

Advances in Applications of Computer Networks

Research in Computing Science

Series Editorial Board

Editors-in-Chief:

Grigori Sidorov, CIC-IPN, Mexico
Gerhard X. Ritter, University of Florida, USA
Jean Serra, Ecole des Mines de Paris, France
Ulises Cortés, UPC, Barcelona, Spain

Associate Editors:

Jesús Angulo, Ecole des Mines de Paris, France
Jihad El-Sana, Ben-Gurion Univ. of the Negev, Israel
Alexander Gelbukh, CIC-IPN, Mexico
Ioannis Kakadiaris, University of Houston, USA
Petros Maragos, Nat. Tech. Univ. of Athens, Greece
Julian Padget, University of Bath, UK
Mateo Valero, UPC, Barcelona, Spain
Rafael Guzmán, Univ. of Guanajuato, Mexico

Editorial Coordination:

Alejandra Ramos Porras
Carlos Vizcaino Sahagún

Research in Computing Science es una publicación trimestral, de circulación internacional, editada por el Centro de Investigación en Computación del IPN, para dar a conocer los avances de investigación científica y desarrollo tecnológico de la comunidad científica internacional. **Volumen 148, No. 2**, febrero de 2019. *Certificado de Reserva de Derechos al Uso Exclusivo del Título* No.: 04-2005-121611550100-102, expedido por el Instituto Nacional de Derecho de Autor. *Certificado de Licitud de Título* No. 12897, *Certificado de licitud de Contenido* No. 10470, expedidos por la Comisión Calificadora de Publicaciones y Revistas Ilustradas. El contenido de los artículos es responsabilidad exclusiva de sus respectivos autores. Queda prohibida la reproducción total o parcial, por cualquier medio, sin el permiso expreso del editor, excepto para uso personal o de estudio haciendo cita explícita en la primera página de cada documento. Distribuida por el Centro de Investigación en Computación, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othón de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, México, D.F. Tel. 57 29 60 00, ext. 56571.

Editor responsable: Grigori Sidorov, RFC SIGR651028L69

Research in Computing Science is published by the Center for Computing Research of IPN. **Volume 148, No. 2**, February 2019. The authors are responsible for the contents of their articles. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of Centre for Computing Research.

Advances in Applications of Computer Networks

**Arnoldo Díaz Ramírez
Carlos M. Tavares Calafate
Verónica Quintero Rosas
Vidblain Amaro Ortega
Juan Pablo García Vázquez (eds.)**



Instituto Politécnico Nacional
"La Técnica al Servicio de la Patria"



Instituto Politécnico Nacional, Centro de Investigación en Computación
México 2019

ISSN: 1870-4069

Copyright © Instituto Politécnico Nacional 2019

Instituto Politécnico Nacional (IPN)
Centro de Investigación en Computación (CIC)
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal
Unidad Profesional “Adolfo López Mateos”, Zacatenco
07738, México D.F., México

<http://www.rcs.cic.ipn.mx>

<http://www.ipn.mx>

<http://www.cic.ipn.mx>

The editors and the publisher of this journal have made their best effort in preparing this special issue, but make no warranty of any kind, expressed or implied, with regard to the information contained in this volume.

All rights reserved. No part of this publication may be reproduced, stored on a retrieval system or transmitted, in any form or by any means, including electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the Instituto Politécnico Nacional, except for personal or classroom use provided that copies bear the full citation notice provided on the first page of each paper.

Indexed in LATINDEX, DBLP and Periodica

Electronic edition

Editorial

This volume of the “Research in Computing Science” journal consists of selected papers related to the computer networks theory and their applications. The relevance of computer networks in modern applications is growing at a fast pace. The Internet of Things and Cyber-Physical Systems, which are enabling technologies for the Industry 4.0 are based on the use of wireless communications. In addition, Smart Cities, Mobile and Ubiquitous Computing, Social Networks, 4G and 5G are prominent technologies that rely on the existence of communication technologies.

This volume contains twelve papers, which were carefully chosen by the editorial board based on the at least two reviews by the members of the technical program committee. The reviewers took into account the originality, scientific contribution to the field, soundness and technical quality of the papers. The acceptance rate was of 52% among the received submissions.

The papers of this volume address interesting topics regarding computer networks applications. The accepted papers covers topics ranging from eHealth, Internet of Things, Cloud Computing, Machine Learning, eLearning and image processing.

I would like to thank the members of Department of Computer Systems of the Instituto Tecnológico de Mexicali by their valuable support, the Organizing Committee, the Technical Program Committee. In addition, I would also like to thank to the National Polytechnic Institute (IPN) and its Center for Computing Research (CIC) for the given support in order to achieve the success of this volume.

The entire submission, reviewing, and selection process, as well as preparation of the proceedings, were supported free of charge by the EasyChair system (www.easychair.org).

Arnoldo Díaz Ramírez
Tecnológico Nacional de México,
Instituto Tecnológico de Mexicali,
México
Guest Editor

December 2018

Table of Contents

	Page
Circuito Cuántico Shor Qiskit	9
<i>Jesús Álvarez Cedillo, Raúl Junior Sandoval, Teodoro Álvarez Sánchez, Adriana Nava Vega, Jacobo Sandoval Gutiérrez</i>	
Diseño de un robot rescatista para terremotos en México	31
<i>Jesús Álvarez Cedillo, Teodoro Álvarez Sánchez, Raúl Junior Sandoval, Jacobo Sandoval Gutiérrez, Adriana Nava Vega</i>	
Detección de sonidos torácicos aplicando detección de actividad de voz	43
<i>Julio A. Valdez, Pedro Mayorga, Gilberto Chavez, Christopher Druzgalski, Vesna Zeljkovic</i>	
RFID Tag for Toll-Collection Systems	55
<i>Sergio Buenrostro Rocha, Roberto Herrera Charles, José Luis Medina Monroy, Andrés Calvillo Téllez</i>	
Sistema con base en el Internet de las cosas para el control de riego en la agroindustria	63
<i>Brenda Lilia Castro Coronado, Arnoldo Díaz-Ramírez, Verónica Quintero Rosas, Mario Alberto Camarillo Ramos, Jorge Antonio Atempa Camacho</i>	
Sujeción de objetos virtuales mediante un gesto manual usando Kinect y software libre	77
<i>Alexis Jair Urquijo-Brito, Abelardo Rodríguez-Leon, Juan Carlos Prince-Avelino, Guillermo E. Ovanco-Chacon, Carlos J. Genis-Triana</i>	
Diseño de una aplicación móvil para la solución de ecuaciones cuadráticas en una variable	87
<i>Oscar Enrique Callejas Melgoza, María Guadalupe Amado Moreno, Kevin Alberto Luna González, Raul Lara Granillo, Manuel Ernesto Robles Hernández</i>	
A Preliminary Approach for Finding Linked Determinants between Diabetes and Caries using Genetic Algorithms	95
<i>Laura A. Zanella-Calzada, Carlos E. Galván-Tejada, Nubia Chavez-Lamas, Miguel A. Cid-Baez, M. del Carmen Gracia-Cortés</i>	

Antena de parche semicircular con elemento parásito para CanSat.....	109
<i>Cruz Ángel Figueroa Torres, José Luis Medina Monroy, Andrés Calvillo Téllez</i>	
Componente del Internet de las cosas para detectar patrones de deambulaje en pacientes con demencia.....	121
<i>Hector Villarreal, Mayra Lizarraga, Arnoldo Díaz-Ramírez, Verónica Quintero Rosas, Juan Pablo García-Vázquez</i>	
Sistema de clasificación SVM de señales electromiográficas extraídas en un sistema embebido	135
<i>Luis Daniel Reyes Crusaley, J. R. Cárdenas-Valdez, Gilberto Enrico Vázquez, Manuel de Jesús García Ortega, Andrés Calvillo Téllez</i>	
Telemedicine Networks: Integrating Different Health Technologies towards a Common Goal.....	143
<i>Roberto Conte-Galvan, Arturo Serrano-Santoyo, Veronica Rojas-Mendizabal</i>	

Circuito Cuántico Shor Qiskit

Jesús Álvarez Cedillo¹, Raúl Junior Sandoval¹, Teodoro Álvarez Sánchez²,
Adriana Nava Vega³, Jacobo Sandoval Gutiérrez⁴

¹ Instituto Politécnico Nacional, UPIICSA, Ciudad de México, México

jalvarez@ipn.mx, rsandova@ipn.mx

² Instituto Politécnico Nacional, CITEDI, Baja California, México

talvarez@citedi.mx

³ Universidad Autónoma de Baja California, Campus Tijuana, México

adriana.nava@uabc.edu.mx

⁴ Universidad Autónoma Metropolitana, Unidad LERMA, Estado de México, México

jacobosandoval@gmail.com

Resumen. La computación cuántica es un campo de investigación dinámico y apasionante que reúne el conocimiento de diversas disciplinas, incluida la física cuántica, la informática y la teoría clásica de la complejidad. Durante la década de 1980, el trabajo de David Deutsch ayudó a definir la teoría de la computación cuántica, actualmente IBM cuenta con una computadora cuántica real. En este artículo se utilizó la plataformas de IBM QISK para realizar la implementación desde cero del algoritmo de Shor. Este algoritmo permite la factorización primaria, se tomó como referencia este algoritmo debido a que ha sido muy criticado en la literatura por mostrar la supuesta intratabilidad de la factorización y la aplicabilidad universal de la teoría de la complejidad clásica, además que poner en duda la seguridad no reversible de los métodos de cifrado actuales.

Palabras clave: computación cuántica, algoritmo de Shor, circuito cuántico, lenguaje ensamblador cuántico.

Shor Qiskit Quantum Circuit

Abstract. Quantum computing is a dynamic and exciting field of research that brings together the knowledge of various disciplines, including quantum physics, computer science and the classical theory of complexity. During the 1980s, David Deutsch's work helped define the theory of quantum computing. Currently, IBM has a real quantum computer. In this article, the IBM QISK platforms were used to implement the Shor algorithm from scratch. This algorithm allows primary factorization. This algorithm was taken as a reference, because it has been highly criticized in the literature for showing the supposed intractability of the factorization and the universal applicability of the theory of classical complexity, in addition to questioning the non-reversible safety of current encryption methods.

Keywords: quantum computing, Shor algorithm, quantum circuit, quantum assembly language.

1. Introducción

La unidad mínima de información de las computadoras clásicas son los bits, estos sólo puede tener un estado en un momento determinado en el tiempo. Si hay varios valores diferentes para ser examinados, cada uno debe procesarse por separado. A diferencia de este tipo de computadoras, las computadoras cuánticas codifican la información usando dígitos binarios cuánticos llamados Qubits. La información está representada por su estado cuántico, y su giro o SPIN [1]. Los Qubits pueden tener de acuerdo al principio de superposición el valor de uno o cero, o ambos.

Como resultado del control de todos los estados al mismo tiempo, las computadoras cuánticas rompen con la no computabilidad y reducen la complejidad computacional de algunos algoritmos, el paralelismo inherente a las computadoras cuánticas las hace muy potentes para procesar información. Por desgracia, aún existen varios obstáculos importantes que los investigadores buscan vencer para convertirlas en una una de su homólogos clásicos [2].

Dos principios de la mecánica cuántica desempeñan un papel importante en el funcionamiento de una computadora cuántica, superposición y entrelazamiento [4].

La superposición es un fenómeno mecánico cuántico que permite a los electrones para evitar que se limita a un solo estado. Por lo que pueden tener alguna combinación de los dos, con diferentes probabilidades para cada estado.

Los Qubits son capaces de mantener su condición de superposición hasta que se realiza una medición del sistema, en cuyo punto el sistema se colapsa y el electrón se limita a un estado o el otro y se dice que se pierde la estabilidad.

Este es el principio de superposición lo que hacen tan potentes a esta computadoras y con esto están revolucionando la informática. Debido a que los Qubits presentan una superposición de estados, los cálculos aplicados a cada estado es capaz de explorar todos los valores de forma simultánea. Lo anterior representa un nivel alto de procesamiento paralelo físicamente posible [11].

El entrelazamiento es una condición que puede ser inducida para vincular un par de partículas cuánticas las cuales aunque no tienen estados cuánticos independientes, dependerán de la otra y afectan su estado cuántico. El entrelazamiento se conserva a pesar de la distancia que los separa. Aunque este avance consiste en una arquitectura de magnifico poder y paralelismo existen problemas significativos que deben superarse. Un problema a resolver es determinar el mejor método para ser utilizado en la fabricación del hardware.

En este momento existen dos tipos de computadoras cuánticas la IBM q y D-wave, diferentes por su modo de operación. Los Qubits pueden ser representados por átomos, iones, electrones o fotones [5].

Los mecanismos para congelar al Qubit son diversos y cada uno tiene ventajas y desventajas, los más usados son los siguientes:

1. Trampas de iones: para su captura se utilizan campos ópticos o magnéticos, o alguna combinación de los dos.
2. Ondas de luz para atrapar partículas.
3. Conversión de Qubits a puntos cuánticos: alternativamente, los átomos de impurezas que se encuentran en un semiconductor también se pueden utilizar para la representación de Qubits.

Cualquier método elegido en la construcción de la computadora cuántica, debe proporcionar una solución práctica a cuestiones tales como la de coherencia y corrección de errores. La de-coherencia es la tendencia a que el estado cuántico de una partícula a decaer a un estado arbitrario como un resultado de la interacción con su entorno [6, 12].

El primer algoritmo que se desarrolló para las computadoras cuánticas fue creado en 1994 por Peter Shor, de los laboratorios Bell. Se desarrolló un algoritmo cuántico de complejidad computacional polinomial que se utiliza para factorizar números muy grandes [8, 9].

Una computadora clásica resuelve este algoritmo con una complejidad computacional exponencial, lo que hace que no puedan usarse para resolver el problema. La mayoría de los algoritmos de criptografía principales se basan en la dificultad para resolver este problema. Es importante mencionar que los algoritmos cuánticos no devuelven una respuesta precisa en la forma en que los algoritmos deterministas clásicos lo hacen, su respuesta tiene cierta probabilidad de ser una solución. El algoritmo de Shor es extremadamente potente, pero todavía tiene la posibilidad de fracasar, y siempre hay una posibilidad de que los factores que devuelve simplemente será el número en sí y uno, que no es muy útil, incluso si existen otros factores.

2. Descripción del método

El diseño de nuestra investigación incluye en primer lugar el desarrollo del algoritmo de Shor usando el lenguaje ensamblador cuántico QASM. En este código se plantean las técnicas de recolección de datos y análisis. En segundo lugar, se realiza el circuito y se simula en la supercomputadora IBM q. En tercer lugar, se obtiene la información necesaria para la comparación con otros códigos y la aportación.

Se realizó comprobación matemática y se comprobó su eficacia y exactitud.

2.1. Diseño del algoritmo

En computación clásica no se conoce ningún algoritmo clásico eficiente para realizar la factorización de una serie de datos de grande. Se dice que un algoritmo es eficiente si su tiempo de ejecución es asimétricamente polinomial en la longitud de su entrada usando bits.

Su algoritmo equivalente es el tamiz cuadrático que tiene una complejidad computacional de

$$O(\exp((\frac{64}{9})^{\frac{1}{3}} N^{\frac{1}{3}} (\ln N)^{\frac{2}{3}}))$$

para factorizar un número binario de N bits.

La multiplicación de números primos grandes es por lo tanto, una función unidireccional, es decir, una función que puede evaluarse fácilmente en una dirección, mientras que su inversión es prácticamente imposible. Las funciones unidireccionales

desempeñan un papel importante en la criptografía y son esenciales para los sistemas criptográficos de clave pública, donde la clave para la codificación es pública y solo la clave para la decodificación permanece en secreto.

En 1978, Rivest, Shamir y Adleman desarrollaron un algoritmo criptográfico basado en el carácter unidireccional de multiplicar dos números primos grandes (100 dígitos), así nació el método RSA y se convirtió en el sistema de clave pública más popular. Se ha demostrado en la literatura que la factorización prima eficiente en una computadora clásica es imposible, Shor propuso un algoritmo eficiente para computadoras cuánticas en 1994.

Se desarrolló un algoritmo secuencial para solucionar este problema en una computadora clásica el algoritmo desarrollado se muestra en el Algoritmo 1. El procedimiento comprueba si el número entero es adecuado para la factorización cuántica, y luego repite el algoritmo de Shor hasta que se haya encontrado un factor.

Algoritmo 1. Algoritmo de Shor clásico.

```
procedimiento shor (número int) {
    int width = ceil (log (número, 2));          // tamaño del número en bits
    qureg reg1 [2 * ancho];                      // primer registro
    qureg reg2 [ancho];                          // segundo registro
    int qmax = 2 ^ ancho;
    int factor;                                  // factor encontrado
    int m; real c;                              // valor medido
    int x;                                       // base de exponenciación
    int p; int q;                               // aproximación racional p / q
    int a; int b;                               // posibles factores de número
    int e;                                       //  $e = x \wedge (q / 2)$  número mod.

    si el número mod 2 == 0 {exit "number debe ser impar"; }
    si testprime (número) {exit "número primo"; }
    si testprimepower (number) {exit "prime power"; };

    {
        // generar base aleatoria
        x = piso (random () * (number-3)) + 2;
        } hasta gcd (x, number) == 1;
        imprimir "elegido al azar x =", x;
        Mix (reg1);                             // Transformada Hadamard
        expn (x, número, reg1, reg2);            // exponenciación modular
        medir reg2;                             // measure 2nd register
        dft (reg1);                             // Transformada de Fourier
        medir reg1, m;                          // medida 1er registro
        Reiniciar;                             // borrar los registros locales

        si m == 0 {                             // falló si se midió 0
            imprimir "cero medido en el 1er registro. Intentar de nuevo ...";
        } else {
            c = m * 0.5 ^ (2 * ancho);           // forma de punto fijo de m
            q = denominador (c, qmax);           // encuentra una aproximación racional
            p = piso (q * m * c + 0.5);
        }
    }
}
```

```

imprimir "medido", m, " , aproximación para", c, "es", p, "/", q;
si q mod 2 == 1 y 2 * q < qmax {           // impar q? intenta expandir p / q
    imprimir "denominador impar, expandiendo por 2";
    p = 2 * p; q = 2 * q; }
si q mod 2 == 1 {                           // falló si es impar q
    imprimir "período impar. Intentar de nuevo ...";
} else {
    imprimir "el posible período es", q;
    e = powmod(x, q / 2, número);           // calcular los candidatos para
    a = (e + 1) número de modulación;       // posibles factores comunes
    b = (e + número-1) número de modulación; // con número
    imprimir x, "^", q / 2, "+ 1 mod", número, "=", a, " , ",
    x, "^", q / 2, "- 1 mod", número, "=", b;
    factor = max(gcd(número, a), gcd(número, b));
}
}
} hasta factor > 1 y factor < number;
número de impresión, "=", factor, "*", número / factor;}

```

3. Implementación cuántica

El algoritmo de Shor busca factorizar un valor dado $M > 0$, que asumimos que es semi-primo $M = pq$ con factores desconocidos. La estrategia es considerar las funciones $f_b(x) = xb \bmod M^2$ potencialmente con varios valores diferentes de $1 < b < M$ y determinar sus períodos en caso de que $\gcd(b, M) = 1$. Cuando se determina que el período es incluso $b^{2\pi} \bmod M = 1$, tenemos $(b^\pi - 1)(b^\pi + 1) \bmod M = 0$, por lo tanto cualquiera $(b^\pi - 1)$ o $(b^\pi + 1)$ debe compartir al menos un factor primo con M .

Si $f^{b\pi} \bmod M \neq -1$, tal factor se puede encontrar utilizando $\gcd(b\pi \pm 1, M)$, de lo contrario, conduce a los factores triviales 1 y M . Cuando se determina que el período es diferente, se prueba otro valor b .

El procedimiento de búsqueda de período se basa en un circuito cuántico (Figura 1), instanciado para un valor dado $1 < b < M$ co primo con M .

El circuito opera en dos registros cuánticos inicializados en 0 con:

1. Un bloque de compuertas Hadamard paralelas en el Registro 1,
2. Un circuito para la exponenciación modular (mod-exp) evaluada con $f(y) = b^y \bmod M$ mediante el mapeo $|y\rangle|0\rangle \rightarrow |y\rangle|f(y)\rangle$, donde y lee el Registro 1 y $f(y)$ escribe para registrar 2; El registro 1 puede modificarse temporalmente, pero debe restaurarse al final,
3. Un circuito para la Transformada Cuántica de Fourier (QFT) en el Registro 1,
4. Un bloque de mediciones paralelas en el Registro 1.

El primer y el último bloque no se pueden optimizar más. Los circuitos QFT se entienden bastante bien y son mucho más pequeños que los circuitos para la exponenciación modular.

Por lo tanto, nuestro enfoque es en los circuitos mod-exp, constan de compuertas reversibles: NOT (N), CNOT (C) y Toffoli (T), que se pueden modelar y optimizar

por completo en términos de lógica booleana. En las implementaciones físicas, las compuertas de Toffoli deben descomponerse en compuertas más pequeñas que son implementables directamente en una tecnología determinada [13].

Los circuitos reversibles para la exponenciación modular comienzan con un inversor en el Registro 2 que cambia el valor de $|000 \dots 0\rangle$ a $|X|000 \dots 1\rangle$, y SE muestran la siguiente estructura:

1. Cada bit (i-ésimo) del Registro habilita un bloque de circuito que multiplica el Registro 2 por $C_i = b^2 \bmod M$ y reduce el resultado del modulo de M.
2. Cuando se conocen b y M, C_i puede pre calcularse sin cálculo cuántico. Por lo tanto, nos referimos a $Cix \bmod M$ blocks.
3. Normalmente se implementan usando circuitos de desplazamiento y adición, y se conocen varios sumadores cuánticos relevantes.

Cada multiplicación modular controlada se implementa tradicionalmente separada. Cuando se trata de circuitos lógicos y cuánticos reversibles, observamos que la co-primidad de C y M hace que $x \rightarrow Cx \bmod M$ sea una transformación reversible [14]. El número de valores co-primos de C es $\varphi(M) = (p-1)(q-1)$, donde $\varphi(M)$ es la Función totient de Euler y da el tamaño de $(\mathbb{Z}/M\mathbb{Z})^\times$ - el grupo multiplicativo de enteros para mod-M.

Para $M = 15$, los circuitos modulares de multiplicación para los ocho valores C co primos se ilustran en la Figura 3.

La figura 4 muestra los circuitos para $f(x) = b^x \bmod 15$, $\gcd(b, 15) = 1$.

Cuando no se conocen p y q , tampoco se debe asumir ningún conocimiento que facilite encontrarlos.

Cuando se trabaja con un valor $M = pq$ pequeño es difícil evitar el uso de los valores conocidos de p y q , pero los resultados obtenidos de esta manera no se escalan necesariamente a valores grandes. Lo mismo puede decirse sobre los resultados producidos a través de una búsqueda exhaustiva.

El procedimiento ajustado se muestra en el Algoritmo 2. El listado se encuentra en Qiskit de IBM.

Algoritmo 2. Algoritmo de Shor cuántico.

```
#-----  
# Importar módulos necesarios  
#-----  
import sys  
from qiskit import QuantumProgram  
import Qconfig  
import Basic_gates  
import math  
from random import randint  
import control_unitaries  
import xlswriter  
#-----  
# variable globales  
#-----
```

```

Counts = 0
A = 0
Ran_Quantum_period_finding = 0
global m
m = 0
#-----
# La función para calcular GCD usando el método de Euclides
# Entrada: dos números a X e Y para los cuales se debe calcular un GCD
# Salida: GCD de dos números dados
#-----
def gcd(x,y):
    while y != 0:
        (x, y) = (y, x % y)
    return x
#-----
# La función para construir el unitario
# Entrada: objeto del programa Quantum, el nombre del circuito y el nombre del registro
cuántico
# Salida: ninguna. El circuito relevante está hecho.
#-----
def cmod(Quantum_program_object,Circuit_name,Quantum_register_name,a):
# Get the circuit and the quantum register by name
    qc = Quantum_program_object.get_circuit(Circuit_name)
    qr = Quantum_program_object.get_quantum_register(Quantum_register_name)

# Construct unitary based on a
    if a == 2:
        qc.cswap(qr[4],qr[3],qr[2])
        qc.cswap(qr[4],qr[2],qr[1])
        qc.cswap(qr[4],qr[1],qr[0])
    if a == 4 or a == 11 or a == 14:
        qc.cswap(qr[4],qr[2],qr[0])
        qc.cswap(qr[4],qr[3],qr[1])
        qc.cx(qr[4],qr[3])
        qc.cx(qr[4],qr[2])
        qc.cx(qr[4],qr[1])
        qc.cx(qr[4],qr[0])
    if a == 7:
        qc.cswap(qr[4],qr[1],qr[0])
        qc.cswap(qr[4],qr[2],qr[1])
        qc.cswap(qr[4],qr[3],qr[2])
        qc.cx(qr[4],qr[3])
        qc.cx(qr[4],qr[2])
        qc.cx(qr[4],qr[1])
        qc.cx(qr[4],qr[0])
    if a == 8:
        qc.cswap(qr[4],qr[1],qr[0])
        qc.cswap(qr[4],qr[2],qr[1])
        qc.cswap(qr[4],qr[3],qr[2])
    if a == 13:
        qc.cswap(qr[4],qr[3],qr[2])
        qc.cswap(qr[4],qr[2],qr[1])

```

```

qc.cswap(qr[4],qr[1],qr[0])
qc.cx(qr[4],qr[3])
qc.cx(qr[4],qr[2])
qc.cx(qr[4],qr[1])
qc.cx(qr[4],qr[0])

#-----
# La función para calcular QFT
# Entrada: circuito, bits cuánticos y cantidad de bits cuánticos
# Salida: ninguna. Circuito creado y guardado
#-----
def qft(Quantum_program_object, Circuit_name, Quantum_register_name,
Smallest_Quantum_register_number, Size_of_QFT):
# Get the circuit and the quantum register by name
qc = Quantum_program_object.get_circuit(Circuit_name)
qr = Quantum_program_object.get_quantum_register(Quantum_register_name)
s = Smallest_Quantum_register_number
for j in range(Size_of_QFT):
    for k in range(j):
        qc.cu1(math.pi/float(2**(j-k)), qr[s+j], qr[s+k])
        qc.h(qr[s+j])

#-----
# La función para encontrar el período usando la computadora Quantum
# Entrada: a y N para la cual se debe calcular el período.
# Salida: período r de la función  $a^x \bmod N$ 
#-----
def period(a,N):
    global Ran_Quantum_period_finding
    Ran_Quantum_period_finding = 1
# Create the first QuantumProgram object instance.
qp = QuantumProgram()
# qp.set_api(Qconfig.APIToken, Qconfig.config["url"])
# TO DO : generalize the number of qubits and give proper security against rogue input.
# Create the first Quantum Register called "qr" with 12 qubits
qr = qp.create_quantum_register('qr', 5)
# Create your first Classical Register called "cr" with 12 bits
cr = qp.create_classical_register('cr', 3)

# Create the first Quantum Circuit called "qc" involving your Quantum Register "qr"
# and the Classical Register "cr"
qc = qp.create_circuit('Period_Finding', [qr], [cr])

# Get the circuit and the registers by name
Shor1 = qp.get_circuit('Period_Finding')
Q_reg = qp.get_quantum_register('qr')
C_reg = qp.get_classical_register('cr')

# Create the circuit for period finding
# Initialize qr[0] to |1>
Shor1.x(Q_reg[0])
# Step one : apply  $a^{**4} \bmod 15$ 

```



```

        Shor1.h(Q_reg[4])
# Controlled Identity on the remaining 4 qubits. Which is equivalent to doing nothing
        Shor1.h(Q_reg[4])
        Shor1.measure(Q_reg[4],C_reg[0])
# Reinitialize to |0>
        Shor1.reset(Q_reg[4])

# Step two : apply a**2 mod 15
        Shor1.h(Q_reg[4])
# Controlled unitary. Apply a mod 15 twice.
    for k in range(2):
        cmod(qp,'Period_Finding','qr',a)
        if C_reg[0] == 1 :
            Shor1.u1(pi/2.0,Q_reg[4])
        Shor1.h(Q_reg[4])
        Shor1.measure(Q_reg[4],C_reg[1])
# Reinitialize to |0>
        Shor1.reset(Q_reg[4])

# Step three : apply 11 mod 15
        Shor1.h(Q_reg[4])
# Controlled unitary. Apply a mod 15
        cmod(qp,'Period_Finding','qr',a)
# Feed forward and measure
        if C_reg[1] == 1 :
            Shor1.u1(pi/2.0,Q_reg[4])
        if C_reg[0] == 1 :
            Shor1.u1(pi/4.0,Q_reg[4])
        Shor1.h(Q_reg[4])
        Shor1.measure(Q_reg[4],C_reg[2])

# Run the circuit
# qp.set_api(Qconfig.APIToken, Qconfig.config['url']) # set the APIToken and API url
simulate = qp.execute(["Period_Finding"], backend="local_qasm_simulator",
shots=1,timeout=500)
simulate.get_counts("Period_Finding")
#print(simulate)
data = simulate.get_counts("Period_Finding")
#print(data)
data = list(data.keys())
#print(data)
r = int(data[0])
#print(r)
l = gcd(2**3,r)
#print(l)
r = int((2**3)/l)
#print(r)
return r

#-----
# La función principal para calcular los factores
# Entrada: el número a factorizar, N

```

```

# Salida: factores del número
#-----
def Factorize_N(N):
    factors = [0,0]
#-----
# Paso 1: Determine la cantidad de bits en función de N;  $n = \lceil \log_2(N) \rceil$ 
#-----
    n = math.ceil(math.log(N,2))
#-----
# Paso 2: comprueba si N es par. En ese caso, devuelve 2 y el número restante como factores
#-----
    If N % 2 == 0:
        factors = [2,N/2]
        return factors
#-----
# Paso 3: compruebe si N tiene la forma  $P^k$ , donde P es un factor primordial. En ese caso
# devuelve P y k.
#-----
# El paso ha sido eliminado con fines de simulación.
#-----
# Paso 4: elija un número aleatorio entre 2 ... (N-1).
#-----
    while True:
        a = randint(2,N-1)
        global A
        A = a
#-----
# Paso 5: Tome GCD de ay N.  $t = \text{GCD}(a, N)$ 
#-----
        t = gcd(N,a)
        if t > 1:
            factors = [t,N/t]
            return factors
#-----
# Paso 6:  $t = 1$ . Por lo tanto, no hay un período común. Encontrar período usando el método
# de Shor
#-----
        r = period(a,N)
        if (r%2 == 0) and (((a**(r/2))+1)%N != 0) and (r != 0) and (r != 8):
            break
        global Counts
        Counts = Counts + 1
        factor_1 = gcd((a**(r/2))+1,N)
        factor_2 = N/factor_1
        factors = [factor_1,factor_2]
        return factors
#-----
# Ejecutando la versión 1 del algoritmo de Shor
#-----
#-----
# Paso 0: Toma la entrada N
#-----

```

```

factors_list = list()
A_used = list()
Ran_QPF = list()
Total_counts = list()
if __name__ == '__main__':
    #global m
    for m in range(100):
        N = 15
        factors_found = Factorize_N(N)
        factors_list.append(factors_found)
        #ws.write(row,col,A)
        #ws.write(row,col+1,Counts)
        #ws.write(row,col+2,factors[0])
        #ws.write(row,col+3,factors[1])
        #ws.write(row,col+4,Ran_Quantum_period_finding)
        #row = row + 1
        #print("The Number being factorized is 15")
        #print("Factors are = ",factors)
        #print("Number of times the quantum circuit did not give correct period =
",Counts)

        #print ("The parameter a used = ", A)
        print("Run ", m)
        A_used.append(A)
        Ran_QPF.append(Ran_Quantum_period_finding)
        Total_counts.append(Counts)
        Counts = 0
        Ran_Quantum_period_finding = 0

    if m == 99:
        wb = xlswriter.Workbook('log.xlsx')
        ws = wb.add_worksheet('Data')
        row = 0
        col = 0
        ws.write(row,col,'a used for factorizing')
        ws.write(row,col+1,'Number of times the quantum circuit did not give correct
period')
        ws.write(row,col+2,'Factor1')
        ws.write(row,col+3,'Factor2')
        ws.write(row,col+4,'Ran_Quantum_period_finding?')
        row = row + 1

        for k in range(100):
            ws.write(row,col,A_used[k])
            ws.write(row,col+1,Total_counts[k])
            ws.write(row,col+2,factors_list[k][0])
            ws.write(row,col+3,factors_list[k][1])
            ws.write(row,col+4,Ran_QPF[k])
            row = row + 1
        wb.close()

```

Después de comprobar su programación se procedió a crear el circuito en IBM *experience* [10]¹. Valores obtenidos de esta manera no se escalan necesariamente a valores grandes. Lo mismo puede decirse sobre los resultados producidos a través de una búsqueda exhaustiva.

4. Resultados

La complejidades de los algoritmos se muestra en la tabla 1. La Figura 1 muestra su curva típica.

Tabla 1. Complejidades de los algoritmos de Shor

Shor clásico	Shor cuántico
$t \sim O(\sqrt{N} \log^3 N)$	$t \sim O(\sqrt{N})$

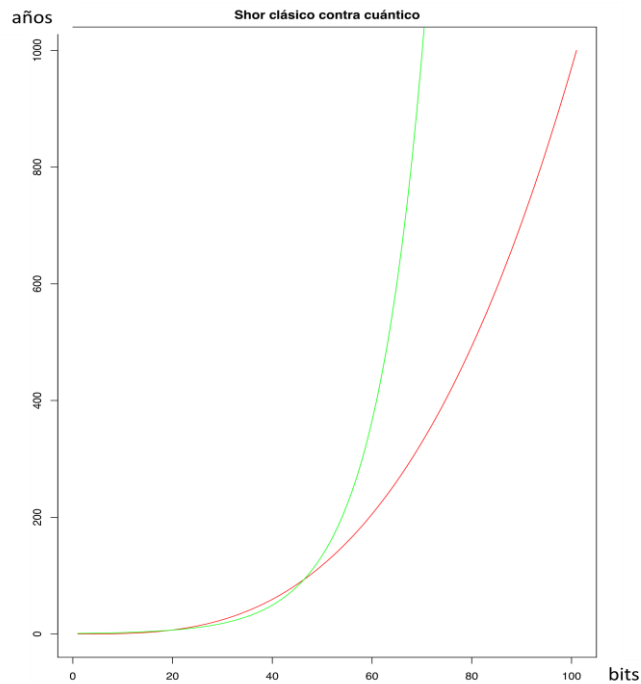


Fig. 1. Comparación de las complejidades computacionales para ambos (años contra bits procesados). Las líneas en color verde representan el algoritmo clásico, las líneas rojas el algoritmo cuántico

La interpretación de los resultados se pueden establecer desde el modelo cuántico a utilizar, los modelos cuánticos son:

¹ <https://quantumexperience.ng.bluemix.net/qx/editor>

1. Modelo basado en compuertas
2. Modelo basado en la maquina de Turín Cuántica [7]
3. Modelo basado en el calculo de Lambda
4. Topología adiabática
5. Modelo basado en medición

Dado que el desarrollo de la investigación esta basada en la creación de los algoritmos de programación establecemos como importante el desarrollo de los procedimientos basándonos en un modelo universal como lo es el modelo basado en compuertas cuánticas (MBCC) [3].

Se realizó la adecuación MBCC del listado 2. Para construir el listado en Qiskit de IBM mostrado en el Algoritmo 3.

Algoritmo 3. Algoritmo de Shor en Qiskit de IBM.

```
//OPENQASM 2.0
IBMQASM 2.0;
include "qelib1.inc";
qreg q[5];
creg c[5];
x q[2];
x q[1];
x q[2];
x q[3];
x q[4];
cx q[3],q[2];
cx q[2],q[3];
cx q[3],q[2];
cx q[2],q[1];
cx q[1],q[2];
cx q[2],q[1];
cx q[4],q[1];
cx q[1],q[4];
cx q[4],q[1];
measure q[1] -> c[1];
measure q[2] -> c[2];
measure q[3] -> c[3];
measure q[4] -> c[4]; }
```

Se realizó la simulación del circuito usando el simulador QX-Studio para definir con más precisión las compuertas a utilizar. Al realizar la simulación se obtuvieron los siguientes valores:

```
-----[Estado cuántico Inicial]-----
(+1.000000,+0.000000) |1101> +
-----
[>>] prediccion en la medida: | 1 | 1 | 0 | 1 |
```

```

-----
[>>] Medida en el registro inicial : | 0 | 0 | 0 | 0 |
-----

-----[ Estado cuántico final]-----
(+1.000000,+0.000000) |1011> +
-----

[>>prediccion en la medida: | 1 | 0 | 1 | 1 |
-----

[>>]Medida en el registro final: | 1 | 0 | 1 | 1 |
-----

Tiempo de ejecución del circuito: +0.000045 sec.
    
```

La figura 1 Muestra el circuito construido por el simulador QX Studio. Con este circuito se realizó el análisis de el comportamiento sobre un registro binario cuántico para determinar y comprobar matemáticamente que los resultados sean coherentes, en la Tabla 3 se muestra el proceso de comprobación por registro de cada valor obtenido en los registros en los 14 pasos observados.

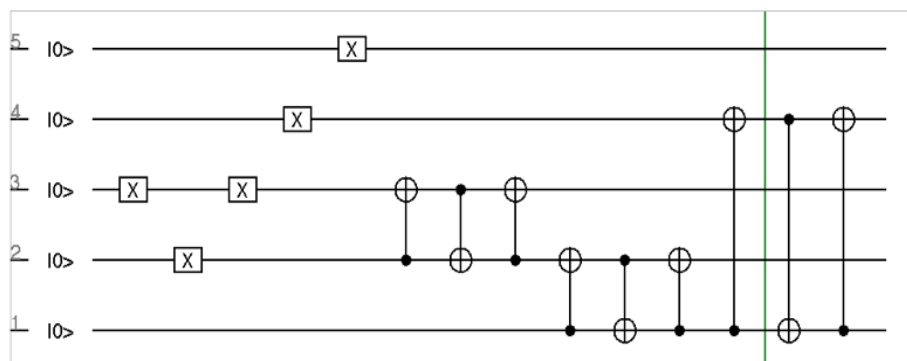
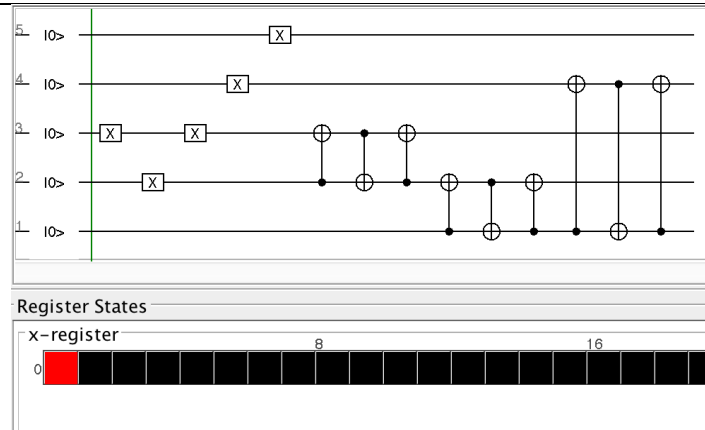


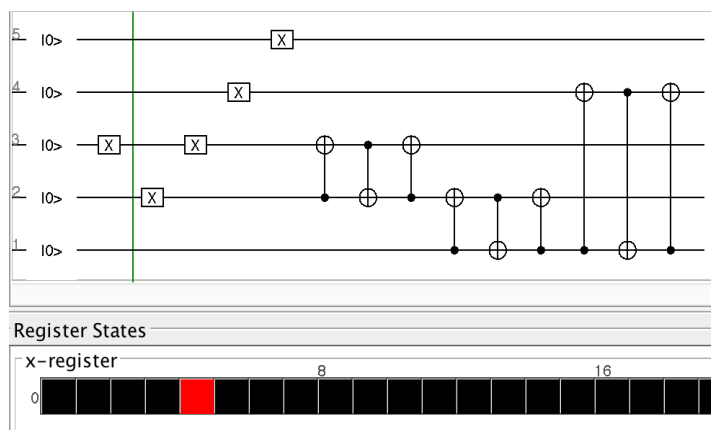
Fig. 2. Circuito cuántico resultante del análisis del código del Algoritmo 3.

Tabla 3 muestra los 14 pasos para comprobar cada compuerta en el circuito cuántico propuesto.

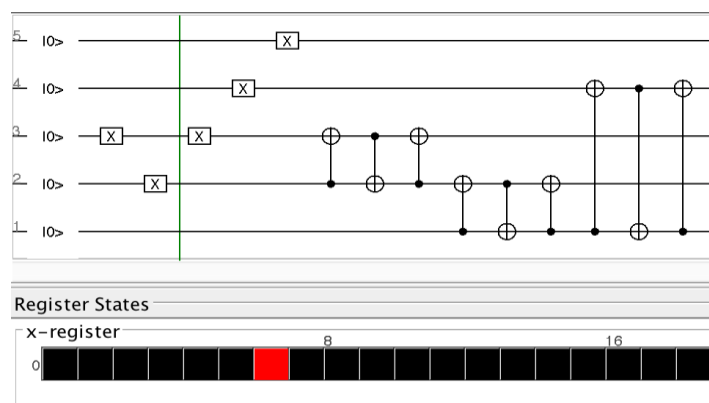
Como último paso, se procedió a implementar las compuertas de acuerdo al listado 3 en la computadora cuántica IBM Q de manera real. La Figura 3 muestra el circuito básico construido en IBM *experience*.



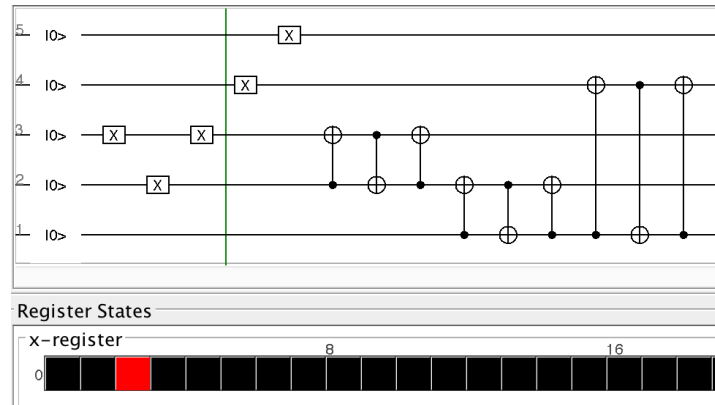
Paso 1



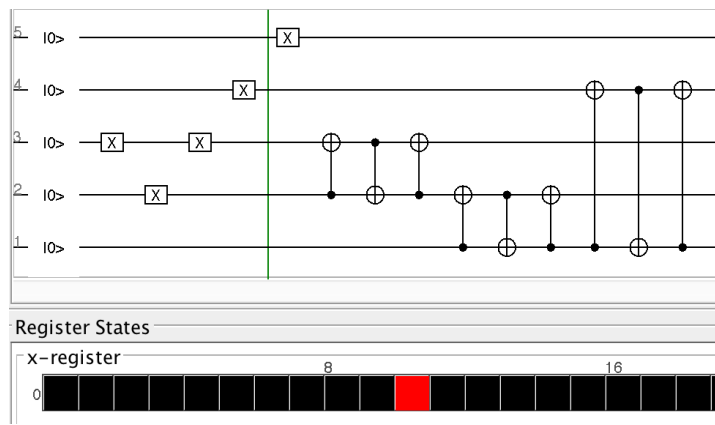
Paso 2



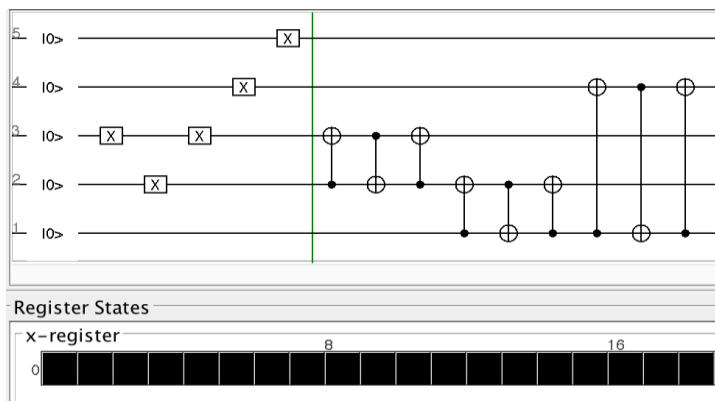
Paso 3



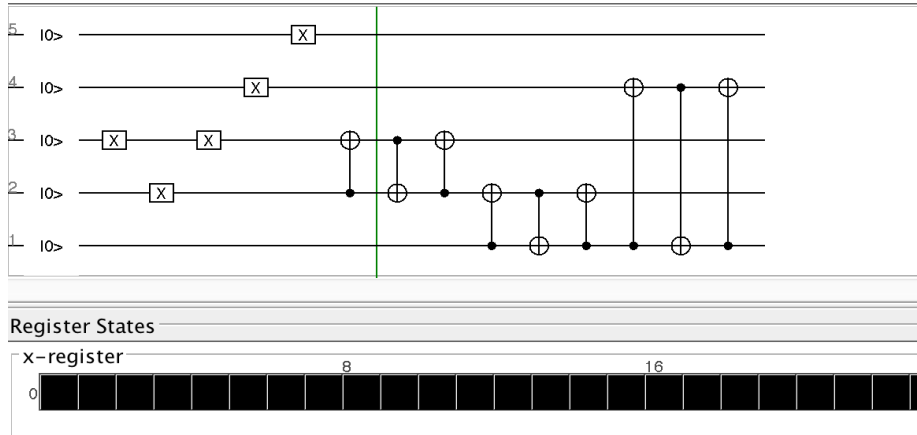
Paso 4



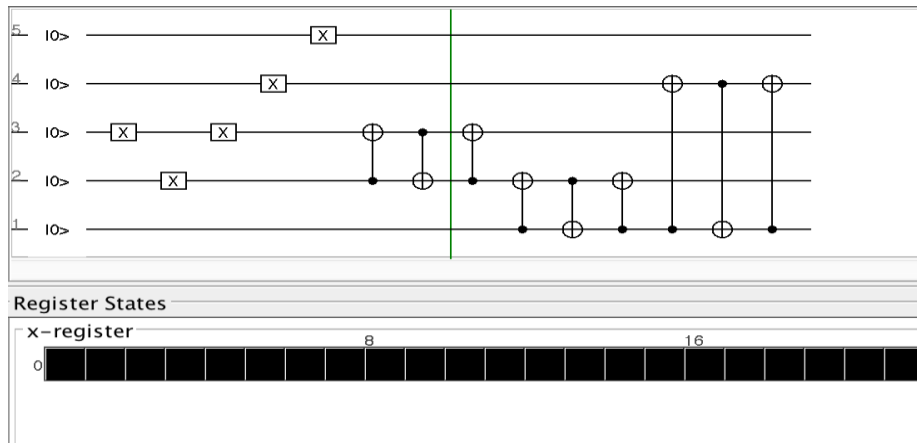
Paso 5



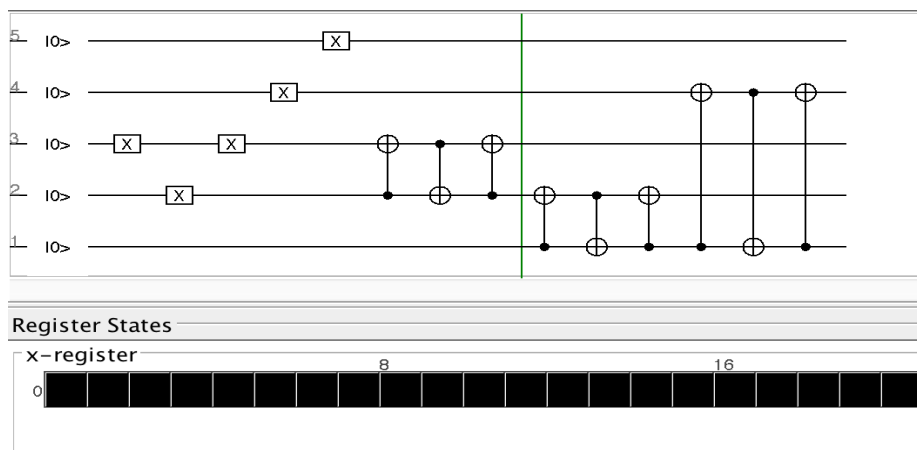
Paso 6



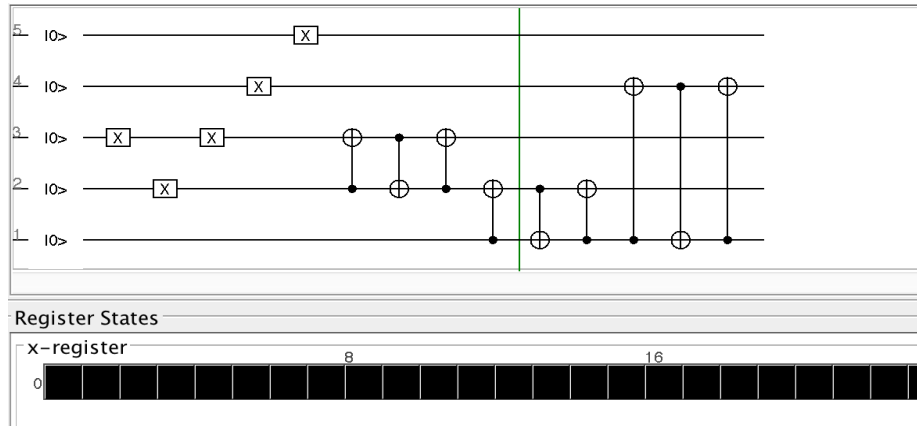
Paso 7



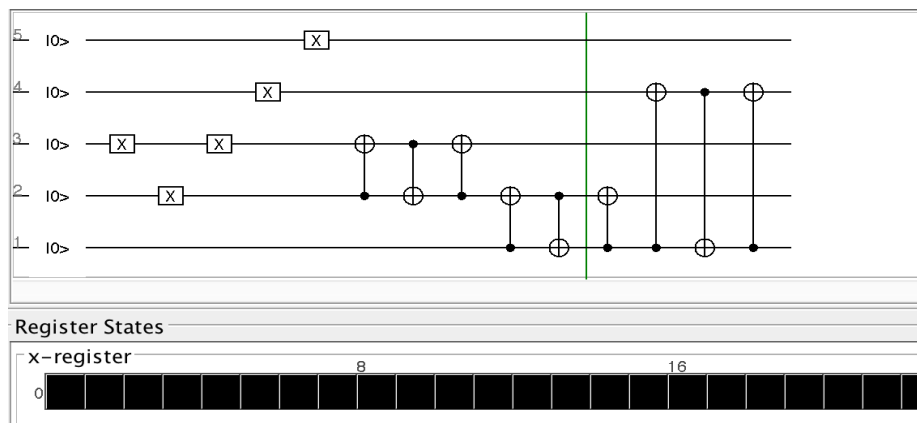
Paso 8



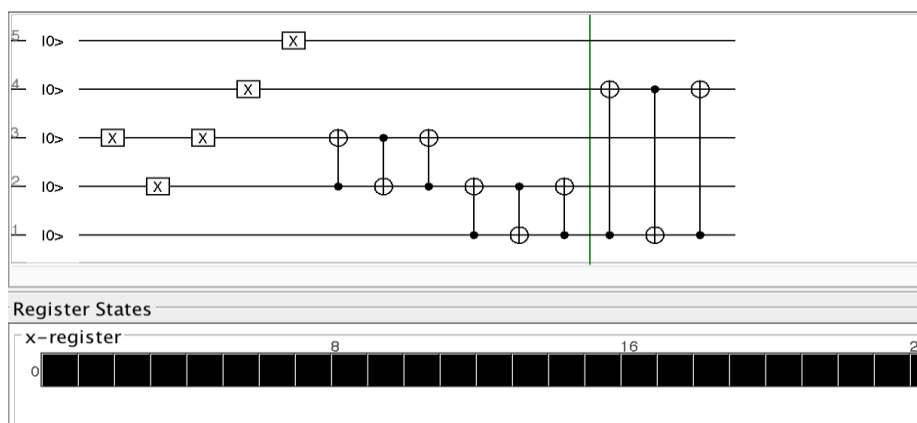
Paso 9



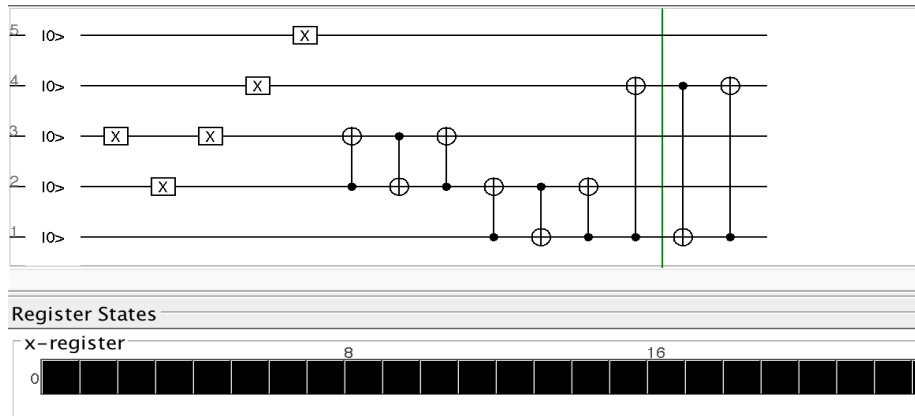
Paso 10



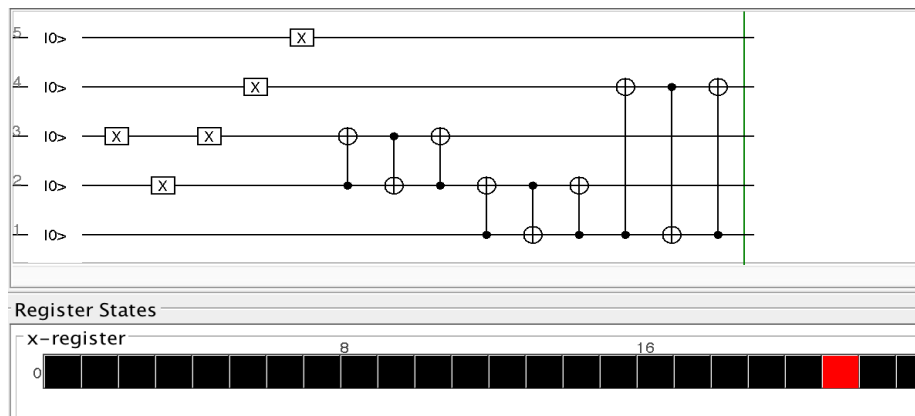
Paso 11



Paso 12



Paso 13



Paso 14

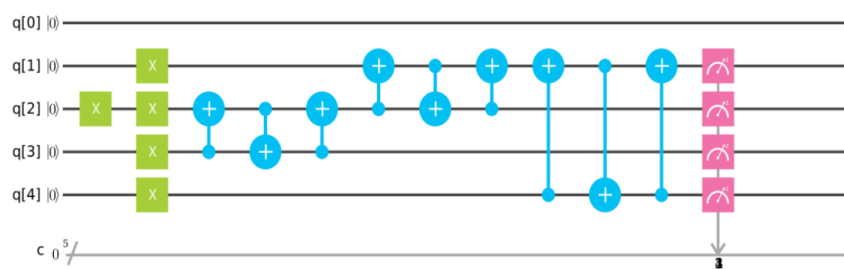
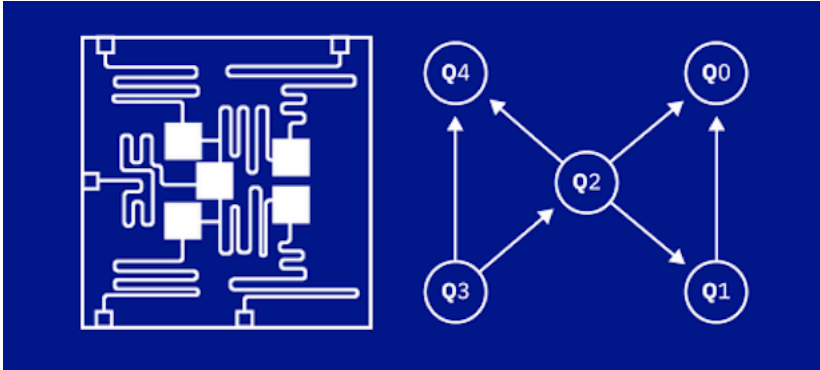


Fig. 3. Circuito cuántico propuesto en IBM Experience.

Este circuito fue compilado en la IBM q de IBM y corrió de acuerdo a los parámetros mostrados en la Tabla 2.

Tabla 2. Parámetros de funcionamiento.

Semilla	2007893536					
Numero de muestras	100					
Temperatura	0.021 K					
Backend	ibmqx4(5 qubits)	Q0	Q1	Q2	Q3	Q4
Error por compuerta	10e-3	0.86	0.69	1.97	1.97	1.8
Error de lectura	10e-2	4.60	5.40	12.80	8.70	4.5
Error de compuerta multi-Qubit	10 e -2	5.85	1.99	2.48	2.68	.94



ibmqx4(5 qubits)

La Figura 4 muestra el cálculo final obtenido.



Fig. 4. Cálculo final para 5 Qubits.

Resumen de resultados. Se pudo comprobar que nuestro algoritmo desarrollado trabaja perfectamente al realizar su comparación con otros

algoritmos desarrollados en la literatura, sin embargo la ventaja que presentamos es su adaptación para su comprobación en la computadora cuántica de IBM (IBM Q).

5. Conclusiones

La invención de una computadora cuántica práctica sea el comienzo de una nueva era en la computación y la tecnología en general. Dicha tecnología nos permitirá romper muchas de las barreras físicas que actualmente limitan las posibilidades de la tecnología informática clásica, mientras que la introducción de las nuevas posibilidades.

Sin embargo, aún existen varios obstáculos que deben superarse. Computadoras cuánticas actuales tienden a ser extremadamente caros, y requieren entornos operativos específicos, tales como ser capaz de refrigerar los componentes de la computadora cuántica cerca del cero absoluto. Muchos de los métodos que se han utilizado con éxito hoy en día no se escala fácilmente para permitir el uso de un mayor número de Qubits.

Sin embargo, hay una tremenda presión en esta zona, y es sólo cuestión de tiempo antes de que se superen estos obstáculos. Se requerirá más años de investigación y desarrollo antes de la edad de la computación cuántica llega, pero los bloques de construcción básicos que ya se han realizado. Las computadoras cuánticas se han construido ya que pueden realizar funciones básicas, y ya hay empresas que tratan de comercializar la tecnología.

Por ejemplo, D-Wave Systems es una empresa canadiense que afirma haber desarrollado una computadora cuántica de 28 Qubits, aunque ha habido críticas a sus demandas. En 2007, demostraron el uso de un equipo de 16 Qubits para resolver problemas tales como el reconocimiento de patrones, disposición de los asientos, y un Sudoku.

Incluso después de que las computadoras cuánticos prácticos se realizan, es poco probable que van a sustituir por completo a las computadoras clásicas. Su ventaja sobre las computadoras clásicas es significativa sólo en áreas específicas de aplicación. Lo más probable es que las computadoras del futuro tendrán lugar una especie de híbrido, que contiene los componentes de ambos tipos de computadoras.

La verificación de artículos que cumplen con este tipo de tipo de clasificación propuesto fue solo el 40 % debido a algunos pierden el enfoque de su investigación hablando de otros elementos que para nuestra propuesta no tienen relación.

Agradecimientos. Agradecemos las facilidades otorgadas para la realización de este trabajo al Instituto Politécnico Nacional a través de la Secretaría de Investigación y Posgrado con el proyecto SIP 20180023. A la Unidad Interdisciplinaria de Ingeniería y Ciencias Sociales y Administrativas y Centro de Investigación y Desarrollo de Tecnología Digital. Asimismo, al Programa de Estímulo al Desempeño de los Investigadores (EDI) y al Programa de Estímulo al Desempeño Docente (EDD).

Referencias

1. Cooper, K.B., Steffen, Matthias, McDermott, R. et al.: Observation of Quantum Oscillations between a Josephson Phase Qubit and a Microscopic Resonator Using Fast Readout. *Physical Review Letters*, 29 October (2004)
2. Devoret, Wallraff, Martinis: Superconducting Qubits: A Short Review. Department of Applied Physics, Yale University (2004)
3. Devoret: Manipulating the Quantum State of an Electrical Circuit. *Science*: 886–889, 3rd May (2002)
4. Han, J., Jonker, P.: On Quantum Computing with macroscopic Josephson Qubits. Pattern Recognition Group, Faculty of Applied Sciences, Delft University of Technology (2002)
5. Gruska, J.: Fundamentos de la informática, cap. 12.: *Frontiers - Quantum Computing* (1998)
6. McDermott, R., Simmonds, R.W., Steffen, Matthias, et al.: Simultaneous State Measurement of Coupled Josephson Phase Qubits. *Science*: 1299–1302, 25 February (2005)
7. Han, M.: Superconducting quantum qubits. *Physics World*, December (2004)
8. Benioff, P.: Modelos de máquinas Quantum Turing. Archivo LANL quant-ph / 9708054 (1997)
9. Shor, P. W.: Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. In: *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, Santa Fe, NM, Nov. 20-22 (1994)
10. Shor, P.W.: Quantum, Information Theory: Results and Open Problems. *Geom. Funct. Anal. (GAFA)*, Special Volume - GAF A2000:816–838 (2000)
11. Keyes, RW: IBM experience. *IBM J. Res. Desarrollar*. 32, 24 (1988)
12. Simmonds, R.W., Hite, D. A., McDermott, R., Steffen, M., Cooper, K. B., Lang, K.M., Martinis, J. M., Pappas, D. R.: Josephson Junctions Materials Research Using Phase Qubits. University of California, Santa Barbara (2005)
13. Simmonds, R.W., Lang, K.M., Hite, S. et al.: Decoherence in Josephson Phase Qubits from Junction Resonators. *Physical Review Letters* 13, August (2004)
14. Strauch, F.W., Johnson, P. R., Dragt, A. J., Lobb, C. J., Anderson, J. R., Wellstood, F. C.: Quantum Logic Gates for Coupled Superconducting Phase Qubits. *Phys. Rev. Lett.* 91, 167005 (2003)

Diseño de un robot rescatista para terremotos en México

Jesús Álvarez Cedillo¹, Teodoro Álvarez Sánchez², Raúl Junior Sandoval¹,
Jacobosandoval Gutiérrez³, Adriana Nava Vega⁴

¹ Instituto Politécnico Nacional, UPIICSA, Ciudad de México, México
jalvarez@ipn.mx

² Instituto Politécnico Nacional, CITEDI, Baja California, México
talvarez@citedi.mx

³ Universidad Autónoma Metropolitana, Unidad LERMA, Estado México, México
jacobosandoval@gmail.com

⁴ Universidad Autónoma de Baja California, Campus Tijuana, México
adriana.nava@uabc.edu.mx

Resumen. La República Mexicana está situada en una de las regiones sísmicamente más activas del mundo, enclavada dentro del área conocida como el Cinturón Circumpacífico donde se concentra la mayor actividad sísmica del planeta. En los últimos 5 años se han producido terremotos de magnitud muy considerable afectando estados como Oaxaca, Puebla, Morelos, Guerrero, Michoacán y la Ciudad de México. El número de personas que murieron atrapadas dentro de los escombros fue muy alta y en los próximos años se seguirá teniendo actividad sísmica constante. La necesidad de desarrollar un robot que permita detectar personas atrapadas no como un prototipo si no como una herramienta profesional de rescate es una prioridad. En este trabajo se propone el desarrollo de un robot rescatista diseñado con diversos sensores para detectar la temperatura, ruidos, detectar respiración de personas en los escombros. El robot propuesto se moverá entre los desechos para detectar a los humanos.

Palabras clave: robótica, aplicaciones de robótica, algoritmos de navegación, robot rescatista.

Design of a Rescue Robot for Earthquakes in Mexico

Abstract. The Mexican Republic is located in one of the most seismically active regions in the world, nestled within the area known as the Circumpacific Belt where the largest seismic activity on the planet is concentrated. In the last 5 years, there have been earthquakes of very considerable magnitude affecting states such as Oaxaca, Puebla, Morelos, Guerrero, Michoacán and Mexico City. The number of people who died trapped inside the rubble was very high and in the coming years, there will continue to be constant seismic activity. The need to develop a robot that allows detecting people trapped not as a prototype but as a professional rescue tool is a priority. This work proposes the development of a rescue robot designed with various sensors to detect temperature, noise, detect breathing of people in the rubble. The proposed robot will move through the waste to detect humans.

Keywords: robotics, robotics applications, navigation algorithms, rescue robot.

1. Introducción

En las últimas dos décadas se estima que los desastres son responsables de aproximadamente 3 millones de muertes en todo el mundo, 800 millones de personas afectadas adversamente y daños a la propiedad que superan los 50 mil millones de dólares. El reciente terremoto en México del 19 de septiembre de 2017 dejó una gran cantidad de personas muertas y heridas. Muchas de estas muertes se debieron a un colapso estructural cuando los edificios cayeron sobre las personas. La Búsqueda y Rescate Urbano involucra la ubicación, el rescate (extracción) y la estabilización médica inicial de las víctimas atrapadas en espacios confinados. Los vacíos que se forman cuando un edificio colapsa es una instancia de un espacio confinado. La Búsqueda y Rescate Urbano puede ser necesaria para una variedad de situaciones, incluyendo terremotos, huracanes, tornados, incendios, actividades terroristas y accidentes con materiales peligrosos.

En la actualidad, un equipo típico de búsqueda y rescate está compuesto por una decena de personas, entre las que se incluyen perros y adiestradores caninos, un paramédico, un ingeniero de estructuras y varios especialistas en el manejo de equipo especial para encontrar y extraer a una víctima. El equipo de búsqueda actual incluye cámaras de búsqueda y dispositivos de escucha. Las cámaras de búsqueda suelen ser cámaras de video montadas en algún dispositivo, como un poste que se puede insertar en huecos y agujeros para buscar signos de personas. A menudo, se perfora un agujero en las paredes obstruidas si se sospecha que existe un vacío en el otro lado. También se utiliza la imagen térmica. Esto es especialmente útil para encontrar cuerpos cálidos que han sido recubiertos con polvo y escombros que camuflajan efectivamente a la víctima. Los dispositivos utilizados que detectan sonidos son micrófonos altamente sensibles que pueden escuchar a una persona que puede estar moviéndose o intentando responder a las llamadas de los rescatistas.

Un temblor, sismo o terremoto, consiste en una serie de vibraciones en las rocas que se propagan en todas direcciones en forma de ondas. Estas ondas sísmicas son producidas por la liberación de energía ocurrida en los movimientos de grandes bloques de roca, los que son generados en grandes fallas (límites de placas tectónicas en separación, en colisión o en deslizamiento lateral); existen también temblores menores o locales (producidos por fallas pequeñas y por explosiones provocadas por volcanes o por el hombre) [1].

La voz popular denomina temblor a aquellos movimientos que son sentidos por el hombre (baja magnitud); sismo, al temblor que se siente y que causa algunos destrozos en construcciones (mediana magnitud); terremoto, a aquellos movimientos que son tan fuertes que producen graves daños al hábitat del hombre (gran magnitud).

La "Sismología" estudia los temblores, sismos o terremotos y sus ondas sísmicas, registradas en forma precisa por instrumentos denominados sismógrafos.

La amplitud de las ondas sísmicas permite clasificar a los temblores dentro de una escala denominada "Escala de Richter" (de uso mundial). Su magnitud está basada en el logaritmo de la amplitud máxima de la onda sísmica, ajustada por un factor que toma en cuenta la intensidad o debilidad de las ondas y su dispersión a partir del foco (centro de dispersión de las vibraciones de un temblor). Debido a que estas magnitudes están basadas en una escala logarítmica, un incremento en magnitud de una unidad, corresponde a un incremento de 10 veces, en el tamaño de un temblor,

por ejemplo un temblor de magnitud 8 será 10,000 veces mayor que un temblor de magnitud 4 (esto es $10^8/10^4$).

Cada año ocurren numerosos temblores registrados por sismógrafos, pequeños de magnitud inferior a 1 hasta magnitud 3.4, sin ser sentidos por el humano; a partir de la magnitud 5 empiezan a ocurrir daños y son sentidos prácticamente por todos, incrementándose estos daños en temblores de intensidad superior hasta llegar a 8 y más, con la casi destrucción total, siendo estos grandes temblores, que ocurren cada 5 a 10 años, registrados en la historia por los grandes disturbios provocados.

Debido a la gran sismicidad reportada de sismos en México, nuestro proyecto de investigación consiste en diseñar, construir y programar un robot autónomo capaz de búsqueda de víctimas en un edificio colapsado por el terremoto y justificar un robot no prototipo.

Los puntos a destacar que deben de ser cubiertos es que debe de desplazarse en un cuadrado de prueba de 20 x 20 metros., inclinaciones de hasta 50 grados, debe de evitar obstáculos y detectar sobrevivientes.

Es posible contar con un mapa de la construcción, pero los obstáculos resultantes del terremoto se dejaron como incógnitas. Nuestro robot tiene que acercarse a las víctimas identificadas y reportar a la central. Un sensor GPS muestra la posición real del robot en el espacio.

Los desastres naturales recientes y las catástrofes provocadas por el hombre han centrado la atención en el área de manejo de emergencias y rescate. Estas experiencias han demostrado que la mayoría de las respuestas de preparación y emergencia del gobierno son generalmente inadecuadas para hacer frente a los desastres. Teniendo en cuenta la gran cantidad de personas que han muerto debido a esfuerzos de rescate reactivos, espontáneos y no profesionales como resultado de la falta de equipo adecuado o la falta de respuesta inmediata, los investigadores naturalmente han desarrollado herramientas de rescate mecánicas y técnicas de planificación estratégica para las operaciones de rescate planificadas. Las actividades de investigación y desarrollo han dado lugar al surgimiento del campo de la robótica de rescate, que puede definirse como la utilización de tecnología robótica para la asistencia humana en cualquier fase de las operaciones de rescate, que son multifacéticas y varían de un país a otro debido a las políticas regionales. Los tipos de desastres y las diferentes composiciones de escombros en las áreas de desastre. Otros aspectos de la robótica de rescate incluyen:

1. Detección e identificación de cuerpos vivos.
2. Enrutamiento y / o limpieza de escombros para acceder a la víctima.
3. Estabilización física, emocional o médica del sobreviviente al proporcionarle primeros auxilios telemétricos y de administración automática.
4. Fortificación del cuerpo vivo para una recuperación segura contra cualquier escombros que caiga y posibles lesiones.
5. Transporte de la víctima.

Estas operaciones también varían en carácter para diferentes tipos de entornos de desastre, como áreas urbanas, subterráneas o bajo el agua, que son desestructuradas y tecnológicamente difíciles de acceder para los humanos. Los problemas críticos en el rescate son la conveniencia y el cumplimiento de las herramientas de rescate. Los otros problemas importantes de rescate encontrados son:

1. Las herramientas no ásperas son generalmente engorrosas, destructivas y, por lo general, se adaptan directamente a los dispositivos de construcción.
2. Las máquinas de limpieza de escombros son dispositivos de construcción pesada que, cuando funcionan sobre escombros, hacen que los escombros se hundan.
3. El funcionamiento de la herramienta es generalmente muy lento y tedioso y no tiene en cuenta los intentos anteriores en el mismo lugar (no aprenden de las pruebas en el lugar), lo que produce muchas repeticiones fallidas.
4. Aunque hay algunos detectores disponibles, la búsqueda de sobrevivientes se basa principalmente en la detección de perros y voces humanas, donde la llamada y la escucha requieren silencio y atención enfocada que es muy difícil debido a los trabajadores de rescate sobrecargados, agotados y deprimidos.
5. El suministro de primeros auxilios solo se puede hacer cuando se encuentra cerca del sobreviviente, una distancia que se alcanza con frecuencia cuando se excede el momento crítico para la supervivencia.
6. La recuperación de cuerpos genera lesiones adicionales, ya que la estabilización profesional de la víctima rara vez se obtiene y no se supervisa continuamente.

Con el objetivo de mejorar la calidad del rescate y la vida después del rescate, el campo de la robótica de rescate busca dispositivos diestros que estén equipados con capacidad de aprendizaje, adaptables a diversos tipos de uso con una funcionalidad lo suficientemente amplia bajo múltiples sensores y que cumplan con las condiciones del medio ambiente y el de la persona rescatada.

2. Descripción del método

Se plantea la necesidad de desarrollar una plataforma robótica que permita acceder y permanecer en entornos naturales, desplazarse sin inconvenientes por los múltiples terrenos [2].

Nuestro método de construcción se basa en dos áreas fundamentales:

1. El cuerpo del robot: Se decidió que para que este tipo de robot tenga éxito, debe ser capaz de viajar en un camino no muy recto, alcanzar alturas de 20 centímetros, y ser capaz de escalar hasta 60 grados. El diseño debe de mantener su integridad estructural resistente. El robot será capaz de viajar en rampas para comprobar si hay víctimas en diferentes niveles. Por último, las herramientas utilizadas se diseñan de manera modular para adaptación de piezas adicionales o modificaciones del diseño.
2. La exploración: Para llevar a cabo los viajes camino recto y giros precisos se creó un sistema de doble accionamiento diferencial. La ventaja de este tipo de sistema es que un motor se utiliza para el movimiento hacia adelante y hacia atrás, mientras que un segundo motor se utiliza para hacer el escalamiento de obstáculos. Esta configuración tiene la intención de hacer que las dos ruedas motrices se mueven exactamente lo mismo si se viaja hacia delante o hacia atrás o exactamente lo contrario si se activa el motor de escalada. Para probar el éxito de nuestra unidad de doble diferencial que

mide la forma recta el robot de conducir sin ningún control de retroalimentación y la precisión con la que sería girar 90 grados. Los resultados para la línea recta eran buenas, pero las vueltas tienen un poco de problema.

2.1. Características mínimas del diseño

Los parámetros usados en el diseño son los siguientes:

1. Capacidad de desplazamiento en lugares de difícil acceso.
2. Posibilidad de desplazamiento en múltiples tipos de terrenos.
3. Movimiento del chasis con el propósito de poder cambiar la posición de este y adaptarlo a las necesidades mediante el control remoto.
4. Utilizar el control remoto para dirigir el robot a los lugares no accesibles para obtener información visual.
5. Capacidad de visión del robot para obtener acercamientos claros de la zona.
6. Consideraciones para realizar cálculos.

3. Mecánica del robot

Los robots de rescate de hoy en día no tienen grandes sistemas de cómputo muy sofisticados en términos de inteligencia artificial (en inglés, AI). La IA es una forma de programar a la computadora que le permite al robot procesar información y tomar algunas decisiones por sí mismo. En lugar de una IA independiente, la mayoría de los robots rescatistas son controlados remotamente por los operadores humanos. Vehículos terrestres no tripulados (UGV) o vehículos aéreos no tripulados (UAV) [3].

Los motores se escogieron en base al cálculo de los momentos críticos en lo que estos se someterían en las peores condiciones y con base a estos cálculos se da un factor de seguridad adecuado en el que se garantice un correcto funcionamiento, igualmente para escoger el material y las dimensiones adecuadas para los ejes se procede a hacer un análisis teniendo en cuenta los esfuerzos cortantes y los momentos vectores en condiciones críticas, entonces se procede a calcular dichas dimensiones y dando un respectivo factor de seguridad para estos, para escoger la relación correcta de engranes para las orugas, se tuvo en cuenta que las velocidades de las orugas y las llantas serán iguales [4].

Como una solución al problema de la plataforma mecánica se planteó inicialmente una plataforma robot explorador terrestre tipo rover de 6 ruedas con sistema de amortiguamiento Rocker bogie, el cual le permite superar obstáculos con facilidad además de garantizar la tracción por lo menos de alguna de las ruedas, acoplándose al terreno, y por lo tanto teniendo ventajas ante sistemas clásicos de todo terreno como la oruga simple. Este sistema permite que la plataforma pueda desplazarse en terrenos donde hay variedad de obstáculos, además de evitar transmisión directa de la energía al chasis debido a este sistema.

Para verificar este sistema se procedió a la construcción de un modelo funcional a escala. La construcción del modelo de menor escala se estableció con el objetivo primordial de analizar la cinemática y la dinámica del sistema mecánico diseñado y

para realizar las pruebas de los sensores necesarios para el funcionamiento de la plataforma antes de la construcción del prototipo final.

El diseño del modelo a escala se basó en la suspensión Rocker bogie utilizada en varios de los robots construidos por la NASA [5], pero usando el eslabón libre como la parte frontal de la plataforma de esta manera permite sobrepasar obstáculos de mayor dimensión, esto se observó en el prototipo a escala. El diseño del modelo se muestra en la Fig. 1.



Fig. 1. Prototipo del robot construido.

Los robots más comunes son robots pequeños y planos montados en huellas de tanques en miniatura. Estos robots son resistentes, capaces de ingresar por huecos y edificios y generalmente tienen una variedad de sensores integrados, que incluyen la vigilancia de audio y video y la detección de vida humana. Estos robots son versátiles, con diferentes sensores o herramientas que se pueden montar. Prácticamente todos son portátiles. El robot propuesto se maneja con un joystick de control, tiene siete configuraciones de velocidad (la velocidad máxima es de 1.8 metros por segundo y puede usar sus huellas para subir escaleras, maniobrar entre los escombros e incluso tomar muestras.

Nuestra versión más simple incluye dispositivos de audio y video y un brazo mecánico. Una versión ligera de 27 kg que omite el brazo.

El robot es controlado por un procesador Pentium con dispositivo empaquetado, un sistema GPS, una brújula electrónica y sensores de temperatura integrados. Puede moverse a más de 8 mph (13 kph), puede desplegarse en minutos y puede soportar una caída de 1,8 metros sobre el concreto, el equivalente a 400 g de resistencia.

Para su fabricación se realizó el modelo en 3D usando el programa Blender para su diseño. La Fig. 2. Muestra diferentes vistas del modelo en 3D.

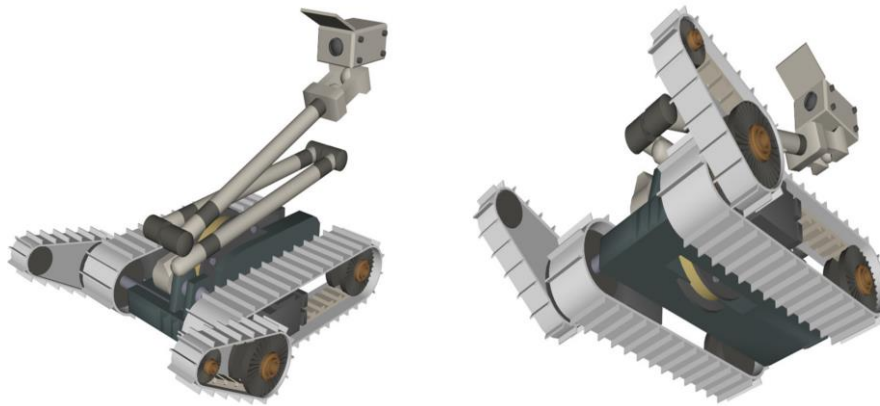


Fig. 2. Vistas del Modelo 3D.

A partir del modelo en 3D creado se maquina cada pieza en aluminio en el Laboratorio de pesados de UPIICSA para tener la estructura y se adquirieron las orugas para el montaje de cada rueda.

El sistema final cuenta de tres motores de 12 volts para accionar las orugas una batería de acido plomo basado en un electrolito de ácido sulfúrico disuelto en agua, 60–110 Wh/l, eficiencia carga/descarga: 50–95% y temperatura de carga: mín. -40°C , máx. 50°C .

El sistema tiene dos puentes H para controlar los motores y un sistema basado en Intel para su control y operación.

Otra de las diferencias de la suspensión Rocker bogie original es que los ejes que soportan el eslabón principal no se encuentran conectados mediante una relación diferencial sino que cada uno dispone de un servomotor que controla su movimiento independientemente [5, 6, 9]. El sistema de transmisión del servomotor al eje se realiza mediante engranajes como se puede ver en la Fig. 3.

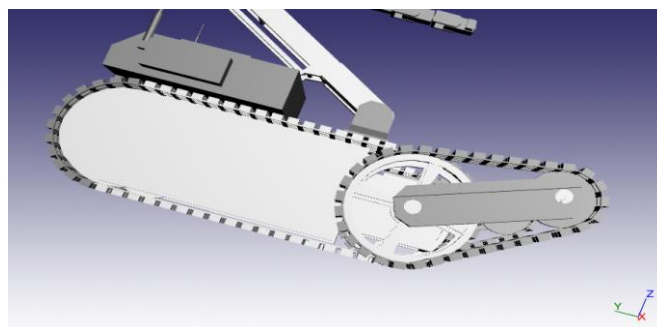


Fig. 3. Sistema de engranajes.

Basados en las ventajas y desventajas halladas mediante las pruebas del modelo a escala se diseño el prototipo a escala real. Se mantiene el acople del armazón al sistema de tracción por medio de un eje a cada lado, ya que esto facilita la absorción

de momentos, además de contribuir al control del posicionamiento del chasis. Debe tenerse en cuenta que para que el sistema de tracción haga que el robot se desplace debe vencer una fuerza determinada por la ecuación (1), donde μ corresponde al coeficiente de rodadura entre las ODS superficies en contacto, m la masa del robot, g la aceleración debido a la gravedad y θ el ángulo de la pendiente del suelo con respecto a la horizontal [7]:

$$r = \mu(mg \cos(\theta)) + mg \sin(\theta) + Fin \quad (1)$$

El primer término de la derecha de la ecuación 1 (Fr) se refiere a la fuerza de fricción entre las llantas y el suelo que se ve afectado por el ángulo de la pendiente por donde se desplaza. El segundo término se refiere a la componente del peso del robot que al entrar en una pendiente se vuelve significativo. El último término corresponde a las fuerzas inerciales debidas a las aceleraciones causadas por el movimiento del robot. El torque que tendría que hacer cada rueda para mover el robot estaría determinado por la ecuación (2), donde N corresponde al número de ruedas a tracción [8]:

$$Tr = Fr \times \frac{R_{rueda}}{N}, \quad (2)$$

$$Tm = f(mg).R_{rueda}. \quad (3)$$

El torque que se le puede pedir a una rueda está determinado por el peso del robot y el coeficiente de fricción f entre el suelo y las llantas según la ecuación (3).

De esta forma para que las ruedas transmitan el torque requerido sin deslizar, este torque debe ser menor al torque máximo tractivo (Tm). Con respecto a las ruedas delanteras, uno de los inconvenientes que se encontró fue la falta de una rueda que completara un triángulo en el eslabón libre de la suspensión Rocker bogie para que cuando este eslabón girara completamente al tratar de sobrepasar obstáculos de tamaño superior al modelo, el robot pueda seguir funcionando correctamente. Para esto se utilizó un sistema de reducción Epicycloide. Para superar los percances detectados se diseñó un prototipo a escala real. El prototipo se muestra en la Fig. 4, donde se aprecian cambios respecto a la versión anterior.



Fig. 4. Robot final.

4. Sistema de control remoto

Tres son los objetivos fundamentales de la estación. El primer objetivo es simplificar la tarea del operador para ello el sistema contará con:

- 1) Interface hombre/ maquina: Nuestro sistema fue desarrollado con LIVE CODE community.
- 2) Visualizar la información que proviene de las cámaras panorámicas del robot.
- 3) Mostrar los datos provenientes de la cámara del brazo manipulador, así como los datos de los sensores.
- 4) Mostrar los datos de los sistemas de posicionamiento global e inercial, así como la posición actual del manipulador.
- 5) Proporcionar una herramienta de manipulación (joystick) que permitan manipular tanto el robot como el manipulador de una manera intuitiva.

5. Resultados

Al haber realizado la plataforma robot rescatador para terremotos y realizar las pruebas pertinentes, se obtienen los diferentes resultados los cuales se encuentran descritos a continuación:

En la modificación realizada al sistema de suspensión Rocker bogie se resalta principalmente los 2 tipos de movimiento, los cuales son el movimiento de desplazamiento plano y el movimiento rotacional para la superación de obstáculos.

La plataforma posee dos tipos de movimientos generales básicos, para su desplazamiento los cuales son el movimiento lineal, en el cual puede llegar a alcanzar velocidades de hasta 1.256km/h y los movimientos rotacionales que pueden ser totales o parciales.

Desplazamiento en lugares de difícil acceso. La plataforma se caracteriza por el poder desplazarse por terrenos en los cuales la dificultad de desplazamiento es alta, como lo pueden ser escaleras, las cuales puede superar gracias al sistema de amortiguamiento Roker bogie modificada, también tuvo la potencia para superar pendientes.

En el proceso de diseño y manufactura del robot se puede apreciar la importancia de la utilización de plataformas robóticas como soporte de las diferentes actividades del ser humano, ya que le brinda una reducción de riesgos a los cuales se debe exponer el individuo. Además de brindar una solución que permite ser portada a diferentes ambientes en los cuales se puede desenvolver de la mejor manera.

Teniendo en cuenta el gran impacto que puede llegar a tener el robot se desean realizar algunos cambios que contribuyan a la disminución de consumo energético para incrementar el tiempo de autonomía de la plataforma, además de realizar pruebas en ambientes extremos que la exijan al máximo y con esto observar las fallas que puedan presentarse en situaciones muy similares a las que se exponga.

6. Conclusiones

El sistema mecánico implementado con una suspensión Rocker bogie modificada, permite superar las diferentes adversidades del terreno. Además de brindar gran información para la toma de decisiones como lo son imagen normal del entorno, detección de vida y temperatura.

Esta información junto con el control remoto del usuario amplía decisiones que se pueden tomar frente a una emergencia.

El sistema Epicycloide con oruga aumenta la tracción y estabilidad del robot, ya que la oruga está en contacto permanente con la superficie. Sin embargo requiere de una mayor potencia en los motores delanteros para que el sistema de rueda triangular mueva el robot con facilidad.

Se resalta que el sistema implementado disminuye el número de motores de la configuración Rocker bogie tradicional en 2 motores, haciendo un control más efectivo de la potencia que necesita el sistema en cada instante de tiempo.

7. Recomendaciones

Los investigadores interesados en continuar nuestra investigación podrían concentrarse en la mejora de los circuitos y métodos de programación y en la estructura Rocker bogie.

Agradecimientos. Agradecemos las facilidades otorgadas para la realización de este trabajo al Instituto Politécnico Nacional a través de la Secretaría de Investigación y Posgrado con el proyecto SIP 20180023. A la Unidad Interdisciplinaria de Ingeniería y Ciencias Sociales y Administrativas y Centro de Investigación y Desarrollo de Tecnología Digital. Asimismo, al Programa de Estímulo al Desempeño de los Investigadores (EDI) y al Programa de Estímulo al Desempeño Docente (EDD).

Referencias

1. Montalvo Arrieta, J.C.: El potencial de riesgo sísmico en el noreste de México y sus implicaciones en la zona metropolitana de Monterrey, N. L. Ciencia UANL, Vol. XII, Núm. 4, octubre-diciembre, pp. 398–408 (2009)
2. Young, E.M., Withrow, T.J., Sarkar, N.: Design of intention-based assistive robot for upper limb. *Advanced Robotics*, 31(11):580–594, DOI: 10.1080/01691864.2017.1295883 (2017)
3. Chang, P., Kang, Y., Cho, G. R., Kim, J. H., Jin, M., Lee, J., Jeong, J. W., Han, D. K., Jung, J. H., Lee, W., Kim, Y.: Control Architecture Design for a Fire Searching Robot using Task Oriented Design Methodology. In: *SICE-ICASE Int. Jt. Conf.* 2006, pp. 3126–3131 (2006)
4. Kim, Young-Duk, Kim, Yoon-Gu, Lee, Seung-Hyun, Kang, Jeong-Ho, An, Jinung: Portable fire evacuation guide robot system. In: *Intelligent Robots and Systems (IROS 2009), IEEE/RSJ International Conference on*, pp. 2789–2794 (2009).
5. Kim, Jong-Hwan, Keller, B., Lattimer, B.Y.: Sensor fusion based seek-and-find fire algorithm for intelligent firefighting robot. In: *Advanced Intelligent Mechatronics (AIM) 2013 IEEE/ASME, International Conference on*, pp. 1482–1486 (2013)

6. Amano, H.: Present status and problems of reghting robots. In: Proc. 41st SICE Annu. Conf. SICE 2002, vol. 2, pp. 880–885 (2002)
7. Harrington, B. D., Voorhees, C.: The Challenges of Designing the Rocker-Bogie Suspension for the Mars Exploration Rover. In: Proc. 37th Aerosp. Mech. Symp., pp. 185–195 (2004)
8. Stauer, C., Grimson, W. E. L.: Adaptive background mixture models for real-time tracking. In: Proc. 1999 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Cat No PR00149, vol. 2, no. c, pp. 246–252 (1999)
9. Nghiem, Anh-Tuan, Bremond, F.: Background subtraction in people detection framework for RGB-D cameras. In: Advanced Video and Signal Based Surveillance (AVSS) 2014 11th IEEE International Conference on, pp. 241–246 (2014)
10. Renno, J.-P.R., Orwell, J., Jones, G.A.: Evaluation of shadow classification techniques for object detection and tracking. In: Image Processing (ICIP '04), International Conference on, vol. 1, pp. 143–146 (2004).
11. Li, Jinping, Guo, Yukun, Ma, Liwen: MCSHM: A simple and practical method for moving objects detection in dynamic scenes. In: Chinese Automation Congress (CAC), pp. 5112–5118 (2017)

Detección de sonidos torácicos aplicando detección de actividad de voz

Julio A. Valdez¹, Pedro Mayorga¹, Gilberto Chavez¹, Christopher Druzgalski²,
Vesna Zeljkovic³

¹ Tecnológico Nacional de México/I.T. Mexicali, Mexicali, México

² California State University, Electrical Engineering Department, California, USA

³ The Lincoln University, School of Science and Technology, Pennsylvania, USA
julito_valdez@hotmail.com

Abstract. La evaluación de los sonidos pulmonares (LS) y los sonidos cardíacos (HS) y la documentación de sus características forman parte de los procedimientos rutinarios de diagnóstico cardiopulmonar. Los componentes acústicos peculiares asociados con las fases inspiratoria, espiratoria, S1 y S2 de la función cardiopulmonar, como por ejemplo las sibilancias asmáticas, sirven como indicadores de diagnóstico útiles. Complementar la auscultación tradicional con la detección automática y la categorización de las señales características proporciona oportunidades de monitoreo mejorado a largo plazo que es indispensable en la EPOC u otras anomalías de origen cardíaco. Además, debido a la edad u otras limitaciones sensoriales de un profesional de la salud, la detección automática de eventos de LS-HS podría mejorar significativamente estos aspectos del diagnóstico médico. En particular, el método novedoso propuesto se basa en los elementos de Detección de Actividad de Voz (VAD), asociados con componentes de frecuencia LS-HS similares, y la utilización de Modelos Mezclados Gaussianos (GMM). La detección automática con GMM está acompañada por los Coeficientes Cepstrales de Frecuencia de Mel (MFCC) y la transformada de Hilbert. Esta técnica permite detectar y clasificar eventos peculiares de LS y HS, y obtener tasas de clasificación más altas en comparación con otras formas de VAD. La eficiencia de clasificación obtenida con Modelos Ocultos de Markov (HMM) fue del 95% para LS y del 92% para HS, documentando su viabilidad clínica.

Palabras clave: Diagnóstico cardiopulmonar, Modelos Mezclados Gaussianos (GMM), Modelos Ocultos de Markov (HMM), transformada de Hilbert, Detección de Actividad de Voz (VAD).

Thoracic Sound Detection Using Voice Activity Detection

Abstract. The lung and heart sounds (LS, HS) assessment and characteristics documentation are part in common procedures to cardiopulmonary diagnostic. Peculiar acoustic components associated to inspiratory and expiratory phases, S1 and S2 of cardiopulmonary function, as in sibilance are useful indicators in diagnostics for asthma. To complete traditional auscultation with automatic detection and categorization signals, give better opportunity of monitoring to long term, even this is indispensable for Chronic Obstructive Pulmonary Disease (COPD) and other anomalies of cardiac origin. Moreover, due to sensorial limitations and age of health professionals, automatic detection could significantly

improve these medical aspects. Particularly, the novel method proposed here, is based on Voice Activity Detections principles (VAD), but they are associated with LS-HS frequency components, which are similarities with voice, and strengthened with Gaussian Mixture Models (GMM). Automatic detection is here reinforced with Mel Frequency Cepstral Coefficients (MFCC) and Hilbert transform. This technique allowed to detect and classify peculiar events in LS and HS, and to obtain higher classification rates compared with other VAD techniques. The classification efficiency reached with Hidden Markov Models (HMM) were 95% for LS and 92% for HS, evidencing clinical viability for our proposed method.

Keywords: Cardiopulmonary diagnosis, Gaussian Mixed Models (GMM), Hidden Markov Models (HMM), Hilbert transform, Voice Activity Detection (VAD).

1. Introducción

La auscultación representa uno de los procedimientos de diagnóstico más comunes y bien establecidos [1]. Sin embargo, todavía se continúa confiando en criterios descriptivos subjetivos más que en la caracterización cuantitativa. Los sonidos cardíacos y pulmonares proporcionan información de diagnóstico muy útil en particular cuando su origen específico se extrae de forma independiente de los sonidos torácicos generales registrados.

Esta es una tarea desafiante, ya que los componentes de frecuencia del corazón y los sonidos pulmonares se superponen, aunque obviamente, sus características periódicas son distintas debido a las diferencias en las frecuencias cardíaca y respiratoria. El S1 y el S2, como dos componentes principales de los sonidos cardíacos, suelen incluir componentes dominantes de frecuencia por debajo de 150 Hz y se pueden correlacionar con el Electrocardiograma (ECG), aunque por razones prácticas rara vez se realiza en un entorno clínico. Los sonidos pulmonares que reflejan las fases inspiratoria y espiratoria incluyen también componentes de mayor frecuencia especialmente cuando, por ejemplo, hay sibilancias. Por las mismas razones prácticas, se registran como parte de la auscultación general, ya que se usa un estetoscopio electrónico como unidad independiente. Por lo tanto, para minimizar el efecto de correlación en el análisis computacional, una separación de los sonidos del corazón de los sonidos del pulmón es útil [1]. En ambos casos, un algoritmo eficiente también puede reducir el ruido.

Diferentes autores [2-6] proponen diversos métodos de extracción de características de LS y HS, que permiten extraer los componentes primordiales de los sonidos cardíacos y respiratorios. Algunos de estos métodos incluyen el Análisis Espectral Singular (SSA) para separar los componentes importantes HS de las señales LS [2]. La extracción HS se mejoró, mediante la localización de los componentes en función de la envolvente de la señal aplicando la transformada de Hilbert. Además, dado que en la envolvente de señal se puede usar la fórmula de Herón para obtener las componentes S1 y S2, por lo cual se puede obtener una mejor eficiencia en detección que con SSA [6]. Debido a que los eventos LS están asociados con características relacionadas con la duración, frecuencia y amplitud, algunos autores [3, 4], aplicaron métodos de Detección de Actividad de Voz (VAD) y Coeficientes Cepstrales de Frecuencias Mel (MFCC),

que podrían ser útiles para extraer eventos importantes de LS. Otros autores han propuesto explorar la realidad virtual, monitorear y detectar la inhalación y la exhalación [4]. En general, los métodos tienen como inconveniente el ruido (ruido ambiental, ruido cardíaco o ruido pulmonar dependiendo del interés en el origen particular de los sonidos). Además, las características principales de extracción todavía se ven obstaculizadas por las condiciones de registro de las señales [3]. Debido a que la detección de actividad de señales LS es bastante compleja, aquí presentamos una técnica centrada en un método novedoso que abarca también la detección automática de eventos acústicos relacionados con la inhalación y la exhalación basados en modelos mezclados gaussianos (GMM). Los modelos VAD normalmente son basados en principios de energía y cruce por cero, pero no en nuestro caso.

La sección 2 explica el conjunto de técnicas aplicadas para llevar a cabo los experimentos, así como los eventos que componen el sonido y la base de datos utilizada. La sección 3 muestra los experimentos llevados a cabo, incluyendo figuras y tablas obtenidas a partir de los resultados. Finalmente, la sección 4 presenta la conclusión de los experimentos y los resultados.

2. Metodología

Esta sección proporciona los fundamentos para la representación, el preprocesamiento y el modelado de eventos que abarcan LS en nuestros experimentos.

2.1. Vectores de Coeficientes Cepstrales en Frecuencia Mel (MFCC)

En MFCC, los sonidos se parametrizan implementando un énfasis previo con filtros FIR, seguidos de una ventana de Hamming aplicada a cada cuadro analizado [7-10]. En este proyecto, los experimentos se llevaron a cabo utilizando ventanas de Hamming de 50 ms (LS) y 130 ms (HS) con un 50% de traslape en ambos casos, a las que se aplicó la Transformada Rápida de Fourier (FFT); posteriormente, se obtuvo el módulo y luego se multiplicó por un banco de filtros cuyo rango de frecuencias y frecuencias centrales se distribuyeron por la escala Mel o Bark. Esto fue seguido por una etapa de cálculo de los valores obtenidos de las salidas de cada filtro y, posteriormente, la Transformada de Fourier Inversa. Como resultado, se obtuvo un vector de características llamado MFCC [4, 11, 12].

2.2. Vectores cuantílicos

El cuantil q_p de una variable aleatoria se define como el valor q más pequeño de tal forma que la función de distribución acumulativa sea mayor o igual que la probabilidad p , donde p es $0 < p < 1$. Esto se puede calcular a partir de una función de densidad de probabilidad continua (PDF) $f(x)$ a través de (1):

$$p = \int_{-\infty}^{q_p} f(x) dx. \quad (1)$$

La duración de la fase de inhalación (~ 1.5 s.) y la fase de exhalación (~ 2.5 s.) para la mayoría de las señales LS utilizadas proviene de [13]. En este proceso, la primera etapa consiste en leer la señal, comenzando desde un archivo *.wav; posteriormente,

se aplica la FFT. Cumpliendo un principio básico para una función de distribución de probabilidad, la distribución espectral se tiene que normalizar (2):

$$F_N(f) = \int_{-\infty}^{\infty} \frac{f(t)e^{-j2\pi ft} dt}{\text{area}(F(f))}. \quad (2)$$

Un ejemplo de cuartil es el uso de cuartiles, calculado por (3), donde cada valor de frecuencia $f_{0.25}, \dots, f_{0.75}$ corresponde a su coeficiente cuartil respectivo [13]:

$$A_{0.25} = \int_{-\infty}^{f_{0.25}} F_N(f) df, \dots, A_{0.75} = \int_{-\infty}^{f_{0.75}} F_N(f) df. \quad (3)$$

2.3. Modelos Mezclados Gaussianos (GMM)

El modelo GMM es una tripleta compuesta por las medias, las covarianzas y las ponderaciones de cada densidad. El modelado GMM usa el algoritmo de maximización de la esperanza (EM) para entrenar los modelos $\Lambda_i = \{m_i, \vec{\mu}, \Sigma_i\}$. El promedio $\vec{\mu}$ representa el promedio de todos los vectores, mientras que la matriz de covarianza Σ_i modela la variabilidad de las características de una clase acústica [14]:

$$p(\vec{x}|\Lambda) = \sum_{i=1}^M m_i b_i(\vec{x}). \quad (4)$$

En la ecuación (4), $\vec{x}$ es un MFCC o vector cuartil, $b_i, \forall i = 1, \dots, M$ son las densidades componentes y $m_i, \forall i = 1, \dots, M$ son las ponderaciones de cada densidad participante en el modelo. Cada densidad componente es una función Gaussiana D-dimensional [8, 11]. Cada densidad gaussiana contiene los parámetros representados en la Ecuación (5):

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp \left[-\frac{1}{2} (\vec{x} - \vec{\mu})^T \Sigma_i^{-1} (\vec{x} - \vec{\mu}) \right]. \quad (5)$$

2.4. Detección de Actividad de Voz (VAD)

Los detectores de actividad de voz (VAD) son fundamentales para el uso eficiente del ancho de banda [3]. Han sido útiles en la transmisión y el procesamiento de audio digital, y ahora en señales LS-HS. En [3] VAD se aplica a una señal aleatoria con ruido de fondo estacionario, su duración es mayor que un fonema, y los eventos en LS tienen una duración promedio mayor que un fonema. Es por este motivo, que aquí se propone una versión de VAD basada en modelos de mezclados gaussianos (GMM). En primer lugar, un segmento de inhalación y exhalación asignado manualmente en muestras seleccionadas de señales LS, y se guardan en un formato wav, además, las zonas de silencio son segmentadas manualmente (que tienen ruido) y se guardan en formato wav. El siguiente paso es calcular los modelos GMM correspondiente a la inhalación y la exhalación. De la misma manera, se calcula otro modelo GMM para zonas de silencio-ruido. Con estos dos modelos se constituye el sistema VAD, que se aplica a todas las señales. Los vectores MFCC se calculan sobre las señales completas, y con VAD propuesto se detecta si cada MFCC corresponde a actividad (en este caso, inhalación o exhalación) o a una región de no actividad (silencio o ruido). Es importante tener en

cuenta que las señales y sus cortes se convirtieron en vectores MFCC para el cálculo de los modelos VAD-GMM. Eventualmente, los vectores MFCC se usarán para distinguir entre las fases de inhalación y exhalación, ya que la media de las sextas componentes de los vectores MFCC corresponde a un umbral para distinguir entre los segmentos de inhalación y espiración [4].

2.5. Análisis de Componentes Principales (PCA)

Uno de los propósitos del Análisis de Componentes Principales (PCA) es reducir la dimensionalidad de p a d , donde $d < p$, y al mismo tiempo preservar la mayor cantidad posible de varianza de los datos originales [15, 16]. En este proyecto, la idea es aplicar PCA para obtener una representación disminuida en la dimensión del vector de características, pero sin disminuir la eficiencia. Para usar PCA, se transforman los datos a un nuevo grupo de coordenadas o variables, que son una combinación lineal de las variables originales. Además, las observaciones en el nuevo espacio de componentes principales no están correlacionadas [17].

2.6. Modelos Ocultos de Markov

El ciclo cardíaco es acústicamente secuencial, normalmente consta de dos eventos principales conocidos como el primer sonido cardíaco, "S1" y el segundo sonido cardíaco "S2". La aparición de los sonidos pulmonares (LS) es también un proceso cíclico formado por dos eventos principales, la inhalación y la exhalación. En ambos casos, las señales HS y LS, tienen silencios entre sus eventos principales, que son diferentes según el evento que tenga lugar antes y después.

Por lo tanto, ambos casos son secuencias de eventos que pueden variar según las circunstancias, las condiciones de salud e incluso el estado de ánimo de la persona. En el análisis de HS, S1 y S2, así como los silencios, son las etapas de la señal a modelar, mientras que en LS, la inhalación, la exhalación y los silencios son las etapas de la señal a modelar. Luego, estas se modelan a través de los Modelos Ocultos de Markov (HMM), que es un autómata de estado finito.

Un HMM puede considerarse como un modelo basado en estados, en el que cada estado no se observa directamente, de hecho, cada estado se caracteriza por un GMM. Formalmente, se explica exhaustivamente en [18]. Como en el caso de los modelos GMM, los HMM se expresan convencionalmente por una tripleta $\lambda = (A, B, \pi)$, como se observa en la Fig. 1.

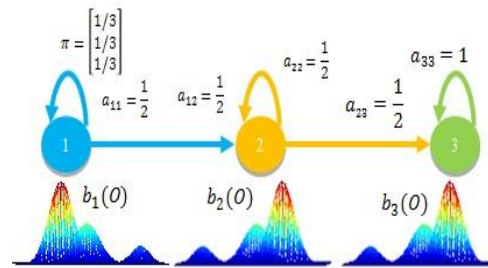


Fig. 1. Modelo HMM con $\lambda = (A, B, \pi)$.

Donde A es una matriz que contiene la probabilidad de transitar de un estado a otro, B (en nuestro caso GMM) es la probabilidad de que un vector acústico (MFCC o vector cuartil) se haya generado a partir de un estado, y π es la probabilidad de iniciar en algún estado. El entrenamiento de los parámetros HMM, dado un conjunto de secuencia de observación (vectores acústicos) $\{O_i\}$, normalmente se hace aplicando el algoritmo Baum-Welch [18], que obtiene los parámetros maximizando la probabilidad $P(O_i | \lambda)$. Para la etapa de evaluación, se requiere calcular $P(O | \lambda)$, dado el modelo λ y una secuencia O de observaciones; aquí se aplicó el algoritmo de avance-retroceso (normalmente conocido en inglés como *forward-backward*) [18].

2.7. Transformada de Hilbert

La transformada de Hilbert de una señal de tiempo continuo con valor real $y(t)$, se define como:

$$H\{y(t)\} = \hat{y}H(s) = \frac{1}{\pi} \oint_{-\infty}^{\infty} \frac{y(t)}{t-s} dt, \quad (6)$$

donde s es real y $H\{\cdot\}$ representa al operador de Hilbert. Aquí, la integración debe llevarse a cabo de acuerdo con el principio de Cauchy, es decir:

$$\oint_{-\infty}^{\infty} \frac{y(t)}{t-s} dt = \lim_{\varepsilon \rightarrow 0} \left(\int_{-\infty}^{s-\varepsilon} + \int_{s+\varepsilon}^{\infty} \right) \frac{y(t)}{t-s} dt, \quad \varepsilon > 0. \quad (7)$$

Sin embargo, las señales en el mundo real son discontinuas y deben ser expresadas como señales de tiempo discreto [6].

2.8. Detección de S1 y S2

Inicialmente, las señales de HS se escalaron en el rango ± 1 , luego se aplicó un filtro pasa bajo Butterworth con una frecuencia de corte de 150 Hz y de orden 10. Después de eso, las zonas S1 y S2 se segmentaron manualmente, y estas secciones se guardaron en formato wav. Del mismo modo, las zonas de silencio (que tenían ruido) se segmentaron, y estas secciones se guardaron en formato wav. El siguiente paso fue calcular un modelo GMM para señales S1 y S2, al mismo tiempo que se calculaba un modelo GMM para zonas de silencio, pero previamente todas las señales se transformaron en vectores de características. Se aplicó PCA sobre todos los vectores obtenidos, y se determinó visualmente que en el caso de MFCC solo los primeros 4 componentes eran necesarios. Los dos modelos constituían el sistema VAD, que se aplicaba a todas las señales para detectar y segmentar automáticamente S1, S2 y las zonas de silencio. Después de la detección automática de S1 y S2, se calculó la envolvente de Hilbert de la señal, donde se comparó la envolvente con la señal original normalizada. A continuación, la envolvente de Hilbert se suavizó con un filtro de Butterworth de quinto orden, con frecuencia de corte de 7-25 Hz (después de varios experimentos se seleccionó 8 Hz como la mejor opción) [6]. VAD detecta S1 y S2, pero sin diferenciarlos entre sí. Para diferenciar S1 de S2 es necesario calcular umbrales. Los umbrales se calculan para discriminar amplitudes y establecer que picos corresponden a S1 o S2. Esto se logra aplicando mínimos y máximos, considerando un mínimo como el inicio de un pico y otro mínimo como el final del mismo pico. Es importante distinguir entre los segmentos sistólico y diastólico,

por eso S1 y S2 podrían identificarse, y poder determinar cuál fue el primero en el registro de la señal. El propósito es separar la sístole y la diástole en pares o señales impares, cuando la señal comienza con S1, la señal corresponde a la sístole, mientras que en las señales pares corresponden a la diástole; si la señal comienza con S2, las señales pares corresponden a la sístole y las impares corresponden a la diástole. A partir de estas señales, es posible obtener las duraciones de tiempo (ancho) de S1 y S2.

2.9. Base de datos

RALE consiste en un conjunto de sonidos LS normales y adventicios como grabaciones en formato *.wav, que se desarrolló en Winnipeg, Canadá [19]. Las señales de RALE se filtraron con paso de banda de 7.5 Hz para suprimir cualquier componente de DC mediante un filtro Butterworth de primer orden. Además, se aplicó un filtro pasa baja Butterworth de octavo orden a 2.5 kHz para evitar traslape. Las señales de la base de datos se muestrearon a 11025 Hz. Las señales de LS Normal de RALE fueron segmentadas manualmente (por los autores) para obtener solo los segmentos de inhalación-exhalación de cada señal, haciendo un total de 13 inhalaciones y 13 exhalaciones.

Las señales HS utilizadas provienen de bases de datos que están disponibles para fines académicos y científicos [9, 20]. Por lo tanto, se usaron 13 señales normales, que se segmentaron manualmente, obteniendo 13 sonidos de S1 y 13 de S2. La frecuencia de muestreo fue de 11025 Hz, con formato *.wav y mono. Las señales utilizadas se dividieron en conjunto de entrenamiento y conjunto de evaluación. La evaluación se realizó con validación cruzada, aquí se deja una señal para evaluar, y se usan las señales restantes para calcular el modelo; este proceso es repetitivo cambiando las configuraciones hasta agotar todas las posibilidades.

Utilizando datos experimentales, se encontró que los intervalos de longitud de tiempo de las fases de señal del repositorio de datos RALE eran aproximadamente de 1.5 segundos para la inhalación y de 2.5 segundos para la exhalación. Con respecto al HS, la fase S1 dura aproximadamente 0.1 a 0.12 segundos; La fase S2 está entre 0.8 y 0.14 segundos [21].

2.10. Clasificación

La evaluación se llevó a cabo con validación cruzada, donde cada señal de inhalación y exhalación es leída y luego procesada con PCA. Debido a que se aplica validación cruzada, los conjuntos de señales de inhalación y exhalación se usan para calcular modelos por clase y su evaluación. Sin embargo, debe tenerse en cuenta que, en la validación cruzada, durante cada prueba se utiliza una señal del conjunto n ($n = 13$) para la evaluación, mientras que las señales $n-1$ restantes se utilizan para calcular el modelo. Se realizan n evaluaciones, pero en cada evaluación, se cambia la señal de prueba y las $n-1$ señales restantes se usan para calcular el modelo.

3. Resultados

Una vez aplicadas las metodologías descritas anteriormente, es importante verificar si son realmente efectivas, por lo cual se llevaron a cabo experimentos de clasificación

aplicando Modelos Ocultos de Markov (HMM). La detección automática con VAD-GMM arroja un gráfico que se muestra en la Fig. 2, al mismo tiempo esto nos da el índice inicial y final de los eventos extraídos de estos datos. Para el primer experimento de clasificación se utilizó una base de datos con 13 señales de inhalación y 13 señales de exhalación.

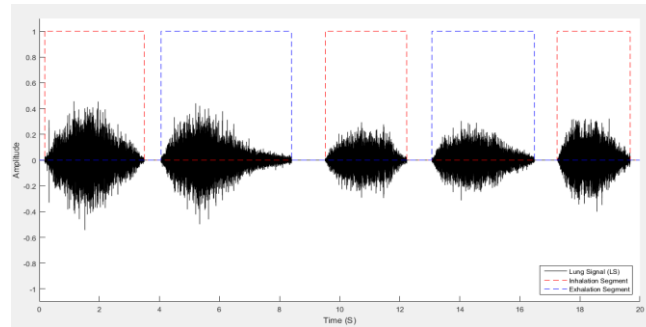


Fig. 2. Segmentación automática de LS con algoritmo VAD-GMM.

Para evaluar la eficiencia de este proceso, se experimentó con diferentes configuraciones de arquitectura de modelos HMM, así como dos tipos de vectores acústicos (cuartiles y MFCC). Los resultados de la eficiencia de clasificación se muestran en la Tabla 1, donde solo se muestran los mejores resultados.

Tabla 1. Eficiencia de clasificación aplicando VAD [4] y detección automática en LS.

Núm. de Estados	Núm. de Gaussianas	Vectores Acústicos	Eficiencia de Clasificación
3	3	Cuartiles	77.5%
3	3	MFCC	76.25%
2	3	Cuartiles	75%
2	3	MFCC	70%

Para el segundo experimento, se utilizaron los mismos conjuntos de señales LS que el experimento anterior, pero en este caso se aplicó VAD-GMM propuesto, como se muestra en la Tabla 2.

Tabla 2. Eficiencia de clasificación con detección automática en LS aplicando VAD-GMM.

Núm. de Estados	Núm. de Gaussianas	Vectores Acústicos	Eficiencia de Clasificación
3	3	Cuartiles	85.625%
3	3	MFCC	95%
2	3	Cuartiles	88.125%
2	3	MFCC	91.25%

En el tercer experimento, se utilizó un conjunto de señales S1 y S2 obtenidas por detección automática de HS (Fig. 3). Así como en la Tabla 2 se aplicó VAD-GMM; el mejor resultado de clasificación fue 92.6223% como se muestra en la Tabla 3. Como se puede ver en la Tabla 3, ambos vectores acústicos se usaron para calcular modelos,

con dos configuraciones de arquitecturas HMM-GMM y los mejores resultados se obtuvieron con 3 estados y 3 gaussianos por estado.

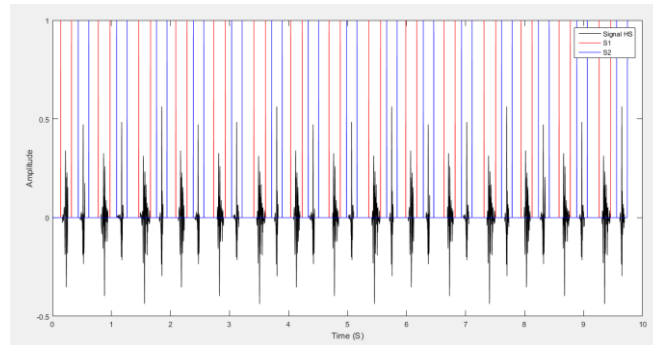


Fig. 3. Segmentación automática de señales HS con el algoritmo VAD-GMM.

Tabla 3. Eficiencia de clasificación con detección automática en HS y con VAD-GMM.

Núm. de Estados	Núm. de Gaussianas	Vectores Acústicos	Eficiencia de Clasificación
3	3	Cuartiles	92.6923
3	3	MFCC	90.3846%
2	3	Cuartiles	87.3077%
2	3	MFCC	91.25%

Como se observa en la Tabla 1, el mejor resultado de clasificación se obtuvo con cuartiles, es 77.5%, mientras que con VAD-GMM se logró 95% como se muestra en la Tabla 2, y esto demuestra la superioridad de VAD-GMM. Incluso en los eventos de HS (S1 y S2) la detección y extracción con VAD-GMM tuvieron un buen rendimiento, lo que significa que VAD-GMM puede determinar a qué evento pertenece un vector, sin importar el tipo de vector (cuartil o MFCC). La utilización de GMM mejora la capacidad de VAD para asociar un vector acústico con su clase. Esto es razonable ya que en los esquemas de reconocimiento de locutor, la identificación de la clase a la cual pertenece un vector acústico, utilizan sistemas VAD. Además de la detección y extracción automáticas, VAD-GMM podría proporcionar un registro documentado para el monitoreo a largo plazo y el análisis comparativo.

4. Conclusiones

La detección automática de eventos acústicos en LS aplicando VAD sería útil en diagnóstico respiratorio, en particular para documentar eventos y monitoreo a largo plazo. Debido al aumento en los casos de EPOC y que no se conoce una cura, un mejor manejo en datos a largo plazo y con medicamentos selectivos aumentaría la calidad general de atención médica.

Los resultados obtenidos son prometedores, ya que la eficiencia de clasificación aumentó aplicando detección automática con VAD-GMM. Con el primer algoritmo de

VAD obtenido de la literatura, se logró el 77.5% de la eficiencia, mientras que con VAD-GMM propuesto se alcanzó un 95% de eficiencia.

Por lo tanto, nuestra propuesta podría mejorar la auscultación tradicional cualitativa apoyada en la evaluación cuantitativa y al diagnóstico automatizado asistido mediante la identificación de indicadores acústicos anormales. Además, este método podría usarse en medicina general.

Referencias

1. Fatma Ayaria, M. K., Alouanib, A.T.: Lung sound extraction from mixed lung and heart sounds FASTICA algorithm. Electrical Engineering Department, National School of Engineers Tunis (ENIT), Electrical Engineering Department, Tennessee Technological University (2012)
2. Ghaderi, F., Mohseni, H. R., Sanei, S.: Localizing Heart Sounds in Respiratory Signals Using Singular Spectrum Analysis. Biomedical Engineering, IEEE Transactions on, vol. 58, pp. 3360–3367 (2011)
3. Sohn, J., Sung, W.: A Statistical Model-Based Voice Activity Detection. IEEE signal processing letters, vol. 6 (1999)
4. Faezipour, A. A. a. M.: Acoustic Signal Classification of Breathing Movements to Virtually Aid Breath Regulation. IEEE journal of biomedical and health informatics, vol. 17 (2013)
5. Wang, Y. X. H. , Li, M.: Study on the MFCC similarity-based voice activity detection algorithm. In: Proc. 2nd IEEE Int. Conf. Artif. Intell. Manage. Sci. Electron. Commerce, pp. 4391–4394 (2011)
6. Ashok Monda, P. B. a. G. S.: An automated tool for localization of heart sound components S1, S2, S3 and S4 in pulmonary sounds using Hilbert transform and Heron's formula. Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, India (2013)
7. Istrate, D. M.: Detection et Reconnaissance des Sons pour la Surveillance Médicale. These pour obtenir le grade de docteur de l'INPG: spécialité Signal, Image, Parole, Télécoms docteur Institut National Polytechnique de Grenoble, Grenoble, France (2003)
8. Mayorga, P., Druzgalski, C., Vidales, J.: Quantitative Models for Assessment of Respiratory Diseases. Pan-American Health Care Exchange (PAHCE), pp. 25–30 (2010)
9. Boston Children's Hospital [Online]. Available: <http://www.childrenshospital.org/>
10. Pearce, D.: Developing the ETSI Aurora advanced distributed speech recognition front-end and what next? In: Automatic Speech Recognition and Understanding, ASRU '01. IEEE Workshop on, pp. 131–134 (2001)
11. Mayorga, P., Besacier, L., Lamy, R., Serignat, J. F.: Audio packet loss over IP and speech recognition. In: Automatic Speech Recognition and Understanding, ASRU '03, IEEE Workshop on, pp. 607–612 (2003)
12. Proakis, J. G., Salehi, M.: Digital Communications. 5th ed. (2008)
13. Mayorga, P., Druzgalski, C., González, O. H., Lopez, H. S.: Modified classification of normal Lung Sounds applying Quantile Vectors. In: Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE, pp. 4262–4265 (2012)
14. Bimbot, F. , Bonastre, J.-F., Fredouille, C., Gravier, G. , Magrin-Chagnolleau, I., Meignier, S., et al.: A Tutorial on Text-Independent Speaker Verification. EURASIP Journal on Advances in Signal Processing, vol. 2004, pp. 1–22 (2004)
15. Martinez W. L., Martínez A. R.: Computational Statistics Handbook with Matlab. Second ed., (2008)
16. Martinez W. L., Martínez A. R., Solka, J. L.: Exploratory Data Analysis with MATLAB. Second Edition (2011)

17. Mayorga Ortiz, P., Druzgalski, C., Criollo Arellano, M. A., González Arriaga, O. H.: GMM y LDA aplicado a la detección de enfermedades pulmonares. *Revista mexicana de ingeniería biomédica*, vol. 34, pp. 131–144 (2013)
18. Rabiner, L. R., Juang, B. H.: *Fundamentals of speech recognition*. Englewood Cliffs, N.J.: PTR Prentice Hall (1993)
19. RALE, Lung Sounds Repository. [Online]. Available: <http://www.rale.ca/Repository.htm>
20. Texas Heart Institute. [Online]. Available: <http://www.texasheartinstitute.org/AboutUs/index.cfm>
21. Echeverry, J. D. , López, A. F., López, J. F.: Reconocimiento de valvulopatías cardíacas en señales de fonocardiografía empleando la transformada Gabor. *Scientia Et Technica*, vol. XIII, pp. 139–144 (2007)

RFID Tag for Toll-Collection Systems

Sergio Buenrostro Rocha¹, Roberto Herrera Charles¹, José Luis Medina Monroy²,
Andrés Calvillo Téllez¹

¹ Instituto Politécnico Nacional, Centro de Investigación y Desarrollo de Tecnología Digital,
Tijuana, Baja California, Mexico

sbuenrostro@citedi.mx, robcharles@citedi.mx, calvillo@citedi.mx

² Centro de Investigación Científica y de Educación Superior de Ensenada, Ensenada, B.C.,
Mexico

jmedina@cicese.mx

Abstract. The installation of an RFID tag on a vehicle's windshield for Automatic Vehicle Identification AVI applications has become a challenge for RFID users and service providers, since in some situations the tag is not read properly because some customers do not follow instructions of the recommended locations for installation. This problem is because each location of the windshield provides different distances to the car hood, car roof and A-pillars, and these metallic surfaces directly interact with the radiation resistance of the tag antenna and consequently change its performance (tag read range, frequency, radiation pattern, etc). There are locations in the windshield where the read range of the tag is severely reduced and in others it is greatly benefited. In the present work, an RFID tag for vehicle windshield with the Higgs 3 chip is developed. The tag is designed to operate at the 865 MHz – 868 MHz frequency band and it can be used in Europe, Asia and South Africa. It was observed that its read range is longer at any point of the windshield than that provided by the commercial RFID tag SMARTRAC DogBone with Monza 4 chip that has been well accepted internationally. In order to design the tag, the modeling of the antenna was performed by electromagnetic analysis using the method of moments, then it was prototyped and characterized directly in vehicles. It was also performed an analysis of the different locations of the windshield in order to find the best and worst locations to install the tag.

Keywords: RFID, tag, toll.

1 Introduction

In recent years, Radiofrequency Identification (RFID) systems have attracted much attention, particularly in the ultra-high frequency (UHF) band [1]. Radio Frequency Identification (RFID) is a wireless technology used for the identification and collection of data of objects, people or animals in digital form and automatically or manually. RFID is an alternative to bar code, but with superior capabilities since the readings are made using electromagnetic waves. A basic RFID system consists of a base station called reader or interrogator and a set of remote transponders called RFID tags. A basic interrogator consists of a reader antenna, a coaxial cable, RFID reader, network cable and a computer with dedicated software to manage the readings of the tags [2, 4, 8, and

9]. A basic RFID system is shown in Figure 1, where it is also possible to observe how an RFID tag is attached to an object to be identified.

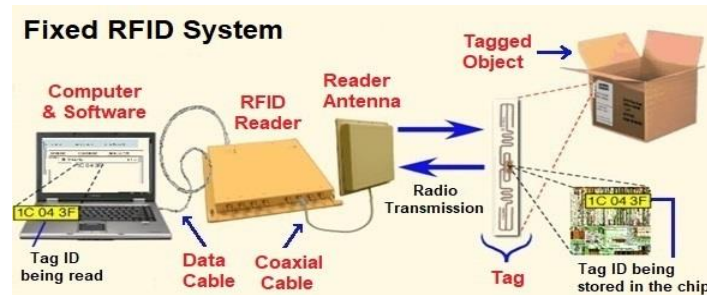


Fig. 1. Basic RFID System.

A typical RFID tag consists of an antenna and an application specific integrated circuit chip (ASIC) which contains an internal read/write memory and can have dimensions of at least 0.25mm^2 . The tag antenna may be printed on a dielectric substrate that may be flexible or non-flexible, planar or non-planar. The size of an RFID tag can measure some few centimeters, and most of the times it is desired to be as small as possible. RFID tags can be passive (they do not require battery and are completely powered by the power emitted by the reader), semi-passive and active (require battery) [2, 8, 9, and 10]. As an example, it is shown in figure 2 the composition of a basic RFID tag. The size of this tag is approximately 98 mm x 12 mm.

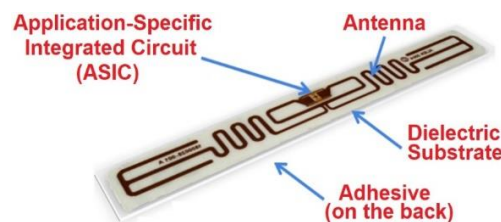


Fig. 2. Composition of an RFID tag.

The number of applications for RFID systems has grown rapidly in industries such as the Internet of Things, access control, inventory control, distribution logistics, security systems, manufacturing process control, medical and biological applications, patient identification, localization and monitoring in hospitals, supermarkets, shops, baggage screening at airports, livestock, human, vehicular application, railway, etc. [1, 18].

Much effort is now being devoted to the development of various UHF applications and solutions. Standardization organizations are working with governments to harmonize UHF frequencies. However, the operating frequency range varies from region to region. UHF bands are currently divided into four globally used regions: China (840-845 MHz), Europe and Asia (865-868 MHz), America (902-928 MHz), Japan (950-956 MHz) [3, 8, 17, 18, and 19].

The most important RFID system performance feature is the read range of the tags, and this is the maximum distance of the RFID reader in which it still can read or write information on the tag. The read range of the tag is defined with respect to a certain speed or successful reading/writing percentage which tends to decrease until it becomes zero as the tag moves away from the reader antenna. Generally, it is preferred that the antenna of the tag has the property of omni-directional radiation to ensure identification from all directions [9].

A good impedance matching between the antenna and the chip must be achieved to assure a maximum power transfer between the antenna and the tag chip. The AVI System that uses RFID emerges as a faster and more effective solution to the manual method of toll collection, where problems of vehicle congestion and delay can be experienced. Automatic collection is done through a pre-paid account assigned to the RFID tag belonging to the owner of the vehicle. This makes the toll transaction more convenient for public use [20].

Figure 3 shows how it works an electronic toll collection (ETC) [21]: When the vehicle approaches the tolling booth, the motion sensors detect the vehicle, then it is initialized the RFID reading system. Once the tag is detected and the vehicle is automatically classified by other sensors, the toll collection is executed and then the traffic barrier is lifted, so that traffic circulates in this way. If there was no tag detected the cameras will take pictures of the license plate for processing.

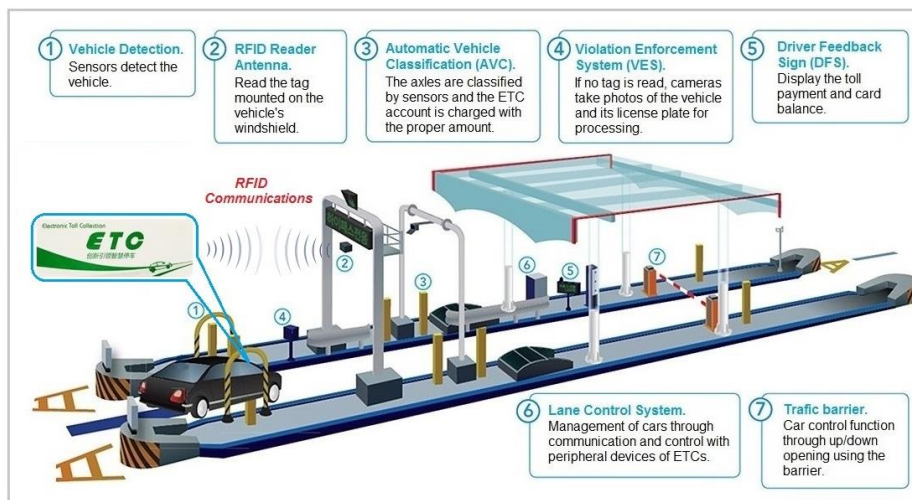


Fig. 3. Toll Booth.

There exist other solutions that can be utilized in Electronic Toll Collection, and these are the use of RFID headlight tags and license plate RFID tags. The inconvenience of the headlight tag is that it is attached at the exterior of the headlight, so its life time is shorter since is exposed to the outdoors environment, humidity, mud, etc. In addition, the disadvantage of the license plate RFID tags is that its cost is higher, since it requires a plastic housing. For these reasons, it has become more popular to use RFID windshield tags, since it is attached in the interior of the vehicle.

2 Antenna Design and Electromagnetic Simulations

Figure 4 shows the proposed windshield tag antenna and it was drawn and modeled in Momentum from Advanced Design System. The dimensions of the tag antenna are 70 mm x 20 mm. The chip used was Alien Higgs 3 with an impedance in its port of $18.58 - j165.94$ ohms for the frequency of 866.6 MHz. So, in order to obtain a good impedance matching between antenna and the chip, the antenna must be designed such that, it presents in its port the conjugate complex impedance of the chip, and in this case is $18.58 + j165.94$ ohms. The antenna tracks were modeled with a conductivity of 5,000,000 Siemens/m, with a thickness of 0.004 mm. The antenna inlay was printed on PET with a thickness of 0.0508 mm and with a relative dielectric constant of 3.2. Also, the tag is modeled on a surface of a glass, with thickness of 4.4 mm and with a relative dielectric constant of 5.0.



Fig. 4. Proposed windshield tag.

Figure 5 shows the simulated impedance matching and return loss of the proposed windshield tag antenna. It can be observed that the antenna was presenting in its ports a very good impedance matching of $19.052 + j165.705$ at 866.6 MHz with a return loss of -37 dB.

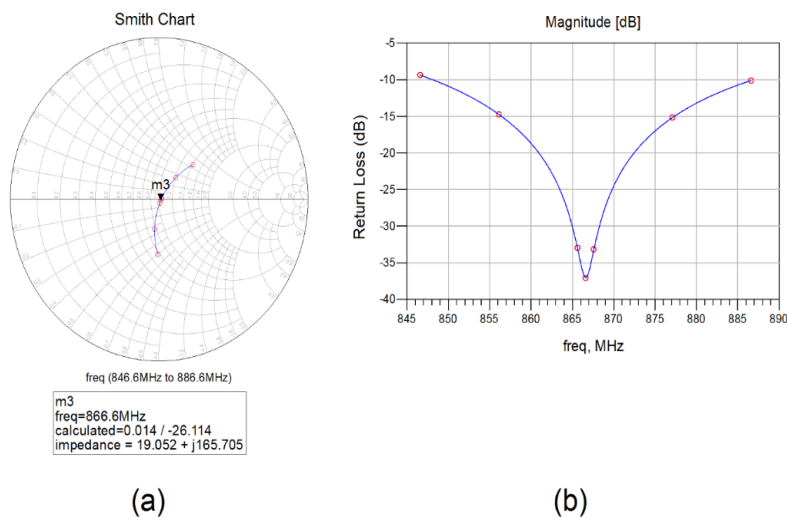


Fig. 5. Impedance Matching (a) Smith Chart and (b) Return Loss.

3 Prototyping and Characterization

As it can be noticed, the tag was only modeled for glass and the real application where it was desired to make it operate properly is on vehicle windshields. After the electromagnetic analysis, the tag was prototyped and tested on different locations of the windshield.

Table 1. Testing conditions and equipment.

Characteristic	Specification
Reader	Sirit INfinity 610 UHF RFID
Frequency	866.6 MHz
ERP	2 Watts
Reader Antenna	Mti - MT-242020/NV; Gain: 9.5 dBi; Linear; 865 - 870 MHz
Reader Antenna Height, angle and localization	56" (floor to the mid center of antenna). Perpendicular to the Pavement. Reader Antenna in front of vehicle.
Vehicle	Acura TL 2016

The optimization basically consists in adding or removing segments of conductor on critical sections of the prototype. So, the CAD drawing have to be updated, re-prototyped, tested and validated until the read range is increased. The criteria for windshield tag optimization were focusing in improve the performance given by the worst location, and this, indirectly, improved the performance given by the rest of locations. So, this way the design fulfilled a commitment of performance between the different critical and important locations on the windshield. On the other hand, in order to optimize the radiation pattern, this process can be repeated for different angles and heights of the reader antenna with respect to the front of the vehicle. Once discovered that the tag cannot achieve even larger reading ranges, the optimization is over.

Figure 6 illustrates the important locations of a windshield where the tag may be installed and table 2 shows the corresponding performances obtained from characterizing the sensitivity of the tag and with this, it was calculated the experimental read range.

The experimental read range RRE [m] was determined as follows [22]:

$$EIRP = P_{Tx} G_{Tx}, \quad (1)$$

where EIRP is Equivalent Isotropic Radiated Power (note that the maximum allowed value for Euro Asiatic RFID standard is 3.28 W), P_{Tx} [dBm] is the transmit power, G_{Tx} [dB] is the reader antenna gain:

$$RRE = d \sqrt{\frac{EIRP}{P_{min} G_{Tg} L_c}}, \quad (2)$$

where d [m] is the distance between reader antenna and tag antenna, and $L_c = -2$ dB is measured cable losses. The RFID reader provides an output power from 5 dBm to 32 dBm at the operating frequency of 865–868 MHz. A query command can be sent to the

tag (mounted face to face to the reader antenna), which replies with its identification code. The minimum transmit power required to activate the tag is P_{min} and G_{Tg} [dB] is the tag antenna gain.

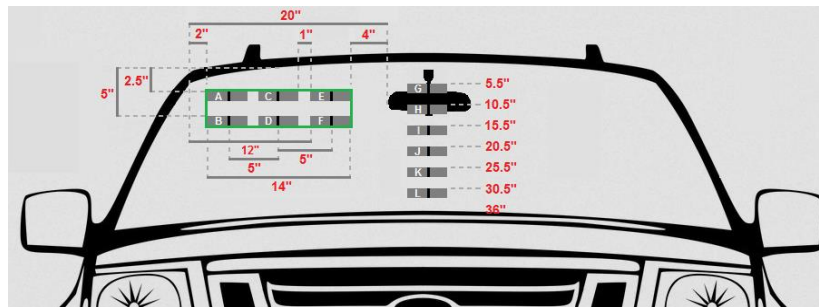


Fig. 6. Locations of the tag on the windshield.

Table 2. Experimental read ranges (meters) of the proposed windshield tag vs Smartrac Dogbone on the different locations of the windshield. (*Med means Median and *Ave means Average).

Tag Location Design	A	B	C	D	E	F	G
Proposed	16.8	13	15.6	11.5	16.6	16	16.6
Smartrac Dogbone	14.4	11.2	11.3	9.1	12.7	12.6	12.7
Tag Location Design	H	I	J	K	L	Med	
Proposed	8.7	11.7	12.1	4.7	7.9	12.6	
Smartrac Dogbone	7.5	9	9.2	4.3	6.7	10.2	

* The performance of the tag may slightly vary due:

- I. Different Windshield Thicknesses and Protective Films on Different Vehicles.
- II. Shape of the vehicle.
- III. Doors and Windows Opened or Closed.
- IV. The Car is Manned or Not.
- V. If Manned, Different Body Mass of the Person and Proximity to the Tag.
- VI. Position, Angle and Height of the Reader Antenna with Respect the Vehicle.
- VII. Position, Angle and Height of the Tag with Respect the Antenna.
- VIII. Etc.

4 Discussion and Conclusion

Evidently, the performance given by the proposed design surpassed the performance given by the tag Smartrac Dogbone. It is a mistake to tune a windshield tag only for glass, due to the existence of metal surfaces around of the windshield that considerably affect the tag antenna performance and these needs to be considered on the design. It was found that the best location to install the tags is close to the top corners of the windshield and the worst is in the center of the windshield.

References

1. Finkenzeller, K.: RFID Handbook: Radio-Frequency Identification Fundamentals and Applications". 2nd ed., Hoboken, NJ: Wiley (2004)
2. Nikitin, P. V.; Rao, K.V.S.: Theory and measurement of backscattering from RFID tags. IEEE Antennas and Propagation Magazine, 2006, vol. 48, no 6, pp. 212–218.
3. Rao, K.V.S., Nikitin, P.V., Lam, S. F.: Antenna design for UHF RFID tags: A review and a practical application. IEEE Transactions on antennas and propagation, vol. 53, no 12, p. 3870–3876 (2005)
4. Marrocco, G.: The art of UHF RFID antenna design: Impedance-matching and size-reduction techniques. IEEE antennas and propagation magazine, vol. 50, no. 1 (2008)
5. Loo, Chye-Hwa, et al.: Chip impedance matching for UHF RFID tag antenna design. Progress in Electromagnetics Research, vol. 81, pp. 359–370 (2008)
6. Chen, Sung-Lin, Lin, Ken-Huang: Characterization of RFID strap using single-ended probe. IEEE Transactions on Instrumentation and Measurement, vol. 58, no 10, pp. 3619–3626 (2009)
7. Mabrouk, M., et al.: A Broadband UHF TAG Antenna for Near-Field and Far-Field RFID Communications. Laboratoire GRESCOM, SUPCOM de Tunis, University de Carthage, Radioengineering, vol. 23, no. 4 (2014)
8. Mun, Byeonggi, et al.: A compact dual-band RFID tag antenna mountable on metallic objects. International Journal of Antennas and Propagation, vol. 2015 (2015)
9. Turcu, C., (ed.): Radio Frequency Identification Fundamentals and Applications: Design Methods and Solutions. [Online] Sciyo.com (2010)
10. Ranasinghe, D. Ch., et al.: Small UHF RFID label antenna design and limitations. In: IEEE International Workshop on Antenna technology: Small Antennas and Novel Metamaterials (2006)
11. Cho, Chihyun, Park, Ikmo, Choo, Hosung: Design of a circularly polarized tag antenna for increased reading range. IEEE Transactions on Antennas and Propagation, vol. 57, no. 10, pp. 3418–3422 (2009)
12. Chen, Hong-Dean, et al.: Compact Circularly Polarized Meandered-Loop Antenna for UHF-Band RFID Tag. IEEE Antennas and Wireless Propagation Letters, vol. 15, pp. 1602–1605 (2016)
13. Lee, J.-W., Lee, B.: Design of high-Q UHF radio-frequency identification tag antennas for an increased read range. IET microwaves, antennas & propagation, vol. 2, no. 7, pp. 711–717 (2008)
14. Chen, Hong-Dean, et al.: Circularly polarized loop tag antenna for long reading range RFID applications. IEEE Antennas and Wireless Propagation Letters, vol. 12, pp. 1460–1463 (2013)
15. Zhang, Jun, Long, Yunliang: A novel metal-mountable electrically small antenna for RFID tag applications with practical guidelines for the antenna design. IEEE Transactions on Antennas and Propagation, vol. 62, no. 11, pp. 5820–5829 (2014)

16. Nikitin, P.V., et al.: Power reflection coefficient analysis for complex impedances in RFID tag design. *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 9, pp. 2721–2725 (2005)
17. Andrenko, A. S.: Conformal fractal loop antennas for RFID tag applications. In: *Applied Electromagnetics and Communications, ICECom 2005, IEEE 18th International Conference on*, pp. 1–6 (2005)
18. Santiago, A. G., Costa, J. R.; Fernandes, C. A.: Broadband UHF RFID passive tag antenna for near-body applications. *IEEE Antennas and Wireless Propagation Letters*, vol. 12, pp. 136–139 (2013)
19. Zamora, G., et al.: Design and synthesis methodology for UHF-RFID tags based on the T-match network. *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 12, pp. 4090–4098 (2013)
20. Li, Chonghua: Automatic vehicle identification (AVI) system based on RFID. In: *Anti-Counterfeiting Security and Identification in Communication (ASID), IEEE International Conference on*, pp. 281–284 (2010)
21. Kalantri, R., Parekar, A., Mohite, A., Kankapurkar, R.: RFID based toll collection system. *International Journal of Computer Science and Information Technologies* 5, no. 2: 2582–5 (2014)
22. Liu, Hsien-Wen, Yang, Chang-Fa, Ku, Chia-Hao: Novel miniature monopole tag antenna for UHF RFID applications. *IEEE Antennas and Wireless Propagation Letters* 9: 363–366 (2010)

Sistema con base en el Internet de las cosas para el control de riego en la agroindustria

Brenda Lilia Castro Coronado, Arnoldo Díaz-Ramírez, Verónica Quintero Rosas, Mario Alberto Camarillo Ramos, Jorge Antonio Atempa Camacho

{a12490431, adiaz, veronicaquintero, mario.camarillo, jatempa}@itmexicali.edu.mx

Resumen. Las limitaciones de agua disponibles para los cultivos representa un tema de gran interés para la agroindustria. Maximizar la producción al mismo tiempo que se utiliza la cantidad óptima de agua es un problema que puede resolverse con el uso de las tecnologías de la información y comunicación. En este artículo se propone un sistema de riego de precisión que hace uso del paradigma del Internet de las Cosas. Los resultados de la evaluación del sistema propuesto permiten comprobar que puede contribuir en el uso eficiente del agua, al mismo tiempo que se mejora la producción.

Palabras clave: Internet de las cosas, agroindustria, control de riego.

System based on the Internet of Things for Irrigation Control in Agribusiness

Abstract. The limitations on the available water supply for the agroindustry is an issue of maximum concern. The goal of maximizing productivity using the optimum amount of water is a problem that can be addressed using the information and communications technologies. In this paper, a precision irrigation system, based on the Internet of Things paradigm, is introduced. The evaluation of the proposed system proves that it could be helpful in providing an efficient use of the available water, while the productivity is improved.

Keywords: Internet of things, agroindustry, irrigation control.

1. Introducción

La cantidad de agua existente en el suelo es fundamental para la producción eficiente de productos agrícolas. Monitorizar el nivel de humedad en el subsuelo de los cultivos es fundamental en la optimización de la producción, el uso racional

del agua, la reducción de costos y el cuidado del medio ambiente. Por lo tanto, existe un gran interés en investigar el uso de la tecnología en el campo de la agroindustria, que ayude en la monitorización de los cultivos y el control de los procesos de producción. Una de las áreas que mas atención han recibido recientemente se denomina riego de precisión, que con base en el uso de la tecnología puede establecer el momento, la frecuencia y el tiempo de riego adecuados según las características del cultivo.

La falta de agua en los cultivos provoca un decremento en la calidad del producto, en el nivel de producción y eventualmente, la muerte de la planta. Por otra parte, el exceso de riego implica un desperdicio de agua, e incluso la muerte del cultivo. El riego de precisión permite determinar la cantidad y frecuencia óptima del riego.

En este artículo se propone un sistema de riego de precisión, que con base en el paradigma del Internet de las Cosas (IoT), pretende ser un apoyo en el control de riego en la agroindustria, optimizando el consumo del agua mediante la monitorización constante de la humedad del suelo.

El resto del documento está organizado de la siguiente manera: la Sección II discute la problemática del riego. En la Sección III se describe brevemente el paradigma del Internet de las cosas, mientras que la Sección IV describe el sistema propuesto. La Sección V está dedicada a discutir los resultados obtenidos de la evaluación del sistema propuesto. Finalmente, en la Sección VI se presentan las conclusiones y el trabajo futuro.

2. Problemática

La agricultura es responsable del 70 % de extracciones de agua dulce a nivel mundial [?], un recurso que en todas sus formas constituye solo el 3 % del agua del mundo entero, mientras que el otro 97 % es agua salada. Del 3 % de agua dulce, aproximadamente del 2 % está confinada en los casquetes glaciares de la Antártida y Groenlandia, a demasiados metros bajo tierra para ser accesible, o está contenida en la humedad del suelo, dejando tan solo el 1 % de agua dulce del mundo disponible para uso humano. Además, el rápido desarrollo socioeconómico y urbanístico en los últimos años han hecho que el agua dulce sea insuficiente para satisfacer las necesidades básicas de la humanidad, en términos de usos agrícolas, industriales y urbanos. La escasez de agua se vuelve más aguda en regiones áridas y semiáridas debido a la distribución desigual de los recursos hídricos, tanto en el tiempo como en el espacio, y la competencia más fuerte para la demanda de agua entre diferentes sectores. Este problema puede reducirse utilizando el riego de precisión.

La agricultura de precisión requiere un conocimiento en tiempo real de los procesos en todas las áreas del campo. Las tecnologías utilizadas en el riego de precisión todavía están en etapa de desarrollo y constante mejora. El uso del riego de precisión implica el desarrollo de un sistema que gestione la máxima eficiencia en el uso del agua y maximizar la productividad [?].

Algunos de los parámetros básicos relacionados con el manejo eficiente del riego se describen a continuación [?], mismo que se muestran en la Fig. 1.

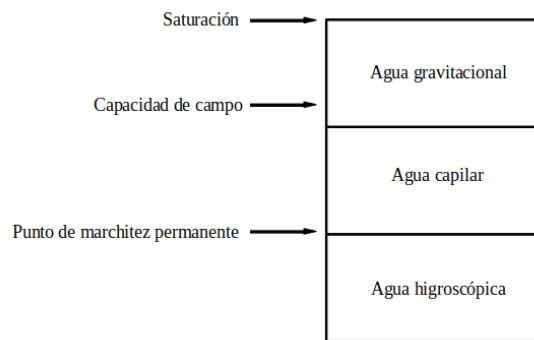


Fig. 1. Parámetros del agua en el suelo y tipos de agua en el suelo.

- Capacidad de campo. Es el contenido de agua o humedad máxima que puede retener el suelo. Esto por supuesto después de haber sido previamente drenado.
- Punto de marchitez permanente. Es el punto de humedad mínima que se alcanza en el que las plantas ya no pueden extraer más agua del suelo. Cuando la humedad está en este punto por mucho tiempo, puede provocarse que la planta ya no se recupere y eventualmente muera.
- Agua disponible para la planta. Es la cantidad de agua retenida en el suelo que esta disponible para que las plantas la extraigan.
- Contenido gravimétrico de agua. Es el contenido directo del agua en una muestra de suelo.
- Densidad aparente. Es la cantidad de masa de una muestra de suelo después de haber sido secada.
- Déficit permitido en el manejo del riego. Es el punto más bajo de contenido de agua en el suelo al cuál no se debe llegar para evitar el estrés en las plantas por falta de agua.
- Profundidad de la raíz. La profundidad a la que llegue la raíz de la planta es la que determina cuál es la cantidad de agua que va a requerir dicha planta.

Un sistema de riego de precisión es capaz de:

- Determinar el tiempo, la magnitud y el patrón espacial del riego para cumplir con el objetivo estacional (es decir, maximizar el rendimiento, la eficiencia del uso del agua o la rentabilidad).
- Aplicar exactamente (o lo más cerca posible) lo que se requiere de agua.

- A través de experimentos de simulación, técnicas de aprendizaje autónomo o medición directa, conocer la magnitud y el patrón espacial de las aplicaciones de riego reales y las respuestas del suelo, de acuerdo a los tipos de cultivos.
- Utilizar estas respuestas para elaborar un plan de riego óptimo.

Adicionalmente, se han propuesto cuatro pasos esenciales en el proceso de control de riego [?], así como el uso algunas tecnologías para automatizar el proceso. Por ejemplo, en [?] se propone una arquitectura con base en IoT para la agroindustria, como puede observarse en la Fig. 2. Dicha arquitectura consta de cuatro componentes:



Fig. 2. Arquitectura propuesta con base en IoT [?].

1. Adquisición de datos. Un sistema de riego de precisión requiere un módulo de adquisición de datos del suelo y las condiciones del cultivo dentro de un campo.
2. Interpretación. Los datos recopilados se requieren interpretar, procesar y analizar.
3. Control. Debe contar con la capacidad de proveer de los recursos necesarios y ajustar el manejo del riego cuando se requiera.
4. Evaluación. Debe de haber una medición del rendimiento del sistema para obtener una retroalimentación que permita realizar mejoras para el próximo ciclo de riego.

La investigación en el riego de precisión inició en Estados Unidos a principios de 1990. Al principio el trabajo se centró en gran medida en la modificación de las máquinas de riego [?, ?, ?]. El interés de los investigadores europeos creció hasta la década de los 2000 y el énfasis se desplazó hacia el rendimiento de riegos.

Los ejemplos de este trabajo incluyen Al-Karadsheh *et al.* [?], Camp *et al.* [?], Chavez *et al.* [?], King *et al.* [?] y Sadler *et al.* [?].

Trabajos recientes (Peters y Evett [?,?,?] y O'Shaughnessy *et al.* [?]) se han expandido para incluir el uso de termómetros infrarrojos montados en pivotes centrales para monitorizar las temperaturas del suelo, desarrollar protocolos para la programación y el control de riego automático en tiempo real. Otro trabajo reciente en la Universidad Estatal de Washington (Camp *et al.* [?]) se ha centrado en el desarrollo y prueba de sistemas de control digital que utilizan una computadora a bordo para implementar instrucciones transmitidas por radio para investigación de parcelas. También un trabajo adicional en Europa (Al-Karadsheh *et al.* [?]) examinó la respuesta de rendimiento a aplicaciones de agua no uniforme bajo sistemas de riego en movimiento. Adicionalmente, en Nueva Zelanda (Yule *et al.* [?], y Hedley y Yule [?,?]) se ha investigado el ahorro de agua y los beneficios económicos del riego de precisión utilizando pivotes centrales.

Algunos proyectos de investigación interesantes relacionados con el riego de precisión han sido desarrollados en Australia. Los grupos clave que participan en éste tema incluyen al Centro Nacional de Ingeniería en Agricultura (CNIA) de la Universidad de Southern Queensland, la Universidad de Melbourne y CSIRO. El CNIA también tiene un amplio programa de trabajo centrado en las tecnologías de detección para el riego de precisión [?].

3. Internet de las cosas

Cuando las computadoras salieron por primera vez al mercado, la tecnología no estaba creciendo tan rápido y no hubo muchas variaciones en los productos por algún periodo de tiempo. A medida que fueron cambiando las necesidades también fue avanzando la tecnología y se crearon soluciones para interconectar las computadoras. En un inicio, tenían que estar conectadas por cables y los usuarios estaban confinados a áreas de trabajo limitadas. Sin embargo, la tendencia se ha desplazado hacia la conectividad inalámbrica debido a la libertad que le da al usuario para moverse sin mucha restricción. Y los últimos años, la conectividad inalámbrica ha sido un área activa de investigación, y se han generado una gran cantidad de iniciativas de parte del gobierno y la industria. Esto ha motivado la creación de nuevos estándares y protocolos para alcanzar el objetivo de uniformidad y cohesión entre todas las tecnologías. Como resultado, hoy se cuenta con un conjunto de diversas de tecnologías de conectividad inalámbrica desde redes satelitales hasta sistemas celulares de área amplia [?,?].

El Internet de las Cosas es un paradigma que permite que los objetos cotidianos, de manera ubicua, sean capaces de monitorizar el entorno y transmitir a las personas u otros dispositivos, información de interés. En el IoT, las *cosas* u objetos pueden ser electrodomésticos, ropa, calzado, vehículos, entre otros. Estos objetos cuentan con sensores, sistemas inmersos, actuadores y radio-transmisores [?].

El paradigma de IoT busca interconectar objetos del mundo físico al mundo virtual, transformando la manera en la que se comunican entre sí, haciendo posible el intercambio de información sin necesidad de la intervención de un humano. Los protocolos de comunicación y estándares comúnmente usados dentro de éste paradigma incluyen: identificación por radio frecuencia (RFID) cerca de campo de comunicación (NFC ej. WLAN, ZigBee, Bluetooth) redes inalámbricas multi-salto/redes de sensores, redes de área personal inalámbrica de bajo consumo bajo IPv6 (6LoWPAN), máquina a máquina (M2M) y otras tecnologías tradicionales como IP, IPv6, etc [?].

En el caso particular de la problemática del riego de precisión, el paradigma de IoT se presenta como una valiosa alternativa para su uso en búsqueda de soluciones eficientes. Los avances en las tecnologías de sensores, sistemas inmersos, sistemas operativos, comunicación inalámbrica y protocolos para integración de IoT con la Computación en la Nube, permiten el diseño de arquitecturas IoT para la agroindustria. En la siguiente sección se presenta un sistema IoT con estas características.

4. Sistema propuesto

El contenido de humedad en las capas superficiales del suelo es un parámetro importante para muchas aplicaciones en las disciplinas de agricultura, hidrología y meteorología. Con la monitorización de la humedad del suelo se puede conocer la cantidad de agua que se encuentra en éste, lo cuál resulta fundamental para los agricultores. Con esta información se puede optimizar la producción definiendo una eficiente programación del riego. Por otra parte, puede determinarse la cantidad de agua a aplicar y cuándo aplicarla, evitando pérdidas de agua excesivas por percolación profunda o por escurrimientos, o bien evitar aplicar una cantidad insuficiente. Además, es posible reducir los impactos ambientales ya que el exceso de irrigación puede incrementar el consumo de energía y el desperdicio de agua, aumentar el movimiento de fertilizantes por debajo de la zona radicular, producir erosión, y provocar transporte de suelo y partículas de químicos a los canales de drenaje [?].

La irrigación es una práctica esencial en muchos sistemas de cultivos agrícolas en áreas semiáridas y áridas, y las aplicaciones y el manejo eficiente del agua son las principales preocupaciones. La determinación de cuándo irrigar requiere estimar el tiempo de riego para que no se produzcan reducciones de rendimiento debido al excesivo agotamiento de la humedad del suelo. Un método para la programación de riego consiste en medir o controlar el contenido de humedad del suelo. Para este propósito, los sensores de humedad del suelo se pueden usar para determinar el intervalo apropiado entre el riego, la profundidad de la humectación, la profundidad de extracción por las raíces y la idoneidad de la humectación [?,?]. Los sistemas de irrigación distribuidos en el campo y con base en el uso de sensores pueden ofrecer una solución potencial para respaldar la administración de irrigación específica.

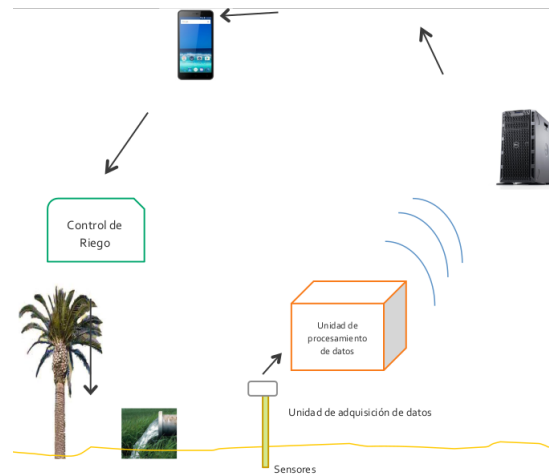


Fig. 3. Arquitectura del sistema propuesto.

En la Fig. 3 puede observarse la arquitectura del sistema propuesto. En primer lugar es necesario determinar los parámetros a monitorizar, relacionados con el nivel de humedad del suelo. Estos parámetros dependerán del tipo de suelo y cultivo. Por ejemplo, puede determinarse un umbral de humedad antes del punto de marchitez permanente, o medir también la capacidad de campo, como se muestra en la Fig. 1. Un vez determinados los parámetros es necesario seleccionar los sensores más adecuados para su monitorización, y en caso de no existir, diseñar una interfaz de hardware que pueda obtener los parámetros previamente determinados. Al módulo que recaba la información de los sensores se le denomina Unidad de Adquisición de Datos. La información recabada por los la Unidad de Adquisición de Datos debe ser pre-procesada o *normalizada*. Estas operaciones se llevan a cabo en la Unidad de Procesamiento de Datos, que es un sistema embebido con capacidades de comunicación inalámbrica. Los datos pre-procesados son enviados a un dispositivo con mayor capacidad de cómputo, en el que se ejecuta un proceso denominado Servidor, que recibirá las mediciones obtenidas por un conjunto de sensores a través de las Unidades de Procesamiento de Datos. El proceso servidor, que recibe las mediciones de los sensores, las almacena localmente y en la Nube, procesa los datos y lleva a cabo un análisis de los mismos. Con esta información decide las acciones a seguir relacionadas con el control del riego. Por ejemplo, notificar al usuario cuando un sector del cultivo tiene exceso o deficiencia de agua, o recomendar un programa de riego. Adicionalmente, se propone el desarrollo de una aplicación móvil que permita al usuario consultar, en tiempo real, los niveles de humedad de sus cultivos, o recibir alertas de situaciones anómalas. Finalmente, se propone que sea un sistema de ciclo cerrado, ya que se considera que el usuario sea capaz de activar el sistema de riego desde la aplicación móvil, con base a las recomendaciones del sistema. De igual manera, que pueda definir o ajustar el programa de riego con

el apoyo de técnicas de aprendizaje autónomo, utilizando los datos obtenidos por el sistema.

5. Prototipo

Para evaluar la arquitectura del sistema propuesto, se desarrolló el prototipo propuesto. Sus componentes se explican a continuación:

5.1. Unidad de Adquisición de Datos

Unidad de Adquisición de Datos se representa en la Figura 4. Para monitorear los niveles de humedad en el suelo se seleccionó un sensor modelo Watermark [?]. Este sensor, patentado por la compañía Irrometer, Inc., es un dispositivo de resistencia eléctrica de estado sólido que se utiliza para medir la tensión del agua del suelo. A medida que la tensión cambia con el contenido de agua, la resistencia también cambia, correlacionando dicha resistencia a centibares (kilopascales) de la tensión del agua del suelo. Está diseñado para ser un sensor permanente, colocado en el suelo a monitorear y «leído» las veces que sea necesario con un dispositivo portátil o estacionario.

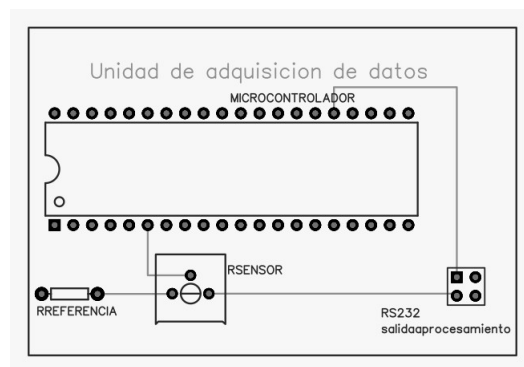


Fig. 4. Unidad de adquisición de datos.

Este sensor permite la conversión directa de la resistencia del sensor en ohms y la temperatura en valores potenciales para definir el agua del suelo. El fabricante proporcionó una tabla con los valores de resistencia y centibares. A partir de la cual se utilizó un ajuste de curvas para acceder a las siguientes ecuaciones:

Para $R \leq 1Kohm$

$$P = -20 * [R * (1 + 0,018 * (T - 24)) - 0,55].$$

Para $1Kohm < R \leq 8Kohm$

$$P = (-3,213 * R - 4,093) / (1 - 0,009733 * R - 8,01205 * T).$$

Para $R > 8K\Omega$

$$P = -2,246 - 5,239 * R * (1 + 0,018 * (T - 24)) - 0,06756 * R^2 * (1 + 0,018 * (T - 24))^2.$$

Aquí P es el potencial hídrico del suelo, R es la resistencia del sensor, T es la temperatura del suelo.

Este sensor se encuentra conectado a un microcontrolador serie PIC16F1779 [?] que cuenta con amplificadores operacionales y un convertidor analógico-digital para acondicionar la señal. Este microcontrolador es el que se encarga de tomar las mediciones del sensor y mandarlas a la unidad de procesamiento de datos vía RS-232.

```

Abrir puerto RS-232
Configurar puerto RS-232
    Baudrate 9600
    Stopbits 1
    Paridad Ninguna
Recibir datos de la unidad de adquisición de datos
    Crear búffer temporal
    Leer puerto RS-232
    Almacenar datos recibidos en búffer temporal
Pre-procesar datos
    JSON {«Fecha»: «AAAA-MM-DD», «Hora»: »HH-MM-SS»,»Humedad»: int}
Crear conexión con el servidor
    Inicializar módulo xbee
    Crear nueva conexión
    Enviar datos
    Cerrar conexión
    
```

Algorithm 1: Unidad de procesamiento de datos.

5.2. Unidad de Procesamiento de Datos

Con la información recibida por el módulo de adquisición de datos, éste módulo realiza una normalización y un pre-procesamiento de los datos. Para el envío de la información se utilizó JSON, que es un formato ligero para el intercambio de datos. Para facilitar su uso, se utilizó la biblioteca de *json-c* [?]. Después del pre-procesamiento de los datos se envían al proceso servidor utilizando los protocolos de comunicación para WPAN denominados ZigBee e IEEE 802.15.4.

En este caso, se hizo uso de la biblioteca *libxbee* [?], la cuál es ligera y especialmente creada para la intercomunicación con dispositivos *xbee*. Ésta biblioteca hace interfaz con un módulo *xbee* modelo S2C de la compañía Digi International, Inc. [?]. La biblioteca implementa el protocolo Zigbee [?], especialmente diseñado

```

Crear conexión con la unidad de procesamiento de datos
    Inicializar módulo xbee
    Crear conexión
    Crear búffer temporal
    Almacenar datos recibidos en búffer temporal
Almacenar datos recibidos en la base de datos
    Crear conexión
    Seleccionar base de datos
    Insertar datos
    Cerrar conexión
Post-procesar datos de acuerdo a las normativas de Carriots
    JSON: { «protocol»: »v2», »at»: »now», »device»:
        »midispositivo@usuario.usuario», »data»: { «Fecha»: «AAAA-MM-DD»,
        «Hora»: »HH-MM-SS»,»Humedad»: int } }
Enviar datos al servidor MQTT de Carriots
    Crear nueva instancia MQTT
    Crear conexión
    Enviar datos
    Cerrar conexión
    
```

Algorithm 2: Servidor

para redes de control y sensores que utiliza en las capas inferiores al estándar IEEE 802.15.4 [?]. Una de las ventajas de ZigBee es que los dispositivos pueden comunicarse entre sí o actuar como repetidores si el servidor esta fuera del alcance del transmisor del nodo. En el Algoritmo 1 se muestra el pseudo-código utilizado en la Unidad de Procesamiento de Datos.

5.3. Servidor

Es un proceso desarrollado en lenguaje C que se ejecuta en una computadora con mayores recursos. En el prototipo se utilizó el sistema operativo Linux (Ubuntu v18.04 [?]), el cuál recibe la información pre-procesada y concluye su procesamiento usando igualmente la biblioteca *json-c*. La información recibida se almacena en una base de datos para que esté disponible para consulta utilizando el sistema administrador de bases de datos MariaDB [?], debido a que es rápido, escalable y robusto, compatible con MySQL, con un amplio ecosistema de motores de almacenamiento. Además de ser código abierto, tiene el soporte de otras herramientas que lo hacen muy versátil. También, con los datos ya debidamente normalizados, procesados y estructurados, se continua con el proceso de almacenamiento en la nube utilizando los servicios de un proveedor de computación en la nube.

En este caso, se utilizó Carriots [?]. El envío de información al proveedor de Computación en la Nube se lleva a cabo por medio del uso de MQTT [?], que es un protocolo de comunicación de tipo *publisher-subscriber*, extremadamente ligero y práctico, diseñado para dispositivos de pocos recursos o redes de bajo

ancho de banda. MQTT es un protocolo diseñado para ser utilizado en aplicaciones IoT que se integran con la Computación en la Nube. En el Algoritmo 2 puede observarse el pseudo-código del proceso servidor.

5.4. Nube

La nube se presenta en la Figura 5. Carriots es una plataforma inteligente de arquitectura abierta, orientada a aplicaciones IoT. Carriots permite recopilar y almacenar datos de dispositivos conectados, crear aplicaciones potentes, implementar y escalar prototipos a miles de dispositivos. Además, permite el análisis de los datos y envío de notificaciones vía correo electrónico o mensaje de texto (ver Fig. 6).. Ésta plataforma si cuenta un costo mensual dependiendo de los dispositivos y los requerimientos del usuario. Para fines de pruebas del prototipo se utilizó la versión gratuita la cuál permite acceder a la plataforma y gran parte de sus herramientas hasta cierto límite.

5.5. Notificaciones

Una de las facilidades que ofrece la plataforma de Carriots, es la de enviar mensajes de texto o correo electrónico para realizar notificaciones con base a los parámetros establecidos. Éstos por supuesto por ser la versión gratuita tienen un límite el cuál fue suficiente para poder realizar las pruebas con el prototipo diseñado.

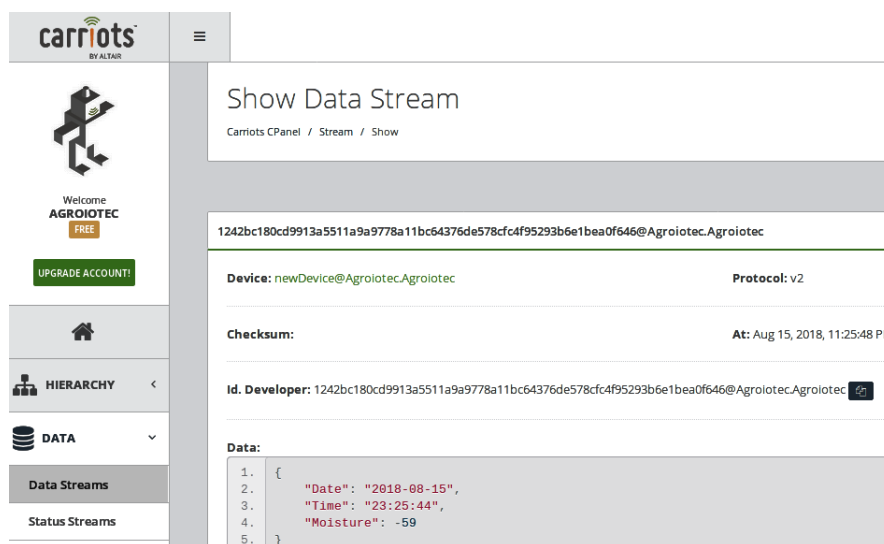


Fig. 5. Nube de Carriots.

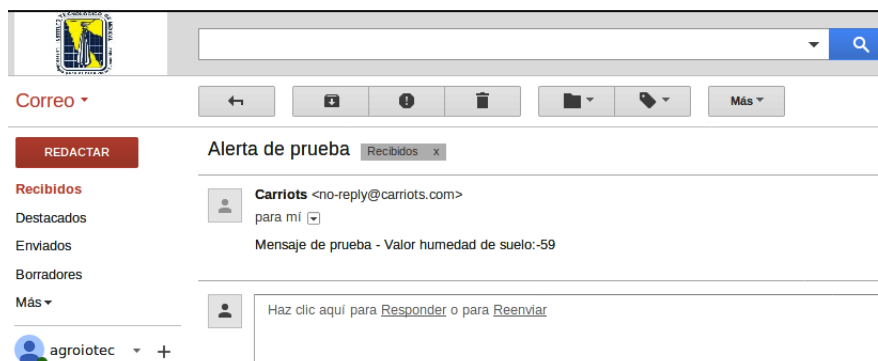


Fig. 6. Notificación de correo electrónico por Carriots.

5.6. Resultados

Al evaluar el desempeño prototipo se obtuvieron resultados satisfactorios, pudiendo realizar el sensado del nivel de humedad del suelo, hasta el envío de notificaciones de prueba vía correo electrónico sobre los estados de la humedad del suelo.

6. Conclusiones y trabajo futuro

En este artículo se presentó un sistema para el control de riego en la agroindustria, con base en el paradigma del Internet de la Cosas. El sistema propuesto fue capaz de monitorizar en tiempo real los niveles de humedad en los cultivos y de emitir alertas en caso de detectar condiciones anómalas. La evaluación de la propuesta sugiere que el uso de IoT en la agroindustria permite hacer un uso racional del agua.

Como trabajo futuro previsto se planea investigar y analizar otros proveedores de servicios computación en la nube para conocer ventajas y desventajas de cada uno de ellos. Adicionalmente, evaluar sistemas operativos para sistemas embebidos e implementar el mas conveniente en el sistema propuesto.

También se tiene previsto integrar técnicas de aprendizaje automático y utilizar la información recabada para hacer una retroalimentación y poder planear mejor los próximos riego de manera inteligente.

Referencias

1. Al-Karadsheh, E., Sourell, H., Krause, R.: Precision irrigation: New strategy irrigation water management. In: Proceedingd of the Conference on International Agricultural Research for Development, Deutscher Tropentag, Wiltzenhausen, Germany. pp. 9–11 (2002)

2. Ashraf, M., Safdar, M.E., Shahzad, S.M., Aziz, A., Piracaha, M.A., Suleman, M., Ahmad, M.B.: Challenges and opportunities for using wastewater in agriculture: a review. *Journal of Applied Agriculture and Biotechnology* 2(2), 1–20 (2018)
3. Atzori, L., Iera, A., Morabito, G.: The internet of things: A survey. *Computer Networks* 54(15), 2787–2805 (Oct 2010)
4. Bronson, K., Booker, J., Bordovsky, J., Keeling, J., Wheeler, T., Boman, R., Parajulee, M., Segarra, E., Nichols, R.: Site-specific irrigation and nitrogen management for cotton production in the southern high plains. *Agronomy Journal* 98(1), 212–219 (2006)
5. json c: github.com/json-c/json-c, <https://github.com/json-c/json-c>
6. Camp, C.R., Sadler, E.J., Evans, R.G.: Precision water management: current realities, possibilities, and trends. *Handbook of Precision Agriculture*, Haworth Press, Inc., Binghamton, NY pp. 153–184 (2006)
7. Carriots: www.carriots.com, <https://www.carriots.com/>
8. Chavez, J.L., Pierce, F.J., Matthews, G.R., Elliott, T.V., Evans, R.G., Kim, Y.: Performance of a continuous move irrigation control and monitoring system. In: 2006 ASAE Annual Meeting. p. 1. American Society of Agricultural and Biological Engineers (2006)
9. Da Xu, L., He, W., Li, S.: Internet of things in industries: A survey. *IEEE Transactions on industrial informatics* 10(4), 2233–2243 (2014)
10. Digi: www.digi.com, <https://www.digi.com/>
11. Enciso, J.M., Porter, D., Peries, X.: Uso de sensores de humedad del suelo para eficientizar el riego. University of Texas System, Austin, TX (2007)
12. Evans, R.G., Han, S., Kroeger, M., Schneider, S.M.: Precision center pivot irrigation for efficient use of water and nitrogen. *Precision agriculture (precisionagricu3)*, 75–84 (1996)
13. Ghodake, R., Jogade, S., Misal, D.: Bluetooth technology: An overview (2018)
14. Haverkort, A., Vos, J., Booij, R.: Precision management of nitrogen and water in potato production through monitoring and modelling. In: XXVI International Horticultural Congress: Potatoes, Healthy Food for Humanity: International Developments in Breeding, 619. pp. 213–224 (2002)
15. Hedley, C.B., Yule, I.J.: Soil water status mapping and two variable-rate irrigation scenarios. *Precision Agriculture* 10(4), 342–355 (2009)
16. Hedley, C., Yule, I.: A method for spatial prediction of daily soil water status for precise irrigation scheduling. *Agricultural Water Management* 96(12), 1737–1745 (2009)
17. IEEE: www.ieee.org, <https://www.ieee.org/>
18. Irrrometer: www.irrometer.com, <http://www.irrometer.com/>
19. Kelly, R., Strong, W., Jensen, T., Butler, D.: Application of probability analysis to assess nitrogen supply to grain crops in northern australia. *Precision Agriculture* 5(2), 95–110 (2004)
20. Kim, Y., Evans, R.: Software design for wireless sensor-based site-specific irrigation. *Computers and Electronics in Agriculture* 66(2), 159–165 (2009)
21. Kim, Y., Evans, R.G., Iversen, W.M.: Remote sensing and control of an irrigation system using a distributed wireless sensor network. *IEEE transactions on instrumentation and measurement* 57(7), 1379–1387 (2008)
22. King, B., Wall, R., Kincaid, D., Westermann, D.: Field testing of a variable rate sprinkler and control system for site-specific water and nutrient application. *Applied engineering in agriculture* 21(5), 847–853 (2005)
23. libxbee: github.com/attie/libxbee3, <https://github.com/attie/libxbee3>

24. MariaDB: mariadb.org, <https://mariadb.org/>
25. Mehta, A.: Bluetooth technology: an exploratory study of the analysis and implementation frameworks. *Riset Journal of Science & Engineering* 1(1) (2018)
26. Microchip: www.microchip.com, <https://www.microchip.com/>
27. MQTT: mqtt.org, mqtt.org
28. O'Shaughnessy, S.A., Evett, S.R., Colaizzi, P.D., Howell, T.A.: Soil water measurement and thermal indices for center pivot irrigation scheduling. In: *Proceedings of the 2008 Irrigation Show and Conference*. pp. 2–4 (2008)
29. Peters, R.T., Evett, S.R.: Complete center pivot automation using the temperature-time threshold method of irrigation scheduling. In: *2004 ASAE Annual Meeting*. p. 1. American Society of Agricultural and Biological Engineers (2004)
30. Peters, R.T., Evett, S.R.: Automation of a center pivot using the temperature-time-threshold method of irrigation scheduling. *Journal of irrigation and drainage engineering* 134(3), 286–291 (2008)
31. Peters, R., Evett, S.: Spatial and temporal analysis of crop conditions using multiple canopy temperature maps created with center-pivot-mounted infrared thermometers. *Transactions of the ASABE* 50(3), 919–927 (2007)
32. Sadler, E., Evans, R., Stone, K., Camp, C.: Opportunities for conservation with precision irrigation. *Journal of soil and water conservation* 60(6), 371–378 (2005)
33. Smith, R., Baillie, J., McCarthy, A., Raine, S., Baillie, C.: *Review of precision irrigation technologies and their application* (2010)
34. Talavera, J.M., Tobon, L.E., Gomez, J.A., Culman, M.A., Aranda, J.M., Parra, D.T., Quiroz, L.A., Hoyos, A., Garreta, L.E.: Review of iot applications in agro-industrial and environmental fields. *Computers and Electronics in Agriculture* 142, 283–297 (2017)
35. Ubuntu: www.ubuntu.com, <https://www.ubuntu.com/>
36. Yule, I., Hedley, C., Bradbury, S.: Variable-rate irrigation. In: *12th Annual Symposium on Precision Agriculture Research & Application in Australasia*. Sydney (2008)
37. Zigbee: www.zigbee.org, <http://www.zigbee.org/>

Sujeción de objetos virtuales mediante un gesto manual usando Kinect y software libre

Alexis Jair Urquijo-Brito, Abelardo Rodriguez-Leon, Juan Carlos Prince-Avelino, Guillermo E. Ovanco-Chacon, Carlos J. Genis-Triana

¹ Tecnológico Nacional de México/I. T. Veracruz, Veracruz, México

arleon@itver.edu.mx

Resumen. El creciente desarrollo de software en el campo de visión por computadora ha sido posible gracias al uso de nuevas herramientas que facilitan la manipulación y procesamiento de imágenes digitales. Un ejemplo de estas herramientas son librerías como OpenCV y los lenguajes como Processing. En este trabajo se describe el uso de estas herramientas de software libre junto con el dispositivo Microsoft Kinect para detectar el gesto de apertura y cierre de manos. Para lograr la identificación de estos gestos se diseñó, implementó y probó un algoritmo que consta de cinco etapas: adquisición de la imagen, segmentación, detección de contorno, búsqueda de puntos estratégicos y reconocimiento. Se comprueba el correcto funcionamiento del algoritmo en un entorno virtual donde se traslada un rectángulo de un lugar a otro de la pantalla, mediante el gesto de sujetar y soltar de la imagen virtual.

Palabras clave: Kinect, OpenCV, Processing, software libre.

Securing Virtual Objects through a Manual Gesture using Kinect and Free Software

Abstract. The growing development of software in the field of computer vision has been possible thanks to the use of new tools that ease the manipulation and processing of images. -an example of these tools are the libraries like OpenCV and languages like Pocessing. In this work describes the use of these free software tools along with the Microsoft Kinect device for the detection of the gesture of opening and closing of hands. To achieve the identification of these gestures an algorithm was designed, implemented and tested, this algorithm is constituted of 5 phases: image acquisition, segmentation, contour detection, search of strategic point and recognition. The correct functioning of the algorithm is verified in a virtual environment where a rectangle is moved from one place to another on the screen, by means of the gesture of holding and releasing the virtual image.

Keywords: Kinect, OpenCV, Processing, free software.

1. Introducción

El reconocimiento de gestos se ha convertido en un área muy estudiada en la actualidad debido al creciente desarrollo de sus campos de aplicación, que van desde Internet de las cosas, pasando por la salud, robótica, realidad virtual, juegos de computadora, entre otras [1]. Ya que en ocasiones una cámara de color no proporciona información suficiente para analizar una escena, muchas aplicaciones requieren el uso de al menos dos sensores, uno para capturar la información de color y otro para la profundidad (Depth). Este tipo de dispositivos es conocido como cámaras TOF (Time Of Flight), un ejemplo de estos dispositivos es el Kinect de Microsoft [2].

Actualmente existen dos versiones en el mercado para este dispositivo, en este trabajo se utiliza la versión uno lanzada por primera vez en el año 2011. En la búsqueda de la manipulación de objetos virtuales este documento se enfoca en resolver el problema de detectar la apertura y cierre de una mano en cualquier entorno, sin importar el ambiente donde se ejecute, tal gesto es de mucha utilidad para sujetar objetos, arrastrarlos y soltarlos en un lugar determinado. La contribución principal de este documento consiste en desarrollar un algoritmo para reconocer el gesto de apertura y cierre de una mano utilizando algunas fórmulas matemáticas. El propósito de todo lo anterior es lograr identificar un gesto de la mano que simule el tomar un objeto. Dicho gesto inicia con la identificación de la mano abierta. A partir de ese momento se hace un seguimiento de la mano para identificar cuando los dedos se cierran (quedando visualmente dentro de la palma de la mano). Esto se interpreta como el agarre de un objeto virtual (en la pantalla de la computadora). Una vez logrado el agarre virtual este se puede aplicar para la manipulación de objetos en diversos tipos de aplicaciones futuras.

2. Trabajos relacionados

Muchos enfoques se han propuesto para buscar reconocer gestos a través de cámaras TOF. Primeramente, se encuentran aquellos que solo usan la cámara de profundidad para realizar el seguimiento y la segmentación de la mano, como propone [5]. La propuesta consiste en un algoritmo que consta de cinco partes: segmentación, extracción de proyección, reducción de error, búsqueda de contorno y aplicación de clasificador basado en redes neuronales. Todo esto usando una antigua y discontinuada librería de software libre llamada OpenNI.

Por otra parte, Dominio [6] propone un algoritmo basado en el reconocimiento de un conjunto de características tridimensionales con la finalidad explotar la información 3D obteniendo la forma y posición de la mano. Específicamente busca explotar cuatro tipos de características: distancia, elevación, curvatura y geometría, para crear un modelo robusto y novedoso usando el Kinect para reconstruir la escena. Posteriormente extrae características descriptivas, dividiendo la mano en palma y dedos. Finalmente propone un clasificador de gestos.

Rem [1] propone el uso de una pulsera en la mano para limitar la forma aplicando el algoritmo de búsqueda RANSAC [1]. Posteriormente la forma de la mano se traslada como una gráfica al dominio de las X donde se busca la posición y tamaño y de cada

dedo. Finalmente propone una nueva métrica de distancia conocida como FEMD, para identificar un gesto de la mano.

Otros modelos como el de Stergiopoulou [7] buscan utilizar redes neuronales para la segmentación y filtrado de color de piel, identificando de manera exitosa las regiones de la mano. Otro enfoque [8] muy utilizado es el uso de mecanismos de detección de articulaciones usando herramientas de software privativo como es el caso de EasyGR que reconoce gestos desde los movimientos del esqueleto usando dos modos de ejecución. El primero modo entrenamiento, donde se almacenan los movimientos en un buffer, este movimiento representa un modelo de gesto que podrá ser reconocido posteriormente.

3. Desarrollo

Para el desarrollo del proyecto inicialmente se preparó el entorno de trabajo necesario. Se utilizó Processign [9] como un lenguaje de programación y entorno de desarrollo ya que es software libre y una excelente herramienta para crear prototipos y herramientas profesionales¹. Se usó también libfreenect [10] como driver y librería para la programación del Kinect. También se usó OpenCV [11] como librería auxiliar en la manipulación de imágenes. Todas estas librerías están bajo licencia GPL.

Posteriormente se vio la necesidad de desarrollar un algoritmo para reconocer el gesto de apertura y cierre de una mano. Se utilizaron fórmulas matemáticas para conseguir una aproximación al centro de la mano. A partir del centro se utiliza la fórmula de la distancia entre dos puntos como métrica de referencia, para saber si todos los dedos se encuentran ya dentro del área de la palma. Este reconocimiento del gesto se hace de manera automática, sin comparaciones con modelos previamente adquiridos, como muchas de las metodologías proponen en la literatura. Estas metodologías comparan las imágenes con modelos previamente almacenados, buscando patrones similares, como es el caso de SP-EMD [2], el cual ha mostrado buenos resultados, aunque se depende en gran medida de los patrones establecidos como referencia.

Como se comentó previamente, para este trabajo se desarrolló un mapeo en dos dimensiones de las imágenes de profundidad y color. La finalidad de este mapeo es eliminar el fondo de la escena tomando ventaja de la cámara de profundidad y estableciendo un umbral válido. Lo anterior nos evita el uso de complejos algoritmos de reconstrucción 3D como los propuestos en otros trabajos de investigación tales como [4] que propone en una comparación cuantitativa de las imágenes. El mapeo de las dos imágenes en una sola, proporciona la ventaja usar solo una imagen 2D, lo que mejora el rendimiento. A diferencia de lo que menciona [5], se ha podido comprobar en este trabajo que la representación 2D sí aporta la información necesaria para rastrear la posición de la mano.

Con la finalidad de mejorar la precisión del algoritmo dentro del umbral de decisión, se usó un guante de látex color azul que permitió eliminar cualquier otro elemento en la escena, por ejemplo, el antebrazo. El reconocimiento de colores se procesa haciendo uso del modelo de color HSV (*Hue, Saturation, Value*), comparando el canal H de este

¹ www.processing.org

modelo en busca del color azul, también utilizan algoritmos para la detección de bordes como el Canny Edge Detector [14] y detección de figuras convexas.

Finalmente, se crea un entorno básico virtual de pruebas donde se busca trasladar un rectángulo de un lugar a otro usando gestos de sujeción y procesando el arrastre.

3.1. Configuración del entorno

Processing es un proyecto de código abierto que podemos descargar en <https://processing.org> y está disponible para plataformas Windows, Mac y Linux. Se basa en lenguaje Java y proporciona una colección de librerías poderosas y robustas, dentro de un entorno de desarrollo (IDE) propio [12].

Dentro del IDE de Processing se instalaron las librerías necesarias para este proyecto (libfreenect y opencv) de la siguiente manera. En el Menú de herramientas de Processing desplegamos “Sketch”, seguido de “Importar biblioteca” y finalmente Añadir biblioteca, esto abre una ventana nueva donde podemos filtrar las bibliotecas por nombre. La biblioteca para el manejo del Kinect se llama “Open Kinect for Processing” y fue escrita por Daniel Shiffman y Thomas Sánchez, antes de proceder con la instalación se recomienda cumplir con los requerimientos. Para un sistema operativo Linux se requiere libusb1.0 y se instala de la siguiente manera.

```
sudo apt-get install libusb-1.0-0 libusb-1.0-0-dev
```

Una vez localizada la biblioteca en el “Contribution Manager” de Processing e instalados los requerimientos procedemos haciendo click en el botón instalar.

De la misma manera buscamos la biblioteca de opencv, la cual se llama “OpenCV for Processing” [13] escrita por Greg Borenstein.

Al finalizar estas instalaciones se obtuvo un entorno funcional para continuar con el proyecto.

3.2. Algoritmo de procesamiento

A continuación, se describe el algoritmo para el proceso de detección de gestos. Como se puede observar en la Figura 1 consta de cinco etapas principales, donde cada una de ellas lleva implícitas actividades adicionales. Las etapas son: Adquisición de imagen, segmentación, detección de contorno, búsqueda de puntos estratégicos y reconocimiento. A continuación, se describe con más detalle cada una de ellas.

Adquisición de imágenes

Como se ha comentado anteriormente, el Kinect es un tipo de dispositivo TOF, lo que significa que obtendremos dos imágenes diferentes. La imagen de profundidad y la imagen de color o RGB. Ya que ambas cámaras están desfasadas (Fig. 2) es necesaria una calibración previa.

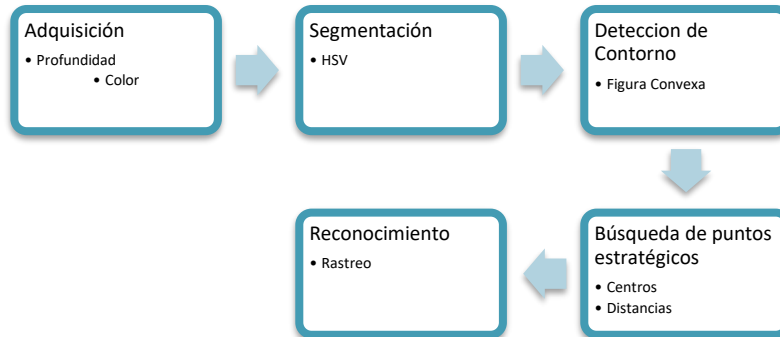


Fig. 1. Algoritmo propuesto. Cada etapa consta de actividades complementarias.



Fig. 2. Mapeo de imagen de profundidad y color sin calibración.

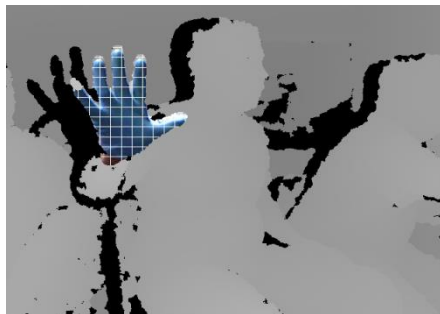


Fig. 3. Calibración y mapeo de imágenes de color y profundidad.

Para llevar a cabo la calibración y mapeo, se realizaron los pasos que se observan en la figura 4. Se realizó la calibración de la cámara, moviendo cada pixel de la imagen de color en una matriz temporal, posteriormente los pixeles se desplazan siete por ciento en el eje 'x' y diecisiete por ciento en el eje 'y'. Finalmente para el mapeo se busca un umbral (1.2 metros en este caso) de decisión, donde cada pixel de la cámara de profundidad se compara contra este valor. Cualquier pixel que contenga un valor de profundidad menor al umbral, se sustituye por los pixeles de la imagen obtenida por la cámara de color. (Fig. 3).

```

// Calibración de imágenes
mientras x < 650 //No permite exceder el límite de la imagen
    imagenCalibrada[nuevoX][nuevoY] = imagenColor

// Mapeo de imágenes
mientras i < numeroPixeles
    si (profundidad < rangoVisible) //Distancia de la mano
        nuevaImagen = imagenCalibrada
    si no
        nuevaImagne = imagenProfundidad
    
```

Fig. 4. Pseudocódigo del proceso de Calibración y mapeo

Segmentación

El siguiente paso es la segmentación de la imagen, que consiste en eliminar todo aquello que no es necesario en la escena, por ejemplo, el antebrazo y el fondo, ya que lo único que nos concierne es la forma de la mano. Para esto se utiliza un guante en color azul (Fig. 5) con la finalidad de limitar el área de la mano. Este guante nos permite enfocarnos en la búsqueda del color en la imagen obtenida. Para ello es necesario transformar la imagen del modelo de color RGB a un modelo HSB, esto permite usar el canal 'H' para hacer un filtrado en la imagen, buscado valores dentro del rango de color azul. Posteriormente a la búsqueda de color, se unifica la imagen eliminando la mayor cantidad de espacios, o huecos. Para lo cual, se usaron algoritmos de erosión y dilatación.

El resultado es una forma clara y limpia de la mano, como la que se muestra en la figura 6.



Fig. 5. Guante de látex, utilizado en el rastreo de la mano.

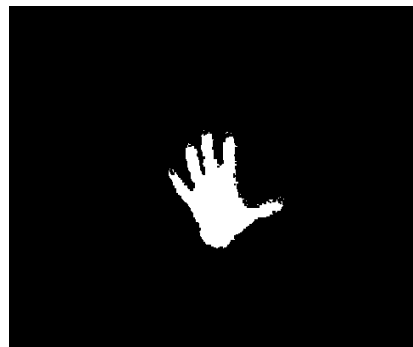


Fig. 6. Resultado de la segmentación.

Detección de contorno en la imagen

Para el análisis del contorno, se utiliza una versión corta del original OpenCV para C++, en el entorno Processing. Aunque esta no es la versión original que contiene cientos de algoritmos para la manipulación de imágenes, fue suficiente para este trabajo, ya que proporciona implementaciones del algoritmo Canny Edge Detector[14]. Un contorno es una lista de puntos que representa una curva en una imagen. Los contornos se pueden manejar de diferentes maneras. OpenCV para Processing los encapsula mediante la clase Contour.

Después de hallar el contorno de la forma (Fig. 7) se reducen el número de puntos, mediante el algoritmo Convex Hull [14]. Lo anterior proporciona un nuevo polígono que representa el contorno de la mano. Convex Hull eliminar las imperfecciones del mapeo de las imágenes, devolviendo una serie de vértices que representan los extremos de todos dedos de la mano y el borde exterior de la palma (Fig. 8).

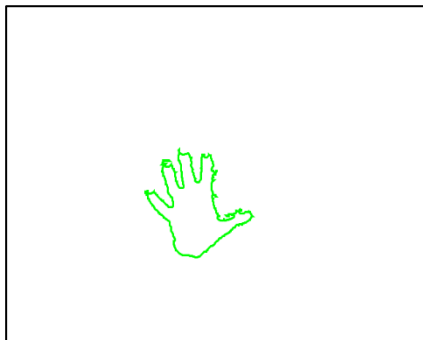


Fig. 7. Contorno de la mano.

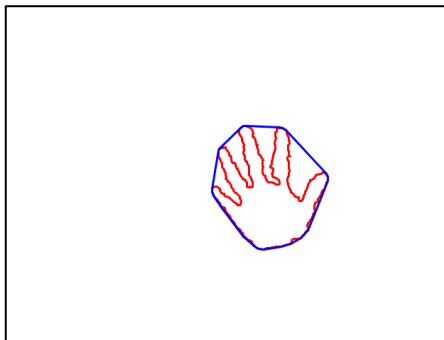


Fig. 8. Aplicación de Convex Hull en color azul.

Búsqueda de puntos estratégicos

A partir de los vértices que forman el contorno de la mano fue posible encontrar el centro de la misma (Fig. 9) con la siguiente formula:

$$C = \left(\frac{\sum_{i=1}^{len} P[i].x}{len}, \frac{\sum_{i=1}^{len} P[i].y}{len} \right),$$

donde:

‘C’ presenta el centro de la mano,

‘len’ representa el número total de puntos que forman el contorno de la mano,

‘P[i].x’ es las coordenadas ‘x’ de todos los puntos del contorno de la mano,

‘P[i].y’ es las coordenadas ‘y’ de todos los puntos del contorno de la mano.

A partir del centro, se logró delimitar ya la palma de la mano, dibujando un círculo de radio proporcional, a la figura de la mano (Fig. 10), en este caso 120 pixeles.



Fig. 9. Centro de la mano.

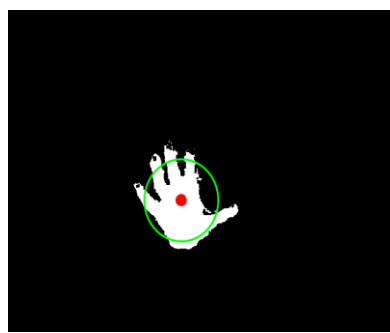


Fig. 10. Aproximación al centro de la mano.

Como se mencionó en el punto anterior (detección de contorno de la imagen), la idea era que cada vértice representara un dedo. Esta suposición resultó no ser cierta ya que

en algunas ocasiones aparecían más de un vértice por cada dedo. Estos errores, se deben principalmente a la reducción de la densidad en la imagen, después de la segmentación por color. Para solucionar este problema se aplicó un algoritmo para la reducción de vértices.

El algoritmo para la reducción de vértices, se basó en medir distancias entre todos los vértices e ir comparando sus valores. Las distancias se miden por la fórmula de distancia entre dos puntos:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Este algoritmo considera dos distancias importantes, la distancia menor y la distancia mayor (Fig. 11). La distancia menor se refiere a los puntos en un mismo dedo y la distancia mayor se refiere a la distancia entre diferentes dedos. Cada distancia menor no debe ser de más de 25 píxeles y cada distancia mayor, debe ser superior a 135 píxeles.

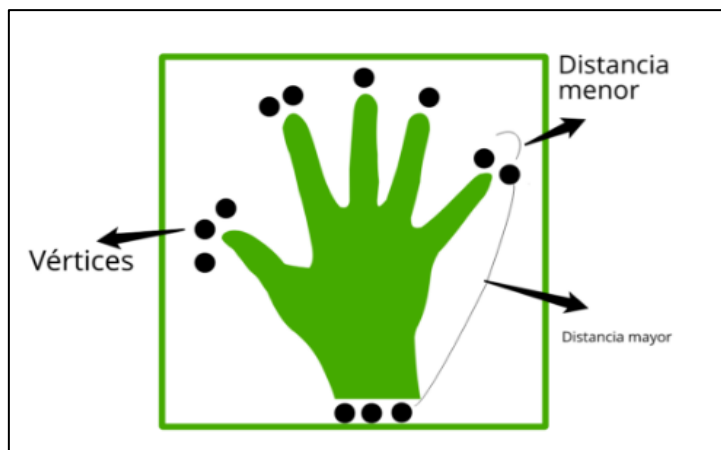


Fig. 11. Algoritmo de reducción de vértices.

Reconocimiento de gestos

En este punto la detección del gesto se vuelve un proceso más simple. Si se tienen identificados los dedos y la palma de la mano, el reconocimiento del gesto se reduce a determinar cuáles dedos se encuentran dentro del área de la palma de la mano, de tal manera que cuando todos los dedos se encuentran dentro hablamos de una mano cerrada.

4. Resultados

Con la finalidad de probar la detección de los gestos, se construyó un entorno virtual, donde se puede apreciar un rectángulo en color azul (Fig. 12), la prueba consiste, en trasladar el rectángulo, de una posición inicial a una posición final, utilizando el gestos de cerrar la mano, para sujetar, y el gesto de abrir la mano, para soltar el objeto.

El experimento anterior se realizó en quince ocasiones, de las cuales, en trece el algoritmo detectó correctamente cuando la mano se cerraba sobre una figura (sujeción). Posteriormente el algoritmo rastreó el recorrido completo de la mano de la mano en las quince ocasiones (incluso en los casos donde no detectó sujeción). Finalmente, el algoritmo detectó la apertura de la mano en las trece ocasiones que si identificó sujeción. Como se muestra en la Tabla 1, el campo “positivos”, indica el número de aciertos del algoritmo, el campo “negativos”, indica el número de errores.

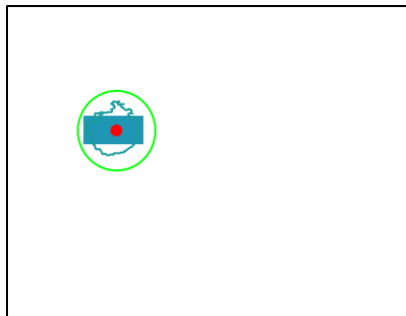


Fig. 12. Posición inicial del rectángulo.

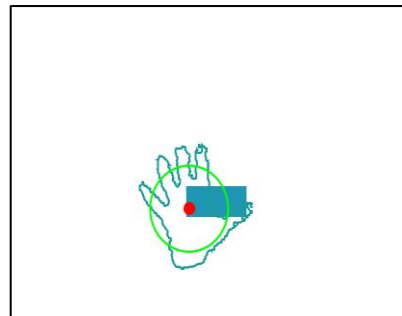


Fig. 13. Posición final del rectángulo.

Tabla 1. Tabla de resultados.

Acción	Positivos	Negativos
Cerrar la mano (sujeción)	13	2
Seguimiento de la mano	15	0
Abrir la mano	13	2

5. Conclusiones

En este trabajo se diseñó e implementó con éxito un algoritmo de cinco etapas que permite manipular las imágenes de color y profundidad proporcionadas por una cámara TOF al momento de escanear una mano. La finalidad de identificar los gestos de cierre y apertura de la misma se interpretan como sujetar y soltar un objeto. El algoritmo presentado en este trabajo se usará en un futuro en aplicaciones para manipulación de objetos 3D en ambientes virtuales. Específicamente para control de cámara en un simulador 3D desarrollado en el LCI [15] que muestra el comportamiento de un flujo de partículas en un muro de visualización.

Para probar el funcionamiento del algoritmo se usó una pequeña aplicación gráfica donde se debía poder tomar, mover y soltar un pequeño rectángulo mostrado en una pantalla. Los resultados fueron buenos siempre y cuando la mano se encuentre dentro del umbral predefinido en el eje Z. En los casos que el algoritmo no detectó correctamente la sujeción del objeto se debieron a que en esta versión no se consideró la velocidad de cerrado de la mano. Dicha situación se corregirá en un futuro cercano.

Referencias

1. Ren, Z., Yuan, J., Meng, J., Zhang, Z.: Robust Part-Based Hand Gesture Recognition Using Kinect Sensor. *IEEE Transactions on Multimedia*, vol. 15, no. 5, pp. 1110–1120 (2013)
2. Herrera C., D., Kannala, J., Heikkila, J.: Joint Depth and Color Camera Calibration with Distortion Correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 10, pp. 2058–2064 (2012)
3. Wang, C., Liu, Z., Chan, S.: Superpixel-Based Hand Gesture Recognition with Kinect Depth Camera. *IEEE Transactions on Multimedia*, vol. 17, no. 1, pp. 29–39 (2015)
4. Villena-Martínez, V., Fuster-Guilló, A., Azorín-López, J., Saval-Calvo, M., Mora-Pascual, J., García-Rodríguez, J., García-García, A.: A Quantitative Comparison of Calibration Methods for RGB-D Sensors Using Different Technologies. *Sensors*, vol. 17, no. 2, p. 243 (2017)
5. Asad, M., Abhayaratne, C.: Kinect depth stream pre-processing for hand gesture recognition. *IEEE International Conference on Image Processing* (2013)
6. Dominio, F., Donadeo, M., Zanuttigh, P.: Combining multiple depth-based descriptors for hand gesture recognition. *Pattern Recognition Letters*, vol. 50, pp. 101–111 (2014)
7. Stergiopoulou, E., Papamarkos, N.: Hand gesture recognition using a neural network shape fitting technique. *Engineering Applications of Artificial Intelligence*, vol. 22, no. 8, pp. 1141–1158 (2009)
8. Ibañez, R., Soria, A., Teyseyre, A., Campo, M.: Easy gesture recognition for Kinect. *Advances in Engineering Software*, vol. 76, pp. 171–180 (2014)
9. Processing.org [Online]. Available: <https://processing.org>. [Accessed: 07- May- 2018].
10. OpenKinect: Openkinect.org. [Online]. Available: https://openkinect.org/wiki/Main_Page. [Accessed: 09- May- 2018].
11. OpenCV library: Opencv.org. [Online]. Available: <https://opencv.org>. [Accessed: 06- May- 2018].
12. White, S: Processing in Eclipse. Processing.org. [Online]. Available: <https://processing.org/tutorials/eclipse/>. [Accessed: 09- Jul- 2018].
13. Borenstein, G.: atduskgreg/opencv-processing, GitHub, [Online]. Available: <https://github.com/atduskgreg/opencv-processing>. [Accessed: 07- Jun- 2018].
14. Kaehler, A., Bradski, G.: *Learning OpenCV 3*. Sebastopol, CA: O'Reilly Media (2017)
15. Lázaro Velasco, R., Rodríguez-Leon, A., Prince-Avelino, J. C., Ovando-Chacón, G. E., Rodríguez-Gonzalez, A.: Migración de un simulador de partículas 3D al renderizador paralelo Equalizer para su despliegue en un muro de visualización. *Revista Coloquio Internacional Multidisciplinario*, Vol. 5. no. 1, pp. 778–786 (2017)

Diseño de una aplicación móvil para la solución de ecuaciones cuadráticas en una variable

Oscar Enrique Callejas Melgoza¹, María Guadalupe Amado Moreno¹,
Kevin Alberto Luna González², Raul Lara Granillo²,
Manuel Ernesto Robles Hernández²

¹ Tecnológico Nacional de México/I.T. Mexicali, Departamento de Ciencias Básicas. Av. Tecnológico, s/n col. Elías Calles, Mexicali B. C. México

² Tecnológico Nacional de México/I.T. Mexicali, Departamento de Sistemas y Computación. Av. Tecnológico, s/n col. Elías Calles, Mexicali B. C. México
oscar.callejas@itmexicali.edu.mx

Resumen. De cada cuatro usuarios tres tienen un teléfono inteligente, es un dispositivo al que tienen acceso en la actualidad la mayoría de los estudiantes. Se diseñó la aplicación móvil “Ecuatica” para resolver ecuaciones cuadráticas de una variable. El usuario ingresa la información a la app a través de la cámara del dispositivo móvil, con tecnología táctil para dibujo o de forma manual con la calculadora del dispositivo. Se recomienda la utilización de la aplicación móvil “Ecuatica” en las clases de matemáticas, ya que disminuye el tiempo de cálculo y permite al estudiante el análisis e interpretación de la solución. Se espera implementar “Ecuatica” como herramienta didáctica en las asignaturas de matemáticas en las que se requiera resolver ecuaciones cuadráticas en una variable e innovarla para que resuelva ejercicios para otros temas de matemáticas.

Palabras clave: aplicación móvil, teléfonos inteligentes, ecuaciones cuadráticas.

Design of a Mobile Application for the Solution of Quadratic Equations of One Variable

Abstract. Of every four users, three have a smartphone, it is a device that most students have access today. The mobile application "Equatica" was designed to solve quadratic equations of one variable. The user enters the information to the app through the camera of the mobile device with touch technology for drawing or manually with the calculator of the device. The use of the mobile application “Equatica” in math classes is recommended, since it reduces the calculation time and allows the student to analyze and interpret the solution. It is expected to implement "Equatica" as a didactic tool in the mathematics courses in which it is required to solve quadratic equations of one variable and innovate it to solve exercises for other mathematics courses.

Keywords: mobile application, smartphones, quadratic equations.

1. Introducción

Desde hace algunos años se ha observado un aumento en el uso de los dispositivos móviles “Las aplicaciones móviles o apps son piezas de software diseñadas para ser instaladas y utilizadas en dispositivos móviles” [1].

Actualmente en la vida cotidiana es común que las personas, principalmente los jóvenes, utilicen teléfonos inteligentes (smartphone) y tabletas, ya que los adaptan a su personalidad y necesidades, sin embargo, el smartphone es el principal medio de conexión y captura de información por ser portable y accesible en todo momento [2].

La interacción virtual entre usuarios genera otra forma de comunicación, es decir, se produce un ambiente de satisfacción personal, motivacional y de indagación [3]. El 86% del consumo de datos en un teléfono inteligente es por la utilización diaria de las aplicaciones móviles y el 75% de uso de estas aplicaciones es para comunicarse y/o acceder a algún contenido [1], de cada cuatro usuarios tres tiene un teléfono inteligente [4].

El diseño e implementación de aplicaciones móviles en la enseñanza de las Matemáticas es reciente, diversos autores lo han utilizado como un medio para promover el aprendizaje en el aula, se continúa la actualización de “Ecuatica” en sus funciones e interacción con el usuario [5], en este trabajo se presenta el uso de la aplicación para la solución de ecuaciones cuadráticas en una variable, favoreciendo el análisis e interpretación de resultados en los ejercicios de la vida cotidiana [6].

Existen identificadas en la web algunas aplicaciones móviles que ofrecen soluciones matemáticas a través de reconocimiento de imágenes (Photomath, MyScriptCalculator, Mathway). Realizando un análisis de competencia se identificaron las siguientes ventajas de “Ecuatica”:

- a) Interfaz intuitiva y práctica.
- b) Amplia variedad de soluciones.
- c) Velocidad de detección del ejercicio.
- d) Procesamiento de ejercicios matemáticos avanzados.

El objetivo de este trabajo fue actualizar el diseño de “Ecuatica” en sus funciones e interacción con el usuario para la solución de ecuaciones cuadráticas en una variable en la resolución de ejercicios de la vida cotidiana.

Las ecuaciones cuadráticas en una variable tienen múltiples campos de aplicación: al jugar fútbol, saltar la cuerda, lanzar una moneda al aire, gota de agua al caer, figuras geométricas, telefonía celular entre otras, se obtiene una ecuación cuadrática [7].

Se actualizó el diseño de la aplicación móvil “Ecuatica” para que el estudiante la implemente como una herramienta digital, en la solución de ecuaciones cuadráticas en una variable que pueda requerir para la interpretación de ejercicios de la vida cotidiana.

2. Materiales y método

Se actualizó el diseño de “Ecuatica”, una aplicación móvil en lenguaje JAVA para el sistema operativo Android, la cual utiliza tecnología de reconocimiento de imagen de teléfonos inteligentes y/o tabletas para encontrar la solución de ecuaciones cuadráticas en una variable.

Se observa en la figura 1 el diagrama de flujo del funcionamiento de “Ecuatica” el cual consta de los bloques de inicio y fin, tres de proceso y uno de decisión.

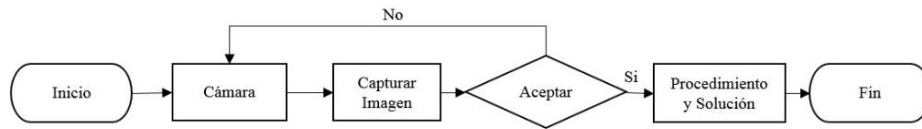


Fig. 1. Diagrama de flujo del funcionamiento de “Ecuatica”.

El algoritmo de procedimiento de una parte del bloque de proceso de “Ecuatica” se observa en la figura 2.

```

1  Procedimiento
2  Se toma una foto al modelo matemático y se envía al servidor
3  El servidor analiza la imagen
4  Se convierte adapta la expresión para el motor matemático
5  Si la expresión es válida
6      El motor matemático procesa el modelo y calcula raíces, integrales y derivadas basado en aproximación
7      Se genera una estructura de árbol con la expresión matemática para obtener el procedimiento
8      Se recorre el árbol, optimizando cada una de sus ramificaciones
9      Por cada optimización se guarda el paso realizado
10     Una vez terminado de recorrer el grafo, el servidor devuelve cada uno de los cálculos
11     La aplicación muestra resultados y procedimientos
12 Si no
13     El servidor no devuelve resultados
14     La aplicación le informa al usuario
15 Fin si
16 Fin procedimiento
    
```

Fig. 2. Algoritmo de procedimiento de “Ecuatica”.

Para utilizar la aplicación móvil se accede a la App de Google Play Store, se descarga e instala, posteriormente puede ingresarse la ecuación cuadrática por medio de la cámara, dibujo o calculadora. Puede observarse en la figura 3 como el usuario ingresa la ecuación cuadrática a través de la calculadora, en la figura 4 el ingreso es con tecnología táctil para dibujo y en la figura 5 con la cámara del teléfono inteligente o tableta.

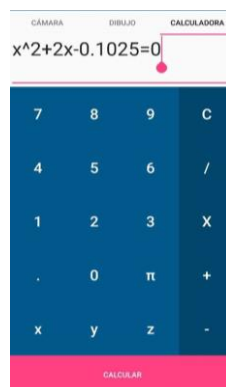


Fig. 3. Ingreso de la ecuación cuadrática con la calculadora.

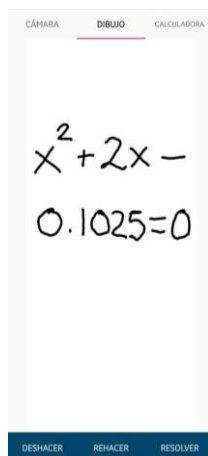


Fig. 4. Ingreso de la ecuación cuadrática con tecnología táctil para dibujo.

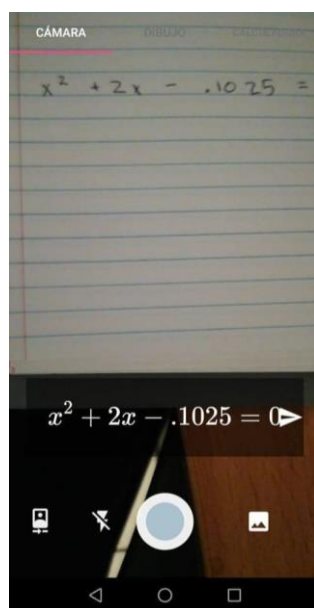


Fig. 5. Ingreso de la ecuación cuadrática con la cámara.

3. Resultados

Se diseñó la aplicación móvil “Ecuatica” para la solución de ecuaciones cuadráticas en una variable en la resolución de ejercicios de la vida cotidiana. Se propuso la aplicación para resolver ejercicios de la vida cotidiana.

Un ejemplo que puede utilizarse para el curso de Matemáticas Financieras se muestra a continuación.

Interés compuesto

La fórmula para calcular el interés compuesto: $A = p \left(1 + \frac{r}{n}\right)^{nt}$ puede utilizarse para determinar el monto A, cuando el capital inicial p, se invierte a una tasa de interés anual r, capitalizable n veces en un año durante t años.

En un principio, Josh Adams invirtió \$1000 en una cuenta de ahorros cuyo interés compuesto se paga una vez al año. Si después de dos años el monto, o saldo, en la cuenta es de \$1102.50, determine la tasa de interés anual r [8].

Analíticamente el estudiante hace lo siguiente para obtener la ecuación cuadrática:

Datos:

$$p = \$1000 \quad A = \$1102.50 \quad n = 1 \quad t = 2$$

Para determinar el interés compuesto se tiene:

$$A = p \left(1 + \frac{r}{n}\right)^{nt}$$

Entonces:

$$1102.50 = 1000(1 + r)^2,$$

$$1000(1 + 2r + r^2) = 1102.50,$$

$$1 + 2r + r^2 = \frac{1102.50}{1000},$$

$$r^2 + 2r + 1 = 1.10250,$$

$$r^2 + 2r + 1 - 1.10250 = 0,$$

$$r^2 + 2r - 0.1025 = 0.$$

La ecuación cuadrática obtenida en el ejercicio se resuelve con “Ecuatica”, el resultado se muestra en la figura 6.

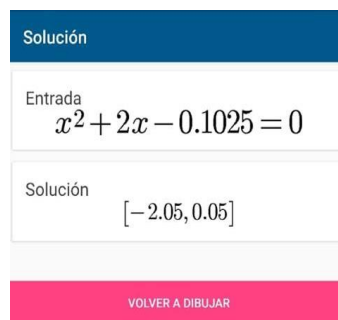


Fig. 6. Pantalla de resultados de “Ecuatica”.

El estudiante obtiene la solución de la ecuación cuadrática en una variable, analiza e interpreta de acuerdo a la pregunta del ejercicio. En este caso la respuesta es: la tasa de interés anual es 0.05 o 5%.

4. Conclusiones

Se actualizó el diseño de la aplicación móvil “Ecuatica” en sus funciones e interacción con el usuario para la solución de ecuaciones cuadráticas en una variable, que puede utilizarse en ejercicios de Matemáticas y de la vida cotidiana cuando se genere una ecuación cuadrática.

Actualmente “Ecuatica” es la única aplicación en el mercado que permite al usuario ingresar la información de tres formas diferentes: a través de reconocimiento de imagen con la cámara frontal, tecnología táctil para dibujo y/o de forma manual con la calculadora del dispositivo.

Se recomienda la utilización de la aplicación móvil “Ecuatica” en las clases de Matemáticas, ya que disminuye el tiempo de cálculo y permite al estudiante el análisis e interpretación de la solución.

El estudiante obtiene la ecuación cuadrática en una variable de un ejercicio de la vida cotidiana, utiliza “Ecuatica” para resolverlo, finalmente interpreta y analiza la solución de acuerdo al ejercicio que está resolviendo.

5. Trabajos futuros

La expectativa es implementar “Ecuatica” como herramienta didáctica en las asignaturas de matemáticas e innovarla para que muestre el procedimiento de los ejercicios matemáticos. Al implementar la aplicación móvil en el aula de clases se buscará adquirir datos que permitan establecer la efectividad de la aplicación.

La ventaja competitiva mundial de “Ecuatica” es que estará próximamente disponible la asistencia profesional matemática 24 horas siete días de la semana, empleando el tiempo universal coordinado.

Referencias

1. Aguando, J., Martínez, I., Cañete, L.: Tendencias evolutivas del contenido digital en aplicaciones móviles. *El profesional de la información* 24(6):787–795 (2015)
2. Rojas, L., Krug, L., Fernández, A., Corrales, J.: Análisis del comportamiento contextual del usuario y su relación con el consumo de aplicaciones móviles. *Revista ingenierías universidad de Medellín* 13(24):214 <https://doi.org/10.22395/rium.v13n24a11> (2014)
3. Contreras Arriaga, J., Herrera Bernal, J.A., Ramírez Montoya, M S.: Elementos instruccionales para el diseño y la producción de materiales educativos móviles. *Apertura* 1 (1) (2009)
4. INEGI: Encuesta Nacional sobre disponibilidad y uso de tecnologías de la información en los hogares 2016. <http://www.beta.inegi.org.mx/app/biblioteca/ficha.html?upc=702825091057> (2016)
5. Díaz Ramírez, A., Calafate, C. T., Quintero Rosas, V., García Vázquez, J. P.: Avances Científicos y Tecnológicos de las Aplicaciones de las Redes de Computadoras. Cap. 3: Diseño e implementación de una aplicación móvil en la solución de sistemas de ecuaciones lineales. México: Instituto Tecnológico de Mexicali, <http://iccna.itmexicali.edu.mx/bundles/app/AvancesCientificosTecnologicosRedesComputadoras.pdf#page=17> (2017)

6. Ramírez, M.: Recursos Tecnológicos para el aprendizaje móvil (mlearning) y su relación con los ambientes de educación a distancia: implementaciones e investigaciones. *Revista Iberoamericana de Educación a Distancia*, 12(2): 57–82 (2009)
7. Pérez Álvarez, P. G.: Aplicaciones de las ecuaciones cuadráticas. http://matematica.cubaeduca.cu/media/matematica.cubaeduca.cu/medias/interactividades/Temas_9no/9422aplicaecuacuadratica_web.publi/web/co/9422aplicaecuacuadratica_web.html (2016)
8. Allen, A.: *Álgebra Intermedia*. México: Pearson Educación (2008)

A Preliminary Approach for Finding Linked Determinants between Diabetes and Caries using Genetic Algorithms

Laura A. Zanella-Calzada¹, Carlos E. Galván-Tejada¹, Nubia Chavez-Lamas²,
Miguel A. Cid-Baez¹, M. del Carmen Gracia-Cortés²

¹ Universidad Autónoma de Zacatecas, Unidad Académica de Ingeniería Eléctrica,
Zacatecas, Zac, Mexico

{lzanellac, ericgalvan, drcidbaez}@uaz.edu.mx

² Universidad Autónoma de Zacatecas, Unidad Académica de Odontología,
Zacatecas, Zac, Mexico

{nubiachavez, gacc005340}@uaz.edu.mx

Abstract. Two of the public health conditions that continue having an increase in their incidence and prevalence worldwide are dental caries and diabetes mellitus. Dental caries is a multifactorial disease that affects the hard tissues of the tooth, being the most widespread disease in man; while diabetes mellitus is a multifactorial chronic disease that involves a group of disorders that alter the production and use of insulin, being of the main causes of disease in high, medium and low income countries. In this work is proposed the use of genetic algorithms in order to select the most significant features from a database of demographic data obtained from the NHANES 2013 – 2014, that best describe the difference between cases (diabetes and caries) and controls (healthy), according to their health status. After the feature selection, it is performed a validation stage, evaluating the multivariate model contained by these features. Within the results achieved shows a fitness > 0.980 and an $AUC = 0.975$, validating that the demographic features selected in this paper present a highly significant capacity to classify subjects that present caries and diabetes simultaneously from healthy subjects.

Keywords: dental caries, diabetes mellitus, health status, genetic algorithm, feature selection, multivariate model.

1 Introduction

There are currently 425 million adults with Diabetes mellitus (DM) in the world, that is, one in 11 adults, 46.5% of whom do not know they have it, and it is estimated that by 2040, this figure will increase to 642 million, with 5.1 million people dying as a direct result of Diabetes [5]. Mexico with 12.03 million people with diabetes ranks 6th worldwide [6] and the 1st. place in obesity worldwide [3]. Most blood glucose control is poor [9] coupled with low adherence to treatment and risk co-morbidities (such as cholesterol and high blood pressure) [1]. The

dramatic increase in the global prevalence of diabetes has led to an inevitable increase in biopsychosocial complications [25]. Mexico is a middle-income country, where a medium-low socioeconomic and educational status predominates (low human development index), it conditions an unfavorable health status (2 to 4 times higher, against a high status) [26]. This has a negative impact on the incidence of co-morbidities of oral pathologies, which are considered public health care, such as caries. [15].

Caries is defined as a preventable disease of chronic multifactorial origin that affects the hard tissues of the tooth, with a prevalence of 94% of individuals worldwide [27]. People with diabetes have an increased risk of gum inflammation (periodontitis) or gingival hyperplasia if their blood glucose is not properly managed. Periodontitis is a major cause of tooth loss and is associated with an increased risk of cardiovascular disease [18]. Other diabetes-related mouth conditions include dental caries 91.4%, gingivitis 97.1%, periodontitis 85.7%, oral candidiasis 17.1%, sensorineural disorders (burning mouth), salivary dysfunction and xerostomia 62.9%, and taste alteration 40%, periodontal abscesses 40% [12]. People with diabetes have a higher risk of developing periodontal disease than people without diabetes [10]. The prevalence of periodontal disease is more common among people with diabetes (92.6%) than among people without diabetes. This complication may be caused by chronic immunosuppression, delayed healing, or salivary hypofunction. The use of a toothbrush and a strict dental maintenance program is important in long-term oral health and for the prevention of complications. Regular oral check-ups provide a guarantee for early diagnosis, especially among previously undiagnosed diabetics and rapid management of any complications among patients with diabetes. On the other hand, oral health is considered to be an integral and essential part of overall health, and this can compromise quality of life. The phenomena of the binomial oral health and diabetes are highly complex, since they interact simultaneously and with variables that correspond to different orders: from biological processes, to complex historical-cultural structures and social relations.

Due to the difficulty that represents the control in the incidence of these conditions, caries and diabetes, which are caused by the large number of factors that influence, recent studies have implemented algorithms and performed analysis based on computer-aided diagnosis (CADx) in order to develop prediction and classification models for the preventive diagnosis and prevalence reduction of them looking for the main factors that affect this condition [21].

This paper is organized as follows, in section 2 is presented a description of the data used for this research and the methods that were carried out. Results and discussion obtained from the genetic algorithm process and the validation step are shown in section 3. Finally conclusions are described in section 4.

2 Materials and Methods

In this work is presented the development of a multivariate model contained by demographic features in order to classify subjects that present two different

conditions, the presence of caries and diabetes or the absence of caries and diabetes.

In this section is described the data from the National Health and Nutrition Examination Survey (NHANES) 2013 – 2014 that were used for this work, as well as the data preprocessing, feature selection and validation of the results.



Fig. 1. Flowchart of the methodology followed. A) Dataset used from the NHANES 2013 – 2014, B) Data preprocessing, C) Feature selection using a genetic algorithm, D) Validation through statistical analysis.

In Figure 1 is presented the methodology followed for the development of this research. Firstly, in (A) is shown the database used, which specifically corresponds to demographic data; then, in (B) is presented a preprocessing step where it was necessary to prepare the data in order to obtain a correct analysis of the features. In (C), the data were subjected to a feature selection through a genetic algorithm, presenting the features that shown better overall performance and individually for the classification of subjects in the two conditions mentioned before. Finally, in (D) is carried out a validation step, trough a statistical analysis, obtaining the Receiver operating characteristic (ROC) curve and the Area under the curve (AUC), with the purpose of evaluating the results obtained from the feature selection.

2.1 Data Description

The NHANES is a national program that design studies to perform a survey to assess the health and nutritional status of adults and children in the United States, including all ethnic groups. The survey combines interviews and physical examinations, allowing to develop studies using clinical, para-clinical and demographic characteristics (features) of individuals. NHANES is a program founded by the National Center for Health Statistics (NCHS), which is part of the Centers for Disease Control and Prevention (CDC) and has the responsibility for producing vital and health statistics for the United States.

Content Description NHANES survey include several types of interviews, to cover a wide range of features, including demographic, socioeconomic, dietary, and health-related questions, described in detail in 1. One examination component that is critical to this study is dental care examination, as well as the

health-related question about the condition of diabetes, which are carried on by trained medical personnel.

In Table 1 is presented the description of the main datasets used in this work, which correspond to Demographic, Examination and Questionnaire.

Table 1. NHANES Data Description

Questionnaire Type	Description
Demographic	The demographics file provides individual, family, and household level information.
Examination	Public health significance in areas of surveillance, prevention, treatment, dental care utilization, health policy, evaluation of Federal health programs.
Questionnaire	information on: Acculturation, Alcohol Use, Diabetes Status, Health Insurance, Income.

Meta Data Health and Nutrition Examination Surveys are comprised by interviews applied to 27,631 subjects. These subjects were submitted to a series of different questionnaires, related to the different features. These subjects were randomly selected with a computer algorithm by NHANES. The algorithm consists of a complex multistage probability design to choose the participants, with a series of stages.

The subjects selected were women and men that belong to different counties in the USA. These counties were divided into 15 groups according to their characteristics, and from each group one county was randomly selected, obtaining the 15 counties for the surveys. Then, a sampling of segments was performed to each county, selecting between 20 and 24 segments, where from all the houses that correspond to the segments, a sample of about 30 households were selected within each each segment. Finally, the sampling of persons was carried out where the NHANES interviewers went to the selected households and asked for the information of the surveys, randomly selecting the members of the households that were part of the final database.

The main target population for NHANES is the non-institutionalized civilian resident population of the USA. The design of the population selection is based on the sampling of a larger number of specific subgroups that present particular public health characteristics, in order to increase the reliability and precision of health status indicators. NHANES started these design changes in 2011, including in its population the oversampled subgroups survey cycle:

- Hispanic persons;
- Non-Hispanic black persons;
- Non-Hispanic Asian persons;
- Non-Hispanic white and other* persons at or below 130% of the poverty level; and
- Non-Hispanic white and other* persons aged 80 years and older.

The subjects contained in the Demographic database, which was the used for this work, were 9,801 (male = 4,826, female = 4,976), and the total of features were 39. Subjects were located in an age range from 0 to 80 years old.

2.2 Data Preprocessing

The data preprocessing consisted on prepare the data for their correct analysis. Initially, it was necessary to unify the database, keeping those subjects that presented the information of the features that made reference to the dental caries and diabetes status. Those features were mixed in one, which was the feature used as output, where subjects that presented a positive condition in both were classified as '1' and subjects that presented a negative condition in both were classified as '0'.

Then, there were removed those features that presented ≥ 70 % of missing data. The remaining data were imputed, normalized and finally, those features that presented singular values after the normalization were removed.

The imputation of the data was performed through the 'rfimpute' function, from the 'randomForest' package (version 4.6-12, 2011-10-18) [14], for R. This function imputes the data replacing all the missing data with the median value of the column, then, the 'randomForest' function is called with the complete data to update the imputation of the missing values.

On the other hand, for the normalization step was used Equation 1, where x represents the value that needs to be normalized, μ represents the mean of the column and σ represents the standard deviation of the column:

$$Z = \frac{x - \mu}{\sigma}. \quad (1)$$

Features that presented singular values were those that had the same value in the whole column.

2.3 Feature Selection

For the feature selection, the data were subjected to an implementation of a genetic algorithm called 'Galgo' [24], which is a package for R.

Galgo performs a procedure for selecting models with a high fitness value and implements a series of functions for the analysis of the population of the selected models, besides the reconstruction and characterization of representative summary models [13].

To begin, from a random population of feature subsets of a given size, which are defined as chromosomes, an assessment for their ability to predict a dependent feature is performed to each one, obtaining a certain level of accuracy. The general principle is to replace the initial population with a new population that includes some variants of chromosomes with higher classification accuracy and to repeat the process enough times to achieve a desired level of accuracy. The continuous improvement of the chromosome population is driven under the principle of the natural selection, through three main steps, selection, mutation and crossover. Besides, in order to increase the proportion of the solution space, independent populations can be evolved in isolated environments, defined as niches, and chromosomes can migrate from one niche to another ensuring that particularly good solutions can recombine.

In summary, Galgo presents features as Gene objects, which are part of a Chromosome object or a set of n features that are included in a multivariate model that is evaluated through a fitness function, using a classification method as cost function. Then, a Niche object organizes chromosomes in populations and a World object is content of several niches. These objects are later ranked in a Galgo object, implementing the general genetic algorithm evolutionary process, saving the best chromosome. Finally, a BigBang object is developed, collecting the result of several iterations or searches for the further analysis. In this process are included most common genetic algorithms methods as reproduction, mutation, crossover, migration and elitism [24].

All the objects in Galgo have properties that have the option to users to control the process, besides they can add custom defined properties in order to add new functionality.

The parameters selected for this process were a chromosome size of five features, a total number of solutions of 300 and a goal fitness of 0.75.

The classification methods that are included in Galgo are “K-nearest-neighbor”, “nearest centroid”, “support vector machines”, “neural networks”, “classification trees” and “discriminant functions”.

For this work the classification methods used was K-nearest-neighbor (K-NN), which is a nonparametric machine learning algorithm that has been widely used in pattern recognition and data mining. Its main purpose is to classify an unknown sample based on the known classification of its neighbors. This method works being fed with a training set that is contained with labeled samples that is used to classify objects. These objects are classified based on the K parameter, which means that they are assigned to a class that results to be the most similar among its closest K neighbors. In K-NN, it's firstly necessary to find the distance between the training and test data, one of the most used distance measure is the Euclidean distance, which can be calculated with Equation 2, where X is the training data with specific parameters (x_1 to x_n), and Y is the test data with specific parameters (y_1 to y_n) [8]:

$$d(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}. \quad (2)$$

Then, the distances are sorted for all the objects in the training set and the nearest neighbor is found according to the minimum distance. The K parameter is the number of the closest neighbors in the place of interest and it can be obtained by the n -fold cross-validation method.

2.4 Validation

The validation consisted on submitting the multivariate model obtained from the feature selection into a statistical analysis, in order to evaluate its performance in the classification of subjects using the set of features selected.

Initially, a logistic regression (LR) is applied to the multivariate model, obtaining the general model that best classifies the subjects. LR is a statistical

analysis technique that consists in the modeling of the relationship between a dependent variable and an independent variable or a set of them. This technique is part to the statistical methods where is measured the contribution of different factors in the occurrence of a simple event. Thus, the main objective of LR is to model the influence of the probability of an event. In Equation 3 is presented the simplest representation of a model, being y from $\text{logit}(y = 1)$ the dependent variable, which represents the outcome feature, that is subjected to a logarithmic transformation (*logit*). This transformation is performed because the initial equation of the model is exponential and through this step the model becomes a lineal function. w is an offset parameter that can be included or not, β_1 is the slope and x is the independent variable or the input feature. Models can be content by n number of independent variables [16]:

$$\text{logit}(y = 1) = w + \beta_1 x. \quad (3)$$

Then, the general model that was obtained through the LR technique was evaluated measuring its true positives - true negatives proportion, obtaining its ROC curve and the AUC value. For these measurements, the *pROC* (version 1.12.1, 2018-05-06) [20] package for R was used.

The AUC value is a standard method used for evaluating the accuracy of the model, and is calculated measuring the relationship between the specificity and the sensitivity [4].

Sensitivity is the proportion of data that belongs to a condition and it's classified as positive. This parameter is calculated with Equation 4, where TP is the quantity of true positives and FP is the quantity of false positives:

$$PPV = \frac{TP}{TP + FP}. \quad (4)$$

Specificity is the proportion of data that doesn't belong to a condition and it's classified as negative. This parameter is calculated with Equation 5, where TN is the quantity of true negatives and FN is the quantity of false negatives:

$$NPV = \frac{TN}{TN + FN}. \quad (5)$$

Finally, a comparison of the ROC curve and AUC value obtained between a model using the 39 initial features and the model of the five features selected was performed in order to do a cost benefit analysis of the results achieved.

All the data analysis were realized with the free software *R* (version 3.3.1) [19].

3 Results and Discussion

In this section the results obtained for each stage of the methodology are exposed, as well as the discussion.

Data preprocessing In the data preprocessing step, the 9,801 subjects were reduced to 3,552 (male = 1,812 , female = 1,740 / controls = 3041, cases = 511)

and the 39 demographic features were reduced to 32. The great reduction of the subjects is because most of them didn't have their complete information of their dental caries or diabetes status, besides, only the subjects that presented a positive result for both conditions were preserved, as well as the subjects that presented a negative results for both conditions. On the other hand, the reduction of the features is related to those features that presented singular values or a high number of missing data. The relationship between controls and cases may look disproportionate, however, the quantity of data for both conditions was enough to reach a significant performance in the classification of subjects.

Feature selection From the feature selection step, Galgo developed a graph of the gene rank and frequency, shown in Figure 2, with which the features appeared in the different chromosomes that were constructed by the iterative process of the genetic algorithm. In this graph, those features that were present in more chromosomes are in black while the features that were present in less are in gray, being possible to observe that the most frequent features were 'RIDEXAGM', 'RIDAGEYR', 'DMDEDUC3', 'DMQMILIZ'.

In Figure 3, it is presented a graph of the performance of fitness throughout the iterations or generations of the Galgo process, where is possible to observe that the fitness reached a significant value of around 0.99, which means that the classification of subjects using the multivariate model created is highly accurate, being possible to classify subjects in the two conditions using a reduced number of demographic features.

Finally, when Galgo finished its iterations, a forward selection and backward elimination were performed, obtaining a model contained by five features, shown in the heat map of Figure 4. This model was the result of the genes combination between the different chromosomes and the reduction of redundant or not significant information, where is shown that the strongest feature is 'RIDEXAGM', followed by 'DMDHSEDU', 'DMDEDUC3', 'DMQMILIZ' and 'RIDAGEYR'. From the final model, four of its features presented the highest frequency according to the graph of Figure 2, and even when the feature 'DMDHSEDU' didn't present a significant frequency of appearance, which means that is not presenting an important contribution for the chromosomes, it may be presenting good behavior together with the rest of the features, being the second strongest feature.

The features of the best chromosome are described in Table 2, exposing that two are referred to the age of the subjects, two are referred to the educational level and one is referred to the service of the country. This means that the age of the subjects represents important information in the presence or the absence of dental caries and diabetes simultaneously, since the age is identified as one risk factor in dental caries as well as in diabetes [11, 22]. On the other hand, the educational level and the participation in the U. S. services are features related with the socioeconomic level, which is one the determinants of both conditions too. Educational level is significant because low-income regions have fewer educational opportunities; while the participation in the U.S. services is a significant feature taking into account that the main target population

for NHANES is the non-institutionalized civilian resident population of the U. S., which may be people that decided to migrate looking for better economic opportunities and one of the requirements for the U.S. Armed Forces, Military reserves and National guard is being a citizen or a permanent resident. Taking into account that people that don't have access to a good educational level are not enough informed about the care to avoid developing these conditions, like learning to have a good oral hygiene, the knowledge that health programs can offer, the eating habits, the diet quality, among others, besides possibly not having access to a health system, it becomes difficult to be under a treatment that allows to combat these conditions [17, 2, 7, 23].

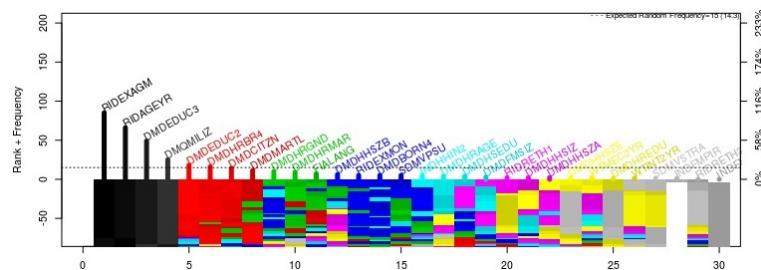


Fig. 2. Graph of the gene rank and frequency of appearance of the features in the developed chromosomes.

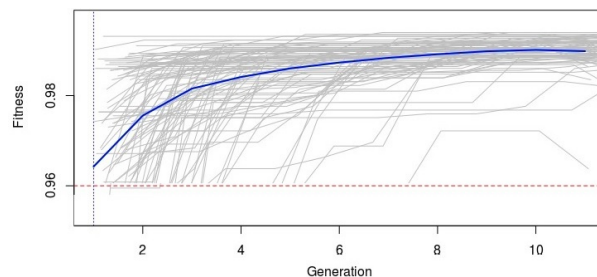


Fig. 3. Graph of the performance of fitness throughout the iterations of Galgo.

Validation Finally, in Figure 5 are shown the ROC curves that were obtained with the general model using the 39 initial features (A) and with the general model using the features selected (B), obtaining an AUC value of 0.979 and 0.975, respectively, which are statistically significant. As is possible to observe

both curves present a very similar performance, as well as the AUC value, which means that the features the are contained in the model are those that present the most significant contribution in the classification of subjects.

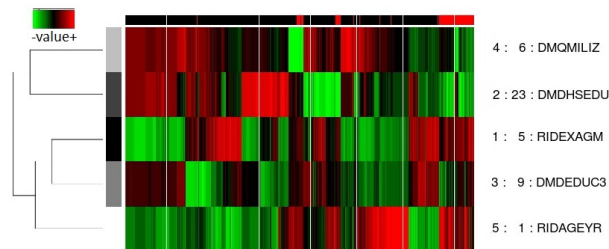


Fig. 4. Heat map of the best chromosome reached by the Galgo process.

Table 2. Best chromosome obtained from the feature selection.

Questionnaire Type	Description
RIDEXAGM	Age in months of the participant at the time of examination.
DMDHSEDU	HH reference person's spouse's education level.
DMDEDUC3	Highest grade or level of school completed or the highest degree received.
DMQMILIZ	As SP ever served on active duty in the U.S. Armed Forces, military Reserves, or National Guard?
RIDAGEYR	Age in years of the participant at the time of screening.

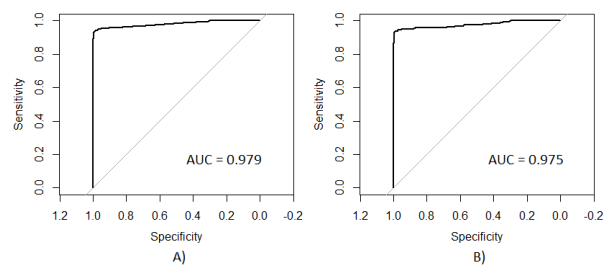


Fig. 5. ROC curves obtained with the modeling of (A) the initial 39 features and (B) the features selected.

4 Conclusion

This paper presents the analysis of a demographic features set in relationship with dental caries and diabetes, in order to develop a model that can present a tool for specialists in the preventive diagnosis of these conditions and the reduction of the incidence of them. This analysis was performed in three main stages, data preprocessing, feature selection and validation. In the feature selection stage was developed a multivariate model contained by features that provided the most significant information in the classification of control and case subjects, since they are important risk factors for these conditions according to literature; while the validation stage allowed to evaluate this multivariate model, obtaining statistically significant results.

Therefore, the model proposed in this work can represent a preliminary tool of low cost that helps in the diagnosis of dental caries and diabetes together, besides make a possible prediction of them, having as main objective the regions of low socio-economic and educational level, in order to reduce their high incidence and avoid the high cost treatments that represents these conditions.

Acknowledgments. We thank to the Convocatoria del Fondo Sectorial for the resources that allowed to finance this investigation: FOSSIS-S0008, CONACyT-SSA/IMSS/ISSSTE-2016/273142.

References

1. Alegre-Díaz, J., Herrington, W., López-Cervantes, M., Gnatiuc, L., Ramirez, R., Hill, M., Baigent, C., McCarthy, M.I., Lewington, S., Collins, R., et al.: Diabetes and cause-specific mortality in mexico city. *New England Journal of Medicine* 375(20), 1961–1971 (2016)
2. Araujo, D., Marquezín, M., Barbosa, T., Fonseca, F., Fegadolli, C., Castelo, P.: Assessment of quality of life, anxiety, socio-economic factors and caries experience in brazilian children with overweight and obesity. *International journal of dental hygiene* 15(4), e156–e162 (2017)
3. Barriguete-Meléndez, J., Hercberg, S., Galán, P., Parodi, A., Baulieux, J.: Nutrinet-salud México. prospective study online: 2018-2028. relationship between nutrition and health and determinants of dietary habits and nutritional status. *Cirugia y cirujanos* 86(1), 8–19 (2018)
4. Bourne, D.A.: *Mathematical modeling of pharmacokinetic data*. Routledge (2018)
5. Cho, N., Shaw, J., Karuranga, S., Huang, Y., da Rocha Fernandes, J., Ohlrogge, A., Malanda, B.: *Idf diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045*. *Diabetes research and clinical practice* 138, 271–281 (2018)
6. Dávila-Torres, J., González-Izquierdo, J.d.J., Barrera-Cruz, A.: Obesity in mexico. *Revista Médica del Instituto Mexicano del Seguro Social* 53(2), 240–249 (2015)
7. Elamin, A., Garemo, M., Gardner, A.: Dental caries and their association with socioeconomic characteristics, oral hygiene practices and eating habits among preschool children in abu dhabi, united arab emirates—the noplax project. *BMC oral health* 18(1), 104 (2018)

8. Fadaei-Kermani, E., Barani, G., Ghaeini-Hessaroeyeh, M.: Drought monitoring and prediction using k-nearest neighbor algorithm. *Journal of AI and data mining* 5(2), 319–325 (2017)
9. Flores-Hernández, S., Saturno-Hernández, P.J., Reyes-Morales, H., Barrientos-Gutiérrez, T., Villalpando, S., Hernández-Ávila, M.: Quality of diabetes care: the challenges of an increasing epidemic in mexico. results from two national health surveys (2006 and 2012). *PloS one* 10(7), e0133958 (2015)
10. Hugoson, A., Thorstensson, H., Faltt, H., Kuylenstierna, J.: Periodontal conditions in insulin-dependent diabetics. *Journal of clinical periodontology* 16(4), 215–223 (1989)
11. Hunter, P.B.: Risk factors in dental caries. *International dental journal* 38(4), 211–217 (1988)
12. Lalla, E., Cheng, B., Lal, S., Kaplan, S., Softness, B., Greenberg, E., Goland, R.S., Lamster, I.B.: Diabetes mellitus promotes periodontal destruction in children. *Journal of clinical periodontology* 34(4), 294–298 (2007)
13. Li, L., Weinberg, C.R., Darden, T.A., Pedersen, L.G.: Gene selection for sample classification based on gene expression data: study of sensitivity to choice of parameters of the ga/knn method. *Bioinformatics* 17(12), 1131–1142 (2001)
14. Liaw, A., Wiener, M.: The randomforest package. *R News* 2(3), 18–22 (2002)
15. Mendoza Romo, M.Á., Padrón Salas, A., Cossío Torres, P.E., Orozco, M.S.: Prevalencia mundial de la diabetes mellitus tipo 2 y su relación con el índice de desarrollo humano. *Revista Panamericana de Salud Pública* 41, e103 (2018)
16. Montgomery, D.C., Peck, E.A., Vining, G.G.: Introduction to linear regression analysis. John Wiley & Sons (2015)
17. Oyedele, T., Fadeju, A., Adeyemo, Y., Nzomiwu, C., Ladeji, A.: Impact of oral hygiene and socio-demographic factors on dental caries in a suburban population in nigeria. *European Archives of Paediatric Dentistry* 19(3), 155–161 (2018)
18. Papapanou, P.N.: Periodontal diseases: epidemiology. *Annals of periodontology* 1(1), 1–36 (1996)
19. Ripley, B.D.: The r project in statistical computing. *MSOR Connections. The newsletter of the LTSN Maths, Stats & OR Network* 1(1), 23–25 (2001)
20. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., Müller, M.: proc: an open-source package for r and s+ to analyze and compare roc curves. *BMC Bioinformatics* 12, 77 (2011)
21. Singh, P., Pal, A.B., Anburajan, M., Kumar, J.: Computer-aided diagnosis of type 2 diabetes mellitus using thermogram of open mouth. In: *Progress in Advanced Computing and Intelligent Engineering*, pp. 167–177. Springer (2018)
22. Stojanovic M, Cvetanovic G, A.A.M.S.D.R.N.: Risk factors in dental caries. *Cent Eur J Public Health* 26(2), 104–110 (2018)
23. Thiele, S., Peltner, J., Richter, A., Mensink, G.B.: Food purchase patterns: empirical identification and analysis of their association with diet quality, socio-economic factors, and attitudes. *Nutrition journal* 16(1), 69 (2017)
24. Trevino, V., Falciani, F.: Galgo: an r package for multivariate variable selection using genetic algorithms. *Bioinformatics* 22(9), 1154–1156 (2006)
25. Twigg, S.M., Wong, J.: The imperative to prevent diabetes complications: a broadening spectrum and an increasing burden despite improved outcomes. *The Medical Journal of Australia* 202(6), 300–304 (2015)
26. Winkleby, M.A., Jatulis, D.E., Frank, E., Fortmann, S.P.: Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *American journal of public health* 82(6), 816–820 (1992)

27. Zanella-Calzada, L.A., Galván-Tejada, C.E., Chávez-Lamas, N.M., Galván-Tejada, J.I., Celaya-Padilla, J.M.: Multivariate features selection from demographic and dietary descriptors as caries risk determinants in oral health diagnosis: Data from nhanes 2013–2014. In: Electronics, Communications and Computers (CONIELECOMP), 2018 International Conference on. pp. 217–224. IEEE (2018)

Antena de parche semicircular con elemento parásito para CanSat

Cruz Ángel Figueroa Torres¹, José Luis Medina Monroy², Andrés Calvillo Téllez³

¹ Universidad Autónoma de Baja California, Ensenada, B. C., México

cruz.figueroa@uabc.edu.mx

² Centro de Investigación Científica y de Educación Superior de Ensenada, Ensenada, B. C., México

jmedina@cicese.mx

³ Instituto Politécnico Nacional, CITEDI, Tijuana, B. C., México

calvillo@citedi.mx

Resumen. En este trabajo se presenta el diseño de una antena de parche de estructura geométrica semicircular y con un elemento parásito, para su utilización en la radiocomunicación de satélites educativos de la categoría lata (CanSat). La antena desarrollada presentó un peso mínimo de 13.2 gramos, realizada físicamente en un material que cubre los requerimientos de tipo electromagnético, impresa en un sustrato FR-4, con dimensiones máximas de 54 mm de diámetro, a una frecuencia resonante de 2.4 GHz. La antena tiene un peso, aproximadamente del 2% del peso total del nano satélite (0.355 Kg). Las pérdidas por retorno S11 logradas son menores a -10 dB. Los resultados obtenidos, como la ganancia, patrón de radiación, ancho de banda, y principalmente sus dimensiones y peso, exhiben a la antena propuesta como una gran opción para ser empleada en nanosatélites.

Palabras clave: CanSat, antena de parche parasito, antena semicircular.

Semicircular Parasitic Patch Antenna for CanSat

Abstract. This paper presents the design of a patch antenna with a semicircular geometric structure and a parasitic element, for its use in the radio communication of educational satellites of the can category (CanSat). The developed antenna presented a minimum weight of 13.2 grams, physically made of a material that covers the requirements of the electromagnetic type, printed on a substrate FR-4, with maximum dimensions of 54 mm in diameter, at a resonant frequency of 2.4 GHz. The antenna has a weight, approximately 2% of the total weight of the nano satellite (0.355 Kg). The return losses S11 achieved are less than -10 dB. The obtained results, such as gain, radiation pattern, bandwidth, and mainly its dimensions and weight, show the proposed antenna as a great option to be used in nanosatellites.

Keywords: CanSat, parasitic patch antenna, semicircular antenna.

1. Introducción

Un CanSat es un prototipo de satélite destinado a la educación en la formación de recursos humanos en matemáticas, ciencia y tecnología. El diseño del CanSat debe ajustarse a las dimensiones y limitaciones de masa, sus componentes y subsistemas deben ser más pequeños en comparación con los satélites convencionales. Actualmente los satélites educativos tienen un papel importante para el desarrollo de tecnología en las agencias espaciales, industria e instituciones educativas.

Se puede conceptualizar a un CanSat como un prototipo educativo a escala de un satélite mayor. Una misión típica consta de: captura de video, toma de fotografías aéreas, y mediciones de las condiciones climáticas y atmosféricas. El CanSat puede realizar las funciones de coleccionar información de la carga útil y transmitirla a una estación terrena. Siendo la antena el elemento esencial para la transmisión formando parte del módulo de comunicaciones que integran el satélite educativo. Este prototipo de CanSat, ha logrado captar la atención de la comunidad académica como una forma de materializar teorías en aplicaciones espaciales. De tal forma que el desarrollo tecnológico forme capital humano en esta área, altamente valorado [1].

Su reducida dimensión permite que se construyan rápidamente, mientras que el peso determina el bajo costo para el lanzamiento. Los elementos básicos que lo constituyen son: fuente de alimentación, subsistema de computadora de abordo, subsistema de adquisición de variables físicas y el subsistema de comunicaciones, constituido por el transpondedor y la antena.

Una forma económica de hacer investigación y desarrollo tecnológico en el área de comunicaciones en especial antenas, es utilizar tarjetas de circuito impreso para señales de radiofrecuencia de material FR-4 sin embargo presenta un nivel de pérdidas significativo cuando opera señales electromagnéticas de alta frecuencia, con respecto al material de alto valor económico, que posee características eléctricas superiores como lo es el de teflón o politetrafluoroetileno (PTFE), esto implica que se espera que el diseño físico presente variaciones de funcionamiento respecto al modelo del diseño [2].

Entre las principales aplicaciones están el realizar investigación espacial, medio ambiente del espacio, investigación atmosférica, astronomía, ciencias planetarias, monitoreo del medio ambiente terrestre, entre otras [3-9].

Dentro del espectro de los nanosatélites está el CubeSat U1 de $10 \times 10 \times 10 \text{ cm}^3$, mientras que el CubeSat de 3U tiene un volumen de $30 \times 10 \times 10 \text{ cm}^3$. El peso total de este nanosatélite es de 1 Kg, aunque pueden existir algunas variaciones [10-14].

Otro tipo de nano satélite es el CanSat, propuesto por el profesor Robert Twiggs de la Universidad de Stanford en 1999, en una lata de refresco (355ml) [1]. El desarrollo lo promueve la Agencia Espacial Mexicana (AEM), para capacitación en el sector Aeroespacial y con el deseo de que los mejores equipos de trabajo mexicanos puedan participar con sus desarrollos a nivel internacional [1].

La agencia encargada de la regulación de la frecuencia de 2.4 a 2.5 MHz para satélites de aficionados es la Unión Internacional de Telecomunicaciones (ITU) [11].

2. Desarrollo

Existen escasos diseños para aplicaciones en nanosatélites, en especial en la categoría de los CanSat. En este trabajo se presenta una antena que puede ser colocada en la cara superior o inferior del CanSat. Las especificaciones de diseño para la antena propuesta son: frecuencia de operación en 2.45 GHz y con un ancho de banda de 100 MHz.

En la figura 1 se puede observar la estructura propuesta en este artículo. Se sabe que la frecuencia de resonancia de un elemento circular está determinada directamente por su diámetro y el punto de alimentación. Para calcular la frecuencia de resonancia del elemento radiador se emplearon las ecuaciones propuestas en [15].

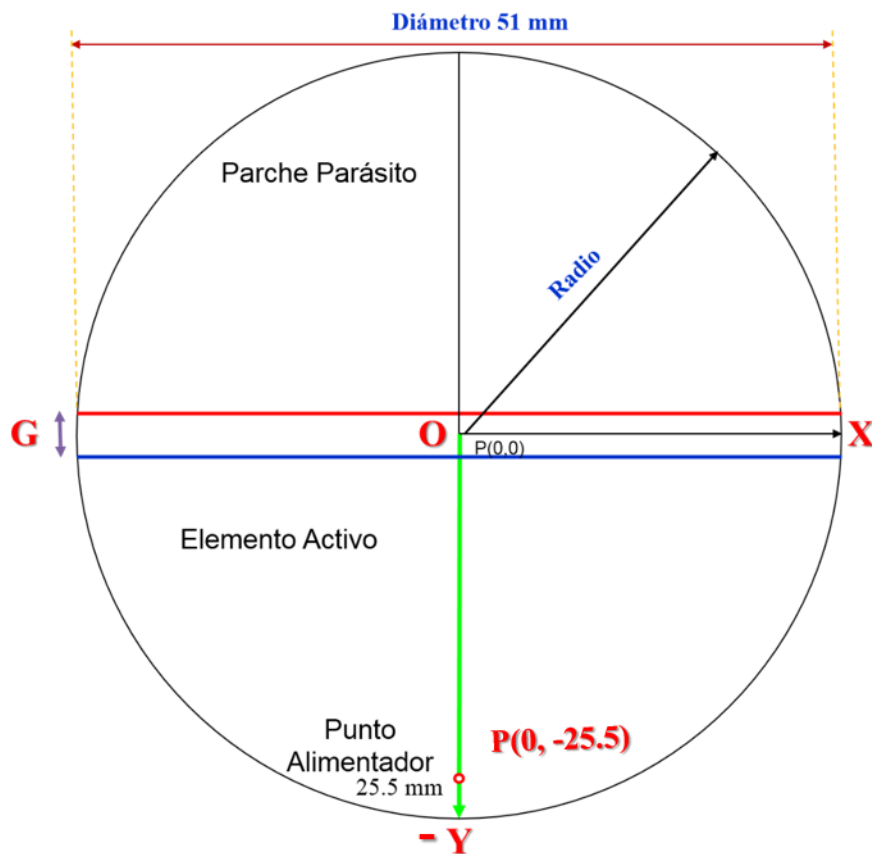


Fig. 1. Antena semicircular de parche con elemento parásito para CanSat.

El parche circular tiene un diámetro $D = 51$ mm. El puerto de alimentación está centrado en el parche, es decir en $P_y = 25.5$ mm. El punto de alimentación es un factor

crítico para establecer la frecuencia de operación de la antena, por esta razón, estos parámetros también son optimizados mediante la técnica de algoritmos genéticos, para determinar la ubicación del punto de alimentación dónde las pérdidas por retorno (parámetro S11) sean adecuadas, es decir que cumplan con la función objetivo. Cabe señalar que el parámetro S11 determina la impedancia de entrada de un dispositivo. El primer paso es definir la función objetivo centrado en las pérdidas por retorno (S11), las cuales deben ser menores a -10 dB en la banda de operación requerida. La Función objetivo queda expresada de la siguiente manera:

$$\text{Función objetivo} = \text{Min}(S11)_{2.45\text{GHz}} \quad (1)$$

La tabla 1 presenta los rangos de las variables a optimizar y el espacio de decisión.

El valor resultante de la abertura G es de 0.5 mm, el cual fue optimizado para sintonizar la frecuencia deseada, mediante el programa de cómputo de análisis electromagnético CST.

Tabla 1. Distancias de la separación de la antena y el elemento parasito y del punto central de la antena sobre el eje Y.

Parámetro	Espacio de decisión (mm)
G	0.1-1
Py	-6 - -25.5

En la figura 2 se puede apreciar el comportamiento de la impedancia de entrada de la antena para diferentes ubicaciones del punto de alimentación con respecto al eje “y”. Se sabe que la ubicación del punto de alimentación de una antena contribuye a determinar su frecuencia de resonancia y a mejorar su acoplamiento ($S11 < -10$ dB). Por lo cual, es necesario realizar un análisis paramétrico para identificar cuál es la ubicación óptima del puerto de alimentación, que garantice el mayor ancho de banda y el mejor acoplamiento en la banda de interés.

La figura 2 muestra algunos de los resultados del análisis paramétrico, los cuales se realizaron con ayuda del software de análisis electromagnético CST. Este análisis fue realizado desde $P_y = 0$ hasta -25.5 mm con pasos de 0.5 mm. Para ello se seleccionaron solo algunas curvas, con el objetivo de mostrar cómo influye la ubicación del puerto de alimentación en la frecuencia de resonancia. Con $P_y = -25$ el ancho de banda se redujo considerablemente, por lo cual se optó por seleccionar $P_y = -25.5$. Con ello se puede observar que cuando el puerto de alimentación está localizado en $P_y = -25.5$ mm se cubre la banda de 2.43 - 2.54 GHz para un $S11 < -10$ dB. Esta banda cumple con las especificaciones de diseño del CanSat.

Además, en base a nuestra experiencia en diseño de antenas, es conocido que cuando la antena sea colocada en la estructura del CanSat la frecuencia de resonancia tenderá a desplazarse hacia abajo un par de MHz, debido a que la estructura de los bordes del satélite hará ver que la geometría de la antena aumentó unos mm, con lo cual la frecuencia de resonancia tenderá a reducirse.

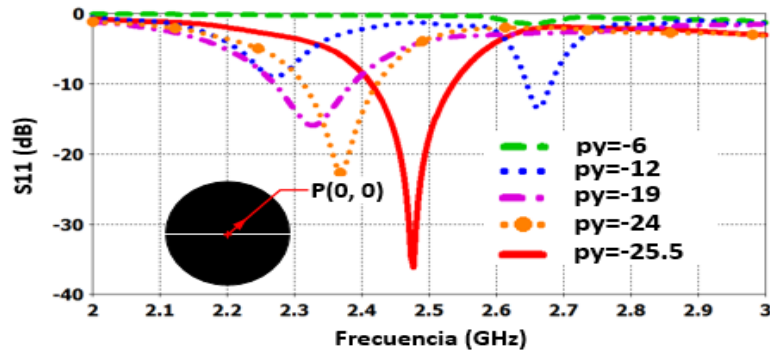


Fig. 2. Análisis del punto de alimentación.

3. Resultados

La antena fue construida en un sustrato FR-4, el cual es económico con respecto al PTFE y fácil de conseguir. La antena está impresa sobre un sustrato con diámetro de 54 mm y un espesor de 3.1556 mm, tal como se observa en la figura 3.

Una ventaja importante de este diseño es el peso, el cual es 13.5g, lo cual representa menos del 2% del peso total de un nanosatélite.

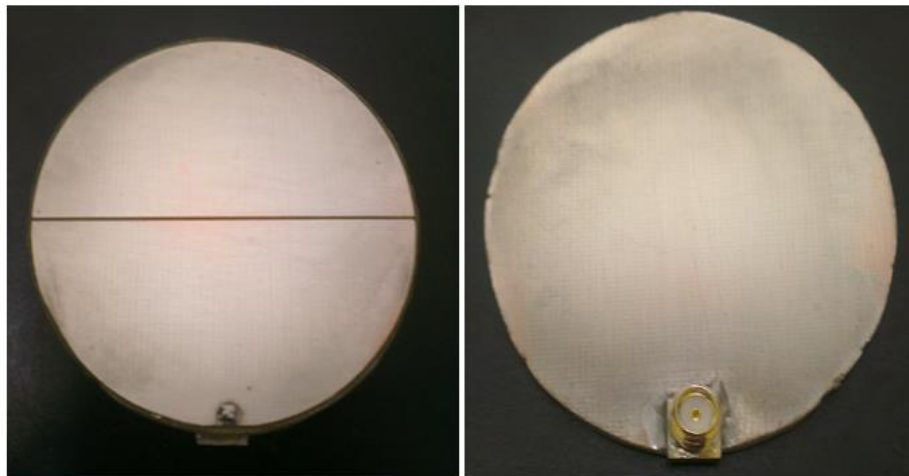


Fig. 3. Antena fabricada: a) vista superior, b) vista inferior.

La antena se diseñó y analizó electromagnéticamente utilizando el Software Computer Simulation Technology (CST). En la figura 4 se realiza una comparación entre los resultados simulados y medidos del parámetro S_{11} , el cual nos determina la frecuencia de resonancia y el ancho de banda de la antena. El ancho de banda se obtiene a partir de $S_{11} \leq -10$ dB. En la figura se puede observar que el rango de frecuencia de

operación medido es de 2.43 GHz - 2.54 GHz. Mientras que el simulado es de 2.41 GHz - 2.54 GHz ambos para un $S_{11} < -10\text{dB}$. Se hace notar que el rango de frecuencia de operación fue medido sin considerar el efecto del satélite. Anteriormente mencionamos que este efecto provocará que la frecuencia de resonancia se desplace hacia abajo un par de MHz. Sería interesante realizar la medición de la antenna una vez que el satélite cuente con todas sus etapas, para ver en qué medida se reduce la frecuencia de resonancia y con esto se puede advertir a los diseñadores de antenas para que consideren esta cantidad en sus diseños. Se puede notar una pequeña variación entre el resultado simulado y el medido, esta se debe a pequeñas variaciones en las dimensiones de la estructura ocasionadas en el proceso de fabricación.

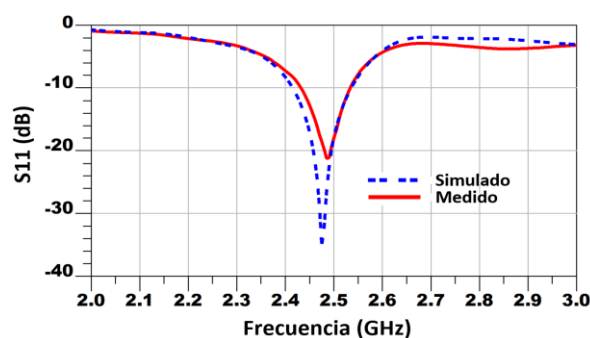


Fig. 4. Parámetro S11 antenna.

La ganancia de la antenna tiene un papel muy importante en el comportamiento de la señal emitida o recibida. Por lo tanto, es necesario conocer la ganancia de la antenna con exactitud, para poder determinar el nivel recibido por un receptor o el nivel emitido por un transmisor.

La ganancia de una antenna depende de: el número de elementos que la forman, las dimensiones del elemento radiador, espesor del sustrato, entre otras. Y se asocia con su directividad.

Una antenna con ganancia alta será más directiva, permitiendo concentrar la energía en una dirección en específico y disminuir el nivel de sus lóbulos laterales. En aplicaciones móviles se requieren antenas con gran cobertura, es decir, con baja directividad.

Existen varios métodos y configuraciones que se utilizan para medir la ganancia, y que dependen de la frecuencia de operación [16-17].

A frecuencias superiores a 1 GHz, usualmente se utiliza el espacio libre para medir la ganancia, puesto que a frecuencias menores es difícil simular las condiciones de espacio libre debido a las longitudes de onda grandes, razón por la cual se mide bajo condiciones con reflexión a tierra [16].

En la figura 5 se puede observar la ganancia de la antenna para el rango de frecuencias de 2.4 GHz a 3 GHz.

La medición de ganancia se llevó a cabo utilizando un analizador de espectros, un generador de señales de microondas y además una antenna patrón o de referencia.

El procedimiento de medición es el siguiente: se realiza un barrido en frecuencia y se efectúa la medición de potencia en cada frecuencia.

La ganancia para la frecuencia de 2.45 GHz es de 3dBi.

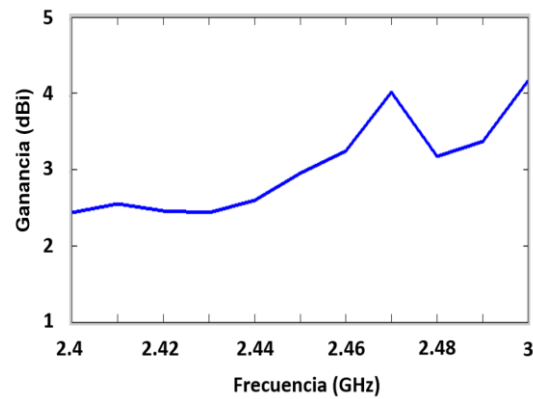


Fig. 5. Ganancia de la antena.

En la figura 6 se muestra la antena colocada sobre una lata de refresco, la cual simula un dispositivo CanSat, esto con el fin de demostrar que la antena propuesta en este trabajo de investigación cumple con los requerimientos de tamaño para ser utilizada en un nano satélite.



Fig. 6. Satélite en lata con antena semicircular de parche con elemento parásito.

El patrón de radiación es una representación gráfica de las propiedades de radiación de la antena en función de las coordenadas espaciales. En la mayoría de los casos, el patrón de radiación es determinado para la región de campo lejano. En la figura 7 se puede apreciar el patrón de radiación para la frecuencia de 2.4 GHz. El ancho del haz de mediana potencia (HPBW) es de 98.2 ° y la magnitud del lóbulo principal es de 5.4

dBi. El nivel de los lóbulos laterales es de -10.4 dB. Debido a que la antena tiene plano de tierra en la parte posterior del sustrato, el nivel de los lóbulos laterales es bajo.

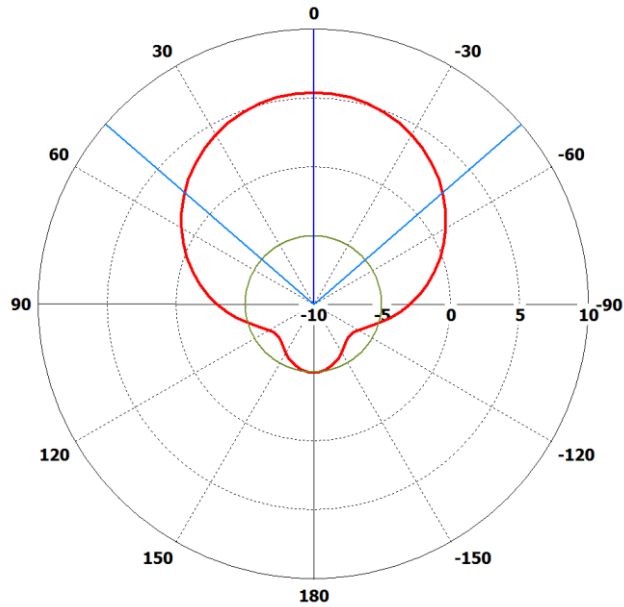


Fig. 7. Patrón de radiación a 2.4 GHz.

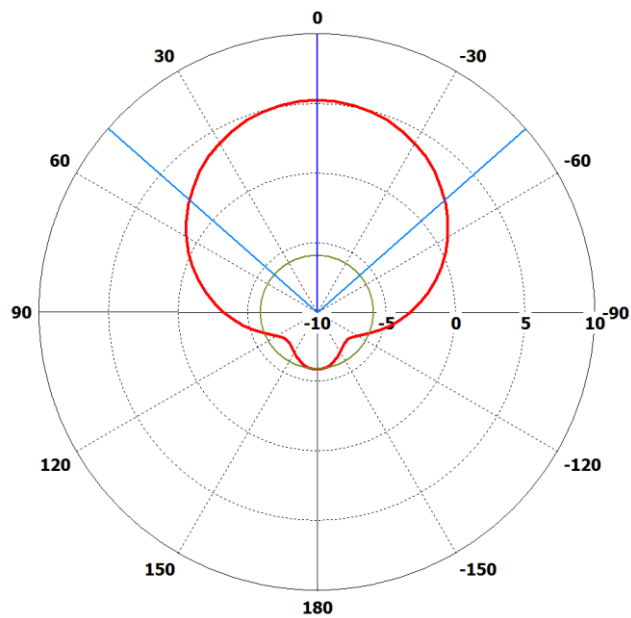


Fig. 8. Patrón de radiación a 2.45 GHz.

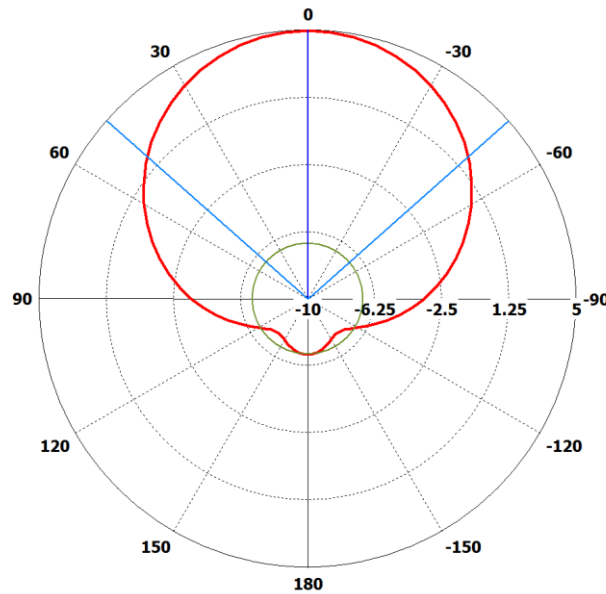


Fig. 9. Patrón de radiación a 2.5 GHz.

Para comprender adecuadamente que la antena sólo radia en una sola dirección, es decir sobre el parche circular, se realiza el análisis electromagnético del patrón de radiación en 3D para la frecuencia de 2.45 GHz, el cual se muestra en la figura 10.

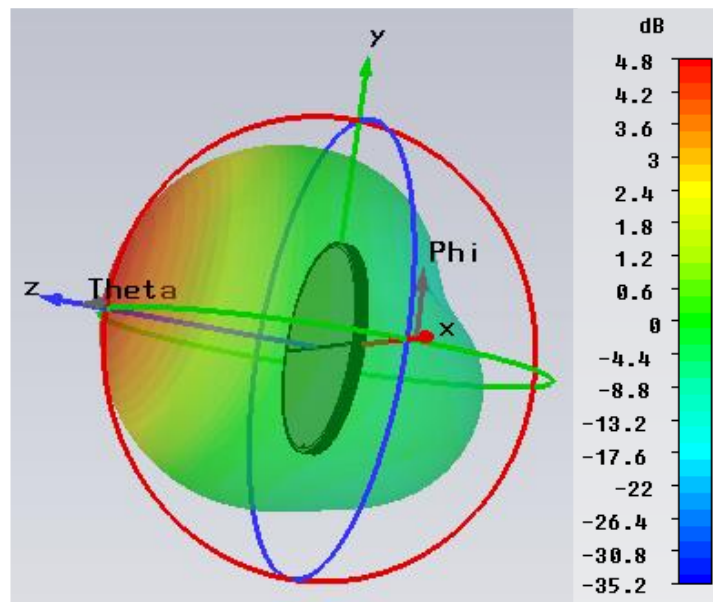


Fig. 10. Patrón de radiación 3D a 2.45 GHz.

En la figura 11 se puede observar que la máxima eficiencia obtenida en el diseño propuesto es de 80 %, siendo esta para la frecuencia de 2.35 GHz. Mientras que la eficiencia mínima obtenida es de 15 % para la frecuencia de 2.65 GHz. Cabe resaltar que la antena tiene buena eficiencia ($> 60\%$) en la frecuencia de operación del CanSat, es decir en 2.4 GHz. La baja eficiencia en algunas frecuencias es debido a que la antena no es diseñada para operar en tales frecuencias.

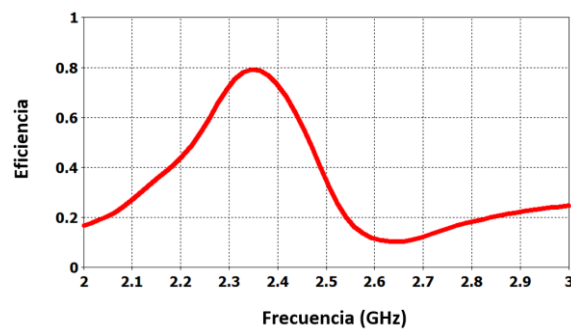


Fig. 11. Eficiencia de la antena.

4. Discusión

La antena mostró parámetros que satisfacen en ganancia, directividad, impedancia de entrada y el patrón de radiación con niveles de lóbulos laterales. Existen algunos diseños de antenas orientadas a nanosatélites, en [13], muestra un diseño de antena rómbica, cuyo porcentaje de ancho de banda es de 7.5 % y sus dimensiones son $10 \times 10 \text{ cm}^2$, con un peso de 157 g, el cual representa el 15.7 % del peso del CubeSat, por lo cual la antena limita en peso la construcción de las etapas de detección, amplificación y procesamiento de la señal. Con el objetivo de aumentar la ganancia y por consecuencia la directividad, en el trabajo presentado en [6], se utiliza un reflector inflable para la antena. Dicha estructura proporciona una ganancia de 9 dBi, sin embargo, el reflector requiere un tanque presurizado, por ello, la antena tiene un periodo de vida corto. En este artículo se presentó una antena que se ha diseñado y construido, la antena es una estructura original, la cual cumple con los requerimientos fundamentales para la integración en nanosatélites, como lo son: el peso y las dimensiones de la antena. En este trabajo se incluyen mediciones de: la impedancia de la antena, ganancia y patrón de radiación.

5. Conclusiones

En este artículo se presentan los resultados del diseño, optimización y construcción de una antena que puede emplearse en nanosatélites. La ventaja de esta antena comparada con otras es: la fabricación en un sustrato de bajo costo, pero sobretodo el tamaño, dado que dicha estructura no aumenta el área ocupada por el sistema.

El peso de la antena es una de las ventajas más importantes de este diseño, el cual es de tan solo 13.5 g. Además, la antena presenta un buen comportamiento en ancho de banda, el cual es de 100 MHz. Por otra parte, la ganancia obtenida es de 3dBi para la frecuencia de 2.45 GHz.

El patrón de radiación mostrado por la antena es omnidireccional, lo cual es muy útil, debido a que se tiene un área de cobertura amplia.

Contrario a lo que normalmente se puede encontrar en la literatura si bien en el diseño de la antena propuesta se emplean metodologías y modelos ya conocidos; la antena de parche semicircular con elemento parásito para CanSat es una novedad, debido a que después de indagar en el estado del arte, no se encontró una topología semejante en aplicaciones en CanSat. Esto la hace una aportación de interés para los nuevos prototipos de satélites pequeños. Actualmente se continúan usando la mayoría de desarrolladores de este tipo de satélites antenas dipolo, las cuales ocupan demasiado espacio, incrementan el peso y sacan de balance al CanSat.

Referencias

1. Navarro Lieja, D.: Integración de un satélite CanSat y propuesta de sus medios de lanzamiento para capacitar recursos humanos a nivel universitario. Tesis de Maestría, UNAM, (2015)
2. Coonrod, J.: Bending and Forming High Frequency Printed Circuits. IPC Printed Circuits Expo, APEX and the Designers Summit (2007)
3. Nascetti, A., Pittella, E., Teofilatto, P., Pisa, S.: High-Gain S-band patch antenna system for Earth-Observation CubeSat Satellites. *IEEE Antennas and Wireless Propagation Letters*, Vol. 14, pp. 434–437 (2015)
4. Swartwout, M.: The first one hundred CubeSats: A statistical look. *Journal of JoSS Small Satellite*, Vol. 2, pp. 213–233 (2014)
5. Sandau, R.: Status and trends of small satellite missions for earth observation. *Acta Astronautica*, Vol. 66, pp. 1–12 (2010)
6. Mahmoud, A., Elazhary, T., Zaki, A.: Remote sensing CubeSat. In: *Proceedings SPIE*, Vol. 7826, pp. 78262I–78262I-8 (2010)
7. Babuscia, A., et al.: Commcube 1 and 2: A CubeSat series of missions to enhance communication capabilities for cubesat. In: *Proceedings IEEE Aerospace conference*, pp. 1–19 (2013)
8. Nugent, R., Munakata, R., Chin, A., Coelho, R., Puig-Suari, J.: The CubeSat: The picosatellite standard for research and education. *Aerospatiale Eng*, Vol. 805, pp. 756–5087 (2008)
9. Toorian, A., Diaz, K., Lee, S.: The CubeSat approach to space access. In: *Proceedings IEEE Aerospace conference*, pp. 1–14 (2008)
10. Heidt, H., Puig-Suari, J., Moore, A. S., Nakasuka, S., Twiggs, R. J.: CubeSat: A new generation of picosatellite for education and industry low-cost space experimentation. In: *Proceedings 14th annual USU Small satellite conference*, SSC00-V-5 (2000)
11. Handbook of frequency allocations and spectrum protection for Scientific uses. Committee on radio frequencies council (2007)
12. CubeSat program and community official. Available at: <http://cubesat.atl.calpoly.edu/>
13. Rudge, A. W., Milne, K., Olver, A. D., Knight, P.: *The handbook of antenna design*. London, Peter Peregrinus Ltd. (1986)

14. Pedroza, J. S. , Peña, S., Martínez-Zúñiga, F.: A 2.4 GHz cross rhombic antenna for a cube satellite application. *International Journal of Antennas and propagation*, Vol. 2014 (2014)
15. Babuscia, A., Van de Loo, M., Wei, Q. J., Pan, S., Mohan, S., Seager, S.: Inflatable antenna for CubeSat: fabrication, deployment and results of experimental test. In: *Proc. IEEE Aerospace conference*, pp. 1–12 (2014)
16. Verma, K., Nasimuddin, N.: Simple expressions for the directivity of a circular microstrip antenna. *IEEE antennas and propagation magazine*, Vol. 44, pp. 91–95 (2002)
17. Balanis Constantine, A.: *Antenna Theory: Analysis and design*. John Wiley and Sons, 2nd ed., New York, p. 941 (1997)

Componente del Internet de las cosas para detectar patrones de deambulaje en pacientes con demencia

Hector Villarreal¹, Mayra Lizarraga¹, Arnoldo Diaz-Ramírez¹,
Verónica Quintero Rosas¹, Juan Pablo García-Vázquez²

¹ Tecnológico Nacional de México/IT de Mexicali,
Departamento de Sistemas y Computación, México
{hevr93, mayra.lizarraga}@gmail.com,
{adiaz, veronicaquintero}@itmexicali.edu.mx

² Universidad Autónoma de Baja California, Facultad de Ingeniería,
México
pablo.garcia@uabc.edu.mx

Resumen. La demencia es un problema de salud que afecta la vida de millones de personas en todo el mundo. Esta enfermedad, además de ser la causa de muerte de muchas personas, provoca que los enfermos sean incapaces de realizar actividades de índole cotidiano. Adicionalmente, el deambulaje es uno de los síntomas de la demencia. Cuando un enfermo deambula hace desplazamientos o caminatas sin supervisión, con la esperanza de encontrar algún lugar o persona que le resulte familiar. Detectar el nivel de avance de la enfermedad de demencia no es una tarea sencilla. Debido a que se trata de padecimientos causados por daños cerebrales, una manera de conocer el estado de la enfermedad es a través de pruebas costosas, tales como tomografías o resonancias magnéticas. Sin embargo, estas pruebas pueden provocar demasiado estrés en los pacientes, lo que puede dificultar su aplicación. Por tal motivo es importante desarrollar métodos alternativos y no invasivos para este fin. Recientemente se ha detectado que existe una relación entre los patrones de deambulaje y el nivel del avance de la demencia. En este artículo se presenta un componente para la detección de patrones de deambulaje de pacientes afectados con demencia. El componente propuesto, desarrollado con base en el paradigma del Internet de las Cosas, no es invasivo ni requiere de infraestructura previamente instalada en los edificios en los que será utilizado. La evaluación del sistema propuesta mostraron una alta tasa de detección.

Palabras clave: demencia, Internet de las cosas, deambulaje.

Component of the Internet of Things for Detection of Wandering Patterns in Patients with Dementia

Abstract. Dementia is a health problem that affects millions of persons around the world. This disease, which is one of the main death causes

of the elderly, does not allow them to perform every day activities. Additionally, wandering is a common symptom of people affected by dementia. When a patient wanders, she walks without supervision seeking for a familiar place or person. Detecting the level of dementia is not an easy task. Since dementia is caused by brain damage, one way to detect it is through the use of invasive techniques, such as tomographies or magnetic resonance, which may cause a lot of stress in patients. Recently, a relationship between the level of dementia and wandering patterns has been found. In this paper, a component based on the Internet of Things paradigm to detect wandering patterns of people affected by dementia is introduced. The proposed component is not invasive and does not require previously installed infrastructure. The evaluation of the component showed a very good detection rate.

Keywords: dementia, Internet of things, wandering.

1. Introducción

La demencia es un problema de salud que afecta la vida de millones de personas en todo el mundo. Debido a su impacto, ha recibido mucha atención de especialistas de diversas áreas del conocimiento [18]. La demencia puede afectar completamente el estilo de vida de quienes la padecen y de los que los rodean, llegando a cambiar la vida de las personas, familias e incluso comunidades. Esta enfermedad, además de ser la causa de muerte de muchas personas, provoca que los enfermos sean incapaces de realizar actividades de índole cotidiano.

La demencia no es una enfermedad que se pueda predecir o identificar con un gran grado de exactitud, ya que tiende a no ser específica; es decir, puede variar de paciente en paciente, e incluso los síntomas pueden ser diferentes en cada uno de los afectados. La demencia es el nombre genérico que recibe una amplia gama de enfermedades que afectan las funciones cognitivas. Cada tipo de demencia afecta una región particular del cerebro [4]. Por ejemplo, la enfermedad del Alzheimer provoca que los niveles de algunas proteínas dentro y fuera de las células cerebrales aumenten [8]. Como consecuencia de esto, las células de la región afectada no pueden mantenerse saludables y tienen problemas para comunicarse. Esta enfermedad ataca principalmente a la región cerebral del hipocampo, el cual es el encargado del aprendizaje y la memoria. Por tal motivo, la pérdida de memoria es una de los principales síntomas de la enfermedad del Alzheimer [2].

Las estadísticas indican que año con año aumentan el número de casos de personas afectadas con demencia. Se estimó que para el año 2017, 5.5 millones de estadounidenses mayores a 65 años sufrirían algún tipo de demencia, de los cuales 200,000 serían personas menores de 65 años [5]. En México, en el año 2009 se registraron 1,200 casos de muertes por algún tipo de demencia. Además, se registraron un poco más de 142, 000 casos confirmados en ese año [17].

La demencia es una enfermedad que no tiene cura, que es progresiva y que afecta no solo la calidad de vida del paciente sino de quienes lo cuidan. El tratamiento para este padecimiento varía de acuerdo al grado de avance de la enfermedad. Un diagnóstico erróneo puede provocar situaciones de riesgo tanto en pacientes como en cuidadores. Un paciente puede cambiar de etapa en el avance de su enfermedad y presentar nuevos comportamientos. Por ejemplo, salir de su casa sin supervisión, extraviarse y poner en riesgo su integridad física o su vida. Por tal motivo, poder saber con anticipación si un paciente corre este riesgo es de vital importancia.

Detectar el nivel de avance de la enfermedad de demencia no es una tarea sencilla. Debido a que se trata de padecimientos causados por daños cerebrales, una manera de conocer el estado de la enfermedad es a través de pruebas costosas, tales como encefalogramas, tomografías o resonancias magnéticas. Sin embargo, estas pruebas pueden provocar demasiado estrés en los pacientes, lo que puede dificultar su aplicación. Por tal motivo es importante desarrollar métodos alternativos y no invasivos para este fin.

El deambulaje es uno de los síntomas de la demencia. Cuando un enfermo deambula hace desplazamientos o caminatas sin supervisión, con la esperanza de encontrar algún lugar o persona que le resulte familiar. Este fenómeno se lleva a cabo a pesar de que el paciente esté en su propio hogar o rodeado de personas conocidas [3]. Es común que mientras una persona deambula sea incapaz de reconocer el entorno que lo rodea, debido a los problemas de memoria y de concentración producidos por la demencia. Como consecuencia del deambulaje los pacientes pueden sufrir accidentes como caídas, golpes, torceduras, etc. Los episodios de deambulaje tienen una duración máxima de 5.41 minutos [1]. Por otra parte, se considera que un episodio concluye cuando la persona pasa mas de 15 segundos detenido en el mismo sitio [10].

Martino-Saltzman *et al.* en [15] identificaron que existe una relación entre el nivel de deterioro de la persona afectada por la demencia y el tipo de trayectoria que presenta al caminar, particularmente cuando deambula. En el desarrollo de los experimentos presentados en [15] se identificaron tres grupos de personas, clasificados por el tipo de avance de la demencia en cada individuo: nula/leve, moderada y avanzada. Durante los experimentos se registraron mas de 5,500 episodios de deambulaje de 40 residentes de 130 estancias geriátricas de la ciudad de Atlanta, Georgia, Estados Unidos. De los 40 residentes, 24 fueron identificados con tendencias a deambular, llamados también *errantes* debido a la naturaleza imprecisa o errónea de sus desplazamientos. Los 16 pacientes restantes fueron identificados como personas sin tendencias a deambular o no *errantes*.

Con base en la observación de los patrones de deambulaje, estos fueron clasificados en cuatro grupos: directo (*direct*), rítmico (*pacing*), cíclico (*lapping*) y aleatorio (*random*). El primero representa un desplazamiento de un punto a otro sin desviación. El segundo, denominado rítmico, representa un desplazamiento repetitivo, de un punto a otro y de regreso, dentro de una área delimitada. En el patrón cíclico la persona se desplaza en movimientos semi-circulares, a menudo en repetidas ocasiones. En el movimiento aleatorio el paciente se

desplaza a diferentes puntos, sin una área delimitada y sin repeticiones. El patrón directo es considerado como un desplazamiento eficiente y síntoma de salud. En contraste, los patrones rítmico, cíclico y aleatorio son considerados desplazamientos ineficientes y relacionados con deterioro cognitivo. En el Cuadro 1 se muestran los resultados de los experimentos realizados en [15] de pacientes errantes y no errantes, clasificados por patrón de deambulaje. Puede observarse que la mayoría de los episodios de deambulaje de personas no errantes fueron eficientes (94 %). Esta tasa disminuyó en los pacientes no errantes, siendo el patrón cíclico el mas común de los episodios no eficientes (17.3 %).

Tabla 1. Comportamientos presentados por los pacientes en contraste errantes y no errantes [15].

TIPO DE PACIENTE	PATRONES DE MOVIMIENTOS			
	<i>Directo</i>	<i>Aleatorio</i>	<i>Rítmico</i>	<i>Cíclico</i>
<i>Errante</i>	81.3 %	1 %	0.1 %	17.3 %
<i>No Errante</i>	94 %	0.4 %	1.3 %	4.3 %

De los 40 residentes participantes en el estudio, a 32 de ellos se les diagnosticó algún tipo de demencia: ligera, moderada o severa. El Cuadro 1 muestra la clasificación de los experimentos realizados en [15] de los episodios de deambulaje de los pacientes afectados por demencia. Los episodios fueron clasificados por patrón de deambulaje. Puede observarse que a mayor daño cognitivo es mayor la tasa de desplazamientos ineficientes.

Tabla 2. Patrón de deambulaje clasificado por avance de la enfermedad [15].

CARACTERISTICAS	NIVEL DE DEMENCIA		
	<i>Ligero/Nulo (N = 12)</i>	<i>Moderado (N = 11)</i>	<i>Severa (N = 9)</i>
<i>Directa</i>	94.7 %	92.5 %	72 %
<i>Aleatoria</i>	0.2 %	1 %	1.7 %
<i>Rítmica</i>	1.5 %	0.0 %	0.3 %
<i>Cíclica</i>	3.6 %	6.5 %	26 %

Los resultados presentados en [15] identifican una relación entre el grado de avance de la demencia y los patrones de deambulaje. Es importante señalar que la identificación de los patrones de deambulaje en [15] se hicieron a través de grabaciones de video y observación. Este es un proceso costoso y lento. Por tal motivo se destaca la importancia de contar con un sistema alternativo para su detección. En este artículo se presenta un componente para la detección de

patrones de deambulaje de pacientes afectados con demencia. El componente propuesto, desarrollado con base en el paradigma del Internet de las Cosas, no es invasivo ni requiere de infraestructura previamente instalada en los edificios en los que será utilizado. La evaluación del sistema propuesta mostraron una alta tasa de detección.

El resto del documento está organizado de la siguiente manera. La Sección 2 presenta algunos de los trabajos relacionados. En la Sección 3 se discute la arquitectura del componente propuesto. La Sección 4 presenta los resultados de su evaluación. Finalmente, en la Sección 5 está dedicada a las conclusiones y el trabajo futuro.

2. Trabajos relacionados

Existen varias propuestas para monitorear actividades de personas afectadas con demencia. En esta sección se presentan algunas de las mas importantes.

El proyecto iWander es un sistema para la monitorización de personas con demencia que pretende ser una opción económica y sencilla de implementar para el apoyo a los cuidadores. El sistema está diseñado para hacer uso de los sensores existentes en un dispositivo móvil, entre los que destacan el GPS [13]. Este sistema se encarga de definir un espacio seguro donde interactúa el usuario, como es el hogar de la persona. Y si el paciente sale de esta zona segura, el programa ejecuta un algoritmo para definir si se trata de un episodio de deambulaje o no [21].

La implementación del algoritmo se divide en 3 fases. La primera se denomina *pre-instalación*, que consiste en la recopilación de datos sobre el usuario, además de la creación de las variables necesarias para la funcionalidad del sistema. Las relaciones y estadísticas entre las variables son las únicas entradas de datos. La segunda es denominada *configuración*. En esta fase se configura el dispositivo para un usuario en particular. Para ello se solicita información como la edad y el nivel de demencia. Otros datos necesarios para el algoritmo se obtienen de manera automática por el sistema. La última fase es la de *aprendizaje extendido*. En esta fase se analizan los hábitos del usuario, se definen las zonas seguras e identifican los episodios de deambulaje. Gracias a esta fase, el sistema aprende *constantemente*. A mayor sea el uso, el sistema es más preciso y eficaz.

Si el paciente sale de la zona segura, se considera que la persona se encuentra en un episodio de demencia, por lo que se toman medidas antes de mandar una alerta. En primer lugar, una notificación le solicita a los pacientes que indiquen si están bien. En caso de ser positivo, el proceso de alerta se detendrá, reduciendo así los falsos positivos. Si el paciente no responde, se deduce que él o ella no está bien y puede perderse. Las coordenadas de GPS se utilizan con las herramientas de navegación de Google Maps [12], para darles instrucciones sobre la zona segura. Si no se realiza ningún progreso después de un período de tiempo determinado, se envía una alerta a los cuidadores.

iSenior [19] es un sistema ciber-físico que soporta funciones como monitorización, alerta y solicitudes de asistencia. Este sistema es una herramienta para el

apoyo a personas mayores. El sistema proporciona apoyo como monitorización y alertas en parámetros como latidos del corazón y nivel de actividad, así como detección de situaciones potencialmente peligrosas. Adicionalmente, proporciona funciones para supervisar las condiciones de salud, que abarcan la recopilación y el almacenamiento de datos, la administración de alarmas, la localización, el acceso remoto, varios perfiles de usuario y la configuración automática.

El sistema cuenta con la capacidad de comunicarse con una unidad central a través de GSM/GPRS [16] [7] o a través de una red inalámbrica de sensores o WSN [11]. Cuenta con una arquitectura especializada, a través de múltiples sensores que se encargan de recolectar información del paciente, quedando a disposición de los cuidadores, los cuales pueden acceder desde cualquier equipo o dispositivo móvil. Además, cuenta con la capacidad de alertar a los cuidadores en caso de que alguno de ellos deje su cuarto o incluso el edificio. Además, es capaz de identificar la dirección en la que se dirige el paciente, pudiendo así el cuidador conocer la ubicación actual evitando accidentes durante el proceso.

En [14], Lin *et al.* propusieron un algoritmo denominado θ_WD , para la detección de trayectorias circulares o cíclicas en pacientes afectados con demencia. El algoritmo utiliza el sensor GPS de los dispositivos móviles para identificar el tipo de trayectoria. Una limitante de esta propuesta es que tan solo funciona en exteriores.

En [22] se propone un algoritmo para la detección de los patrones de deambulaje identificados por Martino-Saltzman *et al.* [15]. Utilizando la potencia de la señal de puntos de acceso WiFi previamente instalados en el edificio, el algoritmo determina la trayectoria del paciente. Para un cálculo sencillo, se supone un tiempo de viaje fijo, entre dos ubicaciones conocidas. Cada episodio de deambulaje (errante y no deseable) consiste en uno o varios movimientos consecutivos. El punto de inicio de un episodio es el punto final del episodio anterior. El algoritmo se inicializa para comenzar en el punto de inicio de un episodio.

Hasta donde los autores conocen, no existe un método para la detección de los patrones de deambulaje de los pacientes con demencia que funcione en interiores y que no requiera infraestructura previamente instalada en el edificio. En la siguiente sección se presenta un componente con esas características.

3. Arquitectura del componente propuesto

En este artículo se propone un componente capaz de detectar y categorizar los patrones de deambulaje de pacientes con demencia. El componente se basa en el paradigma del Internet de las Cosas, en el que sistemas embebidos, equipados con sensores, son colocados en cosas o personas para que de manera ubicua monitoricen y detecten eventos de interés, y a través de tecnologías de comunicación inalámbrica permitan que usuarios o sistemas informáticos reciban la información recolectada y puedan actuar en consecuencia [6]. El componente propuesto no requiere infraestructura previamente instalada en el hogar o residencia geriátrica en la que habita el paciente. La arquitectura general

del componente propuesto se presenta en la Fig. 1. Como puede observarse, el componente consta de cinco módulos, que se describen a continuación.

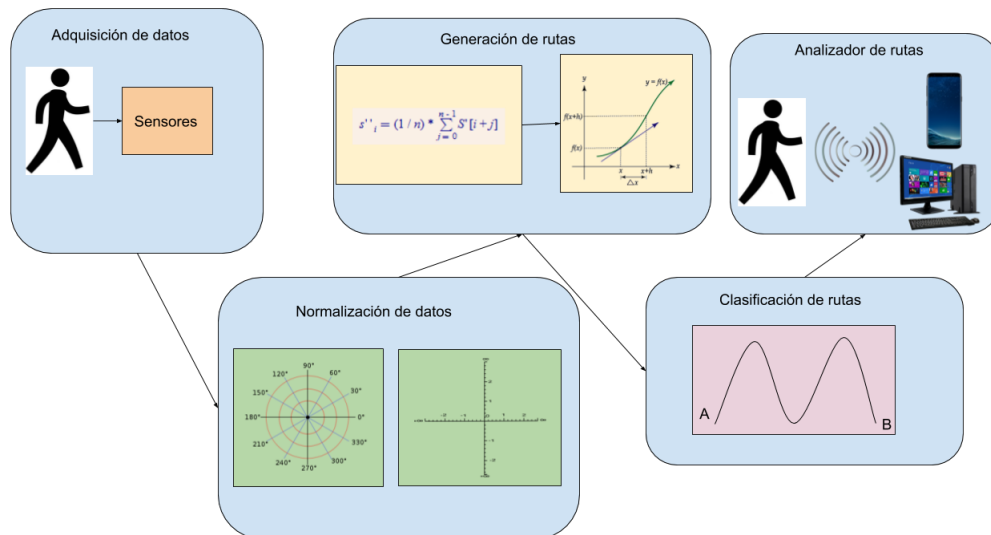


Fig. 1. Arquitectura del sistema.

3.1. Módulo de adquisición de datos

El módulo de adquisición de datos tiene como objetivo tomar mediciones de la aceleración y la dirección de movimiento del paciente, además de la intensidad campo magnético terrestre, con el propósito de tener una geolocalización del sujeto. Para este fin se utilizan sensores como el magnetómetro, el giroscopio y el acelerómetro, en un dispositivo adherido a la persona. Los sensores, a pesar de contar con una gran precisión en la actualidad, en ocasiones registran mediciones erróneas debido al ruido externo o fallos en el mismo sensor. Para tratar dichos errores es común utilizar algún filtro, que es un algoritmo de procesamiento digital de señales. Existen múltiples tipos de filtros, que ayudan con diferentes propósitos. El módulo de adquisición de datos utiliza un filtro paso-bajo para reducir el efecto de las mediciones erróneas de los sensores, tal y como el utilizado en sistema propuesto en [9], que permite la identificación y trazado de rutas de personas en interiores, sin necesidad de infraestructura previamente instalada.

3.2. Módulo de normalización de datos

Con las mediciones realizadas por los sensores del dispositivo, se obtienen valores del ángulo y la distancia del movimiento del paciente. Utilizando estos

datos como coordenadas polares y aplicando un par de funciones matemáticas, se obtienen coordenadas rectangulares del sujeto, pudiendo identificar las posiciones del mismo mientras se desplaza.

Teniendo un vector S conformado por n tuplas de datos obtenidos mediante los sensores, se llamará α a los datos que correspondan a la dirección de un paso y d a los datos que correspondan a la distancia del paso, donde i corresponde a la i -ésima medición. Por lo tanto, se tiene que

$$S = \{s_i = (\alpha_i, d_i) | 1 \leq i \leq n, n \in \mathbb{N}\}. \quad (1)$$

Definiendo estos datos como coordenadas polares, por cuestión de simplicidad, se llevará a cabo una transformación de coordenadas polares a coordenadas rectangulares:

$$x_i = d_i \times \sin(\alpha_i) | 1 \leq i \leq n, n \in \mathbb{N}, \quad (2)$$

$$y_i = d_i \times \cos(\alpha_i) | 1 \leq i \leq n, n \in \mathbb{N}, \quad (3)$$

donde x son coordenadas en el eje horizontal de un plano y y son coordenadas en el eje vertical de un plano, i es el i -ésimo dato de las mediciones registradas por los sensores. Tomando ambos valores, se pueden definir como coordenadas rectangulares en un plano, definiendo entonces que $s'_i = (x_i, y_i)$, por lo tanto

$$S' = \{s'_i = (x_i, y_i) | 1 \leq i \leq n, n \in \mathbb{N}\}, \quad (4)$$

donde S' representa un vector de coordenadas rectangulares.

3.3. Módulo de generación de rutas

Si las distancias recorridas por una persona son mayores a veinte metros, el módulo de adquisición de datos es capaz de detectar las rutas con un aceptable grado de precisión. Sin embargo, si las distancias recorridas son de menor longitud, como es común en desplazamientos de las personas afectadas por demencia, es necesario utilizar otro filtro. En el módulo de generación de rutas se utiliza el filtro de movimiento promedio [20], debido a que este filtro cuenta con la capacidad de ajustarse a una gráfica modelo, por lo que la pérdida de información es mínima y su implementación es sencilla. Para ello se aplica la función de suavizado:

$$s''_i = (1/n) \times \sum_{j=0}^{n-1} S'[i+j]. \quad (5)$$

En esta función n es el número de mediciones totales hechas por los sensores, en donde, i es la i -ésima columna de la matriz de mediciones y j es la j -ésima línea de la matriz. En la matriz de mediciones las columnas son los valores de los ejes de un plano y los renglones son cada una de las mediciones realizadas. En otras palabras, lo que se realiza es promediar un número de puntos de una señal de entrada para producir los puntos de una señal de salida.

Por lo tanto, se creará un nuevo vector de datos que llamaremos S'' , donde $s''_i = (x''_i, y''_i)$, donde x'' y y'' son las coordenadas de salida después de la aplicación de la función de suavizado, por lo tanto, es correcto decir que

$$S'' = \{s''_i = (x''_i, y''_i) | 1 \leq i \leq n, n \in \mathbb{N}\}. \quad (6)$$

El procedimiento de identificaciones de diferencias en los datos de los tres tipos de movimientos (directo, rítmico y cíclico) definidos por [15], se realizó por medio de la pendiente de las rectas que representan las trayectorias de cada paso los pacientes, y los ángulos de las rectas formadas por las trayectorias de cada paso del paciente. Para la pendiente se implementó la función

$$M_j = (y''_i - y''_{i-1}) / (x''_i - x''_{i-1}). \quad (7)$$

M es un vector de pendientes, donde las pendientes representan la razón de cambio de las mediciones, pudiendo diferenciar de esta manera los tipos de movimientos de los pacientes.

Los datos a utilizar corresponden al vector S'' , creando un nuevo vector que llamaremos S''' , donde $s'''_j = M_j | 1 \leq j \leq n - 1$. Por lo tanto, se tiene que

$$S''' = \{s'''_j = M_j | 1 \leq j \leq n, n \in \mathbb{N}\}. \quad (8)$$

Se define como un episodio de deambulaje al movimiento total realizado por un sujeto al perder la capacidad de reconocer el lugar o las personas que lo rodean, ya sea parcial o totalmente, además se conocerá a dicho episodio como Ω .

3.4. Módulo de clasificación de rutas

Para el proceso de clasificación de la ruta obtenida (el conjunto de ángulos representa la ruta del sujeto), se utiliza la definición de los segmentos relevantes, donde se define como un segmento relevante a un conjunto de ángulos del mismo signo matemático. Se define también el concepto de segmentos no relevantes, que están formados por un máximo de dos mediciones de signo diferente al del resto de mediciones del segmento. Los segmentos no relevantes son causados por datos espurios que no fueron eliminados por los filtros aplicados a las mediciones obtenidas por los sensores.

Un segmento es un conjunto de mediciones del mismo tipo de datos (positivos o negativos) consecutivos, que pertenecen a la ruta definida con anterioridad. Estos segmentos nos arrojan información sobre el tipo de ruta seguida por el sujeto.

Se definirá Z como un segmento, el cual está constituido por n elementos positivos o negativos, pero no ambos. Se define entonces:

$$Z = |m_i, m_{i+1}, \dots, m_{i+n}|, n \in \mathbb{N}\}. \quad (9)$$

Por la definición anterior, es correcto decir que un segmento relevante cumple ciertas condiciones, como son el numero de elementos que lo conforman y la consideración del punto donde los elementos del segmento realizan un cambio

de razón, cambiando su signo e iniciando un nuevo segmento. Se dice que, un segmento relevante Z' esta formado por:

$$Z' = \{Z' \subset \Omega | Z' = [a < m_i, m_{i+1}, \dots, m_{i+n} > b] | a, b, i, n \in \mathbb{N} | a \neq b\}, \quad (10)$$

donde a, b son los límites que dictan la razón de cambio de los valores del segmento; es decir, si se supera dicho limite termina el segmento relevante.

Tipos de movimientos Los movimientos directos de acuerdo a [15], son movimientos realizados por el paciente cuando se dirige de un punto A a un punto B sin ningún tipo de desviación. Este tipo de movimiento está conformado por un solo movimiento relevante, ya sea de datos positivos o negativos,. De existir segmentos no relevantes, son ignorados en éste y todos los movimientos:

$$MD = \{Z'_1 \subset \Omega | Z' = [a < m_i, m_{i+1}, \dots, m_{i+n} > b] | a \neq b | a, b, i, n \in \mathbb{N}\}. \quad (11)$$

Los movimientos cíclicos son aquellos movimientos que realiza el paciente cuando se dirige de un punto A a un punto B, después a un punto C y por ultimo a un punto D, regresando de vuelta al punto A. Este movimiento representa un giro en la ruta del paciente, de ahí el nombre de movimiento cíclico. Este tipo de movimiento está conformado por al menos cuatro segmentos relevantes, donde cada uno cuanta con más de cuatro mediciones, variando de datos positivos a datos negativos, además se puede identificar que el primer segmento relevante y el ultimo, generalmente son contrarios:

$$MC = \{[Z'_1, Z'_2, \dots, Z'_n] \subset \Omega | Z' = [a < m_i, m_{i+1}, \dots, m_{i+n} > b] | a \neq b | a, b, i, n \in \mathbb{N} | N \geq 4\}. \quad (12)$$

Los movimientos rítmicos son aquellos donde el paciente se mueve de un punto A a un punto B y después a un punto C, repitiendo el mismo movimiento de manera consecutiva hasta llegar al objetivo. Este movimiento caracteriza por la existencia de al menor tres segmentos relevantes. Además, el primer segmento relevante es del mismo signo del último segmento relevante. Otra característica de este movimiento es que cuenta en cada segmento relevante con un valor máximo o un valor mínimo, donde los valores antes y después de este dato relevante, disminuyen o aumentan hasta cambiar de signo. Es decir, si tenemos un máximo, los valores antes y después de este disminuirán hasta cambiar de signo. Lo mismo ocurre con un mínimo; en este caso aumentarán hasta cambiar de signo.

$$MR = \{[Z'_1, Z'_2, \dots, Z'_n] \subset \Omega | Z' = [a < m_i, m_{i+1}, \dots, m_{i+n} > b] | a \neq b | a, b, i, n \in \mathbb{N} | N \geq 3\}. \quad (13)$$

Por último, los movimientos aleatorios son aquellos donde no existe un orden claro en la forma de movimiento del paciente, volviéndose en pocas palabras impredecible. Por lo tanto, cualquier movimiento que no cumpliera con las definiciones anteriores se considera como movimiento aleatorio.

3.5. Módulo analizador de rutas

Este último módulo del componente presentado en este artículo se encarga de recopilar la información de los módulos anteriores y darles un trato de tal manera que el encargado o especialistas pueda interpretar la información correctamente. En este módulo se recopilará la información y se enviará al encargado/especialista, quien podrá verificar los resultados, filtrando la información por día, semana o mes, pudiendo llevar un control del paciente y realizar diagnósticos o tratamientos.

4. Evaluación de la propuesta

Para la evaluación de la propuesta se decidió utilizar un dispositivo similar al utilizado durante las pruebas realizadas por [9]. Dicho dispositivo es un dispositivo móvil, modelo «Samsung Galaxy s7». Al utilizar un dispositivo parecido 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 los movimientos del sujeto. 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».

Para probar la eficacia del sistema se llevaron a cabo una serie de experimentos donde se eligieron rutas al azar, juntando una muestra de 30 movimientos, de los cuales 10 fueron movimientos directos (D), 10 fueron cíclicos (C), 6 fueron rítmicos (R) y 4 fueron aleatorios (A). Los resultados resultados de los experimentos se muestran en el Cuadro 3.

Tabla 3. Resultados de los experimentos.

<i>Movimientos</i>	D (N=10)	A (N=4)	R (N=6)	C (N=10)
<i>Éxitos</i>	9	4	5	8
<i>Fallos</i>	1	0	1	2

Como se puede apreciar, el sistema obtuvo un error del 13.33 % en las mediciones, donde los mayores problemas identificados, son las rutas rítmicas y cíclicas, esto debido a que ambos movimientos se asemejan al momento de aplicar la definición de los segmentos relevantes, pero a pesar de ellos, los resultados obtenidos son satisfactorios, esto debido a que se obtuvo una tasa de error similar o menor a la reportada por otras propuestas, con la ventaja de que la presentada en este artículo no requiere infraestructura previa y funciona tanto en interiores como en exteriores.

5. Conclusiones y trabajos futuros

La demencia es uno de los principales problemas de salud actuales, debido en gran parte al envejecimiento de la población mundial. La demencia no tiene cura, provoca la muerte de quienes la padecen y son la causa de deterioro económico y emocional de familiares. Para dar el trato adecuado a una persona afectada con este padecimiento es necesario llevar a cabo un diagnóstico preciso de la etapa de desarrollo de la enfermedad. Sin embargo, para determinarla es necesario llevar a cabo pruebas costosas y que provocan mucho estrés en los pacientes.

Recientemente se ha detectado que existe una relación entre los patrones de deambulación y el nivel del avance de la demencia. En este artículo se presentó un componente para la detección de patrones de deambulación de pacientes afectados con demencia. El componente propuesto, desarrollado con base en el paradigma del Internet de las Cosas, no es invasivo ni requiere de infraestructura previamente instalada en los edificios en los que será utilizado. La evaluación del sistema propuesta mostraron una alta tasa de detección.

Como trabajo futuro se tiene contemplado la integración de técnicas de aprendizaje autónomo (*machine learning*) en el módulo analizador de rutas.

Referencias

1. Algase, D.L., Antonakos, C.L., Beattie, E., Beel-Bates, C.A., Yao, L.: New parameters for daytime wandering. *Research in Gerontological Nursing* 2(1), 58–68 (jan 2009)
2. Alzheimer's Association: Types of Dementia. Alzheimer's Association (2017), <https://www.alz.org/dementia/types-of-dementia.asp>
3. Alzheimer's Association: Wandering and Getting Lost. Alzheimer's Association (2017), <https://www.alz.org/care/alzheimers-dementia-wandering.asp>
4. Alzheimer's Association: What Is Dementia? Alzheimer's Association (2017), <https://www.alz.org/what-is-dementia.asp>
5. Alzheimer's Association National Office: Alzheimer's Association National Office. Alzheimer's Association National Office (2017), <http://www.alz.org/facts/overview.asp>
6. Atzori, L., Iera, A., Morabito, G.: The internet of things: A survey. *Computer Networks* 54(15), 2787–2805 (2010), <http://www.sciencedirect.com/science/article/pii/S1389128610001568>
7. Bates, R.J., Bates, R.J.B.: GPRS: General Packet Radio Service (Professional Telecom). McGraw-Hill Professional (2001)
8. Budson, A.E., Solomon, P.R.: Memory Loss, Alzheimer's Disease, and Dementia: A Practical Guide for Clinicians. Elsevier (2015)
9. Camberos, M.L.L., Díaz-Ramírez, A., Marquez-Barja, J., Quintero, V.: Algoritmo de trazado de rutas en interiores sin infraestructura. In: *Proceedings of International Conference on Computer Networks Applications*. pp. 64–69. Tecnológico Nacional de México (2017)
10. Chan, M., Bonhomme, S., Estève, D., Campo, E.: Individual movement trajectories in smart homes. In: *IFMBE Proceedings*, pp. 1014–1018. Springer Berlin Heidelberg (2008)

11. Gavrilovska, L., et al.: Application and Multidisciplinary Aspects of Wireless Sensor Networks: Concepts, Integration, and Case Studies (Computer Communications and Networks). Springer (2010)
12. Google: Google Maps. Google (2018), <https://cloud.google.com/maps-platform/>
13. Kaplan, E., Hegarty, C.J.: Understanding GPS/GNSS: Principles and Applications, Third Edition (Gnss Technology and Applications Series). Artech House (2017)
14. Lin, Q., Zhang, D., Huang, X., Ni, H., Zhou, X.: Detecting wandering behavior based on GPS traces for elders with dementia. In: 2012 12th International Conference on Control Automation Robotics Vision (ICARCV). pp. 672–677 (Dec 2012)
15. Martino-Saltzman, D., Blasch, B.B., Morris, R.D., McNeal, L.W.: Travel behavior of nursing home residents perceived as wanderers and nonwanderers. *The Gerontologist* 31(5), 666–672 (oct 1991)
16. Mouly, M., Pautet, M.B.: The GSM System for Mobile Communications. Telecom Publishing (1992)
17. Organizacion Mundial de la Salud: Indice de Muertes por Demencia. Organizacion Mundial de la Salud (2009), <http://www.who.int/gho/en/>
18. Power, G.A.: Dementia beyond Disease: Enhancing Well-Being. Health Professions Press (2016)
19. Rodrigues, A., Silva, J.S., Boavida, F.: iSenior—a support system for elderly citizens. *IEEE Transactions on Emerging Topics in Computing* 1(2), 207–217 (dec 2013)
20. Sato, H.: Moving average filter (US patent US6304133B1) (2001)
21. Sposaro, F., Danielson, J., Tyson, G.: iWander: An Android application for dementia patients. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. IEEE (aug 2010)
22. Vuong, N.K., Goh, S.G.A., Chan, S., Lau, C.T.: A mobile-health application to detect wandering patterns of elderly people in home environment. In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE (jul 2013)

Sistema de clasificación SVM de señales electromiográficas extraídas en un sistema embebido

Luis Daniel Reyes Crusaley¹, J. R. Cárdenas-Valdez¹,
Gilberto Enrico Vázquez¹, Manuel de Jesús García Ortega¹,
Andrés Calvillo Téllez²

¹ Instituto Tecnológico de Tijuana, B.C., México

reyescrusaley@gmail.com, jose.cardenas@tectijuana.edu.mx
manuel.garcia@tectijuana.mx, gilberto.vazquez@tectijuana.edu.mx

² Instituto Politécnico Nacional CITEDI, Tijuana, B.C., México

calvillo@citedi.mx

Resumen. El presente trabajo presenta el diseño de un sistema de adquisición de señales biomédicas electromiográficas inalámbrico, el cual registra las señales musculares 2 en la tarjeta de desarrollo EKG/EMG, las señales se transmiten mediante el protocolo ZigBee en enlace punto a punto o multipunto, por lo que es escalable para más de un paciente en paralelo. La transmisión de los datos se recibe en la tarjeta de desarrollo Raspberry Pi3 la cual realiza una truncación de la señal recibida y es enviada a la nube para un proceso de clasificación. El sistema desarrollado es una propuesta precisa de bajo costo para análisis de diversos pacientes, la técnica propuesta representa la etapa de adquisición, análisis y truncación de datos para un proceso de clasificación de señal basado en SVM con el fin de predecir el mejor tipo de terapia para un paciente dado. Las pruebas experimentales y de simulación desarrolladas en hardware y clasificadas en software a través de SVM demuestran que el sistema completo es capaz de identificar los límites entre dos señales con una precisión mayor al 95%.

Palabras claves: EMG, sistema embebido, SVM, Raspberry Pi 3, ZigBee.

SVM Classification System of Electromyographic Signals Extracted in an Embedded System

Abstract. The present work presents the design of a wireless electromyographic biomedical signal acquisition system, which records the muscle signals in the EKG / EMG development card, the signals are transmitted through the ZigBee protocol in point-to-point or multipoint link, so it is scalable for more than one patient in parallel. The transmission of the data is received in the Raspberry Pi3 development card which truncates the received signal and is sent to the cloud for a classification process. The developed system is a precise proposal of low cost for the analysis of several patients, the proposed technique represents the stage of acquisition, analysis and truncation of data for a signal classification process based on support of vector machines (SVM) with the In order to predict the best type of therapy for a given patient. Experimental and simulation tests developed in hardware and classified in software through SVM show that the complete

system is able to identify the limits between two signals with an accuracy greater than 95%.

Keywords: EMG, embedded system, SVM, Raspberry Pi 3, ZigBee.

1. Introducción

De acuerdo al Instituto Nacional de Estadística y Geografía (INEGI) 2014, el 6.4% de la población del país reportaron tener al menos una discapacidad, de las cuales el 52.1% de esa población se refiere a una discapacidad motriz [1]. El Plan Nacional de Desarrollo (PND) 2013-2018 del Gobierno de la República, establece que la capacidad del Estado Mexicano para atender las necesidades más urgentes de la población (salud, educación, investigación y desarrollo, infraestructura y seguridad) es limitada debido a la alta demanda de servicio [2], por lo que el desarrollo de sistemas de medición y diagnóstico precisos y de bajo costo que puedan ser escalables para toda la población es crucial en el desarrollo de países tercermundistas. En los recientes años se han realizado esfuerzos importantes en trabajos de clasificación de datos relacionado a señales de electromiografía (EMG), se han presentado estudios de clasificación basados en soporte de máquinas vectoriales SVM y sistema de inferencia neuro-difuso adaptivo (ANFIS) para clasificación de señales de tipo EMG sin el desarrollo de una plataforma física [3], y estudios para el análisis de señales EMG con la finalidad de identificar la patología muscular mediante un método de clasificación [4], así como la comparación entre distintos clasificadores tradicionalmente como K vecinos más cercanos (KNN) utilizados para clasificación de señales EMG [5], además de diseño y desarrollo de sensores que a través de hardware tienen el propósito de medir señales de tipo ECC y EMG [6].

El monitoreo y clasificación de datos basado en señales EMG tienen un auge importante en aplicaciones biomédicas, la literatura reporta el uso de señales EMG para el control de manos robóticas [7], para el control de prótesis robóticas artificiales para desarrollar actividades de su vida diaria [8]. Además de sistemas de clasificación aplicados al reconocimiento de patrones, donde se tienen diferentes movimientos. Posteriormente, cada señal se representa al utilizar un cierto conjunto de características, las cuales se utilizan como entrada para un clasificador. Una vez que el clasificador es entrenado, es capaz de discriminar los movimientos previstos [7], los sistemas de clasificación en este tipo de señales permiten un análisis preciso de lo sucedido durante diversos periodos de tiempo por ejemplo después de una terapia dada.

En este contexto y como forma de definir un sistema de clasificación de señales se propone el desarrollo de una técnica de adquisición de datos precisa y de bajo costo en un sistema embebido mediante el acoplamiento de las tarjetas EKG/EMG-Arduino WiFi, que mediante la adquisición por electrodos se ubicados en la parte inicial, media y final del músculo bajo prueba se implementa para uno o múltiples pacientes en conexión multipunto o punto-punto. La clasificación y truncación de datos se desarrolla en la plataforma Raspberry Pi3. Los datos almacenados en la nube se clasifican mediante un proceso de clasificación de señales basado en soporte de máquinas vectoriales (SVM) con el fin de predecir tendencias y cuál de las terapias sobre un paciente dado resulta la más favorable. La motivación de este trabajo radica en que la investigación actual no tiene un compromiso entre el desarrollo de clasificadores

precisos para señales EMG y su implementación en hardware, el enlace entre ambas etapas es escasa, en este con-texto el trabajo es desarrollado incluso para múltiples pacientes en paralelo por conexión multipunto donde envían la información mediante un protocolo poco agresivo como ZigBee.

Los datos se analizan mediante un sistema clasificador SVM para la detección de cambios en amplitud, fase y frecuencia. Su implementación a corto plazo radica en centros de salud públicos y privados en Tijuana, México para la atención de pacientes en paralelo. La organización de este trabajo se enlista a continuación, en la sección 2 se muestra la metodología para el desarrollo del sistema inalámbrico desde la etapa de adquisición de datos, conexión inalámbrica mediante el protocolo ZigBee y envío de datos vía internet en la tarjeta de desarrollo Raspberry Pi3, en la sección 3 se describen los resultados obtenidos de la adquisición de señales y clasificación de datos. Finalmente, en la sección 4 se muestran las conclusiones y trabajos a futuro.

2. Metodología

Este trabajo es una propuesta de adquisición de señales EMG, envío de datos inalámbricos para enlaces punto-punto y punto-multipunto, truncación y envío a la nube mediante la tarjeta de desarrollo Raspberry Pi3 a través de programación en lenguaje Python. En la Figura 1 se muestra el diagrama del flujo del sistema de adquisición de datos basado en la tarjeta EKG/EMG dicha tarjeta permite la adquisición a través de tres electrodos de una señal EMG, en este caso los tres electrodos se colocan en el músculo deseado en la parte inicial y final, así como en una aproximación intermedia del músculo, en este caso la prueba se realiza sobre el músculo flexor común profundo de los dedos. La señal sensada es discretizada con una frecuencia de muestreo de 256 Hz dado que la mayoría de las señales musculares se ubican por debajo de los 50 Hz y es convertida en una secuencia de bits que serialmente se transmiten través del protocolo ZigBee.

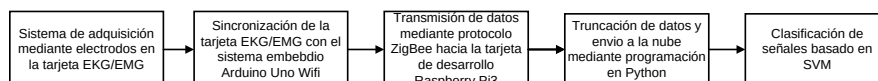


Fig. 1. Diagrama de flujo del sistema de adquisición de datos y transmisión inalámbrica.



Fig. 2. Montaje de la tarjeta de desarrollo Arduino Uno Wifi y tarjeta shield EKG/EMG.

La primera etapa del sistema consiste en la adquisición de las señales biomédicas EMG en particular del brazo, para esto se hace uso de tres electrodos, cables para electrodos, la tarjeta shield EMG/EKG y el sistema embebido Arduino Uno Wifi. Dada

la compatibilidad de ambas tarjetas, se ensamblar la tarjeta shield EKG/EMG con la tarjeta Arduino Uno Wifi como se muestra en la Figura 2, así como los cables de conexión para los electrodos, la implementación de este sistema se basa en el musculo flexor común profundo de los dedos.

Cabe indicar que el sistema basado en la tarjeta EKG/EMG requiere una colocación de los electrodos en la parte inicial, final y en una aproximación intermedia del músculo bajo análisis. El músculo a analizar tiene la función de flexionar los dedos de la mano, específicamente los últimos cuatro dedos.

Para llevar a cabo la transmisión se hizo uso de dos módulos Xbee S1, un shield Xbee para arduino y un Xbee Explorer Dongle USB. La siguiente etapa se refiere a la configuración del Xbee Explorer Dongle de cada módulo basado en la interface del programa X-CTU, este proceso es indispensable para el enlace punto-punto o punto-multipunto para el monitoreo de uno o dos pacientes en paralelo basado en el mismo sistema de adquisición. El sistema a desarrollar es escalable y las primeras aplicaciones consisten en enlazar dos sistemas de adquisición y transmitirlos mediante ZigBee a la tarjeta Raspberri Pi3 para el proceso de truncación de datos y envío de datos a la nube y su posterior etapa de clasificación basado en SVM.

En este caso se realiza una conexión tipo AT entre los dos módulos, es decir, tanto el modulo A como el B intercambiar datos. Para verificar esto se ensambla la tarjeta Arduino transmisora con el shield Xbee y cualquiera de los dos módulos Xbee S1 y se conecta el segundo módulo Xbee S1 al sistema de adquisición. En la Figura 3 se muestra el enlace entre las tarjetas EKG/EMG, Arduino UNO Wifi y el módulo Xbee, así como la transmisión de datos.

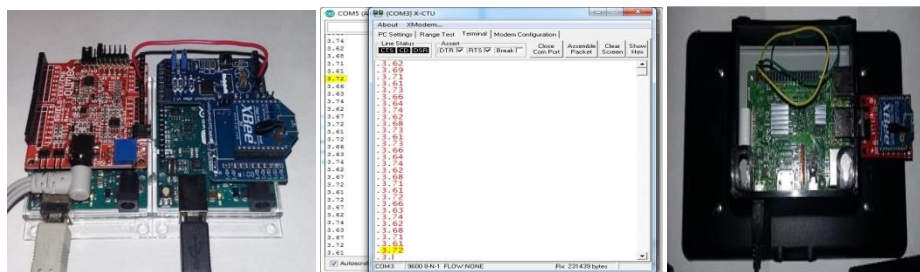


Fig. 3. Enlace de la tarjeta EMG-Arduino y envío datos serialmente mediante ZigBee.

La tarjeta de desarrollo realiza el proceso de truncación de datos y almacenamiento en la nube para una etapa posterior de clasificación basada en SVM. El envío de la señal se realiza mediante programación en Python para almacenamiento de la señal biomédica.

2.1. Clasificador basado en SVM

El sistema de clasificación basado en SVM para sistemas de datos de dos fuentes diferentes se basa en encontrar un hiperplano que separa los puntos de una clase de otra. El mejor de los hiperplanos significa que existe un margen de separación de datos del máximo valor entre ellos. Los datos para el entrenamiento del sistema se establecen

como vectores x_j con sus categorías y_j . Para una dimensión de-finida d , se tiene que $x_j \in \mathbb{R}^d$, además que $y_j = \pm 1$. La ecuación que define el hiperplano es:

$$f(x) = x \cdot \beta + b = 0, \quad (1)$$

donde $\beta \in \mathbb{R}^d$ y b es un número real. La mejor separación se define por el hiperplano si se encuentra β y b que minimiza $\|\beta\|$ tal que todos los puntos del sistema cumplan con:

$$y_j f(X_j) \geq 1. \quad (2)$$

El problema es usualmente dado por el sistema de minimizar $\|\beta\|$. La solución óptima $(\hat{\beta}, \hat{b})$ que habilita la clasificación de un vector z definido como:

$$\text{clas}(z) = \text{sign}(z \cdot \hat{\beta} + \hat{b}) = \text{sign}(\hat{f}(z)), \quad (3)$$

donde $\hat{f}(z)$ es el sistema de clasificación que representa la distancia en la frontera de decisión del sistema [9], en este caso de dos señales EMG bajo análisis.

3. Resultados

En este trabajo se desarrolló satisfactoriamente un sistema de adquisición de señales biomédicas inalámbrico basado en la tarjeta de adquisición EKG/EMG y su transmisión inalámbrica mediante enlace punto-punto y punto-multipunto mediante el protocolo ZigBee para el monitoreo de la actividad eléctrica muscular. En la Figura 4 se muestra la señal extraída por el sistema de adquisición, cuya información es enviada a la tarjeta de desarrollo Raspberry Pi3. En este caso la señal se encuentra almacenada en la nube y se realiza un proceso de clasificación de señal para detectar tendencias, el clasificador basado en SVM detecta dos grupos con tendencia al cambio de amplitud como se muestra en la Figura 5, el sistema es flexible para clasificar cambios de amplitud, frecuencia y fase de dos señales definidas si el médico experto desea hacer alguna otra clasificación como forma de predicción del éxito en terapias musculares, en este caso las pruebas se realizaron sobre el musculo flexor común profundo de los dedos, en la Figura 5b se observa que el clasificador basado en programación en lenguaje C establece los límites de ambas señales.

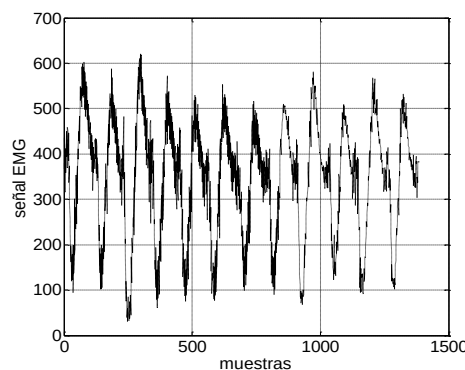


Fig. 4. Señal EMG obtenida con el sistema embebido EKG/EMG enviado vía ZigBee.

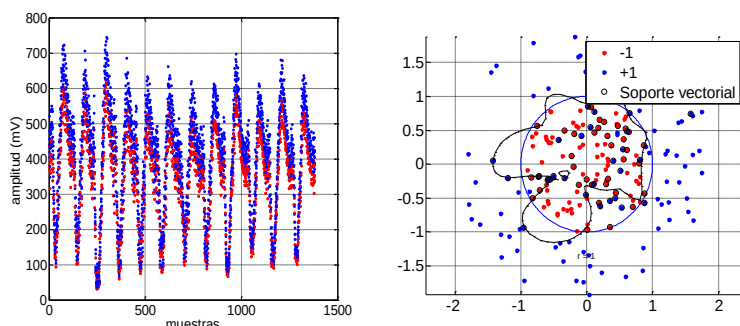


Fig. 5. Comparación de dos señales e implementación del sistema de clasificación SVM.

4. Conclusiones

En este trabajo se desarrolló un sistema de adquisición de señales EMG y una etapa de clasificación basada en SVM, el sistema desarrollado se enlaza como punto-punto o multipunto si más de un paciente debe ser evaluado mediante la tarjeta de adquisición EKG/EMG+Arduino WiFi. En la primera etapa los datos son discretizados y enviados como secuencia de bits mediante el protocolo Zig-Bee. Los datos son recibidos en la tarjeta de desarrollo Raspberry Pi3 y son trun-cados y enviados a la nube para un proceso de clasificación de datos basado en SVM. Los resultados se evaluaron sobre dos muestras de un paciente donde a ser demasiado similares el sistema reconoce particularidades de los rastros, en este caso cambios de amplitud. El trabajo desarrollado permite la adquisición de cualquier señal muscular en donde el proceso final de clasificación SVM define el nivel de mejoría del paciente dada una terapia física aplicada. La capacidad de sensor y clasificar señales del sistema desarrollado es una herramienta de diagnóstico para médicos relacionados a la aplicación de terapias físicas e incluso les permite la atención de más de un paciente a la vez.

Referencias

1. SEGOB: Diagnóstico sobre la situación de las personas con discapacidad en México. CDMX, México (2016)
2. Plan Nacional de Desarrollo 2013-2018 (2013)
3. Caesarendra, W., Tjahjowidodo, T., Pamungkas, D.: EMG based classification of hand gestures using PCA and ANFIS. In: International Conference on Robotics, Biomimetics, and Intelligent Computational Systems (Robionetics), Bali, Indonesia (2017)
4. Sadi-Ahmed, N., Messekher, A., Namane, S., Kadir-Talha, M., Mekaoui, S.: Automatic diagnosis of neuromuscular diseases from electromyographic (EMG) records. In: 5th International Conference on Electrical Engineering - Boumerdes (ICEE-B), Boumerdes, Algeria (2017)
5. Paul, Y., Goyal, V., Avtar Jaswal, R.: Comparative analysis between SVM & KNN classifier for EMG signal classification on elementary time domain features. In: 4th International Conference on Signal Processing, Computing and Control (ISPCC), Solan, India (2017)
6. Jani, A. B., Bagree, R., Roy, A. K.: Design of a low-power, low-cost ECG & EMG sensor for wearable biometric and medical application. IEEE Sensors, Glasgow, UK (2017)

7. Wang, J., Ren, H., Chen, W., Zhang, P.: A portable artificial robotic hand controlled by EMG signal using ANN classifier. In: IEEE International Conference on Information and Automation, Lijiang, pp. 2709–2714 (2015)
8. Zia ur Rehman, M., Gilani, S. O., Waris, A., Niazi, I. K., Kamavuako, E. N.: A novel approach for classification of hand movements using surface EMG signals. In: IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 265–269, Bilbao, Spain (2017)
9. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Second edition, New York: Springer (2008)

Telemedicine Networks: Integrating Different Health Technologies towards a Common Goal

Roberto Conte-Galvan¹, Arturo Serrano-Santoyo¹, Veronica Rojas-Mendizabal²

¹ CICESE Research Center, Ensenada, B.C., Mexico

² CETYS University, Mexicali, B.C., Mexico

conte@cicese.mx

Abstract. People like to live, work, travel and vacation in all kinds of places and weathers. Our active and inquisitive nature takes us to the most extreme locations, and our gregarious behavior allows us to coexist in both isolated as well as overcrowded communities. In all these scenarios, the new paradigm of healthcare becomes significant. The objective of this paper is to highlight the essential elements of both the current and emerging technologies that may constitute a ubiquitous, universal network of healthcare networks, providing an efficient and effective health service distributed globally. Our aim is to describe the current health ecosystem in order to identify the key agents and basic interactions that may contribute to conform a conceptual framework to provide ubiquitous health services using technology as an enabler but not the central objective. We propose an integral scheme that includes technology applications, regulation strategies and collaborative networks in order to provide efficient healthcare services anytime, anywhere, to anyone.

Keywords: telemedicine networks, wireless healthcare, healthcare coverage, eHealth, interoperability.

1 Introduction

Wherever we might be located at any moment, we carry our health and our illnesses along to all those places, and in case any accident or health issue may appear we expect to be efficiently attended and taken care of. Although many digital healthcare devices and services devices are currently available in several regions of the world, and Telemedicine technology helps a remote patient to be reasonably diagnosed, treated and supervised away from qualified physicians and hospitals, our mobile and nomadic nature requires ubiquitous healthcare attention, creating a very complex health challenge to the diverse healthcare systems around the world.

Commonly diagnosed and treated pathologies in some parts of the world may be totally unknown in other traveled places, often placing competent healthcare personnel in challenging situations. Another, often overlooked, problem arises with the increasing growth of natural origin disasters, such as earthquakes, hurricanes, floods or fires, with the capacity to damage or destroy health and medical infrastructure, as well as the electronic medical data of patients not stored in safe,

virtual environments: Are privately-owned, hospital-based, remotely-based or cloud-based health information storage systems safe? Most already-stored digital health information definitely is, but it remains to be seen if this stored information includes the totality (100%) of the health data (health & medical history; text, imaging & audio files; as well as laboratory studies), from all individuals (patients and their families), included into a single, openly compatible, Electronic Medical/Health Records (EMR/EHR), for a complete community, region or country. Even if this already secure health information is safe and available, it remains to be seen if it is also accessible from hospitals, clinics, ambulances, private practice medical or healthcare units, homes and nursing facilities for homecare assistance when necessary, as well as many other places that may require access to such medical and health digital information.

Universal health care and medical attention for the general population are not only considered a basic human right in many countries around the world, but also a reachable goal in many national administrations. Since this is a very complex problem that depends on many factors such as each country's health system, population, climate and territorial size, as well as its general work, economy and educational situation, among others, it requires a very careful analysis that considers the social, regulatory, medical and technological infrastructure available in that country or region.

The main purpose of any health system is to provide efficient health and medical service to a large number of individuals and, often, to their immediate families or relatives, through a network of health practitioners and professionals (physicians, nurses, health administrators, social workers, etc.). These health professionals work in efficient and qualified facilities, with the help of an increasing number of electronic, computing and communications devices. In order to integrate all elements of current and legacy healthcare systems into a digital globalized healthcare network. First, we must describe and integrate all existing and proposed elements of this framework, considering the current and future interoperability of all medical, technical, administrative and socio-economic segments of such complexity.

The aim of this paper is on the one hand to propose a general framework for the provision of universal digital health services. On the other hand, to identify the elements of the health ecosystem emphasizing the relevance of the interactions among their actors in order to support the design and deployment of medical quality services. The context will determine the nature and scope of the digital interventions. In this case, the understanding of the social reality of this context becomes crucial for the success of such interventions.

2 An Evolving Healthcare Framework

The basic proposal for a Healthcare Framework to provide a ubiquitous, global healthcare concept should include the service components shown in Figure 1. One way to address this problem, geographically speaking, is through the use of a diverse combination of networking technologies, considering mainly mobile and wireless

networks applied to healthcare purposes, which will interact with fixed networking technology already positioned in existing, legacy healthcare facilities. This technology combination needs to be fully compatible at the healthcare application level, but it must be also compatible at the lower layers whenever possible, with a maximized interoperability in mind and always considering access to the ubiquitous user or global traveler as mentioned by Brooks [1].

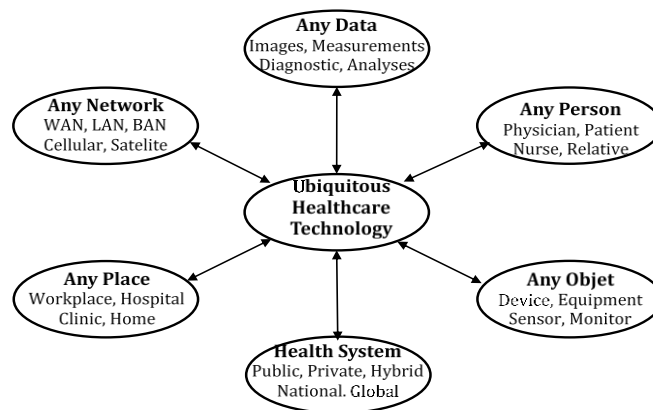


Fig. 1. Ubiquitous Healthcare Service Components

Our aim is to propose a preliminary framework conducive to develop specific strategies and lines of action towards the deployment of universal healthcare services. We expect this initial idea to be defined more precisely with participation of interdisciplinary cadres of scientists, physicians, technologists and healthcare practitioners who might also benefit by developing and creating new, or adapting existing or new technologies.

The prevailing health system in any country is usually a good indicator of that country's development and expected quality of life. This is especially important in rural and challenged settings for large or developing countries, but it is also found in urban settings at developed countries, where the existence and availability of qualified health personnel and facilities may be good, but its geographical distribution may be limited or even critical at times. In fact, Dussault and Franceschini mention in [2] that "A well-balanced distribution of infrastructures needs to go hand-in-hand with a well-balanced distribution of health personnel to be worth the investment, let alone to have an impact on the health of the population." They also report the presence of unemployed, underemployed or inactive medical and nursing personnel in urban areas, while having scarce and unfilled positions in rural facilities at the same time, due to several economic, social, cultural and professional reasons. They mention as justifying issues for this disparity such arguments as inadequate infrastructure, little physician's leisure time and other family-oriented aspects such as children's education and social activities. They even mention that "rural-to-urban physician displacement is usually compounded by public-to-private sector healthcare employment displacement" as it can be seen happening in most countries around the

world, regardless of the size and strength of their economies. Therefore, the uneven geographic distribution of qualified healthcare personnel is a common problem for small and large countries, as well as developed and developing regions around the world, therefore creating a complex dilemma not easy to solve.

The Organization of Economic Cooperation and Development (OECD) recommends a physician density of 2.3 health workers as average (2.2 physicians and 2.6 nurses) per 1000 inhabitants, while the World Health Organization (WHO) recommends a physician density of 3.4 skilled health worker per 1000 [3], and a minimum of 2.3 [4]. These figures are somewhat estimated since there are different criteria regarding professional and vocational health education and training around the world, as well as the regulations, activities, and tasks required for each health worker in order to work in different countries.

Even in highly developed countries, as well as in regions inside developing countries, there can be a similar kind of geographic irregularities regarding medical personnel distribution. There may be certain geographic urban areas in developed countries where the average local distribution of physicians may be much greater than the OECD or WHO recommendations, but also exist other areas in the same country where the physician distribution may be much smaller than the suggested average. The physician density figures published by the health authorities in each country are always an average and not always updated. Therefore, these irregularities always show up in a more detailed analysis [5].

Under this condition, the availability of telecommunications infrastructure with acceptable quality between both ends of the telemedicine link becomes significant. The existence of electric, electronic and computational equipment, as well as qualified personnel able to acquire, transmit, process and safeguard the digital medical information is also crucial. In the same token, we have identified the need to create the necessary didactic material, the desired electronic and printed books, manuals and references, as well as classroom and online presentations for capacity building.

The performance of telemedicine networks depends on the availability of efficient telemedicine systems whose objective is the delivery of digital healthcare services over distance in the form of medical data by healthcare professionals. In order to do this, it is needed to use Information and Communications Technologies (ICT) between the healthcare professional attending the remote patient and the distant healthcare professional at the medical or healthcare facility, over a communications link that works as part of a Telemedicine Network.

A very sensitive aspect on the handling of a patient's electronic medical data (EMR/EHR) is the threat of a cyberbreach in the acquisition, processing, storage or transmission and interconnection stages of a quality telemedicine service. The strong and secure regulation and enforcement of privacy and security laws and regulations is needed in order to provide security measures on health and medical networks, nodes, devices, equipment, personnel and services. With the increasing coverage of health services through the use of electronic and wireless technology, it is imperative that such networks provide availability, confidentiality and integrity on the provision of health and medical information to healthcare professionals and patients, while providing at the same time strong protection to that private personal health and

medical data from possible security breaches. Telemedicine security issues include authorization, authentication and accounting, common with banking and manufacturing applications among other industries, as mentioned in Garg and Brewer [6]. On the same subject, Olanrewaju et al. [7] emphasize the way protection and authentication of medical information and images matter in a telemedicine environment over electronic networks. Protection from undesired access, alteration and/or deletion of confidential health and medical information by intruders, hackers or plain criminals require the identification of loopholes and the implementation of privacy and security measures in order to identify, counter and defeat threats on security and authentication of medical records in telemedicine networks. Both papers [6] and [7] describe the main types of threats and attacks on telemedicine data such as electronic files and images, as well as eavesdropping and monitoring the continuous stream of data in order to illegally collect private health information. Finally, a strong, wide-ranging, standardized regulatory framework with international reach is urgently needed for all wireless, mobile and satellite telemedicine networks, devices and applications.

From a systemic point of view, it is very important to define which device, service or application is found where in the telemedicine system, and how they are interacting with each other, so there will be a fluid and fluent transmission of information within the network. A basic telemedicine system has six main stages, as shown in Figure 2, where each stage may vary among many different options in its own way, depending upon its geographical, infrastructure or medical need.

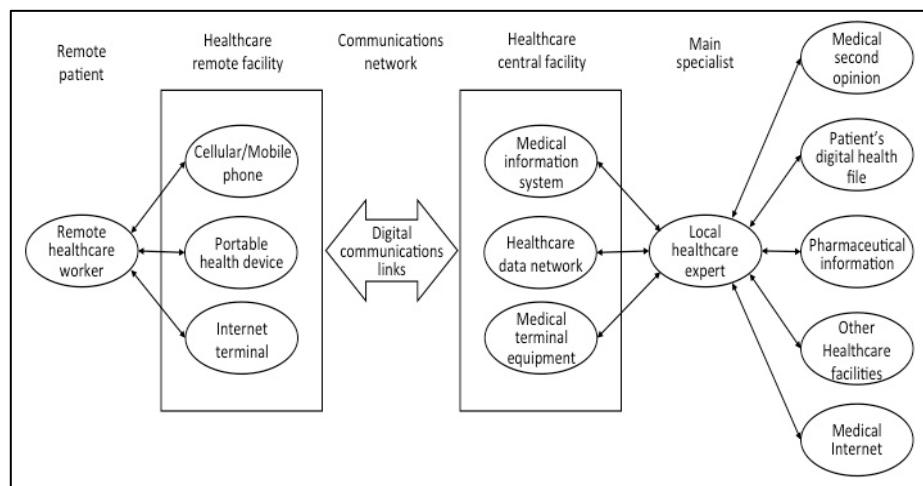


Fig. 2. Basic Elements of a Telemedicine System

The various options of each stage shown in Fig. 2 perform a specific function according to the related healthcare option, as follows:

1. Remote patient/healthcare worker: It is found at one end of the telemedicine service, usually where the remote patient and/or remote healthcare worker is found.

It will provide the patient's physiological health information to the healthcare worker, or to whatever digital device, terminal or communications system is to be used in the telemedicine service for each specific application or service related to a determined illness, health or medical situation.

2. Healthcare remote facility: This facility may be anything from a small measuring health device (electronic stethoscope, thermometer, microphone, image or video camera, computer or videoconferencing terminal) at the hands of a remote health professional, sent from a house, village, rural or small suburban clinic, up to a complex electronic or biomedical equipment that interacts with the patient.
3. Communications network: This is probably the most widely known part of any telemedicine or telehealth system since it is a legacy from the old analog telecommunication systems, now migrated to new digital, Internet-compatible technologies, such as wired, fiber-optical, mobile, wireless and satellite communications systems, ranging from a few meters up to thousands of kilometers apart.
4. Healthcare central facility: Once the remote patient's health data is effectively sent through a secure communications link, there usually is a main healthcare facility at the other end with the corresponding health and medical electrical interfaces to receive the incoming information the most similar possible to the one provided by the remote device or system at the initial end of the link.
5. Main healthcare specialist: This may be a medical expert, or group of experts, who will be attending the remote patient in a number of sessions, monitoring and/or evaluating the medical and health information remotely provided in order to prepare any kind of medical treatment and follow-up indications that may be deemed proper. Depending upon the type, content, and urgency of the medical information provided to this professional, a diagnosis and treatment will be programmed and executed, with the patient's cure, recovery, and well-being as the final goal in mind.
6. Telemedicine or Healthcare application: This final part of the telemedicine system is where a health or medical service or application may be provided through a simple file report (writing and filing electronic medical records), to an average (scheduled treatments, follow-up appointments, teleconsulting sessions, laboratory studies, medical/health files updates), or complex (access to other large medical teams, resources and institutions) process, including related information such as the indicated medicines and dosage, or any other clinical, pharmaceutical, laboratory analysis and report, as well as rehabilitation indications.

Therefore, a typical telemedicine system can be seen as a technical set of common, interoperable and compatible components, interfaces, protocols and systems, both in hardware, software and firmware form, and it must be clearly defined as a global, efficient, low-cost, high-quality e-Health system to be properly deployed, through the use of current and future telemedicine systems and networks.

Since each country has a different medical, technical, regulatory, economic and health system, these telemedicine technologies must be part of an interoperable, versatile, modular, and scalable technology oriented towards the medical sector,

similar to what the Open Systems Interconnection (OSI) model did for the development of the Internet during the 1980s and 90s.

3 Telemedicine Networks

Telemedicine systems usually operate inside the coverage area limits defined by their own telecommunications technology as indicated by Wager et al. [8]. When more than one communications link is used to interconnect several devices exchanging medical data using any communications channel, a Telemedicine Network is created. There are a substantial number of Telemedicine applications and networks operating in the world, with different applications and coverage areas using all kind of channels depending upon the technology used.

Sandhya Rani et al. mention a number of healthcare wireless technologies which use not only Horizontal Hand-Off (HHO) mechanisms to access mobile nodes with similar networks, but also describe Vertical Hand-Off (VHO) mechanisms to allow mobile node access to other heterogeneous access point connections, and describe several cases already considering Quality of Service (QoS) provisions for Telemedicine applications [9].

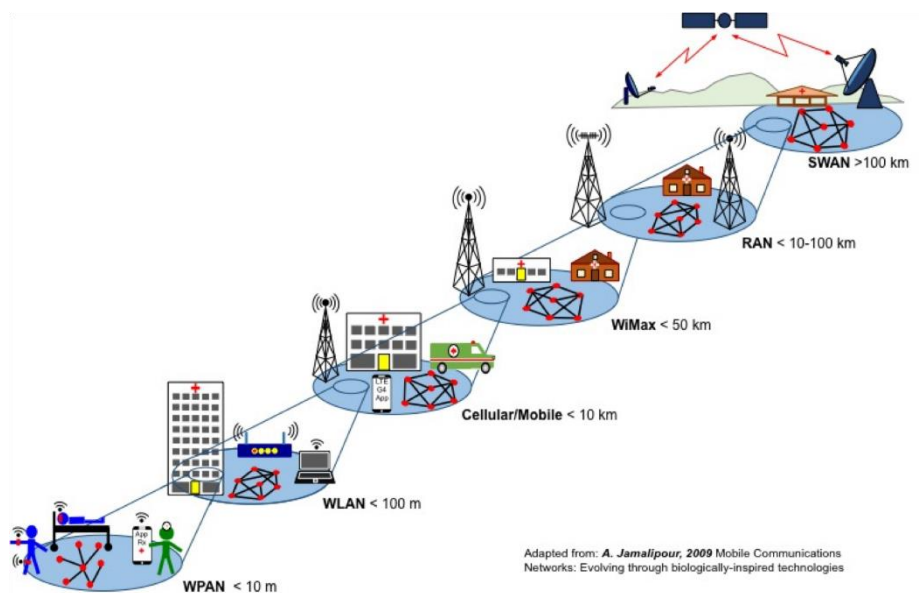


Fig. 3. - Telemedicine Networks, range and area coverage.

Wireless technologies cover all kinds of medical applications, from Bluetooth-based wearable ECG devices as part of a Body Area Network (BAN), to GSM, 3G or LTE mobile teleconsultation and ambulance-based bio-signals transmission, to Wi-Fi-based homecare monitoring and Wireless Local Area Networks (WLAN) clinic or

hospital-based supervision and management, going all the way to Satellite-based earth stations to link far away, remote health facilities for Computer Tomography (CT) scanning and other medical imaging transmissions, as part of a Satellite Wide Area Network (SWAN), as shown in Figure 3.

Keikhosrokiani et al. [10] make a very thorough and concise description of the different wireless technologies and standards used in healthcare communications networks, covering practically all current technologies and representative healthcare and medical applications for wireless, mobile and satellite systems. Next there is a brief description of the main technologies and standards currently used in wireless telemedicine networks.

- Wireless Personal Area Networks (WPAN) are a family of networks designed around the use of the IEEE 802.15 Bluetooth standard in its varied forms: the IEEE 802.15.1 (Bluetooth Low Rate), IEEE 802.15.3 (Bluetooth High Rate) and Zigbee (IEEE 802.15.4 Low Rate) standards for wireless connectivity with fixed, portable and moving devices within a personal space (< 10 meter range). Casillas et al. [11] describe several Wireless Sensor Network (WSN) architectures over WPAN, and WPAN for healthcare applications, based on different available as well as proprietary technologies for mHealth mobile patients with a number of bio-physiological variables in remote monitoring conditions, including a traffic capacity analysis.

- Wireless Local Area Networks (WLAN) are a family of networks designed around the use of the IEEE 802.11 standard in different PHY and MAC layers, initially developed for home and office applications such as Virtual Private Networks (VPN) and Voice over IP (VoIP). Its Wi-Fi version is probably the most used wireless networking technology in use today for broadband applications in homecare and hospital wireless networks for medical and administrative personnel, as well as visiting relatives and patient use.

- Cellular Mobile or Mobile Network is a mobile communications network where one end of the communications link is a mobile radio device inside a vehicle or carried by a person, while the other end of the link may be another such device or a fixed network connection to a public telephone or data network. Although initially developed and implemented for analog telephone service, since the 1990s it evolved to an all-digital telephone network first, and to an IP-compatible data network now. As mentioned above, [10] does an extensive description of several mobile technologies involved in healthcare.

- Worldwide Interoperability for Microwave Access (WiMAX) is an IP-based technology based on the IEEE 802.16 set of standards for wireless communications developed in the 2000s for the delivery of last-mile wireless broadband access as an alternative to cable and Digital subscriber line (DSL) over long distances, as mentioned in [12], and more recently with an interesting telemedicine case study in Ghana, described in [13].

- IEEE 802.22 Wireless Regional Area Network (RAN) is a standard for wireless communications networks aimed at the use of cognitive radio technology to provide wireless broadband access in rural areas between 10 to 60 km, and possibly up to 100 km in good weather conditions, from a Base Station (BS) to the Customer Premises

Equipment (CPE). Magana-Rodriguez et al. [14] describe a comparison between WRAN and WiMAX technologies for telemedicine applications in Mexico.

- Satellite Wide Area Network (SWAN) is a very large area data network, bounded by the limits on the coverage area, where it depends upon its antenna gain and directivity features, as well as its orbit around the Earth. It is usually based on the use of IP-compatible Very Small Aperture Terminal (VSAT) technology between the remote VSAT terminals and a central, large earth station (Hub) connected to the medical Internet. Kocian et al. [15] described the network parameters and user requirements of a hybrid satellite/terrestrial network for a European health project oriented towards the interconnection via satellite of a large number of remote and rural hospitals and health centers in several European countries in order to provide satellite telemedicine service. In another paper Saravanan describes how India, which currently has a large telemedicine program in service [16], provides satellite telemedicine services onboard an ambulance bus, a train and a rain flood water boat.

As it has been discussed, the seamless operation of a Health Network depends on technology factors as well as human and socio-economic aspects. In this respect, Rojas-Mendizabal et al. has proposed an e-Health ecosystem that integrates these key elements [17]. The different elements of the e-Health ecosystem, involving human, technology and socio-economic aspects, are described in Figure 4.

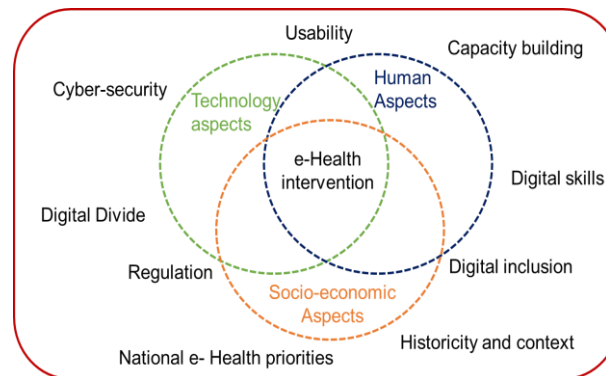


Fig. 4. Elements of an e-Health ecosystem

Rojas-Mendizabal explains in [18] the characteristics of the agents and their attributes in such ecosystem. The human aspects involve the participation of both the patient and the medical specialist. We take into account the digital skills regarding the understanding and adoption of the equipment on the part of the end-users.

The technology aspects consider the medical devices and their operational software and interfaces, as well as the performance of the telecommunications networks in place. The socio-economic aspects of this ecosystem include regulatory concerns, national health priorities, and the cultural and social reality of the particular context, among others. The insights of an in-situ external observer contributed to adjust the weights of the variables employed in the described simulation.

4 Conclusions

Emergent technologies such as Artificial Intelligence, Data and Analytics, as well as, Virtualization and the Internet of Things (IoT), to mention the most relevant, will play a fundamental role in the future provision of medical services. However, the operation of the Health Network is an essential factor to deploy these new applications. The new scenario of advanced applications such as personalized medicine and genomics will require robust, reliable and secure telecommunications channels. One important concern is related to the flow of personal data in Health Networks. Aspects of privacy, transparency, and exploitation of these data have called the attention of governments, academic and social civic organizations. The challenge of establishing adaptive and coherent regulation scenarios as well as strategies for ethical and moral development will play a fundamental role in the conformation of the future Health Networks. Technology is an enabler, but the center of attention must be the individual, the communities, and the institutions.

With a growing and increasingly nomadic population everywhere, ubiquitous health and medical attention with global reach is needed everywhere. The highly irregular geographical distribution of physicians, nurses and other healthcare professionals in both developed and developing countries, as well as the many different approaches to the development of public and private health systems, have created a plethora of new health and medical technology, not always compatible. Current and future wireless telemedicine networks are rapidly evolving, and current health technology allows addressing this problem in many cases, but differences in technical, operative, administrative and regulatory aspects among the different health technology systems and networks around the world difficult the adoption of common, interoperable standards, solutions and applications.

Emergent technologies such as Artificial Intelligence, Data and Analytics, as well as Virtualization and the Internet of Things (IoT), to mention the most relevant, will play a fundamental role in the future provision of medical services. Although these technologies will have a highly disruptive and positive presence in everyday life, their common denominator is the presence of current and future wireless networks everywhere, healthcare and medical networks included. Therefore, the efficient operation of a common Health Network with global reach is an essential factor to deploy these new applications. The new scenarios for advanced applications such as Personalized medicine and Genomics, among others, will require robust, reliable and secure communications channels. One important concern is related to the flow of personal data over current health and medical wireless networks, with weak or non-existent privacy protection regulation.

Aspects of privacy, transparency, and unauthorized access of these data have called the attention of governments, academic and social civic organizations, but a strong regulatory and standardized framework is urgently needed. The challenge of establishing adaptive and coherent regulation scenarios as well as strategies for ethical and moral development will play a fundamental role in the conformation of the future Health Networks. Technology is an enabler, but the center of attention must be the individual, the communities, and the institutions. As a final general framework for

this ubiquitous eHealth ecosystem, we propose the common development of an integral scheme that includes global wireless networking technology; efficient health and medical applications; specialized medical, nursing, health, administrative and engineering human resources; common, secure and compatible health technology regulation and standardization; and the development of collaborative and interoperable applications, devices, equipment, systems and networks in the health technology sector. If the above framework is reached and its components implemented, telemedicine networks and health technology will provide effective and efficient, ubiquitous health services to everyone, anywhere, anytime.

References

1. Brooks, P.: Standards and Interoperability in Healthcare Information Systems: Current Status, Problems, and Research Issues. In: Proceedings of the Fifth Midwest Association for Information Systems Conference, Moorhead, MN, May 21-22, (MWAIS) (2010)
2. Dussault, G., Franceschini, M. C.: Not enough there, too many here: understanding geographical imbalances in the distribution of the health workforce. *Hum Resour Health* 4(12), doi: 10.1186/1478-4491-4-12 (2006)
3. International Labour Organization: World social security report 2010-2011 Providing coverage in times of crisis and beyond. ILO Publication, Geneva (2011)
4. World Health Organization: WHO, World Health Statistics 2014. WHO Press, Geneva (2014)
5. Organisation for Economic Co-operation and Development: Health Care Resources. Retrieved from https://stats.oecd.org/Index.aspx?DataSetCode=HEALTH_REAC (2018)
6. Garg, V., Brewer, J.: Telemedicine Security: A Systematic Review. *J Diabetes Sci Technol*. 5(3): 768–777. doi: 10.1177/193229681100500331 (2011)
7. Olanrewaju1 R.F., Ali, N.B., Khalifa, O., Abd Manaf, A.: ICT in Telemedicine: Conquering Privacy and Security Issues in Health Care Services. *Electronic Journal of Computer Science and Information Technology (eJCSIT)*, Vol. 4, No. 1 (2013)
8. Wager, K. A., Wickham, L. F., Glaser J. P.: Health care information systems: A Practical Approach for Health Care Management. Second Edition, John Wiley & Sons (2009)
9. Sandhya Rani, B. K., Bhat, S., Mukhopadhyay, A.: A Survey of Wireless Technologies and Vertical Handoff Techniques from the Perspective of Telemedicine Scenarios. In: International Conference on Communication and Signal Processing (ICCSP), pp. 1246–1251, doi: 10.1109/ICCSP.2017.8286580 (2017)
10. Keikhosrokiani, P., Zakaria, N., Mustaffa, N., Wan, T. C., Sarwar, M., Azimi, K.: Wireless Networks in Mobile Healthcare. Chapter 30, Springer, doi: 10.1007/978-3-319-12817-7_30 (2015)
11. Casillas M., Villarreal-Reyes, S., Gonzalez, A. L., Martinez, E., Perez-Ramos, A.: Design Guidelines for Wireless Sensor Network Architectures in mHealth Mobile Patient Monitoring Scenarios. In: Mobile Health: A Technology Road Map, pp. 401–428, Springer (2015)
12. Chorbev, I., Mihajlov, M.: Building a Wireless Telemedicine Network within a Wi-Max based Networking Infrastructure. IEEE International Workshop on Multimedia Signal Processing, MMSP '09, Rio de Janeiro, Brazil (2009)
13. Tchao, E. T., Diawuo, K., Ofosu, W. K.: Mobile Telemedicine Implementation with WiMAX Technology: A Case Study of Ghana. *J Med Syst*. 41(1):17 (2017)

14. Magana-Rodriguez, R., Villarreal-Reyes, S., Galaviz-Mosqueda, A., Rivera-Rodriguez, R., Conte-Galvan, R.: Telemedicine Services over Rural Broadband Wireless Access Technologies: IEEE 802.22/WRAN and IEEE 802.16 WiMAX. In: Adibi S. (ed) Mobile Health. Springer Series in Bio-/Neuroinformatics, vol 5. Springer (2015)
15. Kocian, A, De Sanctis, M., Rossi, T., Ruggieri, M., Del Re, E., Jayousi, S., Ronga, L. S., Suffritti, R.: Hybrid Satellite/Terrestrial Telemedicine Services: Network Requirements and Architecture. IEEE Aerospace Conference Proceedings, doi: 10.1109/AERO.2011.5747335 (2011)
16. Saravanan, S, Sudhakar, P.: Telemedicine communication system in Mobile units. In: International Conference on Computer Communication and Informatics (ICCCI -2017), Coimbatore, India (2017)
17. Rojas-Mendizabal, V. A., Serrano-Santoyo, A., Conte-Galvan, R., Gomez-Gonzalez, A.: Toward a Model for Quality of Experience and Quality of Service in e-health Ecosystems. Procedia Technology, 9, pp. 968–974 (2013)
18. Rojas-Mendizabal, V., Serrano-Santoyo, A., Conte-Galvan, R., Villarreal-Reyes, S., Rivera-Rodriguez, R.: Estimation of quality of experience (QoE) in e-Health ecosystems. Ingeniería e Investigación 37(2): 52–59, doi: 10.15446/ing.investig.v37n2.57664 (2017)

Electronic edition
Available online: <http://www.rcs.cic.ipn.mx>

