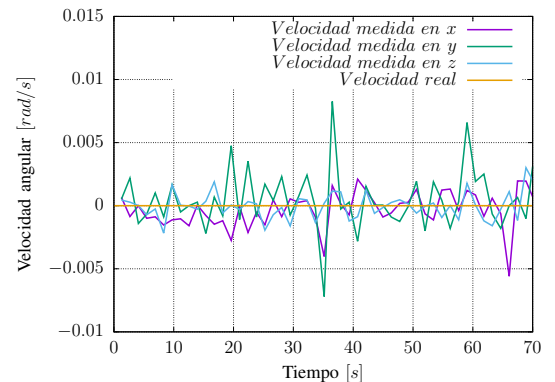
(a) Lectura del ángulo en  $x$ .



(b) Lectura del ángulo en  $y$ .



(c) Lectura del ángulo en  $z$ .



(d) Lectura de las velocidad en los ejes  $x$ ,  $y$  y  $z$ .

Figura 11. Orientación y velocidades del vehículo en estado inmóvil obtenidos del Pixhawk.

### V-C. Activación de propulsores.

Con la ejecución del algoritmo para activar los propulsores, los datos del desplazamiento y velocidad del vehículo se almacenaron para su evaluación. Se realizaron las pruebas para cada eje del submarino, se muestra el desplazamiento y velocidad en  $z$  en la Figura 12. Tomando como origen (0,0,0) a la posición en la que se enciende el submarino, en la Figura 12 se observa el desplazamiento en el eje  $z$  durante la activación de los propulsores 5 y 6 (Figura 7).

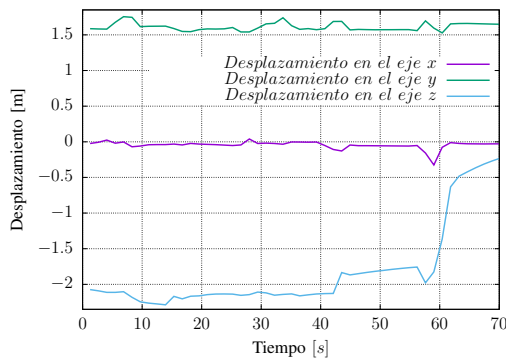


Figura 12. Datos del desplazamiento del vehículo a partir de la activación de actuadores para desplazamiento en el eje  $z$  durante 5 segundos al 50% de potencia.

## VI. CONCLUSIONES

Se ha presentado una nueva configuración de hardware y software del vehículo BlueROV2 para usarlo como plataforma de validación de algoritmos de control. Se incluyó la cámara como sensor utilizando técnicas de visión monocular para reconocer una imagen patrón y a partir de ella establecer la distancia y la orientación con respecto al vehículo submarino.

Con este nuevo uso de la cámara y el algoritmo de control del sistema de propulsión, es posible dotar de autonomía al vehículo, esto es, implementar leyes de control basadas en el modelo que permitan al vehículo cumplir con objetivos de control sin la intervención del usuario.

### AGRADECIMIENTOS

Este trabajo fue financiado parcialmente por el TecNM, contrato 5925.19-P.

### REFERENCIAS

- [1] ARDUSUB, ArduSub GitBook, 2017. Consultado el 16-Abril-2019. URL: [www.ardusub.com](http://www.ardusub.com).
- [2] BLUEROBOTICS. Consultado el 16-Abril-2019. URL: [docs.bluerobotics.com](http://docs.bluerobotics.com).
- [3] CHARGOY, A. (2018). *Control de un vehículo submarino no tripulado: análisis y experimentos*. Tesis de Maestría. Tecnológico Nacional de México, Instituto Tecnológico de La Paz, B.C.S, México.
- [4] CHU-JOU, B. (2018). *6-DoF Modelling and Control of a Remotely Operated Vehicle*. Master's thesis. Flinders University, Adelaide, South Australia.
- [5] GARCÍA-MOLLEDA, D. (2016). *Control de un vehículo subacuático*. Tesis de Maestría. Tecnológico Nacional de México, Instituto Tecnológico de La Paz, B.C.S., México.

- [6] HERNÁNDEZ-ALVARADO, R., GARCÍA-VALDOVINOS, L., SALGADO-JIMÉNEZ, T., GÓMEZ-ESPINOSA, A., FONSECA-NAVARRO, F. (2016). *Neural network-based self-tuning PID control for underwater vehicles*. *Sensors* (Basel, Switzerland), 16(9), 1429.
- [7] HIGUERA, C. (2016). *Implementación del control de un vehículo subacuático*. Tesis de Maestría. Tecnológico Nacional de México, Instituto Tecnológico de La Paz, B.C.S., México.
- [8] HIGUERA C., SANDOVAL J., CORIA L., BUGARÍN E., SANTILLÁN I. (2018). *Configuración experimental de un vehículo submarino no tripulado*. COMRob 2018.
- [9] LOMBARDI, J., GODSEY, W., KELLY, B., UVAROV, N., YUZVIK, A. (2018). *Autonomous Lionfish Harvester*. Bachelor's thesis. Worcester Polytechnic Institute, Worcester, Massachusetts
- [10] NAWAF, M., MERAD, D., ROYER, J. P., BOÏ, J. M., SACONE, M., BEN ELLEFI, M., DRAP, P. (2018). *Fast visual odometry for a low-cost underwater embedded stereo system*. *Sensors* (Basel, Switzerland), 18(7), 2313.
- [11] PIXHAWK. Consultado el 16-Abril-2019. URL: [pixhawk.org](http://pixhawk.org).
- [12] STORESUND, E. (2018). *Surface assisted autopilot for remotely operated vehicle*. Master's thesis, Norwegian University of Science and Technology, Islamabad, Pakistan.

# Identificación Paramétrica de un Quadrotor con Validación Experimental

Ivan Lopez-Sanchez<sup>†</sup>, Javier Moreno-Valenzuela<sup>†</sup>, Jorge Montoya-Cháirez<sup>†</sup> y Ricardo Pérez-Alcocer<sup>§</sup>

<sup>†</sup>Instituto Politécnico Nacional - CITEDI

Av. Instituto Politécnico Nacional No. 1310, Tijuana, B.C. 22435, México

Email: ilopez@citedi.mx, moreno@citedi.mx, jamontoya@citedi.mx

<sup>§</sup>CONACYT - Instituto Politécnico Nacional - CITEDI

Av. Instituto Politécnico Nacional No. 1310, Tijuana, B.C. 22435, México

Email: rrperez@citedi.mx

**Resumen**—En este trabajo se presenta la aplicación del método de identificación de mínimos cuadrados mediante el cual se estiman la masa y los parámetros de inercia correspondientes al quadrotor *QBall 2* de *Quanser*. Primeramente, se implementa un controlador tipo PID para el seguimiento de trayectorias de posición y orientación. Con estos resultados se implementa fuera de línea un algoritmo de identificación paramétrica. Los resultados son validados realizando simulaciones numéricas del controlador PID en lazo cerrado con el quadrotor usando los parámetros del fabricante y los parámetros obtenidos mediante el proceso de identificación presentado. Las simulaciones son comparadas con respecto al experimento. Los mejores resultados son obtenidos con los parámetros estimados mediante el proceso de identificación descrito.

## I. INTRODUCCIÓN

Entre los diversos vehículos aéreos no tripulados disponibles en la actualidad, los quadrotores han sido de los más abundantes en la última década [1]. Un quadrotor es un tipo de helicóptero propulsado por cuatro rotores, los cuales forman los vértices de un cuadrado que a su vez equidistan del centro de masa del quadrotor. Puesto que el quadrotor posee seis grados de libertad y sólo cuatro actuadores, es considerado un sistema subactuado [2]. Las ventajas más importantes de este vehículo son principalmente el tamaño, la maniobrabilidad y agilidad en vuelo. Posee la capacidad de mantenerse estático en un punto en el espacio, despegar y aterrizar verticalmente, lo cual ahorra espacio. Además, es económico y de fácil mantenimiento [3].

Como resultado de la versatilidad de esta clase de vehículos aéreos, su actual auge y sus diversos campos de aplicación, se han implementado distintas técnicas de control como esquemas PD/PID que compensan el modelo dinámico de la planta, como se ha estudiado en [4], [5], [6] y [7]. La implementación de controladores y la sintonización de sus ganancias puede resultar complicada. Por tal motivo una forma rápida de validar esta clase de esquemas de control es mediante simulaciones numéricas, para las cuales el conocer los parámetros de la plataforma es muy importante ya que de estos depende en gran parte la dinámica del sistema. Por lo tanto, contar con los parámetros reales resultaría en un comportamiento más cercano a la realidad [1]. La estimación de estos parámetros

desconocidos es posible mediante métodos de identificación paramétrica, comúnmente basados en técnicas de regresión lineal como se ve en [1], [8] y [9].

En este trabajo se presenta la identificación paramétrica de un quadrotor utilizando el algoritmo de mínimos cuadrados. El propósito de este documento es validar la metodología implementada para esta identificación mediante la comparación en simulación de un controlador de seguimiento de trayectorias de posición y orientación usando los parámetros identificados y los nominales con respecto a datos experimentales. Se presenta un modelo dinámico reducido, el cual podría ser implementado en otras plataformas.

El documento se encuentra organizado como sigue. En la sección II se presenta el modelo dinámico del quadrotor y la parametrización del modelo. La metodología seguida para la identificación, el procesamiento de los datos experimentales, la descripción del controlador implementado y las características del experimento se presentan en la sección III. En la sección IV se presenta la descripción de la plataforma, el algoritmo de mínimos cuadrados y la discusión de los resultados. Finalmente las conclusiones y trabajo futuro se presentan en la sección V.

## II. MODELO DINÁMICO DE UN QUADROTOR

En esta sección se presenta el modelo dinámico del quadrotor como se ve en [4] y [5]. Un quadrotor es un sistema subactuado con seis grados de libertad y cuatro entradas de control, que son el empuje total  $F_z$  y los torques  $\tau_\phi$ ,  $\tau_\theta$  y  $\tau_\psi$  [4]. Bajo la suposición de que el quadrotor es un cuerpo rígido moviéndose en el espacio, es simétrico, tiene centro de masa que coincide con el centro geométrico del cuerpo y que el marco de referencia situado en el cuerpo del quadrotor también coincide con el centro de masa, este sistema puede ser representado como

$$m\dot{\mathbf{v}} + mS(\boldsymbol{\omega})\mathbf{v} + mgR(\boldsymbol{\eta})^T \mathbf{e}_z = \begin{bmatrix} 0 \\ 0 \\ F_z \end{bmatrix}, \quad (1)$$

$$I\dot{\boldsymbol{\omega}} + S(\boldsymbol{\omega})I\boldsymbol{\omega} = \boldsymbol{\tau}, \quad (2)$$

$$\dot{\boldsymbol{p}} = R(\boldsymbol{\eta})\mathbf{v}, \quad (3)$$

$$\dot{\boldsymbol{\eta}} = W(\boldsymbol{\eta})\boldsymbol{\omega}, \quad (4)$$

donde  $m$  es la masa del vehículo,  $g$  es la constante de aceleración gravitacional,  $\mathbf{v} = [u \ v \ w]^T$  y  $\boldsymbol{\omega} = [p \ q \ r]^T$  son las velocidades lineales y angulares en el marco de referencia del cuerpo del quadrotor, respectivamente. La matriz  $I = \text{diag}\{I_{xx} \ I_{yy} \ I_{zz}\} \in \mathbb{R}^{3 \times 3}$  es el tensor de inercia,  $\mathbf{p} = [x \ y \ z]^T$  es el vector de posiciones,  $\boldsymbol{\eta} = [\phi \ \theta \ \psi]^T$  es el vector de ángulos de Euler,  $\mathbf{e}_z = [0 \ 0 \ 1]^T$  es un vector unitario en dirección del eje  $z$  del marco de referencia inercial,  $R(\boldsymbol{\eta}) \in \mathbb{R}^{3 \times 3}$  es una matriz ortogonal de rotación según la convención aeronáutica "ZYX" definida como

$$R(\boldsymbol{\eta}) = \begin{bmatrix} c_\theta c_\psi & s_\theta c_\psi s_\phi - s_\psi c_\phi & s_\theta c_\psi c_\phi + s_\psi s_\phi \\ c_\theta s_\psi & s_\theta s_\psi s_\phi + c_\psi c_\phi & s_\theta s_\psi c_\phi - c_\psi s_\phi \\ -s_\theta & c_\theta s_\phi & c_\theta c_\phi \end{bmatrix}, \quad (5)$$

$W(\boldsymbol{\eta}) \in \mathbb{R}^{3 \times 3}$  es una matriz de transformación definida como

$$W(\boldsymbol{\eta}) = \begin{bmatrix} 1 & s_\phi t_\theta & c_\phi t_\theta \\ 0 & c_\phi & -s_\phi \\ 0 & s_\phi/c_\theta & c_\phi/c_\theta \end{bmatrix}, \quad (6)$$

$s_x$ ,  $c_x$  y  $t_x$  son el  $\sin(x)$ ,  $\cos(x)$  y  $\tan(x)$  respectivamente,  $F_z \in \mathbb{R}$  es una fuerza externa en dirección del eje  $z$  del marco de referencia del cuerpo,  $\boldsymbol{\tau} = [\tau_\phi \ \tau_\theta \ \tau_\psi]^T$  es el vector de torques en el marco de referencia del cuerpo del quadrotor,  $S(\boldsymbol{\omega})$  es la matriz antisimétrica definida como

$$S(\boldsymbol{\omega}) = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}. \quad (7)$$

La acción de control  $F_z$  es el resultado de la sumatoria de los empujes de los rotores del quadrotor. La acción de control  $\boldsymbol{\tau}$  correspondiente al vector de torques, resultado del empuje de los rotores y la distancia entre el eje de rotación de los mismos con respecto al centro de masa del quadrotor.

Figuras que representan el marco de referencia inercial y el marco de referencia en el cuerpo del quadrotor, así como su modelo dinámico pueden ser consultadas en [4] y [5].

## II-A. Parametrización del modelo dinámico.

Se puede observar en (1) y (2) que para describir el comportamiento de este sistema es necesario conocer parámetros como la inercia y la masa. También están presentes otros parámetros relacionados a efectos aerodinámicos, la dinámica de los rotores y efectos giroscópicos, los cuales no son considerados en estos modelos debido a que la estructura del quadrotor es pequeña, se desplaza a baja velocidad y posee propulsores pequeños en comparación a la estructura del mismo [4].

Considérese el modelo dinámico en el marco de referencia del cuerpo del quadrotor en (1) y (2) que puede ser expresado como [1]

$$\begin{aligned} \dot{u} + qw - rv + g_{bx} &= 0, \\ \dot{v} + ru - pw + g_{by} &= 0, \\ m(\dot{w} + pv - qu + g_{bz}) &= F_z, \\ I_{xx}\dot{p} + (I_{zz} - I_{yy})qr &= \tau_\phi, \\ I_{yy}\dot{q} + (I_{xx} - I_{zz})pr &= \tau_\theta, \\ I_{zz}\dot{r} + (I_{yy} - I_{xx})pq &= \tau_\psi, \end{aligned} \quad (8)$$

donde,  $g_{bx} = -g \sin(\theta)$ ,  $g_{by} = g \cos(\theta) \sin(\phi)$  y  $g_{bz} = g \cos(\theta) \cos(\phi)$ . El sistema (8) es lineal respecto a los parámetros y puede ser representado de forma matricial como

$$\Phi(\mathbf{v}, \boldsymbol{\omega}, \boldsymbol{\eta}, \dot{\mathbf{v}}, \dot{\boldsymbol{\omega}}) \boldsymbol{\Theta} = \mathbf{F}, \quad (9)$$

donde  $\Phi(\mathbf{v}, \boldsymbol{\omega}, \boldsymbol{\eta}, \dot{\mathbf{v}}, \dot{\boldsymbol{\omega}}) \in \mathbb{R}^{4 \times 4}$  es la matriz de regresión y está compuesta por funciones conocidas que representan la dinámica del quadrotor,  $\boldsymbol{\Theta} \in \mathbb{R}^4$  es el vector de parámetros constantes y  $\mathbf{F} \in \mathbb{R}^4$  es el vector de entradas de control. Los elementos que componen (9) se definen como

$$\Phi(\mathbf{v}, \boldsymbol{\omega}, \boldsymbol{\eta}, \dot{\mathbf{v}}, \dot{\boldsymbol{\omega}}) = \begin{bmatrix} \phi_{11} & 0 & 0 & 0 \\ 0 & \dot{p} & -qr & qr \\ 0 & pr & \dot{q} & -pr \\ 0 & -pq & pq & \dot{r} \end{bmatrix}, \quad (10)$$

$$\phi_{11} = \dot{w} + pv - qu + g_{bz},$$

$$\boldsymbol{\Theta} = \begin{bmatrix} m \\ I_{xx} \\ I_{yy} \\ I_{zz} \end{bmatrix}, \quad (11)$$

y

$$\mathbf{F} = \begin{bmatrix} F_z \\ \tau_\phi \\ \tau_\theta \\ \tau_\psi \end{bmatrix}. \quad (12)$$

## III. IDENTIFICACIÓN DE LOS PARÁMETROS

Diversos métodos se han presentado en la literatura para la identificación paramétrica [1], [8] y [9]. Estos métodos comparten las siguientes características:

- Se expresa el modelo dinámico del sistema tal que este sea lineal respecto a los parámetros.
- El método de identificación se aplica al modelo parametrizado para un número suficientemente grande de muestras de las entradas y salidas del sistema.
- Se estiman los parámetros mediante técnicas de regresión lineal.

### III-A. Controlador implementado para la identificación

Para realizar el experimento de identificación, es necesario definir una trayectoria deseada tal que se excite totalmente la dinámica del quadrotor para lograr una mejor identificación de los parámetros del sistema. Debido a la inestabilidad de este sistema, tales experimentos no pueden realizarse en lazo abierto. Por lo tanto, el experimento es realizado a lazo cerrado mediante la implementación de un controlador. El controlador utilizado es un control de posición y orientación proporcional-integral-derivativo (PID), descrito como

$$\mathbf{U}_f = m g \mathbf{e}_z + \ddot{\mathbf{p}}_d m + K_{pp} \dot{\mathbf{p}} + K_{ip} \int_0^t \dot{\mathbf{p}} dt + K_{dp} \dot{\mathbf{p}}, \quad (13)$$

$$\mathbf{f}_{Bp} = \begin{bmatrix} 0 \\ 0 \\ F_z \end{bmatrix}, \quad (14)$$

$$\boldsymbol{\tau} = W(\boldsymbol{\eta})^T \left( K_{po} \tilde{\boldsymbol{\eta}} + K_{io} \int_0^t \tilde{\boldsymbol{\eta}} dt + K_{do} \dot{\tilde{\boldsymbol{\eta}}} \right), \quad (15)$$

donde,  $\mathbf{U}_f = [f_x \ f_y \ f_z]^T$  es el vector de fuerzas en el marco de referencia inercial,  $F_z = f_z / (\cos(\phi) \cos(\theta))$  es el empuje total,  $K_{pp} \in \mathbb{R}^{3 \times 3}$ ,  $K_{ip} \in \mathbb{R}^{3 \times 3}$  y  $K_{dp} \in \mathbb{R}^{3 \times 3}$  son las matrices de ganancias diagonales y definidas positivas para las acciones de control de posición proporcional, integral y derivativa, respectivamente,  $K_{po} \in \mathbb{R}^{3 \times 3}$ ,  $K_{io} \in \mathbb{R}^{3 \times 3}$  y  $K_{do} \in \mathbb{R}^{3 \times 3}$  son las matrices de ganancias diagonales y definidas positivas para las acciones de control de orientación proporcional, integral y derivativa, respectivamente,  $\tilde{\mathbf{p}} = \mathbf{p}_d - \mathbf{p}$  denota el error de posición,  $\mathbf{p}_d$  representa la posición deseada,  $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}_d - \boldsymbol{\eta}$  define el error de orientación,  $\boldsymbol{\eta}_d$  denota la orientación deseada, donde  $\phi_d$  y  $\theta_d$  están dadas por [6]

$$\theta_d = \tan^{-1} \left[ \frac{m}{f_z} (f_x \cos(\psi_d) + f_y \sin(\psi_d)) \right], \quad (16)$$

$$\phi_d = \tan^{-1} \left[ \frac{m \cos(\theta_d)}{f_z} (f_x \sin(\psi_d) - f_y \cos(\psi_d)) \right], \quad (17)$$

respectivamente.

La posición y orientación deseados  $\mathbf{p}_d(t)$  y  $\psi_d(t)$  fueron definidos tal que se produzca una trayectoria circular en el espacio cartesiano, la cual debe ser completada por el quadrotor en 5 [s]. Esta trayectoria forma una circunferencia de 0.5 [m] de radio en el plano  $(x,y)$ , a una altitud variable, que está descrita como

$$x_d(t) = 0.5 \cos\left(\frac{2\pi}{5}t\right) \text{ [m]} \quad \forall t, \quad (18)$$

$$y_d(t) = 0.5 \sin\left(\frac{2\pi}{5}t\right) \text{ [m]} \quad \forall t, \quad (19)$$

$$z_d(t) = \begin{cases} 0.3 + 0.0279t^3 - 0.0074t^4 + 0.0005t^5 \text{ [m]} \\ 0 \leq t < 5 \\ 0.9 + 0.1 \sin\left(\frac{2\pi}{5}t\right) \text{ [m]}, t \geq 5 \end{cases} \quad (20)$$

$$\psi_d(t) = 15 \sin\left(\frac{2\pi}{5}t\right) \text{ [}^\circ\text{]} \quad \forall t. \quad (21)$$

### III-B. Proceso de identificación fuera de línea

El proceso de identificación inicia con la obtención de series de tiempo de  $\mathbf{p}(k)$ ,  $\boldsymbol{\eta}(k)$  y  $\mathbf{F}(k)$ , donde  $k = 0, 1, 2, \dots, N-1$  siendo  $N$  el número total de muestras. Los elementos de la matriz de regresión  $\Phi$  están en función de  $\mathbf{v}$ ,  $\boldsymbol{\omega}$ ,  $\dot{\mathbf{v}}$  y  $\dot{\boldsymbol{\omega}}$  que a su vez dependen de  $\boldsymbol{\eta}$ ,  $\dot{\mathbf{p}}$  y  $\dot{\boldsymbol{\eta}}$ . Los quadrotores cuentan con unidades de sensado inercial (IMU del inglés *Inertial Measurement Unit*) que proporcionan información sobre su orientación, esto, por medio de acelerómetros y giroscopios. Sin embargo, determinar la posición del quadrotor a través de la unidad de sensado inercial no es factible para fines de identificación paramétrica. Por tal motivo, en este trabajo se utilizó el sistema de visión *Optitrack* del cual se obtuvo la posición del quadrotor.

Para el procedimiento descrito a continuación se supone que únicamente se encuentran disponibles mediciones de la posición  $\mathbf{p}(k)$ , la orientación  $\boldsymbol{\eta}(k)$  y la acción de control  $\mathbf{F}(k)$ . La información obtenida de ambos sensores puede contener ruido de alta frecuencia debido al tiempo de muestreo y a la naturaleza discreta del sistema. Por lo tanto, las velocidades

y tasas de cambio del quadrotor tienen que ser obtenidas por medio de diferenciación numérica de la posición obtenida del sistema de visión y de la orientación obtenida de la unidad de sensado inercial del vehículo. Cabe mencionar que las derivadas tanto de posición  $\mathbf{p}(k)$  como de orientación  $\boldsymbol{\eta}(k)$  amplifican el ruido de las señales, esto debido a que el proceso de diferenciación actúa como un filtro pasa altas [12]. Para calcular la velocidad y tasa de cambio de los ángulos de Euler  $\boldsymbol{\eta}(k)$  y reducir los efectos descritos anteriormente se filtraron las señales de posición  $\mathbf{p}(k)$  y orientación  $\boldsymbol{\eta}(k)$  siguiendo el procedimiento descrito a continuación:

1. *Filtrado*. Se diseñó un filtro digital pasa bajas  $f_p(z)$  que elimina las altas frecuencias consecuencia del período de muestreo y permite pasar las bajas frecuencias de la trayectoria seguida por el quadrotor. Se implementó el esquema de filtrado fase cero denominado filtrado en ambas direcciones “Adelante-Atrás” para evitar la distorsión de muestras de la posición y orientación. Este esquema de filtrado es fácil de implementar en *Matlab* mediante la función “filtfilt”. De esta manera se obtiene la posición filtrada  $\mathbf{p}_f(k)$  y la orientación filtrada  $\boldsymbol{\eta}_f(k)$ .
2. *Diferenciación numérica*. Para calcular la velocidad y tasa de cambio de los ángulos de Euler se utilizó el algoritmo de diferenciación central con el propósito de evitar el corrimiento de fase [10], [11] y [12]. El algoritmo de diferenciación central se define como

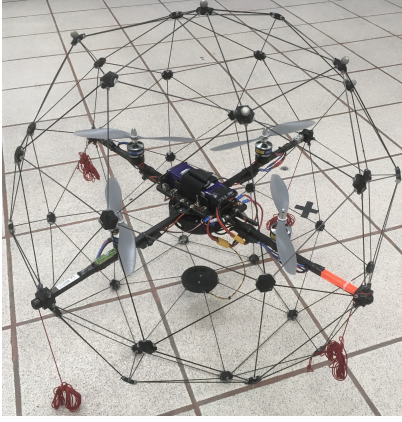
$$\dot{a}(k) = \frac{a_f(k+1) - a_f(k-1)}{2T}, \quad (22)$$

donde  $k$  representa el instante de muestreo,  $T$  es el período de muestreo en segundos y  $a_f(k) \in \mathbb{R}$  denota el vector de elementos filtrados utilizando el filtro pasa bajas  $f_p(z)$  y la función “filtfilt”.

3. *Transformación de las señales*. La identificación se realiza utilizando el modelo parametrizado (9), el cual está expresado en el marco de referencia del cuerpo del quadrotor, por tal motivo hay que transformar las derivadas  $\dot{\mathbf{p}}(k)$  y  $\dot{\boldsymbol{\eta}}(k)$  obtenidas del algoritmo de diferenciación central (22). Para llevar estas señales de un marco de referencia a otro se utilizan las expresiones (3) y (4), y así es como se obtiene la velocidad lineal  $\mathbf{v}(k)$  y la velocidad angular  $\boldsymbol{\omega}(k)$ . Posteriormente se aplica el algoritmo de diferenciación central (22) a ambas señales para obtener sus derivadas, ya que son requeridas en la matriz de regresión del modelo parametrizado (9).
4. *Filtrado de las entradas de control*. Puesto que las señales correspondientes a las entradas de control también presentan ruido, es necesario filtrarlas. Para esto, se utiliza el mismo filtro digital pasa bajas  $f_p(z)$  en conjunto con la función “filtfilt” de *Matlab*.

Los parámetros del vector  $\Theta$  se estiman como la solución por mínimos cuadrados de un sistema lineal sobredeterminado, resultado de las muestras obtenidas del modelo dinámico del quadrotor en múltiples instantes de tiempo a lo largo de las trayectorias  $\mathbf{p}(k)$ ,  $\boldsymbol{\eta}(k)$ ,  $\dot{\mathbf{p}}(k)$  y  $\dot{\boldsymbol{\eta}}(k)$ . Considérese que se tiene  $N$  muestras del vector de entradas de control  $\mathbf{F} \in \mathbb{R}^4$ , de la




 Figura 1: Quadrotor *QBall 2* de *Quanser*.

posición  $\mathbf{p}(k) \in \mathbb{R}^3$  y orientación  $\boldsymbol{\eta}(k) \in \mathbb{R}^3$  a un período de muestreo de  $T = 0.033$  [s]. El modelo parametrizado del quadrotor utilizado para la identificación está dado como

$$\Phi_f(\mathbf{v}, \boldsymbol{\omega}, \boldsymbol{\eta}_f, \dot{\mathbf{v}}, \dot{\boldsymbol{\omega}})\boldsymbol{\Theta} = \mathbf{F}_f \quad (23)$$

donde  $\mathbf{F}_f \in \mathbb{R}^4$  es el vector de entradas de control filtrado,  $\Phi_f(\mathbf{v}, \boldsymbol{\omega}, \boldsymbol{\eta}, \dot{\mathbf{v}}, \dot{\boldsymbol{\omega}}) \in \mathbb{R}^{4 \times 4}$  es la matriz de regresión filtrada compuesta por funciones conocidas y  $\boldsymbol{\Theta} \in \mathbb{R}^4$  es el vector de parámetros a estimar.

La estimación del vector de parámetros del modelo (23) utilizando el algoritmo de mínimos cuadrados se tiene como

$$\hat{\boldsymbol{\Theta}}(k) = \left[ \sum_{j=0}^k \Phi_f^T(j)\Phi_f(j) \right]^{-1} \sum_{j=0}^k \Phi_f^T(j)\mathbf{F}_f(j), \quad (24)$$

donde se ha considerado que  $\hat{\boldsymbol{\Theta}}(k) \in \mathbb{R}^4$  es la estimación de los parámetros reales  $\boldsymbol{\Theta}(k) \in \mathbb{R}^4$  en función del tiempo discreto  $k$ , con  $0 \leq k \leq N-1$  y  $N$  es el número total de muestras.

## IV. RESULTADOS

### IV-A. Descripción de la plataforma experimental

Los resultados presentados en este trabajo fueron obtenidos mediante el quadrotor *QBall 2* que se muestra en la figura 1. La implementación del algoritmo de control se hizo en *Matlab*, *Simulink* empleando el software *Quarc*, necesarios para generar el código ejecutable que se transfiere a la computadora *Gumstix* a bordo del quadrotor. El quadrotor cuenta con un acelerómetro y un giroscopio, ambos capaces de sensar en los tres ejes. La orientación  $\boldsymbol{\eta}(t)$  fue obtenida mediante la unidad de sensado inercial del quadrotor a una frecuencia de muestreo de 500 [Hz]. La posición  $\mathbf{p}(t)$  se obtuvo mediante el uso del sistema de visión *Optitrack* a una frecuencia de muestreo de 30 [Hz]. El sistema de visión *Optitrack* está conformado por seis cámaras conectadas a una estación terrena. Los valores de los parámetros del *QBall 2* utilizados dentro del controlador, se presentan en la tabla I [6]. La distancia entre el centro de masa del quadrotor y el eje de rotación de los rotores está dada por  $l = 0.2$  [m] y la constante de aceleración gravitacional por  $g = 9.8$  [m/s<sup>2</sup>].

 Tabla I: Parámetros de la plataforma *QBall 2* de *Quanser*.

Parámetro	Descripción	Valor	Unidad
$m$	Masa del quadrotor	1.79	[kg]
$I_{xx}$	Momento de inercia respecto al eje $x$	0.03	[kg m <sup>2</sup> ]
$I_{yy}$	Momento de inercia respecto al eje $y$	0.03	[kg m <sup>2</sup> ]
$I_{zz}$	Momento de inercia respecto al eje $z$	0.04	[kg m <sup>2</sup> ]

 Tabla II: Parámetros de la plataforma *QBall 2* de *Quanser* identificados mediante el método de mínimos cuadrados.

Parámetro	Descripción	Valor	Unidad
$m$	Masa del quadrotor	1.7917	[kg]
$I_{xx}$	Momento de inercia respecto al eje $x$	0.0140	[kg m <sup>2</sup> ]
$I_{yy}$	Momento de inercia respecto al eje $y$	0.0148	[kg m <sup>2</sup> ]
$I_{zz}$	Momento de inercia respecto al eje $z$	0.0013	[kg m <sup>2</sup> ]

Las ganancias del controlador PID descrito en (13)-(17) fueron definidas como

$$\begin{aligned} K_{pp} &= \text{diag}\{9.0 \ 9.0 \ 9.0\}, \\ K_{ip} &= \text{diag}\{0.03 \ 0.04 \ 1.3\}, \\ K_{dp} &= \text{diag}\{6.0 \ 6.0 \ 3.5\}, \\ K_{po} &= \text{diag}\{1.25 \ 1.25 \ 2.60\}, \\ K_{io} &= \text{diag}\{0.05 \ 0.05 \ 0.01\}, \\ K_{do} &= \text{diag}\{0.35 \ 0.35 \ 1.0\}. \end{aligned}$$

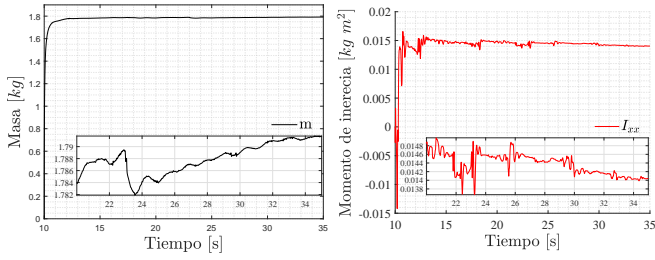
### IV-B. Resultados de la identificación

Con las series de datos de  $\dot{\mathbf{p}}(k)$ ,  $\dot{\boldsymbol{\eta}}(k)$  y  $\mathbf{F}(k)$  obtenidas del experimento descrito en la sección III-A se ha implementado la identificación paramétrica descrita en la sección III-B. Los resultados de los parámetros estimados se muestran en la tabla II. Para la implementación del algoritmo de mínimos cuadrados se utilizaron los datos comprendidos entre  $10 \leq t \leq 35$ , con el propósito de discriminar las muestras contenidas dentro del período transitorio del experimento [9]. En la figura 2 se observa la evolución en el tiempo de la estimación de los parámetros del quadrotor *QBall 2* y se puede apreciar como todos los parámetros convergen a un valor conforme avanza el tiempo. Los valores presentados en la tabla II corresponden al último valor de la serie de datos correspondiente a cada parámetro estimado.

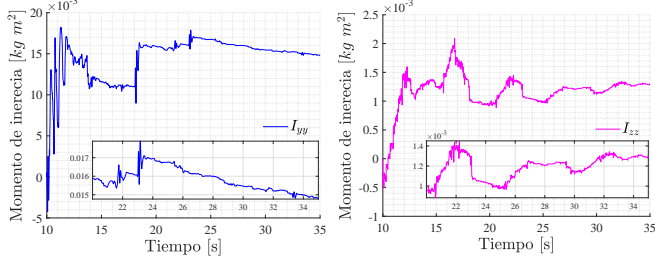
### IV-C. Comparación entre simulaciones numéricas y experimento

Las simulaciones numéricas se realizaron en *Matlab-Simulink*, y se representó la diferencia entre los tiempos de muestreo presentes en el experimento utilizando el bloque *Zero-Order Hold*. Para simular el sensor de posición se utilizó un período de muestreo de 30 [Hz] y para simular el sensor de orientación se empleó un período de muestreo de 500 [Hz], se procedió de igual forma para sus derivadas.

En la figura 3 se presenta la comparación entre experimento y simulaciones de las trayectorias de posición y orientación del quadrotor. Como puede apreciarse, los resultados de posición



(a) Estimación de la masa del quadrotor. (b) Estimación del momento de inercia en el eje  $x$ .



(c) Estimación del momento de inercia en el eje  $y$ . (d) Estimación del momento de inercia en el eje  $z$ .

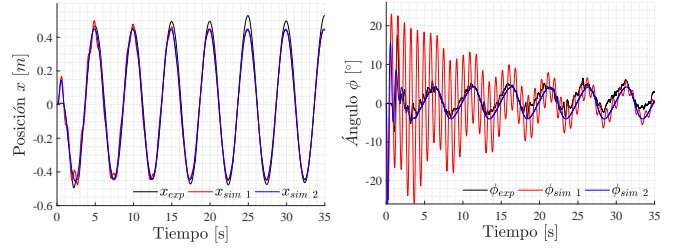
Figura 2: Parámetros estimados mediante algoritmo de mínimos cuadrados.

obtenidos en simulación son similares a lo obtenido experimentalmente, cabe mencionar que los resultados obtenidos en ambas simulaciones para  $x$ ,  $y$ ,  $z$  y  $\psi$  son cercanos, esto se le atribuye a la similitud de los parámetros de masa y los momentos de inercia provistos por el fabricante con respecto de los estimados en este trabajo. Sin embargo, en la comparación correspondiente a los ángulos  $\phi$  y  $\theta$  se puede observar que los resultados obtenidos en simulación utilizando los parámetros estimados resultan más cercanos a los obtenidos experimentalmente. Este es un claro indicador de la similitud de los parámetros estimados respecto a los reales.

En la figura 4 se presenta la comparación entre las acciones de control experimentales y las obtenidas de la simulación. La diferencia entre las señales correspondientes a los torques de control  $\tau_\phi$ ,  $\tau_\theta$  y  $\tau_\psi$  es considerable, mientras que las acciones de control correspondientes al empuje total  $F_z$  guardan cierta similitud entre simulación y lo obtenido experimentalmente. Las diferencias entre las señales de control se atribuyen principalmente a que la dinámica de los actuadores no es considerada en el modelo parametrizado filtrado (23) y a que tampoco se consideró esta dinámica en las simulaciones así como la distribución de acciones de control a cada uno de los rotores del quadrotor.

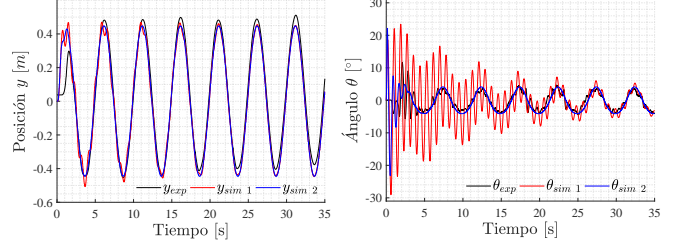
El índice utilizado para la comparación presentada en este trabajo es el valor cuadrático medio ó RMS (*Root Mean Square*, por sus siglas en inglés) del error calculado como la diferencia entre las señales obtenidas experimentalmente y las señales obtenidas en simulación, esto es

$$e_\alpha = \alpha_{exp} - \alpha_{sim}, \quad (25)$$



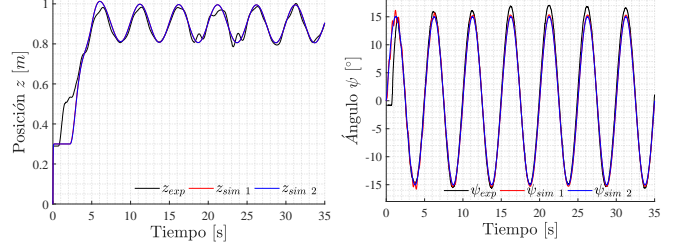
(a) Posición en el eje  $x$ .

(b) Ángulo  $\phi$ .



(c) Posición en el eje  $y$ .

(d) Ángulo  $\theta$ .



(e) Posición en el eje  $z$ .

(f) Ángulo  $\psi$ .

Figura 3: Comparación entre la posición y orientación de los datos experimentales y las simulaciones. El subíndice *exp* representa los datos del experimento, el subíndice *sim 1* representa la simulación con los parámetros de la tabla I y el subíndice *sim 2* representa la simulación con los parámetros identificados presentados en la tabla II.

Tabla III: Valor RMS del error definido como la diferencia entre las señales obtenidas del experimento y las señales obtenidas de la simulación del quadrotor con los parámetros nominales de la tabla I e identificados de la tabla II usando el controlador PID de la sección III-A. Los valores RMS de los errores fueron calculados para  $10 \leq t \leq 35$ .

Error	Unidades	Nominal	Identificado	Mejoría
$e_x$	[m]	0.0308	0.0294	4.3920%
$e_y$	[m]	0.0373	0.0358	4.2622%
$e_z$	[m]	0.01419	0.01415	0.3214%
$e_\phi$	[rad]	0.0633	0.0271	57.2413%
$e_\theta$	[rad]	0.0483	0.0148	69.3485%
$e_\psi$	[rad]	0.0166	0.0177	-6.5383%
$e_{F_z}$	[N]	0.2429	0.2215	8.8160%
$e_{\tau_\phi}$	[N m]	0.3774	0.3625	3.9513%
$e_{\tau_\theta}$	[N m]	0.3617	0.3523	2.5907%
$e_{\tau_\psi}$	[N m]	0.0681	0.0491	27.8788%

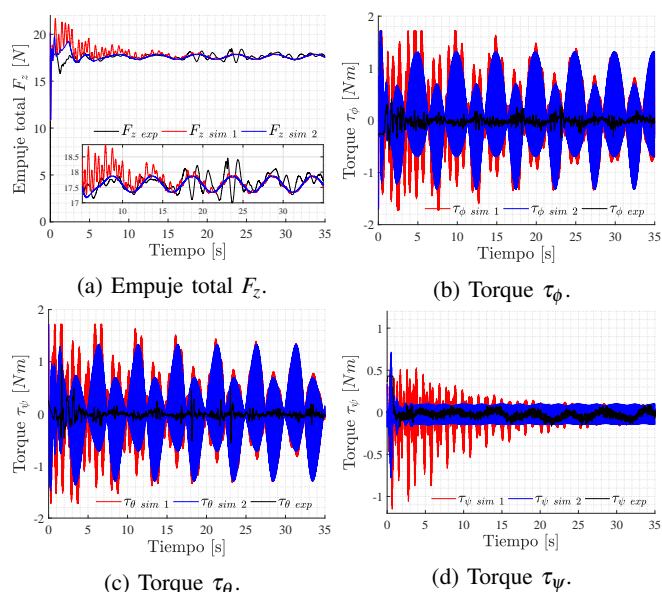


Figura 4: Comparación entre las acciones de control de los datos experimentales y las simulaciones. El subíndice *exp* representa los datos del experimento, el subíndice *sim 1* representa la simulación con los parámetros de la tabla I y el subíndice *sim 2* representa la simulación con los parámetros identificados presentados en la tabla II.

donde,  $\alpha = x, y, z, \phi, \theta, \psi, F_z, \tau_\phi, \tau_\theta, \tau_\psi$ . En la tabla III se presentan los resultados obtenidos al calcular el valor RMS de  $e_\alpha(t)$  cuando se usan los parámetros nominales de la tabla I en simulación y de  $e_\alpha(t)$  cuando se usan los parámetros identificados de la tabla II en simulación. Como se puede observar los errores obtenidos utilizando los parámetros identificados son menores. Por lo tanto, es posible pensar que los parámetros obtenidos reflejan mayor precisión con respecto a la dinámica real del quadrotor.

## V. CONCLUSIONES

En este trabajo se presentó la identificación paramétrica de la masa y los momentos de inercia del quadrotor *QBall 2* de *Quanser*. Se propuso una parametrización del modelo dinámico de un quadrotor. Se implementó un controlador con estructura PID para el seguimiento de trayectorias en posición y orientación. Se presentó una metodología para el filtrado y procesamiento de datos experimentales, de tal forma que se obtuvieran las señales necesarias para construir el modelo parametrizado filtrado. Se implementó el algoritmo de mínimos cuadrados para estimar los parámetros fuera de línea. Se validaron los resultados obtenidos al recrear en simulación el experimento utilizando los parámetros nominales y los estimados, y posteriormente compararlos con los resultados experimentales. La diferencia entre lo obtenido en experimento y simulación se atribuye a los efectos de los distintos períodos de muestreo presentes en el experimento, así como la dinámica de los actuadores que no fue contemplada. Los resultados

obtenidos de los valores RMS de los errores respaldan la metodología utilizada para esta identificación.

Como trabajo futuro, se pretende realizar una nueva identificación de la misma plataforma *QBall 2* suponiendo que el cuerpo del quadrotor no es completamente simétrico, tomando en cuenta la dinámica de los actuadores y utilizando un algoritmo de optimización de trayectorias para definir la trayectoria del quadrotor para el experimento.

## AGRADECIMIENTOS

Este proyecto es apoyado parcialmente por la Secretaría de Investigación y Posgrado - IPN, CONACYT A1-S-24762 y Cátedras 1537.

## REFERENCIAS

- [1] D. Morris, X. Chen y A. Kind, "Real-time system identification of quadrotor dynamics", en memorias del *13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2018, págs. 1563-1568.
- [2] J. Moreno-Valenzuela y C. Aguilar-Avelar, *Motion Control of Underactuated Mechanical Systems*. Champ: Springer International Publishing AG, 2018.
- [3] L. Li, L. Sun y J. Jin, "Survey of advances in control algorithms of quadrotor unmanned aerial vehicle", en memorias del *IEEE 16th International Conference on Communication Technology (ICCT)*, 2015, págs. 107-111.
- [4] F. Kendoul, Z. Yu y K. Nonami, "Guidance and nonlinear control system for autonomous flight of micro-robotcraft unmanned aerial vehicles". *Journal of Field Robotics*, 2010, vol. 27, núm. 3, págs. 311-334.
- [5] R. Pérez-Alcocer, J. Moreno-Valenzuela y R. Miranda-Colorado, "A robust approach for trajectory tracking control of a quadrotor with experimental validation", *ISA Transactions*, 2016, vol. 65, págs. 262-274.
- [6] J. Moreno-Valenzuela, R. Pérez-Alcocer, M. Guerrero-Medina y A. Dzul, "Nonlinear PID-type controller for quadrotor trajectory tracking", *IEEE/ASME Transactions on Mechatronics*, 2018, vol. 23, núm. 5, págs. 2436-2447.
- [7] V.K. Chitrakaran, D.M. Dawson, H. Kannan y M. Feemster, "Vision-based tracking for unmanned aerial vehicles", Dept. of Electrical and Computer Engineering, Clemson University, CU/CRB/2/27/06/1, 2006.
- [8] J. Swevers, W. Verdonck y J. De Schutter, "Dynamic model identification for industrial robots", *IEEE Control Systems Magazine*, 2007, vol. 27, núm. 5, págs. 58-71.
- [9] J. Moreno-Valenzuela, R. Miranda-Colorado y C. Aguilar-Avelar, "A matlab-based identification procedure applied to a two-degrees-of-freedom robot manipulator for engineering students", *International Journal of Electrical Engineering Education*, 2017, vol. 54, núm. 4, págs. 319-340.
- [10] M. Gautier, "Dynamic identification of robots with power model", en memorias del *IEEE International Conference on Robotics and Automation*, Albuquerque, 1997, págs. 1922-1927.
- [11] M. Gautier y P. Pognet, "Extended Kalman filtering and weighted least square dynamic identification of robot", *Control Engineering Practice*, 2001, vol. 9, págs. 1361-1372.
- [12] W. Khalil y E. Dombre, *Modeling, Identification and Control of Robots*, 3.<sup>a</sup> ed. Bristol: Taylor & Francis, Inc., 2002.

# Diseño de un sistema de interacción natural con robots móviles utilizando gestos de la mano

Karina Reyes-Lio\*, Gustavo Castellón-Ramírez\*, Cristina Ramírez-Fernández\*,  
Marlen Meza-Sánchez§, José Luis Monay-Arredondo‡

\* Tecnológico Nacional de México/I.T. Ensenada

Email: {sreyes, gcastillon, cramirez}@ite.edu.mx

§ Email: marlen.meza.sanchez@gmail.com

‡ Tecnológico Nacional de México/I.T. Tijuana

Email: luismonay@gmail.com

**Resumen**—En este trabajo se presenta el desarrollo de un dispositivo de interacción humano-computadora (HCI - Human Computer Interaction) enfocado a la organización de procesos cognitivos a través de gestos de las manos. El sistema propuesto crea un entorno interactivo de reconocimiento de gestos para controlar de manera inalámbrica el movimiento de un robot móvil tipo diferencial. El movimiento implementado es a su vez imitado por otro robot móvil colocado a una distancia fija. Se presentan resultados experimentales de la validación del sistema de interacción natural propuesto, en donde los movimientos ejecutados por cada robot son parametrizados con respecto a sus velocidades lineales y angulares, así como a la trayectoria que describen.

## I. INTRODUCCIÓN

La evolución tecnológica ha permitido que las áreas de aplicación de los sistemas robóticos se diversifiquen. Por ejemplo, se han propuesto sistemas de interacción humano-computadora que abordan problemas de rehabilitación motriz; en [1] se presenta una revisión de robots orientados a las extremidades superiores; una visión más general que aborda las tendencias y condiciones electrónicas, de materiales, y tecnologías de esta clase de robots, es presentado por [2]. Un sistema de interacción natural con robots móviles colaborativos se presenta en [3], en donde gestos de los antebrazos son mapeados a velocidades de los robots. Adicionalmente, utilizan un dispositivo smartwatch para medir los niveles de estrés del usuario mientras utiliza el sistema. Una revisión del impacto del canal de comunicación, los algoritmos utilizados para controlar los robots, la visualización del usuario, y los aspectos básicos de sistemas de interacción humana con sistemas robóticos colaborativos es presentada en [4]. En [5] se propone un juego serio llamado ROBiGAME, basado en el dispositivo robótico REAplan, con el objetivo de mejorar déficits motrices y cognitivos (específicamente hemiparesia y heminegligencia).

Otra área de aplicación que ha motivado muchas investigaciones se concentra en la evaluación de los procesos cognitivos del ser humano al interactuar con sistemas robóticos. Esta clase de sistemas requieren el diseño de interfaces llamativas y de manera particular, que motiven al usuario a enfocarse positivamente. Por ejemplo, una revisión de robots que interactúan socialmente se presenta en [6], en donde se profundiza

la relación de las áreas de investigación con diversas formas de robots sociales, las metodologías aplicadas a su diseño, y el impacto de estos robots en los seres humanos. La influencia de la interacción y respuesta emocional de niños autistas con robots tipo humanoide ha sido evaluado en [7], con el objetivo de establecer criterios para crear sesiones significativas e interacciones exitosas.

De manera general, este tipo de sistemas puede ser enfocado también como una forma entretenida de medir variables fisiológicas del ser humano y a la evaluación de las experiencias con sistemas robóticos. En [8] se propone una metodología aplicada al diseño y supervisión de un experiencia interactiva. Tal metodología es implementada a través de la creación de un juego interactivo con un robot humanoide Nao. Un sistema de detección de gestos de una mano para controlar un robot manipulador PUMA se presenta en [9]. Un sistema similar que utiliza un dispositivo *Leap motion* en el reconocimiento de gestos de la mano para controlar un robot manipulador UR10 se propone en [10]. Otros trabajo enfocados al control por gestos de las manos en robots móviles se proponen en [11] y [12]; en [13] y [14] se presentan resultados utilizando el dispositivo Leap Motion.

La contribución de este trabajo corresponde al diseño de un sistema de interacción natural intuitivo para la realización de tareas de coordinación utilizando gestos de las manos. Este sistema utiliza un servidor de mensajes a través de WiFi para enviar la interpretación de los gestos realizados por el usuario a un robot principal. A su vez, el robot principal debe confirmar la recepción del comando para que el servidor de mensajes envíe el mismo comando a un segundo robot móvil. Al recibir el comando, el segundo robot debe imitar el movimiento del robot principal. Los comandos gestuales han sido definidos para generar los movimientos básicos en robots móviles con ruedas que le permiten construir trayectorias complejas. Los parámetros físicos y el modelo cinemático han sido analizados; y se presenta una validación experimental de la implementación de velocidades lineales y angulares en ellos para lograr las trayectorias solicitadas por el usuario.

Este documento está organizado como se describe a continuación. En la Sección II se presenta las características técnicas y funcionalidad del sistema de interacción natural



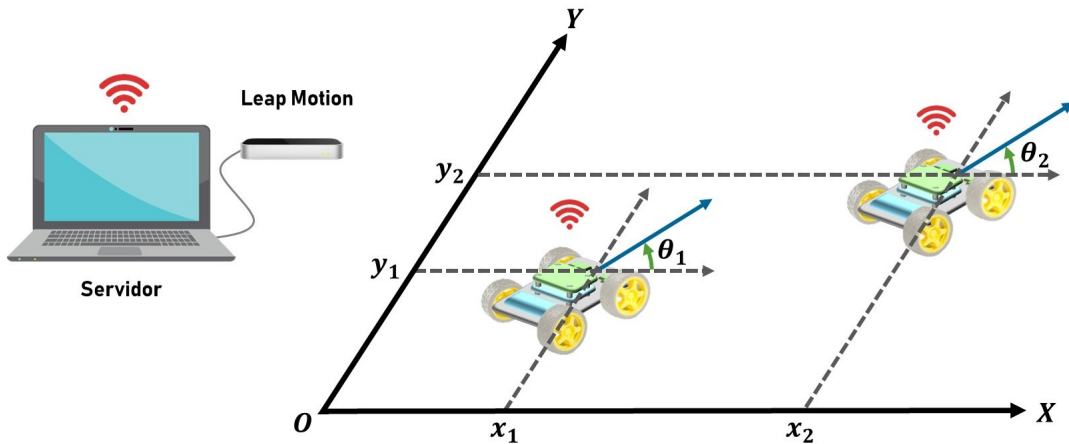


Figura 1: Esquema del sistema interactivo utilizando gestos de las manos para controlar el movimiento de un robot móvil tipo diferencial, y la imitación de este movimiento por un segundo robot colocado a una distancia fija.

propuesto. La validación experimental aplicando una ejecución de prueba y la medición de las variables de interés se presentan en la Sección III. Finalmente, las conclusiones y trabajo futuro se presentan en la Sección IV.

## II. SISTEMA DE INTERACCIÓN NATURAL

El esquema general del sistema de interacción desarrollado se muestra en la Figura 1. La interacción natural del usuario con el sistema requiere el establecimiento de cinco comandos gestuales para guiar el movimiento de un robot móvil tipo diferencial. Los comandos establecidos son

- $\mathcal{A}^+$  Movimiento hacia adelante.
- $\mathcal{A}^-$  Movimiento hacia atrás.
- $\mathcal{G}_d$  Giro a la derecha.
- $\mathcal{G}_i$  Giro a la izquierda.
- $\mathcal{D}$  Detener.

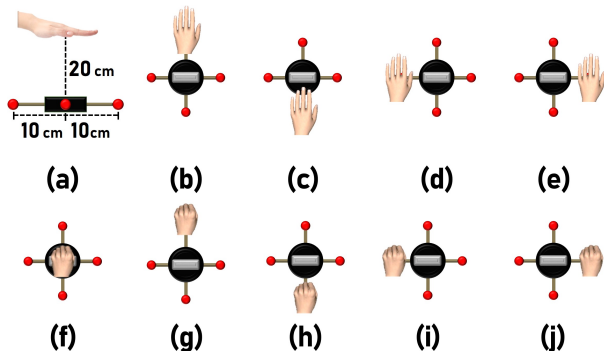


Figura 2: (a) distancias requeridas para el sensor, (b) comando  $\mathcal{A}^+$ , (c) comando  $\mathcal{G}_i$ , (d) comando  $\mathcal{A}^-$ , (e) comando  $\mathcal{G}_d$ , (f-j) comando  $\mathcal{D}$ .

La interpretación del gesto de la mano lo realiza un dispositivo *Leap Motion* [15]. En la Figura 2 se muestra el área de trabajo del dispositivo y los gestos de la mano que corresponden a cada comando de movimiento para los robots.

Esto es, en Figura 2(a) se muestran las distancias que requiere el sensor para reconocer el gesto; en Figura 2(b) el comando de avanzar hacia adelante se activa con mano extendida hacia adelante ( $\mathcal{A}^+$ ); la Figura 2(c) mano extendida giro hacia la derecha ( $\mathcal{G}_d$ ); en la Figura 2(d) muestra mano extendida hacia atrás ( $\mathcal{G}_i$ ); en la Figura 2(e) presenta mano extendida giro hacia la izquierda ( $\mathcal{G}_i$ ); y comando de detenerse ( $\mathcal{D}$ ) con mano en puño en cualquier rumbo o directamente sobre el sensor como se muestra en la Figura 2(f-j). La mano debe colocarse a una distancia de 20 cm por encima del sensor.

El funcionamiento del sistema interactivo desde el servidor se resume en el diagrama de flujo de la Figura 3. La condición de terminación de cada sesión con el sistema puede especificarse en término de tiempo o por número de movimientos. La programación del sistema ha sido realizada en C-Sharp a través de un servidor de mensajes utilizando comunicación inalámbrica WiFi a través del protocolo de control de transmisión (TCP – Transmission Control Protocol). Este sistema se encarga del proceso realizado por el *Leap Motion*; esto es, obtener las imágenes de las cámaras del dispositivo, aplicar una corrección a la distorsión de la imagen, determinar la configuración de cada mano, y ejecutar un algoritmo de visión estereoscópica entre cada par de imágenes para determinar la posición en el plano tridimensional. De los datos arrojados por este proceso, únicamente se toma en cuenta la cantidad de dedos, y el rumbo de la mano con respecto al sensor. Posteriormente, el servidor se encarga de la gestión de la conectividad en la retroalimentación de comandos de los gestos reconocidos por el sensor hacia los robots móviles. La interfaz desarrollada para el servidor se muestra en la Figura 4, en donde se observa la confirmación de la conexión inalámbrica con los robots móviles y en inicio de recepción de datos de tiempo y distancias recorridas.

Los robots móviles utilizados para implementar el movimiento generado por la interacción natural son Xiao R Geek DS Wireless WiFi Robot. Este robot posee cuatro ruedas pero funciona como un robot móvil diferencial, ya que una misma

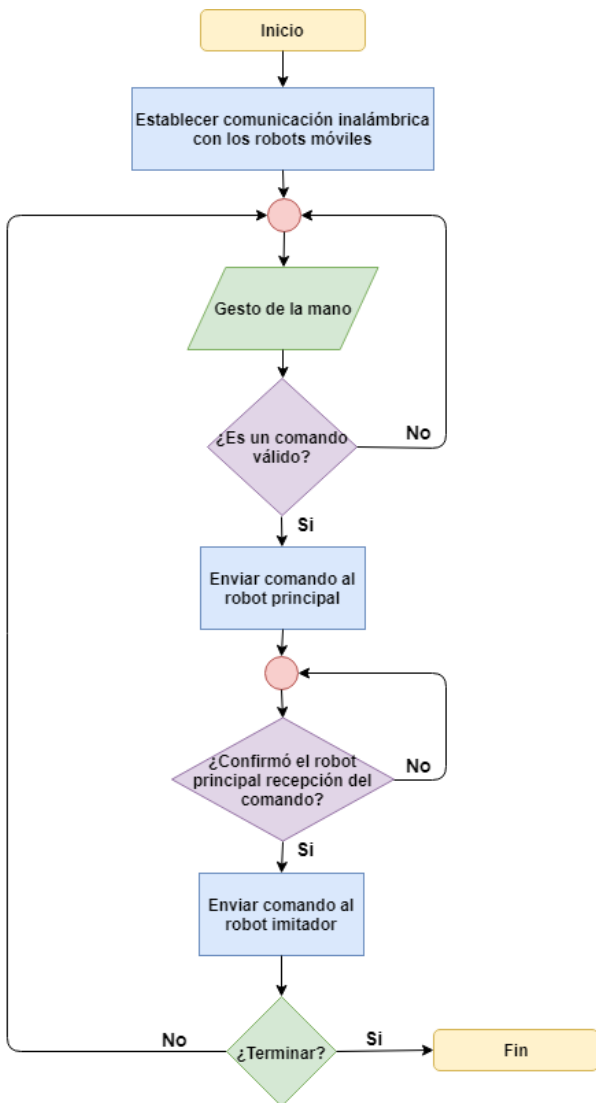


Figura 3: Funcionamiento del sistema de interacción natural implementado por el servidor de mensajes utilizando comunicación inalámbrica.

señal de control activa simultáneamente dos ruedas de cada lado. El diámetro de cada rueda es de  $d = 0.06792$  m y la distancia entre ellas es  $L = 0.12894$  m. Por lo tanto, con cada revolución de la rueda se tiene un desplazamiento de  $\pi d = 0.21337$  m. Considerando un decodificador óptico con una resolución de 20 ppr (pulsos por revolución) para cada rueda, se obtiene una resolución de  $1 \text{ ppr} = 0.01070$  m.

**Observación.** Considerando la resolución del decodificador y la rapidez con la que los robots deben reaccionar instantáneamente de acuerdo a los gestos del usuario, se procede a caracterizar velocidades constantes para implementar cada movimiento. La validación experimental del sistema se realiza mediante la medición de las distancias recorridas por cada rueda.

De acuerdo a los resultados de prueba para evitar derrapes

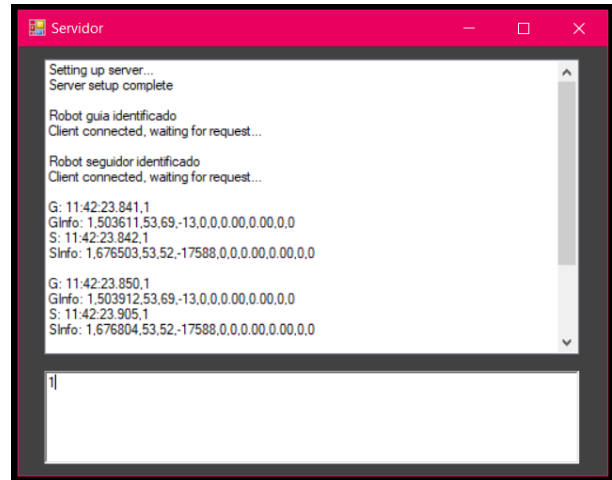


Figura 4: Interfaz del sistema de interacción natural ejecutándose en el servidor.

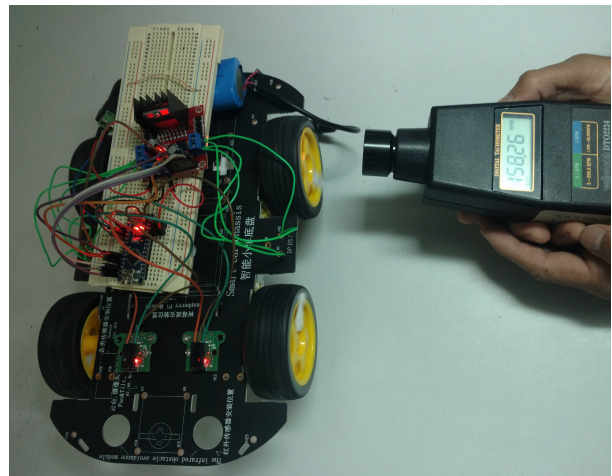


Figura 5: Caracterización experimental de la velocidad en [rpm] (revoluciones por minuto) de las ruedas del robot móvil diferencial aplicando una señal de pulso modulado PWM (Pulse-Width Modulation).

y pérdida de tracción en el movimiento del robot, éste ha sido caracterizado en rev/min (revoluciones por minuto) aplicando una señal de ancho de pulso modulado (PWM – Pulse-Width Modulation) a un 30% del ciclo de trabajo de una batería de 7.4 volts. Bajo estos parámetros, se realizaron mediciones con un tacómetro (como se observa en la Figura 5) para obtener una velocidad aproximada de 144 rev/min. Por lo tanto, cada rueda se desplaza a una velocidad lineal constante de

$$144 \frac{\text{rev}}{\text{min}} \times \frac{0.21337 \text{ m}}{1 \text{ rev}} \times \frac{1 \text{ min}}{60 \text{ s}} = 0.5121 \text{ m/s.}$$

Sean  $v_R$  y  $v_L$  las velocidades de las ruedas derecha e izquierda, respectivamente. Las velocidad lineal a la cual realizará los movimientos el robot está dado como

$$v = \frac{v_R + v_L}{2} = 0.5121 \text{ m/s.} \quad (1)$$



Para realizar los giros, se aplicará la misma velocidad constante a las ruedas, donde los signos de cada voltaje aplicado dependerán del sentido de giro (izquierda o derecha) que se desee. La velocidad angular del robot al girar está dada como

$$w = \frac{v_R - v_L}{L} = 7.9426 \text{ rad/s.} \quad (2)$$

El ciclo de trabajo para cada rueda en ambos robots ha sido escalada para operar bajo las mismas velocidades. De esta manera, se define cada movimiento en términos de velocidades lineales y angulares constantes como se muestra en la Tabla I. El movimiento ejecutado se mantiene hasta recibir un comando diferente o el término de la sesión por parte del usuario.

Tabla I: Definición de los movimientos del robot móvil diferencial en términos de velocidades lineales y angulares.

Movimiento	Velocidad Lineal $v$ m/s	Velocidad angular $w$ rad/s
$\mathcal{A}^+$	0.5121	0
$\mathcal{A}^-$	-0.5121	0
$\mathcal{G}_d$	0	7.9426
$\mathcal{G}_i$	0	-7.9426
$\mathcal{D}$	0	0

Finalmente, la imitación del mismo movimiento en el segundo robot móvil se caracteriza idénticamente. Como se especificó previamente, el servidor del sistema de interacción se encarga del intercambio de mensajes para enviar el comando al robot imitador una vez que el robot principal confirma la recepción correcta del comando.

### III. VALIDACIÓN EXPERIMENTAL

La validación experimental del sistema interactivo se realiza a través de la ejecución de una secuencia de comandos de prueba y la medición de los desplazamientos de las ruedas de cada robot. Los experimentos han sido realizados en un espacio de dimensiones 4.2 [m]  $\times$  4.2 [m].



Figura 6: Escenario experimental para validar el sistema de interacción natural con robots móviles utilizando gestos de las manos.

Considere la siguiente nomenclatura bajo el escenario descrito en la Figura 1, y cuya implementación se muestra en la Figura 6.

**Robot Principal  $R_1$ .** Los comandos de control que recibe son  $v_1$  y  $\omega_1$  de acuerdo a la Tabla I, correspondientes a la

velocidad lineal y angular, respectivamente. Recibe en primera instancia un mensaje del servidor con el comando solicitado por el usuario y ejecuta el movimiento. A su vez, envía mensaje de confirmación al servidor de que ha sido recibido el comando.

**Robot Imitador  $R_2$ .** Este robot recibe el comando a ejecutar después de que el robot principal  $R_1$  ha enviado confirmación de que ha recibido el movimiento solicitado por el usuario.

Se implementan interrupciones programadas en cada robot con la finalidad de llevar un control del tiempo de ejecución y poder evaluar el desempeño del sistema. La medición de las distancias recorridas por cada rueda y la señal aplicada a cada motor ha sido implementada en una tarjeta Arduino Nano basada en el microcontrolador ATmega328. La configuración se describe a continuación.

- El sistema electrónico de cada robot móvil está compuesto de un sistema embebido Raspberry Pi 3 y una tarjeta Arduino Nano. La comunicación inalámbrica con el servidor es manejada por el Raspberry Pi 3 a través de Python; mientras que el manejo del tiempo de ejecución y mediciones es a través de interrupciones programadas en el Arduino Nano. Una comunicación USART permite a ambos sistemas trabajar coordinadamente intercambiando información de forma bidireccional.
- El tiempo de ejecución es contabilizado con un temporizador del microcontrolador, denominado `Timer2`. Se configura con interrupciones por sobreflujo, operación en modo PWM Rápido, con una pre-escala de 64. Con esto, se tiene un conteo de tiempo a una frecuencia de 976.5625Hz equivalente a un periodo de tiempo de 1.024 [ms] de resolución.
- Se implementa una señal de PWM para activar los motores trabajando a una frecuencia de 122.1896 Hz equivalente a un tiempo de 8.2 [ms]. Se habilita el temporizador `Timer1` en modo PWM Fase Correcta con una resolución de 10bits con una pre-escala de 64. El cambio de dirección de cada motor es a través de la habilitación de dos salidas del microcontrolador para realizar la configuración del voltaje aplicado por el puente H.
- Se programa un envío de datos hacia el servidor y la actualización del PWM para los motores, cada 300 muestras del contador del tiempo implementado por el `Timer2`. Considerando la resolución de éste, se envían inalámbricamente información de tiempo y mediciones de distancia a una tasa de 307.2 [ms].

Se realizaron dos experimentos de prueba con el sistema de interacción natural para verificar el movimiento de los robots móviles. En la Figura 7 se muestran las mediciones obtenidas para ambos robots para una secuencia de prueba definida por los comandos  $\mathcal{D}$ ,  $\mathcal{A}^+$ ,  $\mathcal{G}_i$ ,  $\mathcal{A}^+$ ,  $\mathcal{G}_d$ ,  $\mathcal{A}^-$ . Los resultados mostrados corresponden a los datos que envía cada robot al servidor en forma inalámbrica. Esta información está descrita

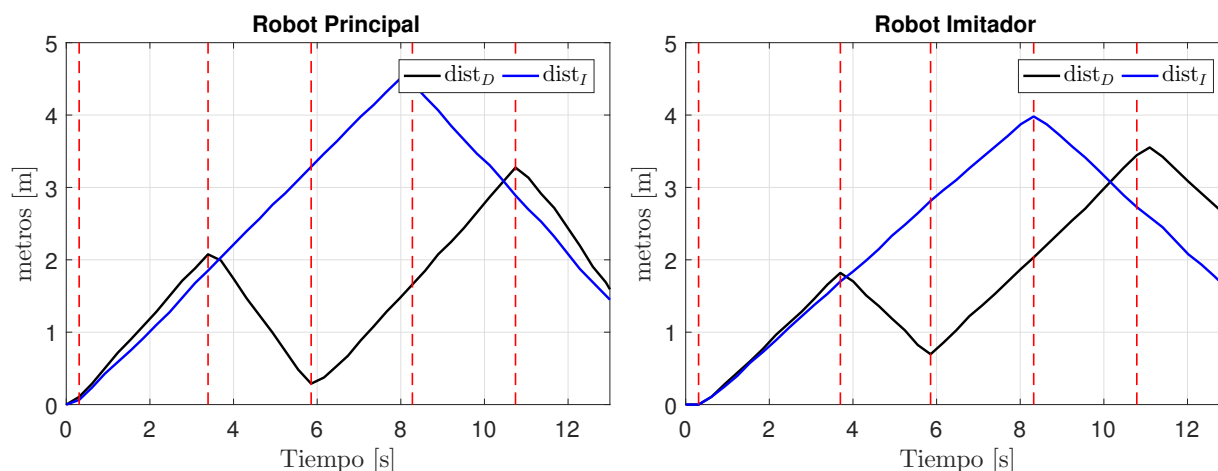


Figura 7: Resultados experimentales de las mediciones de las distancias recorridas por cada rueda (rueda derecha:  $dist_D$ , rueda izquierda:  $dist_I$ ). Izquierda: Robot principal  $R_1$ , Derecha: Robot imitador  $R_2$ . La secuencia de comandos implementada es  $\mathcal{D}$ ,  $\mathcal{A}^+$ ,  $\mathcal{G}_i$ ,  $\mathcal{A}^+$ ,  $\mathcal{G}_d$ ,  $\mathcal{A}^-$ .

por el formato

$$t^i, C^i, p_D^i, p_I^i$$

donde  $i = 1, 2$  es el índice que define a cada robot. La variable  $t^i$  es el tiempo calculado por el temporizador interno de cada robot,  $C^i$  es el comando recibido,  $p_D^i$  y  $p_I^i$  son los pulsos leídos de la rueda derecha e izquierda del  $i$ -ésimo robot, respectivamente.

De acuerdo a los resultados obtenidos, se observa que cada robot reacciona de acuerdo al movimiento recibido. Por ejemplo, para movimiento hacia adelante y atrás ambas ruedas avanzan en la misma dirección de acuerdo al sentido del movimiento solicitado; para movimiento de rotación, ambas ruedas se mueven en dirección contraria. Nótese que aunque los comandos y movimientos son cualitativamente similares, existe una diferencia atribuible al tiempo necesario para que el robot principal responda para enviar el mismo comando al robot imitador.

En la Figura 7 se muestran los resultados experimentales obtenidos en un ejercicio real de 40 segundos, en donde un usuario genera 23 comandos gestuales sobre el Leap Motion. Observe que las diferencias entre ambas mediciones se incrementa con el tiempo. Una de las razones corresponde a la diferencia temporal en el que el comando es recibido, ya que el robot imitador siempre tendrá un retraso con respecto al robot principal. Por otro lado, el sistema requiere reaccionar con mucha rapidez, y los robots móviles exhiben problemas de patinado y deslizamiento sobre la superficie.

#### IV. CONCLUSIONES

Se ha presentado el diseño de un sistema interactivo que permite a un usuario enfocarse en la organización de procesos cognitivos, que a través de gestos de las manos, guía el movimiento de un arreglo de robots móviles diferenciales. Cinco

gestos de las manos han sido establecidos para implementar movimientos básicos en los robots: avanzar, retroceder, girar hacia izquierda, girar hacia la derecha, y detenerse.

El diseño propuesto considera los parámetros físicos y la cinemática de los robots móviles diferenciales utilizados. Con esto se logran velocidades lineales y angulares constantes apropiadas a los gestos establecidos como comandos de movimiento. Se diseñaron temporizadores para garantizar una mejor ejecución y respuesta del sistema ante un envío constante de comandos. La rapidez de la secuencia enviada varía entre la velocidad propia del usuario para realizar los comandos gestuales en el Leap Motion.

Los resultados experimentales de una sesión de prueba demuestran la eficiencia del sistema propuesto y la ejecución correcta de los movimientos solicitados por el usuario. Existe un desfase entre los movimientos de cada robot debido a los retardos que ocurren en el sistema de mensajes, ya que es necesario que el robot principal confirme la recepción del comando, para que éste sea enviado al robot imitador. El funcionamiento del sistema es intuitivo y permite que el usuario se enganche en tareas de coordinación del movimiento de la mano para generar trayectorias deseadas. Como trabajo futuro, se pretende el rediseño de las capacidades físicas de los robots móviles para implementar controladores de velocidad. Asimismo, se considera la implementación de un mecanismo de evaluación del estrés que permita cuantificar la carga de trabajo cognitiva en el usuario.

#### AGRADECIMIENTOS

Karina Reyes-Lío, Gustavo Castellón-Ramírez y Cristina Ramírez-Fernández, desean agradecer el apoyo del Tecnológico Nacional de México y al Instituto Tecnológico de Ensenada a través del proyecto 6798.19-P.

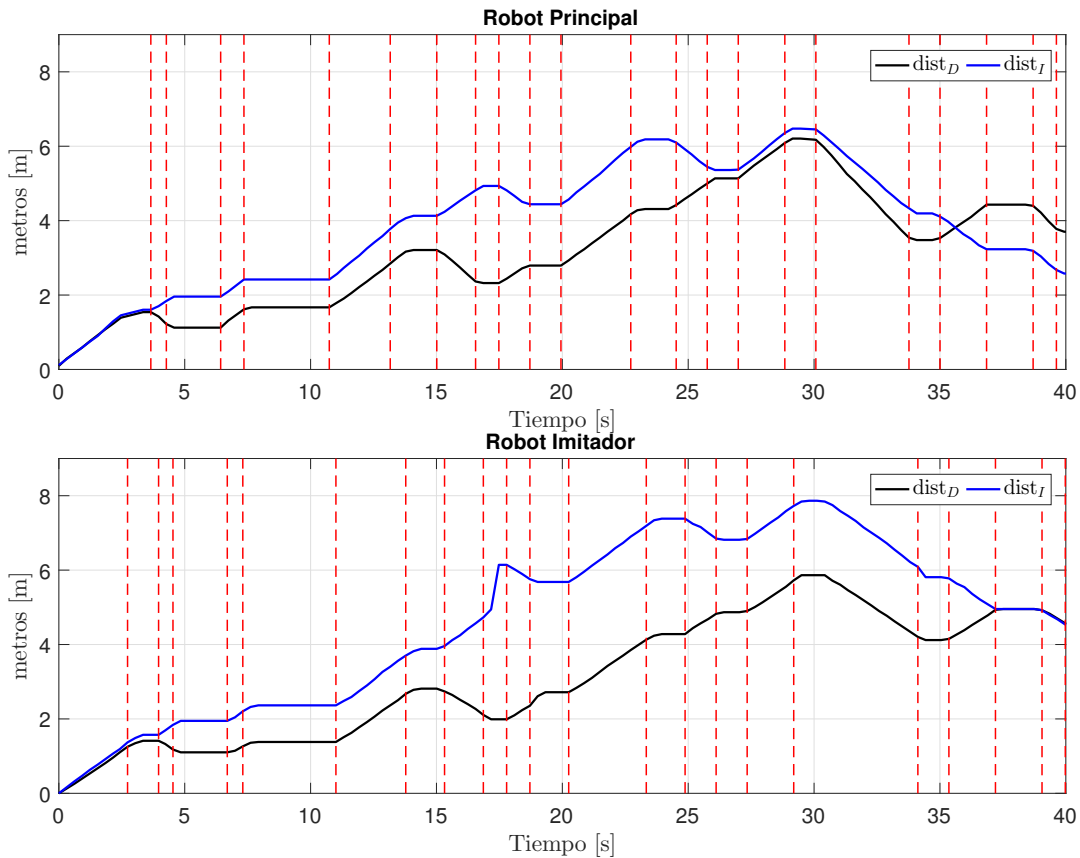


Figura 8: Resultados experimentales de un ejercicio de 40 segundos de duración en donde el usuario genera 23 comandos gestuales. La secuencia de comandos es  $\mathcal{A}^+$ ,  $\mathcal{D}$ ,  $\mathcal{G}_i$ ,  $\mathcal{D}$ ,  $\mathcal{A}^+$ ,  $\mathcal{D}$ ,  $\mathcal{A}^+$ ,  $\mathcal{D}$ ,  $\mathcal{G}_i$ ,  $\mathcal{D}$ ,  $\mathcal{G}_d$ ,  $\mathcal{D}$ ,  $\mathcal{A}^+$ ,  $\mathcal{D}$ ,  $\mathcal{G}_d$ ,  $\mathcal{D}$ ,  $\mathcal{A}^+$ ,  $\mathcal{A}^-$ ,  $\mathcal{D}$ ,  $\mathcal{G}_d$ ,  $\mathcal{D}$ ,  $\mathcal{A}^-$ ,  $\mathcal{D}$ .

#### REFERENCIAS

- [1] B. Ghannadi, R. S. Razavian, and J. McPhee, "Chapter nine - upper extremity rehabilitation robots: A survey," in *Handbook of Biomechanics*, J. Segil, Ed. Academic Press, 2019, pp. 319 – 353.
- [2] M. Bonikowski, M. Hajduk, L. Koukolova, W. J. Klimasara, P. Mrozek, and Z. Pilat, "New trends in the use of robotic devices in motor rehabilitation of upper limbs," in *Challenges in Automation, Robotics and Measurement Techniques*, R. Szweczyk, C. Zieliński, and M. Kaliczyńska, Eds. Cham: Springer International Publishing, 2016, pp. 729–739.
- [3] V. Villani, L. Sabbatini, C. Secchi, and C. Fantuzzi, "Natural interaction based on affective robotics for multi-robot systems," in *2017 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*, Dec 2017, pp. 56–62.
- [4] A. Kolling, P. Walker, N. Chakraborty, K. Sycara, and M. Lewis, "Human interaction with robot swarms: A survey," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 1, pp. 9–26, Feb 2016.
- [5] S. Heins, S. Dehem, V. Montedoro, B. Dehez, M. Edwards, G. Stoquart, F. Rocca, P. H. De Deken, M. Mancas, and T. Lejeune, "Robotic-assisted serious game for motor and cognitive post-stroke rehabilitation," in *2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH)*, April 2017, pp. 1–8.
- [6] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and Autonomous Systems*, vol. 42, no. 3, pp. 143 – 166, 2003, socially Interactive Robots.
- [7] A. A. Aziz, F. F. M. Moganan, M. Mokhsin, N. Sakamat, and A. Ismail, "Humanoid robot intervention with autism: The conceptual model in demonstrating the emotional responses of children with autism," in *Proceedings of the 7th International Conference on Kansei Engineering and Emotion Research 2018*, A. M. Lokman, T. Yamanaka, P. Lévy, K. Chen, and S. Koyama, Eds. Singapore: Springer Singapore, 2018, pp. 574–579.
- [8] D. Mondou, A. Prigent, and A. Revel, "A dynamic scenario by remote supervision: A serious game in the museum with a nao robot," in *Advances in Computer Entertainment Technology*, A. D. Cheok, M. Inami, and T. Romão, Eds. Cham: Springer International Publishing, 2018, pp. 103–116.
- [9] J. L. Raheja, R. Shyam, U. Kumar, and P. B. Prasad, "Real-time robotic hand control using hand gestures," in *2010 Second International Conference on Machine Learning and Computing*, Feb 2010, pp. 12–16.
- [10] S. Chen, H. Ma, C. Yang, and M. Fu, "Hand gesture based robot control system using leap motion," in *Intelligent Robotics and Applications*, H. Liu, N. Kubota, X. Zhu, R. Dillmann, and D. Zhou, Eds. Cham: Springer International Publishing, 2015, pp. 581–591.
- [11] X.-H. Wu, M.-C. Su, and P.-C. Wang, "A hand-gesture-based control interface for a car-robot," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2010, pp. 4644–4648.
- [12] A. M. Faudzi, M. H. K. Ali, M. A. Azman, and Z. H. Ismail, "Real-time hand gestures system for mobile robots control," *Procedia Engineering*, vol. 41, pp. 798–804, 2012, international Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012).
- [13] W. F. B. D., R. B., and F. D., "Analysis of the accuracy and robustness of the leap motion controller," *Sensors (Basez, Switzerland)*, vol. 13, no. 5, 2013.
- [14] A. Sharma, A. Yadav, S. Srivastava, and R. Gupta, "Analysis of movement and gesture recognition using leap motion controller," *Procedia Computer Science*, vol. 132, pp. 551 – 556, 2018, international Conference on Computational Intelligence and Data Science.
- [15] LeapMotion. *Gestures*. [Online]. Available: [https://developer-archive.leapmotion.com/documentation/v2/csharp/devguide/Leap\\_Gestures.html](https://developer-archive.leapmotion.com/documentation/v2/csharp/devguide/Leap_Gestures.html)

# Fuzzy color description on Raspberry PI 3

E. A. Monroy-Sahade, D. Lázaro-Mata, E. A. Vázquez-Rodríguez  
A. I. Barranco-Gutiérrez and J. A. Padilla-Medina

Electronics Department  
Instituto Tecnológico de Celaya  
Antonio García Cubas 600, Col. FOVISSSTE, Celaya,  
Guanajuato, México  
israel.barranco@itcelaya.edu.mx

**Abstract**—A fuzzy system of color description written in Python 3 and implemented in Raspberry PI 3 is presented. The system has the purpose of analyzing and describing thirteen different colors from its three components of RGB tonalities because the digital cameras and screens use that information to produce different colors. The motivation of the work consists of describing the subjectivity of the tonality for applications in automatic color identification in the industry in order to classify color in cars, cables, clothes, plants and others. In this sense, fuzzy logic helps to mathematically model the color perception of the human eye and its brain. The model was implemented in Raspberry PI 3 to take advantage of its portability, low cost and easy programming on Python 3. The system was tested and validated with the survey made to a group of thirty people who were shown a random color and they had to associate it with a list of colors corresponding to which they seemed more from their point of view. This study has the peculiarity that it is implemented in Raspberry PI 3 and uses the video card on it. The most important limitation of this proposal is that the color palette consists of thirteen tonalities. The results show that the proposed system converges to the same decisions of the participants in a high degree of coincidences.

## I. INTRODUCTION

The use of embedded systems is getting more common each day because they can provide mobility and processing reliability, even if they are limited in processing capability. A very important case is the Raspberry PI 3 which is embedded system with Quad-Core Cortex A7 a 900MHZ processor, 1GB DDR2 in RAM, USB ports, HDMI output, ethernet connector, wifi communication and low power consumption and its own Linux distribution as you can see in Figure 1. The PI 3B+ model can consume up to 980mA (5.1W) when 400% is indicated, this figure is approximately 300mA more than the previous model but with an improvement in the processing time of up to 80%. It can be programmed directly in c or python languages. On the other hand, in the automation industry, artificial vision systems are currently in great demand; but even at this time it continues using monochromatic systems to detect a specific wavelength in the visible, infrared or ultra-violet spectrum. Besides, analyzing a monochromatic matrix is easier than interpreting three RGB matrices. This is where the project's relevance stands out. For example, in India there is a documented study about the recognition of tuberculosis bacilli with Digital Image Processing (DIP), where the key characteristics for the identification of this bacilli are color and

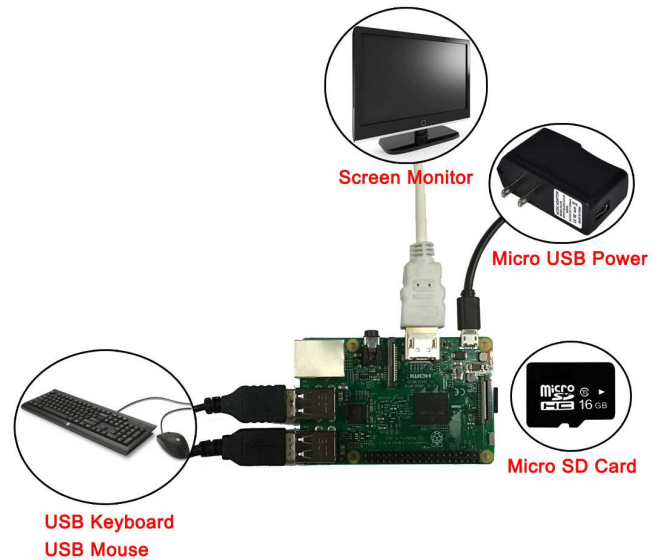


Fig. 1. Keyboard, mouse and screen connected to Raspberry PI

shape. The “regular” way to do it is by manual observation (data obtained unconsciously by the observer). However, in words of the authors “...diagnosis by manual observation is inefficient because of it is time consuming and efficiency depends upon skill of the pathologist and the quality of the smear”. So here is when the DIP comes in handy, using the Hue value they identify the color of the bacilli and with a shape recognition algorithm validates this information [1]. The propose of using a DIP is to obtain the same information we get from our eyes but in a digital system, an example of this is the edge recognition which is one of the basic features we obtain from our surroundings. To this day there are several ways to make an edge detection in an algorithm, the problem is that our brain processes the information so fast that is almost imperceptible, but for a digital system is a heavy process. For this reason, studies exist to enhance these algorithms [2]. In all the cases of DIP, the main objective is to simulate the human visual perception system. The gathering of the images is well advanced, the cameras available currently are capable of perceive most of the information the human eye can. Nevertheless, the processing of this information is

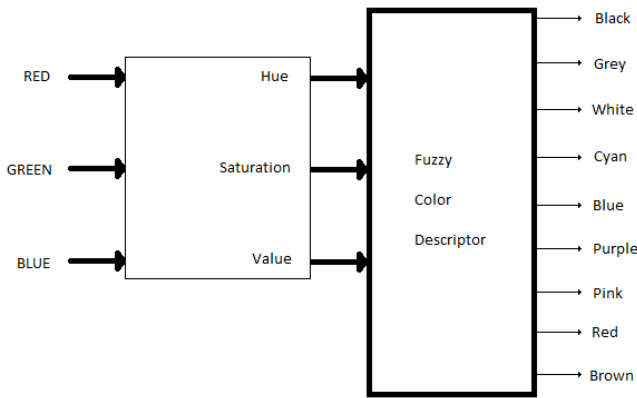


Fig. 2. Information flow

still far behind, the way the digital systems compute this data is still too absolute and not at all consequent i.e. is, relatively speaking, easy to a machine to give you the exact time but not so easy to tell you if it's early in the morning or late in the morning. There is where the fuzzy logic comes into play, this way one can provide the computer system with linguistic variables, convert them to computer language and process them, and convert them again to output linguistic variables [3],[4]. There is a study on fuzzy logic where the system must do the recognition of the color the way a human would do it. Here the authors characterize a palette of 12 colors in a non-uniform segmentation as an input to the fuzzy logic system, this palette was taken from the latest CSS3 specification; same that was used in this paper [5]. In this paper an algorithm, where is made an identification of colors percentages present in a RGB color given is presented. 2 Raspberry PI 3 where previously used in a cable color identifier for the automotive industry this because it has integrated protocols commonly used in the industry such as RS-232 serial communication and internet network protocols and also due to its compact size, connection of cameras and processing capacity superior to that obtained in other platforms such as Arduino or a microcontroller. In this project the colors painted on the cable were verified and it was ensured that a line painted along the cable maintained the tone and thickness of the required color. As the program is pretended to be mounted on a Raspberry Pi 3, the code programming was made in Python 3 language.

## II. SYSTEM DESCRIPTION

### A. Color model

Before the fuzzy logic can be applied to the system, it's necessary to adjust some of the data, since the information needed by the program at the input is the values of the components R, G and B (separately) of the color to identify, the first thing to do is to contrast these values (Figure 2). The normalization is made in order to make a conversion of the RGB color to HSV with values between 0 to 1. The conversion

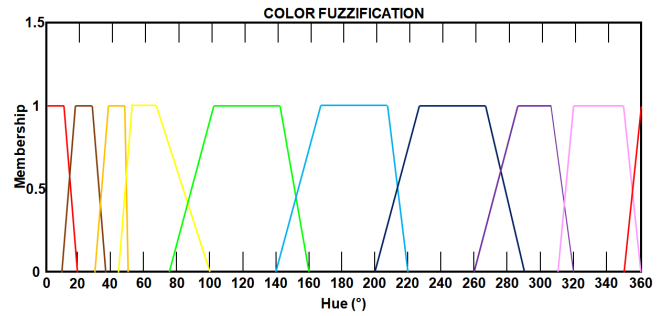


Fig. 3. Separation of colors according to its Hue angle.

was made according to Equations (1), (2) and (3) in order to substituting in Equations (4), (5) and (6) [6].

$$R_1 = \frac{R}{255}; G_1 = \frac{G}{255}; B_1 = \frac{B}{255} \quad (1)$$

$$V_{max} = \text{Max}(R_1, G_1, B_1) \quad (2)$$

$$V_{min} = \text{Min}(R_1, G_1, B_1) \quad (3)$$

$$H = \begin{cases} 0^\circ & \Delta = 0 \\ \left(\frac{G_1 - B_1}{\Delta}\right) * 60^\circ & V_{max} = R_1 \text{ and } G_1 \geq B_1 \\ \left(\frac{G_1 - B_1}{\Delta}\right) * 60^\circ + 360^\circ & V_{max} = R_1 \text{ and } G_1 < B_1 \\ \left(\frac{B_1 - R_1}{\Delta}\right) * 60^\circ + 120^\circ & V_{max} = G_1 \\ \left(\frac{B_1 - R_1}{\Delta}\right) * 60^\circ + 240^\circ & V_{max} = B_1 \end{cases} \quad (4)$$

$$S = \begin{cases} 0 & V_{max} = 0 \\ \frac{\Delta}{V_{max}} & V_{max} \neq 0 \end{cases} \quad (5)$$

$$V = V_{max} \quad (6)$$

where  $\Delta = V_{max} - V_{min}$  in the last equation. At this point the data is correctly formatted to be introduced into the fuzzification process, which is going to be explained in the next sub-section.

### B. Fuzzification

In the last sub-section, the input data introduced in a RGB format was converted to its HSV format to prepare it for the fuzzification, in the Figure 3 is shown the figures and separations considered in the Hue angle. The colors went from red (who shares the last and first angles in the group) to brown, orange, yellow, green, cyan, blue, purple and pink. According to the Figure 3 the tonality depends only of the Hue component of the color, which is not entirely true. For this reason, where introduced two more inputs that was dependent of the Saturation (Figure 4) and the Value (Figure 5), respectively. For the Saturation and Value components, a consideration was made about when does a color starts looking too dark or too bright and start losing its coloration. All the inputs where considered linear equation as there is not much worth gain in comparison with inputs shaped as a Gauss bell [7].

The consideration of the black, white and gray colors was made because of the nature of the HSV representation where



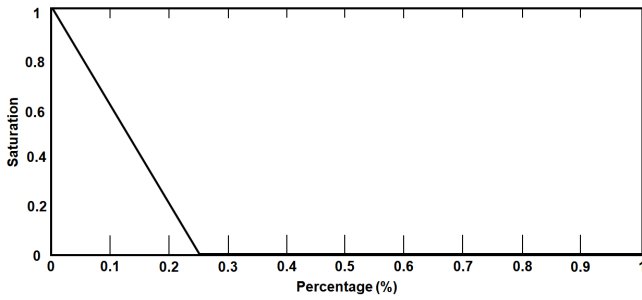


Fig. 4. Fuzzification of the saturation percentage

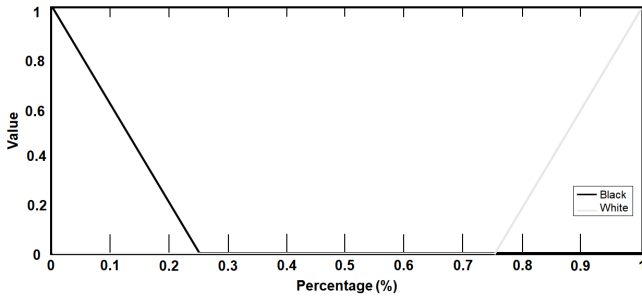


Fig. 5. Fuzzification of the value component in percentage.

levels too low of saturation means that all the colouration is gone (becomes gray). The same way goes for the value parameter, where levels too low represents that the color is too dark and is considered as black, and levels too big means that the color is so bright that can be considered as white.

C. Knowledge Base and Defuzzification

In this particular system the defuzzification is given directly by the output of the knowledge base. According to the Figure 3 the Hue angle, obtained from the conversion of the RGB parameters, is introduced in array of mathematical expressions describing the behavior of the tonality across the angle variation. As the input graph is normalized in the membership axis, the data taken from the knowledge base is directly defuzzified, the only change made to it was a multiplication by one hundred but only for printing proposes. Same process was used for the gray, black and white colors, with the considerations made in the last sub-section the percentage of membership was calculated.

III. RESULTS

In this section, we present the data that was introduced to the system, and the responses of the individuals participating in the survey answered by a group of thirty people, the data obtained are shown in Table I. In the survey process, the participants were shown on the left of the screen a rectangle filled with a color in each question and asked to determine how closely it resembled a color palette in the right (see Figure 6). The colors used in the survey were subsequently introduced into the algorithm to compare the results on each side. The limit



Fig. 6. The color palette used in the poll.

TABLE I  
POLL COMPARISON BETWEEN PARTICIPANTS AND FUZZY SYSTEM.

Participant	Fuzzy system values												Hits	Hits%		
	5	8	11	8	12	12	4	5	7	10	7	9			11	
1	5	8	11	8	12	12	4	5	7	10	7	9	2	11	84.615	
2	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
3	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
4	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
5	5	8	11	8	12	12	4	5	7	10	7	9	2	11	84.615	
6	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
7	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
8	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
9	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
10	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
11	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
12	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
13	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
14	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
15	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
16	5	8	11	10	12	12	4	5	7	10	7	9	2	11	84.615	
17	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
18	5	8	11	10	12	12	4	5	7	10	7	9	2	11	84.615	
19	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
20	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
21	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
22	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
23	5	8	11	8	12	12	4	5	7	10	7	9	2	11	84.615	
24	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
25	5	8	11	10	12	12	4	5	7	10	7	9	2	11	84.615	
26	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
27	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
28	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
29	5	8	11	8	12	12	4	5	7	10	7	9	2	11	84.615	
30	5	8	11	8	12	12	4	5	7	10	7	9	2	12	92.308	
															<b>Mean</b>	<b>90.513</b>

was left open to recover the most natural response from the participants. The group of people selected for the survey has an average age of between twenty-three and twenty-nine, they were men and women. The result of the poll shows that the fuzzy system provided answers similar to the contestants in 90% of the cases. By other hand, the time used to design the fuzzy system in Python 3 took around 10 days. While the time that the Raspberry PI 3 uses to deliver a result from the inputs of the fuzzy system is less than a second.

In the last color identification for the poll in Table I, an inconsistency is observed, because the proposed algorithm shows that the color found is yellow, when, all the participants



TABLE II  
HSV AND COLOR PERCENTAGES FOR THE LAST PRESENTED COLOR IN  
THE POLL.

<b>Hue:</b>	55.38 degree
<b>Saturation:</b>	14.61%
<b>Value:</b>	34.90%
Color percentages are:	
Red:	0.00%
Brown:	0.00%
Orange:	0.00%
Yellow:	100.00%
Green:	0.00%
Cyan:	0.00%
Blue:	0.00%
Purple:	0.00%
Pink:	0.00%
Gray:	41.57%
Black:	0.00%
White:	0.00%

chose it as a gray color. This is due to the fact that the real percentages of this color indicate 100% yellow and 41.57% gray, in addition, its saturation has a value of 14.61% and that is why the participants chose it as yellow, in place of its real color that is gray (see Table II).

#### IV. DISCUSSION

The results show that the classification of colors by a Fuzzy system implemented on Python 3 in Raspberry PI 3 is possible and the method has positive features such as, easy design, high speed execution and portability. Additionally, this system has a big possibility of interconnection using software (TCP/IP) or hardware (GPIOs). Regarding the accuracy of the results we can improve them like most aspects of this world. But more than 90% of correct answers demonstrate a very good quality of the results. Enlarging the color palette would be a good aspect to improve this system to classify common colors among people like “Mexican rose”, “military green” or others. The biggest problem in the results is the confusion between yellow and gray.

#### V. CONCLUSIONS

This document presented a fuzzy model implemented in Python code for color classification in a Raspberry PI 3. The results shows the scope and limits of current technology to design, implement and use classifiers in low cost and highly connected portable systems via software or hardware. In other investigations, designs of fuzzy systems for color classification have been shown, but they do not show the implementation and testing stages. Due to the flexibility of the Python language and hardware, it is possible to make changes in the system quickly. Also, the runtime of the diffuse system shown in this work is small, less than one second. The HSV model is a useful tool to describe colors, although not the only one. It is desirable to increase the number of colors to be classified because the range of colors used in everyday life is very wide. The results of experimental precision were compared with a survey made to a group of people, and showed that the algorithm coincided with the participants 90% of the time.

#### VI. DATA AVAILABILITY

Python program and the poll results from Matlab related to this article can be found at <https://drive.google.com/drive/folders/19qauOxbv7E3dsBge7dV39DSzZ7b2XF1r?usp=sharing>

#### ACKNOWLEDGMENT

The authors would like to thanks to Consejo Nacional de Ciencia y Tecnología for the scholarships granted to Einar Monroy, David Lázaro, and Edgar Vázquez Rodríguez.

#### REFERENCES

- [1] Y. Payasi and S. Patidar, *Diagnosis and counting of tuberculosis bacilli using digital image processing*, IEEE International Conference in Information, Communication, Instrumentation and Control (ICICIC), 2017, pp. 1–5.
- [2] A. Sharma, M. D. Ansari, and R. Kumar, *A comparative study of edge detectors in digital image processing*, 4th IEEE International Conference in Signal Processing, Computing and Control (ISPCC), 2017, pp. 246–250.
- [3] L. A. Cruz-Camacho, R. Miranda-Cervantes, F. D. Ruiz-Huitron, J. A. Téllez-Pérez and A. I. Barranco-Gutiérrez, *Implementación de sistema difuso mandani de salidas triangulares en arduino*, Congreso Internacional de Robótica y Computación, Los Cabos, México 2017.
- [4] A. Cárdenas-León, A. I. Barranco-Gutiérrez, *Implementación de sistema difuso en arduino uno*, Academia Journals, Celaya, 2016.
- [5] J. C. Amante and J. M. Fonseca, “Fuzzy color space segmentation to identify the same dominant colors as users.”, in DMS, 2012, pp. 48–53.
- [6] M. S. Hitam, E. A. Awalludin, W. N. J. H. W. Yussof, and Z. Bachok, “Mixture contrast limited adaptive histogram equalization for underwater image enhancement,” in Computer Applications Technology (ICCAT), 2013 International Conference on. IEEE, 2013, pp. 1–5.
- [7] R. Benavente, M. Vanrell, and R. Baldrich, “Parametric fuzzy sets for automatic color naming,” *JOSA A*, vol. 25, no. 10, pp. 2582–2593, 2008.

# COEFICIENTE EN $\mathcal{F}_{\mathbb{R}}$ BASADO EN SISTEMA DIFUSO PARA ECUACIONES DIFERENCIALES DIFUSAS

Y. Gorrin–Ortega, N. R. Cazarez–Castro, S. L. Cardenas–Maciel  
 Tecnológico Nacional de México–  
 Instituto Tecnológico de Tijuana  
 CP 22414, Tijuana, B.C., México.  
 Email: {nohe,lilletecardenas}@ieee.org,  
 yuney.gorrin17@tectijuana.edu.mx

J. A. Lopez–Renteria  
 CONACyT–Tecnológico Nacional de México–  
 Instituto Tecnológico de Tijuana  
 CP 22414, Tijuana, B.C., México.  
 Email: jorge.lopez@tectijuana.edu.mx

**Resumen**—En este artículo se presenta una metodología para construir coeficientes difusos para modelos matemáticos. Dichos coeficientes se construyen a partir de un Sistema Difuso (SD) previamente diseñado por un experto en el fenómeno de estudio. Se presentan los métodos propuestos para la obtención de los coeficientes difusos a partir del SD. Al introducir los coeficientes difusos en un modelo de Ecuaciones Diferenciales (ED), este se convierte en Ecuaciones Diferenciales Difusas (EDD). Se presentan los principales teoremas de solución de EDD, así como un caso de estudio que presenta el aporte de construcción de coeficiente difuso y la solución de una EDD con componentes difusas (coeficiente, variable y condición inicial).

## I. INTRODUCCIÓN

Innumerables son las aplicaciones de las ED en todas las ramas de las ciencias. Desde las ciencias sociales, biología, química, física, por poner algunos ejemplos, hasta resolver problemas de la misma matemática pura. Desde hace algunas décadas han sido notables los avances en la Inteligencia Artificial (IA) y pilares fundamentales en la resolución de problemas de variada naturaleza. La *Soft Computing* [1] como rama de la IA trata varios temas donde se maneja información con incertidumbre. En la Lógica Clásica un elemento pertenece o no pertenece a un conjunto mientras que en la Lógica Difusa (LD) siempre pertenece pero con cierto grado de incertidumbre [2]. De aquí la importante aplicación de los SD en la representación de conocimiento basados en reglas para resolver problemas de clasificación, predicción, entre otros.

Las EDD surgen por la necesidad de modelar y resolver problemas en donde las condiciones iniciales de la ecuación diferencial presentan incertidumbre y no se conoce el valor exacto [3]. Diversos son los trabajos realizados en la rama de las EDD desde la teoría matemática como por ejemplo en topología difusa [4], [5], en métodos numéricos para solucionarlas [6], [7], [8], hasta aplicaciones en otras ciencias como biología e ingeniería. En varios artículos se ha planteado y solucionados el modelo de Malthus, como por ejemplo [9], mientras que el control automático ha sido otro campo de aplicación [10], [11].

Aunque los trabajos reportados en la bibliografía sobre el tema consideran la ED como difusa debido a la condición inicial de la ED, también se puede convertir en difusa de otras maneras como por ejemplo considerando difuso los coeficientes, las variables del sistema y/o los operadores.

En este artículo se propone una metodología que consiste en la construcción de coeficientes difusos para modelos representados como ED, a partir de un SD. Teniendo un SD de forma tal que las salidas sean coeficientes de algún modelo, obtenemos estos coeficientes, pero en vez de números reales devueltos por el método de defusificación, serán números difusos y de esta manera dicho modelo será difuso también. El resto del artículo está organizado como sigue: en la Sección II se plantea la notación básica, en la Sección III se hace el planteamiento del problema, la Sección IV trata sobre los preliminares, en la Sección V se presenta la forma de determinar los coeficientes difusos para modelos de EDD, en la Sección VI el caso de estudio y la Sección VII las conclusiones.

## II. NOTACIÓN BÁSICA

Se presenta la siguiente notación:

- $\mathbb{R}$  – Conjunto de los números reales
- $\mathbb{R}^+$  – Conjunto de los números reales positivos
- $\mathcal{F}$  – Conjunto de los números difusos
- $\mathcal{F}_{\mathbb{R}}$  – Conjunto de los números difusos con campo real
- $C^k$  – Derivadas continuas hasta orden  $k$
- $[u]^\alpha$  – Representación del número difuso  $u$  en sus  $\alpha$  niveles

## III. PLANTEAMIENTO DEL PROBLEMA

Considérese el modelo:

$$x'_i = f_i(x_i, p_k, t), \quad (1)$$

que corresponde a una EDD, donde:  $i, k \in [1, \dots, n]$  representan los valores que toman los índices  $i$  y  $k$  respectivamente,

$x_i \in \mathcal{F}_{\mathbb{R}}$  la variable de la ecuación pertenece a los números difusos con campo real,  $p_k \in \mathcal{F}_{\mathbb{R}}$  los coeficientes de la ecuación pertenecen también a los números difusos con campo real,  $t \in \mathbb{R}^+$  se refiere al tiempo,  $f_i \in C^k$  se pide que la función  $f$  sea lo suficientemente suave.

El problema consiste en: Dado un modelo matemático expresado como una ED, diseñar con base en conocimiento experto los números difusos tal que caractericen a los coeficientes del modelo de acuerdo con el contexto del fenómeno de estudio.

#### IV. PRELIMINARES

En esta sección se presentan definiciones y teoremas que se utilizan en el resto del artículo y son necesarios para comprender la metodología que se propone.

##### IV-A. Conjuntos Difusos y Sistemas Difusos

Se presentan las definiciones de conjunto difuso y función de membresía que se usará en el resto del artículo:

**Definición 1.** Un conjunto difuso  $\tilde{A}$  es una tupla de dos elementos definida como  $\tilde{A} = \{(x, \mu_A(x)) | x \in \mathbb{R}, \mu_A(x) \in [0, 1]\}$ , donde  $\mu_A(x)$  es llamada la función de membresía del conjunto difuso  $A$  [12].

A continuación se define la función de membresía:

**Definición 2.** Sea  $X$  algún conjunto. La función de membresía  $\mu_A$  de un conjunto difuso no vacío  $A$  es una función:  $\mu_A : X \rightarrow [0, 1]$ . La función  $\mu_A$  es interpretada como el grado de pertenencia de cada elemento  $x$  al conjunto difuso  $A$ , para cada  $x \in X$  [12].

A partir de la Definición 2 se definen las funciones de membresía tipo triangular y gaussiana que se utilizarán en secciones posteriores.

**Definición 3.** [13] Se define como función de membresía triangular de un conjunto  $A$  a la siguiente expresión:

$$\mu_A(x, a, b, c) = \max(\min(\frac{x-a}{b-a}, \frac{c-x}{c-b}), 0), \quad (2)$$

donde los parámetros  $a, b$  y  $c$  cumplen  $a < b < c$  ( $b$  es el núcleo del triángulo).

La función de membresía gaussiana se define como sigue:

**Definición 4.** Se define como función de membresía gaussiana de un conjunto  $A$  a la siguiente expresión:

$$\mu_A(x, \sigma, c) = e^{-\frac{(x-c)^2}{2\sigma^2}}, \quad (3)$$

donde los parámetros  $\sigma$  y  $c$  son la desviación estándar y la media respectivamente.

La operación  $\alpha$  - corte de un conjunto difuso  $u$  denotada por  $[u]^\alpha$  se define como:

**Definición 5.** Sea  $u : \mathbb{R} \rightarrow [0, 1]$ , la función de membresía de un conjunto difuso sobre  $\mathbb{R}$ ,  $\alpha$  - corte o  $\alpha$  - nivel es definido como el conjunto  $[u]^\alpha = \{x \in \mathbb{R} : u(x) \geq \alpha\}$ , para cada  $0 < \alpha \leq 1$ . Entonces  $[u]^0 = \text{sop}\{x \in \mathbb{R} : u(x) > 0\}$ , donde  $\text{sop}$  denota el soporte del conjunto difuso [12].

Una de las propiedades que se usará adelante es la altura de un conjunto difuso y se define como:

**Definición 6.** Sea  $u : \mathbb{R} \rightarrow [0, 1]$ , la función de membresía de un conjunto difuso  $A$  sobre  $\mathbb{R}$ , se define como altura y se denota por  $h(A) = \text{sup}_{x \in A} u(x)$ . Donde  $\text{sup}$  es el supremo de un conjunto.

La operación defusificación sobre un conjunto difuso es quien permite convertir la información difusa en dura. Uno de los métodos de defusificación es la Media de los Máximos (MOM), la cual se define como:

**Definición 7.** [14] MOM se define como:

$$v = \frac{\sum_{x_i \in M} (x_i)}{|M|}, \quad (4)$$

donde:  $M = h(A)$  y  $|M|$  se refiere a la cardinalidad de  $M$ .

Un SD de manera general es una función multivaluada donde a las entradas le corresponden salidas y la función que hace dicha relación se basa en reglas [15].

Un SD tipo Mamdani se define como:

**Definición 8.** Un Sistema de Inferencia Difuso tipo Mamdani es una terna  $I = (E, R, S)$  donde:

- $E$  es el conjunto de variables lingüísticas de la entrada.
- $R$  conjunto de reglas que relacionan la entrada con la salida.
- $S$  es el conjunto de variables lingüísticas de la salida.

Las variables lingüísticas están constituidas por etiquetas lingüísticas que permiten asociar un significado lingüístico a subintervalos en el dominio de cada variable. Una variable puede tener una o más etiquetas lingüísticas, matemáticamente estas se representan a través de funciones de membresía y en consecuencia por conjuntos difusos.

La relación de las variables lingüísticas de entrada y salida se hace a través de reglas Si - Entonces, en este artículo se asume que el SD tiene múltiples entradas y una salida. Por ello, las reglas tienen la forma:

$$R_i : \text{Si } x_1 \text{ es } M_{p1}, x_2 \text{ es } M_{p2}, x_n \text{ es } M_{pk}, \text{ Entonces } y_m \text{ es } N_{lj}, \quad (5)$$

donde  $R_i \in R$  representa las reglas, las  $x_n \in E$  son las variables lingüísticas de la entrada,  $M_{pk}$  son las etiquetas de las variables de entrada,  $y_m \in S$  son las variables de salida y  $N_{lj}$  sus respectivas etiquetas lingüísticas.

El proceso de inferencia que realiza el SD consiste en tomar el mínimo de los conjuntos difusos formados por las activaciones de las reglas y el máximo al operar los conjuntos que están a la salida de las reglas formando un conjunto difuso que se somete a la etapa de defusificación la cual retorna un solo número real.

Existen varios métodos de defusificación, la mayoría se apoya en el área del conjunto difuso de salida, los más usados son: centro de área, bisectriz de área, máximo de los más grandes, media de los máximos (MOM) y máximo más pequeño [16]. Para este trabajo se usa el método MOM definido en (7), debido a que este método es de tendencia

central y por tanto se evita coincidencias con los extremos siendo esto ventajoso para la aplicación.

#### IV-B. Ecuaciones Diferenciales Difusas

En esta sección se presentan definiciones para generalizar el concepto de derivada y tener los elementos para resolver una EDD. Se presenta la definición de número difuso para luego definir la derivada.

**Definición 9.** [17] Sea  $u$  la función de membresía de algún conjunto difuso. se dice que  $u$  es un número difuso si  $u : \mathbb{R} \rightarrow [0, 1]$ , y satisface las siguientes condiciones:

- i)  $u$  es normal; existe al menos un  $x^* \in \mathbb{R}$  tal que  $u(x^*) = 1$ ,
- ii)  $[u]^\alpha$  es cerrado  $\forall \alpha \in (0, 1]$ , y
- iii)  $[u]^0$  es acotado.

La definición de diferenciabilidad en el sentido difuso fue introducida en [18] como:

**Definición 10.** Sean  $u, v \in \mathcal{F}_{\mathbb{R}}$ . Si existe  $w \in \mathcal{F}_{\mathbb{R}}$  tal que  $u = v \oplus w$ , entonces  $w$  es llamado H-diferencia de  $u$  y  $v$ , este se denota por  $u \ominus v$ .

A continuación se plantea la H-diferencia en términos de límites, similar a la derivada clásica.

**Definición 11.** [18] Sea  $F : T \rightarrow \mathcal{F}_{\mathbb{R}}$  y  $t_0 \in T \subseteq \mathbb{R}$ . La función  $F$  se dice diferenciable sobre  $t_0$  si:

(I) Existe un elemento  $F'(t_0) \in \mathcal{F}_{\mathbb{R}}$  tal que, para todo  $h > 0$  es suficientemente cercano a cero, existe  $F(t_0 + h) \ominus F(t_0)$ ,  $F(t_0) \ominus F(t_0 - h)$  y los límites

$$\lim_{h \rightarrow 0^+} \frac{F(t_0 + h) \ominus F(t_0)}{h} = \lim_{h \rightarrow 0^+} \frac{F(t_0) \ominus F(t_0 - h)}{h}$$

son iguales a  $F'(t_0)$ , o

(II) Existe un elemento  $F'(t_0) \in \mathcal{F}_{\mathbb{R}}$  tal que, para todo  $h < 0$  es suficientemente cercano a cero, existe  $F(t_0 + h) \ominus F(t_0)$ ,  $F(t_0) \ominus F(t_0 - h)$  y los límites

$$\lim_{h \rightarrow 0^-} \frac{F(t_0 + h) \ominus F(t_0)}{h} = \lim_{h \rightarrow 0^-} \frac{F(t_0) \ominus F(t_0 - h)}{h}$$

son iguales a  $F'(t_0)$ .

Cabe señalar que si  $F$  es diferenciable en la primera forma (I), entonces no es diferenciable en la segunda forma (II), y viceversa, que se resume en el siguiente teorema:

**TEOREMA (1).** [17] Sea  $F : T \rightarrow \mathcal{F}_{\mathbb{R}}$ , y  $[F(t)]^\alpha = [F_L^\alpha(t), F_R^\alpha(t)]$ , para cada  $\alpha \in [0, 1]$ . Entonces:

(i) Si  $F$  es diferenciable de la primera forma (I) entonces  $F_L^\alpha(t)$  y  $F_R^\alpha(t)$  son funciones diferenciables y

$$[F'(t)]^\alpha = [(F_L^\alpha(t))', (F_R^\alpha(t))'], \quad (6)$$

o

(ii) Si  $F$  es diferenciable de la segunda forma (II) entonces  $F_L^\alpha(t)$  y  $F_R^\alpha(t)$  son funciones diferenciables y

$$[F'(t)]^\alpha = [(F_R^\alpha(t))', (F_L^\alpha(t))']. \quad (7)$$

Teniendo en consideración el resultado que se muestra en el Teorema 1, se enumeran los pasos para resolver una EDD:

- 1) Escribir la ED en forma de intervalos, Ejemplo: si  $x' \in \mathcal{F}_{\mathbb{R}}$  entonces se escribe  $[x'_L, x'_R]$ , si  $x \in \mathcal{F}_{\mathbb{R}}$  en intervalos es  $[x_L, x_R]$ ; donde el subíndice  $L$  y  $R$  hacen referencia al extremo del intervalo, izquierdo y derecho respectivamente.
- 2) Realizar las operaciones pertinentes del álgebra de intervalos para reducir a la forma  $[x'_L, x'_R] = [f(x), g(x)]$  como se muestra en [16].
- 3) Escribir la ecuación en intervalos a su forma extendida igualando los extremos de los intervalos. Considerando la forma I, aplicando la ecuación (6) se obtiene:

$$\begin{cases} x'_L = f(x) \\ x'_R = g(x), \end{cases} \quad (8)$$

o si se considera la forma II, aplicando la ecuación (7) queda:

$$\begin{cases} x'_L = g(x) \\ x'_R = f(x). \end{cases} \quad (9)$$

- 4) Resolver las ecuaciones anteriores por algún método analítico o numérico.

#### V. COEFICIENTES DIFUSOS PARA MODELOS DE EDD

En esta sección se presentan los métodos de construcción de los coeficientes difusos y el procedimiento para aplicar la metodología.

##### V-A. Métodos de construcción de los coeficientes

Para la construcción de los coeficientes  $p_k$  de la ED definida en (1), se asume que son números descritos a través de funciones de membresía triangulares y que satisfacen la Definición 9 para números difusos. La elección se debe a su simplicidad analítica y geométrica. Entonces, cada coeficiente es de la forma  $K = \{x, \mu_K(x) | x \in \mathbb{R}\}$ , donde  $\mu_K(x)$  es la función de membresía definida en (3).

Los métodos que se proponen para construir el coeficiente tienen como condición necesaria que se cuente con un SD. Considere el conjunto difuso de salida del SD denotado como  $Z$ . El primer método que se muestra en la Figura 1 consiste en tomar el valor mínimo y máximo del soporte del conjunto  $Z$  como el primer y tercer parámetro para la función de membresía triangular ( $a$  y  $c$ , respectivamente) y el parámetro  $b$  se toma de lo que resulte del proceso de defusificación del SD (usando (4)).

En el segundo método, la función de membresía del conjunto difuso se le asigna el resultado del proceso de defusificación al parámetro  $b$ , el parámetro  $a$  y  $c$  se obtienen al agregar una cantidad  $\epsilon_i$  definida por el usuario. Es decir, considerando la Figura 2 que ilustra el método  $a = b - \epsilon_1 \geq \min(\text{sop})$  y  $c = b + \epsilon_2 \leq \max(\text{sop})$ .

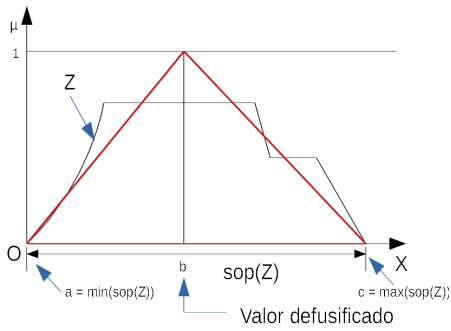


Figura 1. Primer método para construcción de coeficiente difuso.

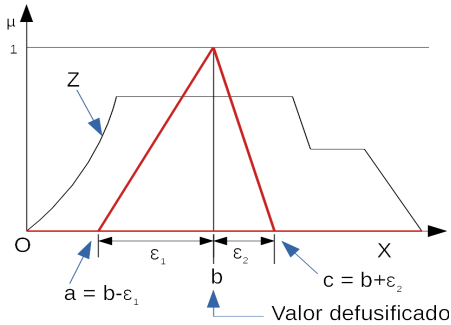


Figura 2. Segundo método para construcción de coeficiente difuso.

V-B. Procedimiento para aplicar la metodología

La metodología que se propone se resume a continuación.

Dado un SD que captura el conocimiento experto de las variables que representa el comportamiento de cada coeficiente de la ecuación diferencial, se puede obtener un número difuso caracterizado por una función de membresía triangular con los parámetros  $a, b$  y  $c$  de (2), usando el siguiente procedimiento:

- 1) Introducir valores a cada variable de entrada del SD y realizar proceso de inferencia y defuzificación este último usando (4).
- 2) Usar el conjunto difuso de salida y el resultado de la defuzificación que se obtienen del paso anterior para aplicar el primer o segundo método que se describe en esta sección.
- 3) Incorporar los coeficientes en la ecuación diferencial definida en (1) y aplicar el método de solución de la ecuación diferencial definiendo la condición inicial y usando la metodología descrita en la sección IV-B o bien, haciendo el análisis de su dinámica como muestran en [19] y [20].

VI. CASO DE ESTUDIO

En esta sección se ejemplifica la metodología antes descrita. El modelo a tratar es el Malthusiano descrito como

$$\begin{cases} x' = kx \\ x(0) = x_0; \quad x, k, x_0 \in \mathcal{F}_{\mathbb{R}}. \end{cases} \quad (10)$$

que representa el crecimiento poblacional donde todas sus componentes son números difusos.

Se aplica la metodología descrita en la sección V-B para crear el número difuso para  $k$  que representa la tasa de crecimiento de la población. Considerando la información proporcionada por el experto quien expresa que la tasa de crecimiento de la población esta relacionada con el Nacimiento, Energía y Muerte, se diseña un SD tipo Mamdani que relaciona estas variables.

Así, el SD consiste de tres variables lingüísticas de entrada que se nombran Tasa de Nacimiento (TN), Tasa de Muerte (TM) y Energía (E), una variable de salida que se nombra Índice de Crecimiento. (IC). Cada variable de entrada tiene tres etiquetas lingüísticas (alto, medio, bajo) que están descritas por funciones de membresía gaussianas. La variable de salida tiene tres etiquetas lingüísticas (crecimiento, constante, decaimiento) que están descritas a través de funciones de membresía gaussianas. Los parámetros de las funciones de membresía se presentan en la Tabla I, la forma de las funciones de membresía para la variables de entrada se muestran en la Figura 3, mientras que en la Figura 4 las de la variable de salida.

Tabla I  
PARÁMETROS DE LAS FUNCIONES DE MEMBRÉSIA.

Tipo de variable	Etiqueta	$\sigma$	Media
Entradas	bajo	1.7	0
	medio	1.7	5
	alto	1.7	10
Salida	bajo	1.7	0
	medio	1.7	5
	alto	1.7	10

Las reglas difusas cumplen con la estructura definida en (5), las 9 reglas se proporcionan en la Tabla II. Por ejemplo el primer renglón en la Tabla I significa Si TN es Alto , TM es bajo, E es alto entonces IC es crecimiento.

Tabla II  
REGLAS DEL SD.

TN	TM	E	IC
alto	bajo	alto	crecimiento
alto	bajo	medio	crecimiento
alto	medio	medio	crecimiento
bajo	alto	bajo	decaimiento
bajo	alto	medio	decaimiento
bajo	alto	alto	decaimiento
medio	alto	bajo	decaimiento
medio	medio	medio	constante
medio	medio	alto	constante

La construcción de coeficiente para el caso de estudio fue programado en *MatLab*<sup>®</sup>. Se utilizó la *ToolBox Fuzzy* para implementar el SD y los métodos de construcción de coeficientes se codificaron mediante *scripts*. Se introdujeron al SD como entrada los valores para las variables de entrada  $TN = 8.5, TM = 4.3$  y  $E = 5$  obteniéndose el número

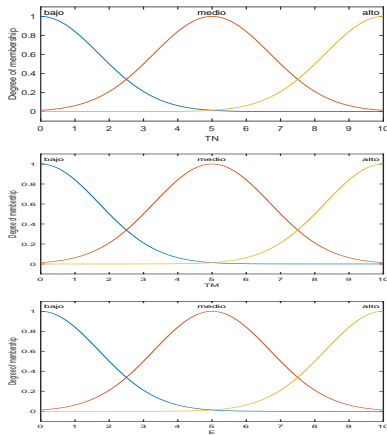


Figura 3. Funciones de membresía de la entrada.

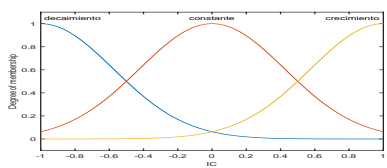


Figura 4. Función de membresía de la salida.

difuso para  $k$  con parámetros en función de membresía triangular  $a = 0.5$ ,  $b = 0.8$  y  $c = 1$ .

Al aplicar el primer paso de la metodología presentada en la sección IV-B para solucionar una EDD, la ecuación queda como:

$$[x'_L, x'_R] = [k_L, k_R] * [x_L, x_R]. \quad (11)$$

Suponiendo que  $x_L, x_R \geq 0$  por tratarse de un modelo biológico, se realizan las operaciones en intervalos que indica el paso 2 y se obtiene:

$$[x'_L, x'_R] = \begin{cases} [k_L * x_L, k_R * x_R]; & k_L, k_R \geq 0 \\ [k_L * x_R, k_R * x_L]; & k_L, k_R < 0 \\ [k_L * x_R, k_R * x_R]; & k_L < 0, k_R \geq 0. \end{cases} \quad (12)$$

Resolviendo la ecuación (12) de forma analítica considerando el número difuso  $k$  obtenido y la condición inicial  $x_0$  definida como el número difuso cuya función de membresía triangular tiene parámetros  $a = 1.7$ ,  $b = 2$  y  $c = 2.3$  se obtiene:

$$[x_L, x_R] = \left[ \frac{37}{20} e^{\frac{67}{100}t}, \frac{43}{20} e^{\frac{91}{100}t} \right] \quad (13)$$

$$[x_L, x_R] = \left[ \frac{37}{20} \cosh\left(\frac{\sqrt{6097}}{100}t\right) + \frac{43}{20} \sqrt{\frac{91}{67}} \sinh\left(\frac{\sqrt{6097}}{100}t\right), \frac{43}{20} \cosh\left(\frac{\sqrt{6097}}{100}t\right) + \frac{37}{20} \sqrt{\frac{67}{91}} \sinh\left(\frac{\sqrt{6097}}{100}t\right) \right]. \quad (14)$$

Las soluciones de la EDD que se muestran en (13) y (14) se pueden observar en la Figura 5 y Figura 6.

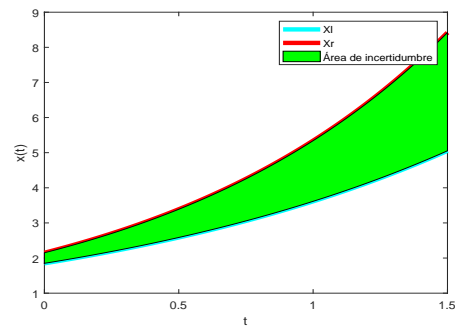


Figura 5. Solución Forma 1.

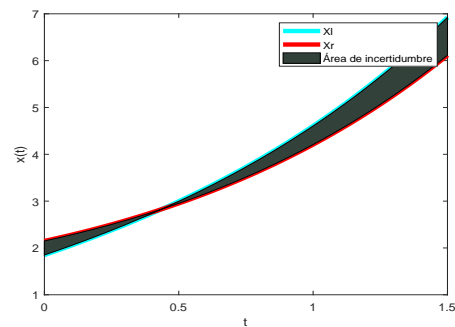


Figura 6. Solución Forma 2.

Nótese que las soluciones tienen un comportamiento creciente similar con un modelo de Malthus estándar con coeficiente positivo, y en este caso el coeficiente difuso se mantiene definido en un dominio de definición positivo. Para ambos casos las soluciones parten alrededor de la condición inicial, dicha condición inicial significa que no se conoce el valor exacto y existe cierta incertidumbre alrededor del núcleo del triángulo difuso.

## VII. CONCLUSIONES

En el artículo, una metodología para construcción de coeficientes de modelos matemáticos expresados en ecuaciones diferenciales fue presentado. La metodología fue probada para diseñar el coeficiente difuso que representa la tasa de crecimiento en el modelo de Malthus. La resolución del modelo de Malthus que utilizó el coeficiente diseñado y condición inicial ambos definidos como números difusos arrojó una familia de soluciones que preservan las trayectorias de crecimiento como las que se obtienen con un modelo en ecuaciones diferenciales considerando coeficiente real positivo.

Los resultados obtenidos muestran que la metodología propuesta para construir coeficientes con base en conocimiento experto representado a través de sistemas difusos podría ayudar a incorporar con mayor fluidez y facilidad las incertidumbres e inexactitudes presentes en los parámetros de los modelos matemáticos.



## ACKNOWLEDGMENT

Este proyecto fue parcialmente financiado por el Tecnológico Nacional de México con los proyectos 6351.17-P y 6827.18-P.

## REFERENCIAS

- [1] R. Patton, F. Uppal, and C. Lopez-Toribio, "Soft computing approaches to fault diagnosis for dynamic systems: a survey," in *4th IFAC Symposium on Fault Detection supervision and Safety for Technical Processes*. Citeseer, 2000, pp. 198–211.
- [2] E. Turunen and E. Turunen, *Mathematics behind fuzzy logic*. Physica-Verlag Heidelberg, 1999.
- [3] O. He and W. Yi, "On fuzzy differential equations," *Fuzzy sets and systems*, vol. 32, no. 3, pp. 321–325, 1989.
- [4] R. P. Agarwal, D. O'Regan, and V. Lakshmikantham, "A stacking theorem approach for fuzzy differential equations," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 55, no. 3, pp. 299–312, 2003.
- [5] S. Bayramov and C. Gunduz, "The cech homology theory in the category of  $s$ -ostak fuzzy topological spaces," *Int. J. Contemp. Math. Sciences*, vol. 5, no. 9, pp. 433–448, 2010.
- [6] O. Abu Arqub, M. AL-Smadi, S. Momani, and T. Hayat, "Numerical solutions of fuzzy differential equations using reproducing kernel hilbert space method," *Soft Computing*, vol. 20, no. 8, pp. 3283–3302, Aug 2016. [Online]. Available: <https://doi.org/10.1007/s00500-015-1707-4>
- [7] A. Ahmadian, S. Salahshour, C. S. Chan, and D. Baleanu, "Numerical solutions of fuzzy differential equations by an efficient runge–kutta method with generalized differentiability," *Fuzzy Sets and Systems*, vol. 331, pp. 47 – 67, 2018, theme: Partial Differential Equations, Fuzzy Numbers and Fractals. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0165011416304006>
- [8] R. G. Sharmila and E. H. Amirtharaj, "Numerical solution of n th-order fuzzy initial value problems by fourth order runge-kutta method based on centroidal mean," *IOSR Journal of Mathematics*, vol. 3, no. 3, pp. 47–63, 2013.
- [9] N. R. Cazarez-Castro, M. Odreman-Vera, S. L. Cardenas-Maciel, H. Echavarría-Heras, and C. Leal-Ramírez, "Fuzzy differential equations as a tool for teaching uncertainty in engineering and science," *Computación y Sistemas*, vol. 22, no. 2, 2018.
- [10] N. R. Cazarez-Castro, S. L. Cardenas-Maciel, M. Odreman-Vera, G. Valencia-Palomo, and C. Leal-Ramrez, "Modeling pd closed-loop control problems with fuzzy differential equations," *Automatika*, vol. 57, no. 4, pp. 960–967, 2016. [Online]. Available: <https://doi.org/10.7305/automatika.2017.12.1554>
- [11] R. Jafari, W. Yu, and X. Li, "Fuzzy differential equations for nonlinear system modeling with bernstein neural networks," *IEEE Access*, vol. 4, pp. 9428–9436, 2016.
- [12] L. A. Zadeh *et al.*, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338–353, 1965.
- [13] J. H. Lilly, *Fuzzy control and identification*. John Wiley & Sons, 2011.
- [14] Y. Bai and D. Wang, "Fundamentals of fuzzy logic control—fuzzy sets, fuzzy rules and defuzzifications," in *Advanced Fuzzy Logic Technologies in Industrial Applications*. Springer, 2006, pp. 17–36.
- [15] G. Chen and T. T. Pham, *Introduction to fuzzy systems*. Chapman and Hall/CRC, 2005.
- [16] —, *Introduction to fuzzy sets, fuzzy logic, and fuzzy control systems*. CRC press, 2000.
- [17] O. Kaleva, "Fuzzy differential equations," *Fuzzy Sets and Systems*, vol. 24, no. 3, pp. 301 – 317, 1987, fuzzy Numbers. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0165011487900297>
- [18] M. L. Puri and D. A. Ralescu, "Differentials of fuzzy functions," *Journal of Mathematical Analysis and Applications*, vol. 91, no. 2, pp. 552 – 558, 1983. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0022247X83901695>
- [19] M. Cecconello, R. Bassanezi, A. Brandão, and J. Leite, "On the stability of fuzzy dynamical systems," *Fuzzy Sets and Systems*, vol. 248, pp. 106 – 121, 2014, theme: Control Theory. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0165011413004958>
- [20] M. Mizukoshi and M. Cecconello, "Dynamics and bifurcations of fuzzy nonlinear dynamical systems," *Fuzzy Sets and Systems*, vol. 319, pp. 81 – 92, 2017, theme: Fuzzy Systems. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0165011416303992>

# Comparativo de técnicas de predicción en la evaluación de indicadores del TecNM

M. López Nava

Departamento de Sistemas y Computación  
TecNM/Instituto Tecnológico de Zacatepec  
Zacatepec, México  
misael.ln@itzacatepec.edu.mx

F. J. Cartujano Escobar

Departamento de Sistemas y Computación  
TecNM/Instituto Tecnológico de Zacatepec  
Zacatepec, México  
francisco.ce@itzacatepec.edu.mx

C. Noguerón González

Departamento de Sistemas y Computación  
TecNM/Instituto Tecnológico de Zacatepec  
Zacatepec, México  
claudia.ng@itzacatepec.edu.mx

A. L. Pineda Méndez

Facultad de Contaduría, Administración e Informática  
Universidad Autónoma del Estado de Morelos  
Cuernavaca, México  
alinda@uaem.mx

**Abstract** — La medición de los indicadores del Tecnológico Nacional de México, propuestos para el periodo 2013-2018, es de gran relevancia para la toma de decisiones, ya que determinan el grado de crecimiento y consolidación de cada una de las instituciones que lo conforman. Utilizando minería de datos, en este artículo se analizan estos indicadores para predecir su comportamiento durante periodos específicos. Estas predicciones se realizan con árboles de decisión y regresión lineal y se determina su grado de confianza. Las herramientas utilizadas para el minado de los datos son Weka® y RapidMiner®, y con ellas se pronostican los valores 2018 de los indicadores a partir de los datos anuales correspondientes al periodo 2012-2017.

**Keywords**—*minería de datos; árboles de decisión; regresión lineal; indicadores; TecNM.*

## I. INTRODUCCIÓN

El TecNM es la institución de educación superior tecnológica más grande Iberoamérica, está constituido por 254 instituciones, todas ellas comparten objetivos e indicadores. Los objetivos están definidos en el Programa Institucional de Innovación y Desarrollo (PIID) 2013-2018 del TecNM, los cuales están orientados hacia el cumplimiento de los objetivos del Programa Sectorial de Educación 2013-2018, documento emanado de la Secretaría de Educación Pública [1]. De estos se derivan 22 indicadores, los cuales se describen en la Tabla I [2].

Aunque, el periodo para el que fueron establecidos los indicadores es el año 2018, a la fecha de redacción del presente artículo, dichos indicadores aún no se han capturado por parte de cada una las instituciones, lo que permite establecer las condiciones para poder predecir los valores correspondientes al presente año, con base a la información de los años anteriores. Es también importante destacar que también se pueden obtener escenarios futuros, como puede ser el fin del sexenio siguiente (año 2024), sin importar que los indicadores actuales no sean contemplados en los documentos rectores derivados del Plan Nacional de Desarrollo (PND) 2019-2024.

Tabla I. Indicadores del TecNM ordenados por objetivo

No.	Indicador
1.1	Porcentaje de estudiantes de licenciatura inscritos en programas acreditados o reconocidos por su calidad
1.2	Porcentaje de profesores de tiempo completo con posgrado
1.3	Porcentaje de profesores de tiempo completo con reconocimiento del perfil deseable
1.4	Porcentaje de estudiantes inscritos en algún curso o programa de enseñanza de lenguas extranjeras
2.1	Matrícula del nivel licenciatura
2.2	Matrícula en posgrado
2.3	Matrícula en educación no escolarizada -a distancia- y mixta
2.4	Eficiencia terminal
3.1	Porcentaje de estudiantes que participan en actividades de extensión: artísticas, culturales y cívicas
3.2	Porcentaje de estudiantes que participan en actividades deportivas y recreativas
4.1	Porcentaje de programas de doctorado escolarizados en áreas de ciencia y tecnología registrados en el Programa Nacional de Posgrados de Calidad
4.2	Profesores de tiempo completo adscritos al Sistema Nacional de Investigadores
4.3	Proyectos de investigación, desarrollo tecnológico e innovación
4.4	Estudiantes de licenciatura y posgrado que participan en proyectos de investigación científica, desarrollo tecnológico e innovación
5.1	Registros de propiedad intelectual
5.2	Porcentaje de egresados incorporados al mercado laboral
5.3	Proyectos vinculados con los sectores público, social y privado
5.4	Estudiantes que participan en proyectos vinculados con los sectores público, social y privado
5.5	Empresas incubadas a través del modelo institucional de incubación empresarial
5.6	Estudiantes que participan en el Modelo Talento Emprendedor
6.1	Personal directivo y no docente capacitado
6.2	Institutos, unidades y centros certificados

Con adecuados modelos de pronóstico de indicadores, las instituciones educativas pueden abordar eficazmente los indicadores de trayectoria académica de los alumnos. Esta es una buena razón para justificar este estudio, ya que se enfocará al análisis de los indicadores de desempeño del Tecnológico Nacional de México (TecNM), debido a que muchas ocasiones las Instituciones de Educación Superior (IES) planean y pocas veces llegan a cumplir esos planes.

La minería de datos es una técnica de gran alcance que permite a las IES asignar mejor los recursos que se tienen, puede construir los modelos de pronósticos con un alto nivel de exactitud. La minería de datos utiliza una combinación de una base de conocimiento explícita, habilidades analíticas sofisticadas y conocimiento de dominio para descubrir tendencias y patrones ocultos. Estas tendencias y patrones forman la base de modelos predictivos que permiten a los analistas producir nuevas observaciones a partir de los datos existentes [3].

Con el uso de la minería de datos, una universidad podría, por ejemplo, predecir con una precisión del 85 por ciento los estudiantes que se graduarán y los que no lo harán. La universidad podría usar esta información para concentrar la asistencia académica en aquellos estudiantes con mayor riesgo [4]. Existen técnicamente dos tipos de técnicas de minería de datos, la descriptiva y la predictiva. El objetivo de la primera es encontrar patrones que resumen relaciones en los datos. El objetivo de la segunda es predecir el valor particular de un atributo basado en otros atributos. En este artículo se emplean las técnicas de árboles de decisión y regresión lineal que pertenecen a las técnicas predictivas.

La técnica de árboles de decisión provee de una herramienta de clasificación muy potente. Como parte de un grupo de modelos analíticos orientados a realizar predicciones, esta técnica no paramétrica clasifica una población en un modelo de segmentos de tipo ramas que construyen un árbol invertido, y luego este modelo se utiliza para predecir una variable objetivo. Sus principales ventajas se asocian a que puede manejar eficientemente grandes y complejos conjuntos de datos, y a su vez, es fácil de usar y sus resultados sencillos de entender [5].

En cada paso el algoritmo divide los datos en diferentes segmentos, considerando que cada segmento debe representar lo mejor posible una de las potenciales clases. Para medir el nivel de la predicción realizada se usa una tabla de contingencia que describe los casos acertados y los casos erróneos de la predicción. Esta tabla de contingencia se denomina matriz de confusión, y contiene información acerca de las clasificaciones reales y las predichas [6].

La regresión lineal permite aprender una función que mapea un elemento de datos a una variable de predicción de valor real, esta variable de predicción se ajusta a las combinaciones lineales de funciones básicas (sigmoides, curvas, polinomios) a combinaciones de las variables de entrada. Aunado a esto, también se le conoce como el cálculo de la ecuación correspondiente a la línea que mejor describe la relación entre la respuesta y la variable que la explica. Dicha ecuación representa la línea que mejor se ajusta a los puntos en un gráfico de dispersión [2].

## II. METODOLOGÍA

Como parte sustancial del proceso de análisis, se considera una etapa de predicción de indicadores, con base en los datos históricos que se han registrado en cada etapa de captura de los mismos. Esta información ordenada de la forma adecuada permite aplicarle técnicas de minería de datos para poder pronosticar los valores esperados para el año 2018, los cuales podrán contrastarse con los valores esperados para el mismo año y determinar el grado de cumplimiento o de desfase; además de tratar de descubrir información valiosa y de vital importancia para la toma de decisiones.

Con la finalidad de tener una aportación de la minería de datos, se proponen dos técnicas para analizar la misma información, contenida en un archivo .csv, el cual es el mismo para todos los casos. Cabe destacar que el orden de la información es de vital importancia para este proceso, pues en gran medida determina el grado de satisfacción de obtener resultados adecuados y de valor para este proceso.

Antes de describir la aplicación de la técnica empleada, se muestra en la Figura 1 la estructura del archivo .csv con la información histórica de los indicadores a analizar. La estructura del archivo es el resultado de integrar los datos de los 126 Institutos Tecnológicos federales y seis Centros especializados, estos valores son capturados de forma anual en un sistema diseñado para tal efecto, de donde se derivará un archivo por cada año (2012-2017).

#	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
																						Objetivo 1
#	Instituto/Centro	Indicador 1.1						Indicador 1.2						Indicador 1.3								
		2012	2013	2014	2015	2016	2017	2018	2012	2013	2014	2015	2016	2017	2018	2012	2013	2014	2015	2016	2017	
10	IT Aguascalientes	35.28	51.36	51.77	45.96	78.21	89.84	91.49	81.33	88.68	96.49	15.45	47.92	13.85	18.18	24.71	17.14	16.8	16.67			
11	IT Agua Prieta	98.07	100	74.9	76.13	82.24	89.7	90	85.56	89.26	79.37	79	74.07	88.66	40	43.73	47.37	42.86	50			
12	IT Aguascalientes	100	100	82.08	76.88	84.85	73.91	81.61	61.36	60.97	60.28	58.76	56.67	17.1	24.89	17.12	29.41	18.83	28.93			
13	IT Ahuacatlán	86.13	83.32	88.99	82.64	81.05	88.14	46.15	43.08	46.15	50	52.94	50	5.56	12.5	11.13	12.53	26.82	29.41			
14	IT Altamira de Tlaxcala	0	26.43	23.83	23.43	100	100	82.61	57.78	68.09	100	61.36	61.36	6.82	11.54	3.12	7.5	18.82	18.82			
15	IT Alvarado Obregon	0	0	0	0	0	0	0	100	100	100	100	100	0	0	0	0	0	0			
16	IT Apaxtepec	100	100	100	75.76	86.5	82.49	37.4	34.91	36.74	42.84	42.82	46.8	41.3	24.89	18.87	46.43	51.79	50			
17	IT Atlixcoatlán	0	0	0	0	0	0	0	37.8	30	30	35.56	35.56	0	0	0	0	0	0			
18	IT Balleza de Banderas	100	68.82	68.65	85.8	100	100	53.33	53.33	53.33	53.33	75	87.5	0	0	0	15.5	16.67	45.45			
19	IT Boca del Río	17.03	21.89	19.73	17.29	16.49	17.31	18.56	13.82	16.76	47.3	31.75	34.71	15.86	26.09	25.89	22.86	25.71	14.71			
20	IT Coahuila	28.12	32.47	33.45	23.87	4.79	23.11	52.63	51.82	51.82	53.95	54.65	70.67	17.5	20.51	28.21	31.71	38.1	38.96			
21	IT Coahuila	44.82	14.62	30	18.03	30.77	30	50	22.36	22.36	42.86	42.86	12.56	12.56	12.56	5.71	4.29	4.29				
22	IT Colima	100	100	100	100	100	100	61.29	42.1	79.54	86.42	79.53	77.19	52.84	51.47	50.68	51.43	51.43	52.07			
23	IT Cuernavaca	82.11	88.93	93.85	23.84	23.84	23.48	16.67	16.73	16.73	76.41	76.41	68.62	0	0	0	2.17	2.174	40.48			
24	IT Chetumal	100	100	96.98	100	87.28	96.38	48.18	57.8	57.8	57.85	57.85	62.39	44.49	13.21	18.87	16.39	15.11	16.39	17.39		
25	IT Chihuahua	88.7	87.89	87.81	86.33	72.23	72.59	100	100.7	100.1	120	48.7	49.6	22.52	19.61	18.83	100	27.27	24.68			
26	IT Chihuahua II	48.69	64.82	64.82	64.81	64.81	73.82	46	43.81	43.81	44.54	44.53	75	3.78	3.81	3.83	6.45	6.45	15.32			
27	IT Chihuahua III	0	0	0	0	0	28	48	51.92	51.92	50.98	50.98	50.98	0	3.7	11.13	15.88	15.88	11.84			
28	IT Ciudad	0	0	74.92	87.18	100	100	69.7	69.19	103.7	47.69	47.69	48.41	0	10	100	64.67	7.41	17.07			
29	IT Ciudad Altamirano	68.85	64.32	70.81	63.89	63.16	0	37.8	39.39	37.8	38.71	41.94	40	0	0	0	0	0	0			

Figura 1. Archivo .csv con la información concentrada de los indicadores del TecNM

Una vez que se generó el concentrado de los 22 indicadores, de las 132 instituciones, de los siete años, se derivó un archivo por cada indicador en formato .csv, considerando que es un tipo de archivo que es compatible con las herramientas de minería de datos utilizadas, de forma que no es necesario realizar ninguna adecuación adicional para que los datos puedan ser procesados.

Esta actividad de generar los archivos para cada indicador corresponde al proceso de pre-procesamiento en el que los datos son organizados de la mejor forma para que sean procesados de mejor manera por el algoritmo utilizado y por la herramienta de minería de datos.

La Figura 2 muestra la estructura de uno de los archivos .csv que se generaron por indicador, en ellos se prescinde de los nombres de las instituciones y solo se consideran los datos ordenados por columna, y cada una de estas representa un año.

	A	B	C	D	E	F
1	Año1	Año2	Año3	Año4	Año5	Año6
2	35.28	31.36	51.77	65.96	73.21	89.54
3	98.57	100	74.9	76.15	82.24	89.7
4	100	80.9	62.08	76.88	84.85	73.91
5	36.13	33.12	58.99	52.64	51.05	50.14
6	100	100	100	73.76	38.5	52.49
7	100	68.82	69.65	83.9	100	100
8	17.03	21.59	19.73	17.29	16.69	17.31
9	28.12	32.47	33.45	23.57	4.78	22.11
10	45.01	16.01	39	85.03	80.77	80
11	100	100	100	100	100	100
12	82.31	38.93	33.85	23.34	23.94	25.48
13	100	100	96.98	100	87.28	90.38
14	88.7	87.89	87.81	86.33	72.25	72.59
15	48.69	84.05	84.32	82.81	86.93	77.92
16	81.2	90.23	87	67.81	68.41	25.9
17	49.98	83.91	60.48	30.02	25.67	25.52
18	79.98	73.06	91.34	92.56	88.92	87.61
19	88.08	85.62	84.03	90.27	38.97	93.28
20	10.95	24.41	25.78	13.98	14.51	73.2
21	42.99	61.49	73.38	69.16	48.79	24.8
22	95	92.58	95.73	76.55	64.48	84.29
23	73.02	54.58	42.05	12.71	10.6	44.84

Figura 2. Archivo .csv con los datos ordenados correspondiente a un indicador

En la Figura 3 se muestra el resultado del proceso de entrenamiento mediante el algoritmo de árbol aleatorio en Weka®, en el que se genera el árbol y se presentan los valores estadísticos del proceso. Entre los valores estadísticos obtenidos se encuentran: coeficiente de correlación, error absoluto medio, error cuadrado medio, error absoluto relativo, error cuadrado relativo y el número de instancias.

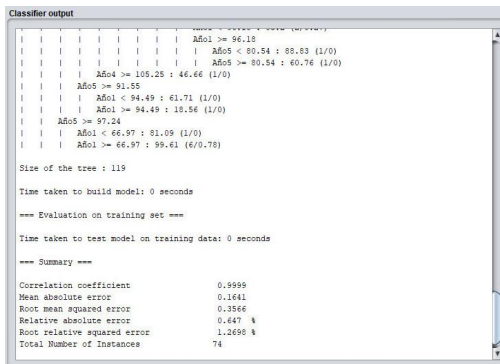


Figura 3. Resultados del proceso de entrenamiento con árbol aleatorio en Weka®

En la Figura 4 se muestra el resultado del proceso de entrenamiento con el algoritmo de regresión lineal, en el que se genera el modelo y se presentan los valores estadísticos del proceso. Entre los valores estadísticos obtenidos se encuentran: coeficiente de correlación, error absoluto medio, error cuadrado medio, error absoluto relativo, error cuadrado relativo y el número de instancias.

Como se puede observar, los valores del modelo de regresión lineal son superiores que los del modelo de árbol aleatorio, lo que indica que el primero permite una mayor diferencia entre los rangos de valores analizados en el conjunto de datos de entrenamiento.

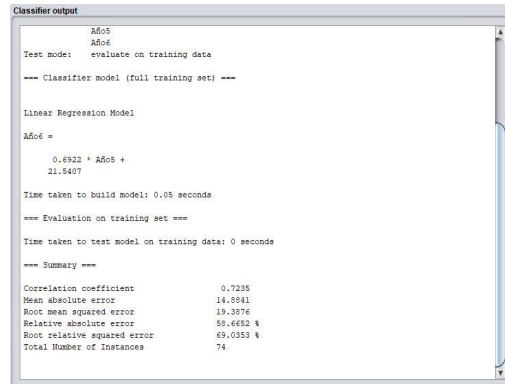


Figura 4. Resultado del proceso de entrenamiento con regresión lineal en Weka®

En la Figura 5 se muestra el resultado del proceso de suministrar un conjunto de prueba, con los valores globales que van a permitir obtener un valor que representa el dato de predicción de todo el indicador. Estos valores son los valores promedio de cada uno de los años, los cuales significan el dato de un indicador a nivel global.

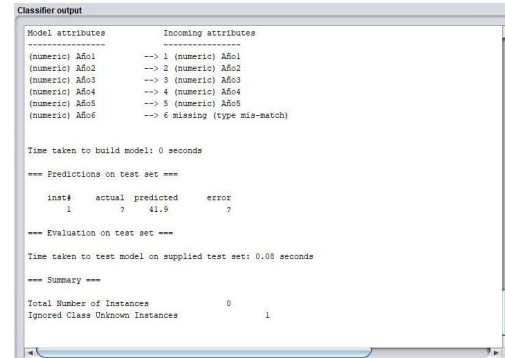


Figura 5. Resultado del conjunto de prueba (predicción) en Weka®

Del resultado del proceso de predicción se obtiene un valor que representa el valor pronosticado para el año 2018, en base a los valores promedio de cada año, de todas las instituciones.

En la Figura 6 se muestra el diseño del modelo de entrenamiento con árbol de decisión en RapidMiner®. En este diseño se muestran los operadores adecuados para poder analizar el conjunto de datos de entrada y producir un modelo de predicción, además de los valores estadísticos evaluables.

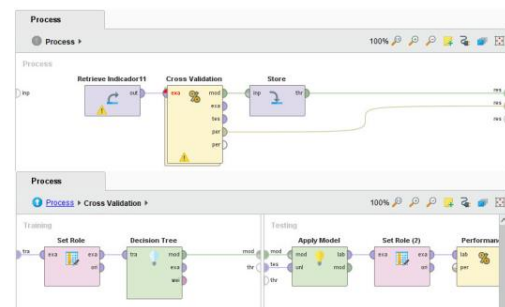


Figura 6. Modelo de entrenamiento con árbol de decisión en RapidMiner®

En la Figura 7 se muestra el diseño del modelo de entrenamiento con regresión lineal en RapidMiner®. En este diseño se muestran los operadores adecuados para poder analizar el conjunto de datos de entrada y producir un modelo de predicción, además de los valores estadísticos evaluables.

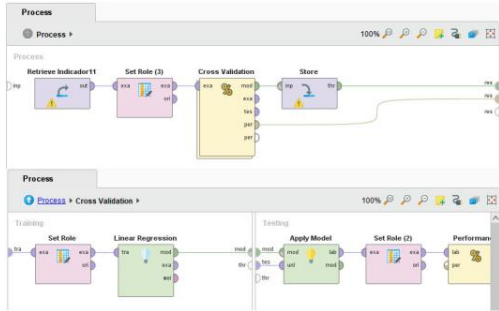


Figura 7. Modelo de entrenamiento con regresión lineal en RapidMiner®

En la Figura 8 se muestra el resultado del proceso de suministrar un conjunto de prueba, con los valores globales que van a permitir obtener un valor que representa el dato de predicción de todo el indicador. Estos valores son los valores promedio de cada uno de los años, los cuales significan el dato de un indicador a nivel global.

Row No.	predicciónA...	Año1	Año2	Año3	Año4	Año5	Año6
1	42.543	49.250	49.840	48.990	41.400	42.400	?

Figura 8. Resultado del conjunto de prueba (predicción) en Weka®

Los valores obtenidos de la predicción con regresión lineal demuestran una mejor adaptación a los datos de entrada, ya que siendo estos muy heterogéneos, las características del algoritmo permiten pronosticar valores más cercanos a los valores reales, para el presente caso de estudio, los datos correspondientes al año 2017.

### III. RESULTADOS

La Tabla II muestra los valores obtenidos de las herramientas Weka® y RapidMiner®, una vez que sea realizó el proceso de minería de datos con las técnicas de árboles de decisión y regresión lineal, utilizando los algoritmos de árbol aleatorio y bosque aleatorio.

Tabla II. Valores obtenidos por algoritmo utilizado

Indicador	Programado	Valores 2018			
		Weka®		RapidMiner®	
		Árbol aleatorio	Regresión lineal	Árbol de decisión	Regresión lineal
1.1	72	41.90	50.89	41.90	49.07
1.2	70	56.83	59.54	57.90	59.78
1.3	28	14.62	22.97	25.10	22.97
1.4	70	38.59	36.29	39.23	34.06
2.1	602,084	10,681	362,952	10,646	358,348
2.2	6,000	325	5,194	325	5,227
2.3	20,000	1,094	12,067	1,094	9,089
3.1	30	63.85	39.91	60.92	39.91
3.2	50	49.83	38.59	68.93	39.63
3.3	60	40.16	34.76	42.39	34.56
4.1	80	100	98.25	100	98.25

4.2	600	47	570	47	545
4.3	900	3	552	3	552
4.4	3,600	129	6,635	379	6,383
5.1	700	98	243	98	185
5.2	65	41.31	53.37	41.82	52.07
5.3	2,700	437	8,207	437	8,338
5.4	8,100	1,711	22,454	683	24,727
5.5	2,222	16	367	16	365
5.6	147,349	100	3,408	100	4,126
6.1	15,000	67	7,394	316	7,394
6.2	696	2.75	112	2.75	112

Los datos presentados en la tabla se entienden desde el sentido que los valores programados están muy elevados y también a que algunos indicadores representan porcentajes (proporciones) y otros valores numéricos son obtenidos de la sumatoria de los datos individuales; además la información que se tiene es parcial pues representa una parte del total de instituciones que integran el TecNM, el cual asciende a 254.

En la Tabla III se presentan los valores obtenidos de la minería de datos, pero comparándolos con los datos correspondientes al año 2017, en esta presentación se observa que hay valores que si corresponden desde una perspectiva simple. Sin embargo, se puede observar que hay una disparidad muy grande entre los valores obtenidos y los valores reales, en los indicadores que tienen que ver con sumatoria, es decir, en donde el resultante es la agregación de todos los valores de los institutos, con las técnicas de árboles, en ambas herramientas.

Tabla III. Valores 2017 vs valores obtenidos

Indicador	Valores 2017	Valores 2018			
		Weka®		RapidMiner®	
		Árbol aleatorio	Regresión lineal	Árbol de decisión	Regresión lineal
1.1	42.40	41.90	50.89	41.90	49.07
1.2	58.57	56.83	59.54	57.90	59.78
1.3	21.07	14.62	22.97	25.10	22.97
1.4	35.78	38.59	36.29	39.23	34.06
2.1	350,365	10,681	362,952	10,646	358,348
2.2	4,614	325	5,194	325	5,227
2.3	10,987	1,094	12,067	1,094	9,089
3.1	38.90	63.85	39.91	60.92	39.91
3.2	39.64	49.83	38.59	68.93	39.63
3.3	34.88	40.16	34.76	42.39	34.56
4.1	20.05	100	98.25	100	98.25
4.2	546	47	570	47	545
4.3	1,678	3	552	3	552
4.4	8,648	129	6,635	379	6,383
5.1	214	98	243	98	185
5.2	49.02	41.31	53.37	41.82	52.07
5.3	9,160	437	8,207	437	8,338
5.4	29,089	1,711	22,454	683	24,727
5.5	493	16	367	16	365
5.6	5,220	100	3,408	100	4,126
6.1	9,136	67	7,394	316	7,394
6.2	144	2.75	112	2.75	112

Sin embargo, se puede observar que el algoritmo de regresión lineal es el que mejor se adapta a los valores de todos los indicadores, sin importar si estos se obtienen por proporción o por agregación.

#### IV. CONCLUSIONES

El análisis realizado sobre los valores históricos de los indicadores 2012-2017, ha permitido obtener valores que representan una predicción de los valores correspondientes al año 2018, y que de manera real se obtendrán de la captura anual, con lo que en su momento se podrán comparar estos valores, los obtenidos del presente documento y los reales que se obtienen del sistema de captura de indicadores del TecNM.

Los valores pronosticados que mejor representan a los indicadores, con base a los datos históricos, y fundamentalmente en comparación con los datos del 2017, son los indicadores derivados de la técnica de regresión lineal. Mientras que datos obtenidos de las técnicas de árboles empleadas solo se ajustaron a los datos que representan valores porcentuales y los datos que representan valores sumariados no son coherentes con los valores del 2017 ni con los indicadores esperados. Esta disparidad puede deberse a la heterogeneidad que existe en los valores que se obtienen de las fórmulas establecidas para generar los valores de cada indicador.

Para validar los datos generados por la predicción de este trabajo, es necesario compararlos con los valores que las instituciones reportarán, los cuales corresponden al año 2018,

cuando se disponga de ellos, se podrá verificar la validez de los datos pronosticados y del grado de confianza de los algoritmos utilizados.

#### REFERENCIAS

- [1] M. López Nava, F. J. Cartujano Escobar, y A. Alday Chávez, "Análisis de indicadores de desempeño del Tecnológico Nacional de México con Lógica Difusa" Revista Coloquio Internacional de Investigación Transdisciplinaria 2018, Instituto Tecnológico de Matamoros, Diciembre 2018, pp. 37-45,.
- [2] Programa Institucional de Innovación y Desarrollo del Tecnológico Nacional de México 2013-2018, Tecnológico Nacional de México, consultado desde [http://www.tecnm.mx/images/areas/planeacion/2014/PIID\\_2013-2018\\_TECNM\\_Final.pdf](http://www.tecnm.mx/images/areas/planeacion/2014/PIID_2013-2018_TECNM_Final.pdf).
- [3] U. Fayyad, G. Piatetsky-Shapiro, y P. Smith "From Data Mining to Knowledge Discovery in Databases", AI Magazine, Vol. 17, 1996, pp.37-54.
- [4] H. Jin, Y. Lu and K. Stone, "Alternative tree structured survival analysis based on variance of survival time", Med Decis Making, 2004, pp. 670-680. doi: <http://dx.doi.org/10.1177/0272989X10377117>.
- [5] Y. Y. Song and L. U. Ying, "Decision tree methods: applications for classification and prediction", Shanghai Archives of Psychiatry, 2015, pp. 130-135.
- [6] P. E. Ramírez, y E. E. Grandón, "Predicción de la Deserción Académica en una Universidad Pública Chilena a través de la Clasificación basada en Árboles de Decisión con Parámetros Optimizados" Formación universitaria, 2018, pp. 3-10, <https://dx.doi.org/10.4067/S0718-50062018000300003>.



# Identificación de Patrones en Antibiogramas Mediante Técnicas de Aprendizaje Automático

Lic. Carlos A. González Rubio Gavarain  
 Instituto Tecnológico de La Paz  
 La Paz B.C.S., México C.P. 23080  
 Email: gonzalezMSC1705@itlp.edu.mx

Dr. Marco Antonio Castro Liera  
 Instituto Tecnológico de La Paz  
 La Paz B.C.S., México C.P. 23080  
 Email: mcastro@itlp.edu.mx

**Resumen**—Se detalla la creación de una solución informática para facilitar la descripción y visualización de los patrones subyacentes en los datos mixtos multidimensionales de los antibiogramas.

El enfoque de desarrollo utilizado es conocido como “descubrimiento de conocimiento en bases de datos” cuyo objetivo es identificar patrones válidos y novedosos a partir de datos existentes, mismos que son suministrados en forma de ejemplos entrenantes al algoritmo de mapa auto-organizado generalizado, utilizado durante la etapa de minado de datos para obtener un modelo que mapea datos de alta dimensión en una malla bidimensional que preserva las propiedades topológicas de los datos originales.

**Palabras clave**—Antibiogramas, Farmacorresistencia, Aprendizaje no supervisado, Minería de datos, Mapas Auto-Organizados, Distancia Jerárquica.

## I. INTRODUCCIÓN

El antibiograma es una prueba microbiológica que se realiza para determinar la susceptibilidad de una bacteria a un grupo de antibióticos, misma que es realizada en las unidades hospitalarias para brindar opciones terapéuticas en el tratamiento de los pacientes.

Los antibiogramas se realizan en el momento del ingreso de un paciente al nosocomio para verificar que no tenga o esté incubando algún microorganismo, a los cinco días después de su ingreso o si presenta picos febriles.

Las IAAS (Infección Asociada a la Atención de la Salud) es un evento adverso durante la prestación de atención sanitaria y son de especial preocupación debido al uso inherente de antibióticos que genera presión selectiva creando microorganismos difíciles de tratar por el fenómeno conocido como farmacorresistencia.

La farmacorresistencia en unidades hospitalarias es un tema que ya ha sido investigado adoptando enfoques estadísticos, sin embargo, la falta de consenso y estructura hacen que se trate de un tópico poco formalizado.

La epidemiología es el estudio de la distribución y los determinantes de los estados o acontecimientos relacionados con la salud en poblaciones específicas y la aplicación de este estudio al control de los problemas sanitarios [1].

El presente trabajo describe la problemática y las soluciones llevadas a cabo durante el desarrollo de un programa informático que permita conocer el número, distribución y

similitud de los microorganismos aislados en las unidades hospitalarias de acuerdo a los campos contenidos en los antibiogramas.

## II. EL PROCESO DE DESCUBRIMIENTO DE CONOCIMIENTO EN BASES DE DATOS

Teniendo como objetivo la generación de un modelo a partir de las características extraídas de los antibiogramas existentes se utilizó un proceso de desarrollo diseñado para transformar datos en conocimiento conocido como KDD [2] (por las siglas en inglés de *Knowledge Discovery in Databases*) donde se encapsulan las responsabilidades en un conjunto de pasos que generan una salida a partir de los datos procesados de la entrada del paso anterior como se muestra en la figura 1.



Figura 1: Pasos del proceso de descubrimiento de conocimiento en bases de datos

### II-A. Creación de la base de datos

Con la finalidad de poder utilizar los datos de los antibiogramas se creó una base de datos relacional *SQL normalizada* (proceso para evitar la redundancia en los datos) utilizando el sistema de gestión de bases de datos *MariaDB*.

### II-B. Extracción, transformación y carga de datos

Esta etapa abarca la selección de pre-procesamiento y transformación de datos.

Los antibiogramas son impresos por el laboratorio de análisis clínicos, los cultivos que salen positivos son enviados a la Unidad de Vigilancia Epidemiológica Hospitalaria donde son capturados en la base de datos.

Los datos de los antibiogramas negativos (donde no hubo crecimiento de microorganismos) o algunos errores de captura de datos (por ejemplo si se aislo una bacteria y se probó con un medicamento antiviral en el antibiograma) no se incluyeron.

### II-C. Minado de datos

Es una de las etapas más importantes en KDD donde se selecciona el algoritmo y la tarea que se realizará (clasificación, asociación, predicción, agrupamiento). Como resultado de este paso se obtiene un modelo de los datos que puede ser utilizado posteriormente para el análisis de los mismos.

## III. PROBLEMÁTICA

### III-A. Dimensionalidad

Uno de los principales problemas al tratar de agrupar las observaciones es el número de características, debido a que cada una de estas representa una dimensión del problema y si aumentamos el número de características el espacio aumenta exponencialmente haciendo que los datos disponibles se vuelvan dispersos.

### III-B. Datos mixtos

El tratamiento de bases de datos con campos mixtos (ordinales y categóricos) es un área de investigación actualmente abierta y representa un problema debido a que la mayoría de los algoritmos para agrupamiento utilizan un concepto de similitud entre dos observaciones en términos de la cercanía que existe entre estas, misma que no está definida para datos categóricos.

### III-C. Interpretación

Para que los datos proporcionados sean útiles, deben de ser presentados de forma que tengan significado para el usuario objetivo, en este caso profesionales de la salud.

## IV. PROPUESTA DE SOLUCIÓN

### IV-A. Aprendizaje no supervisado

Dado que nuestro conjunto de datos no está etiquetado (no hay una columna con las salidas deseadas) es necesario seleccionar un método que nos permita ajustar un modelo de densidad a partir de las observaciones existentes.

### IV-B. Agrupamiento

Es una de las tareas principales de la minería de datos exploratoria, consiste en clasificar observaciones agrupándolas por la similitud que existe entre ellas y la diferencia entre los grupos formados, para llevarla a cabo existen diversos algoritmos que varían en función del problema.

### IV-C. Mapas Auto-Organizados

De acuerdo a la problemática planteada en la sección III-A se seleccionó el algoritmo de aprendizaje no supervisado llamado Mapa Auto-Organizado (SOM por las siglas en inglés de *Self Organizing Map*) debido a su capacidad para generar una representación discreta de baja dimensión (típicamente  $R^2$ ) del espacio de las muestras de entrada de alta dimensión ( $R^n$ ) [3], llamado mapa.

Esta representación se puede utilizar para visualizar correlaciones complejas entre los datos de entrada debido a que durante la etapa de entrenamiento el SOM genera un mapa topológico al acercar entre sí las observaciones que son

similares mediante una función de vecindad.

El algoritmo empleado para entrenar un SOM se describe a continuación:

---

#### Algorithm 1 Algoritmo de entrenamiento de Mapa auto-organizado

---

**Input:**  $D, \sigma_0, \alpha_0$

**Output:** Mapa topológico con pesos adaptados

```

1: Hacer un mapa de neuronas con vectores de pesos aleatorios
2: for  $s \leftarrow 1$  to  $\lambda$  do
3:    $\sigma_s \leftarrow \text{radiusDecayFunction}(\sigma_0, s, \lambda)$ 
4:    $\alpha_s \leftarrow \text{learningRateDecayFunction}(\alpha_0, s, \lambda)$ 
5:   for  $t \leftarrow 1$  to  $\text{size}(D)$  do
6:      $u \leftarrow \text{bmu}(\text{neurons}, D(t))$ 
7:     for  $v \leftarrow 1$  to  $\text{size}(\text{neurons})$  do
8:        $d \leftarrow \text{distance}(\text{neuron}_u, \text{neuron}_v)$ 
9:       if  $d \leq \sigma_t$  then
10:         $W_v(s+1) = W_v(s) + \Theta(d, \sigma_s)\alpha_s(D(t) - W_v)$ 
11:       end if
12:     end for
13:   end for
14: end for
    
```

---

Donde:

- $s$  es la iteración actual.
- $\lambda$  es la cantidad total de iteraciones.
- $D(t)$  es un vector de entrada de índice  $t$  del conjunto de datos de entrada  $D$ .
- $v$  es el índice de una neurona en el mapa.
- $u$  es el índice del BMU en el mapa.
- $W_v$  es el vector de pesos de la neurona  $v$ .
- $d$  es la distancia de la *neurona<sub>u</sub>* a la *neurona<sub>v</sub>*.
- $\sigma_s$  es el radio de vecindad en la iteración  $s$ .
- $\alpha_s$  es un restrictor de aprendizaje debido al progreso de las iteraciones.
- $\Theta(d, \sigma_s)$  es la función de vecindad.

### IV-D. Mapas Auto-Organizados Generalizados

De acuerdo a la problemática planteada en la sección III-B, al utilizar datos categóricos SOM falla en preservar correctamente la topología de los datos originales, por lo que en 2006 un nuevo algoritmo GSOM [4] utiliza una nueva métrica de distancia llamada *distancia jerárquica* compuesta por *árboles de jerarquías* donde las aristas contienen un valor o *peso* y los nodos representan *conceptos* ordenados de lo más general en la raíz, hasta los más particulares en los nodos *hojas* como se muestra en la Figura 2.

Estas estructuras jerárquicas permiten manejar los distintos tipos de datos: categóricos, ordinales e incluso codificación binaria para coincidencias simples de forma unificada de allí el término de *generalización*.

**IV-D1. Distancia Jerárquica:** Cada punto está compuesto por una tupla *ancla* (un nodo hoja) y una ganancia (un valor real que representa la distancia desde el punto a la raíz) por ejemplo el punto  $X = [S. aureus, 2]$  (ver fig. 2a).

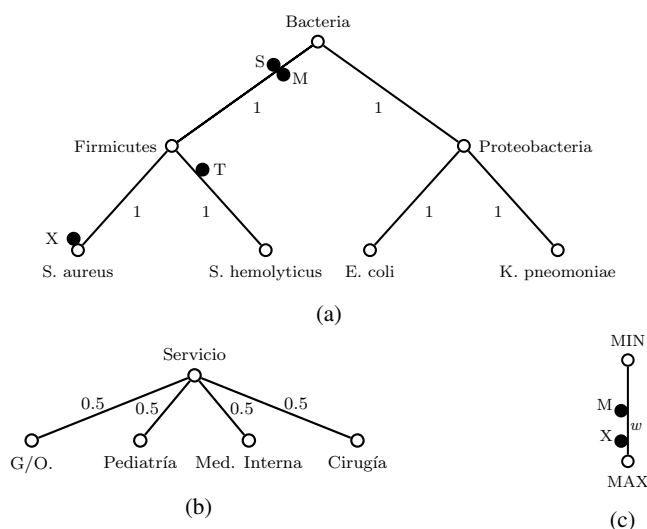


Figura 2: Tipos de distancias jerárquicas: Categórica, Simple y Ordinales.

Los patrones de entrada son mapeados a puntos en la distancia jerárquica, por ejemplo: dado el patrón de entrada (S. aureus, 7) y teniendo los árboles de jerarquías mostrados en la Fig2(a) y Fig2(c) darán como resultado un vector de puntos llamado mapa:  $m = [(S. aureus, 2), (MAX, 7)]$ , estos sustituyen los patrones de entrada y los pesos de las neuronas del algoritmo de SOM original.

La distancia entre dos mapas  $x$  e  $y$  es calculada utilizando (1) donde  $dh$  es un vector de distancias jerárquicas compuesto por árboles de jerarquías y el subíndice  $i$  es cada componente del mapa creado del patrón de entrada.

$$d(x, y) = \left( \sum_{i=1}^n w_i |dh_i(x_i) - dh_i(y_i)|^L \right)^{1/L} \quad (1)$$

De esta forma cada componente del patrón de entrada es mapeado y asociado a un árbol de distancia jerárquica  $dh_i$  correspondiente cuyo valor es calculado utilizando (3).

**IV-D2. Distancia Jerárquica Categórica:** Para poder operar con la distancia jerárquica del tipo categórico se deben de establecer algunos axiomas:

- Dos puntos en distancia jerárquica son equivalentes si se encuentran en la misma posición aunque no tengan la misma ancla.
- Un nodo  $X$  es ancestro de otro  $Y$  si  $X$  se encuentra en el camino de  $Y$  a la raíz.
- El ancestro menos común de dos nodos  $X$  e  $Y$  denotado como  $LCA(X, Y)$  por las siglas en inglés de *Least Common Ancestor*.

- El punto menos común de dos nodos  $X$  e  $Y$  denotado como  $LCP(X, Y)$  por las siglas en inglés de *Least Common Point* (2) se define según sea el caso:

$$LCP(X, Y) = \begin{cases} X \text{ ó } Y, & \text{si } X \equiv Y \\ Y, & \text{si } Y \text{ es ancestro de } X \\ X, & \text{si } X \text{ es ancestro de } Y \\ LCA(X, Y), & \text{de otro modo} \end{cases} \quad (2)$$

- La distancia entre dos puntos es calculada utilizando (3):

$$|X - Y| = d_x + d_y - 2LCP(X, Y) \quad (3)$$

Por ejemplo en la Fig.2a sean los puntos:

$$\begin{aligned} M &= (S. aureus, 0,3) \\ S &= (S. hemolyticus, 0,3) \\ T &= (S. hemolyticus, 1,3) \\ X &= (S. aureus, 2,0) \end{aligned}$$

- La distancia entre los puntos  $M$  y  $S$  es cero ya que son equivalentes.
- La distancia entre los puntos  $T$  y  $M$  es  $(1,3 + 0,3 - 2 * 0,3) = 1$ .
- La distancia entre los puntos  $X$  y  $T$  es  $(2,0 + 1,3 - 2 * 1,0) = 1,3$  ya que el  $LCP(X, T)$  es 'Firmicutes' que tiene una distancia de 1.

Durante la fase de ajuste un punto  $M$  es ajustado a otro punto  $X$  situado en algún nodo hoja con un valor real  $\delta$  y  $N = LCA(M, X)$  según el caso que corresponda:

1. Si  $M$  es ancestro de  $X$  y no se pasa de  $N$  después del ajuste entonces  $d_M = d_M + \delta$
2. Si  $M$  es ancestro de  $X$  y si se pasa de  $N$  después del ajuste entonces  $anchor_M = anchor_X$  y  $d_M = d_M + \delta$
3. Si  $N$  es ancestro de  $M$  y no se pasa de  $N$  después del ajuste entonces  $d_M = d_M - \delta$
4. Si  $N$  es ancestro de  $M$  y si se pasa de  $N$  después del ajuste entonces  $anchor_M = anchor_X$  y  $d_M = 2d_{N_{LCA}} - d_M + \delta$

**IV-D3. Distancia Jerárquica Simple:** Este tipo de distancia jerárquica opera de forma similar a homónima *categórica* pero difiere en dos aspectos: no necesita la operación *Least Common Point* y el cálculo de la distancia es 1 si tienen la misma *ancla* y 0 en caso contrario.

**IV-D4. Distancia Jerárquica Numérica:** En esta distancia el mapeo del patrón siempre retorna un punto donde en *ancla* siempre es  $MAX$  y el valor del *offset* es la distancia de  $X$  a la raíz  $MIN$  (ver fig. 2c) entonces la distancia entre dos puntos es calculada restando los dos valores de *offset*.

En la implementación se crean tres operaciones polimórficas para delegar la responsabilidad de mapear, calcular la distancia y ajustar los pesos a sus respectivas estructuras (ver Fig. 3), es decir que los árboles derivados de la clase *DistanceHierarchy* realizan estas operaciones de forma distinta según sea el caso.

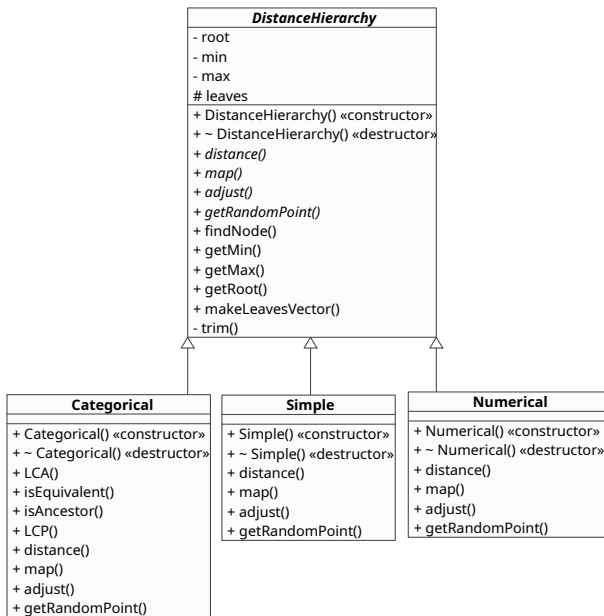


Figura 3: Jerarquía de la clase DistanceHierarchy

#### IV-E. Las Jerarquías en los antibiogramas

Como se vio en el apartado IV-C es posible utilizar el GSOM para visualizar los grupos cuando se tienen datos categóricos; en el caso de los antibiogramas cada patrón de entrada tiene los siguientes componentes: Fecha, Servicio, Género, Muestra, Microorganismo, Antibiótico y Susceptibilidad los cuales se intentan correlacionar, sin embargo para que estas relaciones tengan mayor significancia se crearon dos árboles de distancia categóricos sobre las variables de Microorganismos y Antibióticos utilizando en el primer caso el árbol taxonómico usado en la clasificación compuesto por el Dominio, Filo, Clase, Orden, Familia, Género, Especie (Ejemplo: Bacteria, Proteobacteria, Gammaproteobacteria, Enterobacteriales, Enterobacteriaceae, Escherichia, E. coli); en el caso de los antibióticos se utilizaron los campos de Grupo, Subgrupo y Antibiótico (Ejemplo: Aminoglucósidos, Monobactamicos, Gentamicina), se puede utilizar un árbol del tipo numérico para la fecha y el resto de los campos como una relación simple como se describió anteriormente en el apartado IV-D1.

### V. PRUEBAS Y RESULTADOS

#### V-A. Pruebas

Se creó un mapa auto-organizado generalizado de 10 filas por 10 columnas que se entrenó durante 100 épocas con 608 patrones de pares [microorganismo, antibiótico] sin duplicados, se utilizó una tasa de aprendizaje inicial de 0.9, la función de vecindad utilizada fue gaussiana con un radio inicial que abarca el mapa completo, en este caso  $\sqrt{10^2 + 10^2} \approx 14,1421$ , aplicando el redondeo hacia arriba nos queda 15.

Se crearon dos árboles jerárquicos del tipo categórico uno de microorganismos con un tamaño de profundidad de 7 y

una anchura de 46 y otro árbol para los antibióticos con una profundidad de 3 y un ancho de 48.

Las pruebas se realizaron en una Laptop DELL Inspiron 5559 con 8GB de memoria RAM con sistema operativo Arch Linux x86\_64 Kernel: 5.0.0-arch1-1-ARCH

#### V-B. Resultados

El tiempo de entrenamiento y ejecución fue de 47.18 segundos, los resultados son mostrados utilizando una representación conocida como *Hit map* (ver fig. 4) donde los números dentro de las casillas representan el número de patrones coincidentes con la neurona en la posición  $x, y$  y el color es asignado de acuerdo a dicho número.

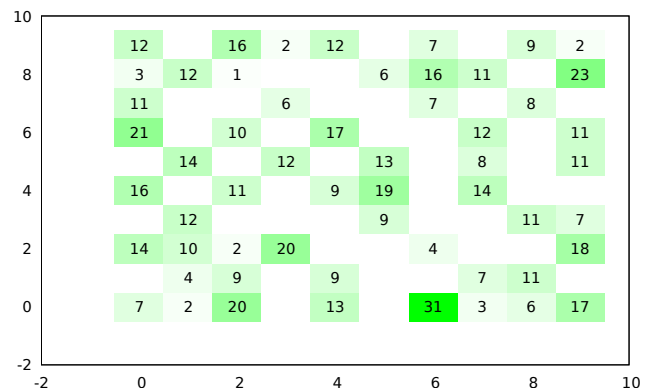


Figura 4: Mapa auto-organizado generalizado de dos distancias jerárquicas (Microorganismos y Antibióticos).

Para verificar la preservación topológica de los datos veremos en detalle el grupo formado por las neuronas (6, 0), (7, 0), (8, 0), (9, 0), (7, 1) y (8, 1).

- La neurona (6, 0) tiene 31 coincidencias y su vector prototipo es: { Acinetobacter iwoffii, 6.00452 }, { Levofloxacina, 2.00839 }
- La neurona (7, 0) tiene 3 coincidencias y su vector prototipo es: { Acinetobacter iwoffii, 6.00498 }, { Levofloxacina, 0.768032 }
- La neurona (8, 0) tiene 6 coincidencias y su vector prototipo es: { Acinetobacter haemolyticus, 6.01374 }, { Ampicilina, 1.05609 }
- La neurona (9, 0) tiene 17 coincidencias y su vector prototipo es: { Acinetobacter haemolyticus, 6.00209 }, { Ampicilina, 1.99333 }
- La neurona (7, 1) tiene 7 coincidencias y su vector prototipo es: { Acinetobacter baumannii, 6.02939 }, { Cefotaxima, 2.00268 }
- La neurona (8, 1) tiene 11 coincidencias y su vector prototipo es: { Acinetobacter baumannii, 6.00723 }, { Cefotaxima, 1.99926 }

Estas seis neuronas agrupan 75 de las 608 observaciones, todas estas corresponden a microorganismos del género Acinetobacter en contraposición por ejemplo con la neurona (0, 6) en donde las 21 observaciones corresponden al género de los Enterococos.

En el caso de las neuronas (7, 0) y (8, 0) las observaciones coincidentes son las siguientes:

- Neurona(7, 0):
  1. Acinetobacter baumannii, Moxifloxacino
  2. Acinetobacter calcoaceticus, Moxifloxacino
  3. Acinetobacter haemolyticus, Moxifloxacino
- Neurona(8, 0):
  1. Acinetobacter baumannii, Imipenem y Cilastatina
  2. Acinetobacter baumannii, Meropenem
  3. Acinetobacter calcoaceticus, Imipenem y Cilastatina
  4. Acinetobacter haemolyticus, Meropenem
  5. Acinetobacter iwoffii, Imipenem y Cilastatina
  6. Acinetobacter iwoffii, Meropenem

En estos casos ambas neuronas están agrupadas de forma contigua, pero preservando las propiedades topológicas de los datos originales. Como se puede ver el primer criterio de agrupamiento es que todos pertenecen al género *Acinetobacter*, pero se encuentran en distintos grupos debido al segundo criterio de *antibióticos*, en este caso el *Moxifloxacino* pertenece al subgrupo de los *Monobactamicos* mientras que *Imipenem* y *Cilastatina* y *Meropenem* pertenecen al subgrupo de los *Carbapenemicos*.

## VI. CONCLUSIONES Y TRABAJO FUTURO

Los pesos de las neuronas ó vectores prototipo fueron ajustados mediante el entrenamiento utilizando una función de vecindad y una similitud de acuerdo a la distancia jerárquica.

En las pruebas realizadas la implementación preserva correctamente la topología de los datos categóricos, lo que nos permite entender la magnitud y distribución de los microorganismos según los criterios dados.

El parámetro  $w$  es un vector utilizado en (1) que puede ser utilizado para darle prioridad a un parametro  $dh_i$ , lo cuál puede conveniente al investigador para encontrar relaciones significativas en valores multidimensionales.

En el presente trabajo se utilizaron 608 patrones correspondientes a los campos categóricos de Microorganismos y Antibióticos para probar la implementación de la distancia jerárquica, sin embargo la base de datos está compuesta por 10,465 observaciones donde cada vector está compuesto por los campos de Fecha, Género, Servicio, Cultivo, Microorganismo, Antibiótico y Susceptibilidad por lo que se implementará una solución que incluya dichos campos.

Se está trabajando en una interfaz gráfica de usuario para poder variar los parámetros de entrenamiento y mostrar los resultados de forma que proporcionen mayor información y significancia.

Los mapas auto-organizados también pueden ser representados utilizando *matrices U* (matrices de distancia

unificada) que son creadas utilizando la distancia promedio entre la neurona y sus vecinos más cercanos, esta distancia se puede representar en una imagen en escala de grises donde los colores claros representan nodos cercanos mientras que los colores más oscuros nodo más distantes. Por lo tanto, los grupos de colores claros se pueden considerar como agrupaciones, y las partes oscuras como los límites entre las agrupaciones. Esta representación puede ayudar a visualizar los grupos en los espacios de alta dimensión.

Una vez probado el funcionamiento de los algoritmos para la tarea dada se puede trabajar en una implementación paralela para mejorar el rendimiento.

## AGRADECIMIENTOS

Se agradece al Tecnológico Nacional de México, a la División de Estudios de Posgrado e Investigación del Instituto Tecnológico de La Paz y al Consejo Nacional de Ciencia y Tecnología (CONACyT), por el apoyo recibido y las facilidades otorgadas para el desarrollo de este trabajo.

Se agradece al comité tutorial por todas sus valiosas observaciones y aportaciones para este trabajo fuera posible.

Se agradece al Benemérito Hospital General con Especialidades Juan María de Salvatierra que mediante el área de Enseñanza e Investigación a través del conducto de su titular el Dr. Gustavo Farías Noyola por haber otorgado el permiso para la utilización de una base de datos de antibiogramas.

Se agradece a la Unidad de Vigilancia Epidemiológica Hospitalaria y al Laboratorio de Análisis Clínicos del Hospital Salvatierra por haber fomentado en mí la investigación en salud y la conciencia sobre la importancia de la farmacoresistencia.

## REFERENCIAS

- [1] L. Gordis, *Epidemiología*, 5th ed., D. Edited by Leon Gordis, MD, MPH, Ed. Barcelona, España: Elsevier, 2015.
- [2] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "The kdd process for extracting useful knowledge from volumes of data," *Communications of the ACM*, vol. 39, no. 11, pp. 27–35, 1996.
- [3] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, no. 1, pp. 59–69, 1982. [Online]. Available: <http://link.springer.com/10.1007/BF00337288>
- [4] C.-C. Hsu, "Generalizing self-organizing map for categorical data," *IEEE transactions on Neural Networks*, vol. 17, no. 2, pp. 294–304, 2006.

# Recoloring Ishihara Plates with PSO algorithm and Proposed Equations

Martín Montes Rivera  
 Postgraduate Studies and Research  
 Universidad Politécnica  
 de Aguascalientes  
 Aguascalientes, Ags.  
 Email: martin.montes@upa.edu.mx

Alejandro Padilla Díaz and  
 Julio César Ponce Gallegos  
 Universidad Autónoma  
 de Aguascalientes  
 Aguascalientes, Ags.  
 Email: apadilla2004@hotmail.com,  
 julk\_cpg@hotmail.com

Juana Canul-Reich and  
 Alberto Ochoa Zezzatti  
 Universidad Juárez  
 Autónoma de Tabasco  
 Villahermosa, Tab.  
 Email: juana.canul@ujat.mx,  
 alberto.ochoa@uacj.mx.

**Abstract**—Human sight is the most important sense, because people use it in their lives to guarantee their safety. Unfortunately, close to 10% of the worldwide population is colorblind, which causes difficulties with the perception of color. There are different tests for diagnosing colorblindness; the most popular of them is the Ishihara test. Several assistants and awareness alternatives have been created, which use image processing and artificial intelligence to aid the colorblind. The development of these partial solutions include the use of artificial neural networks, which are applied with different color spaces reaching up to 99.8% efficiency but increasing computational power required for processing several hidden layers in the neural network or in the color space transformation. In this paper, the application of equations for color classification optimized with PSO algorithm is proposed, since similar results are obtained, besides the fact that computational power is reduced. The equations are tested with Ishihara plates and evaluated using a colorblindness simulation model.

## I. INTRODUCTION

Human sight is the most important sense, because people use it in their daily lives for identifying food and dangers; it is also used in activities that include learning, reading, writing and driving, among others, which are common in modern life [1].

Several body structures are involved in producing the sense of sight. Initially, the eyes use two types of cells (rods and cones), for acquiring electrical signals and then send them to the brain for them to be processed and converted into images. Rod cells perceive the wave length or quantity of light, while cone cells perceive the wave frequency or color in those signals [2], [3].

Unfortunately, almost 10% of the worldwide population has a condition named colorblindness, which causes serious limitations with color perception [4], [5].

These absent or diminished color perception, affects interaction of the colorblind with their environment and with the trichromats (people with normal perception of colors i.e. perceive color variations produced with red, green and blue cones), since several activities require the perception of colors that trichromats perceive [6].

Colorblindness can be classified according to its severity and according to the very cone that causes it [6].

Anomalous trichromacy is the less severe variant of colorblindness characterized by mild deficiencies in the perception of colors associated with a deficient cone, and it is subdivided in protanomaly (a deficiency with the red cones), deuteranomaly (a deficiency with the green cones) and tritanomaly (a deficiency with the blue cones) [6], [7].

If a person is missing one kind of cone, then severe deficiencies in the perception of color are produced; this variant of colorblindness is called dichromacy, which can be subdivided into protanopia for red cones, deuteranopia for green cones and tritanopia for blue cones [6], [7].

The most severe condition is when the entire perception of color is lost, and this condition is called monochromacy [6], [7].

Colorblindness is detected using tests designed for this propose. These can be divided into four areas of diagnosis: The anomaloscopy, the accommodation order disks, the pseudo chromatic plates and the electronic applications tests [3], [8].

The most commonly used test is the Ishihara pseudo chromatic test, whose logo has become an icon in conversations about this mild disability. This test presents patients with plates with numbers, this patients would not see such numbers if they suffer from protanopia, protanomaly, deuteranopia or deuteranomaly(Fig. 1) [3], [7].

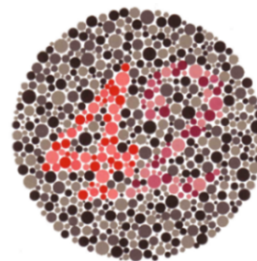


Fig. 1. Ishihara plate with number 42.

There are various alternatives either for assisting the colorblind or for helping raise awareness in trichromats of difficulties experimented by those who suffer from this condition like it is described in [4].



Those alternatives include Enchroma<sup>TM</sup> glasses that can help improve color perception in mild colorblindness by shifting the wavelengths of light, and Variator<sup>TM</sup> glasses that simulate dichromacy [9].

Alternatively, in the field of image processing and computer vision, other researches have appeared either for assisting the colorblind or for helping raise awareness in people with a normal vision [4].

Examples of this processing algorithms include colorblindness simulation models, which allow trichromats to perceive the world the way that colorblind people do, these models are included in [10] and are implemented in smart phones applications like Chromatic Vision Simulator<sup>TM</sup> (CVS) designed for Android<sup>TM</sup>. An example of how colorblind people perceive the color spectrum obtained with CVS is shown in the Fig. 2.

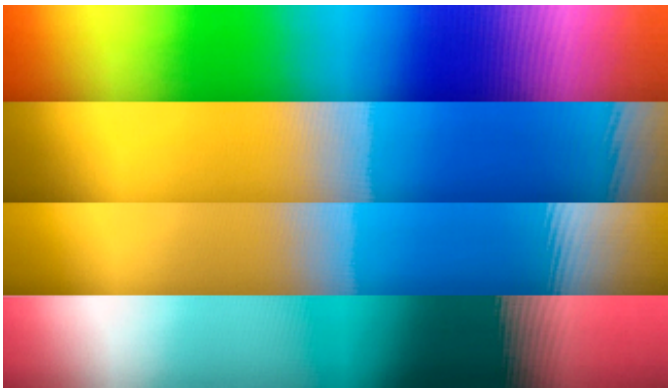


Fig. 2. Color spectrum perceived by trichromats and colorblind people with protanopia, deuteranopia and tritanopia.

Colorblindness models obtained with image processing are also used to modify images and video for showing them to people with normal view for awareness them like in the works in [11]–[13]

Other image processing and computer vision techniques applied for improving conditions of colorblind people include recognize colors and mention them, label them, or transform them into different colors like is described in the works in [4], [12]–[18]

The classification of color like is required in colorblind people assistants is a field of interest in applications of Artificial Intelligence (AI) and the majority of them uses Artificial Neural Networks (ANN) and machine learning algorithms with different color spaces that include RGB, HSV and CIELab, among others, but this represents a problem since computational power is spent it in the transformation of colors even that with today algorithms AI classification can be performed from native spaces like RGB representation, like is described in [19]

Despite the reached results with ANN in the classification of colors with an efficiency of 99.8% like is described in [20], in that research there is required the training of several numerical parameters even with a lower number of layers, which are used in functions that require to be evaluated in several stages one

per layer, for example in the work [20] two hidden layers and transformation to CIELab space are used.

Other faster implementations use Machine Learning Algorithms that reached 92% efficiency like is described in [21] using Particle Swarm Optimization (PSO).

In this work is proposed the use of PSO for optimizing equations that classify colors in the Ishihara plates so that be highlighted for assisting the colorblind while identifying colors in the confusing region, The PSO algorithm was selected since obtains the best results with the used line equations described in [22].

The equations for classifying colors are described in the previous research of color classification and those equations are also used in this project. The equations are evaluated in a single stage and have only 6 numerical parameters per equation like is described in [19], , nevertheless for this work is added a variation of light in the data set for improving the results.

## II. THEORETICAL FRAMEWORK

In this section are presented the acquisition and representation of digital images, the support for the optimizing color equations and the PSO algorithm in order to explain the proposed methodology.

### A. Acquisition and representation of digital images

Digital images are obtained by acquiring voltage signals from electronic cells that react to the quantity of light (Fig. 3), the majority of digital cameras has those cells divided into tree levels of frequency one per basic color Red, Green and Blue (RGB) also known as the RGB representation which is a tern  $(R, G, B) \in R^3$  [23].

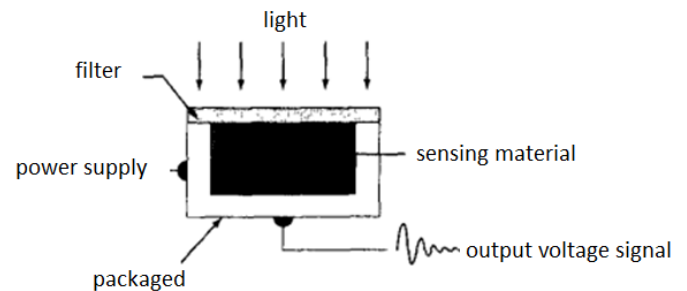


Fig. 3. Camera cells for digital image acquisition taken from [23].

The camera acquisition of those signals are arranged in three matrices one per color (Fig. 4), with numerical values typically of  $2^8$  given from the resolution of its analog digital converter, the sizes of those matrices (height and width) depend on the number of pixels or cells used for acquiring the image [23].

### B. Color Equations

There are several color representations for images and the majority of them are terns in the form  $(A_1, A_2, A_3) \in R^3$ , the native representations are those who are obtained directly from the camera like RGB and YUV [19].

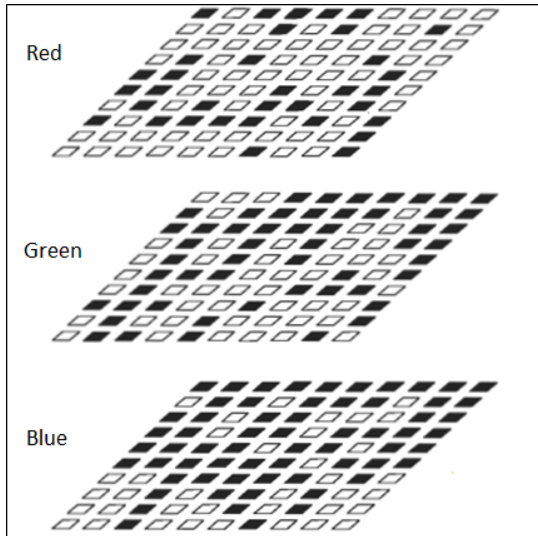


Fig. 4. Camera matrices for RGB color representation [23].

The transformed representations are those generated from this native models by using several non-linear transformations that in HD images could result in a slow transformation. These transformations are used in different applications for example simplifying color classification, with its most common examples HSV, HSI, CIE Lab YIQ and YCbCr [19].

Transformed representations of color are used because its simplicity when classifying colors, even that today machine learning techniques can classify colors without any transformation, examples the entire color spectrum using native RGB and transformed HSV are shown in the Fig. 5 and Fig 6. [19].

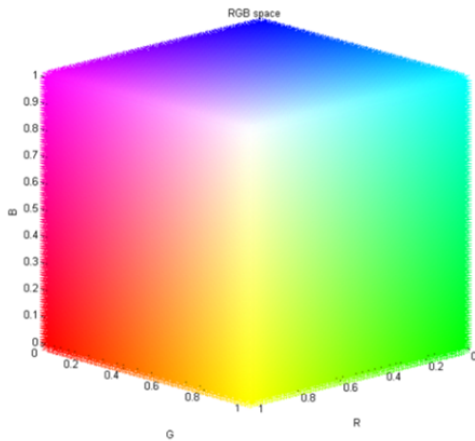


Fig. 5. RGB color spectrum [19].

The complexity separating colors in RGB is greater than in HSV since in RGB are required complex surfaces for limiting the colors while in HSV are required only planes, however if the RGB space is projected into three different planes RG, RB and GB then is possible to separate colors using only three lines per color like is described in [19].

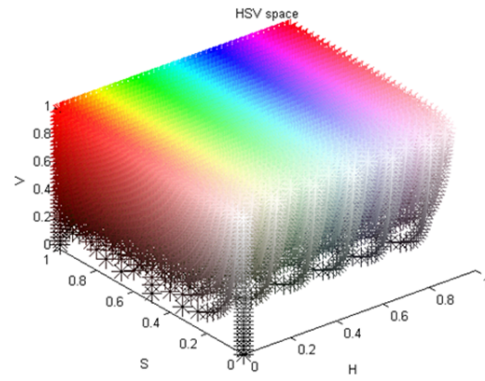


Fig. 6. HSV color spectrum [19].

Classification of colors, require six parameters or two coefficients per line (slope and offset of the line) associated to each plane like is shown in equations 1, 2 and 3 [22].

$$r \leq \alpha_1 \cdot g + \beta_1 \quad (1)$$

$$g \leq \alpha_2 \cdot b + \beta_2 \quad (2)$$

$$\gamma_1 \leq r \leq \gamma_2 \quad (3)$$

where  $\alpha_1$ ,  $\beta_1$ ,  $\alpha_2$ ,  $\beta_2$ ,  $\gamma_1$  and  $\gamma_2$  are the numerical parameters to optimize.  $\alpha_i$  are the related slopes,  $\beta_i$  are the required offsets, and  $\gamma_i$  is the desired range for luminosity for applying those lines with  $i = 1, 2$ .

### C. PSO algorithm

Nowadays there are several variants of PSO but most used PSO uses an inertial coefficient proposed in 1998 as a modification of the original algorithm proposed in 1995 by Kennedy and Eberhart, this is inspired in fish schooling and flock of birds [24].

The PSO describes the behavior of  $n$  particles with position for every  $i$  particle given by  $\vec{X}_i = [x_{i,1}, x_{i,2}, \dots, x_{i,D}]$  where particles movement is performed in a search space of  $D$  dimensions. The  $\vec{X}_i$  positions are initialized to  $i$  random arrays  $\vec{\psi}_i$  in ranges of a lower limit ( $L_l$ ) and an upper limit ( $L_u$ ) i.e.  $[L_l, L_u]$ .

The speed of the particles is described by  $\vec{V}_i = [v_{i,1}, v_{i,2}, \dots, v_{i,D}]$ , which is used for modifying the  $\vec{X}_i$  position and is initialized to a value of zero for each component.

Cost of particles ( $C_i$ ) is determined with cost function  $f(\vec{X}_i)$ , and then some positions must be identified, global best position ( $\vec{P}_G$ ) or the position with the best cost reached ( $G_B$ ), and the best positions known per particle ( $\vec{P}_i$ ) based on ( $G_{P,i}$ ), its best cost obtained [22].

Position of particles it is updated every iteration with its speed, which is determined with the equation 4 using equation 5 [22].

$$\vec{V}_i = w \cdot \vec{V}_i + c_1 \cdot \vec{R}_1 \cdot (\vec{P}_G - \vec{X}_i) + c_2 \cdot \vec{R}_2 \cdot (\vec{P}_i - \vec{X}_i) \quad (4)$$

$$\vec{X}_i = \vec{X}_i + \vec{V}_i \quad (5)$$

where  $w$ , is the inertial coefficient used in range  $[0.9, 1.2]$ , like in the original algorithm, the  $c_1$  and  $c_2$  parameters are personal and social coefficients respectively, which are usually set to their upper limit (2),  $\vec{R}_1$  and  $\vec{R}_2$  are random arrays in ranges  $[0, 1]$  with size  $1 \times D$ .

The position of every  $i$  particle is updated every iteration until a desired cost is reached or a maximum number of iterations.

### III. METHODOLOGY

Like in any other supervised machine learning algorithm, the data set is generated labeling every sample, in order to perceive the numbers in the Ishihara plates, the colors used for classification are: Red, Orange, Green and Brown.

These colors are grouped in two categories:

1) *Protano*: For detecting colors that cannot perceive people with protanopia or protanomaly (red and orange).

2) *Deuterano*: For detecting colors that cannot perceive people with deuteranopia or deuteranomaly (green and brown).

The samples are generated with images of Ishihara plates denoted by  $I$ , which are modified highlighting positives samples with white color and negatives remaining equally or in black, like is shown in Fig. 7, the modified image is denoted by  $M$ , and is reduced in bright to avoid false marks in white colors.

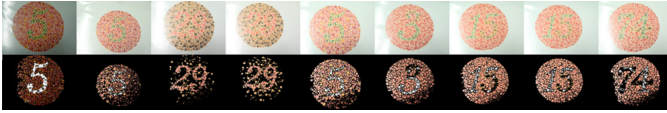


Fig. 7. Labeled Ishihara plates for generation of data set.

Input data of the data set  $DI_i$  are  $IR_i$ ,  $IG_i$  and  $IB_i$  the R, G and B values of the input pixel taken from the original image  $I$ , and the sample is marked as positive in the data set output  $DO_i$  if  $MR_i$ ,  $MG_i$  and  $MB_i$  the R, G and B values of the same pixel but taken from the  $M$  matrix, i.e. are equal to 255, corresponding to a white color, like is described in equation 6.

$$DO_i = \begin{cases} 1 & MR_i = MG_i = MB_i = 255 \\ 0 & \text{else} \end{cases} \quad (6)$$

Cost function for color classification is the Mean Absolute Error (MAE) between  $DO_i$  and the output image  $OI$ , obtained from processing each image with the equations 1,2,3, one equation per each color (in this case the colors red, orange, green and brown). The cost function is described in equation 7.

$$f(\vec{X}_i) = \sum_{i=1}^D |DO_i - OI_i| \quad (7)$$

### IV. RESULTS

The results are obtained with a data set with 30 images per color for a total of 120 images labeled with the four colors. The optimizing of equations 1, 2 and 3 is performed by modifying the original data set with four different levels of light multiplying  $I$  per 0.8, 0.9, 1.1 and 1.2 so that images get better robustness against light variations. These variations of light increase the data set to 480 images.

In order to reach real time execution, images are processed with optimized equations using parallel computing dividing its execution in the number of pixels of each image.

The entire process is made in Matlab<sup>TM</sup> in a Windows 10 computer with processor Intel(R) Core(TM) i7-6700 CPU @ 3.40GHz, 16.0 GB RAM, video card NVIDIA GeForce GTX 970 and a camera with 960x720 pixels resolution.

The PSO algorithm parameters used are shown in table I.

TABLE I  
PSO PARAMETERS USED FOR TRAINING.

Parameter	Value
$D$	6
$[L_l, L_u]$	$[-1, 1]$
max speed	0.4
min speed	-0.4
particles	300
iterations	1000
$c_1$	2
$c_2$	2
$w$	0.9

The training results are shown in Fig. 8 for red color, Fig. 9 for orange color, Fig. 10 for green color and Fig. 11 and for brown color.

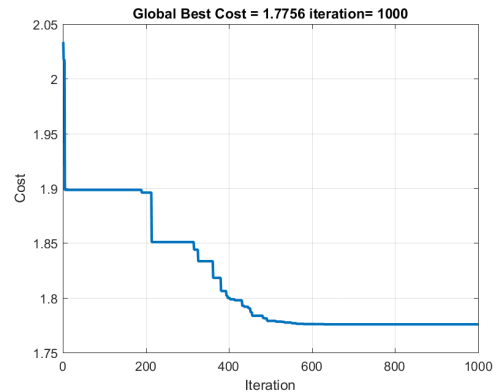


Fig. 8. Training results for red color.

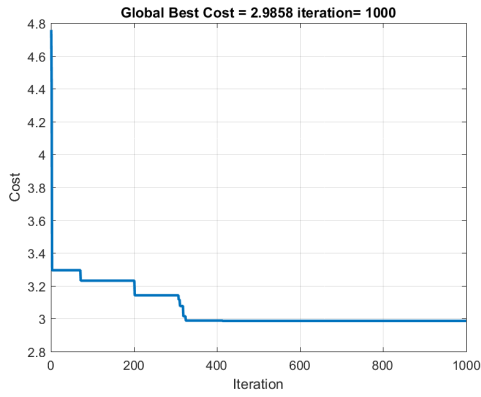


Fig. 9. Training results for orange color.

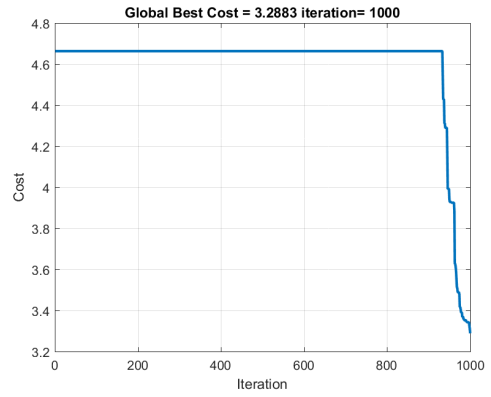


Fig. 11. Training results for brown color.

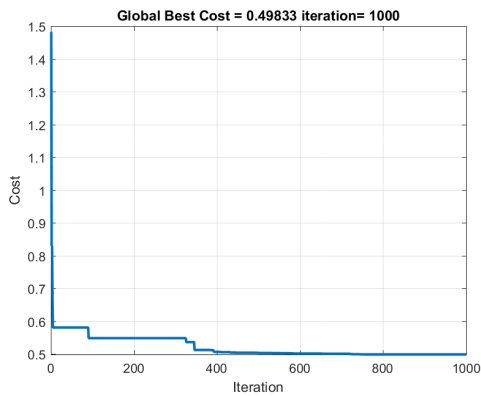


Fig. 10. Training results for green color.

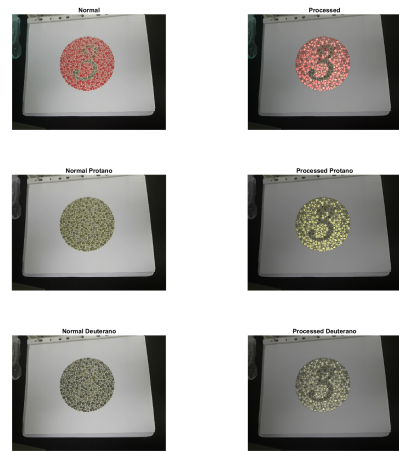


Fig. 12. Comparative of results recoloring Ishihara Plate number 3.

The training is validated using 10 cross fold validation per color obtaining the results described in table II.

TABLE II  
10 CROSS FOLD VALIDATION RESULTS PER COLOR.

Color	Correct Classifications
Red	99.8222%
Orange	99.7013%
Green	99.9457%
Brown	99.6243%

The images of the algorithm are implemented and tested with images directly taken from camera of Ishihara plates. The results are shown in figures 12-14, these obtained images are processed with a speed of 10.5295 frames per second.



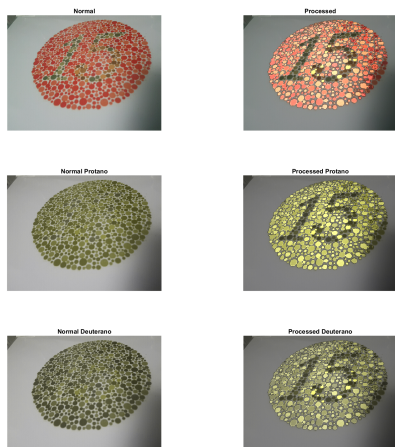


Fig. 13. Comparative of results recoloring Ishihara Plate number 15.

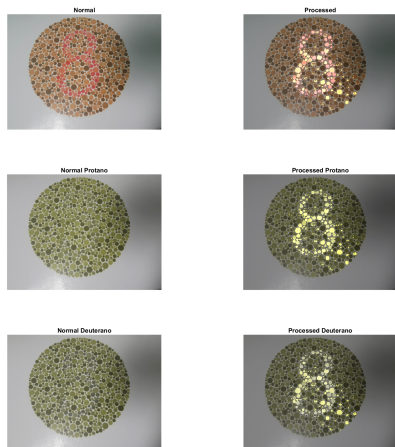


Fig. 14. Comparative of results recoloring Ishihara Plate number 8.

## V. CONCLUSIONS

This paper presents an application of proposed equations for color classification directly from RGB space, applied in the detection of colors in the confusing region for colorblind people.

Those equations are optimized for classifying desired colors using PSO algorithm, the classification reaches 99.8222% for red colors, 99.7013% for orange colors, 99.9457% for green colors and 99.62% for brown colors. The efficiency is validated with 10 cross fold validation. The reached results are similar to the obtained using ANN (99.8%), but using the proposed equations is possible to get a maximum complexity of  $O(n^2)$  when parallel computing is not used, since 2 loops are required (height and width of the image) for highlighting the

positive samples detected in the image. While ANN technique increases complexity depending on the number of layers used.

The results are tested in an Ishihara plates test, comparing the results using a colorblindness recoloring model so that trichomat can perceive the suppressed difficulties.

1) *Future Work*: The results of the work proposed must be tested with colorblind people in not controlled conditions so that be evaluated the quality and safety of the proposed solution and if it could help people to pass the Ishihara test.

## REFERENCES

- [1] J. A. García-Porrero, J. M. Hurlé, and J. García-Porrero Alonso, *Anatomía humana*. McGraw-Hill/Interamericana de España, 2005.
- [2] K. D. Tanaka, "A colour to birds and to humans: why is it so different?," *Journal of Ornithology*, vol. 156, pp. 433–440, dec 2015.
- [3] Colorblindor, *Color Blind Essentials*. 2018.
- [4] Q. U. Dody, L. R. M. Tati, M. Richard, P. G. Andika, and N. A. Tauhid, "RGB Color Cluster and Graph Coloring Algorithm for Partial Color Blind Correction," *MATEC Web of Conferences*, vol. 255, p. 01002, jan 2019.
- [5] S. S. Deeb, "Molecular genetics of color-vision deficiencies.," *Visual neuroscience*, vol. 21, no. 3, pp. 191–6, 2004.
- [6] C. Kato, "Comprehending Color Images for Color Barrier-Free Via Factor Analysis Technique," in *2013 14th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, pp. 478–483, IEEE, jul 2013.
- [7] H. Marey, N. Semary, and S. Mandour, "Ishihara Electronic Color Blindness Test: An Evaluation Study," *Ophthalmology Research: An International Journal*, vol. 3, pp. 67–75, jan 2015.
- [8] T. Bek and M. Kandi, "Quantitative anomaloscopy and optical coherence tomography scanning in central serous chorioretinopathy.," *Acta ophthalmologica Scandinavica*, vol. 78, pp. 632–7, dec 2000.
- [9] J. Velázquez, A. Ferrero, A. Pons Aglio, J. Campos Acosta, M. L. Hernanz, E. Borreguero, and B. Bernard, "Análisis cromático de fuentes oleds," 2016.
- [10] H. Brettel, F. Viénot, and J. D. Mollon, "Computerized simulation of color appearance for dichromats," *Journal of the Optical Society of America A*, vol. 14, p. 2647, oct 1997.
- [11] Bo Liu, Meng Wang, Linjun Yang, Xiuqing Wu, and Xian-Sheng Hua, "Efficient image and video re-coloring for colorblindness," in *2009 IEEE International Conference on Multimedia and Expo*, pp. 906–909, IEEE, jun 2009.
- [12] J.-B. Huang, C.-S. Chen, T.-C. Jen, and S.-J. Wang, "Image recolorization for the colorblind," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1161–1164, IEEE, apr 2009.
- [13] C.-R. Huang, K.-C. Chiu, and C.-S. Chen, "Temporal Color Consistency-Based Video Reproduction for Dichromats," *IEEE Transactions on Multimedia*, vol. 13, pp. 950–960, oct 2011.
- [14] E. Tanuwidjaja, D. Huynh, K. Koa, C. Nguyen, C. Shao, P. Torbett, C. Emmenegger, and N. Weibel, "Chroma," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '14 Adjunct*, (New York, New York, USA), pp. 799–810, ACM Press, 2014.
- [15] M. Chung and H. Choo, "A Real-time Color-matching Method Based on Smartphones For Color-blind People," *UBICOMM 2014, The Eighth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, no. c, pp. 184–188, 2014.
- [16] R. Harwahu and R. F. Sari, "Implementing Speech Feature for Embedded System to Support Color Blind People," in *2011 First International Conference on Informatics and Computational Intelligence*, pp. 256–261, IEEE, dec 2011.
- [17] C.-L. Lai, S.-W. Chang, and J. Sheen, "An integrated portable vision assistant agency for the visual impaired people," in *2009 IEEE International Conference on Control and Automation*, pp. 2311–2316, IEEE, dec 2009.
- [18] T. Ohkubo, K. Kobayashi, K. Watanabe, and Y. Kurihara, "Development of a time-sharing-based color-assisted vision system for persons with color-vision deficiency," in *Proceedings of {SICE} {Annual} {Conference} 2010*, pp. 2499–2503, 2010.

- [19] M. Rivera Montes, A. Padilla Díaz, and J. C. Ponce Gallegos, *Comparative between RGB and HSV color representations for color segmentation when it is applied with artificial neural networks*, pp. 620–638. ALFA-OMEGA, 2016.
- [20] U. Saeed, S. Ahmad, J. Alsadi, D. Ross, and G. Rizvi, “Implementation of neural network for color properties of polycarbonates,” in *AIP Conference Proceedings*, vol. 1593, pp. 56–59, American Institute of Physics, feb 2014.
- [21] J. A. Nasiri, H. S. Yazdi, M. A. Moulavi, M. Rouhani, and A. E. Shargh, “A PSO tuning approach for lip detection on color images,” in *Proceedings - EMS 2008, European Modelling Symposium, 2nd UKSim European Symposium on Computer Modelling and Simulation*, pp. 278–282, IEEE, sep 2008.
- [22] M. Montes, A. Padilla, J. Canul, J. Ponce, and A. Ochoa, “Comparative of Effectiveness When Classifying Colors Using RGB Image Representation with PSO with Time Decreasing Inertial Coefficient and GA Algorithms as Classifiers,” in *Fuzzy Logic Augmentation of Neural and Optimization Algorithms: Theoretical Aspects and Real Applications*, pp. 527–546, 2018.
- [23] R. C. Gonzalez and R. E. R. E. Woods, *Digital image processing*. Prentice Hall, 2008.
- [24] A. Lin, W. Sun, H. Yu, G. Wu, and H. Tang, “Global genetic learning particle swarm optimization with diversity enhancement by ring topology,” *Swarm and Evolutionary Computation*, 2018.