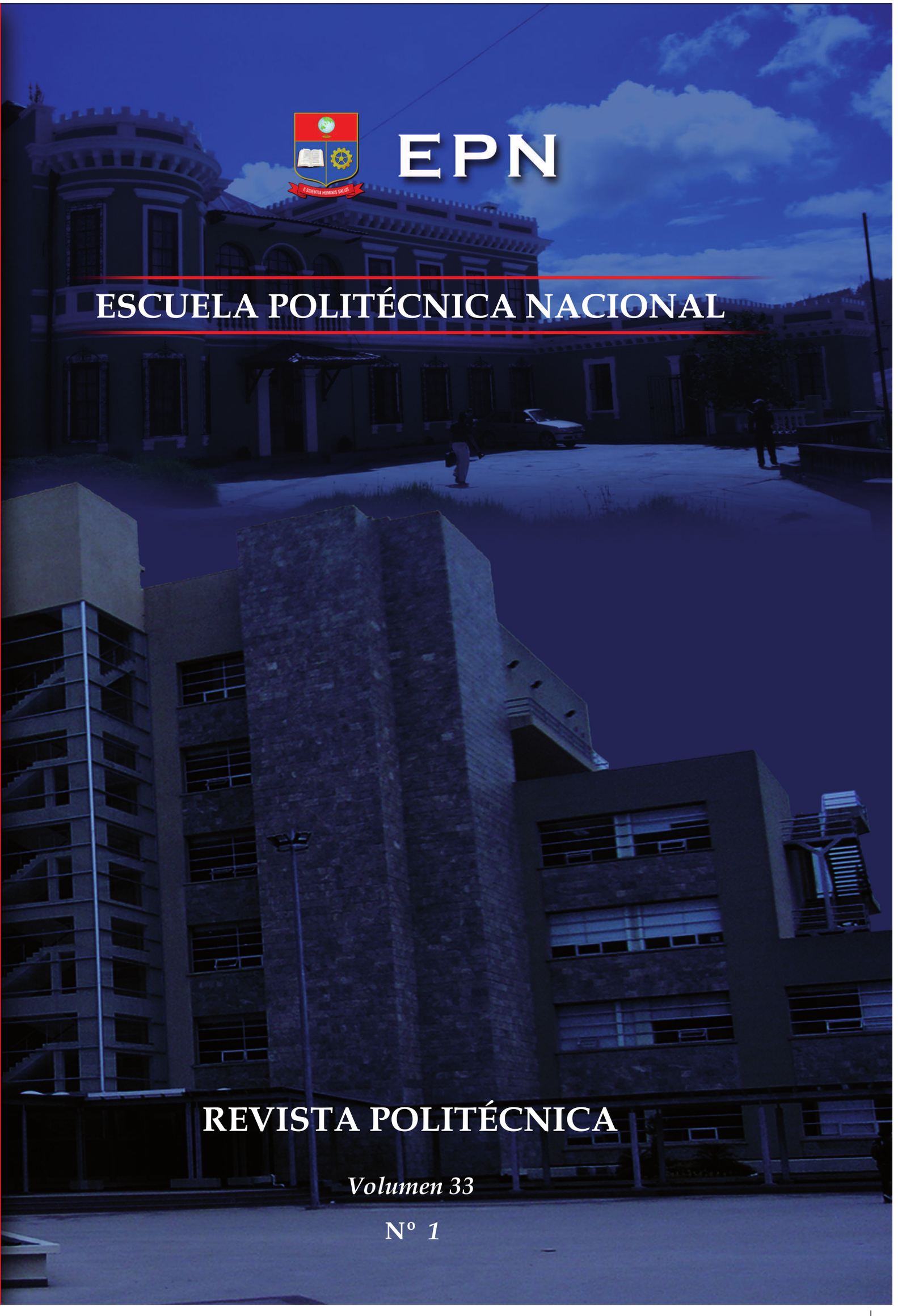


EPN

ESCUELA POLITÉCNICA NACIONAL

REVISTA POLITÉCNICA

Volumen 33

Nº *1*

ESCUELA POLITÉCNICA NACIONAL

REVISTA POLITÉCNICA

VOLUMEN 33 - NÚMERO 1

CAMPUS "JOSÉ RUBEN ORELLANA RICAURTE"
LADRÓN DE GUEVARA OE11-253
www.epn.edu.ec

REVISTA POLITÉCNICA

Volumen 33, Número 1
Enero 2014
ISSN: 1390 - 0129

ESCUELA POLITÉNICA NACIONAL

Rector
Jaime Calderón, MBA

Vicerrector de Investigación y Proyección Social
Alberto Celi, Ph.D.

Vicerrector de Docencia
Tarquino Sánchez, MBA

Editor
Andrés Rosales, Ph.D.

Co Editor - Invitado
Gustavo Scaglia, Ph.D.

Delegado Vicerrector de Investigación y Proyección Social
Eduardo Ávalos, Ph.D.

CONSEJO EDITORIAL

Enio Da Silveira, Ph.D.
Universidad Católica de Río, Brasil.

Farhang Sefidvash, Ph.D.
Federal University of Rio Grande do Sul, Brasil.

Gabriel Salazar, Ph.D.
Organización Latinoamericana de Energía.

Raymundo Forradelas, Ph.D.
Universidad Nacional del Cuyo, Argentina

Ricardo Carelli, Ph.D.
Universidad Nacional de San Juan, Argentina.

István Mezo, Ph.D.
University of Debrecen, Hungary

Vanderlei Bagnato, Ph.D.
Universidad de Sao Paulo, Brasil.

Rui Pedro Pinto de Carvalho, Ph.D.
University of Coimbra, Portugal.

Oscar Ortiz, Ph.D.
Universidad Nacional de San Juan, Argentina

Vicenzo Vespri, Ph.D.
Università degli studi di Firenze, Italia

Carlos Ávila, Ph.D.
Escuela Politécnica Nacional, Ecuador

Chen Ning, Ph.D.
Universidad de Mineralogía y Tecnología de China, China.

Lina Gómez, Ph.D.
Universidad Nacional de Colombia, Colombia.

Alex Ruiz Torres, Ph.D.
Universidad de Puerto Rico, Puerto Rico.

Patricia Garcés, Ph.D.
Pontificia Universidad Católica del Ecuador.

Marcos Villacís, Ph.D.
Escuela Politécnica Nacional, Ecuador

José Luis Rueda, Ph.D.
Universität Duisburg-Essen, Alemania

Danilo Chávez, Ph.D.
Escuela Politécnica Nacional, Ecuador

Hugo Arcos, Ph.D.
Corporación Eléctrica del Ecuador CELEC EP

Patricio Burbano, M.Sc.
Escuela Politécnica Nacional, Ecuador

Cristina Acuña, M.Sc.
Escuela Politécnica Nacional, Ecuador

Yadira Boada, M.Sc.
Universidad Politécnica de Valencia

COLABORADORES

Christian Hurtado
Vanessa San Martín
Iliana Carrera
Diana Barreno
Nicolas Magne
Marcia Acosta

ESCUELA POLITÉCNICA NACIONAL
Campus "José Ruben Orellana" Ladrón de Guevara E11 - 253
Quito – Ecuador

CONTENIDOS

1
Cela A. y Sánchez L.
Unsupervised Classifier for Dynamic Systems Based on ISODATA

7
Atiencia N., Ramírez V. y Bravo Y.
Control de Tráfico Vehicular usando un Sistema Neuro-difuso tipo ANFIS

12
Telenchana B., Villacís M., Rosales A., Chávez D., Ávalos E. y Arcos H.
Diseño y Construcción de una Plataforma Robótica para Control PTZ de un Sistema Electro-Óptico

20
Sotomayor J., Gómez A. y Cela A.
Sistema de Visión Artificial para el Análisis de Imágenes de Cultivo basado en Texturas Orientadas

27
Sotomayor J., Gómez A. y Castillo A.
Visual Control of an Autonomous Aerial Vehicle for Crop Inspection

35
Cela A., Sotomayor J. y Sánchez F.
Diseño de un Controlador Fuzzy para Guiado de un Robot Móvil

42
Gómez A. y Sotomayor J.
Instrument Landing System Lateral Beam Guidance System Based on Sliding Mode Control Technique

49
Boada Y., Morales L. y Sotomayor N.
Control de Seguimiento de Trayectoria y Paletización de un Robot de Tres Grados de Libertad Tipo *SCARA*

58
Chauca F., Llerena F. y Chico P.
Diseño y Construcción de un Inversor Multinivel

68
Molina I., Chico P., Rosales A. y Ávalos E.
Diseño y Construcción de un Sistema Portátil de Generación Eléctrica para Cargar un Banco de Baterías

78
Pozo D., Sotomayor N., Rosero J. y Morales L.
Medición de Ángulos de Inclinación por Medio de Fusión Sensorial Aplicando Filtro de Kalman

88
Leica P., Chavez D., Rosales A., Roberti F., Toibero J. y Carelli R.
Strategy Based on Multiple Objectives and Null Space for the Formation of Mobile Robots and Dynamic Obstacle Avoidance

99
Chaglla S., Terán D., Morales L., Rosales A., Pozo D. y Rosero J.
Apuntamiento y Estabilización Automática de un Sistema Electro-Óptico ante Perturbaciones

107
Narváez V., Yandún F., Pozo D., Morales L., Rosero J., Rosales A. y Auat F.
Diseño e Implementación de un Sistema de Localización y Mapeo Simultáneos (SLAM) para la Plataforma Robótica ROBOTINO®

114
Rosero J., Garza L., Minchala L., Pozo D. y Morales L.
Fuentes de Generación de Energía Eléctrica Convencional y Renovable a Nivel Mundial

127
Andrade L. y Ríos C.
Variabilidad Climática y Caudales Mínimos en los Andes Ecuatorianos

138
Celi C. y Arellano B.
Fiabilidad del Modelamiento Sísmico, en el Rango No Lineal, en Función de la Longitud de las Zonas Plásticas Asumidas Para Vigas

147
Mosquera L. y Hallo M.
Data Mart Para El Sistema De Servicios Sociales Del Conadis

156
Navarrete R. y Luján S.
Accesibilidad web en las Universidades del Ecuador. Análisis preliminar

164
Hernández M. y Gómez J.
Análisis de Sentimientos Aplicado a Referencias Bibliográficas

Unsupervised Classifier for DynamicSystems Based on ISODATA

Cela A.*; Sánchez L.**

**Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica*
Quito, Ecuador (Tel: 593-2-2677-031; e-mail: andres_cela@ieee.org)
***Universidad Carlos III de Madrid, Facultad de Ingeniería Telemática*
Instituto Tecnológico Superior Sucre, Departamento de Electrónica
Madrid, España (Tel: 34-6-00312-583; e-mail: franklin.sánchez@alumnos.uc3m.es)

Resumen:*Elsiguiente trabajo describe el método empleado para diseñar un algoritmo que permita clasificar un conjunto dinámico. Aunque hoy en día los clasificadores son muy bien conocidos, este trabajo muestra la combinación de diferentes técnicas para obtener un clasificador de mejores prestaciones aplicado a un objetivo específico. Los elementos de estudio pertenecen a un conjunto de líneas las cuales se representan con dos parámetros y se han obtenido después de tratar laimagen de una carretera y obtener todas sus líneas. El primer parámetro corresponde a la pendiente de la línea y el segundo a la posición (x,y) dentro de la imagen. Las imágenes se han obtenido de un recorrido real por una carretera, razón por la cual los elementos, es decir, las líneas, van cambiando de posición entre cada imagen, por tanto sus parámetros son dinámicos. Se parte del clasificador básico de Bayes para obtener un clasificador tipo KNN. En los elementos de estudio no se puede determinar las clases ni obtener sus parámetros, por lo tanto es necesario aplicar un clasificador no supervisado basado en KNN. En la fase final se hace una modificación en la semilladel clasificador ISODATA para aplicarlo al conjunto dinámico. Para reducir el trabajo de clasificación se trabaja únicamente con dos conjuntos los cuales correspondena las líneas izquierda y derecha de la carretera respectivamente. Se demuestra que el algoritmo propuesto reduce el tiempo de ejecución en 25% e incrementa los aciertos a un 90% respecto del clasificador ISODATA sin semilla inicial.*

Palabras clave:*clasificador no supervisado, conjuntos dinámicos, estudio de clasificadores.*

Abstract:*This paper describes a method for designing an algorithm to classify a dynamic set. Nowadays, the classifiers are well known, this work shows a combination of different techniques in order to get a better classifier for a specific work. The set of elements correspond to lanes marks of a road and they have been obtained from a road image. The lines have two parameters which are the slope line and the (x,y) position of the line in the image.The lines are obtained for each frame of a video sequence. The lines position changes between frames, then the parameters also change, this mean they are dynamics .The Bayes theory is the initial point to get a KNN classifier. In the set of study there is not possible to calculate the classes or their parameters, then, an unsupervised classifier based on KNN classifier is used for this approach. Finally, in the last step, it is made a modification in the seed for applyingin a dynamic set. It is shown that this approach reduces the process time in 25% and improves the accuracy in 90%.*

Keywords:*unsupervised classifier, dynamic sets, classifiers.*

1. INTRODUCTION

The classification of elements from a population has well been studied for the last two decades. The classifier calculates the probability that an element belongs to one category or another.

A category, called also class or pattern, can represent different things, like a car, an animal, a shape, a face, etc.

Categories can have different features, commonly called dimensions.These categories have the characteristics of the pattern, like color, texture, dimensions, position, velocity, etc. It is easy to think that to have more features is better for the classifier, but that is not true. Therefore, there are complete studies to obtain the better features which allow to classifythe population. To select the minimum features is recommended by some researchers in [1] and [2]. The categories belong to the same set and same features but each one is located in a different place in the d-dimensional space, where d is the number of features.

Some systems are static, but there is also some dynamic population. In these systems the categories change in time[3]. Therefore, a different kind of classifier is applied. The proposal in this work is to use an unsupervised classifier in order to do the classification based on ISODATA Clustering (IDC) method[4]. The data set were collected from a process which obtains images of the road, then, different artificial vision algorithms obtain lines from the road. These lines can belong to the marks of the road lanes. The goal of this

approach is to classify these lines in two categories, the left lane mark and the right lane mark in order to get two estimated lanes marks using the image information.

However, this system is not static but dynamic. This means that, the position of the categories is changing in time.

Classification of dynamic systems has not had many applications because there are other algorithms, like trackers based on Kalman filters, particle filters[5], or other ones[6], that can also do similar classification in these systems.

On the other hand, there are many classifiers, parametric or non parametric, unsupervised or supervised for static systems[7], but for dynamic systemstracking is most used.

This paper shows that an IDC algorithm can classify this kind of systems too.

Next section describes the method to design the classifier. It includes an introduction about Bayesian decision theoryand its application in multidimensional systems. The simplification of the Bayes theory based on the Gaussian shapes is important in order to do the formulation simpler.

Section 3 shows the results in two cases, the first one is applying the basic IDC and the second one is applying the improved IDC.Conclusions and future works are presented in section 4.

2. CLASSIFIER DESIGN

2.1 Bayesian decision theory

The Bayesian decision theory is a fundamental approach for classifying population in categories. It is based on quantifying the probability of each element to belong a category[4]. Also, it uses probabilistic terms and for each category the probability can be measured direct or indirectly. There is a basic rule which allows to find this probability. If avariable x , or element that belongs to a set C , satisfies (1), then it will be classified in class w_1 , instead class w_2 because the probability that belongs to w_1 is higher [7].

$$p(x \mid w_1) > p(x \mid w_2) \tag{1}$$

Bayes' formulais the general way to understand this theory. It converts the prior $P(w_i)$ probability to the posterior probability $P(w_i \mid x)$. The $p(x \mid w_i)$ function is the likelihood of w_i respect to x and *i* is the number of categories[8], as it is described in (2).

$$P(w_i \mid x) = \frac{p(x \mid w_i).P(w_i)}{p(x)} \tag{2}$$

If (2) is combined for two categories it will obtain (1). The evidence factor $p(x)$ is the same for all categories and can be viewed as a merely scale factor to normalize to 1 the posterior probability.

In some cases the prior probability can be observed like a modulation factor, but if this is not known it can be calculated by using (3), where $i_{\max}$ is the total number of categories.

$$P(w_i) = 1/i_{\max} \tag{3}$$

The prior probabilities modulate the $p(x \mid w_i)$ functions. However, it is important to know the prior probability because it can give more weight or cost to any category.

Generally, the Gaussian shape, called also Normal density, is used to study the basic classifiers. The Probability Density Function (PDF) is represented by $p(x \mid w_i)$ and it will be defined by a Gauss function. In the Fig.1 the PDF of each category is multiplied by the prior probability, whichis why the Gaussians have different altitudes.

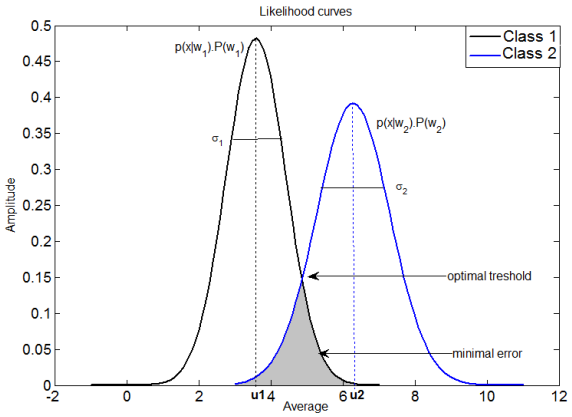


Figure 1.Decision theory.

The normal shape function for univariateis defined by (4) where the probability is calculated by the normal function. In the mono variable function u is the mean and σ is the variance. Therefore, the mean and the variance are the parameters which specify the normal function[9].

$$p(w \mid \sigma) = 1/(\sigma(2\pi)^{1/2}).\exp(-(x-u)^2/(2\sigma)) \tag{4}$$

2.2 PDFmultidimensional simplification

The normal function parameters can be found by different parametric techniques like Maximum Likelihood (ML), Bayesian Learning (BL) or by non parametric techniques like Parzen Window (PW) and K-Nearest Neighbors (KNN)[4].

In this work the population is a dynamic system, so the parameters of each category are dynamic too.

Applying a logarithmic function to (4) it will be simplified as it is shown in (5).

$$\ln(\theta(u,\sigma)) = \ln(u,\sigma) = -\ln(2\pi)^{1/2} - (x-u)^2/(2\sigma) \quad (5)$$

ThePDF for multidimensional systems is defined by (6), where $\bar{x}$ is the n-dimensional variable, $\bar{u}$ is the mean vector and Σ is the covariance matrix.

$$L(\bar{u},\Sigma) = -Ln((2\pi)^{1/d} |\Sigma|^{1/2}) - (\bar{x}-\bar{u})^T \Sigma^{-1}(\bar{x}-\bar{u})/2 \quad (6)$$

With this transformation it is possible to see that the normal shape implied to use the distance between the variable and the mean, as it is showed by the factor $(\bar{x}-\bar{u})^T \Sigma^{-1}(\bar{x}-\bar{u})$. This is known as the Mahalanobis Distance (MD)[9].

It is assumed that the variables do not have correlation between them and the covariance matrix is an identity matrix. Then, the MD became Euclidean Distance (ED). Besides, the first factor in (6) is a constant for all variables and it can be eliminated. Thereby the final PDF is

$$L(\bar{u}) = (\bar{x}-\bar{u})^T I(\bar{x}-\bar{u}) \quad (7).$$

The normal shape hasyielded a PDFbasedon distances.

However, this technique is already a parametric technique and it is necessary to know the static $\bar{u}$, but it has been defined that this will be a dynamic mean instead.

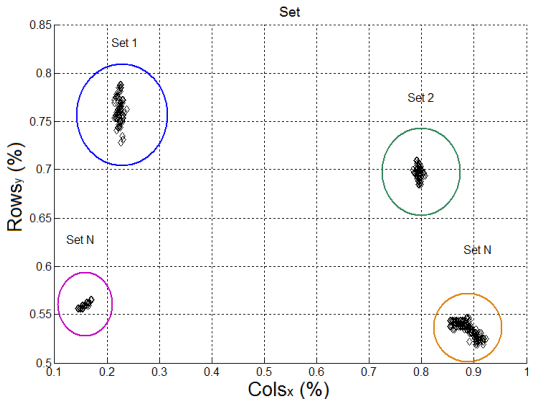
In this paper it is proposed to use an unsupervised classifier to find the dynamic mean.

2.2Finding the mean $\bar{u}$.

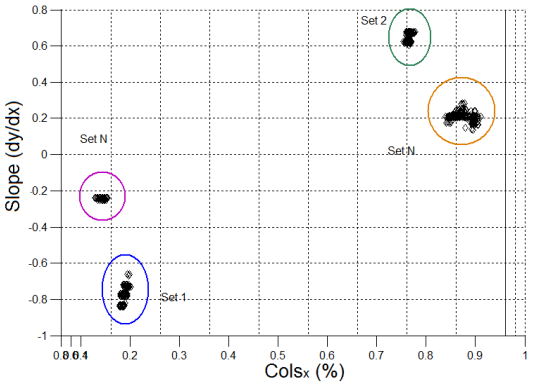
The population in this research will start in a known space. There are 4 categories in this set, but only thetwo central categories are considered for classification, as it is shown in Fig.2, where set 1 and set 2 are these categories. These have been considered because their represent the left lane mark and the right lane mark of respectively. The others categories, represented by set N, correspond to lanes marks of others lanes of the road because the road had 3 lanes.

The population is in the 3D space, where X axis represents columns, Y axisrepresents rows and Z axis represents the slope. The population is a set of lines which has been obtained from a sequence of video of road lanes and represent possible lane marks on the road. The (x,y) position is the line location on the image, and the Z is its slope.

The variables, the population, can seem an easy set to classify, but it is necessary to remember that it is only the first frame of a dynamic system and then the categories will change their position on the 3D space. Additionally, the mean of each category is not known yet. This is the aim of this work.



(a)



(b)

Figure 2.Population. (a) (x,y) plane (b). (x,z) plane.

Equation (7) is used for KNN. In this kind of classifier the variable is assigned to the nearest category [10]. However, the means are not known yet and they have to be calculated by another method.

The unsupervised classifier ISODATA Clustering (IDC) is usedto solve this problem. In this method it is not necessary to know the categories because it uses the KNN spectral distance to define the clusters and to find their means [11].

The IDC is an iterative method and it starts with random categories.The KNN classifies the population in the categories defined in the previous step. Then, the means of the categories are calculatedand new means are founded. Whole process is repeated until satisfying a threshold defined by the programmer[11].

Table 1. describes the IDC process.

Table1.ISODATA Clustering process.

1 Cluster centers are randomly placed. $\bar{u}_1, \ldots, \bar{u}_c$.
2 Simple classify by KNN.
3 Calculating new means $\bar{u}$ for each category.
4 If $\Delta\bar{u}$ > min. →go to step 1
Else,Stop.

In this research a data set has been classified and the results are shown in Fig. 3. In some cases the classifier converges to false categories. These results are presented in the next section.

This process calculates the dynamic parameters for categories, in this case the parameter is the mean $\bar{u}$.

3. RESULTS

3.1 ISODATAapplied to a dynamic system.

Fig.2 shows the classification of the first set, but the category means change in time.Hence, the IDC is applied to each frame. The KNN has much delay because it measures the distances for each variable. Besides, if the IDC does not converge in the first frame the whole process will fail.

To solve this issue it is proposed to use a prior mean in step 1 of IDC process. Then, the means obtained in the last step will be the prior mean for the next frame of the dynamic set.

The real lane marks have initial positions in the image which is taken from the first image. To define the first $\bar{u}$ the mean in the expected categories has been measured, which always starts in the same regions with means defined by (8) and (9) which represent the initial positions of categories in the image.These is the initial seed for using the IDCclasiffier.

$$\bar{u}_{1_0}(x,y,z) = (0.2 * x\max, 0.8 * y\max, -y\max / x\max) \quad (8)$$

$$\bar{u}_{2_0}(x,y,z) = (0.8 * x\max, 0.8 * y\max, y\max / x\max) \quad (9)$$

The result to applying IDC without the prior parameters (8) and (9) to population is depicted in Fig. 3 where the blue elements have been classified in set 1, but some of these elements do not belong to set 1, these are positive false. The red ones have correctly been classified in set 2. The initial seeds u_1 and u_2 have been randomly selected. We have tested this process 200 times and it fails around 60 %.

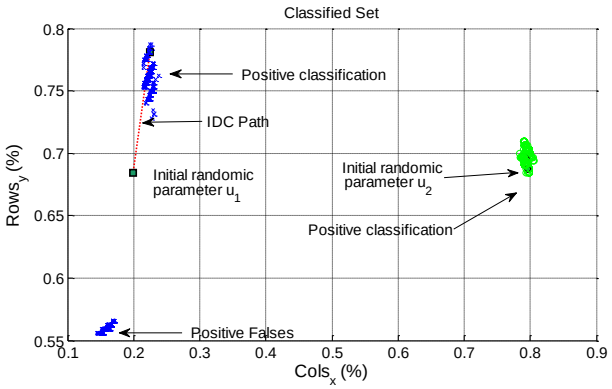


Figure3.ISODATAClustering without $\bar{u}$ prior.

On the other hand, in Fig. 4 is depicted the IDC results with prior parameters. The blue elements have been classified in set 1 and the green ones in set 2. The initial seed in category 1 is $\mathbf{u_{10}}$ which change to $\mathbf{u_{11}}$. The initial seed in category 2 is $\mathbf{u_{20}}$ and the recalculated seed is $\mathbf{u_{21}}$. The black elements have not been classified in any set.In this case IDC has 99% reliability.

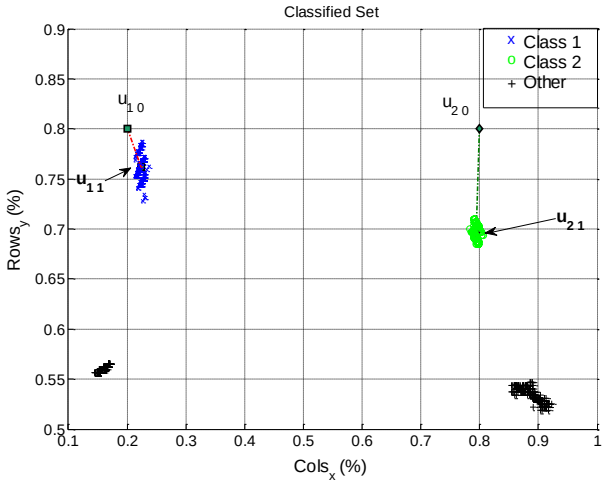


Figure 4.ISODATAClusteringwith $\bar{u}$ prior.

Fig. 5 shows the convergence of four means which started in a random position. Insecond iteration, x1 and x4 means have been changed, and in iteration 3 the means have converged.

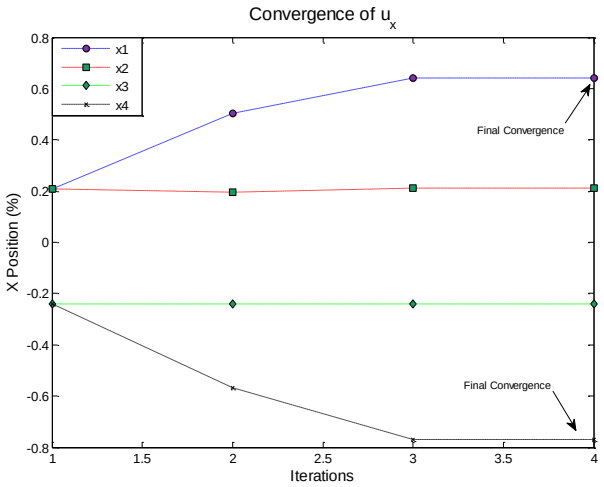


Figure 5.Convergence of means in IDC without initial seed.

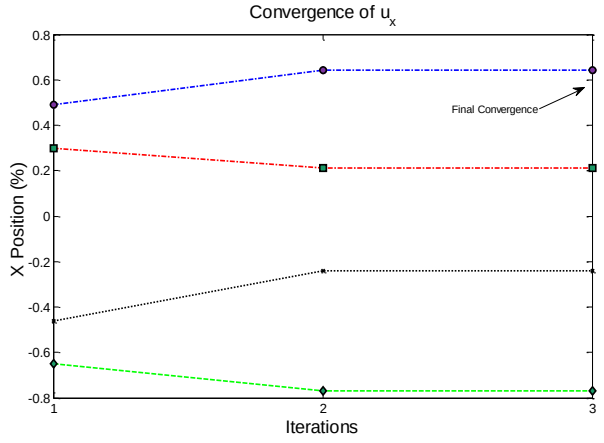


Figure6.Convergence of means in IDC for dynamics sets.

Fig. 6 shows the convergence of the meansusing IDC with initial seed. In iteration 1 the means started in the initial seed. In iteration 3 the means have converged. This method needs just 3 iterations, while typical IDC needs 4.If our method is applied to the whole dynamic system the classification will be 25% faster.

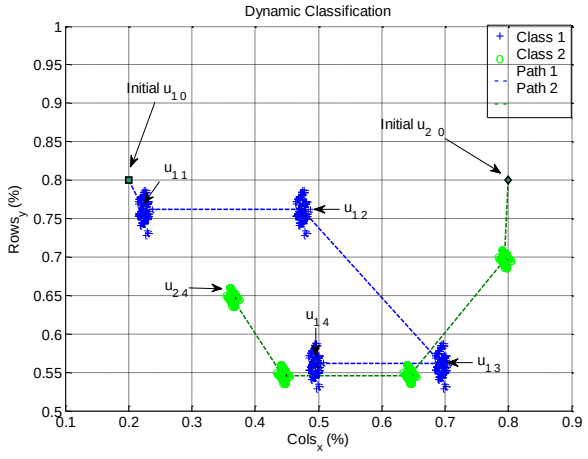


Figure 7.Classifier applied to dynamic system.

The classification of 3D dynamic systems is depicted in Fig.7.Here, the first frame is classified based on the prior parameters (8) and (9). Then, IDC classifies the population in two categories, in green and blue respectively, and it obtains the newer means $\bar{u}_{1_1}$ and $\bar{u}_{2_1}$. This algorithm is recurrent until the last frame of population. The blue line and green lines are the paths taken by category 1 and 2. This example shows a classification of a 3D dynamic system for the first 4 frames.

This process can resemble a tracking mechanism, but the difference is that in this case the categories are classified for each frame.

Table 2 shows the algorithm of the improved IDC process applied in dynamic systems. The***k*** variable is the frame number.

Table2.ISODATA Clustering process for dynamic systems.

1 For ***k*** == 0

Cluster centers $\bar{u}_{1_0}$ = (8) and $\bar{u}_{2_0}$ = (9).

2 For ***k***> 0 **until** whole frames.

3 Clustercenters placed at: $\bar{u}_{1_k-1}$ and $\bar{u}_{2_k-1}$.

4Simple classify by KNN.

5Calculating new means $\bar{u}_{1_k}$ and $\bar{u}_{1_k}$.

6If $\Delta\bar{u}_{1_k}$ > min&& $\Delta\bar{u}_{k_2}$ > min. → go to step3

Else →go to step 2, *k*= *k*+ 1

4. CONCLUSIONS ANDFUTURE WORKS

The Bayesian classifier is the basis for understanding the scope of decision theory.

An unsupervised classifier allowsto findthe categories in sets.Thiskind of classifier does not require training or supervision, but in some cases the algorithm does not converge to expected categories. In this paper we have used an unsupervised classifier called ISODATA IDC. We have tested IDC with the sets in this paper and it fails around 60%, as we have shown in the results. IDC uses a randomic initial means but we have used an initial seed in order to improved IDC. For each image the improved IDC is 25% faster than the typical IDC and has 99% reliability.

The sets in this work change in time and categories are dynamics which can be classified using this new method, but it is important to know the prior parameter, as in this case.

In future works, we will test other kinds of classifiers and filters. It can also apply a Kalman filter to the means to refine the results and tracking the means.

Although this algorithm is applied to specific sets, we will test the improved IDC in other kinds of sets.

REFERENCES

[1] M. B. A. Haghighat y E. Namjoo, «Evaluating the informativity of features in dimensionality reduction methods», en *2011 5th International Conference on Application of Information and Communication Technologies (AICT)*, 2011, pp. 1-5.

[2] Laurens van der Maaten y ic Postma, «Dimensionality Reduction: A Comparative Review». Tilburg University, 26-oct-2009.

[3] D. Gillies y Y. Zheng, «Dynamic Interactions with the Philosophy of Mathematics», *Theor. Rev. Teoría Hist. Fundam. Cienc.*, vol. 16, n.º 3, pp. 437-459, jun. 2008.

[4] R. O. Duda, P. E. Hart, y D. G. Stork, *Pattern Classification*. John, 2nd ed., Wiley & Sons, 2012, pp. 20-200.

[5] J. Carpenter, P. Clifford, y P. Fearnhead, «An Improved Particle Filter for Non-linear Problems», 2004, pp. 2–7.

[6] M. Isard y A. Blake, «Contour Tracking By Stochastic Propagation of Conditional Density», 1996, pp. 343–356.

[7] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed., Academic Press, 1990, pp. 120-300.

[8] Y.-H. Cai, «The comparative study of different Bayesian classifier models», en *2010 International Conference on Machine Learning and Cybernetics (ICMLC)*, 2010, vol. 1, pp. 309-313.

[9] C. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 2007, pp.100-200.

[10] J. M. Kleinberg, «Two algorithms for nearest-neighbor search in high dimensions», en *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, New York, NY, USA, 1997, pp. 599–608.

[11] J. C. Bezdek, «A Convergence Theorem for the Fuzzy ISODATA Clustering Algorithms», *Ieee Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, n.º 1, pp. 1-8, 1980.

Control de Tráfico Vehicular usando un Sistema Neuro-difuso tipo ANFIS

Atiencia N. [*], Ramírez V.[*], Bravo Y. [*]

**Escuela Politécnica Nacional, Departamento de Automatización y Control Industrial*
Quito, Ecuador (Tel: 593-2-2507-144; e-mail: neymar.atiencia; yadira.bravo @epn.edu.ec; vic.ramireza@gmail.com

Resumen: *La intención de este trabajo es proporcionar un enfoque en el manejo de la semaforización en la ciudad de Quito de acuerdo a los requerimientos de tránsito en una vía, basado en técnicas de control inteligente como lo es ANFIS. Se escoge un modelo de tráfico adecuado y se desarrolla un simulador, además se realiza la recolección de datos para alimentar al mismo y obtener los datos necesarios para entrenar la estructura ANFIS. Por último se valida el modelo en un software profesional dedicado para tráfico, AIMSUN*

Palabras clave: Control, Semaforización, Mesoscopico, Neuro-difuso

Abstract: *The intention of this paper is to provide a focus on managing the traffic lights in the city of Quito according to the requirements of traffic on a road, based on intelligent control techniques such as ANFIS. Choosing a suitable traffic model, a simulator is developed, and collected data feeds the same and then get the data needed to train the ANFIS structure. Finally the model is validated in a professional software dedicated to traffic, AIMSUN*

Keywords: Control, Traffic Light, mesoscopic, Neuro-fuzzy

1. INTRODUCCIÓN

Con el pasar del tiempo en las grandes capitales del mundo, el flujo vehicular ha incrementado constantemente, sin embargo, el desarrollo de la infraestructura vial ha sido insuficiente; esto se ha reflejado en situaciones donde el volumen de tránsito en uno o más puntos de la red vial excede el volumen máximo de circulación. Por lo tanto, el control de tráfico inteligente es una herramienta muy importante para el alivio de la congestión vehicular y consecuentemente beneficia la calidad de vida de las personas, el medio ambiente y la economía de la ciudad.

En los últimos años el arquetipo de control del flujo vehicular mediante la programación de los semáforos ha presentado un sin número de variaciones que van desde la programación con tiempos fijos de las diferentes fases de los semáforos hasta la coordinación de grupos semafóricos desde una central computarizada, demostrando que se han obtenido grandes avances en el tema.

Ergo, existen diferentes propuestas de solución al problema de tráfico vehicular, pero a su vez existe un gran campo de trabajo, una muestra de ello es el modelo presentado en este artículo para optimizar el tiempo de sincronismo de los semáforos.

2. DISEÑO DEL CONTROLADOR DE TRÁFICO

Inicialmente se presenta un modelo de tráfico mesoscópico, el cual es el punto de partida para diseñar el control de los tiempos de desfase de verde entre los semáforos de cada intersección utilizando el ANFIS.

2.1 Descripción del Modelo

En "Modeling traffic through a sequence of traffic lights" [1] se introduce un modelo microscópico de tráfico vehicular; prácticamente es un modelo ideal basado en el comportamiento cinemático de un solo vehículo que viaja a través de una secuencia de semáforos que encienden y apagan con una frecuencia específica y que presentan distancias fijas (para este caso 100 m).

Partiendo de esta investigación, el primer objetivo es analizar una corriente vehicular (varios vehículos) a través de una secuencia de semáforos que puedan presentar distancias variables entre ellos y luego controlar estos semáforos para pretender mantener la velocidad máxima del flujo vehicular permitida en la vía. A continuación se describen los parámetros que componen la función:

La separación entre el n-ésimo y el (n+1)-ésimo semáforo es Ln. La n-ésima luz es verde si sen $(\omega nt + \varphi n) > 0$ y roja en caso contrario, donde ωn es la frecuencia del semáforo y φn es el desfase entre semáforos, positivo si el verde del n-ésimo semáforo enciende luego de que haya encendido el verde del (n-1)-ésimo semáforo y negativo en caso contrario. Un automóvil en esta secuencia de semáforos puede tener:

- Una aceleración positiva a_+ hasta que su velocidad alcanza la velocidad máxima de la vía $v_{máx}$.
- Una velocidad constante $v_{máx}$ con aceleración nula.
- Una aceleración negativa a_- hasta que se detiene.

$$\frac{dv}{dt} = \begin{cases} a_+\theta(vmáx - v), & acelera \\ a_-\theta(v) & frena \end{cases} \quad (1)$$

En la Ecuación (1), θ(v) es la función escalón unitario, la cual impide que el modelo arroje velocidades superiores a la máxima e inferiores a cero.
Como el auto se acerca al n-ésimo semáforo con velocidad v el conductor debe tomar una decisión, frenar o no en la distancia Xc:

$$X_c = \frac{V^2}{2a_-} \quad (2)$$

La diferencia entre L_n y X_c (L_n - X_c), es la distancia a la cual el vehículo puede acelerar hasta alcanzar la $v_{máx}$.
Para el modelo mesoscópico en el que se maneja un flujo vehicular, cada vehículo acelera con una aceleración a_+, se supone que el primer vehículo en la fila posee una $a_+ = 1.39\, m/s^2$, este dato se obtiene de la velocidad máxima de circulación en el perímetro urbano y de medidas experimentales tomadas del tiempo promedio de aceleración de un vehículo.

$$a_+ = \frac{v_{máx}}{t_{promedio_aceleración}} = \frac{50\times\frac{1000m}{3600\,s}}{10s} = 1.39\, m/s^2 \quad (3)$$

Se estima que cada vehículo que precede al primero disminuye su aceleración a_+ en un 2% [2], y se evalúa cada 0.1 segundos Δt la diferencia entre Xy X_c(X es la distancia recorrida por el vehículo), cuando ésta sea menor a 1m se analiza el estado del semáforo, es decir, *sen* ($\omega_n t + \varphi_n$), si el resultado es positivo y $v < v_{máx}$ se continúa con a_+, ó se mantiene v si $v = v_{máx}$; si el resultado es negativo se actualiza el valor de aceleración a a_-, que se calcula para que el *n-ésimo* vehículo frene dentro de la distancia correspondiente. Para limitar la aceleración mínima y máxima de frenado se hace referencia a datos medidos, de donde se obtienen las siguientes condiciones:

$$a_{-mín} = \frac{-v_{máx}}{t_{máximo_{frenado}}} = \frac{-50\times\frac{1000m}{3600\,s}}{3.6s} = -3.9\, m/s^2 \quad (4)$$

$$a_{-máx} = \frac{-v_{máx}}{t_{mínimo_frenado}} = \frac{-50\times\frac{1000m}{3600\,s}}{3s} = -4.63\, m/s^2 \quad (5)$$

En base a estos parámetros se calculan los datos de velocidad instantánea con los cuales se obtienen los datos de la velocidad promedio para los m autos del flujo vehicular.

La velocidad v de la ecuación del modelo inicial se reemplaza por el promedio de la velocidad promedio de los m vehículos ($\overline{V}_t$), es decir:

$$\overline{V_t} = \frac{\sum_{i=1}^{m} V_i}{m} \quad (6)$$

En la Ecuación (6) Vi es la velocidad promedio del *m-ésimo* vehículo.
Esta velocidad promedio se obtiene a partir de la velocidad instantánea del vehículo (V_{in}), la que se calcula cada 0.1 segundos (Δt), mientras el vehículo se desplaza en el flujo vehicular del *n-ésimo* hasta el *(n+1)-ésimo* semáforo.

Donde:

$$V_{in} = V_o + a_+\Delta t \quad (7)$$

$$V_i = \frac{\sum_{i=1}^{k} V_{in}}{k} \quad (8)$$

En la Ecuación (7), V_o es la velocidad que lleva el vehículo en el intervalo de tiempo y en (8), k representa el número de intervalos (Δt) en los que se pueda dividir la sección. Luego de describir el modelo a utilizarse para caracterizar el flujo de vehículos a través de una vía, se utiliza (9) para sincronizar los tiempos de verde (t_v) y de desfase (t_d) de los semáforos:

$$\sin\left(\frac{2\pi}{t_{Vn}+t_{An}+t_{Rn}}t + t_{dn}\omega_n\right) > 0 \quad (9)$$

Donde:
t_{Vn}es el tiempo de verde del *n-ésimo* semáforo.
t_{An}es el tiempo de amarillo del *n-ésimo* semáforo.
t_{Rn}es el tiempo de rojo del *n-ésimo* semáforo.
t_{dn}es el tiempo de desfase de la luz verde entre el *n-ésimo* semáforo y el *(n-1)-ésimo* semáforo.

Se debe considerar que la longitud promedio de un vehículo es de 4 metros y la separación entre vehículos así como la reducción de la velocidad promedio de circulación por presencia de transporte público se encuentran a partir del modelo lineal planteado en "Ingeniería de Tránsito: Fundamentos y aplicaciones" [3].
Partiendo de información recolectada mediante mediciones manuales en la zona en estudio, el volumen máximo vehicular (q_m) que se presenta sobre la avenida Colón es de alrededor de 2956 vehículos/hora, la velocidad máxima ($v_{máx}$) permitida en la vía es de 50 km/h (13,89 m/s) y debido a la presencia de transporte público en dos de los 4 carriles de la avenida en estudio esta se reduce en un 30%, se tiene entonces que la densidad de congestionamiento vehicular (k_c) es:

$$k_c = \frac{4 * q_m}{v_{máx} \times 0.7} = 337.8\, veh/km \quad (10)$$

El espaciamiento promedio a flujo máximo (s_m) es:

$$\boldsymbol{s_m = \frac{1}{k_m} = \frac{1}{\frac{k_c}{2}} = 6\ m/veh} \quad (11)$$

Y la separación entre vehículos a flujo máximo (s_p) es:

$$\boldsymbol{s_p = s_m - l_v = 6 - 4 = 2\ m/veh} \quad (12)$$

Adicionalmente se deben incluir las restricciones consideradas en este modelo, para el posterior desarrollo del simulador, las cuales consisten en:

- Tener un número máximo de 10 vehículos en un carril partiendo del reposo, esta restricción aparece debido a la geometría de la red vial (longitud entre intersecciones) que presenta la zona en estudio.
- Cada auto calcula si se encuentra en la zona en que puede frenar de forma normal para detenerse en la intersección y acomodarse uno a continuación de otro, manteniendo una separación (entre vehículos) igual a la inicial.
- Si el pelotón de vehículos se separa en más pelotones, cada vehículo debe redefinir la posición que ocupa dentro de su pelotón actual para mantener una separación adecuada entre vehículos al detenerse en una intersección.
- De acuerdo a la distancia de frenado y la velocidad que lleva el vehículo se calcula la aceleración de frenado, misma que no puede ser menor en modulo a $a_{-mín}$.
- Los autos se detendrán un metro después del semáforo de la última intersección en estudio.

2.2 Desarrollo de la Arquitectura ANFIS

Para el entrenamiento del modelo neuro-difuso, que se implementará, las reglas difusas del tipo Takagi-Sugeno se representan de la siguiente forma:

Regla n: Si X es An y Y es B_n, entonces $f_n = p_nX + q_nY + r_n$, la arquitectura de las reglas implementadas es una combinación lineal entre los parámetros del antecedente de cada regla:

Regla n: Si X es A_ne Y es B_n, entonces $f_n = p_nX + q_nY + r_n$,

Donde A_ny $B_n$son los conjuntos difusos de entrada y p_n, q_ny $r_n$son constantes, obteniendo un modelo Takagi-Sugeno de primer orden, para una salida lineal.

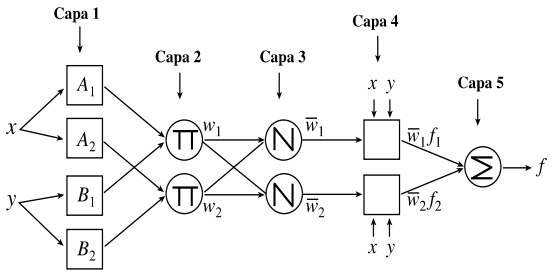


Figura. 1. Arquitectura ANFIS equivalente [3].

Se deben generar los datos utilizados para construir la estructura de inferencia difusa (FIS), es decir una matriz de datos entrada / salida que sea lo suficientemente representativa del comportamiento del modelo. Para construir la matriz mencionada se utilizó una matriz de 560x3 para los datos de entrada y una matriz de 560x1 para los datos de salida.

Para generar la estructura ANFIS se utilizaron los comandos de MATLAB que permiten escoger los parámetros adecuados para primero construir el Sistema de Inferencia Difuso FIS y luego entrenarlo.

Ahora la estructura ANFIS generada tendrá la información suficiente en sus 5 capas para poder ser evaluada ante un conjunto de datos representativos del sistema en estudio y así poder alcanzar la meta de lograr la velocidad máxima permitida en la avenida Colón.

Una vez que se obtenga la salida de velocidad deseada, los valores de tiempos de desfase serán guardados para luego realizar la simulación tanto en el Simulador implementado en MATLAB en base al modelo matemático descrito anteriormente como en el software profesional AIMSUN.

3. RESULTADOS

3.1 Resultados del Modelo Desarrollado en MATLAB

Para validar el modelo matemático desarrollado se hicieron pruebas con varios vectores de tiempos de desfase, tal que se obtengan diferentes tipos de comportamiento del tráfico vehicular.

Primero se comprobará que usando los datos recolectados el simulador de tráfico desarrollado replique un comportamiento similar al real del tráfico vehicular presente en la zona en estudio.

El vector de tiempos de desfase cargado para esto, obtenido de mediciones manuales en la zona en estudio, es:

$$td=[0\ -35\ 33.2\ -2.5\ -1.8\ 6.2\ 2.8]$$

Se utilizó como punto de comparación la gráfica de la velocidad de los vehículos en función del tiempo, Fig. 2.

Mediante una prueba de campo de observación se probó que las interacciones de los vehículos con los semáforos y entre ellos en la vida real son comparables con las obtenidas en la gráfica velocidad vs tiempo.

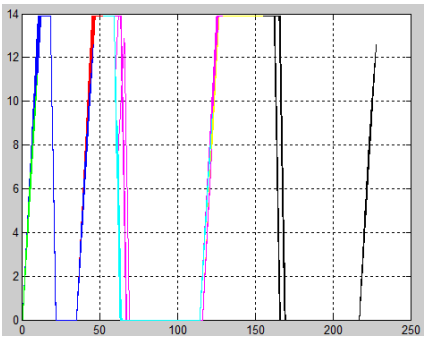


Figura. 2. Gráfica de Velocidad vs Tiempo para el Modelo Original.

Ahora se configura un modelo controlado, en base a la estructura ANFIS desarrollada, en dicho modelo controlado

se logra que las velocidades promedio del flujo vehicular que circula en las secciones de la zona en estudio sean máximas. De diferentes pruebas realizadas usando el controlador ANFIS generado, se obtuvieron como datos de ingreso para alimentar al modelo controlado los valores que forman el siguiente vector de tiempos de desfase:

$$td=[0\ 25\ -24\ 4\ -1\ 37\ 10]$$

Las velocidades promedio obtenidas en las secciones de la zona en estudio, Tabla 1., fueron máximas (13.8889m/seg=50km/h correspondiente a la velocidad máxima permitida) a excepción de la velocidad promedio de la primera sección, esto se debe a que en esta los vehículos del pelotón se encuentran acelerando hasta llegar a la velocidad máxima.Gráficamente los vehículos que parten del reposo en la simulación se mantienen en un solo pelotón y pasan a través de todas las intersecciones sin detenerse ni disminuir su velocidad.

Tabla 1. Velocidades promedio por sección para el modelo controlado

Sección	v [m/seg]
6 Diciembre-D.Almagro	12.7653
D. Almagro-R. Victoria	13.8889
R. Victoria-J. León Mera	13.8889
J. León Mera-Amazonas	13.8889
Amazonas-9 Octubre	13.8889
9 Octubre-10 Agosto	13.8889

3.2 Resultados en el Simulador Profesional AIMSUN.

Primero se trabaja omitiendo las intersecciones no semaforizadas para poner a prueba los resultados del controlador ANFIS desarrollado, evaluar la respuesta del sistema ante este control, y posteriormente generar perturbaciones en el modelo tales como el incremento de un plan de transporte público y un incidente en la vía, para así mostrar cómo se ven afectados los resultados del plan de control fijado inicialmente.

A través de los resultados obtenidos de las simulaciones interactivas y presentados en la Fig. 3 y en la Tabla 2, se prueba que el controlador ANFIS mejora la velocidad promedio de la zona en estudio.

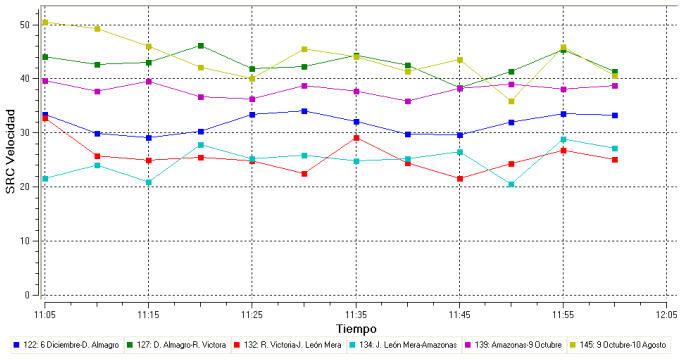


Figura.3. Velocidades por Sección del Sistema Controlado con un Plan de Transporte Público como Perturbación.

Nótese que dicho controlador se desarrolló en base a los datos de un modelo determinístico, el mismo que sirvió para implementar el Simulador de tráfico creado en MATLAB, y se validó en el software profesional de tráfico vehicular AIMSUN, que maneja modelos estocásticos en sus réplicas de simulación.

Tabla 2. Velocidades promedio por sección para el modelo controlado

	Velocidad Media [km/h]
Modelo Original	32.7
ModeloControlado	37.2
TransportePúblico	34.5
T. Público + Incidente	33.6

Luego se construye en AIMSUN una réplica más real de la Av. Colón entre las avenidas 6 de Diciembre y 10 de Agosto donde se consideren todas las intersecciones, semaforizadas y no semaforizadas. Para esta nueva condición se debe ajustar el plan de control, y los resultados de la simulación se evaluarán ante las mismas perturbaciones que en el modelo de simulación original.

En la Tabla 3 se nota que la velocidad media de la sección donde se configuró el incidente cae respecto a la mostrada en la simulación del modelo controlado. Sin embargo, dicha disminución no representa un cambio drástico en el comportamiento de las secciones en estudio como conjunto.

Tabla 3. Velocidad del Flujo Vehicular que Viaja a través de la Zona en Estudio

	Velocidad Media [km/h]
Modelo Original	31.1
ModeloControlado	34.5
Incidente en la vía	33.5

4. CONCLUSIONES

Un modelo mesoscópico, determinístico de tráfico vehicular puede servir como punto de partida para el diseño, implementación y ajuste de planes de control fijos, que sean aplicados a modelos de tráfico vehicular reales de carácter microscópico. Los modelos determinísticos facilitan el procesamiento matemático de las variables que intervienen en sus interacciones.

La precisión de la arquitectura ANFIS al emular un sistema, al igual que la de cualquier técnica adaptativa, depende de la correcta selección del conjunto de datos con los cuales se genere y entrene la estructura, pues estos deberían recoger todas las características representativas del sistema.

La información recogida, sobre el comportamiento del tráfico vehicular de la zona en estudio, alimentó correctamente el modelo de simulación de AIMSUN para obtener una réplica aproximada del sistema real; y así evaluar diferentes planes de control.

No se presentó ningún inconveniente al aplicar los resultados, obtenidos en el controlador ANFIS y evaluados en el

simulador desarrollado en MATLAB, al modelo de tráfico vehicular construido en AIMSUN; puesto que las velocidades medias por sección mejoran en ambos casos.

Una característica favorable para el alivio del tráfico vehicular, que se puede apreciar en los resultados, es el alcance de velocidades medias por sección que además de un incremento respecto a las originales presenten uniformidad a lo largo de la zona en estudio.

5. REFERENCIAS

[1] TOLEDO, B.A.; MUÑOZ, V.; ROGAN, J.; TENREIRO C., "Modeling Traffic through a sequence of traffic lights". Physical Review E. Vol. 70. pp. 0161071-0161076. 2004.

[2] PEDRAZA, L. F.; HERNÁNDEZ, C. A.; LÓPEZ, D. A., "Vehicular Traffic Control using ANFIS". Ingeniare. Vol 20. pp.79-88. 2012.

[3] CAL Y MAYOR**,** Rafael; CÁRDENAS, James, *"INGENIERÍA DE TRANSITO, Fundamentos y Aplicaciones"*. SéptimaEdición. Alfaomega. México 2000. Capítulo 10. Págs. 279 – 291.

Control PTZ de un Sistema Electro-Óptico utilizando una Plataforma Robótica

Telenchana B.*;Villacís M.*; Rosales A.*; Chávez D.*; Ávalos E.*; Arcos H.**

**Escuela Politécnica Nacional, Departamento de Automatización y Control Industrial,Quito, Ecuador*
e-mail: braulyo22@hotmail.com; marcoruben59@hotmail.com;
{andres.rosales; danilo.chavez; eduardo.avalos}@epn.edu.ec
***Corporación Eléctrica del Ecuador – CELEC EP*
e-mail: harcosm@gmai.com;

Resumen: *En el presente trabajo se ha desarrollado el control PTZ y la adquisición de imagen desde un sensor CMOS, para un sistema electro-óptico, el cual se prueba en una plataforma robótica terrestre.El control PTZ comprende el control del giro e inclinación del sensor CMOS y zoom de la imagen. El movimiento en los ejes horizontal y vertical se efectúa mediante eluso servomotores, gobernados por un microcontrolador ATmega8, el cual es comandado inalámbricamente desde un computador.Para la adquisición de imagen se utiliza la µCAM-TTL que cuenta con el sensor CMOS. Por medio de un microcontrolador ATmega164PA se envían los comandos de configuración al sensor para la adquisición de la imagen en los formatos RAW y JPEG, y se transmite inalámbricamente los datos de imagen al computador.Adicionalmente, se desarrolla un sistema de adquisición de imagen para la cámara analógica C-208A, con la finalidad de realizar una diferenciación en el desempeño y funcionalidad con la µCAM.Finalmente, se ha incorporado un sensor de orientación, el mismo que le permite al usuario conocer la orientación del sensor electro-óptico y representarlo gráficamente en el espacio tridimensional.*

Palabras clave:*Control PTZ, Sistema Electro-Óptico, Plataforma Robótica*

Abstract:*In this work, it has been developed PTZ control and image acquisition from a CMOS sensor for an electro-optical system, which is tested on a robotic platform ground.PTZ control includes the control of pan and tilt movements for CMOS sensor and zoom of the image. The movement in the horizontal and vertical axis is performed by servomotors, which are controlled by an ATmega8 microcontroller, and it is commanded wirelessly from a computer. For image acquisition, we have used a µCAM-TTL, which has the CMOS sensor. An ATMEGA164PA microcontroller sends the commands to the sensor, in order to set the image acquisition in RAW and JPEG formats, and it transmits wirelessly the image data to the computer.Additionally, it develops an image acquisition system for C-208A analog camera, in order to make a difference in the performance and functionality with µCAM.Finally, it has been incorporated an orientation sensor in the system, which allows the user to know the orientation of the electro-optical sensor and represent it graphically in 3D.*

Keywords*: PTZ Control, Electro-Optic System, Robotic Platform.*

1. INTRODUCCIÓN

Este trabajo surge como parte de un proyecto entre la Escuela Politécnica Nacional y la Fuerza Aérea Ecuatoriana, para diseñar y construir un prototipo nacional de un avión no tripulado (UAV), con un costo menor a los UAVs comerciales, partiendo de la ingeniería que utilice tecnología desarrollada en nuestro país.

Un UAV es una aeronave sin tripulación humana a bordo. Son usados mayoritariamente en aplicaciones militares. Actualmente los UAVs militares realizan tanto misiones de reconocimiento como de ataque. Adicionalmente tienen aplicaciones civiles, como en labores de lucha contra incendios y vigilancia de oleoductos.

Figura1. UAV equipado con un Sistema Electro-óptico

Para efectuar este tipo de misiones, el UAV lleva incorporado en la parte inferior de su fuselaje un sistema electro-óptico, como se muestra en la Fig. 1.

En particular, estetrabajo se encuentra enfocado al desarrollo del control PTZ y la adquisición de imágenes en tiempo real para un sistema Electro-óptico.

Debido a la complejidad de construcción de un sistema electro-óptico, este trabajo se enfoca en un prototipo

básico pero funcional, sobre el cual se podrán desarrollar En el país no se ha desarrollado este tipo de tecnología de manera integral, por lo que éste prototipo sentará las bases necesarias para realizar muchas mejoras o modelos nuevos y más avanzados.

2. DISEÑO E IMPLEMENTACIÓN DEL HARDWARE

El prototipo consta físicamente de dos partes: la estructura Pan &Tilt y los módulos del sistema.

2.1 Estructura Pan &Tilt

Al ser este un prototipo se procedió a analizar las distintas alternativas de sistemas robóticos existentes en el mercado, con el objetivo de seleccionar aquella quepermita realizar los movimientos en los grados de libertad requeridos, de tamaño y peso pequeños, pero a su vez con capacidad de carga suficiente para el montaje de los módulos del sistema. Como mejor opción se eligió la estructura SPT200 Pan &TiltSystemde componentes electrónicos "SERVOCITY".
En base a la estructura adquirida, se procede a seleccionar los servomotores considerando características como: velocidad, torque, tipo de acoples, rango de movimientos y dimensiones. A continuación se presentan los dos tipos de servomotores empleados:

- HS-645MG Ultra Torque
- GWS S125 1T 2BB Sail Winch Servo

Finalmente se realiza el ensamblaje de las piezas y subpiezas de la estructura con los servomotores, como se indica en la Fig. 2.

Figura2. Estructura Pan &Tilt ensamblada

2.2 Montaje de los módulos del sistema

En la parte superior se instala la cámara analógica C-208A, la µCAM y el sensor inercial IMU, adaptando una caja plástica para su montaje, como se indica en la Fig.3.
Para la realización del zoom, es necesario quitar los lentes iniciales de la cámara serial y en su lugar acoplar los lentes motorizados.

modelos mejorados a futuro.

(a) Vista trasera (b) Vista delantera

Figura3. Módulos del sistema instalado

Además, se evidenció la necesidad de una base que sirve de soporte para todo el prototipo, la misma que se la construye en madera de acuerdo a las medidas del servomotor y las dimensiones de la plataforma robótica que se utilizará en las pruebas.

Figura4. Prototipo del sistema electro – óptico

3. DESARROLLO DEL SOFTWARE

El funcionamiento de todo el prototipo se encuentra basado en dos microcontroladores. El primer microcontrolador es el ATmega 8 y es el que realiza el control PTZ del sistema. El otro microcontrolador es el ATmega 164P y es el encargado de recibir los datos de la imagen y enviarla hacia el computador para la adquisición de la imagen.
En la Fig.5 se presenta el diagrama esquemático del funcionamiento de todo el sistema.

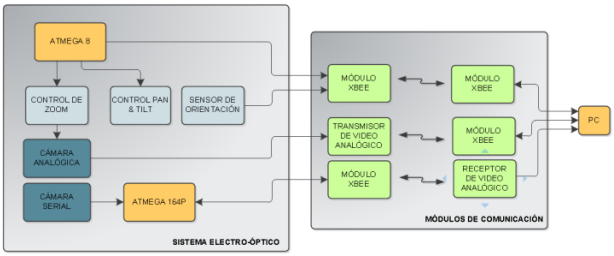


Figura5. Esquema general del sistema

3.1 Adquisición de Imagen

Existe un gran número de cámaras en el mercado, pero al tratarse este proyecto como un prototipo de un sistema electro-óptico, se comienza utilizando la µCAM – TTL, la cual está incorporada por el sensor de imagen y su respectivo procesador. Luego se emplea la cámara analógica C-208A, con la finalidad de analizar la complejidad de los dos sistemas y diferenciar los resultados en cuanto a la calidad de imagen.

1) µCAM-TTL

Mediante esta cámara es posible adquirir la imagen en dos formatos: RGB565 (RAW) y JPEG. Previamente a la configuración de la *µCAM*es necesario el estudio en detalle de los comandos y protocolos [7] utilizados por la cámara en los dos formatos.
La cámara se conecta directamente al puerto serial 1 del microcontrolador debido a su nivel de voltaje TTL, como se muestra en la Fig. 6. Se utiliza el regulador de voltaje LM1117T para obtener una alimentación de 3.3V, la cual es conectada a la cámara para su funcionamiento. Además se intercala un interruptor entre la salida del regulador y el pin Vcc, con la finalidad de reiniciar la cámara en caso de que no responda a los comandos enviados desde el computador.

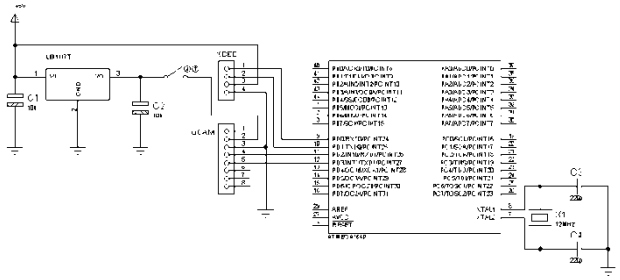


Figura6. Conexión µCAM TTL - Microcontrolador

La µCAM-TTL utiliza comunicación serial para la transferencia de datos de imagen. Mediante un puerto serial del microcontrolador ATmega164PA se transmite y recibe los comandos hacia y desde la cámara, los datos de imagen provenientes de la cámara son enviadosal ordenador, a través de otro puerto serial delmicrocontrolador.
Los datos son transmitidos inalámbricamente a una velocidad de 57600 baudios, mediante módulos XBee entre el microcontrolador y el computador. El computador envía las señales de control al microcontrolador, recibe los datos de imagen y los acondiciona para reconstruir la imagen.

2) Cámara Analógica C-208A

Mediante este tipo de cámaras la adquisición de imagen es sencilla, pues trabajan con el sistema de codificación NTSC que LabVIEW reconoce fácilmente, permitiendo de esta manera adquirir video en tiempo real.
En la Fig. 7 se muestra el esquema dela adquisición de imagen con la cámara analógica, donde:

- Cámara analógica: captura la imagen/video.
- Transmisor de video analógico: inalámbricamente el video en tiempo real, a una frecuencia entre 2.4GHz – 2.483GHz.
- Receptor de video analógico: recibe la señal inalámbrica de video y la envia por puertos RCA,.
- PC: Se comunica con la cámara analógica mediante el módulo NI VisionDevelopment de Labview, el cual adquier las señales analógicas de video y las convierte a formato digital, para su visualización.

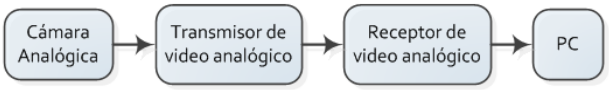


Figura7.Esquema de la adquisición de imagen con la Cámara Analógica

3.2 Control PTZ

En primer lugar, PTZ es un acrónimo de Pan, Tilty Zoom, y se refiere al movimiento horizontal, vertical y zoom, los cuales son controlados inalámbricamente desde un computador.
El microcontroladorATmega 8 recibe las señales de control provenientes del computador indicándole el tipo de control que debe realizar, ya sea control Pan, Tilto Zoom.
Para poder realizar el control Pan yTilt, el microcontrolador debe controlar la posición de los servomotores de giro e inclinación con los que cuenta la estructura Pan &Tilt. En cambio, para la realización del control de zoom se debe controlar un motor paso a paso bipolar.

1) Control de servomotores

El control de un servomotor se lo realiza mediante la técnica PWM (modulación por ancho de pulso), la cual consiste en enviar continuamente una serie de pulsos de relación de trabajo variable. La variación del ancho de pulso permite posicionar el servomotor en un determinado ángulo. El rango de variación del ancho de pulso está dado por el fabricante. El periodo de la señal de control por lo general tiene un valor de 20ms, y el ancho de pulso de la señal de control comúnmente se la debe variar entre 0.6ms y 2.4ms.
En la Fig. 8 se muestra la señal de control y la posición resultante del eje del motor para los valores de operación mínimo, central y máximo del servomotor.

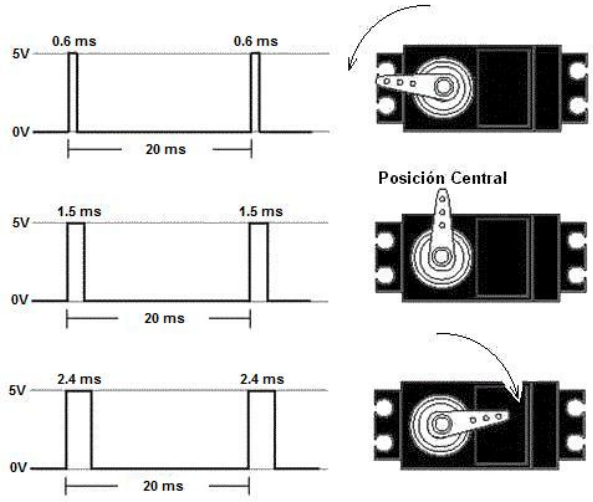


Figura8.Señales de control y posición del servomotor[8]

2) Control de motores paso a paso bipolares

Un motor paso a paso está diseñado para moverse en pequeños incrementos, por lo que es eficaz para un trabajo preciso. El ángulo de giro (también llamado paso) es muy variado pasando desde los 90° hasta los 1.8°, cada paso se efectúa aplicando un pulso a sus bobinas. La velocidad de giro del motor depende del tiempo entre cada uno de los pulsos.

Los motores paso a paso bipolares se encuentran conformados por dos bobinas y únicamente poseen cuatro cables de salida, como se muestra en la Fig. 9.

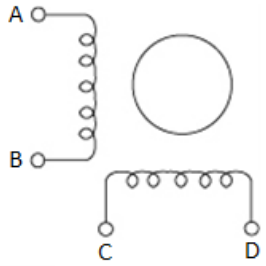


Figura9.Conexionado interno de un motor paso a paso bipolar

Este tipo de motores necesitan la inversión de la corriente que circula por sus bobinas en una secuencia determinada. Cada inversión de la polaridad provoca que el eje se mueva un paso, a continuación en la Tabla 1 se muestra la secuencia necesaria.

Tabla1. Secuencia normal de un motor paso a paso bipolar

Paso	Bobina A	Bobina B	Bobina C	Bobina D
1	ON	OFF	ON	OFF
2	ON	OFF	OFF	ON
3	OFF	ON	OFF	ON
4	OFF	ON	ON	OFF

3.3 Sistema de Orientación

El sistema de orientación proporciona información acerca de la rotación con respecto a tres ejes: el eje lateral, el eje vertical y el eje longitudinal. Esta rotación produce la variación de los tres ángulos de Euler que son: la dirección (yaw), elevación (pitch) y el ángulo de alabeo (roll).

Estos ángulos permiten conocer los movimientos realizados por el sensor electro-óptico en el espacio tridimensional, y de esta manera conocer su orientación en cada instante.

1) Sensor de Orientación

El sensor de orientación utilizado en este prototipo es el CHR-UM6, el cual combina sensores de medición de giro, acelerómetros, y sensores magnéticos, para entregar al usuario mediciones de orientación.[3, 4, 9, 10]

Figura.10.Sensor CHR-UM6

2) Obtención de ángulos de Euler

El sensor de orientación envía serialmente la información en forma de paquetes que se encuentran formados por 13 bytes hexadecimales. Cada paquete se encuentra estructurado como se muestra en la Tabla 2.

Tabla 2. Estructura de un paquete de datos

Secuencia de inicio	Tipo de Paquete	Dirección	Angulos de Euler			Checksum
			Roll	Pitch	Yaw	
3 Bytes	1 Byte	1 Byte	2 Bytes	2 Bytes	2 Bytes	2 Bytes

Los datos de los ángulos son enviados por el sensor como enteros de 16 bits en complemento a 2, por lo que para obtener una estimación del ángulo, se realiza el proceso inverso al cálculo del complemento a 2. Al concluir este cálculo, se tiene un rango de valores enteros con signo, que van desde -32768 a 32767.

Finalmente para obtener el ángulo real en grados se debe multiplicar el dato de la estimación del ángulo por un factor de escala cuyo valor es dado por el fabricante del sensor[10], como se muestra en la siguiente ecuación:

$$Ángulo\ real\ (°) = 0.0109863 * estimación\ del\ ángulo$$

Una vez realizado el procedimiento para la obtención del ángulo en grados, se tendrán variaciones de ángulos desde $-180°$ hasta $180°$ para yaw, y de $-90°$ a $90°$ para pitch y roll.

4.INTERFAZ GRÁFICA

El HMI (Interfaz Hombre-Máquina) del prototipo fue desarrollado mediante el software LabVIEW. Este software es una herramienta de programación gráfica.
El HMI desarrollado permite que el usuario se pueda comunicar con el prototipo, de tal modo que mediante una interfaz gráfica el usuario puede realizar el control PTZ, conocer la orientación y realizar la adquisición de la imagen del prototipo.

Figura11. Interfaz gráfica desarrollada

5.PRUEBAS Y RESULTADOS

El prototipo final fue sometido a varias pruebas, y en base a los resultados obtenidos, se comprobó el cumplimiento de los objetivos planteados para este proyecto.
Para la realización de pruebas en varios ambientes, el prototipo fue montado en un robot móvil todo terreno. En la Fig. 12se indica las imágenes de los distintos ambientes en los que fue probado el sistema.

Figura 12.Pruebas en distintos escenarios

5.1Pruebas de Adquisición de Imagen

Las pruebas de imagen están enfocadas a comprobar el funcionamiento de la cámara serial y el trabajo de los componentes que conforman el sistema, ya que éste sistema fue desarrollado en su totalidad tanto para el formato RAW como para el formato JPEG. El formato RAW contiene la información de cada pixel del sensor de imagen al momento de la captura, mientras que el formato JPEG comprime los datos para tener imágenes de menor tamaño. Además se muestran los resultados y diferenciaciones entre la imagen reconstruida con la cámara serial y la imagen de la cámara analógica.

1) Formato RAW

Las primeras pruebas con la cámara serial son realizadas con los lentes propios con los que cuenta la cámara.
Una vez verificado el correcto envío de los comandos hacia la cámara, se continúa con la adquisición de los datos de imagen. Luego se comprueba el adecuado funcionamiento del SubVI de pixelado de la imagen, para finalmente obtener como resultado la imagen formada con una resolución de 256x256 pixeles, la misma que se indica en la parte superior derecha de la Fig. 13.
La resolución de la imagen obtenida por el sensor es de 128x128 píxeles, mediante la realización de zoom digital de 2x es posible obtener una resolución dos veces mayor a la original, es decir, 256x256 píxeles[7].

Figura 13.Imagen RAW

En la Fig. 13 se puede apreciar la diferencia entre las imágenes obtenidas. Se tiene como resultado una imagen digital de baja calidad debido al zoom digital realizado y a la transformación de RGB565 a RGB888. El formato de color RGB565 representa cada pixel de la imagen mediante 2 bytes (16 bits), donde los 5 bits más significativos son para el rojo, los 6 bits siguientes para el verde y los 5 bits menos significativos para el azul, entonces para el color rojo y azul se tiene una resolución de 0 a 31 y para el color verde se tiene una resolución de 0 a 63. Labview emplea el formato RGB888 para procesar el color de los pixeles, es decir, se debe tener 3 bytes de 8bits con una resolución de 0 a 255, por lo tanto es preciso realizar la transformación.
El tiempo de trasmisión de una imagen en este formato es de 5.69 segundos.
Con el objetivo de comprobar la realización de elzoom óptico, se utilizaron lentes motorizados que fueron adaptados al sensor de imagen de la cámara serial, en la Fig.14 se muestran las pruebas realizadas.

Figura 14.Pruebas de zoom óptico

Uno de los puntos importantes a considerar es el desenfoque de la imagen al acoplar los nuevos lentes. La cámara con los lentes originales tiene una distancia focal predeterminada, mientras que con el cambio de lentes se produce la alteración de la distancia focal, es decir, todos los rayos de luz no coinciden en un solo punto y los lentes al no tener control de enfoque se presentanlos resultados de la Fig.14.

2) Formato JPEG

En primer lugar se adquiere una sola imagen para comprobar el correcto funcionamiento de los algoritmos de programación tanto para el microcontrolador y el HMI, para luego realizar la adquisición sucesiva de imágenes.
En la Fig. 15 se indica la imagen obtenida en este formato. La adquisición fue realizada con los lentes originales de la µCAM, para la obtención de mejores resultados.

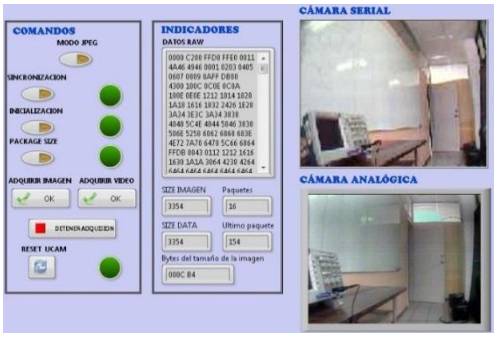


Figura 15.Imagen JPEG

Al ser un formato de imagen comprimida, el número de datos transmitidos se reduce considerablemente. El tiempo de transmisión aproximado por cada imagen es de 0,58 segundos.
Es de considerar que al ser un formato de compresión, el tamaño de los datos de imagen está en función del color y brillo de la escena a captar, es decir, el tamaño de los datos de cada imagen es variable. Los tiempos de transmisión de cada imagen son relativamente bajos, por lo cual en este formato es posible presentar las imágenes en modo video.

3) Cámara analógica

Para la adquisición de imagen se detecta la dirección de la cámara en el computador y se comienza a adquirir la imagen en modo video mediante las herramientas virtuales de Labview. La imagen obtenida se presenta en la Fig. 16.

Figura 16.Imagen de la cámara analógica

Como se observa en la Fig.16, la imagen obtenida por ésta cámara tiene mejores características como: color, definición y tiempo de trasmisión de la imagen.
La cámara analógica funciona adecuadamente gracias a una fina sintonización del receptor analógico. Además en la Fig.17 se muestran las pruebas de zoom digital, las cuales también presentan el desempeño deseado.

Figura17.Pruebas de zoom digital

4) Diferenciación entre cámaras y formatos

Uno de los puntos predominantes de diferenciación entre las dos cámaras es el tiempo de transmisión. La cámara analógica transmite video en tiempo real a una resolución de 682x452, mientras que la µCAM transmite una imagen de 128x128 en un tiempo de 5.72s en formato RAW, y en el mejor de los casos se tiene el formato JPEG con imágenes de 160x128 y un tiempo de 0.9s.
La cámara analógica adquiere una imagen de mejor calidad, por lo cual, se la utilizará para realizar las pruebas sobre la plataforma móvil. Además se utilizará el formato JPEG de la µCAM para observar su desempeño.

5.2 Pruebas Control PTZ

Estas pruebas consisten en comprobar el correcto funcionamiento de los movimientos horizontal, vertical y realización de zoom óptico. Estos movimientos pueden ser controlados mediante los botones situados en el panel frontal del HMI, o directamente desde los botones del joystick que se encuentra conectado al computador. Para tener una clara apreciación del control PTZ, en estas pruebas no se realizará el movimiento del móvil todo terreno.
Cabe mencionar que debido a la alta precisión en cuanto al posicionamiento de los servomotores, el control PTZ fue realizado en lazo abierto.

1) Control de giro

Estas pruebas se realizaron moviendo el sistema 90° desde el origen hacia el lado derecho, y 90° hacia el lado izquierdo. Los resultados de las pruebas de movimiento de giro se pueden apreciar en la Fig.18.

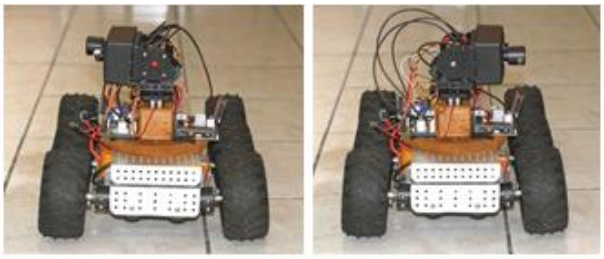


Figura 18.Movimientos hacia derecha e izquierda

El prototipo tiene la capacidad de girar una vuelta completa, es decir, desde la posición inicial puede realizar un movimiento de 180° en sentido horario, y 180° en sentido contrario. Este movimiento de 360° le permite al prototipo visualizar objetos que se encuentren en la parte posterior del sistema, sin la necesidad de mover el móvil todo terreno.

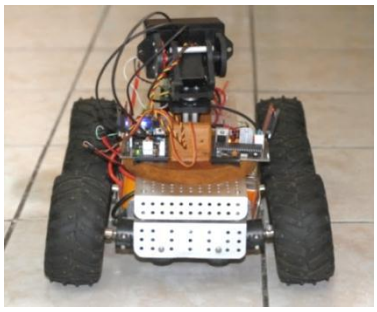


Figura19.Movimiento hacia la parte posterior del sistema

2) Control de inclinación

Las pruebas del movimiento vertical se muestran en la Fig.20, en la cual se puede ver el resultado del movimiento de la estructura hacia una posición de 90°.

Figura 20.Movimiento vertical hacia la posición de 90°

Los límites del movimiento vertical están dados por los topes mecánicos que posee la estructura, haciendo que el rango del control de inclinación se encuentre entre 0° y 135°,como se muestra en la Fig. 21.

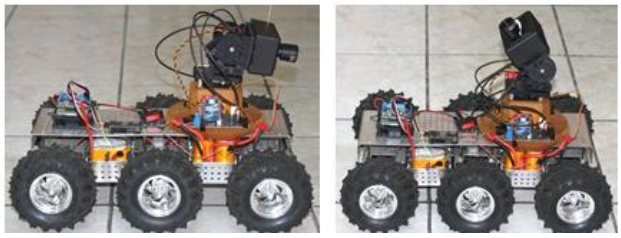


Figura. 21.Límites del movimiento vertical

3) Control de zoom

El control del zoom se realiza mediante lentes motorizados que fueron adaptados al sensor de imagen, estos lentes permiten realizar zoom óptico de una manera remota desde el computador.

En la parte izquierda de laFig.22 se puede observar la posición del lente motorizado para poder visualizar imágenes en tamaño normal.

Figura. 22.Lentes motorizados

Los lentes motorizados permiten la variación continua del zoom en pequeños pasos, gracias a que tiene incorporado un motor paso a paso. En la parte derecha de laFig. 22 se muestra el máximo desplazamiento posible del lente para el cual realiza el máximo cantidad de zoom permitida.

5.3 *Pruebas de Orientación*

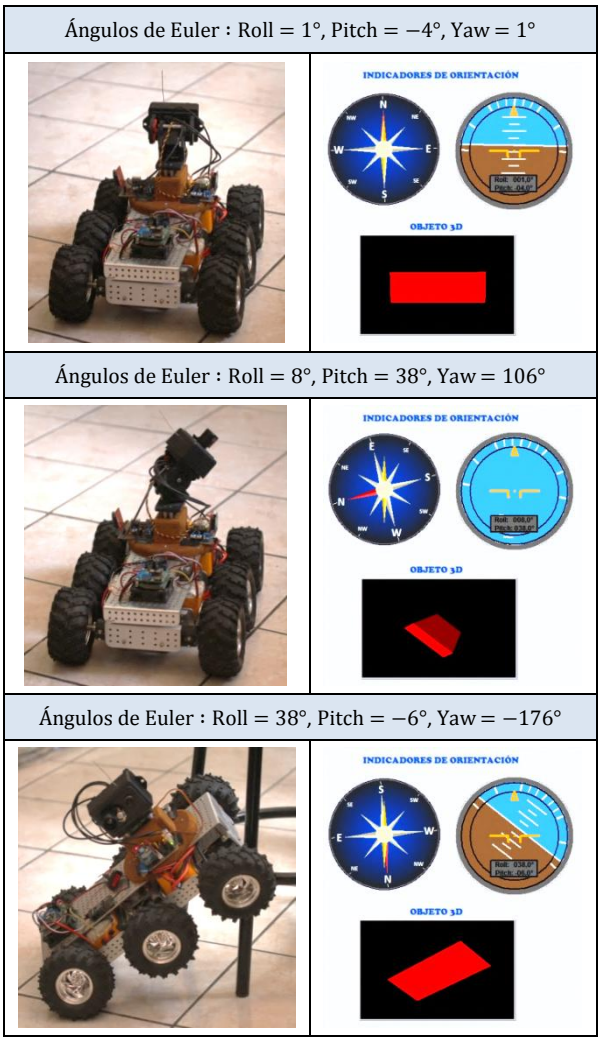


Figura 23.Pruebas del Sensor de Orientación

Las pruebas del sensor de orientación se realizaron colocando el prototipo en varias posiciones, como se muestra en la Fig. 23.

En la parte izquierda de la Fig. 23 se puede observar la posición real del sensor electro-óptico para cada caso, en la parte derecha se muestra el indicador de tipo brújula, el indicador de actitud y la representación gráfica en el espacio tridimensional. Adicionalmente se tienen los ángulos de Euler para cada posición del sensor electro-óptico.

6. CONCLUSIONES

Al tratarse este proyecto de un prototipo básico pero funcional, el control PTZ fue realizado en lazo abierto, beneficiándonos de esta manera del controlador de posición interno que poseen los servomotores.

La estructura mecánica cumple con los requisitos de movimiento que demanda el sistema, ya que cuenta con dos grados de libertad, permitiendo así controlar la inclinación y giro del sensor electro-óptico.

La selección de los servomotores es de gran importancia, ya que se debe considerar el peso total que va ser montado en la estructura mecánica y los rangos de movimientos que ésta tendrá.

La comunicación inalámbrica es muy confiable en distancias menores a 30m cuando existe línea de vista entre los módulos. A medida que aumenta la distancia o en ambientes cerrados, existe pérdida de información provocando errores en el sistema.

El formato RAW representa el color de cada pixel mediante 2 bytes, esto implica tener una gran cantidad de datos por cada imagen, produciendo un mayor tiempo en la transmisión.

El formato JPEG comprime los datos de la imagen sin perder la calidad de la misma, obteniendo de esta manera un menor tamaño por cada imagen. Esto provoca una reducción del tiempo de transmisión, haciéndola más eficiente.

Si se toma en cuenta los resultados obtenidos en la adquisición de imagen mediante los dos tipos de cámaras, es evidente que la utilización de una cámara analógica presenta muchas ventajas en la transmisión y adquisición de la imagen.

Para el desarrollo de un sistema electro-óptico capaz de cumplir con las funciones básicas que poseen equipos existentes en el mercado, es necesario contar con los recursos tecnológicos y financieros, y líneas de investigación enfocadas a las distintas áreas del sistema.

REFERENCIAS

[1] Delgado García J. PhD, Fotogrametría Digital, Depto. Ingeniería Cartográfica, Geodésica y Fotogrametría, Universidad de Jaén, 2010.

[2] Documentación del Master de Integración de Sistemas de Aeronaves. EADS-CASAyUniversidad Carlos III de Madrid, 2009.

[3] Escalante Ramírez B. PhD, Procesamiento Digital de Imágenes, Agosto 2006.

[4] Escobar M. I.,Control de sensores electro-ópticos en aviones no tripulados y tratamiento de imágenes utilizando máquinas de vectores soporte, Junio 2010.

[5] Fernández I. A., Las localizaciones geográficas y la proyección utm (universal transversemercator), Universidad de Valladolid, 2001.

[6] Hecht E., Óptica, Addison Wesley Iberoamericana, Madrid, 2000.

[7] MicroCAM Serial JPEG Camera Module, Data Sheet, Junio 2010.

[8] Servomotores, Caracterísitcas básicas. [En línea]. Disponible: http://robots-argentina.com.ar/MotorServo_basico.htm

[9] Siouris George M., Aerospace Avionics Systems, EEUU, 1993.

[10] UM6 Ultra-MiniatureOrientation Sensor, Datasheet, Agosto 2012.

[11] Woodman, Oliver, An introduction to inertial navigation, University of Cambridge, Agosto 2007. http://robots-argentina.com.ar/MotorServo_basico.htm

Sistema de Visión Artificial para el Análisis de Imágenes de Cultivo basado en Texturas Orientadas

Sotomayor J.; Gómez A.; Cela A.
Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica, Quito, Ecuador
e-mail: (juan.sotomayor; alejandro.gomez; andres.cela) @epn.edu.ec

Resumen: *Dentro del área de la Agricultura de Precisión al obtener la mayor cantidad de información sobre los campos de cultivo se puede optimizar los procesos de siembra, control de malas hierbas y cosecha. Obtener imágenes con cámaras sobre vehículos aéreos no tripulados está siendo actualmente una solución viable para esta tarea, por lo que el procesamiento y análisis de las imágenes aéreas brindaría mayor información al navegar sobre un cultivo. El presente trabajo muestra el desarrollo de un sistema de visión para detectar la dirección de las líneas de cultivos de trigo de imágenes aéreas basándose en el análisis de texturas orientadas; tres estrategias son propuestas para afrontar las diferentes situaciones de las imágenes de cultivo: poca capa vegetal, rodales de mala hierba, variabilidad de la iluminación, escalado de la imagen, entre otras. Las pruebas han sido realizadas con un repositorio de imágenes propio, logrando resultados muy satisfactorios que indican una respuesta en tiempo real, manteniendo un alto nivel de eficiencia en la detección de la dirección de las líneas de cultivo.*

Palabras clave: *Visión por Computador, Procesamiento de Imágenes en Tiempo Real, Visión artificial, Texturas Orientadas, Agricultura de Precisión.*

Abstract: *In the area of Precision Agriculture to get most information about the farmlands is possible to optimize the processes of: planting, weed control and harvesting. Obtaining images with cameras on unmanned aerial vehicles is currently an able solution for this task. The processing and analysis of aerial images would provide more information when we browse on a crop. This paper shows the development of a vision system to detect the direction of the lines of wheat crops using aerial images based on texture analysis oriented, three strategies are proposed to tackle different situations in crop images: little topsoil, weed zones, variability of the illumination, image scaling. The tests were conducted with an own repository of images with highly satisfactory results. Also, the results indicate a response in real time, while the detection of the crop rows direction is keep at high level of efficiency.*

1. INTRODUCCION

Dentro del concepto de Agricultura de Precisión (AP) la ubicación y eliminación de malas hierbas dentro del cultivo es muy importante para garantizar una buena producción. Y además provee información de mucha utilidad para cosechas futuras mediante la utilización de mapas de riego. Adquirir toda la información necesaria de los campos de cultivo mediante vehículos no tripulados, transmitir la información para su posterior procesamiento en un laboratorio con la ayuda de un sistema experto para la toma de decisiones y por último, enviar las acciones correspondientes al campo para ejecutar la aplicación de herbicida de manera óptima, está permitiendo automatizar todas estas, lo que favorece la viabilidad económica.

La visión por computador ha sido de mucha utilidad para el procesamiento de imágenes de cultivo, aunque los logros son poco generalizables pues la variabilidad del campo junto con las condiciones ambientales hace que en cada caso se utilicen técnicas puntuales. En la mayoría de los trabajos desarrollados se realiza la discriminación de la capa vegetal del resto (suelo desnudo, sombras, paja, etc.) o lo que es lo mismo la segmentación, para a continuación distinguir los pixeles de la capa vegetal que pertenecen a mala hierba. En [1] [2] se realiza la segmentación de imágenes de cultivos de cebada en tres etapas: en la primera se transforma la imagen en color (canales RGB) a una imagen en escala de grises obtenida mediante la combinación lineal de los 3 canales. La imagen obtenida se transforma mediante una binarización en una imagen en blanco y negro, donde la capa de vegetación se aísla del resto (suelo desnudo, sombras, paja, etc.). A continuación se identifican los píxeles de la imagen que son parte de las hileras de cultivo lo que permite a su vez identificar el resto de píxeles como parte de las zonas de mala hierba.

En [3] se propone el uso del exceso de verde (ExG) como forma de discriminación de la capa vegetal, suministrando buenos resultados en imágenes que presentaban humedad en el suelo y variación de las condiciones de iluminación.

Por otra parte, en [4] [5] se trabaja en el espacio de color HSI, consiguen una buena segmentación entre la capa vegetal y el suelo en distintos tipos de cultivo con diversas condiciones ambientales.

Los tres procedimientos mencionados son algunos de los más utilizados para segmentación de imágenes de cultivo porque han arrojado buenos resultados en la segmentación de la capa vegetal, suministrando robustez frente a cambios en la iluminación y en el tipo de cultivo. Además es importante resaltar que de esta segmentación depende el éxito de las siguientes etapas de procesamiento de la imagen.

Una vez separada la capa vegetal en las imágenes de cultivo se puede extraer la información necesaria para distinguir la línea de cultivo, el objetivo sería utilizar esa información para el guiado de un vehículo autónomo. Según [6] se aprovecha la orientación de las líneas de cultivo de las imágenes

obtenidas en un plano paralelo al suelo (cenital) tomadas desde un trípode, para poder detectar la posición y la cantidad de píxeles pertenecientes a la línea de cultivo. En otros trabajos como en [7] se enfrentan a las dificultades de procesar imágenes que se adquieren de una cámara montada sobre un tractor. En este caso, la imagen se ve afectada por las vibraciones y el cabeceo propio del vehículo debido a las irregularidades del terreno. Para discriminar la línea de cultivo se utiliza la transformada de Hough con buenos resultados.

En [4] [8] se procesa imágenes adquiridas con una perspectiva aproximadamente cenital desde una cámara montada sobre un vehículo terrestre de pequeña envergadura (un vehículo de aeromodelismo). El procedimiento utilizado en la detección de la línea de cultivo divide a la imagen binaria en 5 franjas horizontales. Seguidamente se utiliza un vector con tantas componentes como columnas tiene la imagen segmentada y en cada componente se almacena el número de píxeles en blanco de la columna asociada. De esta manera se forma una imagen definida en bloques, determinando los centros de estos bloques se obtienen los puntos con los que se puede definir la recta en la que se encuentra el cultivo.

En la imagen obtenida desde un vehículo como un tractor se pueden encontrar distintas zonas como suelo, capa vegetal, cielo, puentes, caminos, líneas de alta tensión, etc. En estos casos claramente se pueden definir tres texturas: la capa vegetal, el suelo, y en algunas imágenes, el cielo. Las dos primeras texturas brindan cierta información acerca del estado del cultivo y del suelo. Al considerar las imágenes como una combinación de texturas es posible realizar la clasificación e identificación de las diferentes zonas que las componen. En [9] se propone combinar varias técnicas ya existentes con el objeto de afrontar los problemas presentes en la segmentación de este tipo de imágenes adquiridas en exteriores y en muchos casos en condiciones extremas y muy variables de iluminación. Y para mejorar aún más la segmentación de la capa vegetal, se propone utilizar un clasificador fuzzy que identifique los diferentes tipos de verde a los que dan lugar los distintos tipos de planta (cultivos y malas hierbas) o incluso las sombras en plantas.

Este trabajo presenta el desarrollo de un innovador sistema de visión por computador basado en el análisis de texturas orientadas que se encarga de obtener el **ángulo de orientación de la estructura línea de cultivo de imágenes aéreas** donde existe: poca capa vegetal, rodales de mala hierba, variabilidad de la iluminación, escalado de la imagen, entre otras.

En la sección 2 se explica el problema y el método propuesto para la detección del ángulo de orientación de las líneas de cultivo. En la sección 3 se presenta el procesamiento de la imagen aérea de cultivo. En la sección 4 se muestra las pruebas y resultados obtenidos. Y en la sección 5 se exponen las conclusiones más relevantes de este trabajo.

2. DESCRIPCIÓN DEL PROBLEMA Y MÉTODO PROPUESTO

Las imágenes aéreas de campos de cultivo de trigo utilizadas en el presente trabajo se adquirieron con un plano paralelo al suelo mediante una cámara Olympus Digital montada sobre un vehículo aéreo no tripulado (UAV).

2.1 Características de las imágenes aéreas

Las imágenes se tomaron a alturas de 30 y 100 metros con una resolución de 4032x3024 píxeles. La Fig. 1 muestra algunos ejemplos en los que se observan ciertas características que deben tomarse en cuenta a la hora de procesar la imagen como son: la homogeneidad, las fronteras con otros cultivos, la poca capa vegetal, las rodadas de tractor, los rodales de mala hierba, las líneas de cultivo no perpendiculares a la base de la imagen, variabilidad en la iluminación, distintas escalas, etc.

Por otra parte, si el elemento de interés a detectar en la imagen es la línea de cultivo, en este caso cultivo de trigo, también se observa que la separación entre líneas es poca alrededor de los 15 centímetros, lo que dificultará la detección de la línea.

Además es importante tener en cuenta que el método de una respuesta en tiempo real.

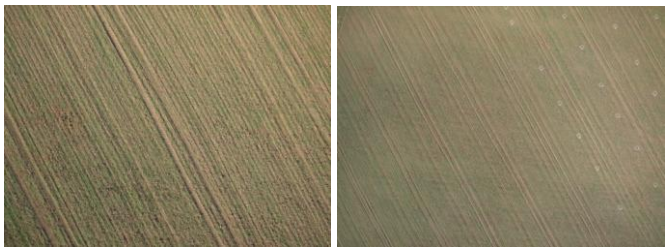


Figura 1. Imágenes capturadas a dos alturas diferentes.

2.2 Método propuesto

El método propuesto en el presente trabajo se basa en el análisis de texturas orientadas y consta del siguiente conjunto de etapas:

- Segmentación de la imagen aérea de cultivo.
- Suavizado de la imagen con un Filtro de Gauss.
- Cálculo del Gradiente de la imagen suavizada.
- Cálculo de los ángulos de orientación sobre una región en la imagen.
- Cálculo del ángulo de orientación total de la imagen.
- Cálculo de la coherencia del flujo de orientación.

3. PROCESAMIENTO DE LA IMAGEN DE CULTIVO

Inicialmente sobre la imagen se realiza un re-escalado a un tamaño de 320 x 320 pixeles, con el fin de reducir la cantidad de información a procesar consiguiendo mayor conectividad entre las líneas de cultivo. Para la segmentación se han hecho pruebas en los dos espacios de color (RGB y HSI) que han

dado buenos resultados en los trabajos citados en las secciones anteriores. Sin embargo en el presente trabajo se ha optado por seleccionar el espacio RGB debido a que tiene mayor robustez ante la variabilidad de la iluminación.

3.1 Segmentación de la imagen

Se obtiene una imagen en escala de grises mediante la combinación lineal de los tres canales RGB [1].

$$\mathrm{Im}G = r \times ChannelR + g \times ChannelG + b \times ChannelB \quad (1)$$

Los valores de las constantes de (1) son los valores propuestos en [1] r = -0.884, g = 1.262 y b = -0.311; obtenidos con un procedimiento de optimización basado en algoritmos genéticos. A la imagen en escala de grises se le aplica un umbral de 0,48 obtenido mediante un procedimiento heurístico, con el fin de conseguir una imagen binaria (imagen en blanco y negro). En los pasos siguientes se analiza el comportamiento de la imagen resultante de multiplicar la imagen en escala de grises, con la imagen binaria, esto se realiza para segmentar y definir de mejor manera sólo las líneas de cultivo. La Fig. 2 describe el procedimiento para conseguir algunas de las imágenes descritas.

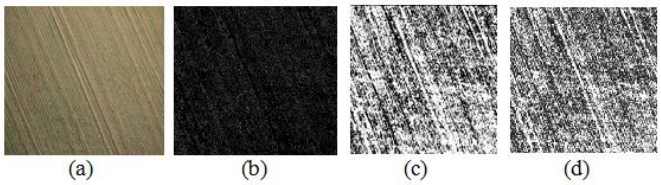


Figura 2. Imagen de cultivo en diferentes etapas: (a) imagen original; (b) imagen en escala de grises; (c) imagen binaria y (d) imagen resultante de la multiplicación binaria-escala de grises.

3.2 Análisis de la imagen como textura orientada

Para determinar el ángulo de orientación de las líneas de cultivo, se considera la imagen como una textura orientada [10]. Así, se aplica a la imagen el gradiente del gausiano (GOG). En concreto la imagen se filtra con un Kernel de convolución gausiano de tamaño 3 con una sigma (σ) igual a 0,5 con el fin de conseguir una imagen con transiciones suaves entre el borde de las líneas y el suelo.
A continuación a la imagen resultante se le aplica el operador gradiente. Se define el vector gradiente de una función $f(x, y)$ como la siguiente expresión:

$$\nabla \mathrm{f}(\mathrm{x}, \mathrm{y}) = \begin{bmatrix} \partial \mathrm{f}(\mathrm{x}, \mathrm{y})/\partial \mathrm{x} \\ \partial \mathrm{f}(\mathrm{x}, \mathrm{y})/\partial \mathrm{y} \end{bmatrix} = \begin{bmatrix} \mathrm{Gx} \\ \mathrm{Gy} \end{bmatrix} \quad (2)$$

Donde la magnitud viene representada por:

$$Mag[\nabla \mathrm{f}(\mathrm{x}, \mathrm{y})] = \sqrt{\left(\partial \mathrm{f}(\mathrm{x}, \mathrm{y})/\partial \mathrm{x}\right)^2 + \left(\partial \mathrm{f}(\mathrm{x}, \mathrm{y})/\partial \mathrm{y}\right)^2} = G \quad (3)$$

Y el ángulo o argumento se expresa como:

$$\theta = \tan^{-1}\left(\partial \mathrm{f}(\mathrm{x}, \mathrm{y})/\partial \mathrm{y} \Big/ \partial \mathrm{f}(\mathrm{x}, \mathrm{y})/\partial \mathrm{x} \right) \quad (4)$$

De esta forma para cada pixel (i, j) de la imagen se pueden calcular las componentes "x" e "y" del vector gradiente, es decir, $G_x(i, j)$ y $G_y(i, j)$. En el vector gradiente cada pixel (i, j) apunta en dirección del incremento máximo posible de la intensidad, y la magnitud del vector gradiente corresponde a la cantidad de cambio de la intensidad en esa dirección. La Fig. 3 presenta una zona de la imagen obtenida usando (2).

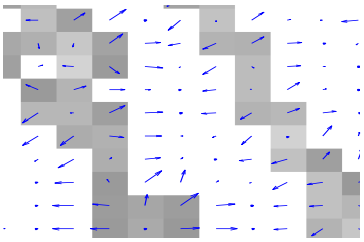


Figura 3. Zona de imagen con Gradiente.

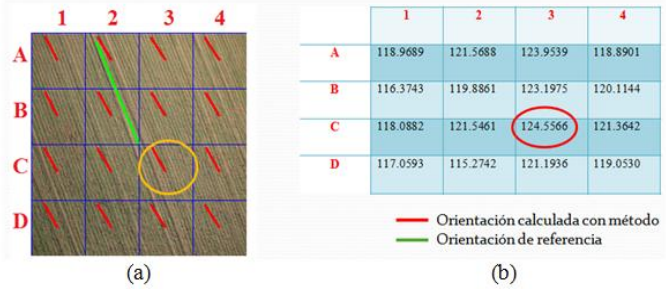
3.3 Obtención del ángulo de orientación

Se puede pensar que el problema estaría resuelto sumando vectorialmente cada uno de los segmentos y hallando la dirección resultante, sin embargo esto no funciona por dos razones: 1) un único segmento puede tomar en un mismo punto dos direcciones diferentes pero opuestas entre sí (θ o θ+π) y 2) incluso si los segmentos tienen direcciones diferentes no opuestas, existe el riesgo de que se contrarresten, en lugar de realizar su aporte a la orientación dominante. Para solventar los problemas anteriores se utiliza el cálculo del ángulo dominante de una textura a partir de la representación polar del gradiente mediante la siguiente expresión:

$$\theta_{\mathrm{nn}} = \tan^{-1}\left(\frac{\sum_{i=1}^{i=N} \sum_{j=1}^{j=N} G(i, j)^2 \sin 2\theta_{(i,j)}}{\sum_{i=1}^{i=N} \sum_{j=1}^{j=N} G(i, j)^2 \cos 2\theta_{(i,j)}} \right) \Big/ 2 \quad (5)$$

En donde se tiene el ángulo de orientación θ para una región de NxN de la imagen, como $\theta_{nn} + (\pi/2)$, ya que el gradiente calculado es perpendicular a la dirección de las líneas. Al no tener una imagen de cultivo uniforme parece adecuado dividirla y trabajar con orientaciones locales. Siguiendo esta estrategia se considera dividir la imagen en 16 regiones de 80x80 píxeles distribuidas de forma matricial, en cuatro filas por cuatro columnas. La Fig. 4 muestra los resultados obtenidos donde se observa el mayor error

cometido, zona C3. Se pinta con verde el ángulo de referencia que es obtenido manualmente (113.12°) y con rojo los ángulos calculados.



	1	2	3	4
A	118.9689	121.5688	123.9539	118.8901
B	116.3743	119.8861	123.1975	120.1144
C	118.0882	121.5461	124.5566	121.3642
D	117.0593	115.2742	121.1936	119.0530

Figura 4. Resultados; (a) imagen dividida en 16 regiones con orientaciones locales; (b) tabla de valores de los ángulos para cada región.

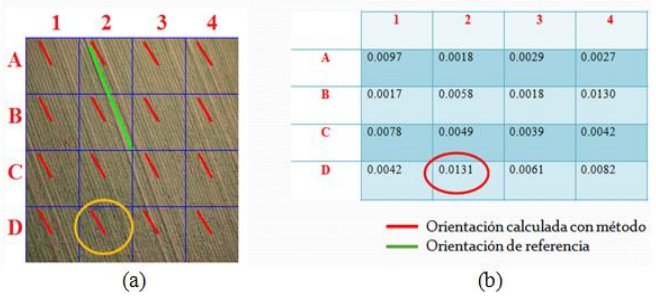
Una vez obtenidas las orientaciones locales se procede a determinar las regiones que mejor representan la orientación global de la imagen, para ello se usa el termino de coherencia del flujo de orientación, que es el valor que permite conocer en una región, la proporción de orientaciones locales que están en la misma dirección de la orientación global.

Así, teniendo el ángulo de orientación θ_{nn} en la región de análisis W, dato obtenido usando (5), y considerando para cada pixel (i, j) el valor de gradiente, en magnitud $G(i, j)$ y en orientación $\theta(i, j)$ se puede calcular el índice de coherencia de flujo usando (6). Se proyecta la magnitud del gradiente con su orientación sobre el vector unitario en la dirección θ_{nn} y se suman todas las proyecciones de cada pixel.

Así, se presenta la siguiente expresión:

$$\rho = \frac{\sum_{(i,j)\in W} \left\| G(i,j)\cos\left(\theta_{nn} - \theta_{(i,j)}\right) \right\|}{\sum_{(i,j)\in W} G(i,j)} \qquad (6)$$

El valor absoluto usando (6) evita la cancelación de los vectores que apuntan en direcciones opuestas. Un índice de coherencia alto indica el mejor candidato (región) para representar la orientación de la imagen, mientras que un índice bajo representa una baja uniformidad en la tendencia de las líneas de cultivo. En la Fig. 5 se muestran los resultados obtenidos y se observa al mejor candidato con un círculo rojo.



	1	2	3	4
A	0.0097	0.0018	0.0029	0.0027
B	0.0017	0.0058	0.0018	0.0130
C	0.0078	0.0049	0.0039	0.0042
D	0.0042	0.0131	0.0061	0.0082

Figura 5. Resultados; (a) imagen dividida en 16 regiones con orientaciones locales; (b) tabla de valores del índice de coherencia del flujo de orientación.

Con la medida de coherencia del flujo de orientación es posible minimizar efectos de líneas de cultivo desordenadas, dando más peso a los ángulos cuyas regiones presentan un valor elevado.

Por último, se han considerado 3 estrategias diferentes para calcular el ángulo global de orientación. El ángulo de la región con mayor índice de coherencia (método E1).

Se debe tener en cuenta que las líneas de cultivo no son exactas, habrá en ciertas regiones segmentos de línea desconectados y seguidos uno de otro, los principios y finales de estos segmentos aportarán información de un ángulo de orientación totalmente diferente al que en realidad se tiene. También es probable que se tenga regiones con líneas continuas y definidas pero con una ligera desviación debido a irregularidades del terreno. Todo lo comentado lleva a plantear una segunda estrategia (método E2), que considera el máximo valor de coherencia del flujo de orientación y todos los valores no inferiores al 70% del máximo. En la Fig. 4 se involucran a las regiones A1, D2 y B4, cuyos ángulos de orientación son: 118.9689°, 115.2742° y 120.1144° respectivamente. Para calcular el ángulo de orientación global de la imagen, los ángulos elegidos se promedian de forma ponderada, considerando como pesos los valores de coherencia del flujo de orientación.

Los campos de cultivo cubren amplias zonas y en las fronteras colindan con otros terrenos, por lo que dentro del campo de visión de la cámara podrían aparecer estas zonas que carecen de interés para el análisis. Para evitar este tipo de problemas se propone una tercera estrategia (método E3). En la Fig. 4 únicamente se toma en cuenta las regiones centrales B2, B3, C2 y C3, cuyos ángulos de orientación son: 119.8861°, 123.1975°, 121.5461° y 124.5566° respectivamente. El ángulo de orientación global de las líneas de cultivo se obtiene mediante el promedio de forma ponderada de los ángulos anteriores.

En la Fig. 6 se visualizan las imágenes con los resultados obtenidos donde: los círculos azules señalan la o las zonas consideradas para cada una de las estrategias, la línea roja representa el ángulo de orientación calculado, y la línea verde representa la orientación de referencia de la imagen.

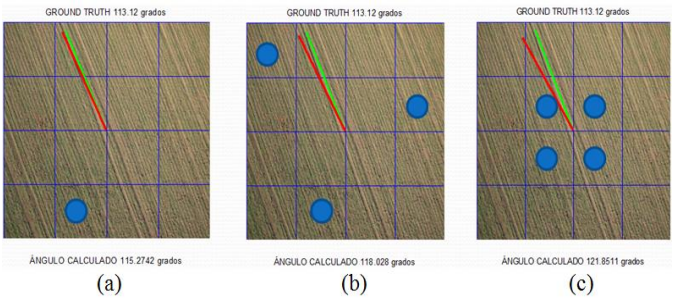


Figura 6. Imagen de cultivo con resultado del ángulo de orientación; (a) imagen con resultado del método E1; (b) imagen con resultado del método E2; (c) imagen con resultado del método E3.

En la Tabla 1 se resumen los errores obtenidos al aplicar las tres estrategias a la imagen junto con el tiempo de procesamiento para cada una de ellas.

Tabla 1. Error absoluto y tiempo de procesamiento

Medida	Método E1	Método E2	Método E3
Error absoluto (°)	-2.15	-4.91	-8.73
Tiempo de procesamiento (ms)	63.7	64.6	63.3

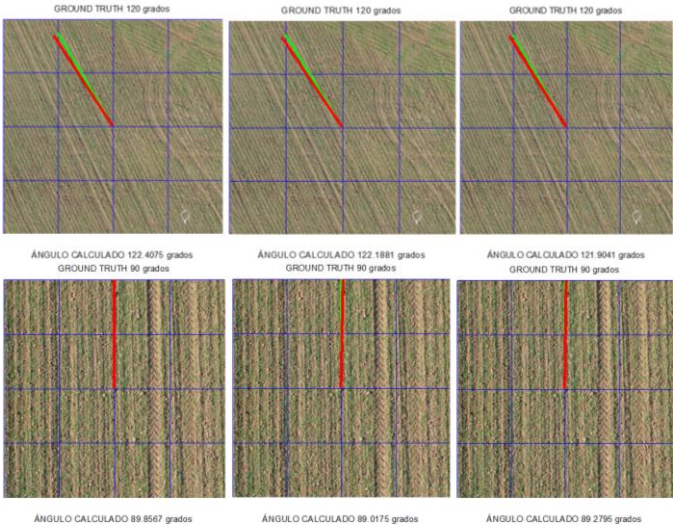
Con el método E1 se aprecia que siendo la región D2 la más representativa con un ángulo de 115.2742°, es la que más se aproxima al ángulo de referencia 113.12°; mientras que al promediar los ganadores más prometedores (método E2) se tiene 118.028°; y al trabajar solo con la parte central (método E3) el ángulo es de 121.8511°, en estos dos últimos casos el error es mayor.

Finalmente el tiempo de cómputo requerido para procesar la imagen con las tres estrategias propuestas es aproximadamente igual a 64 milisegundos en media.

4. PRUEBAS Y RESULTADOS

Se han realizado varias pruebas considerando las tres estrategias propuestas en base a la coherencia del flujo de orientación: método E1 que toma el mejor candidato que representa la orientación global del cultivo en la imagen, método E2 que además de tomar el mejor candidato toma los no inferiores al 70% del valor del mejor candidato y método E3 que únicamente considera la parte central de la imagen.

El repositorio utilizado para las pruebas constaba de 129 imágenes y los resultados para alginas de estas imágenes se muestran en la Fig. 7 donde la línea verde es el ángulo de referencia y la línea roja es el ángulo de orientación calculado.



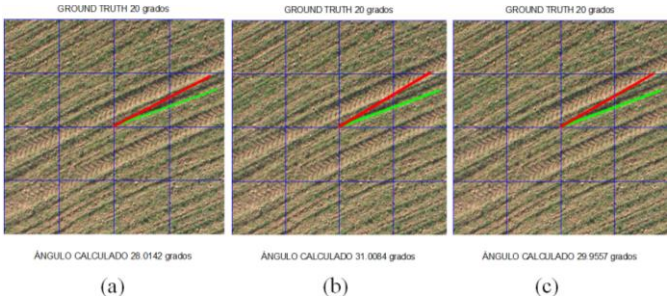


Figura 7. Algunos casos de imágenes de cultivo con resultado del ángulo de orientación; (a) columna de casos aplicando el método E1; (b) columna de casos aplicando el método E2; (c) columna de casos aplicando el método E3.

La Tabla 2 resume los resultados obtenidos de las pruebas realizadas como el Error Medio Cuadrático (EMC) y el tiempo de procesamiento.

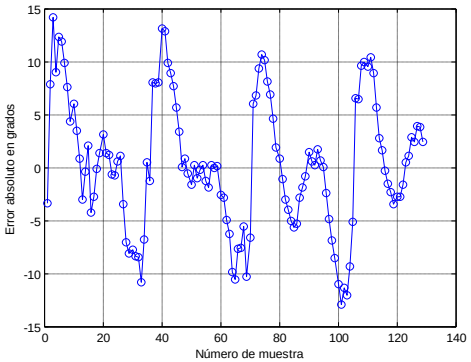
Tabla 2. Resultados para las tres estrategias

Medida	Método E1	Método E2	Método E3
Error Medio Cuadrático (°)	4.90	4.68	5.21
Tiempo de procesamiento (ms)	64.5	65.4	64.0

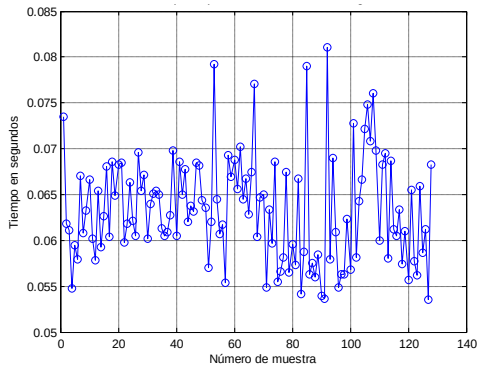
Se observa que las tres estrategias son válidas para solucionar el problema, lo cual es muy satisfactorio si se tiene en cuenta la heterogeneidad de las imágenes del campo de cultivo, por eso en cada imagen por separado funciona mejor una estrategia concreta.

La estrategia E2 es la que muestra una mayor precisión al determinar la orientación del cultivo con un tiempo procesamiento de 65.4 milisegundos.

La Fig. 8a y 8b muestran los resultados de los errores absolutos cometidos y el tiempo de procesamiento para cada imagen con el método E1 respectivamente. En la Fig. 8a se observa que errores absolutos máximos rodean aproximadamente los 15° que son los casos de imágenes en donde los cultivos tienen frontera con otros terrenos. Se observa que los errores absolutos mínimos están cercanos a 0° y además el tiempo de procesamiento máximo por imagen (Fig. 8b) es de aproximadamente 80 milisegundos y el mínimo rodea los 55 milisegundos.



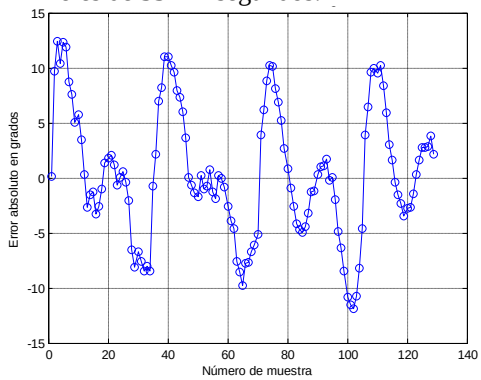
(a)



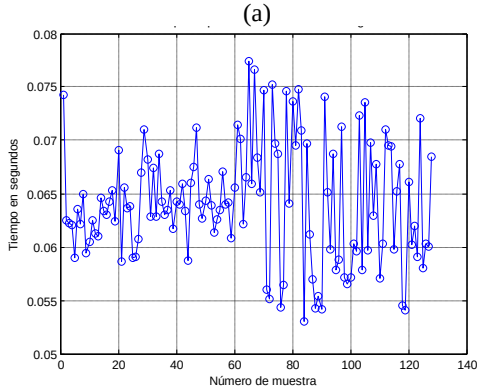
(b)

Figura 8. Resultados de las pruebas método E1: (a) error absoluto en grados y (b) tiempo de procesamiento.

La Fig. 9 muestra los resultados de los errores absolutos y el tiempo de procesamiento para cada imagen usando el método E2. En la Fig. 9a se observa que los errores absolutos máximos rondan aproximadamente los 10° que son los casos de imágenes en donde los cultivos tienen poca capa vegetal. Se observa que los errores absolutos mínimos están cercanos a 0° y que el tiempo de procesamiento máximo por imagen (Fig. 9b) es de aproximadamente 75 milisegundos y donde el tiempo mínimo es de 55 milisegundos.



(a)

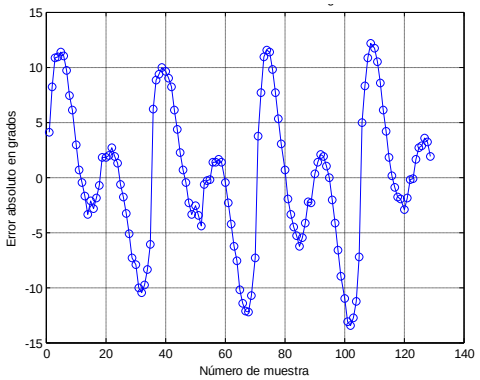


(b)

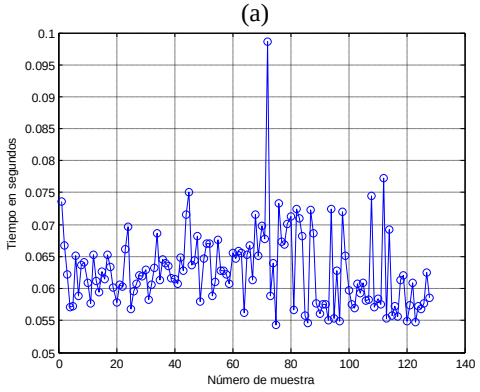
Figura 9: Resultados de las pruebas método E2: (a) error absoluto en grados y (b) tiempo de procesamiento.

Finalmente la Fig. 10 muestra los resultados de los errores absolutos y el tiempo de procesamiento para cada imagen con el método E3. En la Fig. 10a se observa que errores absolutos máximos rodean aproximadamente los 10° que son los casos de imágenes en donde los cultivos tienen rodadas de tractor en la zona central mientras que los errores mínimos cercanos a 0°.

Además el tiempo de procesamiento máximo por imagen (véase Fig. 10b) es de aproximadamente 100 milisegundos y donde el tiempo mínimo es de 55 milisegundos.



(a)



(b)

Figura 10: Resultados de las pruebas método E3: (a) error absoluto en grados y (b) tiempo de procesamiento.

5. CONCLUSIONES

En este artículo se ha presentado un innovador y original sistema de visión por computador que detecta el ángulo de orientación de las líneas de cultivos de trigo en imágenes aéreas basándose en el análisis de texturas orientadas. Los métodos propuestos han conseguido resultados de manera general muy satisfactorios considerando la amplia gama de imágenes utilizadas para las pruebas realizadas; así se tiene un error cuadrático medio aproximado global de 4.93° y un tiempo de procesamiento por imagen global de 64,6 milisegundos en media, lo que permite dar una respuesta en tiempo real consiguiendo procesar aproximadamente 15 tomas por segundo.

Al ser estos primeros resultados muy alentadores es posible, como trabajo futuro, integrar el sistema de visión mediante técnicas de control visual al modelo de un vehículo aéreo autónomo y de esta manera realizar el control sobre la navegación del vehículo para que en su trayecto se pueda adquirir tomas del cultivo de mejor calidad.

REFERENCIAS

[1] X. P. Burgos-Artizzu, A. Ribeiro, A. Tellaeche, G. Pajares. "Optimisation of natural images processing using different evolutionary algorithms", 2008 IEEE CONGRESS ON EVOLUTIONARY COMPUTATION 1-8, pp. 1268-1275. 2008.

[2] A. Tellaeche, X. P. Burgos-Artizzu, G. Pajares, A. Ribeiro. "A vision-based classifier in precision agriculture combining bayes and support vector machines", Intelligent Signal Processing 2007 (WISP) IEEE International Symposium on, pp. 1-6. 2007.

[3] Li Chen, Jin-Guo Zhang, Hai-Feng Su, Wei Guo. "Weed identification method based on probabilistic neural network in the corn seedlings field", Proceedings of the Ninth International Conference on Machine Learning and Cybernetics 3, pp. 1528-1531. 2010.

[4] R. Gottschalk, X. P. Burgos-Artizzu, A. Ribeiro, G. Pajares, A. Sainchez-Miralles. "Real-time image processing for the guidance of a small agricultural field inspection vehicle", 15th International conference on Mechatronics and Machine Vision in Practice (M2VIP08) 2-4, pp. 493-498. 2008.

[5] W. Lanlan, W. Youxian. "Application of support vector machine for identifying single corn/weed seedling in fields using shape parameters", Information Science and Engineering (ICISE) 2010 2nd International Conference on, pp. 1-4. 2010.

[6] A. Ribeiro, C. Fernández-Quintanilla, J. Barroso, M.C. García-Alegre. "Development of an image analysis system for estimation of weed", Proceedings 5th European Conf. On Precision Agriculture (5ECPA), pp. 169–174. 2005.

[7] J. Barreda. "Seguimiento visual de líneas de cultivo", Universidad de Murcia, Tutores: Ángela Ribeiro y Alberto Ruiz, España. 2009

[8] J. Conesa-Muñoz, R. Gottschalk, X. P. Burgos-Artizzu, A. Ribeiro. "Método de detección visual de líneas de cultivo para el control en dirección de vehículos agrícolas", Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA) 2011. Workshop: Computer Vision in Natural Environments: agriculture and forestry (CVNE). Universidad de la Laguna. 2011.

[9] M. Guijarro, G. Pajares, I. Riomoros, P. J. Herrera, X. P. Burgos Artizzu, A. Ribeiro. "Automatic segmentation of relevant textures in agricultural images", Computers and Electronics in Agriculture 75, pp. 75-83. 2011.

[10] A. R. Rao, B. G. Schunck. "Computing oriented texture fields", Graphical Models and Image Processing 53, pp. 157-185. 1991.

Visual Control of an Autonomous Aerial Vehicle for Crop Inspection

Sotomayor J*; Gómez A.*; Castillo A.**

**Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica*
Quito, Ecuador (e-mail: juan.sotomayor@epn.edu.ec; alejandro.gomez@epn.edu.ec)
***Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros Industriales*
Madrid, España (e-mail: alexander.castillo.quintero@upm.es)

Resumen: *El presente trabajo muestra el control visual en un vehículo aéreo no tripulado para navegar a lo largo de las líneas en un campo de cultivo. El sistema integra un método de visión por computador basado en el concepto de texturas orientadas para determinar la orientación de las líneas de cultivo. El ángulo obtenido por el sistema de visión es la referencia para el control visual de modo que conduce a un conjunto de acciones sobre los motores del vehículo, así éste es orientado de forma paralela a las líneas. Finalmente el trabajo es complementado con el diseño y desarrollo de un planificador de trayectorias el cual determina la ruta de vuelo que asegure la cobertura total del campo a ser inspeccionado. Las pruebas realizadas por simulación han llevado a excelentes resultados en la navegación, lo que permite la realización de una retroalimentación en tiempo real desde el sistema de visión mediante el uso de un bucle de control y de esta manera tomar imágenes aéreas de buena calidad.*

Palabras clave: *Control Visual, Visión por computador, Procesamiento de Imágenes en Tiempo Real, Agricultura de Precisión.*

Abstract: *This paper presents a visual control applied on an unmanned aerial vehicle to navigate along the crop rows. The system integrates a computer vision method based on the oriented textures concept to determine the orientation of the crop rows. The angle obtained by the vision system is the reference for visual control, so that leads to a set of actions on the vehicle´s motors, so it is oriented in parallel to the rows. Finally the work is complemented with the design and development of a path planner which determines the flight plan that assures a whole cover of the field to be inspected. Results have been obtained by simulation, obtaining excellent responses in navigation, which allows realizing a real-time feedback from the vision system by using a control loop and in this good quality aerial images are attained.*

Keywords: *Visual Control, UAV, Computer Vision, Image Processing Real-Time, Precision Agriculture.*

1. INTRODUCTION

The concept of Precision Agriculture in the 80 listed in the U.S. as agricultural management adapted to the spatial and temporal variability of the crop. The implementation of a management accurate scheme requires the use of technologies to acquire information using Global Positioning Systems (GPS), sensors, satellite and aerial images together with Geographic Information Systems (GIS) to estimate, evaluate and understand the variations that occur along a crop. The information, which is collected from the crops, was used to evaluate aspects such as density, estimate fertilizers to crops, generation of risk maps and even to predict the level of production of the land.

The collection of technical information requires that experts go to the field and move trying to produce the least possible impact on the crop. Now, commercial crops have several acres (>4ha); therefore, sampling becomes costly and impractical in many cases. For this reason, the usage of an aircraft to perform the inspection or field data collections is highly valuable.

A camera mounted on an autonomous vehicle provides visual information (crops' images) very important for the management of a plot. The information collected in the field is focused on generating different types of maps. Thus, several crop aerial images are captured and integrated with mosaic process [1]; the success of this process depends directly on the quality of input images, i.e. minimum angular rotations and distortions geometric, that in small crops can be corrected with ground reference. However, for large areas of

Manuscript received September 15, 2013; revised July XX, 2013.
This paper has received support from the Government of the Republic of Ecuador by the National Secretary of Higher Education, Science, Technology and Innovation (SENESCYT), under the "Scholarship Program 2011-2012".

crop (> 4 ha) often is considered neither the rotation nor the distortion of the image, due to these techniques cause high percentage of error in the mosaic avoiding taking pictures with a high degree of overlap. It is enough to get images with a 60% overlap in advance (also called longitudinal or horizontal) and 30% lateral overlap (also called vertical). Ultimately, the map generation process becomes a tedious process and therefore costly.

Applications developed in agriculture based on image processing have a common operation whose goal is to separate the vegetal layer from other elements that appear on the image (bare soil, shades, straw, etc.). When it is needed to know the distribution of infestations, the pixels of the vegetation that belong to the weed must be distinguished, so we can find different approaches based on RGB color space [2],[3] and [4] and others working in the HSI color space [5],[6]. By combining various techniques [7] may be performed three textures segmentation (vegetal layer, ground and sky) by a fuzzy classifier unsupervised; in all cases shown, good results approaches for different types of crops and infestations.

By discriminating the vegetation layer in crop images, The information that distinguishes the crop line can be extracted, [8] taking advantage of the vertical orientation lines present in the images obtained in a plane parallel to the ground (overhead) with a tripod, achieving satisfactory results. In an alternative way, images are processed by a perspective acquired using a camera mounted on a tractor. In this case the image is affected by vibration and pitching of the vehicle due to uneven ground. To discriminate the crop line using the Hough transform with good results. In this paper, processed aerial images acquired from a camera mounted on an unmanned aerial vehicle (UAV), specifically a quadrotor Fig. 1. In this case the direction of the crop is obtained with an approach based on oriented texture analysis.

Figure 1. quadrotor

Regarding to the visual control to guide the autonomous vehicles, visual control techniques based on PID controllers can be used [5], [10] these techniques are applied on land vehicles of small size. In order to position and guide the vehicle autonomously so that it moves within the grooves established for transit. It is also possible to use other information from the vision system which [11] describes the self guided, based on a fuzzy visual inspection for commercial vehicles (Citroën C3). The vehicle follows a line painted on the ground with 49 fuzzy rules reaching speeds of 50 km / h. In the case of UAVs, control is very complex, because it is a multivariable nonlinear system naturally very unstable. It is therefore of great importance to have a reliable and robust controller [12]. Working with visual information for controlling speed, obtaining the absolute or relative position serves to control the orientation of the aerial robot. In [13], [14] visual control strategies are applied on a quadrotor to follow a 3-D object that has certain characteristics of color and form. In other work [15], [16] visual information is used to maintain altitude, takeoff and landing. In [17] performed a comprehensive review on the use of data from a computer vision system for controlling the movement of a robot providing taxonomy.

This paper presents the **design and development of visual control of a UAV based on visual discrimination** crop line structure as a guide for navigation. The camera is mounted on the vehicle. It discusses two general tasks: 1) obtaining the orientation angle of line structure in the crop aerial image taken with the camera on board (vision system) and 2) from the angle obtained by the vision system with information UAV sensors, generating the control actions on the UAV to navigate along the crop lines and according to the flight plan generated by a path planner. The ultimate goal is that the UAV cover the entire area to be inspected with a crop line-guided navigation.

The next section describes the problem and the proposed method for guided navigation crop line. Section 3 presents the design and development of visual control. Section 4 shows the experiments and results obtained. To conclude with Section 5 displays the most relevant conclusions of this work.

2. PROBLEM DESCRIPTION AND PROPOSED METHOD

The collection of images from a UAV of extensive cultivation areas to generate risk maps motivates the realization of a system able to obtain images that provide further analysis and integration on a map. Keep in mind that to generate risk maps are needed two types of operation on the pictures: 1) combining all the individual images collected form a single image (mosaic generation) and 2) the discrimination of vegetal layer, then distinguishing between weeds and crop.

2.1 Acquisition of images

When working with large areas crops, the collected images often imply a complicated integration task mosaic process, which is tedious and costly. The images acquired in an ordered manner facilitate the process of mosaic and even allow its full automation. In brief, the hypothesis is that navigation along the lines of crop will take all images with roughly the same orientation crop lines (approximately vertical to the base image or x axis). Thus the crop line can be

used as ground reference to compose the total image and increase the overall accuracy of the generated maps.

2.2 Proposed method

As it was already mentioned the objective of this work is to develop visual control of a UAV to navigate along the crop lines. Fig. 2 displays a global scheme of the developed system. The system consists of two main tasks: 1) obtaining the signal input through perception, treatment and processing of the acquired image and 2) visual control or piloting the UAV taking into account the information provided by perception.

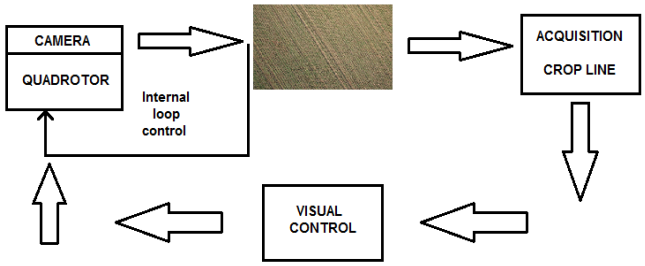


Figure 2. Overall system diagram

Crop lines can provide the necessary information to control the UAV and that it is placed in a flight path parallel to the crop. The idea is to achieve an organized sequence of snapshots and use the element in the crop line for mosaic generation process in order to automate the full procedure and improve the accuracy of the resulting map.

3. VISUAL CONTROL

3.1 Computer vision system

Crop lines orientation are obtained from an approach based on texture analysis where the input is the crop image grayscale. The grayscale image is obtained from the linear combination of the RGB channels, refer to (1). In Fig. 3 shows an example of an crop wheat image and in Fig. 4 one can see the image obtained by applying the linear combination with the values for the constants, r = -0884, g = 1.262 and b = -0311, given by [2] and obtained with an optimization method based on genetic algorithms.

Figure 3. Picture of a wheat crop

$$\mathrm{Im}G = r \times CanalR + g \times CanalG + b \times CanalB \qquad (1)$$

Figure 4. Grayscale image, refer to (1)

Then, the grayscale image is rescaled to a size of 320 x 320 pixels, in order to reduce the amount of information to be processed achieving greater connectivity between the crop rows. After, the image is applied to the gradient of the Gaussian (GOG). Specifically, the image is filtered with a Gaussian convolution kernel with a size 3 sigma (σ) is equal to 0.5 in order to obtain an image with smooth transitions between the edge of the lines and the ground, refer to (2) .

$$\nabla \mathrm{f(x, y)} = \begin{bmatrix} \partial \mathrm{f(x, y)}/\partial \mathrm{x} \\ \partial \mathrm{f(x, y)}/\partial \mathrm{y} \end{bmatrix} = \begin{bmatrix} \mathrm{Gx} \\ \mathrm{Gy} \end{bmatrix} \qquad (1)$$

Thus, for each pixel (i, j) of the image the components "x" and "y" can be calculated. These components belong to the gradient vector, i.e. $G_X(i, j)$ and $G_Y(i, j)$. In each pixel (i, j) gradient vector pointing in the direction of increasing the maximum possible intensity and gradient vector module corresponds to the amount of change of intensity in this direction. Then calculate the dominant angle from a texture polar gradient representation, refer to (3).

$$\theta_{\mathrm{nn}} = \tan^{-1}\left(\frac{\sum_{i=1}^{i=N} \sum_{j=1}^{j=N} G(i, j)^2 \sin 2\theta_{(i,j)}}{\sum_{i=1}^{i=N} \sum_{j=1}^{j=N} G(i, j)^2 \cos 2\theta_{(i,j)}} \right) \Big/ 2 \qquad (2)$$

Where θ is the angle of orientation for a region (W) of NxN pixels in the image, such as $\theta_{nn} + (\pi/2)$, due to the calculated gradient is perpendicular to the direction of the lines.

Thus, taking the angle of orientation θ_{nn} in the region of analysis (W), data obtained by (3), and considering for each pixel (i, j) value gradient in magnitude $G(i, j)$ and orientation $\theta(i, j)$ can calculate the coherence index by flow orientation, refer to (4), i.e. projecting the gradient magnitude

with its guidance on the unit vector in the direction of θ_{nn} and adding all the projections of each pixel.

$$\rho = \frac{\sum_{(i,j)\in W} \left\| G(i,j)\cos\left(\theta_{nn} - \theta_{(i,j)}\right) \right\|}{\sum_{(i,j)\in W} G(i,j)} \quad (3)$$

With the coherence measure flow guidance is possible to minimize the effects of disordered crop lines, giving more weight to the angles which regions have a high value.

Finally, we have considered three different strategies to calculate the overall angle of orientation in the image: 1) the angle of the region with the highest consistency index (E1 method) 2) the angle that is the weighted average of the local orientations maximum consistency index and all values not lower than 70% of the maximum (E2 method) and 3) finally the angle that is the weighted average of the central local orientations of the image (E3 method).

3.2 Flight control with visual feedback

Regarding to control, the aim is to design and develop a control strategy that allows the UAV navigate parallel to the lines of the crop. In this paper the UAV is a quadrotor. Fig. 5 shows the flight control architecture used in this work and based on the dynamic model defined for a quadrotor in [15], [16].

With information from onboard sensors and through interaction cascaded controllers maintaining the stability of the UAV in flight. The parameters of the PID controllers used are summarized in Table 1.

In this paper, both the quadrotor dynamic model as the flight control system has been implemented with Matlab Simulink tool. It was considered the computer control-quadrotor as a single block whose entries are the references to the quadrotor flight path and outputs the position and posture of the same. Consider the set has, among others, the advantage of being more flexible and easily allow the exchange platform flight.

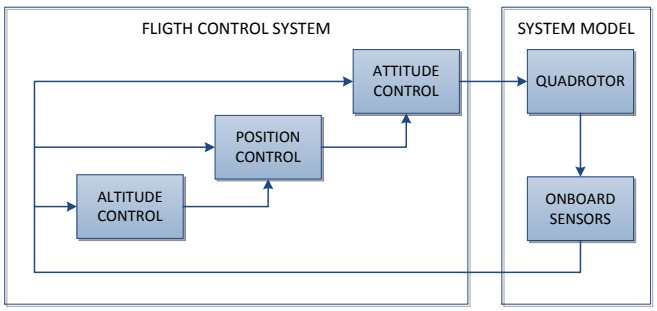


Figure 5. Flight control architecture implemented for the quadrotor

Table 1. Parameters of PID controllers for flight system

Description	Value
Roll PID Control Gain (proportional; derivative; integral)	(1; 0.2; 0.5)
Pitch PID Control Gain (proportional; derivative; integral)	(1; 0.2; 0.5)
Yaw PID Control Gain (proportional; derivative; integral)	(0.5; 0.5; 0.45)
X-axis PID Control Gain (proportional; derivative; integral)	(0.1; 1; 0)
Y-axis PID Control Gain (proportional; derivative; integral)	(0.1; 1; 0)
Z-axis PID Control Gain (proportional; derivative; integral)	(2.5; 5; 0)

It was also necessary to create a virtual world in order to check the response of the entire system, so we developed a 3-D environment in which integrated the quadrotor model and soil crop is represented on the which make the flight. It also included the ability to define and test lighting and a wide range of scenarios. For the generated virtual world was compatible with the implementation of the quadrotor model in Simulink, the block was used 3-D animation (VR Sink). Fig. 6 shows the quadrotor and configuration parameters, while Fig. 7 shows a developed view of the virtual world.

The virtual world is used to simulate the displacement of the quadrotor while the captured image is generated by the camera connected to the body of the UAV. The latter is the information used by the vision system to generate the signal that closes the control loop, in this case visual control. In short, the visual controller is responsible for guiding the quadrotor make for it to maintain its position parallel to the crop rows.

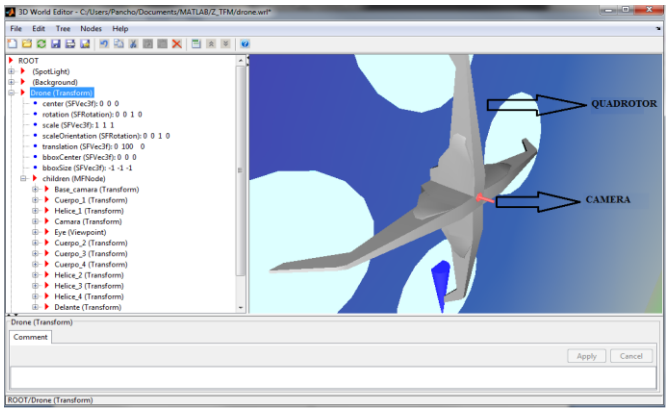


Figure 6. View quadrotor developed in world editor interface 3-D

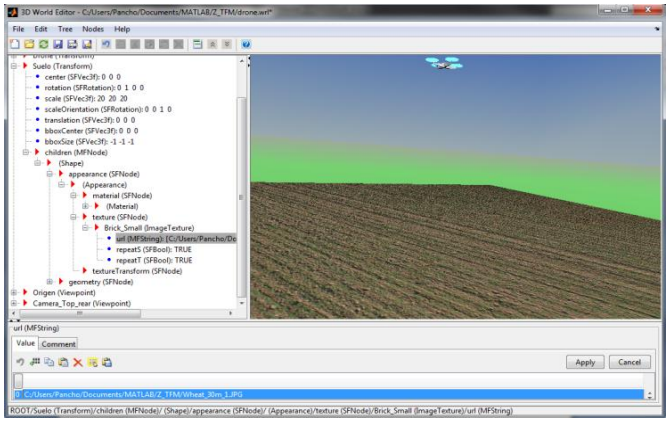


Figure 7. View virtual world developed in the world editor interface 3-D

Summarize the most important characteristics of the proposed control are: eye-in-hand camera configuration, watch and move dynamics and strategy of image-based visual control (IBVS).

Ultimately the strategy used is an indirect visual high level as the corresponding references are generated for the low-level controller (flight control) of quadrotor. Fig. 8 shows the architecture proposed visual control.

For flight control (inner loop) is considered an autonomous flight at low speed (5m/s) with constant height, while visual control (outer loop) the vision system provides the orientation angle of the crop lines over the UAV. Fig. 9 shows how to adjust the relative position of the quadrotor with respect to the lines.

Due to the versatility of the vision algorithm can be used as visual controller a PID classical architecture. Proportional parameters, derivative and integral are: 0.5, 0.5, and 0.45 respectively, these values are adjusted by a trial and error procedure.

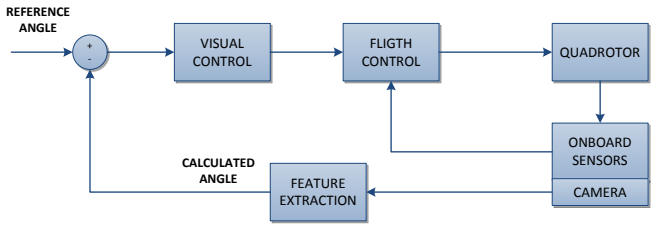


Figure 8. Visual control scheme

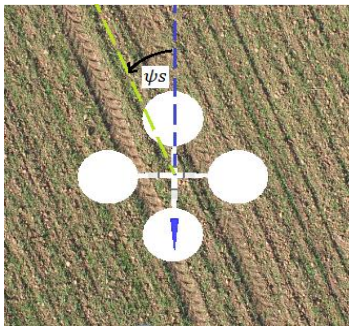


Figure 9. Direction of rotation on the UAV with visual control

It is also included a Kalman Filter to mitigate the effect on sharp turns camera on quadrotor orientation movements because the dynamics of the system. By toolbox of computer vision systems can be implemented in Simulink the Kalman Filter from the value of the variance of the test measurement, which is 0.0051. Thus, the filter can predict future states and minimizing the error optimally. Also sharp variations in the motion of the UAV are minimized.

After guiding the UAV resolved parallel to the lines of culture, a path planning was developed, which ensures that the quadrotor perform a full coverage of the area to explore/inspect.

The developed planner preliminary receives information as input like the shape and dimensions of the field, namely the vertices of a polygon in UTM, and some other features of the camera such as resolution and field of view. With this information the strategy followed by full coverage of the area to be inspected is:

- Location of the quadrotor at constant height, knowing his position (P0).
- Guidance on the initial point P0 of quadrotor placing it parallel to the crop rows.
- Calculation of the path to navigate from the initial position to the final position (Pf), which is the closest border edge of farm field in the direction of the crop rows.
- Rotate and displacement of quadrotor, considering an area of overlap so that the camera does not lose visual information from the crop field in the return path.
- Checking termination condition. If it has reached full coverage of the area to inspect the flight ends and otherwise the process is repeated along the lines of crop to reach the opposite side of the crop edge.

4. EXPERIMENTS AND RESULTS

There have been two types of experiments in order to analyze the system developed from two perspectives: 1) the operation of the vision system and 2) visual control operation with integrated vision system.

In the part of perception, yet tested considering the three proposed strategies based on flow orientation coherence, namely: 1) E1-making the best candidate that represents the overall direction of image cropping, 2) E2-making the best candidate and not lower than 70% of the value of the best candidate, and 3) E3 - considers only the central part of the image. In the study we used a repository of 308 wheat crop images, resolution 1280x960 pixels that collect a variety of possible situations, for example with other land borders, thin vegetation regions, regions with poor stands excess grass or tractor ruts. It should be noted that the crop line spacing is 15 cm. For each of the images from the repository was manually defined ground truth, i.e. actual crop orientation.

Table 2 summarizes the results obtained with the proposed approach showing the mean square error (MSE) with respect to ground truth expressed in degrees and the required processing time in milliseconds.

Table 2. Results for the three strategies

	E1 Method	**E2 Method**	**E3 Method**
Mean Square Error (°)	4.38	4.11	4.15
Average processing time (ms)	64.4	65.2	66.8

Examining Table 2, the E2 method is the best results supplied with MSE of 4.11° and a response time that can process approximately 15 frames per second. In Fig. 10 one can find the error obtained for each image and in Fig. 11 the processing time.

In Fig. 10 shows that the maximum absolute error ranges approximately 15 ° which are the cases in which culture lines are not continuous due to errors in the planting and also shows that the minimum absolute errors are close at 0 °, which is generally satisfactory considering the wide range of images used for the tests.

In Fig. 11 shows that the maximum processing time per image is about 90 milliseconds and the minimum is around 55 milliseconds, which means that it can process at least 11 frames per second.

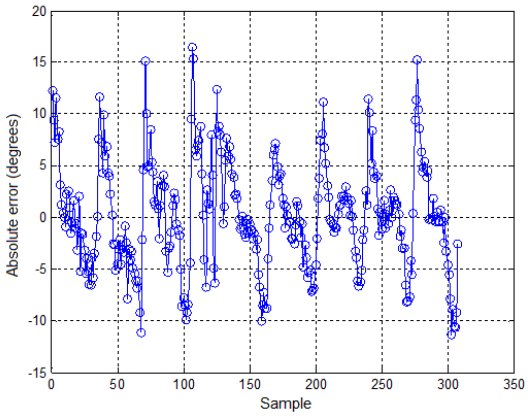


Figure 10. Absolute error using the E2 method

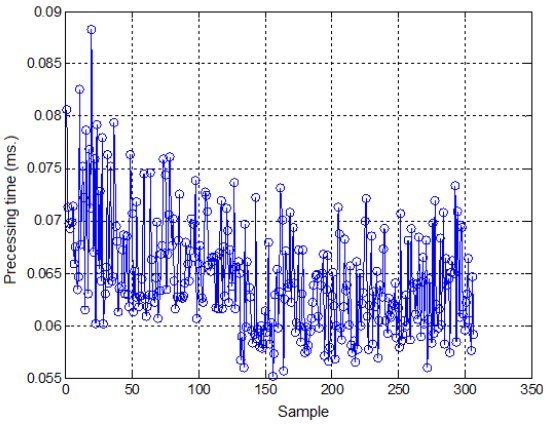


Figure 11. Processing time using the E2 method

As it is stated in E2 method visual control is integrated with the vision system and joined the path planning, two scenarios are analyzed exploring a wheat crop by simulation, one to 30 meters high and over 100 meters in height.

To analyze visual control operation yielded the highest percentages of errors, one on 30 meters has 8.80% and one on 100 meters 9.68%. Even with these errors, due to the actions executed control is achieved by capturing aerial images in a plane parallel to the direction or trend of crop lines. In Fig. 12 are shown on the right side aerial images captured by the camera at two different times and in close proximity.

To reach the end point of the line which coincides with the crop edge rotation phase begins, where it is considered a visual overlap area to prevent the loss of information from the crop field as shown in Fig. 13.

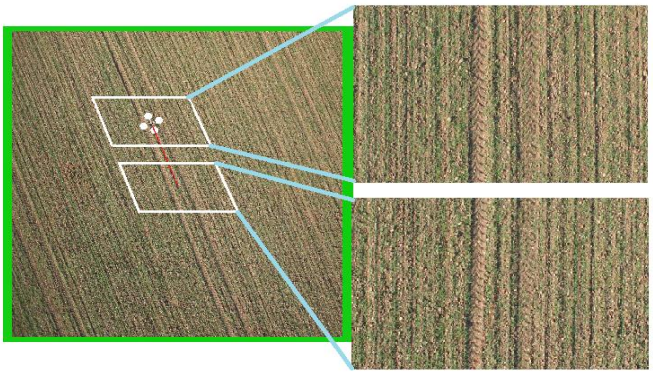


Figure 12. Quadrotor navigation on the farm field. Close view

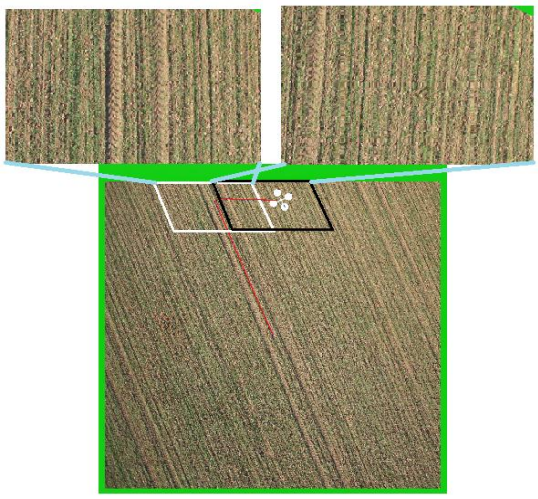


Figure 13. Quadrotor navigation on the farm field. Turning operation overlap

The references generated by the trajectory planner, Fig. 14 shows the path traveled by the quadrotor during scanning of the crop needs successfully completing navigate a constant height parallel to the cultivation lines, plus the properly executed swing action.

Also in Fig. 15, it shows the value of the rotation angles which determine the position of quadrotor during the various stages that execute the scan job. Shows two moments (about 1s, 30s and 43s) coinciding with the beginning of the navigation and the change line rotation in which pitch and roll angles have a high value since this depends on the UAV move in a given direction in the XY plane. In other periods can be seen that these angles take a very low value allowing certain stability in the chamber and therefore the optimal imaging during navigation air parallel to the crop lines.

Figure 14. Quadrotor navigation task on wheat crop field

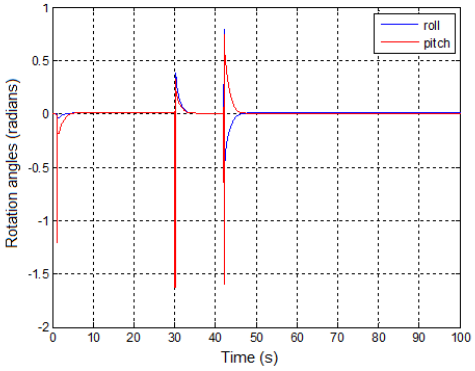


Figure 15. Rotation angles defining the movement in the XY plane

Finally, Fig. 16 shows the visual control response to the reference given in the scan job. In other words, the UAV is maintaining its angle of orientation (yaw) in parallel with the crop lines while it is moving on the plane flight maintaining constant height.

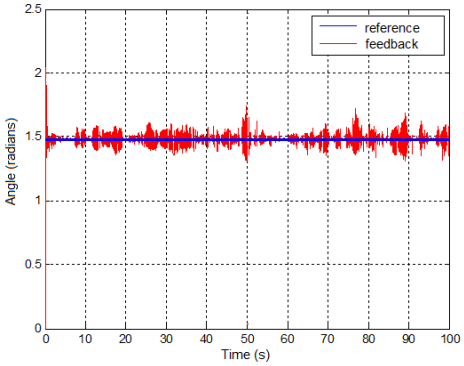


Figure 16. Visual Response Control

5. CONCLUSION

This paper presents the design and development of visual control of a UAV in order to navigate it in the direction of the crop lines through a computer vision system that takes care of finding the angle of orientation of these lines. It has solved the problem of detecting the orientation angle of the wheat breeding lines in an image captured by a camera mounted on a quadrotor. Three methods have been proposed for the vision system (E1, E2, E3) that operate in a similar fashion when there is little vegetal layer regions or areas with weed stands, resulting in a mean square error of approximately 4.21° and a response which can process approximately 15 frames per second. When the quadrotor explores the culture using visual information how navigational guidance has a maximum error of 9.68% on the actual angle presenting crop lines. This result is satisfactory because in the aerial images captured have a perpendicular orientation of the crop rows relative to the base of the image, and this is expected to improve the mosaic process.

Perception problem is more complex in wheat crops because the line spacing seed is small (15cm) which greatly hinders

the operations to perform on the image. Also, the fact that the aerial images taken at high elevations (>100m) depend on the resolution of the camera, but this behavior is desirable and necessary, since the quadrotor has little energy independence so the aerial shot at higher heights implies a greater coverage of the crop in less time. Finally it is proposed a scheme for modular simulation in which all parts are easily modified to replace each module allows simple and thus flight test different platforms, control strategies, or other scanning modes.

The results obtained allow proposing as future work the generalization of vision algorithm corn and sunflower, which have a different planting pattern of wheat with greater separation between crop rows.

REFERENCES

[1] D. Gómez-Candón, S. Labbé, M. Jurado-Expósito, J. M. Peña-Barragán, G. Rabatel, F. López-Granados, "How the spatial resolution can affect the quality of mosaics and assessment of optimum number of tie points necessary to obtain good quality images", RHEA-2011 Robotics and Associated High-Technologies and Equipment for Agriculture, pp. 61-72, 2011.

[2] X. P. Burgos-Artizzu, A. Ribeiro, A. Tellaeche, G. Pajares. "Optimisation of natural images processing using different evolutionary algorithms", 2008 IEEE CONGRESS ON EVOLUTIONARY COMPUTATION 1-8, pp. 1268-1275. 2008.

[3] A. Tellaeche, X. P. Burgos-Artizzu, G. Pajares, A. Ribeiro. "A vision-based classifier in precision agriculture combining bayes and support vector machines", Intelligent Signal Processing 2007 (WISP) IEEE International Symposium on, pp. 1-6. 2007.

[4] Li Chen, Jin-Guo Zhang, Hai-Feng Su, Wei Guo. "Weed identification method based on probabilistic neural network in the corn seedlings field", Proceedings of the Ninth International Conference on Machine Learning and Cybernetics 3, pp. 1528-1531. 2010.

[5] R. Gottschalk, X. P. Burgos-Artizzu, A. Ribeiro, G. Pajares, A. Sainchez-Miralles. "Real-time image processing for the guidance of a small agricultural field inspection vehicle", 15th International conference on Mechatronics and Machine Vision in Practice (M2VIP08) 2-4, pp. 493-498. 2008.

[6] W. Lanlan, W. Youxian. "Application of support vector machine for identifying single corn/weed seedling in fields using shape parameters", Information Science and Engineering (ICISE) 2010 2nd International Conference on, pp. 1-4. 2010.

[7] M. Guijarro, G. Pajares, I. Riomoros, P. J. Herrera, X. P. Burgos Artizzu, A. Ribeiro, "Automatic segmentation of relevant textures in agricultural Images", Computers and Electronics in Agriculture 75, pp. 75-83. 2011.

[8] A. Ribeiro, C. Fernández-Quintanilla, J. Barroso, M.C. García-Alegre. "Development of an image analysis system for estimation of weed", Proceedings 5th European Conf. On Precision Agriculture (5ECPA), pp. 169–174. 2005.

[9] J. Barreda. "Seguimiento visual de líneas de cultivo", Universidad de Murcia, Tutores: Ángela Ribeiro y Alberto Ruiz, España. 2009

[10] J. Conesa-Muñoz, R. Gottschalk, X. P. Burgos-Artizzu, A. Ribeiro. "Método de detección visual de líneas de cultivo para el control en dirección de vehículos agrícolas", Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA) 2011. Workshop: Computer Vision in Natural Environments: agriculture and forestry (CVNE). Universidad de la Laguna. 2011.

[11] M. A. Olivares-Mendez, I. Mellado, P. Campoy, I. F. Mondragon, C. Martinez, "A visual AGV-Urban car using fuzzy control". Automation, Robotics and Applications (ICARA), 2011 5th International Conference on, pp. 145-150. 2011.

[12] L. Mejías, "Control visual de un vehículo aéreo autónomo usando detección y seguimiento de características en espacios exteriores", Universidad Politécnica de Madrid, España. 2006.

[13] I. F. Mondragón, P. Campoy, J. F. Correa, L. Mejias, "Visual model feature tracking for UAV control". Intelligent Signal Processing. WISP 2007. IEEE International Symposium on, pp. 1-6. 2007.

[14] I. F. Mondragón, P. Campoy, M. A. Olivares-Mendez, C. Martinez, "3D Object following based on visual information for unmanned aerial vehicles". Robotics Symposium, 2011 IEEE IX Latin American and IEEE Colombian Conference on Automatic Control and Industry Applications (LARC), pp. 1-7. 2011.

[15] R. W. Beard, "Quadrotor dynamics and control", Brigham Young University. 2008.

[16] J. D. Colorado, "Towards miniature MAV autonomous flight: a modeling & control approach". Universidad Politécnica de Madrid. España. 2009.

[17] B. Siciliano, O. Khatib (eds). Springer Handbook of Robotics. Springer, Berlin, Heidelberg, Ch. 24, pp. 563–583. 2008.

Diseño de un Controlador Fuzzy para Guiado de un Robot Móvil

Cela A.*; Sotomayor J.*; Sánchez F.**

**Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica, Quito, Ecuador*
e-mail: {andres.cela, juan.sotomayor}@epn.edu.ec
*** Universidad Carlos III de Madrid, Facultad de Ingeniería Telemática, Madrid, España*
e-mail: franklin.sánchez@alumnos.uc3m.es

Resumen: *En este trabajó se diseñó un controlador Fuzzy de tipo Mamdani para guiar un robot móvil por un entorno desconocido. El controlador tiene cuatro entradas y dos salidas. Las entradas contienen información de los sensores del robot mientras que las salidas corresponden a la velocidad lineal y angular del robot. Como parte del controlador se diseñó un conjunto de reglas en base a un conocimiento heurístico del comportamiento del robot y de los sensores. El control permite que el robot pueda evitar obstáculos los cuales son detectados por sensores ultrasónicos. El sistema ha sido simulado en el entorno virtual de Player Stage y probado en el robot móvil Amigobot el cual se mueve con tracción diferencial. Los resultados son satisfactorios en vista de que el diseño ha sido implementado en un robot real y se conoce con certeza que el objetivo se ha cumplido plenamente.*

Palabras clave: *control Fuzzy, robótica móvil, control, electrónica*.

Abstract: *In this work a Mamdani Fuzzy controller has been designed and implemented in a mobile robot Amigobot. The goal is guiding the robot in an unknown environment avoiding obstacles. Controller has four inputs and 2 outputs. The inputs are based on sensors information and outputs are angular and lineal velocities of robot. A set of rules has been designed as part of controller. The rules were designed using heuristic information about robots and sensors behavior in different maps. The control allows the robot avoids obstacles, which are detected by ultrasonic sensors. The whole system has been simulated in Player Stage environment and tested in the real mobile robot Amigobot which has differential traction. The aim has been accomplished satisfactory and real robot avoids obstacles.*

Keywords: *Fuzzy logic, robotic movil, control, electronic*.

1. INTRODUCCION

Hoy en día se puede observar el auge de la robótica móvil tanto en el ámbito académico como en el industrial. Aunque este campo es relativamente nuevo, 20 años [1], se observa un desarrollo acelerado debido al auge también de otras disciplinas como la electrónica y la inteligencia artificial. El robot móvil ha pasado de ser un mero elemento de estudio, investigación, competición [2] y hasta juego, a un dispositivo útil en el hogar, como se evidencia en las múltiples empresas que han desarrollado robots móviles de limpieza, [3] y [4]. Tomando este ejemplo, se puede notar que este tipo de robots de servicio necesita un sistema computacional que a más de controlar la tarea para lo cual está diseñado, también realice tareas básicas de guiado y navegación, lo cual se considera como elementos básicos que un robot móvil debe tener.

La terea de guiado se considera esencial ya que describe las trayectorias que el robot puede seguir o las dirección que puede tomar. La navegación permite que robot siga trayectorias reconociendo o no el entorno pero es fundamental que lo haga sin colisionar con obstáculos o paredes.

En este sentido existen diferentes formas de "*sensar*" el entorno. En [5] se propone usar una cámara de tiempo de vuelo y en combinación de odometría se consigue planificar trayectorias para un robot móvil. Otra forma de detectar el entorno es usando sensores infrarrojos. Se han usado los sensores ultrasónicos que son parte del robot móvil Amigobot ya que el software de programación ofrece suficiente soporte para ello.

Para generar las trayectorias existen diferentes métodos, como se muestra en [6] y [7], para los cuales el objetivo es encontrar el camino más eficiente en términos de distancia y seguridad en base a un mapa conocido y representado en grillas, del cual se obtiene un mapa topológicamente organizado.

Para entornos desconocidos esta tarea es irrelevante, y el sistema se limita, como en este trabajo, a la navegación, la cual puede tener objetivos como reconstrucción de entornos [8], [9] conocido como mapeado, con el cual se obtiene un mapa a partir de la información de los sensores y de la odometría del robot, o reconocimiento de objetos [10].

En cuanto a los métodos usados para controlar un robot móvil se encuentran desde los más comunes, como controladores PID [11] o Híbridos [12], hasta los más actuales como algoritmos genéticos y redes neuronales [13].

En este trabajo se propone usar un controlador borroso para controlar las velocidades asociadas al robot móvil, siendo

estas la velocidad lineal y la angular. Además, como se concluye en [14], este controlador es fácil de implementar en robots para aplicaciones de tiempo real. Por otro lado, el controlador borroso puede ser diseñado en programas como Matlab, o X-Fuzzy y luego ser exportado como un archivo de extensión .fis, al cual se puede acceder mediante una llamada desde cualquier compilador en C++. De esta forma se libera carga del programa principal que controla todos los dispositivos del robot y se accede al controlador cuando se tiene datos disponibles de los sensores.

El entorno de simulación Player Stage (PS) permite probar los algoritmos diseñados antes de llevarlos al robot real. Player Stage funciona bajo el sistema operativo Linux y permite conexión con algunos robots como Amigobot o Pionner. La ventaja de este simulador es que permite cargar diferentes mapas y estudiar el comportamiento del robot en cada uno de ellos. Si bien es cierto que los datos de sensores, velocidades, posiciones, etc. tienen ciertos errores por el hecho de ser señales simuladas, el poder de abstracción es tal que permite simplificar los algoritmos con resultados aceptables en el entorno real como se observará en los siguientes apartados de este trabajo.

En el siguiente apartado se muestra el diseño del controlador. En el apartado 3 se estudia las pruebas y en el 4 los resultados del sistema, tanto del real como del simulado. Finalmente en el apartado 5 se dan conclusiones sobre el trabajo.

2. ROBOT AMIGOBOT

2.1 Robot móvil Amigobot

El robot móvil que se usa para las pruebas es el Amigobot, aunque también se realizaron pruebas con el robot Pionner 2. Amigobot tiene un arreglo de 6 sensores ultrasónicos en su parte frontal con alcance de aproximadamente 4 [m]. Se mueve mediante un sistema de tracción diferencial a una velocidad máxima lineal de 0.9 [m/s]. Se usa odometría para determinar la posición relativa del robot mientras está en movimiento. El control total del robot se realiza usando una computadora enlazada a la red LAN mediante la cual se transmiten los datos [15]. La Fig. 1 muestra el robot Amigobot en el cual se aprecian los sensores ultrasónicos.

Figura 1. Robot Móvil Amigobot [15].

2.2 Robot simulado en PS

Una ventaja de PS sobre otro tipo de simuladores como Matlab, es que el robot móvil, cualquiera sea el fabricante, se simula como un objeto rectangular de 3 dimensiones, de esta manera se toma en cuenta aspectos como la inercia al acelerar o frenar, y las esquinas sólidas del robot. Una variable que no se puede simular es la reflexión del ultrasonido en los objetos, sean éstos paredes u obstáculos.

En la Fig. 2 se observa la el robot Amigobot simulado en PS. Se usan todos los sensores del robot, siendo estos S1 y S6 los laterales, S2 y S5 los esquineros y S3 y S4 los frontales.

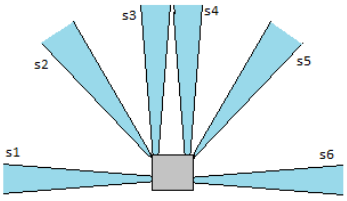


Figura 2. Robot Móvil Amigobot.

Cuando el robot choca con algún objeto la simulación genera un aviso de colisión. El usuario o programador debe detener la simulación para evaluar el problema.

3. DISEÑO DEL CONTROLADOR FUZZY

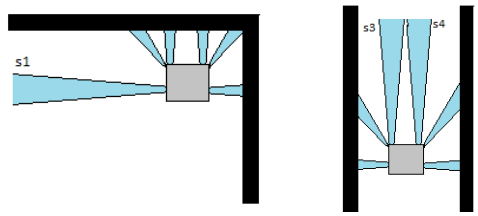
Se ha seleccionado el controlador Fuzzy debido a su operatividad en sistemas MIMO (Multiple Input - Multiple Output) por sus siglas en inglés, lo cual es una ventaja frente al control clásico.

Se ha seleccionado el método de Inferencia Borrosa Mamdani, el cual es usado en casos prácticos y cuando el número de variables lingüísticas es reducido. La estructura de este método se basa en condiciones de mínimos y máximos, lo cual hace que el controlador sea rápido al computar operaciones simples [16].

3.1 Selección de Entradas y Salidas

El controlador Fuzzy permite controlar una planta sin necesidad de tener el modelo de la misma, por esta ventaja se ha decidido usar este método. En un controlador Fuzzy se usa como entradas conjuntos borrosos, los cuales dependen del estado del robot, que en este caso está dado por la distancia medida por los sensores. Las salidas del controlador también son borrosas las cuales se desborrosifican para obtener un valor lógico para el sistema de tracción.

Considerando algunos posibles escenarios, como los de la Fig. 3, se decidió usar 4 entradas y dos salidas borrosas La entrada 1 corresponde al sensor S3, la entrada 2 al sensor S4, la entrada 3 corresponde a la diferencia sensor S2 - sensor S5 y la entrada 4 a la diferencia sensor S1- sensor S6. Las salidas son dos y corresponden a las velocidades lineal y angular del robot.



(a) (b)

Figura 3. Dos Posibles escenarios.

En la Fig. 3(a) se observa el robot cuando está frente a la esquina de una pared. La información de los dos sensores frontales S3 y S4 es prácticamente la misma porque están a la misma distancia de la pared, lo mismo sucede con los sensores esquineros, de tal manera que la información útil para salir la ofrece el sensor S1. La ubicación del los sensores laterales permite saber si existe espacio disponible para que el robot vaya en esa dirección, en cuyo caso el robot gira para tomar ese camino. Para reducir el número de entradas, y por consiguiente de reglas del sistema, no se usan los sensores S1 y S6 como entradas independientes, sino que su diferencia genera una sola entrada. De la misma forma se usa la diferencia de los sensores esquineros S2 y S5 como otra entrada, En la Fig. 3 (b) se observa que los sensores frontales aportan mucha información acerca del entorno ya que indican si existe espacio disponible para navegar, por tanto, los sensores S3 y S4 se usan como entradas independientes. De esta manera quedan definidas las 4 entradas del controlador Fuzzy que se muestran en la Fig. 4 y que se describen a continuación.

3.2 Diseño de los Conjuntos de Entrada y Salida Borrosos

La entrada 1, con información del sensor S3, tiene dos subconjuntos trapezoidales, Pequeño (P) y Grande (G), y un triangular central, Mediando (M), esto permite que en el rango central de la variable exista más dinámica que en los extremos, en donde se hacen correcciones extremas como: girar lo máximo a derecha o izquierda, etc.
La entrada 2, con información del sensor S4 tiene la misma configuración y distribución que la entrada 1.
La entrada 3, con información de la diferencia entre S2 y S5, se distribuye en tres subconjuntos similares a la entrada 1 y con nombres Negativo (N), Cero (Z) y Positivo (P).
La entrada 4, con información de la diferencia entre los sensores S1 y S6 tiene dos subconjuntos trapezoidales ubicados de tal forma que el valor borroso es prácticamente un valor digital, así cuando el subconjunto Negativo (N) está en 1 el subconjunto Positivo (P) está en 0, y viceversa.
La escala de las variables está dada en cm ya que es la escala usada en el simulador.

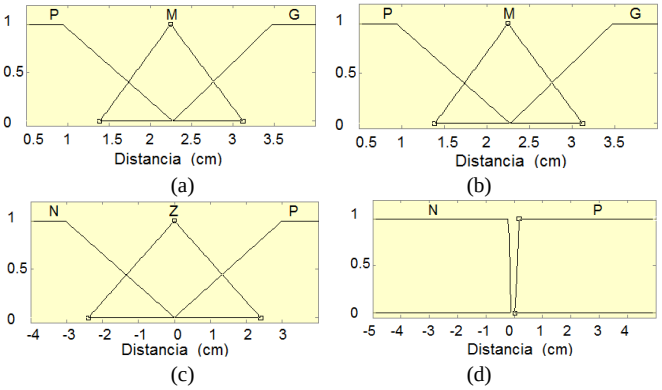


Figura 4. Conjuntos borrosos de las entradas. (a) Entrada 1,sensor frontal S3. (b) Entrada 2, sensor frontal S4. (c) Entrada 3, S2 - S5. (d) Entrada 4, S1- S6.

En la Fig. 5(a) se observa el conjunto borroso de la salida1 correspondiente a la velocidad lineal, para este conjunto la escala varía de 0 a 1 donde 0 es el mínimo y 1 el máximo.
Se han considerado 7 subconjuntos triangulares distribuidos uniformemente a lo largo del rango de la velocidad lineal, lo que permite tener una salida dinámica para todo su rango. Los subconjuntos borrosos son: Pequeño Pequeño (PP), Pequeño Medio (PM), Medio (M), Grande Medio (GM) y Grande Grande (GG).
Para la salida de velocidad angular se consideró los siguientes 7 subconjuntos borrosos: Negativo (N), Negativo Medio (NM), Cero Negativo, (ZN), Cero (Z), Cero Positivo (ZP), Positivo Medio (PM) y Positivo (P), como se observa en la Fig. 5 (b). Los subconjuntos N y P se ubican en los extremos del rango de la velocidad angular mientras los demás subconjuntos están alrededor de la velocidad nula. Con esta distribución la dinámica de la respuesta es tal, que es más sensible a cambios alrededor de 0 (rad/s).

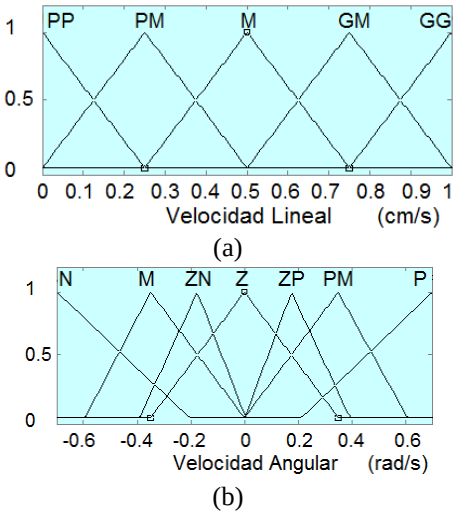


Figura 5. Conjuntos borrosos de las salidas. (a) Velocidad Lineal. (b) Velocidad angular.

3.3 Diseño del Conjunto de Reglas de Control.

Un controlador Fuzzy se considera completo si existe un número de reglas igual a todas las posibles combinaciones de los subconjuntos de entrada. En este trabajo el controlador completo debería tener 2^{11} reglas, pero considerar este número de reglas resulta intratable, razón por la cual, usando conocimiento heurístico del sistema se diseñó el conjunto de reglas basado en la forma general expresada en (1), en donde ***x*** representa los conjuntos y **A** los subconjuntos de entrada y, ***y*** los conjuntos y **B** los subconjuntos de salida.

$$\mathrm{R}^{\mathrm{i}}:\mathrm{IF\,f}(x_1 \text{ es } \mathrm{A}_1,\ldots,x_k \text{ es } \mathrm{A}_{\mathrm{k}}) \text{ THEN } y=\mathrm{B}_{\mathrm{m}} \quad (1)$$

Un subconjunto está activo cuando el valor del rango de la variable da un valor borroso mayor de cero. Así por ejemplo, si la entrada 1 tiene un valor de 2 (cm), los subconjuntos activos son P y M con valores borrosos de 0.2 y 0.6 aproximadamente. El subconjunto G tiene un valor borroso de cero, por tanto está desactivado.

La regla debe considerar la combinación de las diferentes entradas. Es decir, pueden estar activados subconjuntos de diferentes entradas al mismo tiempo, como se indicó arriba. Esta combinación se define mediante el operador **AND** que reemplazando en (1) da como resultado

$$\begin{aligned}\mathrm{R}^{\mathrm{i}}:\mathrm{IF\,f}(&x_1 \text{ es } \mathrm{A}_{\mathrm{k}} \text{ AND } x_2 \text{ es } \mathrm{A}_{\mathrm{k}} \text{ AND } x_3 \text{ es } \mathrm{A}_{\mathrm{k}} \ldots \\ &\text{AND } x_4 \text{ es } \mathrm{A}_{\mathrm{k}}) \text{ THEN } y_1 \text{ es } \mathrm{B}_{\mathrm{m}}, y_2 \text{ es } \mathrm{B}_{\mathrm{m}}\end{aligned} \quad (2).$$

El set de reglas se compone de 26 combinaciones en donde ***I1*** es la variable lingüística ***Distancia 1*** de la entrada 1, ***I2*** es la variable ***Distancia 2*** de la entrada 2, ***I3*** es la diferencia de distancias correspondiente a la entrada 3, ***I4*** es la diferencia de los sensores laterales, entrada 4. Las premisas básicas para la construcción del set de reglas son las siguientes:

- Si los sensores frontales no detectan obstáculos se debe incrementar la velocidad lineal para avanzar rápido.
- Si se detecta un obstáculo con el sensor frontal izquierdo a media distancia se debe girar un poco, "medio", hacia la derecha y reducir la velocidad lineal. Si el obstáculo se detecta más cerca, se reduce aun más la velocidad lineal.
- Si se detecta obstáculo muy cerca, "pequeña distancia", por los dos sensores frontales, entonces se debe reducir la velocidad lineal al mínimo y girar hasta encontrar vía libre de obstáculos con alguna otra entrada, como I3 o I4, así el robot gira hacia la derecha o la izquierda buscando una vía alternativa.
- Si ningún sensor detecta obstáculos, se debe dar prioridad a las reglas de los sensores delanteros, es decir, se incrementa la velocidad lineal para ir más rápido hacia adelante.

De ésta forma se define el set de 26 reglas que se muestra a continuación.

R[1]: If (***I1*** is P) and (***I2*** is P) and (***I3*** is N) and (***I4*** is N)then(***V*** is PP)(***W*** is N)
R[2]: If (***I1*** is P) and (***I2*** is P) and (***I3*** is N) and (***I4*** is P)then (***V*** is PP)(***W*** is P)
R[3]: If (***I1*** is P) and (***I2*** is P) and (***I3*** is Z)and (***I4*** is N) then (***V*** is PP)(***W*** is N)
R[4]: If (***I1*** is P) and (***I2*** is P) and (***I3*** is Z) and (***I4*** is P) then (***V*** is PP)(***W*** is P)
R[5]: If (***I1*** is P) and (***I2*** is P) and (***I3*** is P) and(***I4*** is N) then (***V*** is PP)(***W*** is N)
R[6]: If (***I1*** is P) and (***I2*** is P) and (***I3*** is P) and (***I4*** is P) then (***V*** is PP)(***W*** is P)
R[7]: If (***I1*** is M) and (***I2*** is P) and (***I3*** is N) then (***V*** is PM)(***W*** is N)
R[8]: If (***I1*** is M) and (***I2*** is P) and (***I3*** is Z) then (***V*** is M)(***W*** is PM)
R[9]: If (***I1*** is M) and (***I2*** is P) and (***I3*** is P) then (***V*** is GM)(***W*** is P)
R[10]: If (***I1*** is M) and (***I2*** is M) and (***I3*** is N) then (***V*** is M)(***W*** is N)
R[11]: If (***I1*** is M) and (***I2*** is M) and (***I3*** is Z) then (***V*** is GM) (***W*** is Z)
R[12]: If (***I1*** is M) and (***I2*** is M) and (***I3*** is P) then (***V*** is M)(***W*** is P)
R[13]: If (***I1*** is M) and (***I2*** is G) and (***I3*** is N) then (***V*** is GG)(***W*** is NM)
R[14]: If (***I1*** is M) and (***I2*** is G) and (***I3*** is Z) then (***V*** is GG)(***W*** is Z)
R[15]: If (***I1*** is M) and (***I2*** is G) and (***I3*** is P) then (***V*** is M)(***W*** is NM)
R[16]: If (***I1*** is P) and (***I2*** is M) and (***I3*** is N) then (***V*** is GM)(***W*** is N)
R[17]: If (***I1*** is P) and (***I2*** is M) and (***I3*** is Z) then (***V*** is M)(***W*** is NM)
R[18]: If (***I1*** is P) and (***I2*** is M) and (***I3*** is P) then (***V*** is PM)(***W*** is P)
R[19]: If (***I1*** is G) and (***I2*** is G) and (***I3*** is N) then (***V*** is GG)(***W*** is N)
R[20]: If (***I1*** is G) and (***I2*** is G) and (***I3*** is Z) then (***V*** is GG)(***W*** is Z)
R[21]: If (***I1*** is G) and (***I2*** is G) and (***I3*** is P) then (***V*** is GG)(***W*** is P)
R[22]: If (***I1*** is G) and (***I2*** is M) and (***I3*** is N) then (***V*** is M)(***W*** is PM)
R[23]: If (***I1*** is G) and (***I2*** is M) and (***I3*** is Z) then (***V*** is GG)(***W*** is Z)
R[24]: If (***I1*** is G) and (***I2*** is M) and (***I3*** is P) then (***V*** is GG)(***W*** is PM)
R[25]: If (***I1*** is G) and (***I2*** is P) then (***V*** is M)(***W*** is ZP)
R[26]: If (***I1*** is P) and (***I2*** is G) then (***V*** is M)(***W*** is ZN)

Las Variables lingüísticas de velocidades lineal y angular están representadas por ***V*** y ***W*** respectivamente. Sus subconjuntos borrosos toman el nombre de consecuentes.

Las 6 primeras reglas ayudan a tomar decisiones extremas cuando el robot está en circunstancias similares a las de la Fig. 3(a). En esos casos prima el giro del robot a altas velocidades angulares y bajas lineales.

Las siguientes 18 reglas trabajan en el rango dinámico de las combinaciones de los subconjuntos de entrada premiando o castigando las velocidades dependiendo de las distancias a los obstáculos.

Finalmente las dos últimas reglas, premian el sistema con altas velocidades lineales cuando los sensores frontales detectan que existe "*Grande*" vía libre para la navegación, como se indicó en las premisas básicas.

3.4 Características del Controlador

El Operador **MIN** se usa en casos prácticos para ajustar el Operador **AND** que une las variables lingüísticas. El operador **MIN** se define por

$$\mathrm{MIN}(\mathrm{r_x} = \min(\mathrm{A}_1, \mathrm{A}_2, \ldots \mathrm{A}_{\mathrm{k}})) \quad (3).$$

El ajuste del operador **AND** se hace con el operador producto (**PROD**) en casos de estudio de estabilidad de sistemas [17]. La agregación, es decir cuando dos o más reglas actúan sobre el mismo consecuente, se resuelve por máximos y se define por

$$\mathrm{r_y} = \max\left[\mathrm{V}^{\mathrm{N}}{}_{\mathrm{i}=1}\ \mathrm{R}^{\mathrm{i}} \rightarrow B_m\right] \quad (4).$$

En donde **V** es el conjunto de valores de los consecuentes repetidos de la Regla $\mathbf{R}^{i}$ correspondiente [17].
La *desborrosificación* está dada por el método de los centroides de acuerdo al método Mamdani [16].
El controlador borroso tiene las siguientes características:

- Entradas: 4
- Salidas: 2
- Método Borroso: Mamdani
- Método AND: Mínimos
- Método OR: Máximos
- Agregación: Máximos
- Desborrosificación: Método de los Centroides.

Una vez diseñado el controlador borroso se simuló los resultados de las variables lingüísticas en el programa X-Fuzzy, el cual permite diseñar, simular y evaluar controladores borrosos.
Las superficies de control son superficies que indican las tendencias del sistema de control para corregir el error. Estas están formadas por una combinación de variables de entrada y salida.
En la Fig. 6 se observa dos superficies de control de tres dimensiones. El plano xy lo forman los rangos de las entradas I1 e I2, mientras el eje Z lo forma la Velocidad Lineal del robot. Se puede observar que, si las distancias de los sensores frontales son pequeñas, alrededor de 1, la velocidad es 0.5 (cm/s), pero conforme las distancias crecen la velocidad aumenta de acuerdo a lo establecido por las reglas. La máxima velocidad lineal se da cuando los dos sensores miden una distancia de 4 (cm), lo que indica que el robot puede navegar de frente a máxima velocidad ya que no existe obstáculos.
Existen tres superficies adicionales que se forman por la combinación de las entradas I3 e I4 y la velocidad lineal.

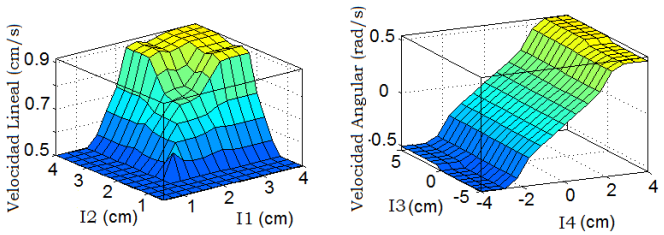


Figura 6. Superficies de control. (a) Respecto de la Velocidad Lineal. (b) Respecto de la Velocidad Angular

En la Fig. 6(b) se observa la superficie de control formada por las entradas I3, I4 y la velocidad angular del robot. Cuando la entrada I4 indica 4 (cm) de distancia hacia la izquierda, lo que sugiere que no hay obstáculos, el robot gira en esa dirección para tomar ese camino. Lo mismo ocurre si no hay obstáculos al otro lado del robot.
La velocidad angular sigue una distribución aproximadamente lineal a lo largo de las distancias de I3 e I4, lo cual indica que el robot no gira cuando no hay distancia disponible, como ocurre en I3=0 (cm) e I4=0 (cm).

Existen otras tres superficies de control formadas por las entradas I1, I2 y la velocidad angular que tienen similares características.
El lazo de control se representa en la Fig. 7, en donde se observa que el Punto de consigna (Set Point) forma parte del controlador Fuzzy el cual entrega las variables de velocidad. Las 4 entradas del controlador vienen de la realimentación de los 6 sensores que miden la distancia del robot al obstáculo. Los bloques Vel. L y Vel A. acondicionan y escalan las variables de velocidad lineal y angular respectivamente, al rango de velocidades del robot real. El sistema es MIMO.

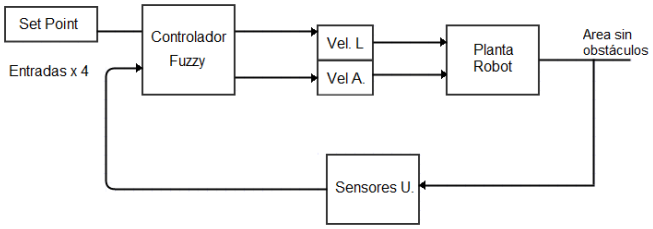


Figura 7. Lazo de control del Robot (Sistema MIMO).

4. PRUEBAS Y RESULTADOS

Las pruebas de simulación se realizaron en PS en el cual se pueden enlazar el programa de simulación, el controlador borroso, el robot y el mapa, todo bajo SO Linux.
El programa de simulación tiene como tarea principal leer los sensores y enviar los datos del controlador al robot. Además, como tareas adicionales se pinta puntos en el mapa en el lugar de detección de los sensores y en el camino recorrido por el robot.
La Fig. 8 muestra la simulación realizada, en donde se observa el robot como un cuadro de color rojo y el alcance de sus sensores en color celeste. El recorrido realizado por el robot se ha pintado como finos puntos de color rojo, mientras los obstáculos detectados se han pintado en puntos de color azul. Este mapa es básico ya que tiene un camino en forma de rectángulo en donde la detección de los sensores frontales es la máxima en todo el recorrido excepto en las esquinas. Se puede observar que en el camino el robot no ha colisionado con ningún obstáculo o pared. El tiempo de recorrido del robot en una vuelta fue de 50 (s).

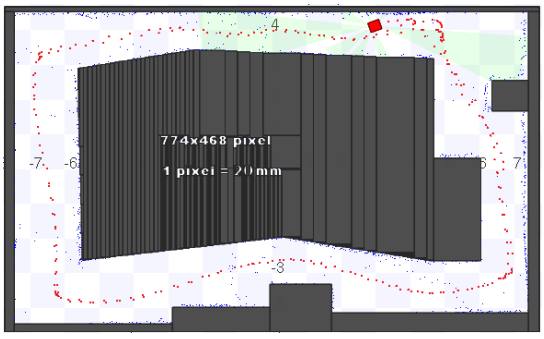


Figura 8. Pruebas en Mapa 1.

La Fig. 9 muestra el recorrido en el Mapa 2, el cual tiene una dificultad mayor que el anterior por las curvas en "U" y las paredes multiformes. La acumulación de puntos rojos, especialmente en las curvas, indica que el robot ha reducido su velocidad para no colisionar. El tiempo para recorrer este mapa fue de 85 (s). No se detectaron colisiones.

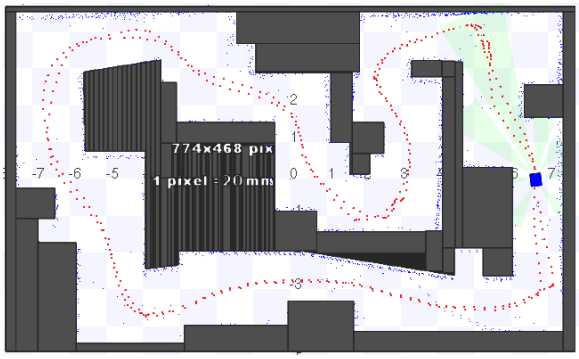


Figura 9. Pruebas en Mapa 2.

Una vez pasadas las pruebas de simulación se enlazó el sistema al robot Amigobot. Para ello se hizo un programa de conexión, similar al de simulación pero que no pinte puntos en el recorrido ni en la detección de los sensores.

El entorno de recorrido fue el Laboratorio de Inteligencia Artificial de la Politécnica de la Universidad de Alcalá.

Se observó que los sensores no detectan objetos muy delgados como las patas de una mesa. También tienen problemas de detección de objetos con una superficie de textura lisa.

El recorrido se hizo poniendo como obstáculos objetos grandes como: cajas, personas, otro robot, etc. y los resultados fueron satisfactorios ya que el robot no colisionó con estos objetos sino que tomó rutas alternativas.

5. CONCLUSIONES

Las conclusiones de este trabajo son las siguientes:

Un controlador Fuzzy permite usar varias entradas para supervisar varias salidas. La ventaja de este tipo de control en sistemas MIMO es evidente ya que un controlador clásico PID necesitaría un lazo para cada variable de control, mientras que el Fuzzy trabaja con todas las variables como un solo bloque. Por otro lado, el algoritmo del control Fuzzy ocupa más memoria de programa, y por tanto necesita más tiempo para ejecutarse que un PID, aunque este punto es irrelevante cuando el controlador está funcionando sobre un procesador de altas prestaciones como los de un computador personal.

Aunque este tipo de control trabaja fácilmente con sistemas MIMO hay que considerar seriamente el número de variables de entrada y salida, pues esto repercute en el número de reglas que definen el sistema. Además, mientras más variables tiene el controlador mayor debe ser el conocimiento del comportamiento del sistema ya que las reglas se construyen en base a este conocimiento.

Un set de reglas completo ofrece mayor dinámica de control por cuanto para cada combinación entrada-salida ofrece una regla consecuente, pero un set de reglas reducido, como en este caso, permite eliminar algunas combinaciones entrada-salida que no se cumplen, y por tanto simplifica la solución del controlador.

La facilidad que presentan programas como XFuzzy o Matlab para diseñar controladores Fuzzy permite que el desarrollo de este tipo de sistemas sea más rápido. Además, con la ayuda del software de simulación Player Stage se observa de primera mano los resultados del controlador sobre el robot. Aunque el medio de simulación tenga sus limitaciones, es una herramienta poderosa para depurar el algoritmo, empezando desde las entradas, salidas, reglas y método de control Fuzzy.

Los sensores ultrasónicos ofrecen una buena solución para medir distancia entre los rangos de 2 (cm) a 4 (m). En el robot Amigobot están distribuidos de tal forma que se puede "observar" obstáculos alrededor de los 180º frontales del robot a una distancia mínima de 5 (cm) y máxima de 4 (m).

Antes de seleccionar el tipo de sensores para medir distancia se debe conocer el tipo de obstáculos a detectar. Por ejemplo, los sensores del robot Amigobot no detectan obstáculos con superficies de textura lisa como madera lacada, baldosa lisa, etc., u obstáculos de grosor inferior a 3 o 4 (cm), pero ofrecen buenos resultados al detectar personas, cajones u otros robots similares.

Este trabajo demuestra que un controlador Fuzzy puede ser implementado exitosamente en un robot móvil para controlar su dinámica. Aunque podría tener limitantes en el tipo de procesador que use dicho robot, pues, como se ha comentado anteriormente, este tipo de control necesita muchos recursos computacionales que en robots básicos no se tiene.

Una forma de implementar este tipo de control en procesadores básicos de 8 bits y 20 [MHz], es diseñar los conjuntos borrosos como líneas que se intersecan, o hacer tablas de valores. El problema con la primera opción es la dificultad de la programación, el de la segunda es la cantidad de memoria, aun más, en cualquiera de los dos casos, si se desea mover los subconjuntos se los debe reprogramar. Por eso se recomienda usar procesadores de mayores prestaciones como los ARM o Core 2 DUO, o como en este trabajo, usar un computador que realice el trabajo de control y que envíe datos al robot mediante una red inalámbrica.

Aunque el objetivo de este trabajo fue diseñar el controlador para que el robot evite obstáculos, un aporte adicional se puede observar en las Figs. 8 y 9, en dónde, con los puntos azules que representan los obstáculos detectados por los sensores, se puede reconstruir el mapa del entorno recorrido. Este avance proyecta un trabajo futuro de reconocimiento y reconstrucción de entornos desconocidos, también llamado mapeado o "mapping" por su traducción al inglés, en el cual se usará el controlador Fuzzy diseñado en este trabajo.

REFERENCIAS

[1] A. Aksamovic, M. Hebibovic, y S. Konjicija, «Similarities in development of digital computers and mobile robots», en *IEEE EUROCON 2009, EUROCON '09*, 2009, pp. 740-745.

[2] J. Green y S. Plumb, «Mobile robot competition», en *AFRICON, 2011*, 2011, pp. 1-6.

[3] «Robot aspirador LG: limpieza automática de tu casa | LG España». [En línea]. Disponible en: http://www.lg.com/es/robot-aspirador. [Accedido: 13-sep-2013].

[4] «Aspiradoras | SAMSUNG». [En línea]. Disponible en: http://www.samsung.com/latin/consumer/home-appliances/vacuum-cleaners/?cid=ec_search_google_alwayson_ha_20130304. [Accedido: 13-sep-2013].

[5] H. Fukai, Y. Mitsukura, y G. Xu, «The calibration between range sensor and mobile robot, and construction of a obstacle avoidance robot», en *2012 IEEE RO-MAN*, 2012, pp. 737-742.

[6] A. R. Willms y S. X. Yang, «An efficient dynamic system for real-time robot-path planning», *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 36, n.º 4, pp. 755-766, 2006.

[7] N. Sedaghat, «Mobile robot path planning by new structured multi-objective genetic algorithm», en *2011 International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, 2011, pp. 79-83.

[8] R. Ouellette y K. Hirasawa, «Mayfly: A small mapping robot for Japanese office environments», en *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2008. AIM 2008*, 2008, pp. 880-885.

[9] P. Mirowski, R. Palaniappan, y T. K. Ho, «Depth camera SLAM on a low-cost WiFi mapping robot», en *2012 IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*, 2012, pp. 1-6.

[10] K. Balasubramanian, R. Arunkumar, J. Jayachandran, V. Jayapal, B. A. Chundatt, y J. D. Freeman, «Object recognition and obstacle avoidance robot», en *Control and Decision Conference, 2009. CCDC '09. Chinese*, 2009, pp. 3002-3006.

[11] T.-Y. Kuc y W.-G. Han, «Adaptive PID learning of periodic robot motion», en *Proceedings of the 37th IEEE Conference on Decision and Control, 1998*, 1998, vol. 1, pp. 186-191 vol.1.

[12] X. Lai, S. Zhu, y W. Wu, «Research on driving wheel control of cleaning robot based on fuzzy adaptive tuning PID», en *International Conference on Mechatronics and Automation, 2009. ICMA 2009*, 2009, pp. 535-540.

[13] X. Yang y M. Meng, «A neural network approach to real-time motion planning and control of robot manipulators», en *1999 IEEE International Conference on Systems, Man, and Cybernetics, 1999. IEEE SMC '99 Conference Proceedings*, 1999, vol. 4, pp. 674-679 vol.4.

[14] S. H. Lian, «Fuzzy logic control of an obstacle avoidance robot», en , *Proceedings of the Fifth IEEE International Conference on Fuzzy Systems, 1996*, 1996, vol. 1, pp. 26-30 vol.1.

[15] «Map Making». [En línea]. Disponible en: http://www8.cs.umu.se/~c06aag/mapmaking/. [Accedido: 13-sep-2013].

[16] K. Michels, *Fuzzy control: fundamentals, stability and design of fuzzy controllers*. Springer, 2006.

[17] Cela, A, Yebes, J.J, Arroyo, R, Bergasa, L.M, Barea, R, y López, E, «Complete Low-Cost Implementation of a Teleoperated Control System for a Humanoid Robot», *Sensors 2013*, vol. 13, pp. 1385-1401, ene. 2013.

Instrument Landing System Lateral Beam Guidance System Based on Sliding Mode Control Technique

Gómez A.*; Sotomayor J.*

Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica, Quito, Ecuador

e-mail: (alejandro.gomez; juan.sotomayor)@epn.edu.ec

Resumen:*En este artículo se muestra un entorno basado en Matlab-Simulink para el análisis de un sistema de aterrizaje de un avión que puede ser utilizado para el aterrizaje automatizado o asistido. Se describe la construcción de un modelo matemático que posteriormente es simulado para validar su funcionamiento. Finalmente, el controlador por modos deslizantes es implementado y se analiza el comportamiento del sistema controlado incluso tomando en cuenta perturbaciones.*

Palabras clave:*SlidingMode Control, Control No-lineal, Teoría de Control, Aeronaves.*

Abstract:*This paper shows an environment based on Matlab-Simulink to analyze a landing system of an aircraft which can be used for automated or assisted landing. Firstly, it is described the development of a mathematical model which is simulated to validate its performance. Finally, the sliding mode controller is implemented, and the behavior of the controlled system is analyzed even taking into account disturbances.*

Keywords:*Sliding Mode Control, Non-linear control, Control Theory, Aircrafts.*

1. INTRODUCTION

The ground-based elements of an Instrument Landing System (ILS) are composed of a localizer transmitter, a glideslope transmitter and marker beacons. These provide the azimuth, vertical and distance signals respectively (Fig. 1). [1]

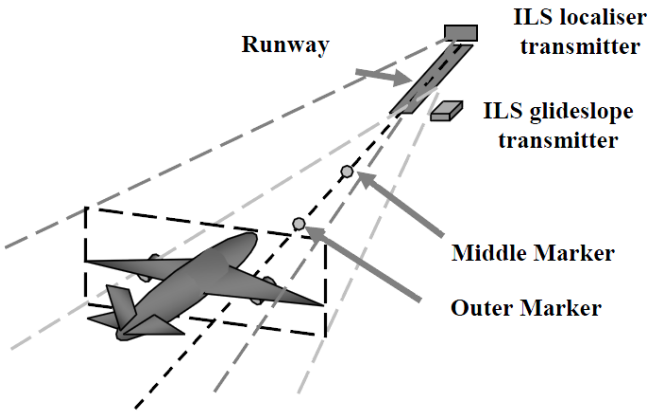


Figure1. ILS Diagram

On the aircraft isinstalled a localizer antenna, a glideslope antenna, an ILS receiver unit and a marker beacon antenna and receiver. The position of the aircraft relative to the localizer and glideslope is displayed on an indicator in the cockpit and is used to land safely. [1]

The Localizer System

"The localizer transmitter is located at the end of the runway which the aircraft is approaching. It transmits on a given frequency in the band 108 MHz to 112MHz. The signals radiate to the left and right of the centre line of the runway as shown in Fig. 2. The signalto the left is modulated by a 90 Hz component while the corresponding frequency for the signal on the right is 150 Hz. The two signals overlap in the middle"[1]. If there is predominance either of 90Hz or 150 Hz means that the aircraft out of the centre and it must move to the right or the left until reach the centre of the line.

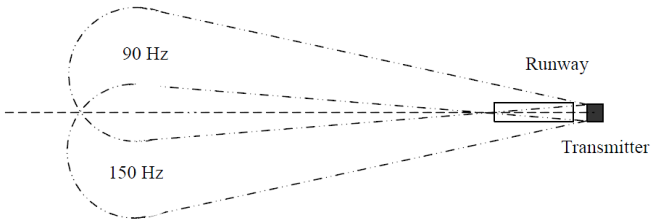


Figure2.The pattern of the localizer

The Glideslope System

The glideslope transmitter is an antenna located near the point of touchdown on the runway and transmits on a given frequency in the range 329.3 MHz and 335.0 MHz. The radiated signal pattern is similar to that of the localizer but

provides vertical guidance relative to a descent path (Fig. 3) [2]

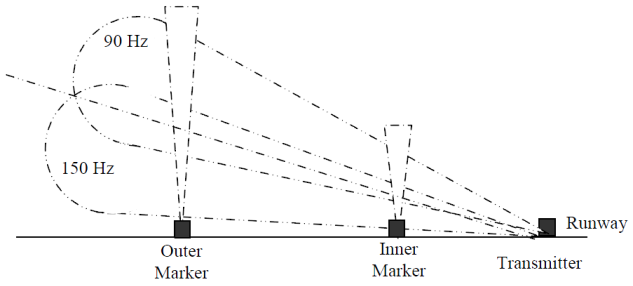


Figure3. The Glideslope system

Marker Beacons

Marker beacons give 75 MHz signals beamed vertically. These markers display the distance of the aircraft to the runway, so that the speed during the descent can be controlled by the pilot.

2. MODELLING THE SYSTEM

The system of the lateral beam guidance is shown in Fig. 4.

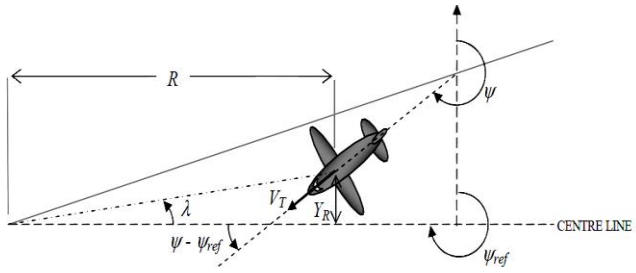


Figure4.System of Lateral Beam Guidance

The system which is going to be used in Matlab is presented in Fig. 5.

The system contains three feedback loops. The outermost loop involves the feedback of the heading angle. The other two loops provide feedback signals needed to ensure that the whole system has a reasonable dynamic response.

The reference is the heading angle, it could be provided by the pilot or as an electronic signal from other system in the aircraft [1].

The transfer function that relates the roll rate (p) to the aileron deflection angle delta (δ) is stated in the equation (1).

$$\frac{p(s)}{\delta(s)} = \frac{Ka}{sTa + 1} \quad (1)$$

Where Ka is a gain factor and Ta is a time constant associated with this simplified model. Similarly, the relation between δ and the roll angle PHI Φ is the equation (2).

$$\frac{\emptyset(s)}{\delta(s)} = \frac{Ka}{s(sTa + 1)} \quad (2)$$

The aileron servo is represented by the following relationship:

$$\frac{1}{sT + 1} (3)$$

The three transfer functions described above can be converted into a state space form as in the equations (4), (5) and (6).

$$\frac{dp}{dt} = -\frac{1}{Ta}p + \frac{Ka}{Ta}\delta (4)$$

$$\frac{d\emptyset}{dt} = p (5)$$

$$\frac{d\delta}{dt} = -\frac{1}{T}\delta + \frac{1}{T}e_s \quad (6)$$

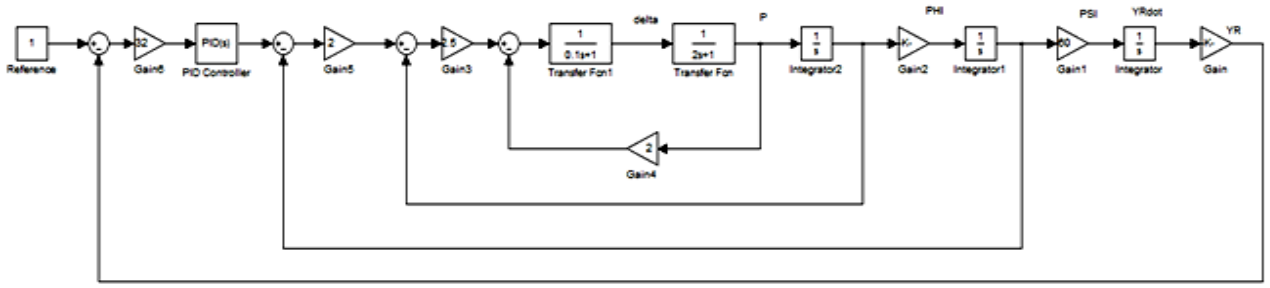


Figure 5.Block diagram of an aircraft directional-control system

The roll angle and the heading angle are related by the following equation:

$$\frac{d\psi}{dt} = \frac{g}{Vt}\emptyset (7)$$

where g is the gravitational constant. The structure of the feedback system is defined by the following equations:

$$e_s = K_v e_v - K_r p (8)$$

$$e_v = K_d e_d - \emptyset (9)$$

$$e_d = \psi c - \psi (10)$$

Fig. 5 shows that the Lateral Beam Guidance has a reference which is compared with the output. This comparison is made by a PI controller. The form of this PI controller is the following:

$$C = G_c\left(1 + \frac{k_i}{s}\right)(\lambda_{ref} - \boldsymbol{\lambda}) (11)$$

In order to estimate the angular error between the aircraft and the centre line of the runway, the equation (12) is going to be considered.

$$\sin(\boldsymbol{\lambda}) = \frac{\boldsymbol{Y_R}}{\boldsymbol{R}} (12)$$

For small angle approximation, the equation (12) becomes equation (13)

$$\lambda = \frac{Y_R}{R} \quad (13)$$

From the equations described above is possible to obtain the state space model for this system. The first step is to define the state vector, which is the following:

$$\begin{bmatrix} \lambda \\ Y_R \\ \psi \\ \emptyset \\ p \\ \delta \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}$$

Once defined the vector state, the second step is to take the derivative of every state of the vector.

$$\begin{bmatrix} \dot{\lambda} \\ \dot{Y}_R \\ \dot{\psi} \\ \dot{\emptyset} \\ \dot{p} \\ \dot{\delta} \end{bmatrix} = \begin{bmatrix} \dot{x_1} \\ \dot{x_2} \\ \dot{x_3} \\ \dot{x_4} \\ \dot{x_5} \\ \dot{x_6} \end{bmatrix}$$

Finally, equations (1) to (13) are combined and the state variables are substituted and reduced into the derivative equations.

The first state $\boldsymbol{\lambda}$is obtained from the derivative of the equation (13)

$$\dot{\lambda} = \frac{\dot{Y_R}}{R}$$

where $\dot{Y_R} = Vt\psi$, thus:

$$\dot{\lambda} = \frac{V_t}{R}\psi$$

or

$$\dot{x_1} = \frac{V_t}{R}x_3 \quad (14)$$

Since

$$\dot{Y_R} = Vt\psi$$

Then

$$\dot{x_2} = Vtx_3 (15)$$

From equation (7)

$$\dot{x_3} = \frac{g}{Vt}x_4 (16)$$

$$\dot{\emptyset} = p$$

$$\dot{x_4} = x_5 (17)$$

From equation (4)

$$\dot{x_5} = -\frac{x_5}{Ta} - \frac{K_a}{T_a} \quad (18)$$

From (6), (8), (9) and (10)

$$\dot{x_6} = -\frac{x_6}{\tau} + \frac{KvKd}{\tau}x_3 - \frac{k_v}{\tau}x_4 - \frac{k_r}{\tau}x_5 \quad (19)$$

Equations (14) to (19) are the state space model of the system that is going to be used in Matlab.

3. SIMULATION OF THE SYSTEM

Fig. 6 to 11 show the behavior of the system when the reference is 15 degrees. $\boldsymbol{\lambda}$does not have equilibrium point, it grows without control as it is shown in the Fig. 6. Fig.7 shows that the altitude of the plane is not set to 6000m, which is the initial condition for this. Fig.8 describes the heading angle of the plane. This angle should be zero. It is seen in the figure that there is a jump from zero to 15 degrees and slowly it comes back to zero after 450 seconds. Fig.11 can be seen a deflection in the aileron during 10 seconds. This leads to a

change in the roll rate and in the roll angle. As it is shown in Fig.9 the roll angle has a transient during 10 seconds reaching a maximum value of 4.4 degrees after that the plane tries to be stable at zero degrees again. In general the behavior of the system is unstable even if the system does not present any external disturbance. The next step is to include a controller to make the system stable.

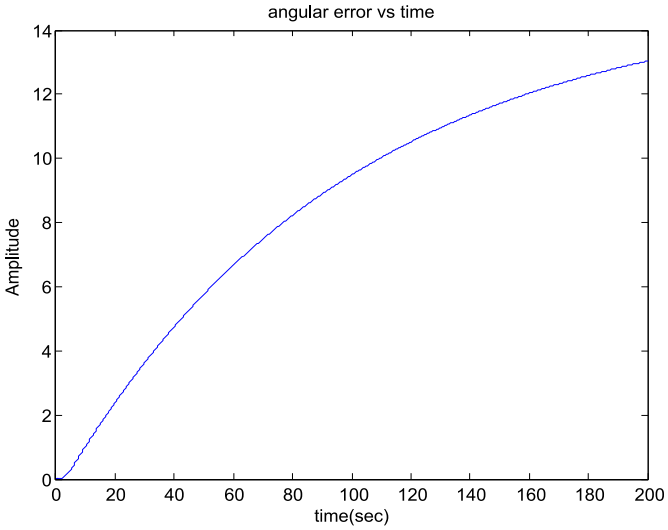


Figure6. The error angle**λ** when the reference is 15 degrees without controller

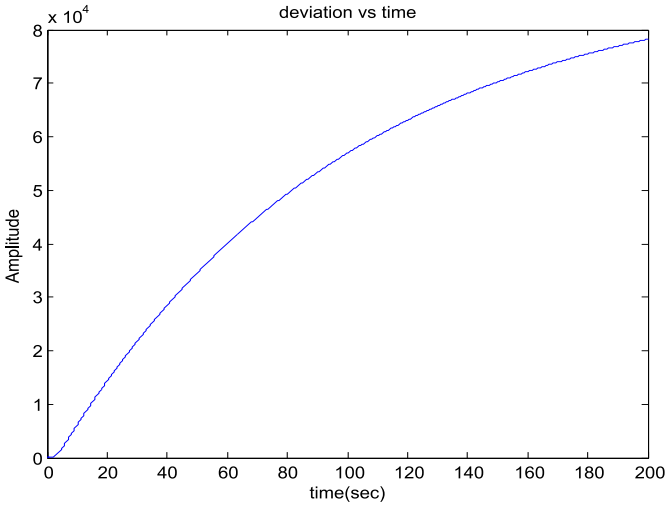


Figure7. The deviation YR when the reference is 15 degrees without controller

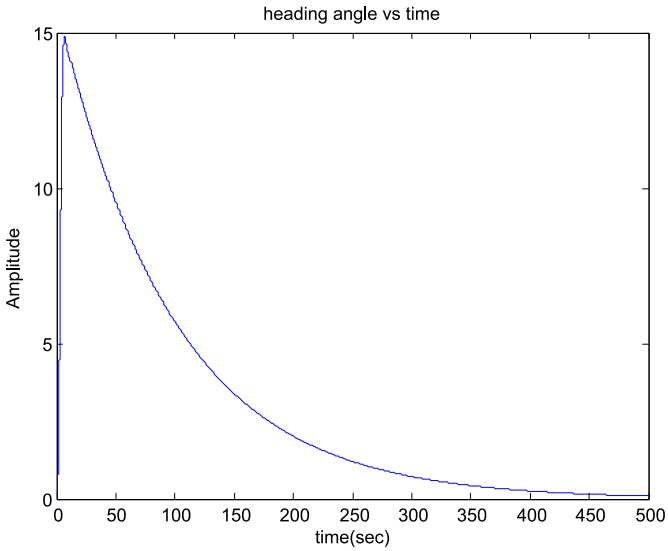


Figure8. The heading angle ψ when the reference is 15 degrees without controller

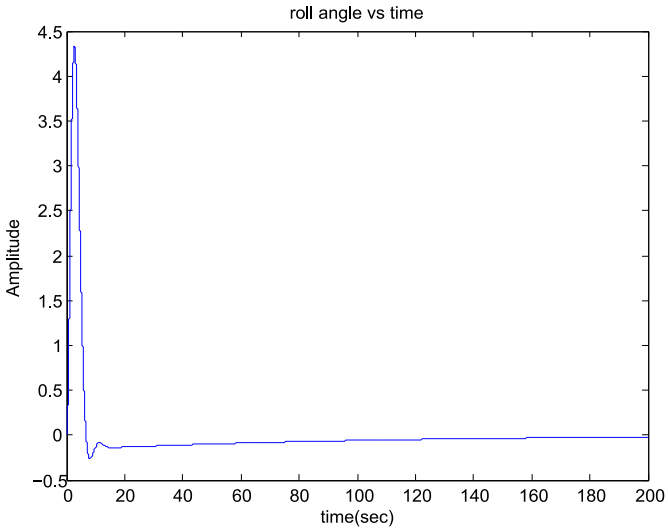


Figure 9. Roll angle ψ when the reference is 15 degrees without controller

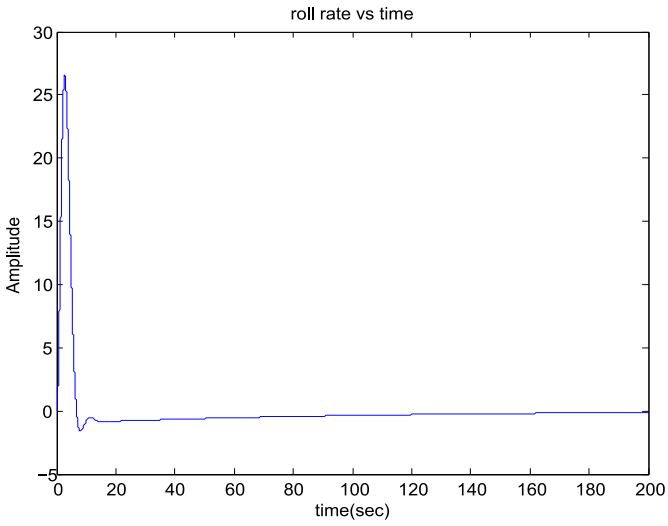


Figure10. Roll rate p when the reference is 15 degrees without controller

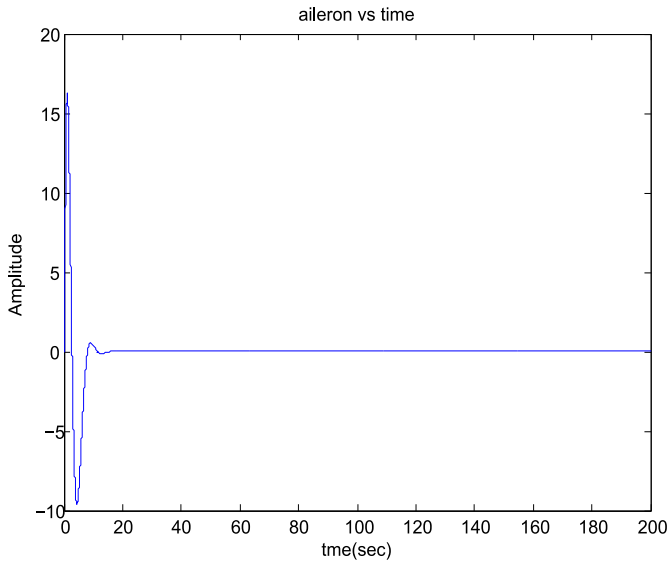


Figure11. aileron deflection δ when the reference is 15 degrees without controller

4. LONGITUDINAL DYNAMICS

In principle, the system becomes stable by using a PI controller when the reference is a small value. Now, the longitudinal dynamics of the aircraft is going to be considered. See in table 1how the range of an aircraft on ILS approach changes with time:

Table1. Variation in time

t(s)	R(m)
0	6000
10	4675
20	3640
30	2835
40	2210
50	1720
60	1340
70	1045
80	810
90	630
100	495

The linear interpolation routine is going to be implemented to determine the range values that fall between these data points.

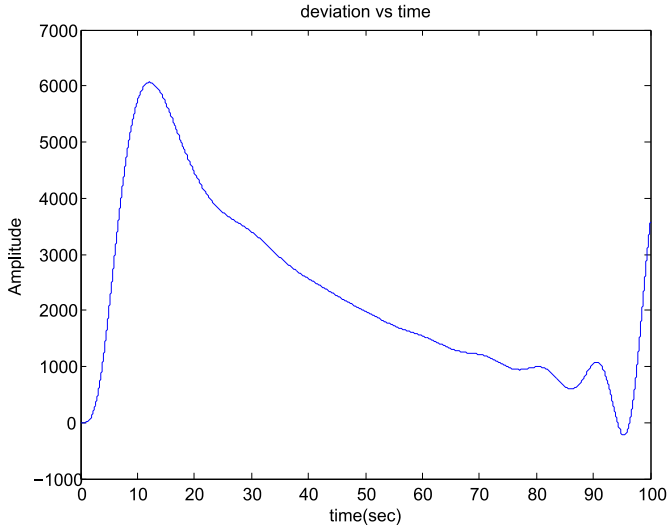


Figure12. Deviation YR in time

Fig. 12 shows the deviation of the system, when Gc = 16.63 and Ki = 0.0028 in the PI controller. After some tests, was found that the system behaves properly using these values of Gc and Ki. The initial position is 6000m. The beginning of the curve shows a jump from zero to 6000m because the input signal is a step function used only for simulation. For these values of coupler, the system tries to follow the trajectory in correct manner going down from 6000m, but at around 85 seconds, the system shows oscillations, which means that the system becomes unstable.

Looking at Fig. 4 and the equation (13),it is concluded that the system is more sensitive when the plane comes to O, because $\boldsymbol{\lambda}$ increases while R decreases. Fig. 12 clearly shows that the system is stable for large values of R, but it becomes unstable when R becomes smaller.

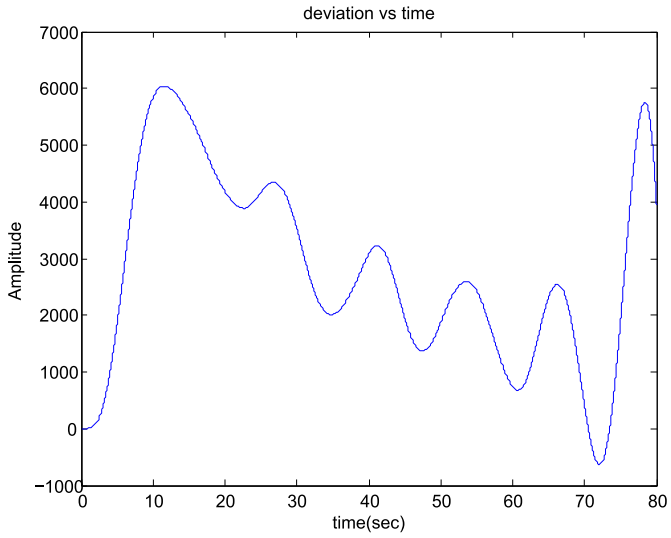


Figure13. Deviation in presence of disturbances.

According to Fig. 13, the system is unstable when external disturbances are present. It is clearly seen that during the

trajectory the system presents high oscillations and at 70 seconds it loses completely the stability.

5. SLIDING MODE CONTROL

In order to reduce the effect of disturbances and to make the system more robust, a non linear control technique is going to be implemented, which is the sliding mode control technique.

The main idea to control a system by using sliding mode control is to make the states of the system converge to a sliding surface and make them stay on it, as it is shown in the Fig. 14. In this way, the dynamics of the system is defined by the equations that determine this surface.

Establishing these equations and making them act on the system makes possible to obtain the stabilization of the system, a precise set point following and the regulation of the variables.

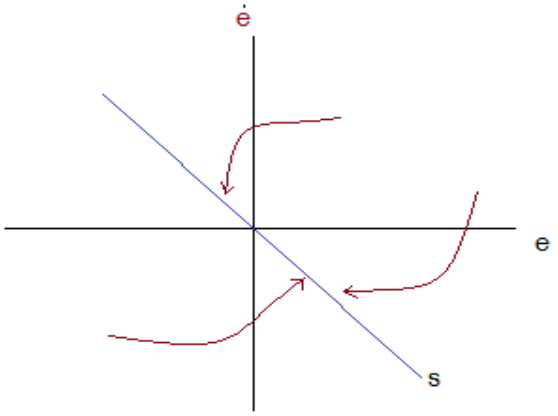


Figure14.Sliding Surface

The sliding surface, noted as "s" is defined by the equation (20)

$$s = \dot{e} + \alpha e \quad (20)$$

whereα is a parameter design greater than zero.
In order to reach the main idea of the sliding mode control, the control law will be defined by the equation (21)

$$u = ueq + ucr. \quad (21)$$

The equation (21) shows that the control law has two elements which are ueq and ucr. ueq is the equivalent control and this is the part of the controller that maintains the state of the system restricted to the sliding surface. ucr is the correction control and this is the part that make the state to converge the sliding surface. Hence, the controller is going to have the structure of the equation (21).

The sliding surface is defined in the equation (20), and e is the error which is defined as the difference between the measured state and the desired state. For this system the angular error $\boldsymbol{\lambda}$ is going to be the variable that is wanted to be controlled, so the error for our system will be:

$$e = \boldsymbol{\lambda} - \boldsymbol{\lambda}d \quad (22)$$

where$\boldsymbol{\lambda}$ is the angle measured and $\boldsymbol{\lambda}$d is the desired angle.
Replacing (22) in (20), the sliding surface is the equation (23).

$$s = (\dot{\lambda} - \lambda\dot{d}) + \alpha(\lambda - \lambda d) (23)$$

The equation (24) is the law for the attractive surface, which is the derivative of (23).

$$\dot{s} = \left(\ddot{\lambda} - \lambda\ddot{d}\right) + \alpha\left(\dot{\lambda} - \lambda\dot{d}\right) \quad (24)$$

From (5), (7) and (13) the following relationship is obtained.

$$\ddot{\lambda} = \frac{Vt}{R}p \quad (25)$$

Replacing (25) in (24) the following relation is found.

$$\dot{s} = \left(\frac{Vt}{R}p - \lambda\ddot{d}\right) + \alpha(\dot{\lambda} - \lambda\dot{d}) \quad (26)$$

In this case p = ueq+ucr. In order to calculate ueq $\dot{s}$ and ucr are going to be considered as zero; therefore, ueq is defined by the equation (27).

$$u_{eq} = -\frac{R}{Vt}\alpha\left(\dot{\lambda} - \lambda\dot{d}\right) + \frac{R}{Vt}\lambda\ddot{d} \quad (27)$$

Once defined the sliding surface s, ucr is the relation stated in the equation (28).

$$u_{cr} = -sign(s).\eta \quad (28)$$

where η is another design parameter and greater than zero.
The implementation of this gives the result plotted in Fig. 15.

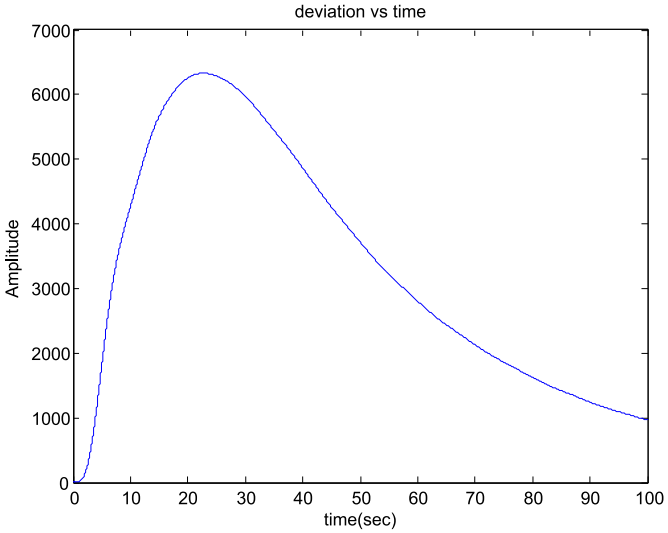


Figure15. Deviation using sliding mode control.

The constants α and η found were 0.5 and 20 respectively. With this, the system is stable even for low values of R, which was a problem by using a PI controller.

When there is a disturbance, the system responses in a better way, during the trajectory, it is possible to see oscillations, but this time they are very small as it is presented in Fig. 16.

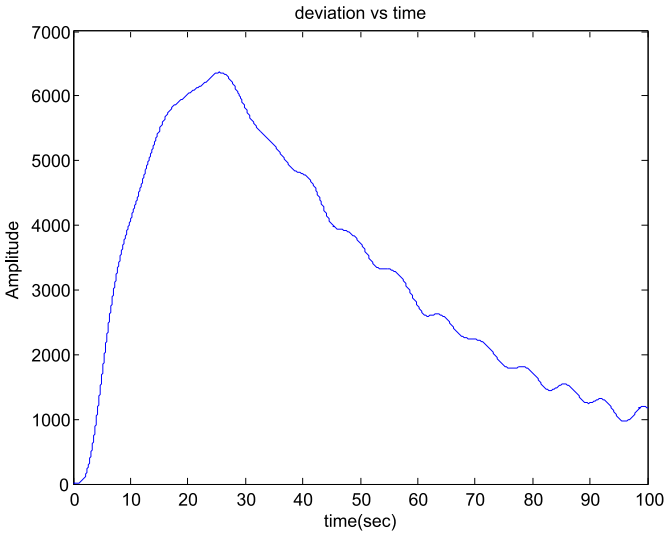


Figure 16. Deviation using sliding mode control and disturbances.

6. CONCLUSIONS

In order to make the system more stable, the sliding mode control technique was implemented. The results showed a stable system even for low altitudes and considering external disturbances. The sliding mode controller is precise because it was designed considering the non linearities of the plant. The sliding mode control gave better results. It is recommendable to try more non linear techniques like fuzzy or H infinity which would improve the stability and the robustness as well.

REFERENCES

[1] E.McGookin,"Instrument Landing System Lateral Beam Guidance System,"University of Glasgow, Oct. 2012

[2] J. Colorado, A. Barrientos, Senior Member, IEEE, A. Martinez, B. Lafaverges, and J. Valente. Mini-quadrotor Attitude Control based on Hybrid Backstepping&Frenet-Serret Theory. 2010.

[3] Rubén Rojas, Alfredo R. Castellano, Ramón O. Cáceres, Oscar Camacho.Una propuesta decontrol por modo deslizante para convertidores electrónicos de potencia.

[4] Slotine, Li. Applied nonlinear control. 1991

[5] YigengHuangfu, S. Laghrouche, Weiguo Liu, *Senior Member, IEEE*, and A. Miraoui, *A Chattering Avoidance Sliding mode Control for PMSM Drive.* 8th IEEE International
Conference on Control and Automation Xiamen, China, June 9-11, 2010.

[6] Mehmet ¨OnderEfe. *Robust Low Altitude Behavior Control of a Quadrotor RotorcraftThrough Sliding Modes.* Mediterranean Conference on Control and Automation July 27-29 2007Athens-Greece.

Control de Seguimiento de Trayectoria y Paletización de un Robot de Tres Grados de Libertad Tipo *SCARA* (*Selective Compliance Assembly Robot Arm*)

BoadaY.*; Morales L.*; Sotomayor N.*

**Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica*
Quito, Ecuador (Tel: +593-2-2507-144; e-mail: yadira.boada@epn.edu.ec;
luis.moralesec@epn.edu.ec; nelson.sotomayor@epn.edu.ec)

Resumen:*El presente trabajo se centra en el ensamblaje de un robot SCARA de tres grados de libertad y su posterior control para el seguimiento de trayectorias y tareas de paletización. Este manipulador cuenta con tres eslabones unidos por dos articulaciones de tipo revoluta y una prismática, las cuales le proporcionan movilidad en los tres planos. Utilizando programación de alto nivel, se ha diseñado una interfaz de usuario o HMI en el computador desde el cual se envían y se recibe información del robot a través de una tarjeta de adquisición de datos. El programador puede definir tareas como la colocación de una pinza en una posición, generar trayectorias y tareas de paletización deseadas. La regulación y seguimiento trayectorias se han obtenido como resultado del diseño de controladores mediante técnicas de control PID.*

El sistema fue diseñado con un módulo de potencia compacto y de alta eficiencia, de fácil acceso a todas las variables del sistema tanto en su lectura como escritura, proporcionando vía libre para aplicar cualquier tipo de técnica de control moderno y permitiendo el uso del sistema como plataforma de pruebas para controladores en el entorno académico.

Palabras clave:*Robótica, Controladores PID, Robot SCARA.*

Abstract:*The present work has the objective of assembling a SCARA robot with three degrees of freedom and, it will be controlled to path tracking and transport pallets. This robot arm has three links connected by two revolute joints and one prismatic joint, which allows mobility in the three planes. We used high level programming to design a HMI that communicates the robot with the operator through a data card. The programmer can defines some tasks such as: placing of an end-effector in a decided position, path tracking and executing palletizing tasks. The control of SCARA robot was made using PID techniques. Moreover, the robot has a compact and high efficiency power module to enable the easy reading and writing of all system variables. The previous features provide freedom to apply any type of modern control and this robot will be used like a test academic platform.*

Keywords:*Robotics, PID Controllers, SCARA robot.*

1. INTRODUCCIÓN

La arquitectura de robots tipo *SCARA*, está equipada de libertad total de movimientos en los ejes X e Y, pero limitados severamente en sus desplazamientos en el eje Z, es decir, permite ubicar el efector final en cualquier lugar con gran rapidez y precisión sobre el plano de trabajo, mientras que en el eje vertical solo se realizan desplazamientos en unos pocos centímetros.

Estos manipuladores son una opción diseñada y ampliamente utilizada, en tareas tales como soldadura, pulido, atornillado, clasificación de productos, fabricación de productos de electrónica de consumo, entre otras tareas propias del sector productivo, donde el ser humano debe desarrollar labores más complejas y aprovechar al máximo las bondades que los prototipos robóticos ofrecen.

Se pretende con este proyecto, la realización de tareas de paletización y el seguimiento de trayectorias predefinidas, que fueron escogidas heurísticamente para la demostración del correcto funcionamiento de dicho módulo. Se utilizó una herramienta nueva del *software*Labview 8.6, el *Toolkit*PID[5] para la creación de un controlador de este mismo tipo. La calibración e implementación del compensador requiere de conocimientos de la teoría de Control y del funcionamiento de cada uno de los parámetros que incluye esta herramienta.

2. ENSAMBLAJE DEL ROBOT SCARA

2.1 Estructura Mecánica

Los planos de la estructura mecánica del robot se basan en el Cuarto Curso Internacional en Robótica Aplicadadesarrollado en México DF 2008 [1]. El robot está compuesto por 23 piezas, diseñadas en lenguaje CAD (*ComputerAsistedDesign*)mediante el paquete computacional SolidWorks Professional. Esteprogramaproporciona funciones para la modelación de piezas complejas, superficies, laminado, facilita el diseño de grandes ensambles y la creación automática de planos de ingeniería, diseño en 2D y 3D, traductores CAD, análisis de esfuerzos, entre otras bondades.

Tomando en cuenta la dinámica de las piezas del robot, se simularon sus movimientos en SolidWorks Professional tanto de manera individual como colectiva, y se obtuvieron las partes más vulnerables al desgaste y a la deformación que se muestran en la Tabla 1.

Tabla 1. Piezas críticas del robot SCARA

PIEZA	FUNCIÓN
Base Inferior	Soporta todo el peso del robot
Brazo	Primera articulación rotacional
Antebrazo	Segundaarticulación rotacional
Brida	Unión entre cuerpo y brazo y apoyo de rodamientos
Eje de hombro	Transmite el movimiento del motor M_1
Eje de codo	Transmite el movimiento del motor M_2

Las piezas de la Tabla 1 se manufacturaron en una aleación de acero inoxidable, mientras que las piezas restantes fueron elaboradas en grilón o poliamida 6 y aluminio 1020. Una pinza realiza la función de efector final y el robot *SCARA*terminado se muestra en la Fig.1, donde aparecen las 2 articulaciones revolutas y 1 prismática que brindan los 3 grados de libertad al manipulador.

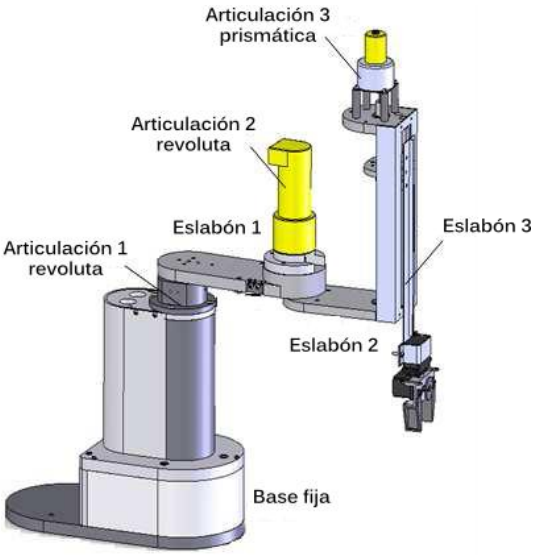


Figura 1. Robot *SCARA* de tres grados de libertad**.**

2.2 Diseño Electrónico

La parte electrónica comprende los circuitos necesarios para el manejo de: 3servomotores(motores de DC) de las tres articulaciones del manipulador, 2 servomotores de menor torque para el control del efector final, 6 fines de carrera electromecánicos, 2*encoders* en cuadratura y 1 absoluto y finalmente la conexión de un pórtico DB25 para el fácil acceso a todas las variables anteriormente mencionadas.

Estos elementos y circuitos permiten la comunicación entre la PC y el robot utilizando la tarjeta de adquisición de datos PCI-6221M [7], que fue escogida acorde las características de sensores y actuadores del manipulador, la Fig.2 muestra la interacción de todos los módulos electrónicos que conforman el control del robot *SCARA*.

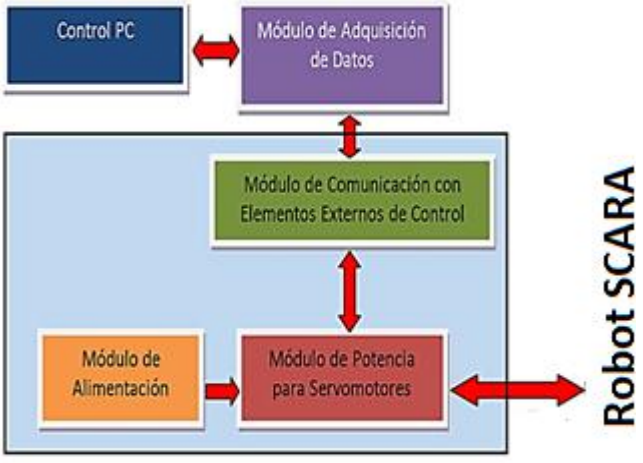


Figura 2. Diagrama de módulos electrónicos del Robot *SCARA.*

3. PROGRAMA DE CONTROL

3.1 Modelo Geométrico Directo del robot SCARA

Los manipuladores se consideran un conjunto de *n* articulaciones y *n* eslabones, donde un par de articulación-eslabón le otorga un grado de libertad al robot. Para determinar la ubicación de un eslabón *i* con respecto al eslabón *i*-1, se puede asociar a cada uno un sistema de referencia.A cada sistema de referencia se le asocia una matriz de Transformación Homogénea(TH)[2] que representa la rotaciónytranslación relativa entre el eslabón *i*y el *i*-1 para consecutivamente ubicar el elemento final del robot en la posición deseada.

Las matrices de TH requieren parámetros de traslación y de rotación que deben ser obtenidos en base a la trayectoria o posición deseada, dependiendo también de la geometría propia del robot *SCARA*. El Modelo Geométrico o Modelo Cinemático Directo (MGD) permite calcular los pares de puntos cartesianos (x, y) para cada una de las articulaciones del robot y al que el efector final debe llegar, utilizando las coordenadas de rotación o ángulos a moverse $q_1,q_2,\ldots,q_n$ (*n* depende del número de grados de libertad).

Estas matrices de THcontienen las coordenadas de rotación y traslación de dos sistemas asociados a dos eslabones consecutivos de un robot, y se denota como ${}^{i-1}T_i$. Por ejemplo, ${}^{0}T_1$ denota la orientación y posición del sistema 1(primer eslabón móvil) con respecto al sistema ubicado en la base fija del robot o sistema 0.En la matriz ${}^{i-1}T_i$se distinguen los ángulos de rotación en torno a los ejes X, Y y Z, además el vector de posición del *i-ésimo* eslabón en esos mismos ejes.

$$ {}^{i-1}T_i = \begin{bmatrix} \cos\theta_i & -\sin\theta_i\cos\alpha_i & \sin\theta_i\sin\alpha_i & a_i\cos\theta_i \\ \sin\theta_i & \cos\theta_i\cos\alpha_i & -\cos\theta_i\cos\alpha_i & a_i\sin\theta_i \\ 0 & \sin\alpha_i & \cos\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1) $$

donde: $\boldsymbol{\theta_i}$ es el ángulo en torno al eje Z_{i-1} para que los ejes X_{i-1} y X_i queden paralelos y es positivo si rota en sentido anti-horario, $\mathbf{d_i}$ es la distancia medida a lo largo de Z_{i-1} desplazando el origen del sistema S_i, $\mathbf{a_i}$es la distancia entre el origen del sistema S_i y S_{i-1}medida a lo largo del eje X_i, $\boldsymbol{\alpha_i}$es el ángulo en torno a X_i para que Z_{i-1} y Z_i queden paralelos y es positivo si rota en sentido anti-horario.Las distancias vienes dadas en centímetros (cm) y los ángulos en grados.

Los parámetros α, θ, *d* y *a* se hallan utilizando el algoritmo de Denavit-Hartenberg (D-H) [2],y se los reemplaza en (1) para obtener el MGD, es decir, las coordenadas P_x, P_y y P_z del efector final en cada instante de tiempo. Estos parámetros sedefinen en la Tabla 2 recordando que a_2, a_3 y d_1 son constantes físicas del robot, mientras que θ_1 y θ_2 son los ángulos de las articulaciones rotacionales y $q_3$es la coordenada de tipo prismática que varía según el desplazamiento del efector final.

.

Tabla2. Parámetros del robot SCARA obtenidos con D-H

ESLABÓN	a_i	α_i	d_i	θ_i
1	0	0	d_1	θ_1
2	a_2	0	0	θ_2
3	a_3	0	q_3	0

donde: a_i: longitud del eslabón, α_i: ángulo de torsión del elemento (medida según el sentido de las agujas del reloj entre el eje i−1 y el eje i), d_i: offset del eslabón, θ_i: ángulo entre elementos adyacentes.

Para obtener estos parámetros se siguió la siguiente metodología [2]:

a. Numerar los eslabones comenzando con 1 (primer eslabón móvil) y acabando con n (último eslabón móvil). Se numerará como 0 a la base fija del robot.
b. Numerar cada articulación comenzando por 1 (la correspondiente al primer grado de libertad) y acabando en n.
c. Localizar el eje de cada articulación. Si es rotativa, el eje será su propio eje de giro, si es prismática, será el eje a lo largo del cual se produce el desplazamiento.
d. El sistema S_0 se escoge localizando el origen sobre el eje Z_0; los ejes X_0 y Y_0 se toman arbitrariamente, pero formando un sistema ortogonal. Este sistema coincide con la base.
e. Para **i** desde 1 hasta n-1, situar el sistema S_i (perteneciente al eslabón i) en la intersección del eje Z_i con la línea normal común a Z_{i-1} y Z_i. Si ambos ejes se cortasen se situaría S_i en el punto de corte, si fuesen paralelos S_i se situaría en la articulación i+1.
f. Situar X_i en la línea normal común a Z_{i-1} y Z_i con la dirección de la articulación i a la i+1 (hacia donde se aleja).
g. Situar Y_i de modo que forme un sistema ortogonal con X_i y Z_i.
h. Repetir los pasos e, f y g desde i=1 hasta i=n-1.
i. Situar el sistema S_n en el extremo del robot de modo que si la articulación **n** es revoluta, se debe alinear Z_n con Z_{n-1}; si la articulación*n* es prismática, se debe tomar arbitrariamente a Z_n. El eje X_n se escoge como en el paso f.
j. Obtener $\boldsymbol{\theta_i}$ como el ángulo que se debe girar en torno a Z_{i-1} para que X_{i-1} y X_i queden paralelos, siendo positivo si la rotación es en sentido antihorario.
k. Obtener $\mathbf{d_i}$como la distancia medida a lo largo de Z_{i-1} desplazando el origen del sistema S_i.
l. Obtener $\mathbf{a_i}$como la distancia entre el origen del sistema S_i y S_{i-1}, medida a lo largo del eje X_i.
m. Obtener $\boldsymbol{\alpha_i}$ como el ángulo que habría que girar en torno a X_i para que Z_{i-1} y Z_i queden paralelos, siendo positivo si la rotación es en sentido antihorario.

En la Fig. 3 se muestra gráficamente la forma de obtención de los parámetros de la Tabla 2, mientras que en la Fig. 4 se observan los sistemas de referencia S_i (fijo y móviles) para cada articulación según el algoritmo de D-H.

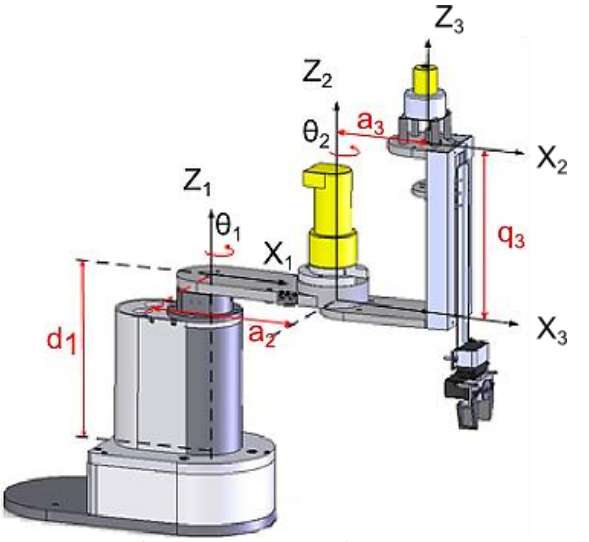


Figura 3. Parámetros y ejes X_i,Z_i según el algoritmo de D-H.

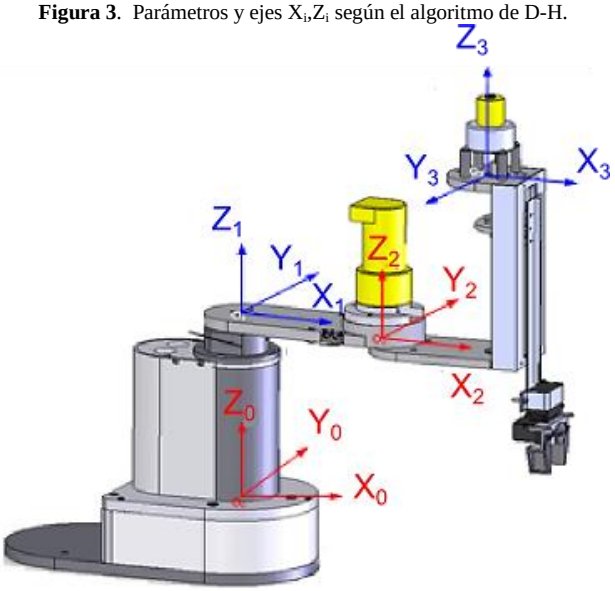


Figura 4. EjesX_i, Y_i y $Z_i$de las 3 articulaciones según el algoritmo de D-H.

Si se reemplazan los datos de la Tabla 2, de manera iterativa en la matriz deTH se pueden obtener las matrices de transformación correspondientes a los tres eslabones: dos rotacionales${}^0T_1$ (θ_1) y${}^1T_2$ (θ_2), y uno prismático${}^2T_3$ (q_3). Finalmente se realiza una post-multiplicación para obtener una nueva HT que representa la posición del efector final con respecto a la base fija del robot ${}^0T_3 = {}^2T_3\,{}^1T_2\,{}^0T_1$.

$$ {}^0T_1(\theta_1) = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & \sin\alpha_i & 1 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2) $$

$$ {}^1T_2(\theta_2) = \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 & 0 & a_2\cos\theta_2 \\ \sin\theta_2 & \cos\theta_2 & 0 & a_2\sin\theta_2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3) $$

$$ {}^2T_3(d_3) = \begin{bmatrix} 1 & 0 & 0 & a_3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4) $$

$$ {}^0T_3 = \begin{bmatrix} C_{12} & -S_{12} & 0 & a_3C_{12} + a_2C_1 \\ S_{12} & C_{12} & 0 & a_3S_{12} + a_2S_1 \\ 0 & 0 & 1 & d_3 + d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5) $$

donde, C_{12}=Cos(θ_1+θ_2), S_1=Senθ_1, C_1=Cosθ_1.

De la última columna de la matriz (5) se obtiene la posición final del actuador o efector final (P_x, P_y y P_z).

$$ P_x = a_3C_{12} + a_2C_1 \quad (6) $$
$$ P_y = a_3S_{12} + a_2S_1 \quad (7) $$
$$ P_z = d_3 + d_1 \quad (8) $$

Reemplazandod_1, a_2 y a_3 por sus medidas físicas en centímetros se obtienen las ecuaciones (9-11) que son el MGD:

$$ P_x = 14\,C_{12} + 14C_1 \quad (9) $$
$$ P_y = 14\,S_{12} + 14S_1 \quad (10) $$
$$ P_z = d_3 + 15 \quad (11) $$

3.2 Modelo Geométrico Inverso del robot SCARA

A partir del MGD (9-11) se deducen expresiones en función de P_x, P_y y P_z que definen en cualquier instante los ángulos q_1, q_2 y $q_3$que las articulaciones rotacionales y prismáticadeben moverse hasta alcanzar la posición deseada del efector final. Después de un análisis geométrico en el espacio de trabajo del manipulador y la utilización de la Ley de cosenos se obtuvieron las siguientes ecuaciones:

$$ q_3 = P_z - 15 \quad (12) $$
$$ q_2 = \cos^{-1}\left(\frac{P_x+P_y-392}{392}\right) \quad (13) $$

$$ q_1 = \sin^{-1}\left(\frac{P_y}{\sqrt{{P_x}^2+{P_Y}^2}}\right) - \tan^{-1}\left(\frac{14\sin(q_2)}{14+14\cos(q_2)}\right) \quad (14) $$

$Six \geq 0,$

$$ q_1 = -\sin^{-1}\left(\frac{P_y}{-\sqrt{{P_x}^2+{P_Y}^2}}\right) + \tan^{-1}\left(\frac{14\sin(q_2)}{14+14\cos(q_2)}\right) \quad (15) $$

$Si\ x < 0.$

Estas ecuaciones representan el Modelo Geométrico Inverso (MGI) del robot *SCARA*. Según la geometría propia de nuestro robot se efectuaron ciertas correcciones necesariasenla presente aplicación. El MGI es de gran importancia en este trabajo, ya queda lugar a la generación de trayectorias y tareas de paletización utilizando un control en lazo cerrado permitiendo el cálculo de los ángulos y desplazamientos que deberán tener los eslabones para llegar a los puntos deseados dentro del área de trabajo.

3.3 Estrategia de Control

Se implementó un control PID para garantizar un error de posición nulo, es decir, la diferencia entre la posición deseada y la calculada por el manipulador será cero. Tomando en cuenta el tipo de actuadores del robot, la señal de control *u(t)* será un PWM (*Pulse WidthModulation*) para cada uno de los motores DC, teniendo como mínimo valor de ancho de pulso un 0.1% de su período debido a la inercia de los motores. La ley de control está definida por la expresión (16) [3].

$$u(t) = W_n{}^2 e(t) + 2\varepsilon W_n \dot{e(t)} + k_i \int_0^t e(t)\, dt \quad (16)$$

donde:*u(t)* = señal de control, *e(t)* = señal de error.

La frecuencia natural W_n está directamente relacionada con la constante K_p o ganancia proporcional del controlador. Esta ganancia tiene la característica de aumentar la velocidad del control, además el tiempo de subida es más rápido pero a costa de un mayor sobre-impulso que puede desestabilizar el sistema. La constante de amortiguamiento ξ está relacionada con el valor de la constante derivativa K_d. Si la constante K_d aumenta, el control es más lento, el sobre-impulso decrece y la estabilidad mejora. La constante de integración K_i provoca consecuencias similares que la proporcional.

Para el controlador PID propiamente dicho, se utilizó el *software*LabView 8.6, que cuenta con herramientas especiales para su construcción. La sintonización del PID, fue realizada siguiendo las reglas de ZIEGLER-NICHOLS puesto que no se disponía de la función de transferencia de los servomotores colocados en las articulaciones, la Fig. 5muestrael lazo de control PID para las dos articulaciones de tipo revoluta.

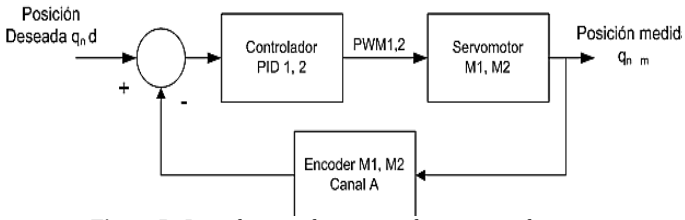


Figura 5. Lazo de control para articulaciones revolutas.

Para el motor perteneciente a la articulación prismática, se implementó un control tipo ON/OFF debido a la simplicidad de su movimiento, es decir, deslizar el tercer eslabón paralelamente al eje Z después de posicionar a los dos eslabones revolutos anteriores en la ubicación exacta en el plano de trabajo. La articulación prismática únicamente permite que la pinza que es el efector final, se aproxime o se separe del plano de trabajo.

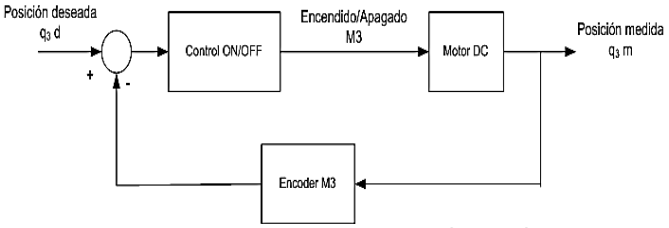


Figura 6. Lazo de control para articulación prismática.

3.4 Desarrollo de la interfaz de usuario en el computador

Para controlar las posiciones y trayectorias seguidas por el robot *SCARA*, se diseñó en LabView una interfaz que cuenta con controles e indicadores para la ejecución de los modelos geométricos Directo e Inverso (MGD y MGI respectivamente), la generación de 3 trayectorias escogidas heurísticamente (círculo, espiral y rosa de los vientos) con dimensiones definidas por el operador, y la ejecución de tareas de paletización a manera de ejemplo.

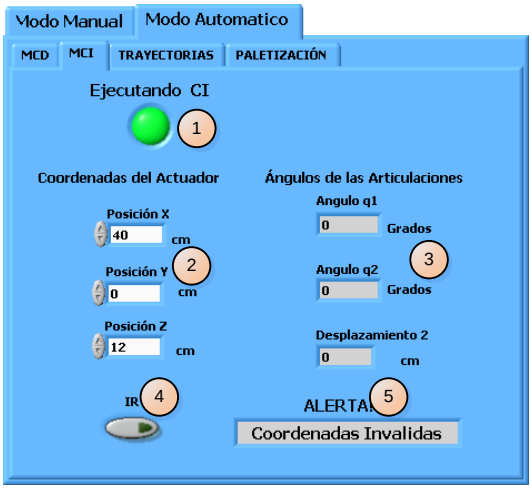


Figura 7. Pantalla inicial de interfaz del MCI.

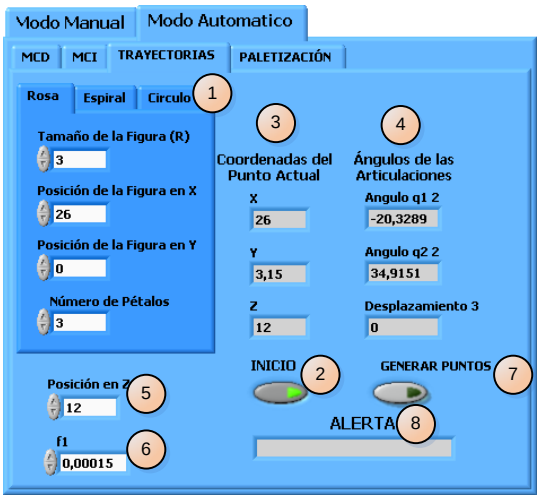


Figura 8. Pantalla de interfaz para generación de trayectorias.

Para el desarrollo de las tareas descritas en el párrafo anterior, se programaron las ecuaciones vistas en la sección 3

empleando lenguaje C, tratando en todo momento de disminuir el coste computacional para realizar control en tiempo real. Dentro de la interfaz, se aprecian también las respuestas de los motores DC frente a la señal de entrada o *setpoint*, que es la posición a la que el efector final debe llegar. En la programacióndelcontrolador PID se incluye el cálculodel ancho de pulso de la señal PWM o *u(t)* que se actualiza y disminuye hasta que el robot alcance la posición deseada.

Una vez que el controlador PID ha calculado el ancho de pulso del PWM, este dato es enviado al algoritmo de generación de la señal PWM que debe tener en cuenta la frecuencia y el número de muestras para un correcto funcionamiento del controlador (ver Fig. 9).

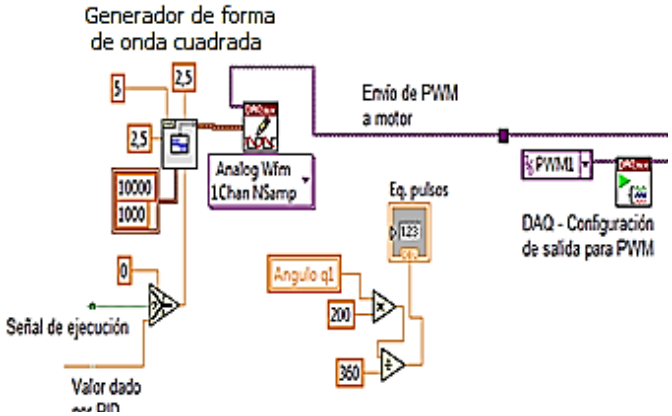


Figura 9. Algoritmo de generación de la señal de control PWM.

4. PRUEBAS Y RESULTADOS

Una vez ensamblado el robot *SCARA*, se pudo observar que la estructura mecánica construida en poliamida y aleación de hierro para ciertas piezas, resultó robusta con articulaciones lo suficientemente rígidas para realizar las tareas que nos planteamos inicialmente.

En la Fig. 10 se muestra la estructura mecánica completa del manipulador y que se utilizópara realizar las pruebas del control implementado en el computador.

Figura 10. Robot *SCARA* completo con módulo de entradas y salidas.

4.1 Pruebas de Control Cinemático

Dentro del programa de control se procedió a la calibración de los parámetros de ganancia proporcional K_p, tiempo integral T_i, y tiempo derivativo T_d deducidos de las ecuaciones (17, 18 y 19), buscando alcanzar de forma rápida y sin error, los puntos o ángulos deseados en el plano de trabajo del manipulador.

$$K_c = \frac{1}{K_p} \qquad (17)$$

$$K_i = \frac{K_p}{T_i} \qquad (18)$$

$$K_d = K_p T_d \qquad (19)$$

Para esta aplicación en particular del robot *SCARA*, se definió un tiempo de levantamiento de t_d=2segundos, un tiempo de establecimiento t_s=4 segundos, eliminar por completo el sobre-impulso o M_p de la respuesta del sistema y obtener un error de posición o erros de estado estable E_{ss} nulo.

En la Tabla 3 se aprecian las constantes para el controlador PID que después de ser calibrado, produjo la mejor respuesta al escalón (ángulo o posición (x,y) de referencia), mientras que en la Fig. 11 se observa la respuesta escalón de uno de los servomotores de las articulaciones rotacionales.

Tabla3. Constantes de los controladores PID de las articulaciones rotacionales

	PID_M$_1$	**PID_M$_2$**
K$_c$	20	30
T$_i$	2	1.1
T$_d$	0.2	0.13

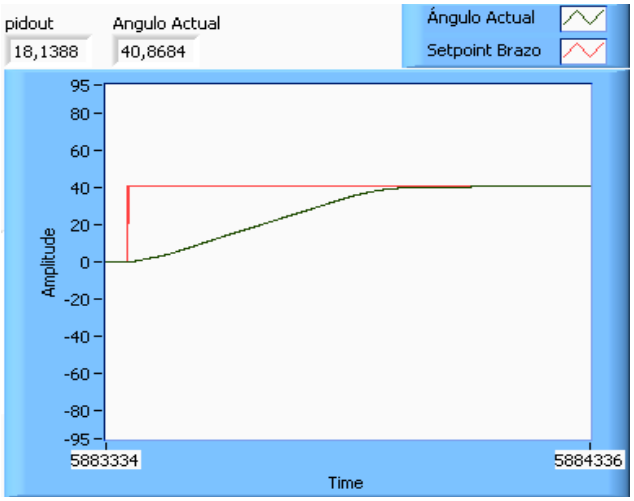


Figura 11. Respuesta del servomotor M_1 (brazo del robot), referencia (línea roja) y respuesta (línea negra).

4.2 Respuesta del robot SCARA a los modelos cinéticos Directo e Inverso

Cuando el robot *SCARA* alcanza los puntos (x,y) o ángulos entregados por los modelos cinemáticos directo e inverso respectivamente, lo hace de forma rápida y exacta, tratando en todo momento de mantener una diferencia casi nula entre el punto o ángulo deseado y el punto o ángulo medidos. Para las articulaciones tipo revoluta, se muestran los datos de precisión, exactitud y repetibilidad en la Tabla 4.

Tabla 4. Parámetros del robot *SCARA*

MEDIDA	VALOR (°)
Precisión	± 0.017
Exactitud	± 0.045
Repetibilidad	± 0.029

4.3 Pruebas de control en la generación de trayectorias y tareas de paletización

Para el ajuste de parámetros de los compensadores PID's en la generación de trayectorias se debe tomar en cuenta la variación continua, rápida y permanente de los ángulos. Para este efecto se aumentaron progresivamente las constantes K_c (recíproco de K_p)y$T_i$hasta obtener la salida deseada,y manteniendo fijo el valor del tiempo T_d para el motor M_1. Para el motor $M_2$$T_d$fue decrementado y se obtuvieron los datos mostrados en la Tabla 5.

Tabla5. Constantes de los controladores PID de las articulaciones rotacionales para generar trayectorias

	PID_M_1	PID_M_2
K_c	100	100
T_i	7	8
T_d	0.2	0.08

En la Fig. 12se muestra una de las trayectorias programadas, la generación de una rosa de los vientos como trayectoria predefinida desde la interfaz de usuario.La configuración de parámetros como: número de pétalos, coordenadas (h,k) del centro, radio y resolución de la figura pueden aumentar o disminuir los puntos que formarán latrayectoria.Las ecuaciones de los puntos cartesianos (x, y) que permiten generar dicha figura son:

$$x = r[\sin(2\pi f_2 t)\sin(2\pi f_1 t)] + h \qquad (20)$$

$$y = r[\sin(2\pi f_1 t)\cos(2\pi f_2 t)] + k \qquad (21)$$

donde:**r** es el radio de la figura, $\mathbf{f_1}$es la constante de número de puntos guía, $\mathbf{f_2}$es el número de pétalos de la rosa, **t** es el tiempo variable que permite la generación de puntos, **h** es el desplazamiento de la figura en eje X, y **k** es eldesplazamiento de la figura en eje Y.

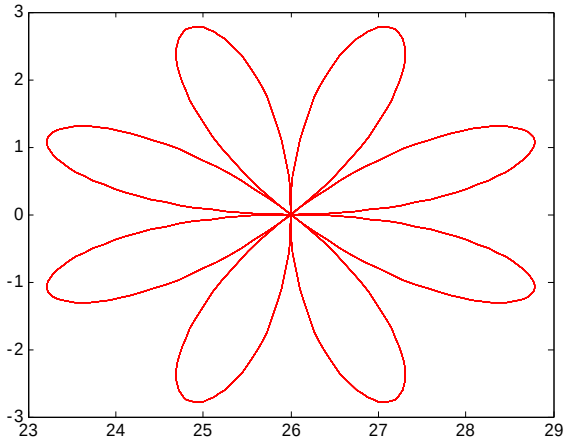


Figura 12. Trayectoria"Rosa de los Vientos" programa en la interfaz del robot *SCARA*.

El robot *SCARA* al poseer una pinza como efector final puede sujetar firmemente un lápiz que le permita dibujar cualquiera de las trayectorias disponibles en la interfaz, esto a manera de mostrar el verdadero movimiento sobre el plano de trabajo. La Fig. 13 indica la trayectoria dibujada por el manipulador que se asemeja a la presentada en la Fig. 12 que es la de referencia.

LasFigs. 14 y 15 muestran las trayectorias esperada y generada de una espiral cuya ecuación matemática es:

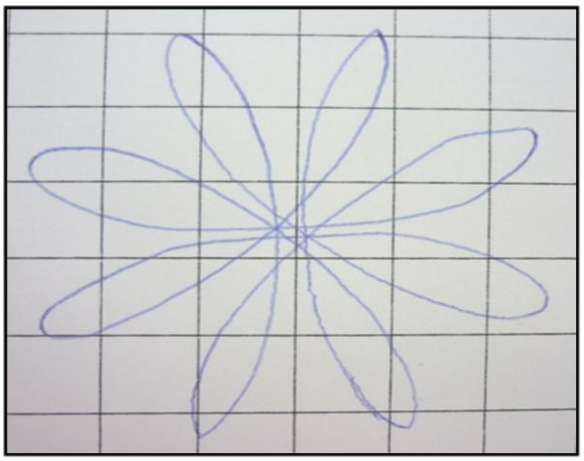


Figura 13. "Rosa de los Vientos"generadapor el robot *SCARA*.

$$x = atf_1[\cos(2\pi f_1 t)] + h \qquad (22)$$

$$y = atf_1[\sin(2\pi f_1 t)] + k \qquad (23)$$

donde:**a** es la distancia de separación de los brazos de la espiral, $\mathbf{f_1}$es laconstante de número de puntos guía, **t** es la variable de generación de puntos, **h** es eldesplazamiento de la figura en eje X, y **k** es eldesplazamiento de la figura en eje Y.

En la Fig. 16 se puede observar al robot implementado, siguiendo la trayectoria de la Fig. 14 y marcándola en su espacio de trabajo.

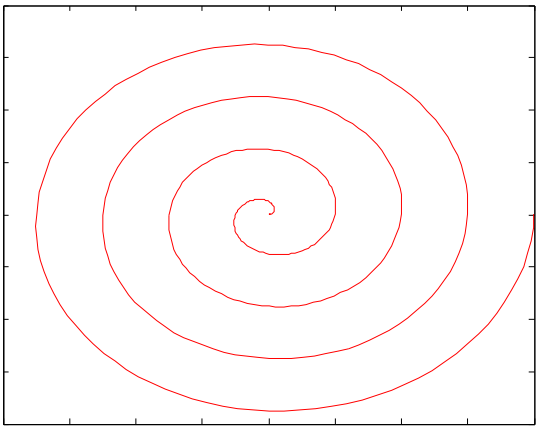


Figura 14. Trayectoria "Espiral" programa en la interfaz del robot *SCARA*.

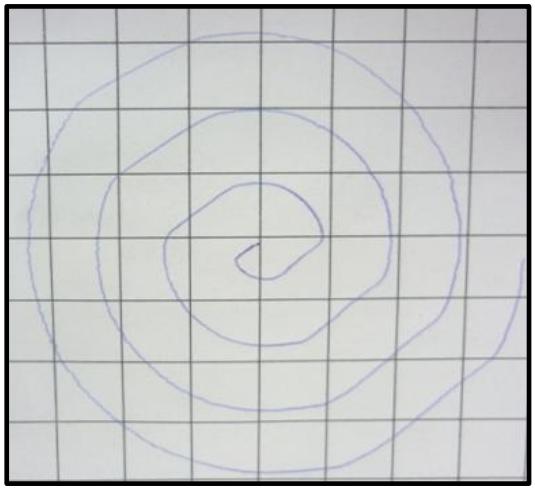


Figura 15. "Espiral" generada por el robot SCARA.

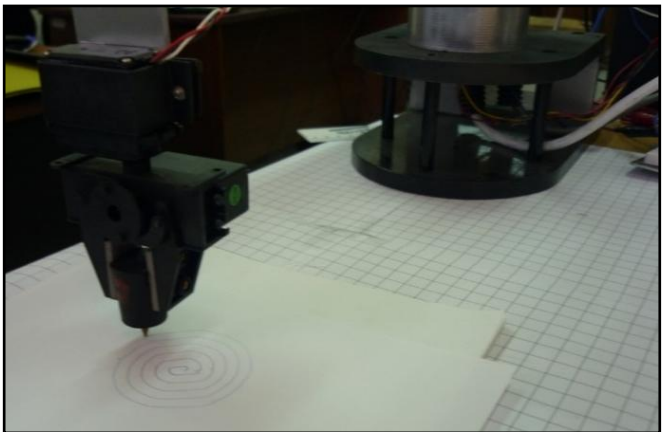


Figura 16. Ejecución de una trayectoria programada para el robot *SCARA*.

Las tareas de paletización son muy comunes en los manipuladores y constan de un movimiento continuo y repetitivo en el tiempo, se trata de ubicar el efector final en una posición deseada, tomarun objeto o palet automáticamente, trasladarlo y soltarlo en otra posición deseada. Durante las pruebas, se escogieron fusibles de 3 cm de largo aproximadamente como palets u objetos a transportar, el agarre de la pinza permite levantar hasta 3 fusibles sin peligro a soltarlos y los datos técnicos para la ejecución de esta tarea se muestran enla Tabla 6.

Tabla 6. Datos para tareas de paletización

Máxima altura de carga	150 mm*
Zona de trabajo hasta	280 mm
Peso de *pallet* (unidad)	350 gr **
Máxima apertura de pinza	30 mm

*milímetros, **gramos.

5. CONCLUSIONES

El movimiento suave de las articulaciones del robot *SCARA*, se debe al uso de rodamientos o cojinetes que disminuyen la fricción y evitan el desgaste de las piezas móviles acopladas a un eje, además le sirven de apoyo y facilitan su rotación. Es apropiado el uso de rodamientos del tipo de bolas rígidas que sirven para operar incluso a altas velocidades y requieren de muy poco servicio de mantenimiento, por este motivo son los más populares y aptos para este tipo de aplicación.

El algoritmo de Denavit–Hartenberg para la definición de ejes y obtención del Modelo Geométrico Directo (MGD), es una herramienta útil y una forma simplificada de definir los puntos cartesianos de ubicación del efector final, definiendo previamente los ángulos para cada una de las articulaciones rotacionales y/o prismáticas del manipulador.

Las características principales de un robot *SCARA* son la precisión y la velocidad con que ejecuta sus movimientos, para cumplir estos requerimientos se utilizaron motores que llevan acoplados *encoders* de alta resolución, otorgándole a éste, rotaciones en el orden de las centésimas de grado en cualquier posición o trayectoria.

Para el motor perteneciente a la articulación prismática, se implementó un control tipo ON/OFF puesto que el movimiento a lo largo del eje Z consiste únicamente en deslizar el efector final verticalmente.

Se implementaron dos compensadores tipo PID para los dos servomotores que controlan el movimiento rotacional del manipulador. Con la calibración y sintonización de estos compensadores (siguiendo un método heurístico), se logró obtener una respuesta sobre-amortiguada y con error en estado estable de 0.05 milímetros para ambos motores.. Esto se traduce en una veloz respuesta del sistema, sin oscilaciones y con una excelente aproximación entre el punto deseado y el medido.

Las tareas de paletización constituyen una de las principales aplicaciones de los brazos robóticos porque introducen automatismos y mecanizado en línea para una amplia gama de procesos industriales, reduciendo costos de producción y tiempos de ejecución de casi cualquier manufactura.

6. AGRADECIMIENTOS

Un agradecimiento especial a la Escuela Politécnica Nacional por otorgar parte del financiamiento para la culminación de este trabajo.

7. REFERENCIAS

[1] R. Silva, F. Riveros, M. Cordero, *Cuarto Curso Internacional en Robótica Aplicada, Diseño Mecánico*. México DF, Septiembre 2008.

[2] A. Barrientos, L.Peñín, C.Balaguer,*Fundamentos de Robótica*. Segunda edición. 2007.

[3] K. Ogata, *Control Automático*.Edición 2006.

[4] W., Etienne, Modeling, *Identification and Control of Robots*. Terceraedición, 2002.

[5] *Toolkit PID Manual*, National Instruments,2009.

[6] *LabView 8.6 Manual*,National Instruments, 2009.

[7] *Low-Cost M Series Multifunction DAQ PCI 6221-M Manual*, National Instruments, 2009.

[8] O. Vivas. (2006). *Control Predictivo De Un Robot Tipo ScaraPredictive Control of a Scara Robot*. Revista Chilena de Ingeniería, Vol. 14 Nº 2.

[9] Dessaint L., Saad M., HCbert B. (April 1992). Al-Haddad K.*AnAdaptive ControllerFor A Direct-Drive ScaraRobot.*IEEETransactions On Industrial Electronics, Vol.39, No *2.*

[10] I. Minchala. (Agosto 2009). *Curso de Robótica Industrial*. Escuela Politécnica Nacional.

Diseño y Construcción de un Inversor Multinivel

Chauca F.*; Llerena F.*; Chico P.*
**Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica*
Quito, Ecuador (e-mail:franklinchauca@hotmail.com ; faustollerena@hotmail.com; patricio.chico@epn.edu.ec)

Resumen:*Muchos de los inversores comercialesusan modulación de dos o tres niveles y se utilizan filtros de salida para obtener una forma de onda aproximada a la sinusoidal. El presente trabajo muestra los resultados de la construcciónde un inversor multinivel (81 niveles) en base a inversores de tres niveles conectado, mediante transformadoresen serie. Se usa una fuente de corriente continua común para la entrada de los inversores y transformadores con distinta relación de transformación a la salida. El control es diseñado en base a algoritmos implementados en microcontroladores. Los circuitos de potencia son diseñados con MOSFET de potencias adecuadas en función de cada inversor.*

Palabras clave: *Inversor, MOSFET, multinivel, Puente H, transformador, THD.*

Abstract:*Commercial inverters commonly use two or three levels modulation and output filters are used to obtain a waveform approximate to sine wave. This project is a multilevel inverter (81 levels) based on three level inverters connectedusing transformers. It uses a common DC source and transformers with different input-output ratio. The control circuit is designed based on algorithms implemented in microcontrollers. The power circuits are designed with adequate power MOSFET, depending on the power handled of each inverter.*

Keywords: *Inverter, MOSFET, multilevel, H bridge, transformer, THD.*

1. INTRODUCCIÓN

Un inversor es un conversor estático de energía de corriente continua a corriente alterna [10]. Los inversores multinivel se los puede usar en muchas aplicaciones, por ejemplo: fuentes de energía ininterrumpida UPS (*UninterruptiblePowerSupply*), en compensadores de energía reactiva, en filtros activos, en la generación eléctrica desde fuentes no convencionales (plantas fotovoltaicas o eólicas), en la industria a través de control de motores, etc.

Las ventajas que presentan frente a los inversores convencionales de dos o tres niveles es que los inversores multinivel presentan una baja distorsión armónica de voltaje en la salida, una mayor capacidad de potencia.

Existen varias topologías para el diseño de inversores multinivel como la de capacitores flotantes, diodos enclavados o inversores tipo puente en cascada [5]. Ésta última es la topología utilizadaya que se dispone de una entrada de corriente continua (CD) de 12 Vdc y mediante el adecuado diseño de transformadores se puede llegar a obtener un voltaje requerido de 120Vrms a la salida.

Se usan cuatro inversores de tres niveles (uno principal y tres auxiliares). Cada uno de los inversores entrega a su salida un nivel con un valor específico de voltaje que, sumado algebraicamente con la salida del resto de inversores, va formando cada uno de los 81 niveles.

2. INVERSOR DE TRES NIVELES DE PUENTE COMPLETO

La topología usada es un puente de MOSFET (denominado Puente H) dispuesto como se muestra en la Figura 1. El voltaje pico de salida es Vs (valor de la fuente DC) y la salida de voltaje alterno está entre los puntos A y B.

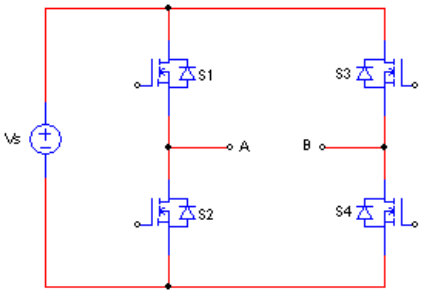


Figura 1. Inversor monofásico en configuración puente completo o puente H

En función del tipo de control utilizado se tienen diferentes formas de onda de voltaje de salida. Por ejemplo en la Figura 2 se observa una onda de voltaje y corriente de salida tipo cuasi-cuadrada, mientras que la Figura 3 muestra voltaje de salida con control tipo PWMSenoidal (SPWM) [10].

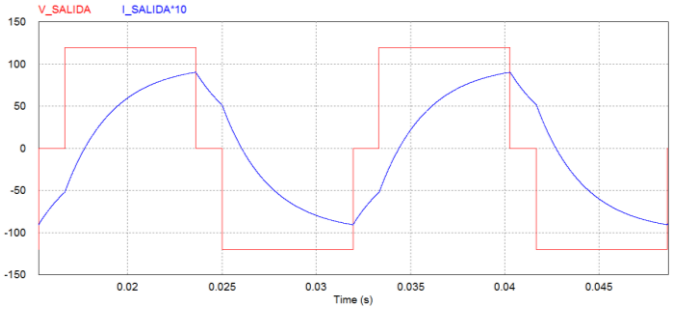


Figura 2. Salida del inversor monofásico. Ondacuasi-cuadrada

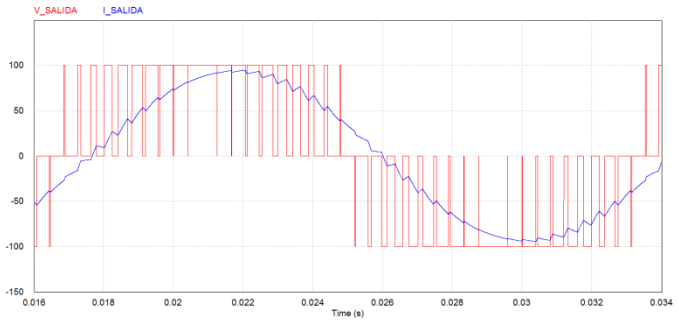


Figura 3.Salida del inversormonofásico.Onda SPWM

La diferencia entre estos tipos de control es la calidad de la energía a la salida. La Figura 4 muestra la respuesta de frecuencia del voltaje de salida cuasi-cuadrado, en el que se observa que los armónicos representan gran parte de la señal de salida, convirtiéndose de esta manera en pérdidas. La Figura 5 muestra lo propiopara la señal SPWM. Básicamente la salida debería tener el menor contenido armónico reflejado en la Distorsión Armónica Total (THD) más baja posible.

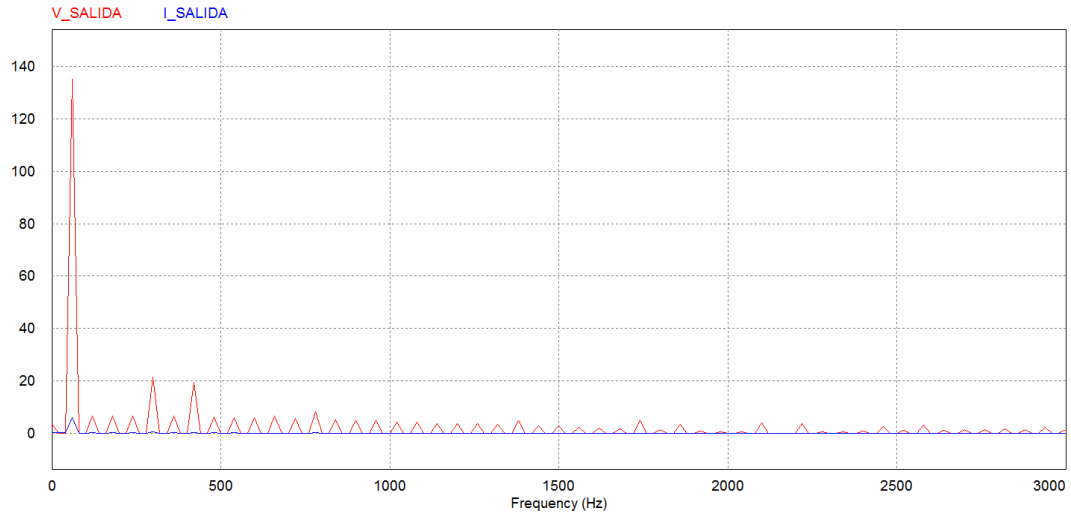


Figura 4. Respuesta de frecuencia de la onda cuasi-cuadrada

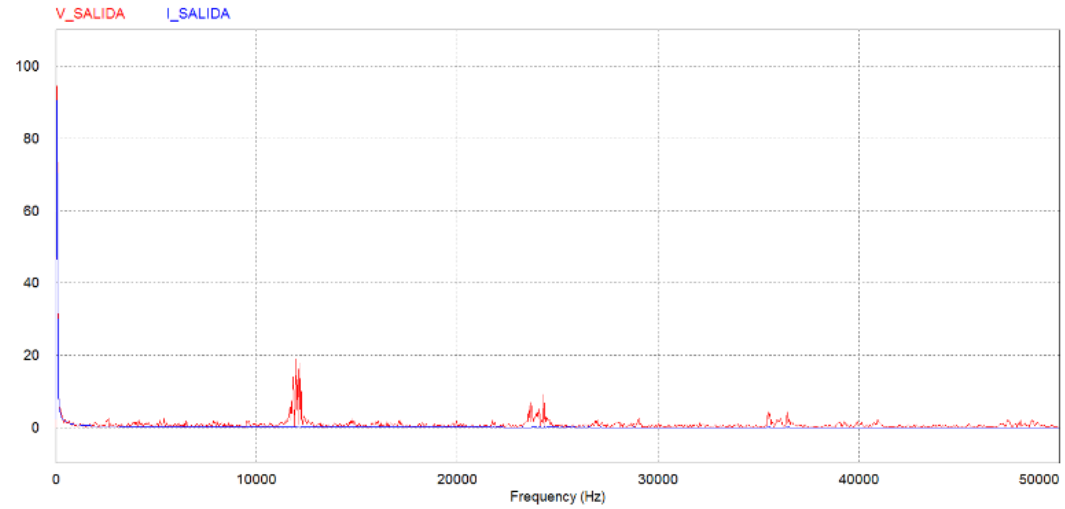


Figura 5. Respuesta de frecuencia de la onda SPWM

2.1 Control del inversor tipo puente de tres niveles [11]

Para obtener la onda de salida deseada en un inversor de tres niveles (cuasi-cuadrada o PWM) es necesaria cierta secuencia de disparo en sus cuatro interruptores de potencia.La lógica a seguir para conseguir la onda alterna a la salida se detalla en la Tabla 1.

Para que la carga pueda tener un voltaje positivo o negativo en sus terminales es necesario encender a la vez dos MOSFET en diagonal (S1 y S4 o S2 y S3), mientras que para que en la carga haya 0V se deben encender dos MOSFET opuestos (S1 y S3 o S2 y S4), esto es por la corriente que se almacena en las cargas inductivas, la cual se descargará en la

misma carga cerrando el circuito con los interruptores de potencia y los diodos en anti-paralelo.

La Figura 6 detalla gráficamente la secuencia de disparos para obtener una onda cuasi-cuadrada a la salida del inversor. Se debe tener en cuenta que en un mismo ramal no pueden estar ambos MOSFET cerrados al mismo tiempo, ya que se genera un corto circuito a la fuente. Para evitar esto es necesario generar *tiempos muertos*, esto es, un tiempo en que se garantice que los dos interruptores de un mismo ramal estén completamente apagados [3].

Tabla 1. Lógica de control para un inversor de tres niveles tipo puente

	SEÑALES DE CONTROL DE LOS MOSFET			
SN / VO (A-B)	S1	S2	S3	S4
+VS	1	0	0	1
0	1	0	1	0
-VS	0	1	1	0
0	0	1	0	1

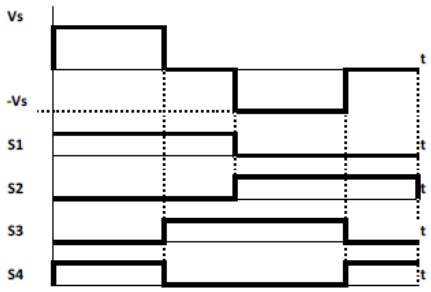


Figura 6. Secuencia de disparos para un inversor de tres niveles tipo puente

Se observa en la Figura 7 la señal de control a dos MOSFET de un mismo ramal, existiendo un tiempo en que ambas señales están en 1L, que corresponde al tiempo muerto. Este tiempo muerto se o genera de esta forma porque el driver usado invierte las señales al aplicarlas a las compuertas..

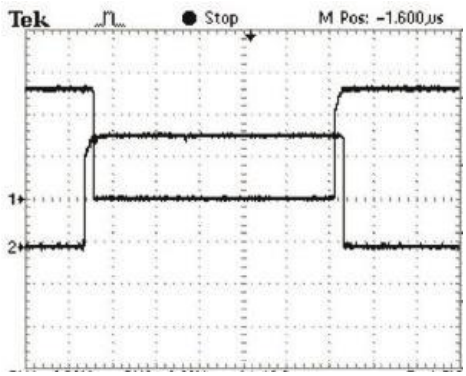


Figura 7. Señales de un ramal del puente inversor con tiempos muertos.

3. INVERSORMULTINIVEL

Un inversor multinivel presenta varios niveles de voltaje o escalones en su onda de salida, pudiendo los mismos ser generados por distintos métodos. La técnica utilizada en este proyecto se basa en inversores conectados a transformadores individuales cuyos bobinados secundarios se disponen en serie para obtener los niveles deseados como en la Figura 8.

Las características de esta técnica son las siguientes:

- El uso de una sola fuente de alimentación CD.
- Con el uso de una adecuada relación de transformación se obtiene una gran cantidad de número de niveles con pocos inversores conectados en el sistema.
- Menor frecuencia de conmutación, por lo tanto elementos más lentos, siendo éstos de mayor potencia.

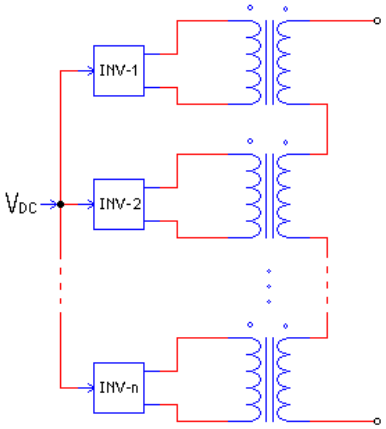


Figura 8. Topología de inversores multinivel tipo puente monofásico

3.1 Control de inversores multinivel tipo puente [8] [4]

El control se puede hacer en función de la relación de transformación de los transformadores de salida. Por ejemplo, al usar transformadores de relación 1:1, el control se lo puede implementar como muestra la Figura 9, en la que en base a las ondas de salida de cada inversor, sumándolas, se obtiene cinco niveles a la salida del sistema.

Así mismo si se cambia de relación de transformación en cada transformador de salida y la secuencia de disparos, se pueden obtener más niveles a la salida con el mismo número de inversores.

Al observar las distintas configuraciones que se pueden hacer a partir de un inversor monofásico de una sola etapa de puente H (3 niveles E, -E, 0) se puede incrementar de manera geométrica el número de niveles según vaya aumentando el número de etapas. De esta forma con dos etapas se pueden lograr hasta 9 niveles (3^2) y así sucesivamente. En la Tabla II se puede ver el número máximo de niveles según el número de etapas.

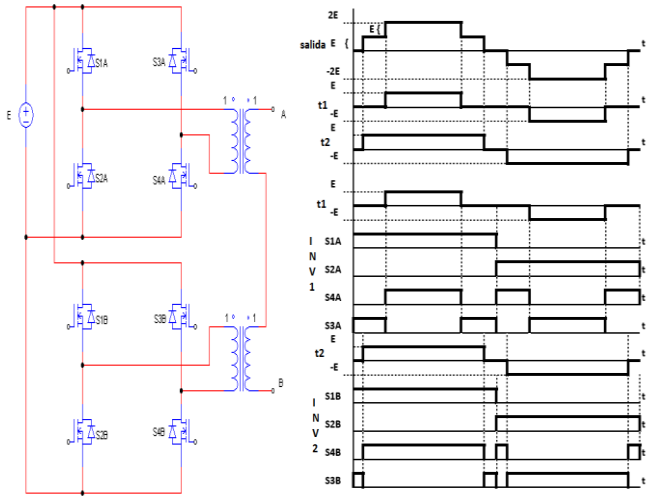


Figura 9. Circuito de potencia y secuencia de disparos para un inversor de cinco niveles

Tabla 2. Número de niveles según el número de etapas

ETAPAS (N)	NIVELES (3^N)
1	3
2	9
3	27
4	81

3.2 Eliminación de armónicos

Al generar varios escalones a la salida del inversor, es importante tomar en cuenta el ángulo de disparo de cada nivel, debido a que basándose en relaciones matemáticas es posible eliminar armónicos.

El análisis de Fourier permite obtener el valor de los coeficientes de la tensión de salida como la suma de cada uno de los coeficientes de cada forma rectangular.

$$H(n) = \frac{4}{\pi n}[E_1\cos(n\theta_1) + E_2\cos(n\theta_2) + E_3\cos(n\theta_3) + \cdots + E_s\cos(n\theta_s)] \quad (1.1)[9]$$

Donde:

Hes la función de la onda de voltaje en términos de Fourier.
n es un número impar.
s es el número de fuentes.
$\theta_1, \theta_2, \theta_3, \theta_s$son los ángulos de disparo de cada escalón.
E_1, E_2, E_3, E_sson los valores de voltaje de cada escalón respecto al anterior.

Debido a la simetría de la forma de onda en la tensión de salida, sólo existen armónicos de orden impar. La ecuación (1.1) permite obtener el sistema de ecuaciones para calcular los ángulos de conmutación. Como ejemplo, para un inversor de 7 niveles trifásico se tiene la ecuación 1.2 [4]:

$$H(n) = \frac{4E}{\pi n}[\cos(n\theta_1) + \cos(n\theta_2) + \cos(n\theta_3)] \quad (1.2)$$

En este ejemplo se toma al voltaje VCD=E constante para todas las 3 fuentes. A partir de esta ecuación se define las siguientes 3 ecuaciones [4]:

La componente fundamental con un índice de modulación M:

$$\frac{sM\pi}{4} = [\cos(n\theta_1) + \cos(n\theta_2) + \cos(n\theta_3)] \quad (1.3)$$

Donde:

Mes el índice de modulación
n es un número impar
s es el número de fuentes

Las ecuaciones no lineales para las componentes impares:

$$0 = [\cos(5\theta_1) + \cos(5\theta_2) + \cos(5\theta_3)] \quad (1.4)$$
$$0 = [\cos(7\theta_1) + \cos(7\theta_2) + \cos(7\theta_3)] \quad (1.5)$$

Con estas ecuaciones se pueden calcular los ángulos que deben estar comprendidos entre 0 y π lo mismo que se puede obtener a partir de un programa elaborado en MATLAB® basándose en aproximaciones sucesivas de las variables. La siguiente tabla muestra los ángulos de conmutación calculados para distintos índices de modulación. En este caso se eliminan los armónicos 5^{to} y 7^{mo}.

Tabla 3. Ángulos de disparo para diferentes índices de modulación [9]

Índice de modulación M	α_1	α_2	α_3
0,8	29,5°	54,53°	64,56°
0,9	17,53°	43,08°	64,14°
1,0	11,68°	31,18°	58,58°

Donde:
$\propto_1, \propto_2, \propto_3$son los ángulos de disparo para un inversor de siete niveles con eliminación selectiva de armónicos.

4. INVERSOR DE 81 NIVELES [5] [8] [4]

Para obtener 81 niveles a la salida del inversor (40 niveles positivos, 40 niveles negativos y el nivel 0) se usan cuatro inversores de tres niveles, conectados a transformadores cuyas ganancias sean potencias de 3, es decir, la salida del transformador de un inversor debe ser tres veces la salida del anterior.

4.1 Diseño de los Transformadores [12]

El voltaje RMS a la salida del inversor es de 120V, con este dato se debe determinar el valor de cada escalón, para esto, en primer lugar se calcula el voltaje pico de la onda:

$$V_p = \sqrt{2}\cdot V_{rms} \quad (1.6)$$
$$V_p = \sqrt{2}\cdot 120V = 169.7V$$

Este voltaje se lo divide en 40 partes iguales debido a que se tiene solo una fuente de voltaje CD.

$$V_{escalon} = \frac{V_p}{40} \qquad (1.7)$$

$$V_{escalon} = \frac{169.7V}{40} = 4.24V$$

El $V_{escalon}$es la base del cálculo para los voltajes de salida de los transformadores, basándose en este valor, se calcula los valores para la salida de cada inversor:

$$V_{inv-x} = 3^n \cdot V_{escalon} \quad (1.8)$$

Donde n = 0, 1, 2, 3.

$$V_{inv\ PRINCIPAL} = 3^3 \cdot 4.24V = 114.48V$$

$$V_{inv\ AUXILIAR1} = 3^2 \cdot 4.24V = 38.16V$$

$$V_{inv\ AUXILIAR2} = 3^1 \cdot 4.24V = 12.72V$$

$$V_{inv\ AUXILIAR3} = 3^0 \cdot 4.24V = 4.24V$$

El voltaje de entrada pico para la entrada de todos los transformadores es de 12V, por lo tanto para obtener las relaciones de transformación se lo hace de la siguiente manera:

$$N_x = \frac{V_{in}}{V_{inv-x}} = \frac{I_o}{I_{inv-x}} \qquad (1.9)$$

Donde:

Nes la relación de transformador
$x$es el número de inversor

$$N_{PRINC} = \frac{12V}{114.48V} = 0.1048$$

$$N_{AUX1} = \frac{12V}{38.16V} = 0.3145$$

$$N_{AUX2} = \frac{12V}{12.72V} = 0.9434$$

$$N_{AUX3} = \frac{12V}{4.24V} = 2.8302$$

También es importante determinar las corrientes de trabajo de cada uno de los transformadores ya que en base a esto se determina el calibre del conductor, el tamaño del núcleo y por ende el tamaño del transformador. Se toma como premisa que a la salida el inversor debe operar con una corriente de 5A, por lo tanto, utilizando la ecuación 1.9, las corrientes de entrada de los transformadores son las siguientes:

$$I_{inv-x} = \frac{I_o}{N}$$

$$I_{PRINC} = \frac{5A}{0.1048} = 47.71A$$

$$I_{AUX1} = \frac{5A}{0.3145} = 15.90A$$

$$I_{AUX2} = \frac{5A}{0.9434} = 5.30A$$

$$I_{AUX3} = \frac{5A}{2.8302} = 1.77A$$

La distribución de la potencia que consume la carga entre los diferentes transformadores es aproximadamente:

$$T_{PRINCIPAL} = 67.5\%$$

$$T_{AUXILIAR1} = 22.5\%$$

$$T_{AUXILIAR2} = 7.5\%$$

$$T_{AUXILIAR3} = 2.5\%$$

4.2 Diseño de los MOSFET[11]

Con los parámetros eléctricos obtenidos para el diseño de los transformadores, se realiza el diseño de los interruptores de potencia para los puentes inversores. Tabla IV muestra los MOSFET elegidos con los requerimientos de diseño y con sus características técnicas.

Tabla 4. Elementos utilizados para cada uno de los inversores.

Etapa	**Auxiliar 3**	**Auxiliar 2**	**Auxiliar 1**	**Principal**
Necesidad	**V_{DSS}**=Tensión drain – source: 12 V **I_D**=Corriente continua de colector: 1,77 A	**V_{DSS}**=Tensión drain – source: 12 V **I_D**=Corriente continua de colector: 5,32A	**V_{DSS}**=Tensión drain – source: 12 V **I_D**=Corriente continua de colector: 16,13A	**V_{DSS}**=Tensión drain – source: 12 V **I_D**=Corriente continua de colector: 47,62A
Elemento	IRFZ46N	IRFZ46N	IRFZ46N	IRFP2907
Caracterís-ticas	VDS=55V ID=53A	VDS=55V ID=53A	VDS=55V ID=53A	VDS=75V ID=209A

4.3 Diseño del Software [6]

Para el diseño del software es necesario referirse a la forma de la onda de salida la misma que se la obtiene al dividir un periodo de una onda sinusoidal en los 81 niveles de amplitud constante que se espera conseguir con este tipo de inversor. En la Figura 10 se observa la forma de onda de voltaje de salida para el semiperiodo positivo.

Los ángulos para cada uno de los pasos se calcularon discretizando un periodo de una onda sinusoidal en 81 pasos de voltaje iguales, 40 positivos, 40 negativos y el 0. Para lograr esta discretización se utiliza un programa realizado en

MATLAB® que se basa en la ecuación 1.1 para calcular los ángulos para eliminar los 40 primeros armónicos impares.

El dispositivo que se utiliza para realizar el programa es el microcontroladorATMEGA 8. Esto debido a que el número de puertos que este dispone es suficiente para el propósito de este trabajo. También se utiliza un microcontrolador adicional, el ATMEGA 16 que se encarga de colocar los tiempos muertos y de generar los complementos de cada una de las señales de control. Esto se debe a que unas señales son el complemento de las otras. La Figura 11 muestra cómo funcionan los microcontroladores para realizar el control del inversor [1].

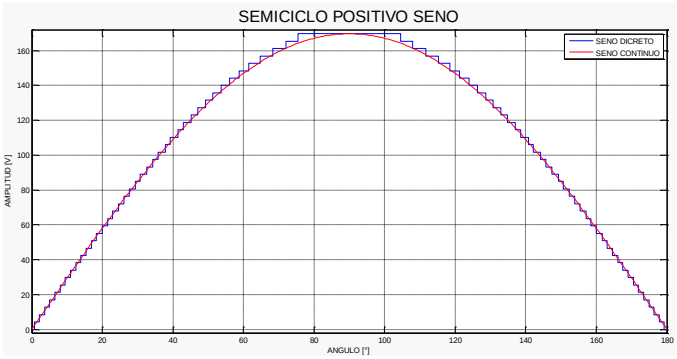


Figura 10. Semiperiodo positivo del voltaje de salida

La frecuencia a la que trabajan los sistemas microprocesados es 16MHz debido a la gran cantidad de señales que deben generar. La utilización de dos microcontroladores permite calibrar los tiempos muertos que se van a utilizar de manera independiente a la duración del programa que genera las señales de control.

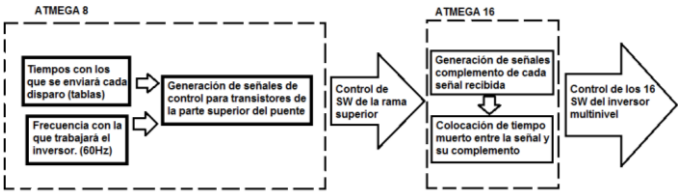


Figura 11. Esquema del funcionamiento del control

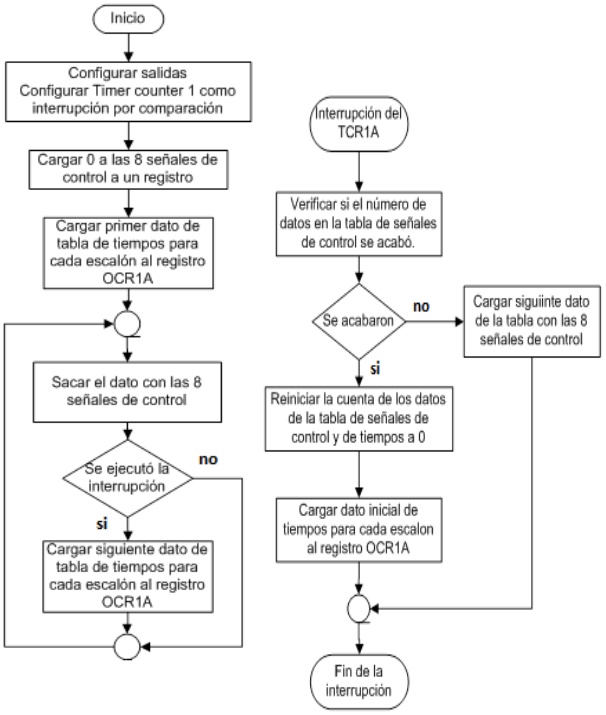


Figura 12. Diagrama de flujo del programa que genera 8 señales de control

Se observa en la Figura 12 diagrama de flujo del algoritmo de control implementado para generar 8 señales de control, las cuales mediante el segundo microcontrolador y basándose en el diagrama de flujo de la Fig. 13, son tomadas, invertidas y colocadas tiempos muertos para de esta manera obtener 16 señales de control (4 para cada puente H) que manejarán al bloque de potencia del inversor. Estas señales deben ser aisladas y acondicionadas debidamente para que puedan ser receptadas por los MOSFET que componen los puentes.

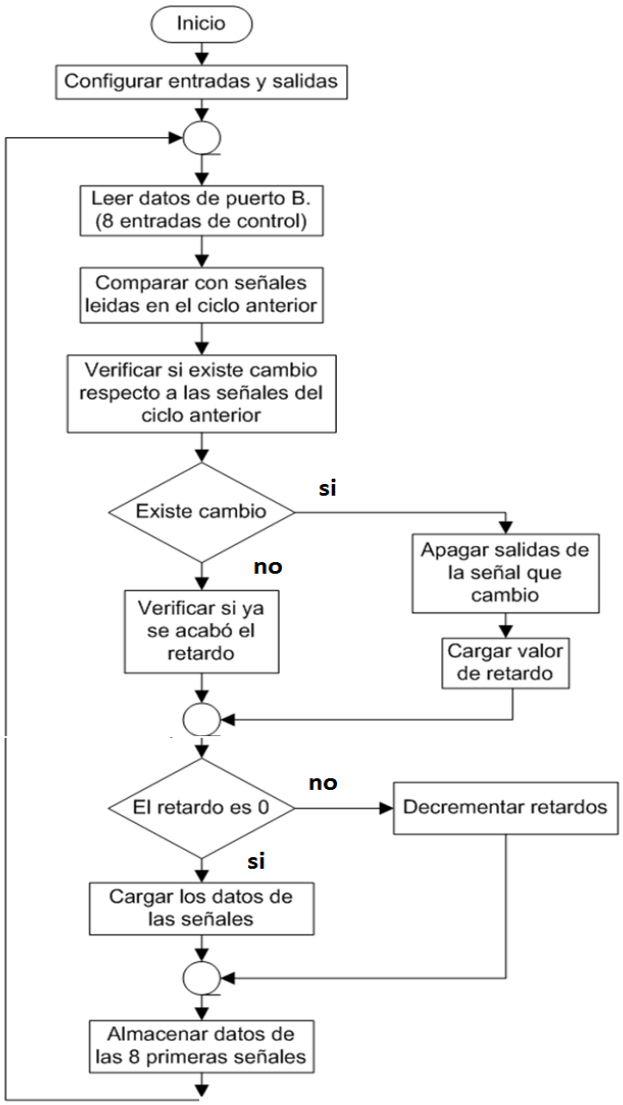


Figura 13. Diagrama de flujo del programa de inversión de señales y tiempos muertos

5. Módulo Implementado

Para la construcción del prototipo hay que tomar en cuenta que cada una de las partes debe estar diseñada para funcionar en conjunto evitando que sean susceptibles a interferencias electromagnéticas y a pérdidas de las señales de control debido a la excesiva longitud de los conductores [2].

Se observa en la Figura 14 el módulo didáctico de inversor multinivel implementado completamente.

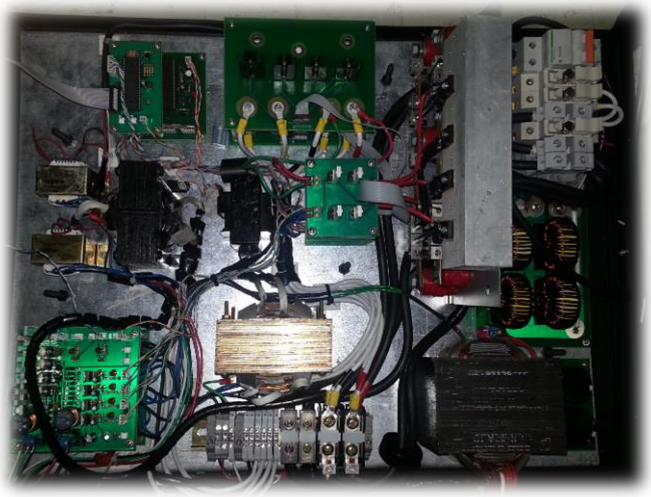


Figura 14. Módulo didáctico del inversor multinivel

6. Resultados Obtenidos

Para que el funcionamiento del módulo didáctico se pueda observar de manera adecuada, el control de éste permite elegir entre un número de niveles en potencias de 3 para observar los cambios en la distorsión armónica que se genera cuando se trabaja con un mayor número de escalones.

6.1 Pruebas con carga R

A continuación se presentan las formas de onda y el análisis de armónicos para la salida del inversor con diferente número de niveles con carga R (en este caso la forma de onda de voltaje es igual a la de corriente y ambas están en fase).

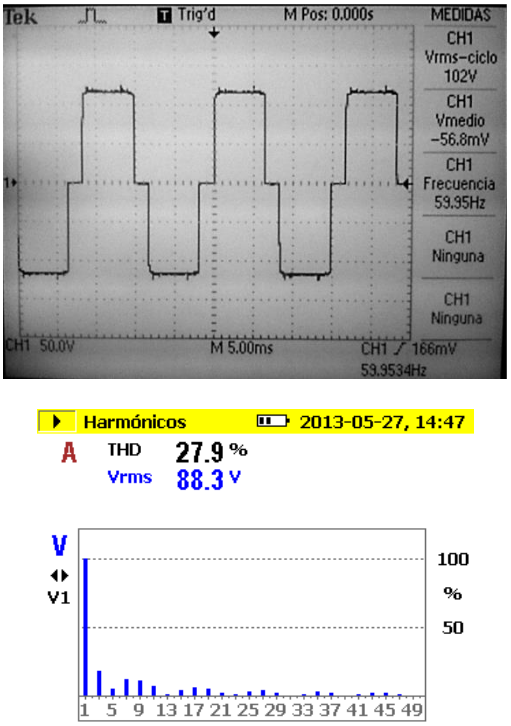


Figura 15. Formas de onda y armónicos para 3 niveles

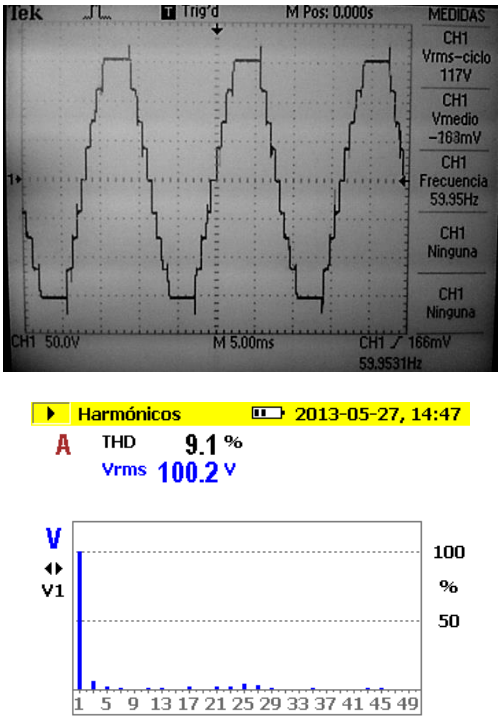


Figura 16. Formas de onda y armónicos para 9 niveles

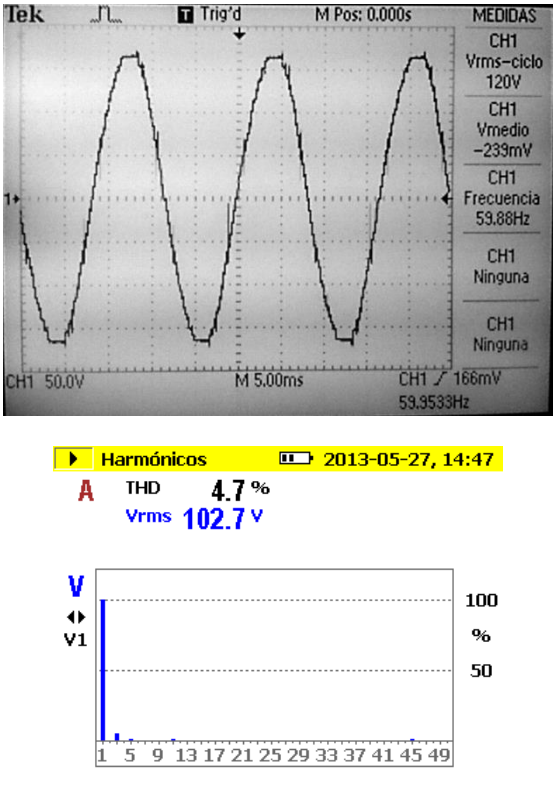


Figura 17. Formas de onda y armónicos para 27 niveles

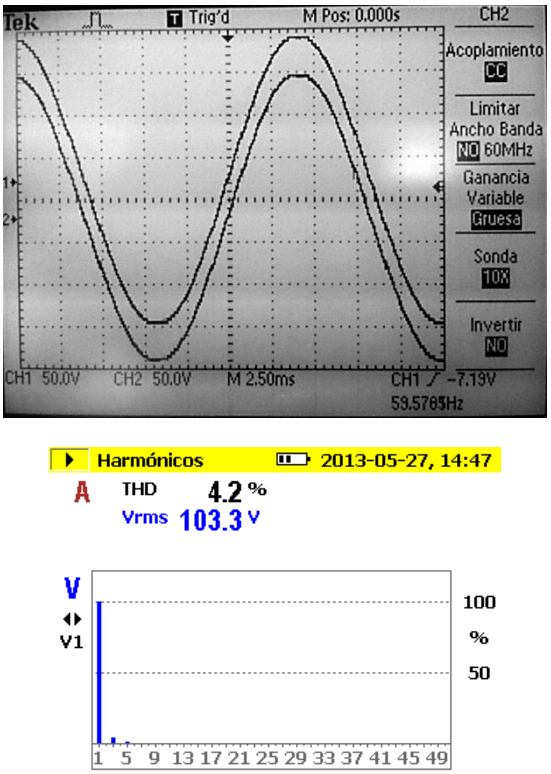


Figura 18. Formas de onda y armónicos para 81 niveles

Al realizar las pruebas se observó que conforme varían los niveles de la señal de salida la onda tiende a parecerse cada vez más a una señal sinusoidal. El voltaje de salida aumenta conforme aumenta el número de niveles esto se debe a que al tener una onda más sinusoidal el área bajo la curva aumenta.

Si se trata de medir la distorsión armónica esta disminuye conforme aumenta el número de niveles. La variación del número de niveles de la salida se logra gracias al control de algunos relés que permiten que los secundarios de los transformadores se conecten a la salida del prototipo o no.

6.2 Pruebas con carga RL

A continuación se presentan las formas de onda para la salida del inversor con diferente número de niveles con carga RL.

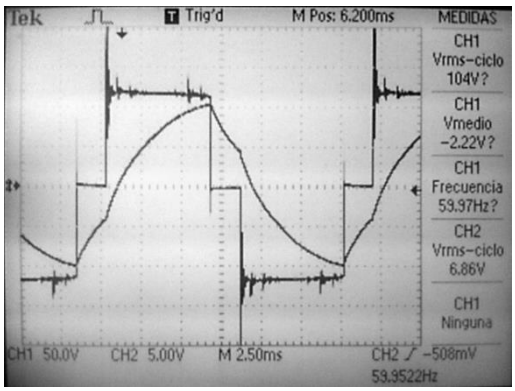


Figura 19. Formas de onda para 3 niveles

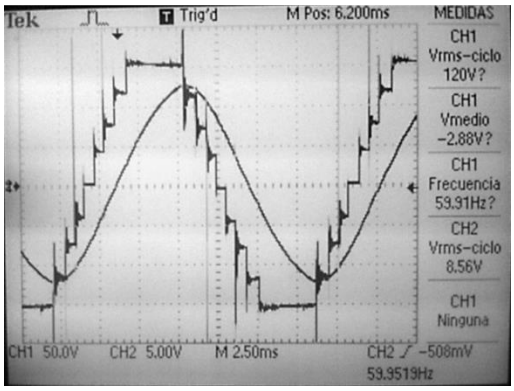


Figura 20. Formas de onda para 9 niveles

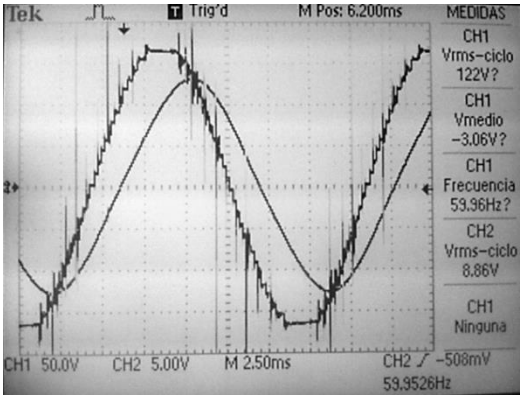


Figura 21. Formas de onda para 27 niveles

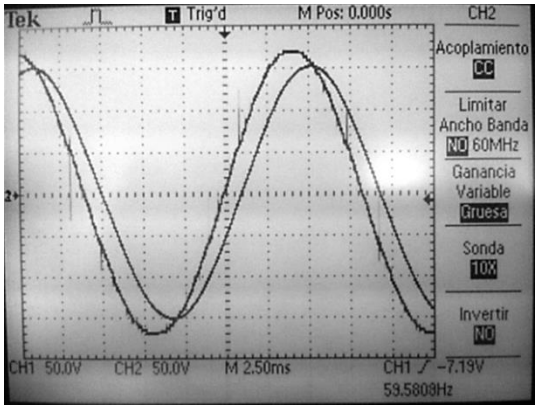


Figura 22. Formas de onda para 81 niveles

En este caso la potencia activa disminuye debido a que parte del voltaje cae en la inductancia. Conforme aumenta el número de niveles a la salida del inversor la forma de onda de corriente tiende a ser más parecida a sinusoidal esto debido a la carga de tipo inductivo.

En el caso del voltaje éste presenta una forma de onda muy similar a una función seno cuando se trabaja a 81 niveles. Esta onda presenta picos de voltaje debido a que la carga inductiva afecta al disparo de los MOSFET

Al realizar mediciones de armónicos el THD disminuye conforme aumenta el número de niveles, la carga inductiva hace que la corriente tenga una onda similar a una sinusoidal a pesar que se esté trabajando a 3 niveles.

Cabe recalcar que se está realizando el análisis con un analizador industrial Fluke 1735 que trabaja a una frecuencia de muestreo de 10,24kHz [7].

La Tabla 5 muestra la distorsión armónica total (THD) calculada en función del número de niveles a la salida del inversor, se observa que mientras el número de niveles se incrementa, el THD reduce considerablemente teniendo a la salida una onda de muy buena calidad.

Tabla 5. THD según el número de niveles

NÚMERO DE NIVELES	THD (%)
3	24.6
9	6.8
27	1.9
81	0.3

7. CONCLUSIONES

El objetivo del proyecto es construir un módulo didáctico que soporte una carga de 500VA a 120Vrms. Todos los elementos han sido dimensionados para soportar dicha carga por lo que se cumple dicho objetivo.

Se obtiene una señal de 81 niveles de voltaje CA que tiene una baja distorsión armónica.

Los filtros que usan los conversores CD/CA comerciales se basan en redes de elementos pasivos o activos, mientras que el usado en este trabajo es un mayor número de inversores simples a los que se les aplica una técnica de control que pretende sumar niveles conectando a los inversores en cascada, por lo que no es necesario un filtro a la salida del mismo.

El contenido armónico a la salida del inversor de 81 niveles depende del ángulo de disparo de cada nivel (en el caso de que cada nivel tenga la misma amplitud), es por eso que se hace una eliminación selectiva de armónicos sintetizando un programa en MATLAB® para la solución de un sistema de cuarenta ecuaciones para encontrar los ángulos en que se debe activar cada nivel.

Según se aumenta el número de niveles de la señal de salida la distorsión armónica disminuye lo cual cumple con el hecho de que conforme la señal de salida se parece más a una sinusoide pura el THD es cada vez menor.

REFERENCIAS

[1] ATMEL, "ATMEGA 164P/324P/644P" [Online]. Available: http://www.atmel.com/images/doc7674.pdf.

[2] Audio-Technia, "Tipos de interferencia" [Online]. Available:http://www.audio-technica.com/cms/site/26971e97bf9429d3/index.html.

[3] BALOGH, LASZLO, "Design And Application Guide For High Speed MOSFET Gate Drive Circuits" [Online]. Available:http://www.ti.com/lit/ml/slup169/slup169.pdf.

[4] BARCENAS, Ernesto, "Análisis y desarrollo de un inversor multinivel" [Online]. Available:http://www.cenidet.edu.mx/subaca/web-elec/tesis_mc/126MC_ebb.pdf.

[5] BRETON, Alberto, "Diseño y construcción de un Inversor trifásico multinivel de cuatro etapas para compensación armónica y de reactivos" [Online]. Available:http://web.ing.puc.cl/~power/paperspdf/dixon/tesis/Breton.pdf.

[6] CALDERON, Wilber, "Inversor de cuatro etapas" [Online]. Available:http:// ewh.ieee.org/r9/el_salvador/concapan/descargas/memoria_secciones/ Viernes_11/boqueron/P65.pdf.

[7] FLUKE, "Fluke 1735 Three-Phase Power Logger" [Online]. Available:http://support.fluke.com/find-sales/download/asset/2634932_6112_eng_a_w.pdf.

[8] NEGRONI VERA, Juan José, "Análisis y Diseño de Controladores para Inversores Multinivel en Sistemas Fotovoltaicos Conectados a Red" [Online]. Available:http://pels.edv.uniovi.es/pels/pels/Pdf/Tesis/Negroni_thesis.pdf.

[9] QUISHPE, Enrique, "Efectos del Desbalance y los Armónicos de Voltaje en la Operación de los Motores de Inducción" [Online].Available:http:// www.enriquequispe.com/.../2004Quispe_Armonicos_Desequilibrio_I Motores.pdf.

[10] RASHID, Mohamad, "Electrónica de Potencia, Circuitos, dispositivos y aplicaciones". México. 3ª Edición. Prince Hall, 2004.

[11] SAEI, "Inversores Modulados" [Online]. Available:http://pels.edv.uniovi.es/pels/pels/Pdf/inversores_modulados.pdf.

[12] UNIVERSIDAD DE NAVARRA, "Diseño y Fabricación de un Transformador Monofásico". [Online]. Available: http://www.tecnun.es/asignaturas/SistElec/Practicas/Presentacion_09_10.pdf.

Diseño y Construcción de un Sistema Portátil de Generación Eléctrica para Cargar un Banco de Baterías

Molina I.*, Chico P., Rosales A.**, Ávalos E.*****

**Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica*
Quito, Ecuador e-mail: gatoims@yahoo.es
*** Escuela Politécnica Nacional, Departamento de Automatización y Control Industrial*
Quito, Ecuador e-mail: {patricio.chico; andres.rosales}@epn.edu.ec
****Escuela Politécnica Nacional, Departamento de Física*
Quito, Ecuador; e-mail: Eduardo.avalos@epn.edu.ec

Resumen:*El trabajo desarrollado consiste en el diseño y construcción de una máquina portátil que permita cargar un banco de baterías tanto desde un grupo motor-generador, así como también desde la red eléctrica. Además incluye el dimensionamiento de dicho banco de baterías que se encargará de suministrar la suficiente energía para alimentar a un láser de nitrógeno y un computador portátil por medio de un inversor comercial. Estos equipos son necesarios para realizar estudios de espectroscopia en lugares donde la red eléctrica no está disponible.*

Palabras Clave: *grupo electrógeno, baterías, conversor dc/dc reductor, inversor, alternador, cargador de baterías.*

Abstract:*Thispaperdescribes the design of hardware, software and the construction of a portable device that charges a battery bank from a motor-generator unit and also from the power grid. It also includes the design of the battery bank, since it will supply enough energy to power a nitrogen laser unit and a laptop using a commercial inverter. This equipment is required to perform spectroscopy studies in places where the power grid is not available*.

Keywords:*generator, batteries, buck converter, inverter, alternator, battery charger*

1. INTRODUCCIÓN

En el presente trabajo se abordan los procesos de diseño y construcción de un prototipo para cargar un banco de baterías desde dos fuentes de energía diferentes: por medio de un grupo motor - generador, o desde la red de distribución eléctrica. El banco de baterías debe ser capaz de suministrar durante dos horas aproximadamente la suficiente energía para alimentar a un láser de nitrógeno y un computador portátil, los cuales consumen una potencia de hasta 150W máximo.

Para reestablecer la energía del banco de baterías de manera adecuada y sin comprometer la vida útil de las mismas, ya sea desde el grupo motor-generador o desde la red eléctrica, se requiere de un sistema de control que esté monitoreando durante el proceso de carga las siguientes variables en las baterías: voltaje, temperatura y corriente; y en base a ellas tomar las respectivas acciones de control.

Se debe tener en cuenta que el tiempo de autonomía del grupo motor-generador es limitado debido a la capacidad del tanque de combustible, motivo por el cual se implementa una técnica de carga rápida. También se tiene la opción de cargar las baterías desde la red de distribución si está disponible, y este caso se utiliza una técnica de carga lenta.

2. DESCRIPCIÓN DEL SISTEMA

Según se muestra en el diagrama de la Fig. 1. la máquina consiste fundamentalmente en el sistema portátil de generación eléctrica, el sistema de control y el banco de baterías.

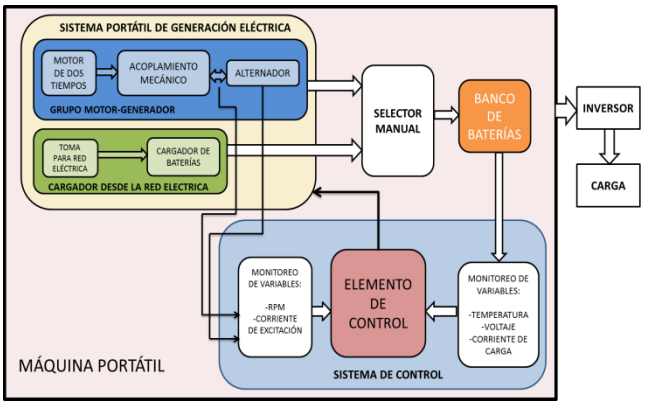


Figura. 1 Diagrama general de la estructura del proyecto

Dentro del sistema portátil de generación eléctrica, se encuentran el grupo motor-generador y el cargador de baterías que utiliza la red eléctrica como fuente, mientras que el sistema de control incluye el monitoreo de las variables del sistema las cuales son: voltaje, corriente de carga, temperatura, velocidad a la que gira el grupo motor-generador y corriente de excitación del alternador, y en base

a estas variables toma las respectivas acciones de control que permitan cargar de manera eficiente al banco de baterías.

3. SISTEMA PORTÁTIL DE GENERACIÓN ELÉCTRICA

El sistema portátil de generación eléctrica es el encargado de reestablecer la energía del banco de baterías cuando éste se encuentra descargado, la Fig. 2. muestra un diagrama correspondiente a este sistema.

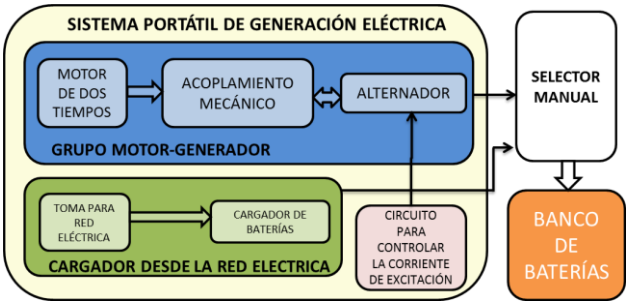


Figura. 2 Diagrama del sistema portátil de generación eléctrica

Dimensionamiento del Banco de Baterías

El banco de baterías se dimensiona para abastecer a la carga (láser de nitrógeno y un computador portátil), por un tiempo máximo de dos horas.

Se calcula el consumo total de energía en (W) de las cargas, como se muestra en la Tabla 1.

Tabla 1. Cálculo del consumo de las cargas en (W)

Tipo de equipo	Cantidad	Potencia [W]	Consumo [W]
Laptop	1	65	65
Láser de Nitrógeno	1	50	50
		Potencia total	**115**

Los equipos se conectan a través de un inversor comercial que en su hoja de datos especifica una eficiencia del 85% durante el proceso de conversión de corriente continua a corriente alterna. Además se sobredimensiona un 10% en capacidad, para contrarrestar la pérdida de eficiencia por deterioro de las baterías con el transcurso del tiempo; esto se nuestra en la Tabla 2. [2]

Tabla 2. Cálculo del consumo total de las cargas en (W)

Potencia (W)	Potencia/0.85	Factor de seguridad	Potencia total
115	135.3W	10%	148.8W

Para calcular la capacidad del banco de baterías se trabaja con la corriente de consumo de la carga, el tiempo de autonomía, así como la tasa de descarga. [1]

De acuerdo a las especificaciones de la energía que requiere la carga se calculó que la capacidad del banco de baterías debería ser de 24.8Ah para alimentar a la carga durante dos horas de respaldo, el valor comercial más cercano a la capacidad del banco es de 24Ah, para lo cual se utilizan dos baterías de plomo ácido VRLA (baterías recargables de plomo acido con válvula regulada para la gasificación y libres de mantenimiento) de 12V, 12Ah, que conectadas en paralelo da una capacidad de 24Ah.

Las baterías utilizadas se muestran en la Fig. 3.

Figura. 3 Batería de plomo-ácido VRLA. [2]

Diseño y Construcción del Grupo Motor-Generador

Para la construcción del grupo motor-generador, se destacan las siguientes partes fundamentales:

- Motor a combustión interna
- Alternador
- Acople mecánico (unión entre el motor y el alternador)
- Estructura de montaje (bancada)

1) *Motor a combustión interna*

El motor de dos tiempos de una motoguadaña modelo FS85 marca STIHL (mostrado en la Fig. 4), se ha seleccionado como motor de combustión interna, el cual entrega la suficiente potencia mecánica para mover al alternador y a su vez éste pueda convertir la energía mecánica en energía eléctrica.

Figura. 4 Motor de motoguadaña STIHL FS85

2) *Alternador*

Es una máquina eléctrica que transforma la energía mecánica en energía eléctrica, generando corriente alterna mediante inducción electromagnética. [3]

Para el trabajo desarrollado, se utiliza un alternador de automóvil, que suministra corriente trifásica rectificada para cargar al banco de baterías.

Considerando que la corriente máxima que se requiere para cargar el banco de baterías de acuerdo a la referencia [2] corresponde al 30% de la capacidad del banco (0.3*24Ah=7.2A), se utiliza un alternador compacto correspondiente al de un vehículo Daewo Tico (véase Fig. 5) cuyas características principales son: Voltaje nominal: 12V, corriente nominal: 50A, velocidad máxima de rotación: 18000rpm.

Figura. 5 Alternador compacto

Se aprecia que la corriente que puede entregar el alternador supera notablemente el requerimiento de la corriente necesaria para carga el banco de baterías, razón por la cual el alternador está sobredimensionado, pero es el más pequeño que se encontró en el mercado local y que se utiliza en sistemas eléctricos de vehículos basados en baterías de 12V.

El sobredimensionamiento tanto del motor de combustión interna, así como el alternador, permitiría cargar un banco de baterías de capacidad superior al utilizado para este trabajo, además de ello podría también alimentar directamente a una carga.

3) Acople mecánico

El acople mecánico constituye el dispositivo para unir el eje del motor de combustión interna con el eje del alternador. Para el desarrollo de este trabajo se diseña un acople directo por la facilidad que presenta en la construcción y adaptación tanto para el motor de combustión interna como para el alternador en un espacio físico limitado. [4] (Véase Fig. 6)

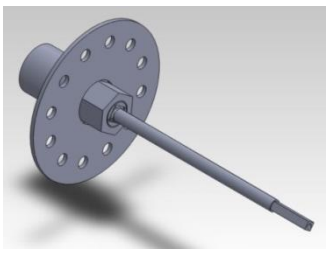


Figura. 6 Acople diseñado

Este tipo de acoplamiento requiere de una alineación precisa de los ejes, para que no pueda haber movimiento relativo entre ellos. Aprovechando el dispositivo de acople se diseña un disco que funcione como encoder con el fin de poder utilizarlo para monitorear la velocidad de giro del motor.

4) Estructura de montaje

Los equipos son montados en una plataforma que permite alinear y unir de manera correcta los ejes del motor de combustión interna así como del alternador a través del acople mecánico. La Fig. 7 muestra la plataforma sobre la cual se realiza el montaje de los equipos, cabe destacar que esta base es lo suficientemente robusta y pesada puesto que debe soportar el peso del alternador, el motor, el acople mecánico, el banco de baterías y las placas electrónicas.

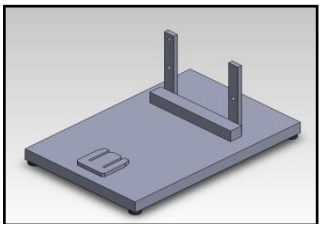


Figura. 7 Estructura de montaje

Finalmente en la Fig. 8 se muestra el montaje de los diferentes componentes. Se observa la palanca con la cual se comanda al motor de combustión interna que es sustituida por un mecanismo basado en un servomotor de alto torque marca Hitec HS-485HB, que se acopla perfectamente a las necesidades del sistema y será el encargado de variar la velocidad del motor de acuerdo a las necesidades de la aplicación(véase Fig. 9).

Figura. 8 Estructura de montaje incluido los equipos

Figura. 9 Servomotor Hitec HS-495HB

Cargador del Banco de Baterías

A más de haber diseñado el grupo motor-generador para cargar al banco de baterías cuando éste se encuentre en condiciones de descarga, se hace necesario el diseño y construcción de un cargador que use la red eléctrica como fuente; para el efecto se utiliza el diseño característico de las fuentes de alimentacion de voltaje, cuya estructura se muestra en la Fig. 10.

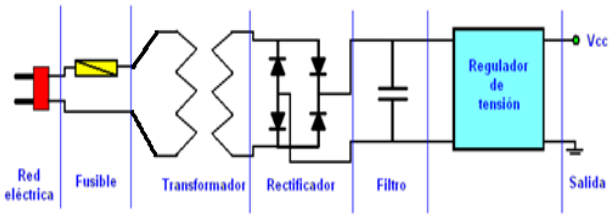


Figura. 10 Estructura de una fuente de alimentación

Este cargador está basado en un regulador de tension integrado LM338T. Entre las aplicaciones que se mencionan en las hojas de datos de este dispositivo se encuentra un cargador de baterías, y es el diseño que se tomó como referencia. [5]

El cargador de baterías implementado, entrega un voltaje constante de 14.5V y una corriente aproximada del 10% de la capacidad del banco de baterías (0.1*24Ah=2.4A), de tal manera de realizar una carga lenta. La ventaja de realizar una carga lenta es que se evita el sobrecalentamiento y por ende la gasificación de las baterías.

Circuito para controlar la Corriente de Excitación del Devanado de Campo del Alternador.

Para controlar la corriente de carga hacia el banco de baterías se debe controlar la corriente suministrada al devanado de campo del alternador. Para este fin se utiliza un conversor estático dc-dc reductor cuya carga es el devanado de campo y una resistencia R para limitar la corriente. (véase Fig. 11).

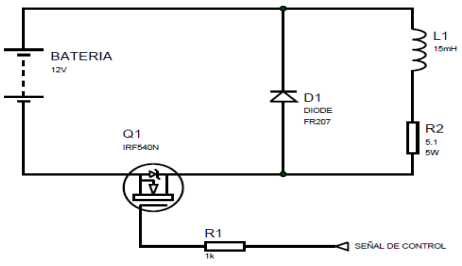


Figura. 11 Conversor dc-dc reductor carga RL

El circuito implementado que se muestra en la Fig. 11 ha sido diseñado de tal forma que la corriente que circula por el devanado de campo se encuentre dentro este rango: 0<I<1A. El dispositivo que permite controlar esa corriente de excitación es el MOSFET IRF540N, mediante una señal PWM de un microcontrolador. [6]

4. DESARROLLO DEL SISTEMA DE CONTROL

El sistema de control es el encargado de hacer cumplir los algoritmos de carga, y se muestra en la Fig. 12. Este sistema de control, permite reestablecer de manera eficiente y segura la energía que ha sido consumida por la carga desde el banco de baterías, teniendo siempre como prioridad salvaguardar la vida útil de éstas.

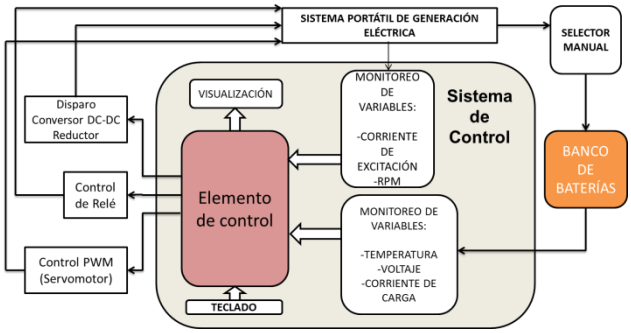


Figura. 12 Diagrama de la estructura del sistema de control

A. Variables Monitoreadas.

Las variables que requiere el sistema de control son:

- Voltaje
- Temperatura
- Corriente de carga
- Velocidad a la que gira el motor
- Corriente de excitación del alternador

El sistema de control toma las decisiones que permitan cumplir con el algoritmo de carga para cargar el banco de baterías, en base a las variables del sistema ya mencionadas anteriormente.

B. Elemento de Control

Para la ejecución de los algoritmos de carga, tanto el que constituye carga rápida a través del grupo motor-generador, así como la carga lenta a través de la red eléctrica, se requiere utilizar un dispositivo de control que se encargue de realizar las tareas u operaciones respectivas mediante el uso de software y hardware. Se utiliza el microcontrolador ATMEGA16, el cual se encargará de ejecutar los algoritmos de carga para el banco de baterías, además con el objetivo de visualizar los valores de las variables del sistema y conocer el estado en el que se encuentra el proceso utiliza un display de cristal líquido LCD, así como un teclado numérico para que el usuario del prototipo pueda manipularlo.

C. Diseño del Software del Sistema de Control

El software corresponde al programa desarrollado para la ejecucion del algoritmo de carga. Este programa maneja al

microcontrolador para que se realicen las tareas necesarias para el correcto funcionamiento del hardware.

El microcontrolador Atmega16 se programa para que procese las cinco señales monitoreadas del sistema, las cuales son voltaje, temperatura, corriente de carga, velocidad del motor (RPM), corriente de excitación y mediante el análisis de estas señales se toman las acciones de control para realizar la carga eficiente del banco de baterías evitando que se acorte la vida útil de las mismas.

El programa implementado también se encarga de manejar los periféricos externos como son el visalizador y el teclado, ambos necesarios para que el usuario pueda observar los parámetros del sistema y el proceso de carga del banco de baterías.

La carga del banco de baterías se realiza en dos etapas las cuales son carga a corriente constante (etapa bulk) y la etapa de absorcion. Dependiendo si el algoritmo de carga corresponde a una carga rápida o una carga lenta, los procesos son diferentes en cada una de ellas.

1) Algoritmo de carga rápida

Durante la primera etapa (etapa bulk o carga máxima) el grupo motor-generador entrega una elevada corriente de carga en corto tiempo con el fin de reestablecer la energía que ha sido descargada. Esta corriente de acuerdo a las especificaciones técnicas de las baterías no debe sobrepasar el valor del 30% de la capacidad del banco de baterias [2]. En esta etapa las baterías recuperan la mayoría de su capacidad (80%), y la temperatura de las baterías no debe sobrepasar del valor especificado por el fabricante (-10ºC a 50ºC).

En la etapa bulk el voltaje de las baterías se va incrementando, hasta alcanzar un voltaje de referencia (14.9V) para pasar a la siguiente etapa (absorción) que corresponde a una sobrecarga controlada. Lo que hace el algortimo de control una vez detectado el voltaje de referencia, es disminuir la corriente de carga, para así mantener un voltaje adecuado en el banco de baterías, esto se logra disminuyendo la velocidad de giro del motor y la corriente de excitación del devanado de campo del alternador. El proceso de absorción finaliza cuando la corriente de carga disminuye a un valor de 3-10mA/Ah cuando se carga a 2,4 vpc (voltios por celda) recomendado por el fabricante. El sistema carga a 2,49 vpc por lo que extrapolando para el banco implementado corresponde aproximadamente a un máximo de 480mA que es un 2% de la capacidad de las baterías, con lo cual la batería ya se encuentra cargada de acuerdo a las especificaciones del fabricante [2][10].

2) Algoritmo de carga lenta

El cargador de baterías que utiliza la red eléctrica como fuente, ha sido diseñado para una carga lenta, con lo cual la corriente inicial esta limitada a un 10% de la capacidad del banco de baterías (2.4A), y entrega un voltaje constante de carga de 14.5V.

Durante la etapa bulk, el voltaje del banco de baterías se va incrementando lentamente hasta alcanzar el voltaje aplicado por el cargador de baterías, cuando ello sucede se pasa a la etapa de absorcion, con lo cual la corriente de carga irá disminuyendo paulatinamente, y lo que hace el sistema de control es mostrar estas variables hasta que la corriente disminuya hasta un valor aproximado de de la capacidad del banco. Una vez que la corriente carga disminuyó al 2% de la capacidad del banco de baterías el sistema desconecta el cargador de baterías a través de un relé de conexión/desconexion del banco.

3) Diagrama de flujo del programa principal

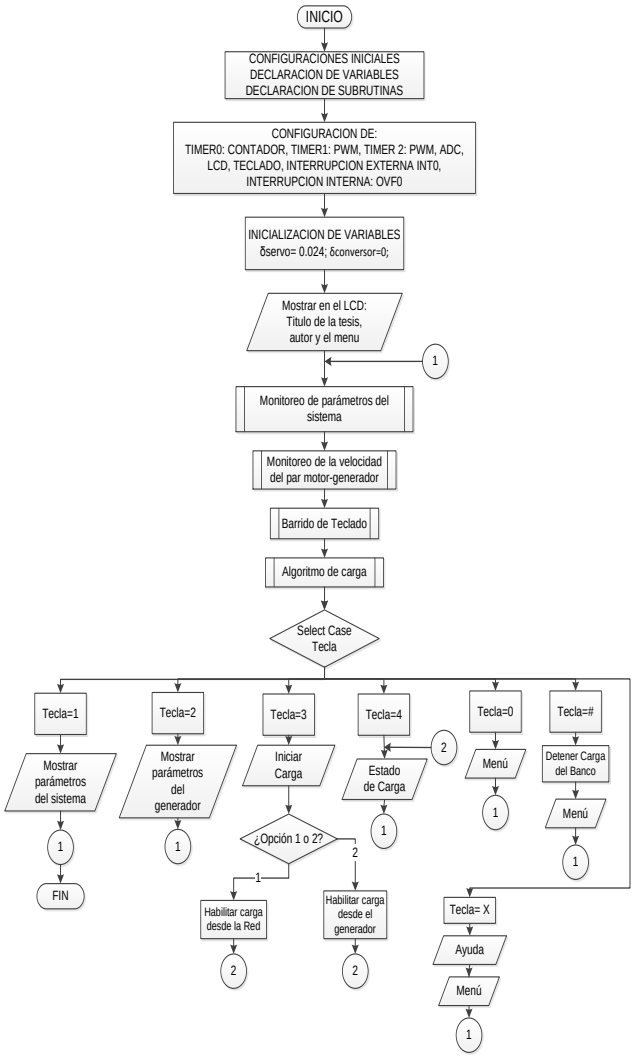


Figura. 13 Diagrama de flujo del programa principal

El diagrama de flujo del programa principal se muestra en la Fig. 13, donde se ha resumido el programa implementado en el microcontrolador, el cual permite el correcto funcionamiento del sistema. En este diagrama se han

utilizado subrutinas las cuales son parte del programa y que realizan una función específica de acuerdo al algoritmo de carga de las baterías.

5. PRUEBAS Y RESULTADOS

Las pruebas y resultados del proyecto se basan en la carga y descaga del banco de baterías, monitoreando las variables que intervienen durante el proceso de carga, las cuales son voltaje, temperatura y corriente principalmente.

A. Visualización de las variables del sistema

Figura. 14 Visualizacion de las variables del sistema

B. Formas de onda

La Fig. 15 muestra las forma de onda de las señales de control, la primera forma de onda (parte superior) corresponde a la señal PWM generada por el microcontrolador que se entrega al servomotor de alto torque marca Hitec HS-485HB, acoplado al dispositivo mecánico que varía la velocidad del motor de combustión interna, esta señal tiene una frecuencia de 50Hz de acuerdo a las especificaciones del fabricante, y el rango del ancho de pulso de acuerdo a las pruebas realizadas está comprendido entre 480us y 1.72m para lograr los desplazamientos angulares requeridos.

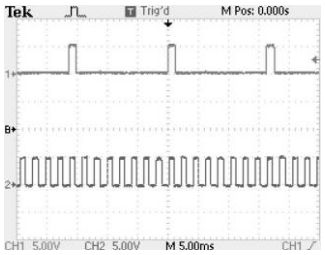


Figura. 15 Formas de onda de las señales de control

En la misma Fig. 15 la onda inferior corresponde a la señal PWM que se utliza para controlar la corriente de excitación del devanado de campo del alternador, la frecuencia de esta señal es de 1KHz. Esta frecuencia produce una salida de voltaje sin rizado apreciable, adecuada para la siguiente etapa del cargador.

En la Fig. 16 se muestran los pulsos obtenidos del encoder, señal digital utilizada para medir las revoluciones por minuto a las que gira el motor.

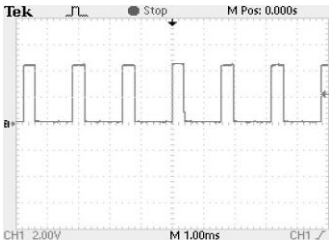


Figura. 16 Pulsos enviados por en encoder

En la Fig. 17 se muestra la forma de onda de la corriente de carga hacia el banco de baterías, utilizando el grupo motor-generador como fuente; mientras que la Fig. 18 muestra la forma de onda de la corriente de carga cuando se usa el cargador desde la red eléctrica.

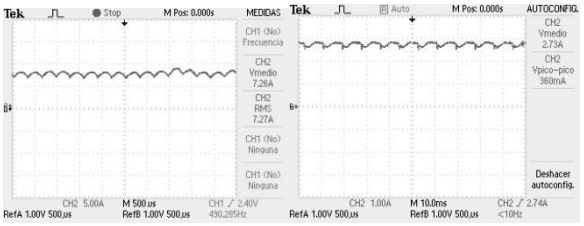


Figura. 17 Figura. 18

C. Resultados obtenidos durante el proceso de carga del banco de baterías a través de la red eléctrica.

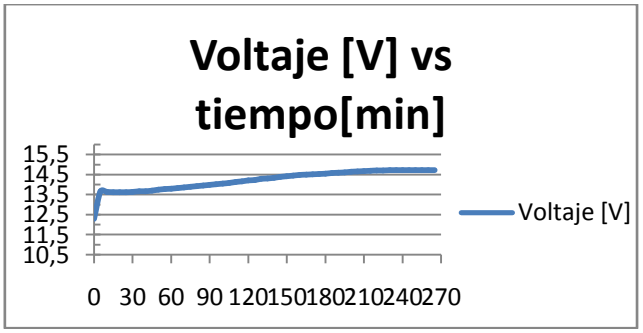


Figura. 19 Voltaje del banco de baterías en función del tiempo

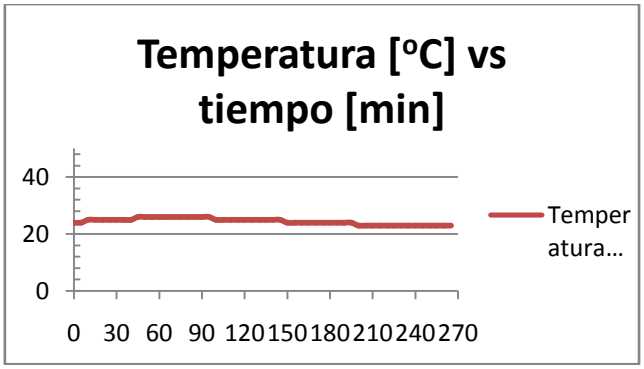


Figura. 20 Comportamiento de la temperatura durante el proceso de carga

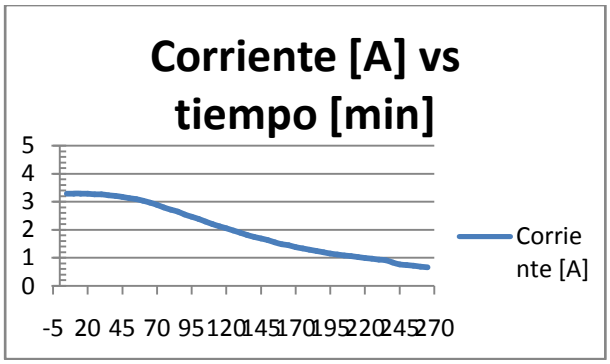


Figura. 21 Corriente de carga en función del tiempo

El proceso de carga duró 4 horas y 30 minutos, para un banco de baterías que se encontraba descargado aproximandamente un 40%. Durante el proceso de carga se puede observar que la temperatura se encuentra estable, y el voltaje se incrementa paulatinamente hasta alcanzar el voltaje de referencia, mientras que la corriente se mantiene constante hasta la primera hora de carga, para luego ir disminuyendo hasta llegar a un valor del 2% de la capacidad del banco de baterías, valor en el cual las baterías se encuentran completamente cargadas.

D. Resultados obtenidos durante el proceso de descarga del banco de baterías.

La prueba de descarga se realizó utilizando dos computadores portátiles que consumen 90W y 65W, respectivamente.

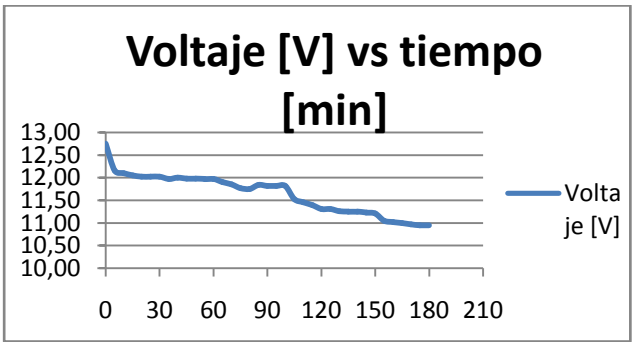


Figura. 22 Voltaje del banco de baterías en función de tiempo

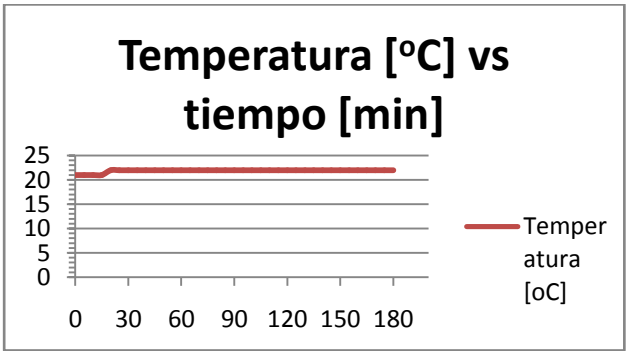


Figura. 23 Comportamiento de la temperatura durante el proceso de descarga

El tiempo de autonomía para las cargas conectadas fue de 3 horas, descargando las baterías hasta un voltaje de 10.9V, y claramente se observa en la Fig. 23 que la temperatura se mantiene estable.

E. Resultados obtenidos durante el proceso de carga del banco de baterías a través del grupo motor-generador.

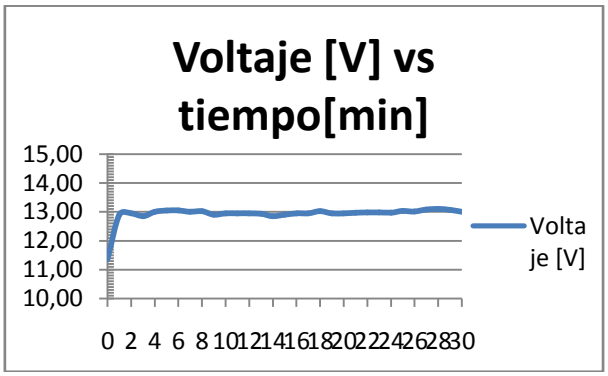


Figura. 24 Voltaje del banco de baterías en función del tiempo

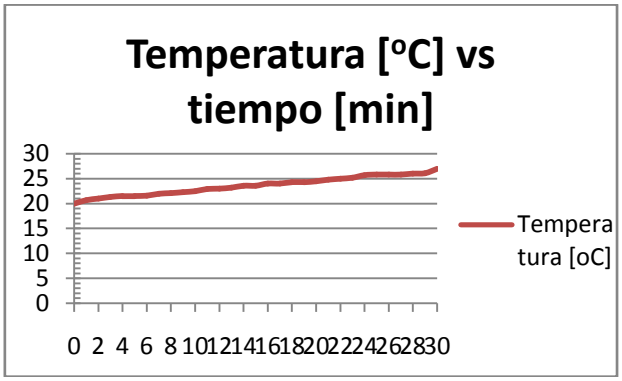


Figura. 25 Comportamiento de la temperatura durante el proceso de carga

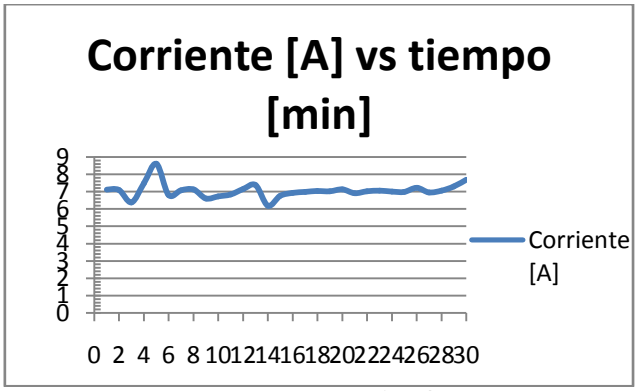


Figura. 26 Corriente de carga en función del tiempo

El proceso de carga duró 30 minutos, consumiéndose el 90% del combustible del motor. Al realizar la carga rápida con el grupo motor-generador durante ese periodo de tiempo se logró recuperar aproximandamente un 30% de la capacidad del banco de baterías. Durante el proceso de carga rápida se observaron ciertos picos de corriente (véase Fig. 26), esto se

producía cuando el movimiento del eje del alternador se afectaba por la vibración que ocasiona el motor de combustión interna, motivo por el cual el algoritmo de control considera un rango de histéresis para no desconectar el grupo motor-generador durante el proceso de carga.

En la Tabla. 3 se muestran los valores correspondientes a los ciclos de trabajo de las señales PWM generadas por el microcontrolador y también las rpm a las que giró el motor durante el proceso de carga.

Tabla. 3 Datos tomados durante el proceso de carga

Ciclo de trabajo (delta conversor)	**Ciclo de trabajo (delta servo)**	**RPM**
0,35	0,064	5020

F. Fotografías del proyecto

Figura. 27 Prototipo de la máquina portátil de generación eléctrica

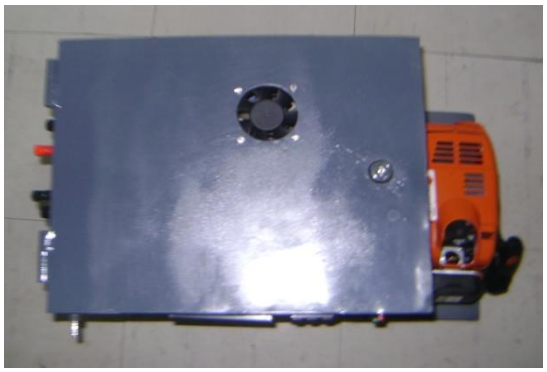


Figura. 28 Cara superior del prototipo

Figura. 29 Cara lateral izquierda del prototipo

Figura. 30 Cara lateral derecha del prototipo

Figura. 31 Interior del prototipo parte inferior

Figura. 32 Interior del prototipo parte superior

Figura. 33 Láser de nitrógeno. [9]

CONCLUSIONES

En este trabajo se realizó el diseño y construcción de un sistema portátil de generación eléctrica para cargar un banco de baterías desde dos fuentes de energía diferentes.

El banco de baterías es capaz de alimentar a una carga que consume una potencia de 150W. Esto lo demuestran las distintas pruebas realizadas, donde se hacen las experimentaciones de carga y descarga al banco de baterías.

El grupo motor-generador diseñado es capaz de entregar la suficiente cantidad de corriente para cargar las baterías en el modo de carga rápida.

El diseño del cargador de baterías que utiliza la red eléctrica como fuente que permite cargar las baterías en el modo de carga lenta.

El acople mecánico que constituye el dispositivo para unir al eje del motor con el eje del alternador, ha sido diseñado de tal manera que pueda ser implementado en un espacio físico limitado y que presente facilidad en la construcción y adaptación tanto para el motor como para el alternador, motivo por la cual se realizó un acoplamiento directo, y que funciona correctamente dentro del sistema.

La potencia requerida para cargar el banco de baterías es pequeña en comparación con la potencia que pueden entregar el alternador y el motor de combustión interna, motivo por el cual la velocidad con la que trabaja el grupo motor-generador no se ve afectada de manera significativa durante el proceso de carga. Es por esta razón que se decidió hacer trabajar al grupo motor-generador a una velocidad constante, sin embargo, si existe alguna falla o malfuncionamiento ocasionada por sobrevoltaje, sobrecorriente o sobretemperatura durante el proceso de carga del banco de baterías el lazo de control ajusta la velocidad del motor para evitar daños al sistema.

El grupo motor-generador que es parte del sistema portátil de generación eléctrica, ha sido sobredimensionado debido a que en el mercado local se encontraron únicamente componentes de mayor potencia, no obstante, cumplen con los requisitos necesarios para obtener la energía suficiente para cargar al banco de baterías.

El algoritmo de carga implementado, restaura la energía del banco de baterías en dos etapas las cuales son carga máxima o etapa bulk, y sobrecarga o absorción, teniendo siempre en cuenta salvaguardar la vida útil de las baterías.

La eficiencia del algoritmo de carga implementado depende principalmente de la lectura de las variables sensadas en el sistema las cuales son voltaje, corriente y temperatura. De acuerdo a estas variables el sistema toma las respectivas acciones de control que permitan cargar de manera eficiente el banco de baterías, ya sea desde el grupo motor-generador o desde la red eléctrica.

El grupo motor-generador se lo utiliza exclusivamente para cargar el banco de baterías, sin embargo debido al sobredimensionamiento de sus partes puedríaalimentar directamente una carga, es decir funcionar como un generador comercial. Como una mejora a futuro para lograr que esto sea posible se requiere modificar el sistema de control de velocidad en lazo cerrado, de tal manera que cuando se conecte una carga al generador, el sistema de control compense la caída de voltaje y la reducción parcial de la velocidad de giro del motor, como consecuencia de conectar dicha carga. En el diseño actual se usa el motor a velocidad constante y el sistema de control de velodidad toma acciones solamente cuando hay fallas o mal funcionamiento.

REFERENCIAS

[1] TODOPRODUCTIVIDAD Website (2011), [Online].Disponible: http://todoproductividad.blogspot.com/2011/02/como-calcular-la-capacidad-de-las.html.

[2] "FIRST POWER BATTERIES 12120 data sheet," [Online]. Disponible: http://www.tema-energy.ro/img/manuals/FP12120.pdf

[3] DR. ING. K. G. BURGER, ALTERNADORES BOSH, electricidad y electrónica del automóvil, edición 2000.

[4] ECURED homepage, 2010, Acoplamientos, [Online]. Disponible:http://www.ecured.cu/index.php/Acoplamiento_%28Mec%C3%A1nica%29.

[5] "5-Amp Adjustable Regulators LM338T data sheet," [Online]. Disponible: http://www.ti.com/lit/ds/symlink/lm138.pdf

[6] "International IOR Rectifier, IRF540N, Power MOSFET data sheet,"[Online]. Disponible: http://www.irf.com/product-info/datasheets/data/irf540n.pdf

[7] "Allegro, Hall Effect-Based Linear Current Sensor ACS715 data sheet," [Online]. Disponible: http://www.pololu.com/catalog/product/1186 http://www.allegromicro.com/~/media/Files/Datasheets/ACS715-Datasheet.ashx

[8] ROBÓTICA, encoder, website (2011), [Online]. Disponible: https://sites.google.com/site/proyectosroboticos/encoder

[9] "SRS Stanford Research Systems, NL100 Nitrogen Laser data sheet," [Online]. Disponible: http://www.thinksrs.com/downloads/PDFs/Manuals/NL100m.pdf

[10] "FirstPower Rechargeable Batteries Products Manual" [Online]. Disponible: http://www.efirstpower.com/download/FirstPower_2012.pdf

Medición de Ángulos de Inclinación por Medio de Fusión Sensorial Aplicando Filtro de Kalman

Pozo D. *; Sotomayor N. *; Rosero J.*;Morales L.*
**Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica*
e-mail: (david.pozo; nelson.sotomayo; jorge.rosero; luis.moralesec)@epn.edu.ec

Resumen:*Con el fin de obtener medidas angulares con el mínimo error y desviación posible se establece la fusión sensorial mediante el uso del Filtro de Kalman de las señales del acelerómetro y el giroscopio para un sistema en reposo sometido a perturbaciones (vibraciones). Finalmente el algoritmo es implementado en una plataforma que se basa en un microcontrolador dsPIC30F4011 que permite medir ángulos de inclinación en cabeceo y alabeo.*Además e*n este documento se presenta una metodología que permiteentender el manejo de los sensores inerciales (acelerómetro y giroscopio), así como también se da a conocer sus principales características ylosproblemas inherentes a su funcionamiento.*

Palabras clave:*Giroscopio, acelerómetro, Filtro Kalman, fusión sensorial, cabeceo, alabeo, IMU.*

Abstract:*In order to obtain the minimum error and deviation of an angular measurement, is established sensory fusion using the Kalman Filter between the accelerometer and the gyroscope signals for a stationary system under disturbances (vibrations).Finally, the algorithm is implemented on a platform that is based on a microcontroller dsPIC30F4011 to measure angles in roll and pitch. Additionally, this paper presents a methodology that enables a better understanding of the management of inertial sensors (accelerometer and gyroscope) and also discloses its main characteristics and problems inherent in their operation.*

Keywords:*Gyroscope, accelerometer, KalmanFilter, sensor fusion, pitch, roll, IMU.*

[1]

1. INTRODUCCION

Desde hace algunos años la industria de los sensores inerciales [1] ha ido creciendodebido al bajo coste que estos representan, bajo consumo de energía y al nivel de miniaturización al que se está llegando. En la actualidad el ámbito de uso de los sensores inerciales es muy amplio, cada día se presentan innovadoras propuestas y aplicaciones para su uso (Fig. 1), tanto así que estos sensores se encuentran presentes en dispositivos tan comunes para las personas como los celulares y sistemas de entretenimiento, en sistemas de seguridad automotriz como por ejemplo liberación del airbag, control de cinturones de seguridad, suspensión activa y control de tracción, en sistemas de estabilización de plataformas para video cámaras, en aplicaciones para realidad virtual y en sistemas de monitoreo de impacto como los existentes en dispositivos de almacenamiento de memoria (discos duros) [2].

Las áreas de investigación en las cuales se está trabajando están ligadas especialmente a sistemas de análisis de movimiento como el presentado en [3], donde los sensores inerciales son usados para el rastreo peatonal (caminata), además existen otros sistemas donde los investigadores se han enfocado en el estudio y análisis del movimiento del cuerpo humano [4],[5] o sólo en las extremidades superiores [6], en base a una serie de sensores inerciales ubicados de manera estratégica sobre una persona.

Por otra parte la utilización de sensores inerciales está presente en sistemas de posicionamiento y navegación [7] en ambientes tanto internos [8] como externos, siendo éstos últimos objeto de investigación especialmente en el campo militar y de seguridad con el uso de los así llamados UAV's (UnmannedAerialVehicle) [9].

Si bien los sensores inerciales se utilizan en varias aplicaciones, éstos presentan una serie de errores en sus medidas debido principalmente a su construcción [10], ya que mientras la miniaturización va en incremento (MEMS, ***MicroelectromechanicalSystems***), su desempeño puede verse disminuido. Es por esta razón que los sensores inerciales se utilizan junto a métodos y algoritmos que permiten obtener medidas fiables ya sean éstas de posición angular, desplazamiento, velocidad lineal, velocidad angular o aceleración[5],[6],[8],[11].
Uno de los métodos usados con mayor frecuencia para corregir errores de desviación en el tiempo en medidas inerciales es el denominado Filtro de Kalman[12].

.

En muchas ocasiones las técnicas presentadas en trabajos previos sobre aplicaciones del Filtro de Kalman a fusión sensorial resultan un tanto complejas ya que involucran la utilización de técnicas adicionales al FK aumentando la dificultad para el entendimiento del comportamiento del Filtro, es por tal motivo que en el presente trabajo se muestra una introducción a la fusión sensorial entre acelerómetro y giroscopio con el fin de obtener una medida angular de fácil implementaciónen base a la utilización del Filtro de Kalman.

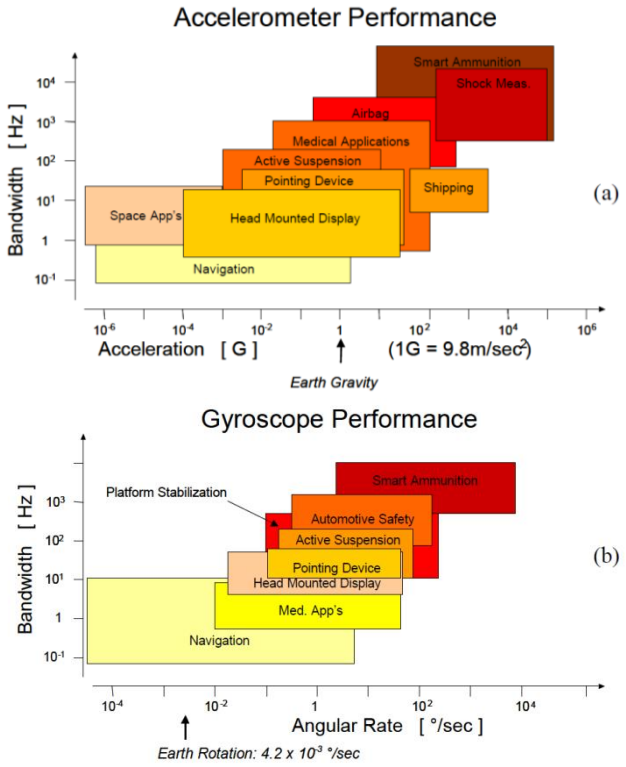


Figura 1. Desempeño requerido para sensores inerciales según la aplicación. a) Acelerómetro b) Giroscopio, tomado de [2]

1.1 Giroscopio

Es un sensor inercial que mide velocidad de rotación sobre su propio eje (velocidad angular). La manera como el giroscopio logra medir este movimiento rotacional puede ser mediante dos principios básicos: Vibratorios y Ópticos como se menciona y trata en [1], [10], [13], [14].

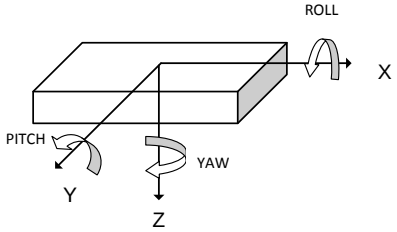


Figura 2. Esquema de un giroscopio de 3 ejes.

La resolución, la deriva, la salida en estado estacionario y el factor de escala son características que definen el desempeño del giroscopio. La resolución del sensor está definida por el ruido blanco y está expresada en términos de la desviación estándar o velocidad de rotación equivalente por la raíz cuadrada del ancho de banda. $\left[(°/s)/\sqrt{Hz}\ \ o\ \ (°/h)/\sqrt{Hz}\right]$, el denominado "angle random walk" puede ser usado también con este fin $[°/\sqrt{h}]$. La pequeña variación en la medida del sensor en estado estacionario (sin rotación) define la deriva a corto o largo plazo y está expresada en $°/s$ o $°/h$. El factor de escala se define como el rango de cambio en la señal de salida por unidad de cambio en la velocidad de rotación y está expresada en $V/(°/s)$. El último factor a ser tomado en cuenta es el ZRO (zero – rate output) o salida en estado estacionario, el mismo que representa la salida del dispositivo en ausencia de velocidad de rotación sobre su eje. En la ausencia de rotación, la señal de salida de un giroscopio está afectada por el ruido blanco y la deriva [10].

En general los giroscopios pueden ser clasificados según su desempeño en tres grandes grupos como se muestra en la Tabla 1.

Tabla 1. Clasificación Giroscopios, tomada de [10]

Parameter	Rate Grade	Tactical grade	Inertial Grade
Angle Random Walk	>0.5	0.5 – 0.05	<0.001
Bias Drift, $°/h$	10 - 1000	0.1 - 10	<0.01
Scale factor Accuracy, %	0.1 -1	0.01 – 0.1	<0.001
Full Scale Range $(°/s)$	50 – 1000	> 500	>400
Bandwidth, Hz	>70	~ 100	~ 100

Si se desea obtener posición angular (orientación) con respecto a un ángulo inicial conocido es necesario realizar un proceso de integración en el tiempo de la velocidad angular dada por el sensor. Hay que tomar en cuenta la variación de ángulo en intervalos cortos de tiempo como se muestra en la expresión (1).

$$d\theta = \omega * dt \quad (1)$$

Donde:
$\omega = velocidad\ angular\ medida\ por\ el\ giroscopio.$
$d\theta = ángulo\ medido.$
$dt = tiempo\ de\ muestreo.$

La expresión mostrada en (1) sólo es válida para instantes definidos de tiempo y no indica el ángulo total de giro alcanzado por el sensor sobre un eje debido a un movimiento en un tiempo "t". Por tanto es necesario realizar un algoritmo de carácter acumulativo en tiempo discreto que permita conocer el ángulo total alcanzado (actitud) en un rango de tiempo $t_0 \rightarrow t_f$ como se muestra a continuación.

$$\theta_k = \theta_{k-1} + \omega_k * dt \qquad k = 0,1,2,3 \ldots \ldots \qquad (2)$$

Donde:
$\theta_k = ángulo\ actual.$
$\theta_{k-1} = ángulo\ anterior.$
$\omega_k = velocidad\ angular\ actual.$

Existen dos inconvenientes en el momento de encontrar la actitud medianteeluso del giroscopio:el primero es la deriva, la misma que es propia del dispositivo y que debido a la recursividad de la expresión (2) provoca que la desviación del valor real en la medida del ángulo vaya incrementándose a lo largo del tiempo;el segundo inconveniente es, la falta de conocimiento de un estado inicial para θ_{k-1} o ángulo inicial de partida, el cual en el presente trabajo será asumido como "0"[1].

1.2 Acelerómetro

Un acelerómetro lineal es un sensor inercial que mide la diferencia total entre la aceleración traslacional y la componente de la aceleración gravitatoria a lo largo de su eje de acción [1]; típicamente tiene como unidad de medida [g], donde $1g = 9{,}8\ m/s^2$. Un acelerómetro incluye una masa suspendida mediante un muelle sobre un marco fijo, donde una fuerza externa producida por una aceleración lineal ocasiona un desplazamiento relativo entre la masa de prueba y el marco, creando al mismo tiempo cambios de tensión sobre el muelle de suspensión. Tanto el desplazamiento relativo como la tensión que sufre el muelle de suspensión al deformarse pueden ser usados para medir una aceleración externa[10], [13].Por otra parte, físicamente el acelerómetro puede ser modelado por un sistema de segundo orden: masa, amortiguador, muelle como se muestra en la Fig.3, en donde, además de la masa de prueba "M" y la constante de muelle "K" aparece un factor de amortiguamiento D (Disminución progresiva del movimiento del muelle) [10], [15].

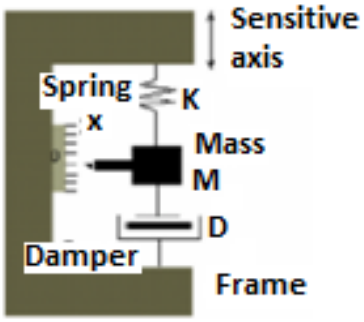


Figura 3. Representación mecánica de un acelerómetro, tomado de [15]

Los acelerómetros típicamente están especificados por su sensibilidad, máximo rango de operación, resolución, frecuencia de respuesta, no linealidad a máxima escala, offset ysupervivencia de choque[10],[16],[17].En la Tabla 2, se muestran parámetros de desempeño para acelerómetros en aplicaciones automotrices y de navegación.

Tabla 2. Especificaciones típicas de acelerómetros según su aplicación, tomada de [10]

Parameter	Automotive	Navigation
Range	$\pm 50g\ (airbag)$ $\pm 2g\ (vehicle\ stability\ system)$	$\pm 1g$
Frecuency Range	DC – 400Hz	DC – 100Hz
Resolution	$< 100\ mg\ (airbag)$ $< 10mg\ (vehicle\ stability\ System)$	$< 4\mu g$
Off – axis Sensitivity	< 5 %	< 0.1%
Nonlinearity	< 2%	< 0.1%
Max. Shock in 1msec	2000g	10g
Temperature range	$-40°C\ to\ 85°C$	$-40°C\ to\ 85°C$
Tc of offset	$< 60mg/°C$	$< 50\mu g/°C$
Tc of Sensitivity	$< 900ppm/°C$	$< \pm 50ppm/°C$

Estos sensores son capaces de medir ángulos (cabeceo y alabeo) con mucha precisión en condiciones estáticas, es decir, sin presencia de vibraciones ni aceleración lineal.

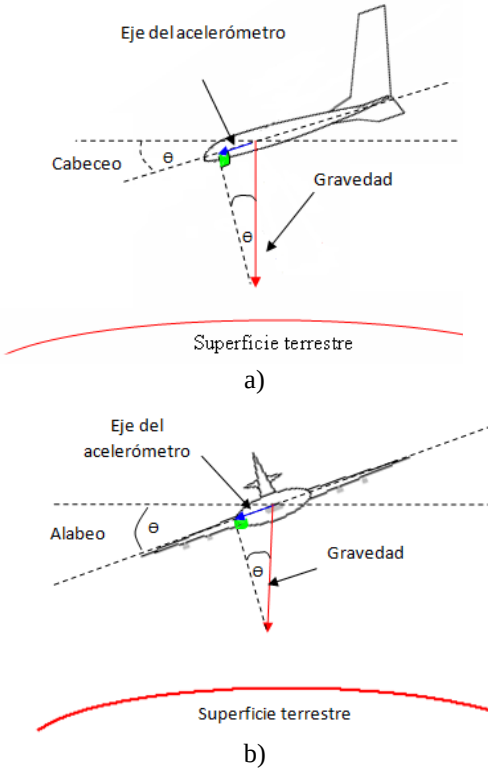


Figura 4. a) Cabeceo con un ángulo θ en sentido anti-horario. b) Alabeo con un ángulo θ en sentido anti-horario.

Si el acelerómetro se encuentra paralelo a la superficie terrestre, entonces los ejes de medición se encuentren perpendiculares a la gravedad, en cuyo caso en estado estacionario no deberían registrar aceleración alguna. Por otra parte si el sensor cambia su inclinación y si se toma en cuenta que no existe aceleración lineal (sistema en reposo), entonces aparecerán componentes de aceleración debido a la gravedadsobre los ejes "X" e "Y" del sensor, como se ilustra en las Figuras 4a y 4b. Para obtener el valor de inclinación en estado estacionario (reposo) se utilizan las expresiones (3) y (4) sobre cada eje, basadas en los ángulos de Euler representados en la Fig. 5.

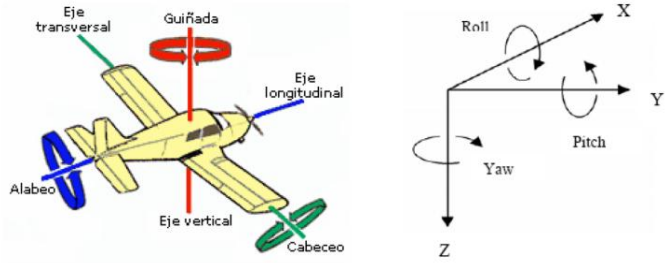


Figura 5. Sistema de Coordenadas y ángulos de Euler asociados [18].

$$Pitch = \sin^{-1}\left(\frac{\mathrm{Ax}}{1\mathrm{g}}\right) \quad (3)$$

$$\mathrm{Roll} = \sin^{-1}\left(\frac{Ay}{1g}\right) \quad (4)$$

donde:
$Ax = Componente\ de\ la\ gravedad\ medida\ sobre\ el\ eje\ x\ del\ acelerómetro$
$Ay = Componente\ de\ la\ gravedad\ medida\ sobre\ el\ eje\ y\ del\ acelerómetro$

El principal inconveniente al utilizar el acelerómetro para medir ángulos de cabeceo y alabeo es la presencia de perturbaciones (vibraciones) y componentes de aceleración lineal en el caso de existir movimiento traslacional en dirección de cualquiera de sus ejes, portanto,lavibración y la aceleración lineal provocarían que al utilizar las expresiones (3) y (4) se obtengan medidas de ángulo alteradas.

En secciones posteriores se establece el modelo matemático necesario para la medición angular en base a las características del giroscopio y acelerómetro y su posterior implementación con el Filtro de Kalman.

2. FUSIÓN SENSORIAL CON FILTRO DE KALMAN

El FK es un estimador óptimo que puede implementarse de manera sencilla en sistemas de carácter tanto lineal como no lineal, y cuyo procesamiento de datos es de carácter recursivo [19]. El FK realiza un proceso de predicción y otro de corrección mediante la medición y observación de un grupo de variables presentes en el sistema a tratar, dicho conjunto de variables forman el vector de estados y el observador.El algoritmo para el FK puede dividirse en dos grupos de ecuaciones a utilizarse[20]: las primeras son las ecuaciones que se actualizan en el tiempo o también llamadas de predicción y el segundo grupo se refiere a las ecuaciones de actualización mediante observaciones conocidas también como de corrección. Ambos grupos de Ecuaciones se muestran en la Tabla 3.

Tabla 3. Ecuaciones de predicción y Corrección para el Filtro Discreto de Kalman, tomada de [20]

Ecuaciones de predicción	
$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k + w_{k-1}$	(5)
$P_k^- = AP_{k-1}A^T + Q$	(6)
Ecuaciones de actualización	
$K_k = P_k^- H^T \left(HP_k^- H^T + R\right)^{-1}$	(7)
$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$	(8)
$P_k = (I - K_k H)P_k^-$	(9)

donde:
$k = 0,1,2,3, \ldots\ldots\ Intantes\ de\ tiempo$
$\hat{x}_k^- = estado\ estimado\ para\ el\ instante\ k$
$\hat{x}_{k-1} = estado\ en\ el\ instante\ k-1$
$A = Matriz\ de\ realimentación$
$B = Matriz\ de\ entrada$
$w_{k-1} = Representa\ el\ ruido\ inherente\ al\ proceso$
$P_k^- = Matriz\ estimada\ de\ covarianza\ en\ el\ instante\ k$
$P_{k-1} = Matriz\ de\ covarianza\ en\ el\ instante\ k-1$
$Q = Matriz\ de\ covarianza\ de\ la\ perturbación\ del\ proceso$
$R = Matriz\ de\ covarianza\ de\ la\ perturbación\ de\ la\ observación$
$K_k = ganancia\ de\ Kalman$
$z_k = medida\ tomada\ por\ el\ observador$

Del grupo de ecuaciones de la Tabla 3, la ecuación (5) representa el modelo de comportamiento del sistema que contiene las variables de estado a ser estimadas. La ecuación (6) representa la estimación de la matriz de covarianza del error. La ecuación (7) representa la denominada ganancia de Kalman "K" que busca minimizar la covarianza del error para la nueva estimación del estado mediante la expresión (9). Finalmente la Ecuación (8) representa la corrección del estado estimado $\hat{x}_k^-$en base a la ganancia de Kalman y a la innovación en la medida $(z_k - H\hat{x}_k^-)$, dando como resultado un estado corregido $\hat{x}_k$.

La ecuación (10) representa la medición de las variables de estado por parte de un observador externo al sistema representado por la ecuación (5), siendo v_k el ruido presente en la medición para cada estado y H la variable que define la observabilidad de los estados.

$$z_k = H\hat{x}_k^- + v_k \quad (10)$$

Los vectores w y v son independientes el uno del otro y representan el ruido gaussiano blanco con media cero presente en el proceso y en las observaciones respectivamente. Además estos vectores$w$ y $v$traen consigo asociadas las matrices de covarianza Q y Rrespectivamente que en general son diagonales, pudiendo también no serlo.

$$p(w) \sim N(0, Q)$$
$$p(v) \sim N(0, R)$$

En la práctica la matriz de covarianza de la perturbación del proceso Q y la matriz de covarianza de la perturbación de la observaciónR, pueden variar en el tiempo pero en general pueden ser consideradas como constantes [20], [21].

En el caso de la matriz de la covarianza, la estimación inicial con la cual dará inicio el filtro de Kalman puede ser la matriz identidad, mientras que en el caso de la estimación inicial para el estado, podría asignarse cualquier valor a conveniencia del proceso en donde se esté aplicando el filtro.

En el presente trabajo se ha asumido que el modelo de sistema en estudio es de carácter lineal y que está sometido únicamente a perturbaciones estáticas (vibraciones) más no a aceleraciones de tipo lineal.En el caso que se requiera tratar un modelo de tipo no lineal es necesaria la utilización del Filtro Extendido de Kalman(EKF) como se describe en [20] y una aplicación prácticase puede encontrar en [22].

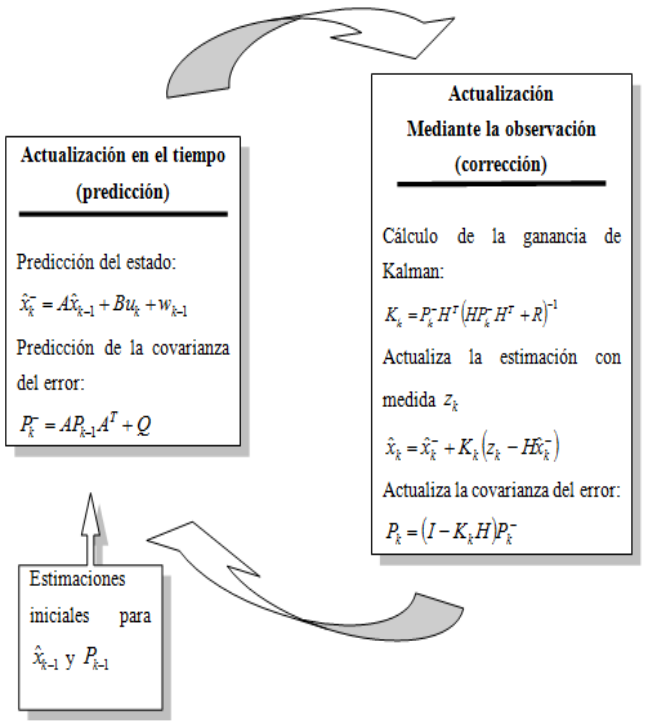


Figura 6. Operación completa del Filtro de Kalman

Con los antecedentes mencionados sobre la medición de ángulos en cabeceo y alabeo en base a la utilización de sensores inerciales, se puede apreciar que tanto el giroscopio como el acelerómetro por sí solos presentan dificultades en el momento de obtener una medida angular.

Tómese como ecuación que describe la evolución del sistema (medida angular) a la expresión (2) más la presencia de deriva en el giroscopio a la que se denominará w_k. De igual manera sea la expresión (10) aquella que representa observaciones sobre el sistema en instante definidos de tiempo (medidas angulares). Si el ángulo θ representa la variable de estado que se desea corregir entonces se tiene que:

$$\theta_k^- = A * \theta_{k-1} + B * \omega_k * dt + w_{k-1} * dt \quad (11)$$

$$z_k = H * \theta_k^- + v_k \qquad (12)$$

donde:
A=1
B=1
$dt = periodo\ de\ muestreo\ [s]$
$\omega_k = u_k = rapidez\ angular\ entregada\ por\ el\ giroscopio\ [°/s\ o\ rads/s]$
$\theta_k^- = ángulo\ estimado\ [°\ o\ rads]$
$z_k = media\ angular\ dada\ por\ el\ acelerómetro\ [°\ o\ rads]$
$w_{k-1} = ZRO\ (zero\ rate\ output\ )\ [°/s\ o\ rads/s]$

Hay que tomar en cuenta que las expresiones (11) y (12) representan las ecuaciones que definen al sistema y al observador respectivamente, expresadas en términos de una sola variable de estado θ,por tanto, debido a que en el presente documento se desea fusionar las medidas en cabeceo y alabeo entre el giroscopio (modelo sistema) y el acelerómetro (observador) es necesaria la inclusión de un modelo que contenga dos variables de estado quedando matricialmente las expresiones de la siguiente manera:

$$\begin{pmatrix}\theta_k^- \\ \varphi_k^-\end{pmatrix} = \begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix}\begin{pmatrix}\theta_{k-1} \\ \varphi_{k-1}\end{pmatrix} + \begin{pmatrix}dt & 0 \\ 0 & dt\end{pmatrix}\omega_k + W_{k-1} \qquad (13)$$

$$\begin{pmatrix}z1_k \\ z2_k\end{pmatrix} = \begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix}\begin{pmatrix}\theta_k^- \\ \varphi_k^-\end{pmatrix} + V_k \qquad (14)$$

Donde:

$$A = \begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix} \qquad B = \begin{pmatrix}dt & 0 \\ 0 & dt\end{pmatrix}$$

$$W_{k-1} = \begin{pmatrix}ZRO * dt * rand(-1,1) \\ ZRO * dt * rand(-1,1)\end{pmatrix}$$

$$H = \begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix} V_k = \begin{pmatrix}v_k \\ v_k\end{pmatrix}$$

$$\hat{x}_k^- = \begin{pmatrix} \theta_k^- \\ \varphi_k^- \end{pmatrix} \hat{x}_{k-1} = \begin{pmatrix} \theta_{k-1} \\ \varphi_{k-1} \end{pmatrix} z_k = \begin{pmatrix} z1_k \\ z2_k \end{pmatrix}$$

Además se tiene que:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad R = \begin{pmatrix} {v_k}^2 & 0 \\ 0 & {v_k}^2 \end{pmatrix}$$

$$Q = \begin{pmatrix} (ZRO * dt)^2 & 0 \\ 0 & (ZRO * dt)^2 \end{pmatrix}$$

Se logra conseguir de esta manera todos los parámetros para realizar la implementación de Filtro de Kalman basado en las expresiones de la Tabla 3.

Siendo R y Q funciones del ruido en la salida en el acelerómetro y ZRO en el giroscopio respectivamente y manteniendo su valor constante a lo largo del tiempo.

Nótese además que, en el caso del sistema propuesto las matrices de covarianzas R y Q son diagonales debido a que se asume que no existe correlación alguna entre las mediciones realizadas para los ángulos de alabeo y cabeceo. Si se conociera que una de las variables tiene influencia sobre la otra, entonces las matrices dejarían de ser diagonales y los elementos con "0" pasarían a tener valores dependientes de la covarianza entre las variables de estado.Q es una matriz que puede ser encontrada en base a modelamiento del comportamiento del sistema en estudio, normalmente es muy complicada de conseguir y muchas veces es colocada entre rangos ya establecidos previamente. Un criterio válido para colocar valores en la matriz Q sería en base a la suposición de que el sistema posee medidas muy fiables, caso en el cual los valores de Q serán pequeños caso contrario si se conoce que el proceso es poco fiable, entonces Q contendrá valores altos.

El término w_{k-1} en muchas ocasiones es despreciado en el momento de la implementación práctica ya que al representar el ruido en el proceso o sistema modelado es implícito a su comportamiento. Por otra parte, si se desea realizar una simulación del sistema modelado, el término w_{k-1} resulta de vital importancia ya que permite al algoritmo conocer que perturbaciones pueden presentarse en el sistema en la etapa de predicción y así poder compensarlos en la etapa de corrección. Sin embargo, ya sea de manera práctica o por simulación el término Q siempre debe estar presente en el algoritmo FK, ya que interviene de manera directa en la estimación de la covarianza del error.

La Covarianza del errorestimadaproviene del error que puede existir entre la variable estimada y su posterior corrección, siendo esto un factor preponderante en el cálculo de la ganancia K que sirve para corregir la variable estimada. Además, nótese que esta covarianza de error es nuevamente recalculada mediante la expresión (9), esta actualización o corrección de P_k^- para conseguir una nueva P_k implica que a lo largo de tiempo debe existir una tendencia a minimizar este parámetro cuyo caso ideal sería 0, pero que normalmente y en un proceso donde todas las variables han sido correctamente modeladas puede llegar a valores muy pequeños. En muchos casos puede presentarse que P_k, además de llegar a tomar valores cercanos a cero, puede estabilizarse en un valor fijo después de varias iteraciones del algoritmo, es entonces cuando se puede decir que el Filtro a convergido a un valor estable de P_k, si esto sucede el proceso de predicción y corrección de P_k puede omitirse y las ecuaciones (6) y (9) pueden ser reemplazadas por una única constante P_k.El número de iteraciones necesarias para que se estabilice el Filtro depende en gran medida de R mientras más grande sea R mayor número de iteraciones se necesitarán para que P_k logre alcanzar un valor estable.

Por otra parte refiriéndose al término v_k, éste desaparece en el algoritmo presentado en la Figura 6, pero su acción queda representada por la matriz R que en la expresión (7) implica una variación en el comportamiento de K que no es más que un peso o ganancia que se da a la relación entre lo medido por el observador y lo estimado en la etapa de predicción como se puede apreciar en la expresión (8).

Si bien R es un término que puede ser obtenido mediante un proceso de pruebas en base a medidas del observador, también es factible que éste pueda ser variado de manera heurística en base a las necesidades del comportamiento del FK, tomando como referencia que mientras más alto sea el valor colocado en R menor será el peso que se da al término de innovación en la ecuación (8), viéndolo como algo práctico podría decirse que si R es alta entonces implicaría que el observador es muy poco confiable en sus medidas y por tanto se debe dar el mínimo peso posible al término de innovación, llegando al límite donde K=0, R=∞ y el factor de innovación sea ignorado completamente, caso en el cual según la expresión (8) se presente que $\hat{x}_k = \hat{x}_k^-$.

3. EXPERIMENTOS Y RESULTADOS

Para realizar un estudio del comportamiento del Filtro para la fusión sensorial, la fase de experimentación se ha divido en dos secciones; la primera se remite al análisis del comportamiento del FK mediante la simulación del modelo matemático representado por las expresiones (13),(14) y las ecuaciones de la Tabla 3, la segunda corresponde a la implementación del FK para la fusión sensorial entre el acelerómetro ADXL-203 y el giroscopio IDG-300 con el fin de obtener medidas angulares en cabeceo y alabeo en un sistema en reposo (sin aceleraciones debido a desplazamiento) pero que puede ser afectado Perturbaciones leves (Vibraciones).

3.1 Simulación del modelo

La simulación se ha realizado en el Software MATLAB y se ha tomado como información base para el FK las características técnicas de los sensores IDG-300 y ADXL-203resumidasen la Tabla4.

Tabla 4. Características principales de los sensores ADXL-203 e IDG-300

Acelerómetro ADXL-203	
Ancho de banda [23]	50 [Hz]
Ruido (rms) [24]	1 [mg]
Sensibilidad [24]	1000 [mV/g]
Giroscopio IDG-300	
Densidad de ruido[25]	0.014 $[°/s/\sqrt{Hz}]$
Ancho de banda externo [26]	2 [KHz]
ZRO [25]	$\pm$ 100 $[mV]$
Sensibilidad [25]	2 $[mV/°/s]$

En el caso del acelerómetro dados los datos de la Tabla 4. y tomando en cuenta las expresiones (3) y (4) es posible obtener el ruido en la medida angular del sensor en alrededor de los 0,057 [°]. En el caso de Giroscopio se calcula la desviación como una medida de velocidad angular en base al ZRO y la sensibilidad dando aproximadamente un valor de 50 $[°/s]$, por otra parte si se toma en cuenta la densidad de ruido mostrada en la Tabla 4, seobserva que existe una presencia de ruido en la medida de velocidad angular de alrededor de 0,626 $[°/s]$. Comparando los valores del ZRO y la densidad de ruido dados por el fabricante, se aprecia que el ZRO influye de mayor manera en el comportamiento de la salida del sensor, por esta razón, para la implementación del modelo se ha tomado en cuentaúnicamente éste valor.Sin embargo este parámetro es modificable conforme se explicó en el apartado 2.

En las Figuras 7,8,9,10 y 11 se muestran simulaciones del comportamiento del Filtro en una posición de reposo 0 [°] para los sensores inerciales y un horizonte de tiempo de 5 [s];el acelerómetro está sometido a vibraciones de amplitud $\pm$87mgy con R distinto para cada caso. En azul se tiene la medida angular en base únicamente al modelo del giroscopio, en verde la medida angular perteneciente al acelerómetro sujeto a vibraciones y en rojo la medida angular obtenida mediante la fusión sensorial del FK.

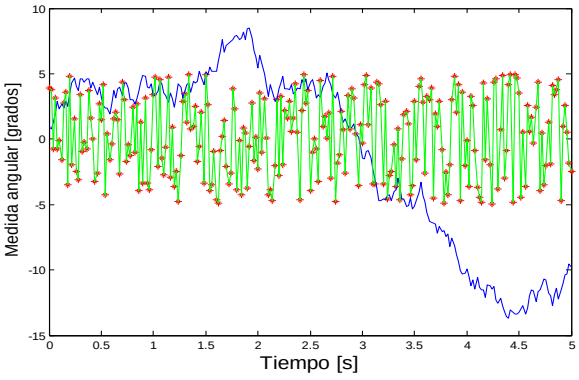


Figura 7. Simulación para $R = 0{,}0032$ y $Q = 50\text{^}2$

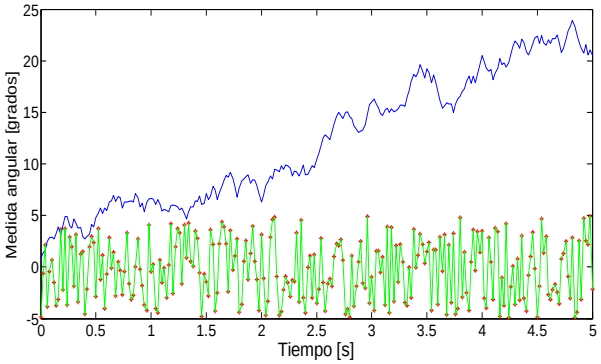


Figura 8. Simulación para $R = 0{,}01$y $Q = 50\text{^}2$

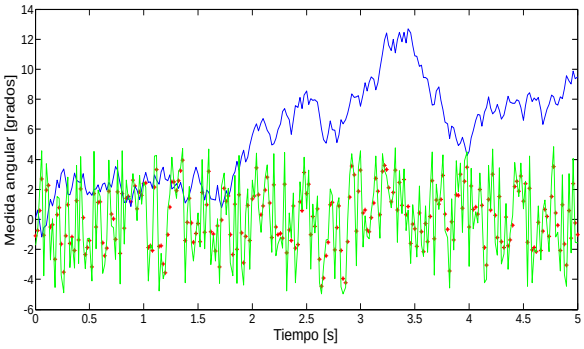


Figura 9. Simulación para $R = 1$ y $Q = 50\text{^}2$

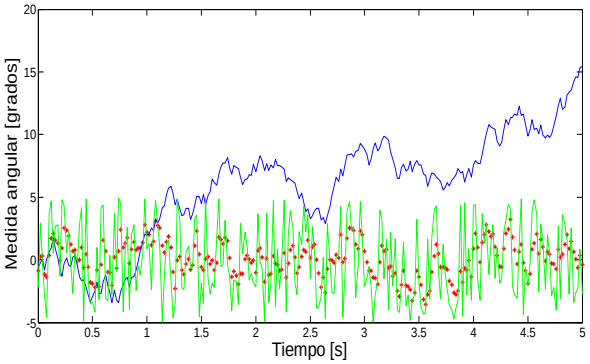


Figura 10. Simulación para $R = 100$ y $Q = 50\text{^}2$

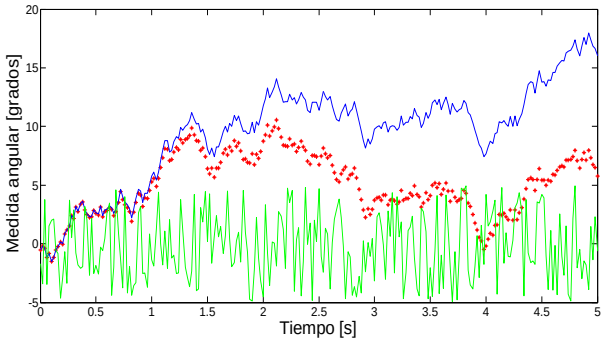


Figura 11. Simulación para $R = 900$ y $Q = 50\text{^}2$

La Fig. 7 representa el caso en el cual $R = v_k^2 = 0{,}057^2$ (dato del fabricante), se aprecia que el acelerómetro al ser sometido a vibraciones provoca que el ángulo conseguido mediante el FK se vea afectado , obteniendo como salida del Filtro una medida angular casi idéntica a la medida del acelerómetro (observador). Este comportamiento está plenamente justificado, pues como se explicó en el apartado 2 mientras más pequeño sea el valor de R más credibilidad se da a la medida del observador. En el caso de la Fig. 11, se ha fijado un valor de R alto equivalente a asumir que el acelerómetro tiene un ruido asociado de almenos$\pm$500[mg], por tanto la medida angular entregada por el FK, tenderá a hacia la medida dada por el modelo del giroscopio y tendrá muy poca influencia por parte del acelerómetro. En la Fig. 10, se puede apreciar que el valor R propuesto logra que la señal entregada por el filtro no sea influenciada de mayor manera por vibraciones y que tampoco se vea desviada de su valor debido a la influencia del giroscopio.

En las Figuras 12, 13 y 14 se analiza el efecto del parámetro Q en el comportamiento del filtro, para ello se tomará como base el valor de R=100.

Claramente se puede apreciar que Q influye de manera drástica en el comportamiento del modelo del giroscopio representado por la expresión (11), además de que el Filtro responde de mejor manera cuando la deriva del giroscopio ZRO es pequeña. Como se mencionó en el apartado 2 el valor Q muchas ocasiones es complicado de obtener con certeza, razón por la cual normalmente se lo asigna de manera heurística, basados en un conocimiento básico del proceso o sistema a modelar.

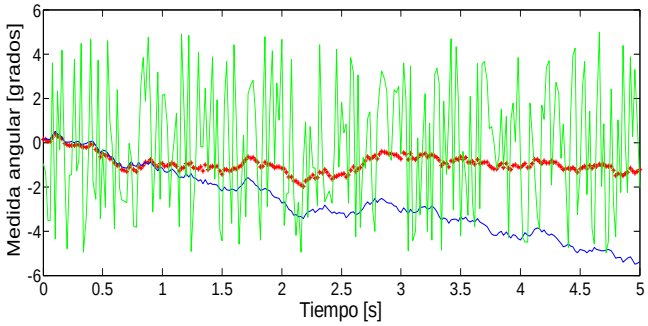


Figura 12. Simulación para $R = 100$ y $Q = 10\text{^}2$

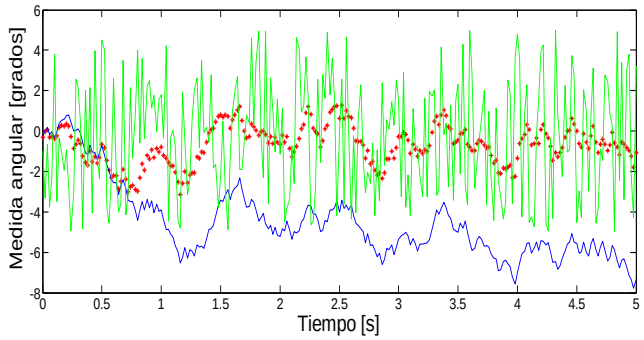


Figura 13. Simulación para $R = 100$ y $Q = 30\text{^}2$

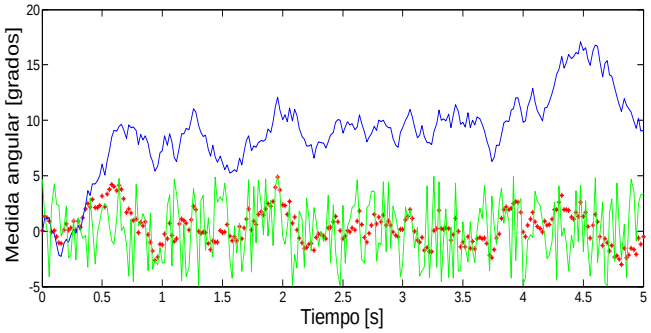


Figura 14. Simulación para $R = 100$ y $Q = 70\text{^}2$

3.2 Implementación del sistema de medición angular.

Con el fin de comprobar la eficacia del algoritmo presentado, se ha implementado un sistema basado en el acelerómetro ADXL-203 y el Giroscopio IDG-300, los cuales mediante un Filtro de Kalman procesado en un DSPIC30F4011, permiten obtener una medida angular tanto en cabeceo como alabeo.

Figura 15. Sistema Implementado.

El sistema implementado es de carácter didáctico y posee un sistema de visualización y manipulación de la variable R para observar el comportamiento de la medida angular cuando es sometida a pequeñas perturbaciones (vibraciones), además mediante un menú es posible observar independientemente los ángulos medidos, ya sean sólo del acelerómetro, sólo del giroscopio o los ángulos filtrados mediante fusión sensorial.

Con el fin de realizar las pruebas de medición de ángulo mostradas en lasFiguras 16, 17, 18 y 19 se toma como valores referenciales los mencionados en el apartado 3.1 con R=100 y Q=50^2, siendo éstos manipulables. Además, se dispone de un instrumento Patrón (graduador) de medición angular marca CEMCO que sirve para establecer una comparación con el ángulo medido por el prototipo desarrollado.

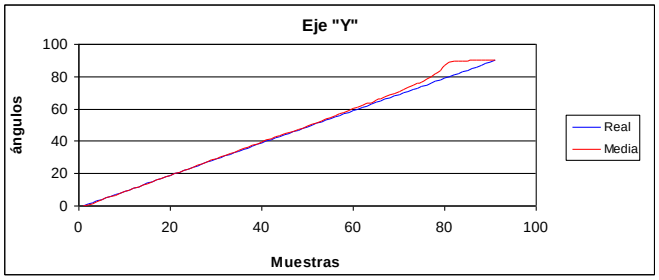


Figura 16. Comparación de la media del ángulo positivo en cabeceo medido por el prototipo y el valor real tomado del instrumento patrón.

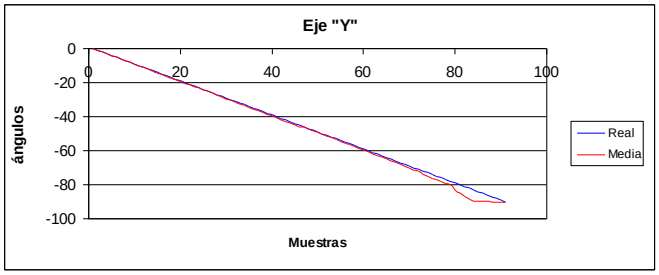


Figura 17. Comparación de la media del ángulo negativo en cabeceo medido por el prototipo y el valor real tomado del instrumento patrón.

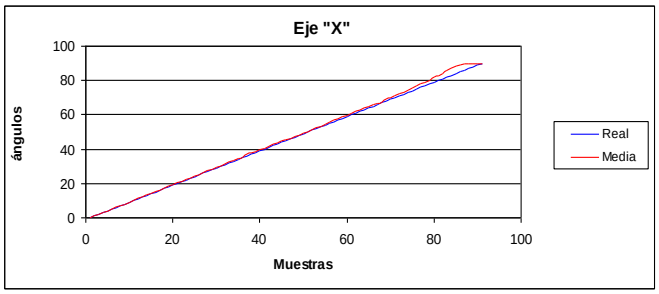


Figura 18. Comparación de la media del ángulo positivo en alabeo medido por el prototipo y el valor real tomado del instrumento patrón.

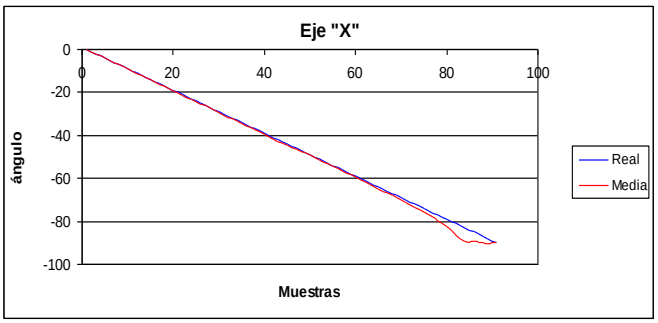


Figura 19. Comparación de la media del ángulo negativo en alabeo medido por el prototipo y el valor real tomado del instrumento patrón.

Además, en las Tablas 5 y 6 se muestra un resumen de los errores relativos existentes entre la medida dada por el instrumento patrón y la medida del prototipo.

El sistema implementado utiliza una fuente de 9V DC (batería), con una corriente de consumo de 108mA promedio. Experimentalmente se comprueba que el tiempo de funcionamiento óptimo del dispositivo con una batería alcalina es de aproximadamente 1 hora y 15 minutos, después de lo cual las prestaciones del dispositivo comienzan a disminuir presentándose problemas especialmente con el rango de medición y la exactitud del ángulo que se está midiendo,esto debido a que la estabilidad de las señales de los sensores dependen en gran medida del voltaje de alimentación. Para comprobar que el sistema necesita un cambio de batería simplemente se debe colocar el mando en una posición de 90° o -90° ya sea en cabeceo o alabeo y si la medición tiene un error relativo porcentual de un 6% o superior, entonces un cambio de batería es necesario.

Tabla 5. Características principales de los sensores ADXL-203 e IDG-300

Rango ángulos en cabeceo	Error relativo en porcentaje (%)
De 1° a 50°	1,65
De 51° a 70°	1,86
De 71° a 80°	5,99
De 81° a 90°	5,36
De -1° a -50°	1,44
De -51° a -70°	1,29
De -71° a -80°	3,29
De -81° a -90°	4,66

Tabla 6. Características principales de los sensores ADXL-203 e IDG-300

Rango ángulos en alabeo	Error relativo en porcentaje (%)
De 1° a 50°	1,58
De 51° a 70°	1,27
De 71° a 80°	2,56
De 81° a 90°	3,37
De -1° a -50°	1,25
De -51° a -70°	1,08
De -71° a -80°	2,97
De -81° a -90°	4,54

4. CONCLUSIONES

La búsqueda de las variables a ser utilizadas en el Filtro de Kalman, requiere un amplio estudio matemático, especialmente en lo referente a los valores de Q y R, ya que el hecho de implementarlas por ensayo y error puede provocar que el conseguir una correcta calibración del FK se convierta una tarea difícil y demorosa.

El parámetro R es de gran importancia para estimar el posible comportamiento del FK y que a más de brindar una medida de credibilidad al término de innovación mediante la ganancia K, también es necesario mencionar que la

disminución del valor de R en el sistema provoca que éste se vuelve vulnerable a las más leves vibraciones, dando como resultado que el ángulo medido tenga variaciones súbitas en su valor, sin embargosiR disminuye también se presenta la ventaja de que el sistema responde de manera más rápida al cambio en el ángulo convergiendo casi instantáneamente. Por otra parte cuando la variable R tiende a un valor alto el sistema es menos susceptible a vibraciones (acelerómetro) que son las causantes de una variación brusca en el ángulo medido,pero puede provocarque en cambios bruscos en la medidaangularésta converja de manera más lenta hacia el ángulo deseado, generando retardos en las medidas.

Se logra dar una metodología del uso y aplicaciones de los sensores inerciales así como de las principales diferencias presentes entre estos.

La utilización de la señal del giroscopio presenta un grave problema debido a la deriva propia de este tipo de sensores y al no poderse eliminar por completo implica que un giroscopio para una aplicación de inclinación no puede usarse solo. Por otra parte el acelerómetro es una estupenda solución para aplicaciones en donde se necesite tener una medida de inclinación, siempre y cuando el lugar donde se lo instale no esté sometido a vibraciones significativas o a movimientos que provoquen aceleraciones a lo largo de los ejes del sensor.

Si bien en el presente proyecto se presenta una aplicación de los sensores inerciales junto con el Filtro de Kalman orientado hacia el cálculo de un ángulo de inclinación, sería de gran importancia profundizar en el estudio y desarrollo de INS (Sistemas de navegación Inercial) que en la actualidad tienen gran importancia y gran número de aplicaciones.

REFERENCIAS

[1] Curey, R.K.; Ash, M.E.; Thielman, L.O.; Barker, C.H., "Proposed IEEE inertial systems terminology standard and other inertial sensor standards," *Position Location and Navigation Symposium, 2004. PLANS 2004*, vol., no., pp.83,90, 26-29 April. 2004.

[2] M. Kraft, "Micromachined Inertial Sensors State of the Art and a Look into the Future", University of Southampton, Highfield, Southamptom, SO17 1BJ, 2000.

[3] R. Feliz, E. Zalama, J. Gómez, G. Bermejo, "Pedestrian tracking using inertial sensors", Journal of Physical Agents, Vol. 3, No. 1, January 2009.

[4] D. Roetenberg, H. Luinge, P. Slycke, "Xsens MVN: Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors", XSENS TECHNOLOGIES, April 3, 2013.

[5] D. Roetenberg, P. Slycke and Peter H. Veltink. "Ambulatory Position and Orientation Tracking Fusing Magnetic and Inertial Sensing", IEEE Trans. Biomedical Engineering, Vol. 54, No. 5, May. 2007.

[6] H. Zhou, H. Hu, "Inertial sensors for motion detection of human upper limbs", Emerald Sensor Review, 2007.

[7] H. Chao, C. Coopmans, L. Di and Y. Chen, " A comparative Evaluation of Low- Cost IMUs for Unmanned Autonomous Systems", IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, University of Utah, Sept. 2010.

[8] F. Li, C. Zhao, G. Ding, j. Gong, C. Liu, F. Zhao, " Areliable and Accurate Indoor Localization Method using Phone Inertial Sensors", Microsoft Research Asia, Beijing, China, 2012.

[9] M. Stuart, A. Friedland, C. Friendland, "A Survey of Unmanned Aerial Vehicle (UAV) Usage for Imagery Collection In Disaster Research And Management", Louisiana State university, 2011.

[10] N. Yazdi, F. Ayazi and K. Najafi, "Micromachined Inertial Sensors", IEEE, Procedings, Vol. 86, No. 8, 1998.

[11] Lucian T. Grigorie, Ruxandra M. Botez, "Modeling and Numerical Simulation of an Algorithm for the Inertial Sensors Errors Reduction and for the Increase of the Strapdown Navigator Redundancy degree in a Low Cost Architecture", Transactions of the Canadian Society for Mechanical Engineering, Vol. 34, No. 1, 2010.

[12] Honghui Qi; Moore, J.B., "Direct Kalman filtering approach for GPS/INS integration," *Aerospace and Electronic Systems, IEEE Transactions on* , vol.38, no.2, pp.687,693, Apr 2002

[13] Shaeffer, D.K., "MEMS inertial sensors: A tutorial overview," *Communications Magazine, IEEE* , vol.51, no.4, pp.100,109, April 2013.

[14] Barbour, N.; Schmidt, G., "Inertial sensor technology trends," *Sensors Journal, IEEE* , vol.1, no.4, pp.332,339, Dec 2001

[15] A. Manzanares del Moral, A. Luque, "Estudio de Modelos matemáticos de Acelerómetros Comerciales", Universidad de Sevilla, España, 2008.

[16] Silicon Designs, Inc. Advanced Accelerometer Solutions, "Glossary", 2011,Disponible en: http://www.silicondesigns.com/glossary.html

[17] MSY. SivaPrasad, PVR. Murthy, "Design Simulation &. Fabrication of Micromachined Acceleration Sensor", PhD. Dissertation, Dept. Mech. Eng., Jawaharlal Nehru Technological University.,Kukatpally, 2010.

[18] A. Candelas, A. Corrales, "Giroscopios en el Sistema GypsyGyro-18"., Publicación interna No. 4., Grupo de Automática , Robótica y Visión Artificial, Universidad de Alicante, España, 2007.

[19] Peter S. Maybeck., "Chapter 1: Stochastic models, estimation, and control", Volume 1. Department of electrical engineering air force institute of technology Wright-Patterson air force base, OHIO, 1979.

[20] G. Welch, G. Bishop., "An Introduction to the Kalman Filter", Department of computer Science, University of North Carolina, Chapel Hill, EEUU., July, 2006.

[21] A. Solera., "El Filtro de Kalman", Banco Central de Costa Rica, División Económica, Departamento de Investigaciones Económicas., Julio, 2003.

[22] Begoña F. Aguado, Janelcy A. Castaño, Eduardo Z. Casanova, Jaime G. García-Bermejo., "Diseño y Simulación de un Filtro Kalman para un Robot Móvil"., XXV Jornadas de Automática, Ciudad Real., Septiembre, 2004.

[23] Dual Axis Accelerometer Evaluation Board ADXL203EB, Analog Devices, Norwood, U.S.A., 2004.

[24] Precision $\pm 1.7\,g$ Single -/Dual – Axis iMEMS® Acelerometer ADXL103/ADXL203, Analog Devices, Norwood, U.S.A., 2006.

[25] Integrated Dual-Axis Gyro IDG-300, InvenSense Inc., Santa Clara, U.S.A., 2006.

[26] IDG – 300 Dual – Axis Gyroscope Evaluation Board Specification, InvenSense Inc., Santa Clara, U.S.A., 2007.

Strategy Based on Multiple Objectives and Null Space for the Formation of Mobile Robots and Dynamic Obstacle Avoidance

Leica P. *; Chavez D. **; Rosales A. **; Roberti F. †*; Toibero J. †*; Carelli R. †*
†Consejo Nacional de Investigación Científicas y Técnica
**Universidad Nacional de San Juan, Instituto de Automática, San Juan, Argentina*
e-mail: (pleica,froberti,mtoibero,rcarelli)@inaut.unsj.edu.ar
***Escuela Politécnica Nacional, Facultad de Ingeniería Electrónica y Electrónica, Quito, Ecuador*
e-mail: (danilo.chavez,andres.rosales)@epn.edu.ec

Resumen:*En este trabajo se presenta un nuevo algoritmo para el control de formación flexible de robots móviles basado en múltiples objetivos de control. La estrategia contempla el uso del espacio nulo de una matriz Jacobiana para el control de forma y postura. La estrategia de evasión de obstáculos está basada en la definición de energía potencial ficticia. Se establece como objetivo primario el control de forma y evasión de obstáculos, y como objetivo secundario el control de postura y seguimiento de trayectoria de la formación de los robots. Se analiza la estabilidad de los controladores implementados y se presentan los resultados obtenidos por simulación que muestran el correcto desempeño de los controladores.*

Palabras clave:*Formación de robots, evasión de obstáculos, función potencial, seguimiento de trayectoria, espacio nulo.*

Abstract:*In this paper, a new algorithm for controlling mobile robot flexible formation based on multiple control objectives is presented. The strategy includes the use of null space for shape and posture control. The obstacle avoidance strategy is based on the definition of fictitious potential energy. The primary objective established is to shape control and obstacle avoidance, whereas the secondary objective includes the posture control and trajectory tracking of the robot formation. Stability analysis of the proposed control system is proven. Simulation results show the performance of the proposed controllers.*

Keywords:*robot formation, obstacle avoidance, potential function, trajectory tracking, null space*.

1. INTRODUCTION

One of the current challenges in robotics is the cooperative control of multiple robots with a common goal. Thepossibility of making several smaller capacity robots to perform tasks that would be impossible or inefficient individually, has motivated the scientific community to encourage the development of innovative control strategies. The use of multiple robots versus one offers several advantages, which could include reduced costs, greater robustness, improved performance and efficiency [15]. Instead of designing a single specialized powerful robot, a multi-robot system can be simpler and less costly [12]. This feature has allowed formation control to be successfully used in various applications: military [18], civil [17] and service robotics. There are three basic structures in the bibliography. The control of multi-robot systems: leader-follower strategy, methods based on behavior and virtual structures, each with their respective advantages and disadvantages. In the leader-follower structure, an agent is considered the leader and the remaining agents are considered followers of the designated leader [7] and [8]. In this structure, only the follower has information about the leader, so if it fails there is no possible mechanism that would ensure the compliance of the control target. However, this structure is easy to understand and implement.In the structure based on behavior, group behavior is defined as a combination of individual behavior of its members [5]. The main problem with this approach, is the difficult mathematical formalization and therefore it is not easy to ensure the convergence of the formation to the desired setting. In virtual structures geometry maintains a rigid connection between the robots and the reference system, which can be a virtual point or a virtual agent. One advantage of this method is that the virtual leader would never fail, so training will be maintained during the execution of the task.

Formations are categorized as rigid or flexible [14]. Working with rigid formations is advantageously less complex in terms of representation and control. The main disadvantage is that it may suffer collisions and encounter mobility problems, especially in corners and narrow passages, where the formation is larger than the available space [6]. In [13] a control scheme based on virtual structure is called cluster space control. Position control (or trajectory tracking) is carried out considering the centroid of the geometrical structure (a triangle) corresponding to the formation of three robots.In [16] the problem inherent in centralized control systems is addressed. More specifically, a technique to

extrapolate intrinsic generalization capabilities not discussed in [13] is developed, allowing application of the control approach based on the centroid of the formation in formations with a number equal to or greater than three robots. It also analyzes the ability of the formation in obstacle avoidance, whereby it can modify its structure momentarily, allowing an elastic behavior.At present the implementation of tasks in which robots are used requires extensive data processing in real time, while meeting a variety of tasks (manipulation, exploration, obstacle avoidance, etc.). This means that one must achieve several control goals simultaneously, sometimes causing conflict between them and the assigned order of priority. In [1] a number of control schemes are discussed that decompose the control problem into several sub-problems that are eventually solved individually.

Among the options, the control based on null spaces is enunciated, where the primary and most important objective is considered a minimum norm solution obtained by the pseudo-inverse of the Jacobian associated with the problem whereas the secondary objectives are posed in the null space of the aforementioned Jacobian. The main advantage is that this control scheme guarantees the fulfillment of the primary or higher priority, while the lower-level objectives should be analyzed in each case, but are projected in a space (null space) where it does not conflict with the main objective [3]. This concept was introduced in [2] to control generic robotic systems and in [4] to control multi-robotic systems. By interacting in dynamic and non-structured environments, multi-robot systems need to preserve their integrity, thus necessitating tools for obstacle avoidance. In bibliography, several proposals were found to solve this problem. One of them is the use of potential fields, as proposed in [9] and [10]. In this approach the obstacles generate repulsive forces on the robot, while the target generates attractive forces. The sum of all forces produces a resulting force that determines the direction and speed of the movement. Reference [11] analyzes the main limitations, among which the most important is the existence of local minimums that trap the robot, thus making it unable to reach the target.Another major limitation is the complication of passing through small spaces between obstacles, since they can generate repulsive forces greater than the attractive forces of the target. This paper presents a control scheme for tracking the formation of mobile robots, based on the null space of Jacobian matrix and the implementation of fictitious potential fields for obstacle avoidance. It is considered as a zero potential region to the entire environment with no barriers and non-zero in those regions containing obstacles. Then two control objectives are posed: a primary objective is to maintain the shape in areas of zero potential (without obstacles), and a secondary objective is to control trajectory and posture training. When an obstacle is found, the fictitious potential is different from zero and the formation deforms to avoid collisions with obstacles (static and dynamic). For obstacle avoidance fictitious potential fields are used, by characteristically not presenting local minimums, thereby offering an advantage. In this work, the temporal variation of the potential field (which is not covered in the literature of potential fields) is contemplated, allowing introduction of the dynamic of moving obstacles.A new approach in the analysis of potential fields is presented, using the derivatives of the field in the trajectories of the system.Consequently, this work aims to solve the problem of controlling a formation of mobile robots in unstructured environments, using null spaces for multiple control objectives, incorporating the dynamics of the obstacles in a single controller.The main contribution of the paper is to present a methodology considering multiple- control objectives using the null space of a Jacobian matrix associated with the definition of an artificial potential and its temporal variation linked to the primary target.

The paper is organized as follows: Section 2 presents statement of the problem, the structure of the robot, the potential function and the robots formation implemented. In Section 3 the system is modeled. Section 4 develops control laws and their stability. Section 5 presents the simulation results. Finally, Section 6 concludes.

2. PROBLEM STATEMENT

2.1Control Problem

While considering that two control objectives exist: the first combines obstacle avoidance (*Task 1*) and control shape (*Task 2*), whereas the second control objective combines trajectory tracking and posture angle of the formation (*Task 3*). This paper focuses on solving the trajectory tracking of mobile robotic formation with obstacle avoidance, using the concept of multiple-control objectives within the null space of the system. It is proposed that the centroid of the robot formation moves through a desired trajectory in an environment with obstacles.Trajectory tracking is part of one of the control tasks, which will be detailed later. The obstacle avoidance here is based on a fictitious potential field $\phi_{t,x}$ which includes the positions of the robots in the formation and obstacles external to it. In the absence of obstacles $\phi_{t,x} = 0$ and robots of the formation can navigate fulfilling the control objectives for shape and posture. In the presence of obstacles $\phi_{t,x} \neq 0$, the formation deforms to avoid hitting obstacles. For potential field generation only fictitious position obstacles and robots are required. However, if the variation of the potential field time $\partial\phi_{t,x}/\partial t$ is considered, it includes the dynamic behavior of the obstacles.

2.2 Potential Function

One of the control objectives is obstacle avoidance, requiring definition of a fictitious potential field that can describe a region of repulsion over the static or mobile obstacles. Potential function $\phi_{t,x}$ must describe the size of different obstacles within the environment, for which the following function has been adopted:

$$\phi_{t,x} = \begin{cases} r_v\, e^{-\left(\frac{x(t)-x_o(t)}{l_v}\right)^2-\left(\frac{y(t)-y_o(t)}{p_v}\right)^2} + \varepsilon\left(x(t)+y(t)\right) - r_o \;;\; C_1 \\ 0 \;;\; C_2 \end{cases} \quad (1)$$

where $\varepsilon \ll r_o;\ 0 < r_o \ll r_v;\ C_1 = l_v < l_o,\ p_v < p_o;\ C_2 = l_v > l_o;\ p_v > p_o;\ r_v, l_v and p_v$ are parameters that describe the object size; $x_o(t)$ and $y_o(t)$ are the coordinates of the obstacle in the world, $x(t)$ and $y(t)$ are the coordinates in the world. Figure 1 shows the shape of the potential function for an obstacle.

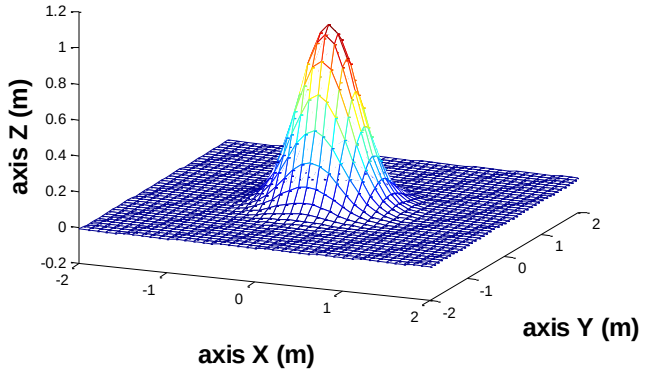


Figure 1. Shape of the potential function$\phi_{t,x}(l_o = p_o = 0.35;\ r_v = 1)$

2.3Mobile Robot

This paper uses unicycle-like mobile robots (see Figure 2) whose kinematic model is given as follows by:

$$\begin{bmatrix}\dot{x}_i \\ \dot{y}_i \\ \dot{\psi}_i\end{bmatrix} = \begin{bmatrix} cos(\psi_i) & -a\ sin(\psi_i) \\ sin(\psi_i) & a\ cos(\psi_i) \\ 0 & 1 \end{bmatrix}\begin{bmatrix}u_i \\ \omega_i\end{bmatrix} \quad (2)$$

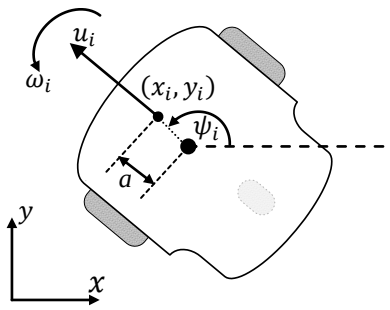


Figure 2. One-cycle type mobile robot

where a is the displacement of this point of interest (x_i, y_i) on the longitudinal axis of the i-th robot to the midpoint between the wheels; ψ_i is the orientation of the i-th robot; u_i and ω_i are the linear and angular velocities of the i-th robot respectively. For control purposes, the kinematics can be described in a compact form with (3) and (4) without the $\dot{\psi}_i$, , because of the non-holonomic characteristic of the mobile robot, the only way to navigate the zero position error path is when the robot has the same orientation as the path or trajectory. Where $\dot{X}_i = [\dot{x}_i \quad \dot{y}_i]$are the temporal variations of the i-th position robot; $U_i = [u_i \quad \omega_i]$are the i-th speeds.

$$\begin{bmatrix}\dot{x}_i \\ \dot{y}_i\end{bmatrix} = \begin{bmatrix} cos(\psi_i) & -a\ sin(\psi_i) \\ sin(\psi_i) & a\ cos(\psi_i) \end{bmatrix}\begin{bmatrix}u_i \\ \omega_i\end{bmatrix} = J_{ri}\begin{bmatrix}u_i \\ \omega_i\end{bmatrix} \quad (3)$$

$$\dot{X}_i = J_{ri}U_i \quad (4)$$

2.4Robot Formation

This paper proposes to work with the formation disposed according to Figure 3, where d_1 is the distance between robots R_1 and R_3;d_2 is the distance between robot R_1 and R_2, d_3 is the distance between the robot and R_2 and R_3, β is the angle opposite to the $d_3$segment; x_c and $y_c$are the positions of the centroid of the formation in reference to the world;θ is the formation posture angle. Thus the shape variables are defined by $q_f = [d_1 \quad d_2 \quad \beta]^T$and the posture variables by $q_p = [x_c \quad y_c \quad \theta]^T$, where $q = [q_f \quad q_p]$, being:

$$q_f = \begin{bmatrix}d_1 \\ d_2 \\ \beta_1\end{bmatrix} = \begin{bmatrix} \sqrt{(x_1-x_3)^2-(y_1-y_3)^2} \\ \sqrt{(x_1-x_2)^2-(y_1-y_2)^2} \\ acos\left(\dfrac{{d_1}^2+{d_2}^2-{d_3}^2}{2\,d_1 d_2}\right) \end{bmatrix} \quad (5)$$

$$q_p = \begin{bmatrix}x_c \\ y_c \\ \theta\end{bmatrix} = \begin{bmatrix} \dfrac{x_1+x_2+x_3}{3} \\ \dfrac{y_1+y_2+y_3}{3} \\ atan\left(\dfrac{\frac{2}{3}x_1-\frac{1}{3}(x_2+x_3)}{\frac{2}{3}y_1-\frac{1}{3}(y_2+y_3)}\right) \end{bmatrix} \quad (6)$$

where $R_1 = (x_1, y_1);\ R_2 = (x_2, y_2)$ and$R_3 = (x_3, y_3)$.

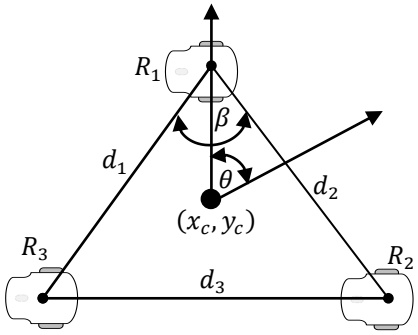


Figure 3.Robot formation diagram

This formation was selected because it can express the shape variables and posture variables separately.This allows to obtain multiple control objectives using the null space of the system. It should be clarified that, the formation can have more than three robots, but it is should consider changes in $q_f$and q_p.

3. SYSTEM MODELING

3.1Kinematic Modeling of the System

The structure of the control system is shown in Figure 4, where the generation of multiple tasks is performed by the

concept of multiple control targets using the null space of the system.$J_1$is the Jacobian that relates the tasks velocities with velocities of each robot.

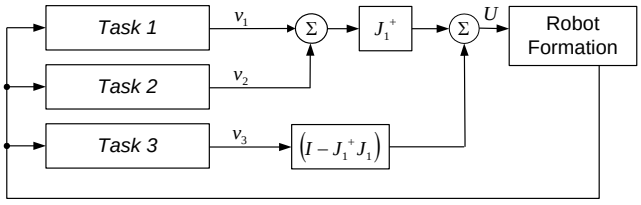


Figure 4. Control diagram structure

Task 1 is to avoid obstacle collision, *Task 2* is to keep the shape q_f of robot formation and *Task 3* is to follow a desired trajectory while maintaining the posture angle of the formation. The variables v_1, v_2 and v_3 represent the velocities generated by each task to meet the control objectives.Vector $U = [u_1\omega_1 u_2\omega_2 \ ... \ u_n\omega_n]$contains linear and angular velocities of the robots included in the formation and can be defined as:

$$U = J_1^{+}(v_1 + v_2) + (I - J_1^{+}J_1)v_3 \quad (7)$$

Equation 7 describes the overall control system structure in general, the same that will be placed depending on the variables such as shape q_f and posture q_p, which is detailed as follows.

Firstly,*Task 2* is analyzed and is defined as a function of posture variables q_f:

$$\dot{q}_f = J_f \dot{X} \quad (8)$$

where:

$$J_f = \begin{bmatrix} \frac{\partial d_1}{\partial x_1} & \frac{\partial d_1}{\partial y_1} & \frac{\partial d_1}{\partial x_2} & \frac{\partial d_1}{\partial y_2} & \cdots & \frac{\partial d_1}{\partial x_n} & \frac{\partial d_1}{\partial y_n} \\ \frac{\partial d_2}{\partial x_1} & \frac{\partial d_2}{\partial y_1} & \frac{\partial d_2}{\partial x_2} & \frac{\partial d_2}{\partial y_2} & \cdots & \frac{\partial d_2}{\partial x_n} & \frac{\partial d_2}{\partial y_n} \\ \frac{\partial \beta}{\partial x_1} & \frac{\partial \beta}{\partial y_1} & \frac{\partial \beta}{\partial x_2} & \frac{\partial \beta}{\partial y_2} & \cdots & \frac{\partial \beta}{\partial x_n} & \frac{\partial \beta}{\partial y_n} \end{bmatrix} \quad (9)$$

From linear algebra the minimum solution norm is defined by the right pseudo-inverse as$J_f^{+} = J_f^{T}\left(J_f J_f^{T}\right)^{-1}$, thus the solution in the row space of $J_f$is defined by the system inverse kinematics as:

$$\dot{X} = J_f^{+}\dot{q}_f \quad (10)$$

where$\dot{X} = [\dot{x}_1\dot{y}_1\dot{x}_2\dot{y}_2 \ ... \ \dot{x}_n\dot{y}_n]$.Replacing (10) according to structure (4) for *n* robots reveals the relationship between the shape variables and the velocities of the robots, defined by:

$$U = J_r^{-1}J_f^{+}\dot{q}_f \quad (11)$$

Now, to enter the *Task 3* in (11) it is projected position variables in the null space of J_f. This will allow the formation to retain its shape, and also to retain its posture along a desired trajectory. Thus (11) can be rewritten as:

$$U = J_r^{-1}\left(J_f^{+}\dot{q}_f + \left(I - J_f^{+}J_f\right)\dot{X}_p\right) \quad (12)$$

being:

$$\dot{X}_p = J_p^{+}\dot{q}_p \quad (13)$$

where$J_p^{+} = J_p^{T}\left(J_p J_p^{T}\right)^{-1}$; $\dot{X}_p$ are temporal variations of positions for the posture, and J_p is the Jacobian defined by:

$$J_p = \begin{bmatrix} \frac{\partial x_c}{\partial x_1} & \frac{\partial x_c}{\partial y_1} & \frac{\partial x_c}{\partial x_2} & \frac{\partial x_c}{\partial y_2} & \cdots & \frac{\partial x_c}{\partial x_n} & \frac{\partial x_c}{\partial y_n} \\ \frac{\partial y_c}{\partial x_1} & \frac{\partial y_c}{\partial y_1} & \frac{\partial y_c}{\partial x_2} & \frac{\partial y_c}{\partial y_2} & \cdots & \frac{\partial y_c}{\partial x_n} & \frac{\partial y_c}{\partial y_n} \\ \frac{\partial \theta}{\partial x_1} & \frac{\partial \theta}{\partial y_1} & \frac{\partial \theta}{\partial x_2} & \frac{\partial \theta}{\partial y_2} & \cdots & \frac{\partial \theta}{\partial x_n} & \frac{\partial \theta}{\partial y_n} \end{bmatrix} \quad (14)$$

The system design is completed including obstacle avoidance (*Task 1*) to which a term is added to the shape variables $\dot{q}_{ob}$, moreover allowing static and dynamic obstacle avoidance by robots. Thus (12) is rewritten as:

$$U = J_r^{-1}\left(J_f^{+}\left(\dot{q}_f + \dot{q}_{ob}\right) + \left(I - J_f^{+}J_f\right)\dot{X}_p\right) \quad (15)$$

In the followingthe $\dot{q}_{ob}$ is obtained.

3.2Obstacle Avoidance

The control problem is to design a controller so multiple mobile robots can maintain their shape while its centroid follows a desired trajectory in a dynamic environment (with fixed and moving obstacles) and suggests that only robots can move into positions where the fictitious potential field $\phi_{t,x}$ is less than or equal to zero. This requires finding a relationship that associates the fictitious potential field of each obstacle to the movement of each robot in the formation. This can be obtained by setting the variation of the potential field on the basis of temporal variations in the robot position. Deriving $\phi_{t,x}$ in the $\dot{X}_{ob}$ trajectories obtains:

$$\frac{d\phi_{t,x}}{dt} = \nabla\phi_{t,x}\dot{X}_{ob} + \frac{\partial\phi_{t,x}}{\partial t} \quad (16)$$

where $\dot{X}_{ob}$are temporal variations of the position of each robot to avoid collisions; $\nabla\phi_{t,x}$ is the partial derivative of $\phi_{t,x}$ with respect to the positions of each robot in the formation; $\partial\phi_{t,x}/\partial t$ is the potential field variation in time, this component provides dynamic information of the movement of the obstacles within the environment. Hence forth, $\nabla\phi_{t,x} = J_\phi$ is the Jacobian that relates the temporal variations of the field with temporal variations of the positions of each robot of the formation, and is expressed as follows:

$$J_\phi = \left[\frac{\partial\phi_{t,x}}{\partial x_1}, \frac{\partial\phi_{t,x}}{\partial y_1} \ldots, \frac{\partial\phi_{t,x}}{\partial x_n}\frac{\partial\phi_{t,x}}{\partial y_n}\right] \quad (17)$$

The kinematics of the system is defined by:

$$\dot{\phi}_{t,x} = J_\phi\dot{X}_{ob} + \frac{\partial\phi_{t,x}}{\partial t} \quad (18)$$

From linear algebra it is known that the minimum norm solution is defined by the right pseudo-inverse $J_\phi^{+} =$

$J_\phi{}^T\left(J_\phi J_\phi{}^T\right)^{-1}$, then the solution in the row space of J_v is defined by the inverse kinematics of the system as:

$$\dot{X}_{ob} = J_\phi{}^+\left(\dot{\phi}_{t,x} - \frac{\partial \phi_{t,x}}{\partial t}\right) \quad (19)$$

Now, from (14) $\dot{q}_{ob} = J_f\,\dot{X}_{OB}$ is obtained and replacing (19) gives:

$$\dot{q}_{ob} = J_f J_\phi{}^+\left(\dot{\phi}_{t,x} - \frac{\partial \phi_{t,x}}{\partial t}\right) \quad (20)$$

4. CONTROLLERS

4.1Proposed Controller

To meet the three tasks, the following controlleris suggested.

$$U_c = J_r{}^{-1}\left(J_f{}^+\left(\dot{q}_{fc} + \dot{q}_{obc}\right) + \left(I - J_f{}^+ J_f\right)\dot{X}_{pc}\right) \quad (21)$$

being:

$$\dot{q}_{fc} = \dot{q}_{fd} + K_f\ \tanh \tilde{q}_f \quad (22)$$

$$\dot{q}_{obc} = J_f J_\phi{}^+\left(\dot{\phi}_d + K_\phi \tanh \tilde{\phi}_{t,x} - \frac{\partial \phi_{t,x}}{\partial t}\right) \quad (23)$$

whereK_f and $K_\phi$are positive definite diagonal matrixes. The variables $\dot{q}_{fd}$ and $\dot{\phi}_d$ are the temporal variations of the shape and obstacle avoidance variables respectively. The shape errors are defined as $\tilde{q}_f = q_{fd} - q_f$, where q_{fd} is the desired value of the shape $q_{fd} = [d_{1d} \quad d_{2d} \quad \beta_d]^T$.

Fictitious potential error $\tilde{\phi}$ is generated by the presence of obstacles and is defined by $\tilde{\phi} = \phi_d - \phi$ where ϕ_d is the desired potential function, in this case $\phi_d = 0$ (obstacle free potential), which means that the robot can only move in the position $(x(t), y(t))$ that are obstacle-free. The position controller is defined by:

$$\dot{X}_{pc} = J_p{}^+\left(\dot{q}_{pd} + K_p\ \tanh \tilde{q}_p\right) \quad (24)$$

whereK_p is a diagonal matrix defined as positive, the error position is defined by $\tilde{q}_p = q_{pd} - q_p$; where $q_{pd} = [x_{cd} \quad y_{cd} \quad \theta_d]^T$; temporal variations desired are defined by $\dot{q}_{pd} = [\dot{x}_{dc} \quad \dot{y}_{dc} \quad \dot{\theta}_d]^T$, where$\dot{x}_{dc}$ and $\dot{y}_{dc}$ are trajectory reference speeds.

4.1Stability Analysis

This section analyzes the stability of the proposed controller. To clarify this analysis, the following definition is presented:

Definition*:* For any*m* by *n* matrix A, the null space and the row space are orthogonal sub-spaces of$\mathcal{R}^m$. Similarly the left null space and the column space are orthogonal sub-spaces of $\mathcal{R}^n$.

This means that $J_2 J_f{}^+ = 0$because $\mathcal{R}(J_2{}^+) \subseteq I - J_f{}^+ J_f = \mathcal{N}(J_f)$, J_2 being the projection matrix in the null space of J_f.

First, it is analyzed the secondary objective (*Task 3*) that does not affect the primary endpoint (*Task 1* and *Task 2*). To demonstrate this (4) in (8) for *n* robots:

$$\dot{q}_f = J_f\ J_r\ U \quad (25)$$

Assuming perfect velocitytracking$U_c \equiv U$, replacing (21) in (25):

$$\dot{q}_f = J_f\ \left(J_f{}^+\left(\dot{q}_{fc} + \dot{q}_{obc}\right) + \left(I - J_f{}^+ J_f\right)\dot{X}_{pc}\right) \quad (26)$$

Since it is known that $J_f{}^+ = J_f{}^T\left(J_f J_f{}^T\right)^{-1}$by replacing in (26) and developing results in:

$$\dot{q}_f = \left(\dot{q}_{fc} + \dot{q}_{obc}\right) \quad (27)$$

This implies that the secondary target does not affect the primary objective.

Now the stability of the shape errors is analyzed. Assuming that there are no obstacles $\dot{q}_{obc} = 0$, and assuming perfect velocity tracking, by replacement of (22) in (27) results in:

$$\dot{q}_f = \dot{q}_{fd} + K_f\ \tanh \tilde{q}_f \quad (28)$$

The shape errors are defined by:

$$\dot{\tilde{q}}_f + K_f\ \tanh \tilde{q}_f = 0 \quad (29)$$

where$\dot{\tilde{q}}_f = \dot{q}_{fd} - \dot{q}_f$; as$K_f$is a positive defined diagonal matrix $\tilde{q}_f \to 0$asymptotically in the absence of obstacles.

Now the stability of the position errors is analyzed. From (4) for *n*robots $\dot{X}_p = J_r\,U$, assuming perfect velocity tracking$(U \equiv U_c)$ thus:

$$\dot{X}_p = J_f{}^+\left(\dot{q}_{fc} + \dot{q}_{obc}\right) + \left(I - J_f{}^+ J_f\right)\dot{X}_{pc} \quad (30)$$

Replacing (13) in (30):

$$J_p{}^+\dot{q}_p = J_f{}^+\left(\dot{q}_{fc} + \dot{q}_{obc}\right) + \left(I - J_f{}^+ J_f\right)\dot{X}_{pc} \quad (31)$$

Multiplying both members of equality by $J_2 = I - J_f{}^+ J_f$and developing them results in (see Appendix A):

$$J_A\left(\dot{\tilde{q}}_p + K_p\ \tanh \tilde{q}_p\right) = 0 \quad (32)$$

where $J_A = J_2 J_p{}^+$, which moreover verifies that the null space $N(J_f) = \{\dot{X}/J_f\dot{X} = 0\}$has a dimension $dim\,N(J_f) = 3$, which is logical because three shape variables are used in the primary objective and three variables are left free for posture. Now as in the secondary objective (posture) the three remaining free variables are used, it is observed that the null space $N(J_A) = \{\dot{X}/J_A\dot{X} = 0\}$ has a $dim\,N(J_A) = 0$, this means that there are no free variables, therefore further object of control cannot be increased. In addition, its algebraic meaning signifies that the J_A matrix is a full range and the posture error system is defined by:

$$\dot{\tilde{q}}_p + K_p\ \tanh \tilde{q}_p = 0 \quad (33)$$

where$K_p$is a positive defined diagonal matrix, then$\tilde{q}_p \to 0$.

5. SIMULATION RESULTS

Simulations were performed using the Matlab computer program. The simulated robots incorporate a kinematic model of a non-holonomic robot (Pioneer 2 [19]). Two experiments were conducted: 1) It is proposed that the centroid of the robot formation moves through a desired trajectory in an environment with three obstacles (two staticobstacles and one dynamicobstacle, see Figure 5), maintaining the formation and avoiding obstacles as a primary objective; and 2) It is proposed that the centroid of the robot formation moves through a desired trajectory in an environment with threerobots (one fixed and two mobiles).The objective of this second experiment is to demonstrate control strategy behavior when the formation interacts with other robots having the same collision-avoidance strategy. All robots (six robots) react to avoid collisions between them, thus reducing control errors with respect to first experiment.

5.1 Experiment 1

The initial conditions of the robots formation are: $R_1 = (-1.5, 1)$; $R_2 = (-1,\ 1)$; $R_3 = (-1, -1.5)$; $\psi_1 = \psi_2 = -30^{\circ}$ and$\psi_3 = 0$ the initial positions of the objects are: $X_{ob_1} = (2.5,\ 2)$; $X_{ob_2} = (2, 0)$ and$X_{ob_3} = (10,\ 2)$, the speed of the dynamic object is 0.4 (m/s). The reference positions are: $\theta_d = 0^{\circ}$and$(x_{dc}, y_{dc}) = (x_{ref}, y_{ref})$, where x_{ref}, y_{ref} corresponds to the reference trajectory given by $y = A\sin(\omega x)$ with $A = 1$, $\omega = \pi/4$.Shape references are: $d_{d1} = d_{d1} = 1.2\ (m)$ and $\beta_d = 60^{\circ}$.

Figure5 shows the evolution of one experiment. Figures 6 and 7show the evolution of shape errors and verify that in the absence of obstacle control, errors tend to be zero. In$t = 9\ (s)$ to $t = 17\ (s)$ errors occur due to the presence of two static obstacles, and in $t = 34\ (s)$to$t = 37\ (s)$ the errors are due to the presence of the dynamic obstacle.

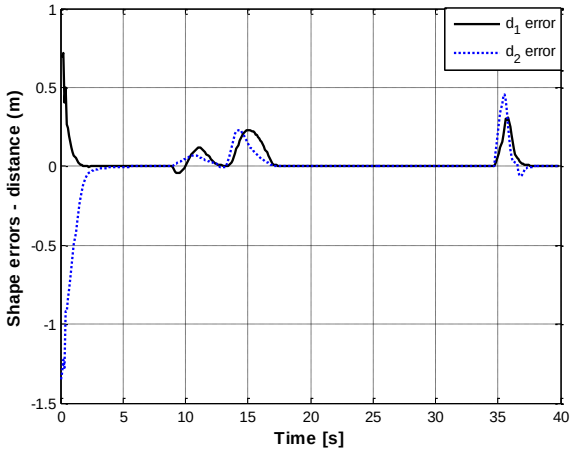


Figure 6. Evolution of the shape errors $\tilde{d}_1$and$\tilde{d}_2$

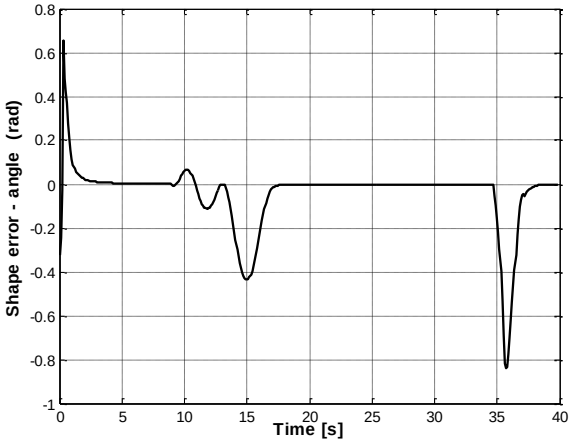


Figure 7. Evolution of the shape error $\tilde{\beta}$

Figure 8 shows the evolution of the posture error $\tilde{x}_c$ and$\tilde{y}_c$corresponding to the trajectory reference errors. Figure 9 presents the evolution of $\tilde{\theta}$. It is observed that at $t = 9\ (s)$ to $t = 17\ (s)$ and from $t = 34\ (s)$ to $t = 37\ (s)$errors appear in the presence of obstacles. In Figure10 the evolution of the potential field is shown for the three formation robots, which verifies that the potential field isn't zero in presence of obstacles.

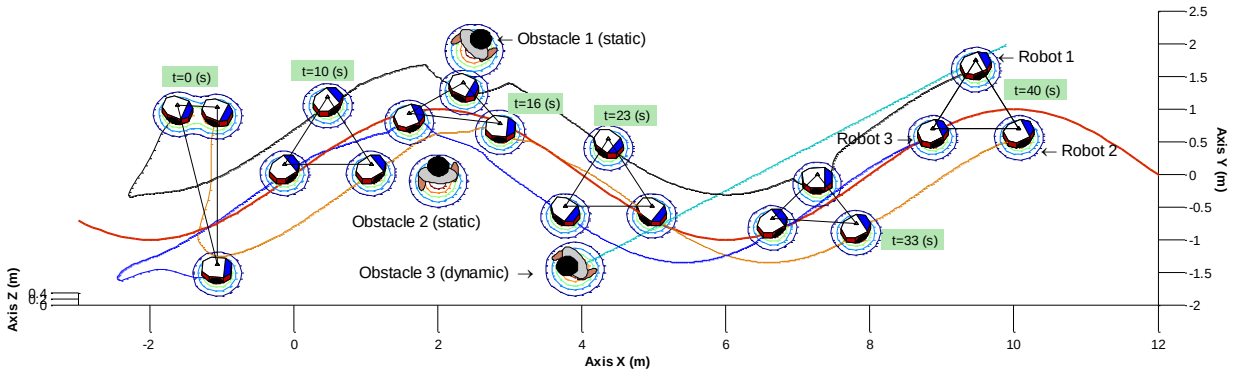


Figure 5.Positions and trajectories of obstacles and robots for experiment one

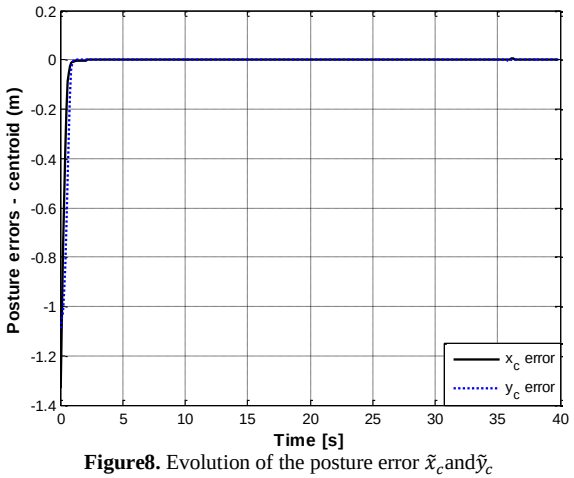


Figure8. Evolution of the posture error $\tilde{x}_c$and$\tilde{y}_c$

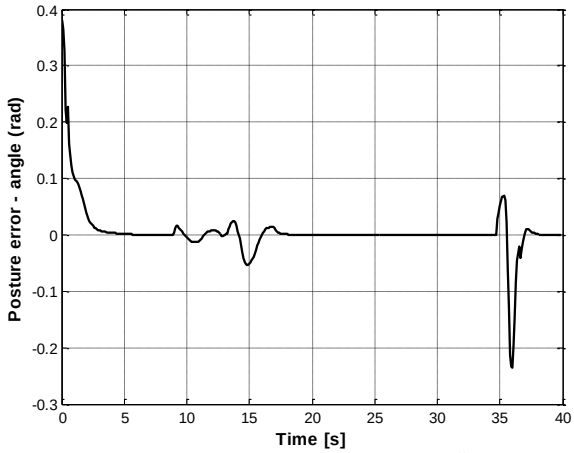


Figure9. Evolution of the posture error $\tilde{\theta}$

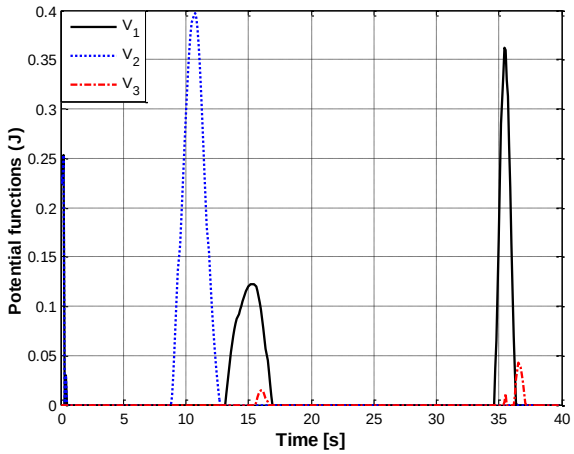
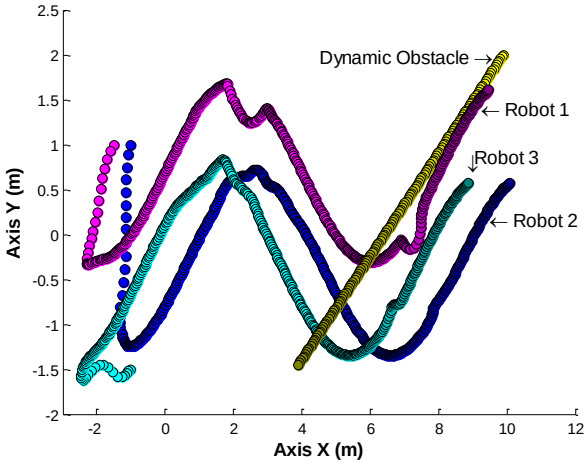
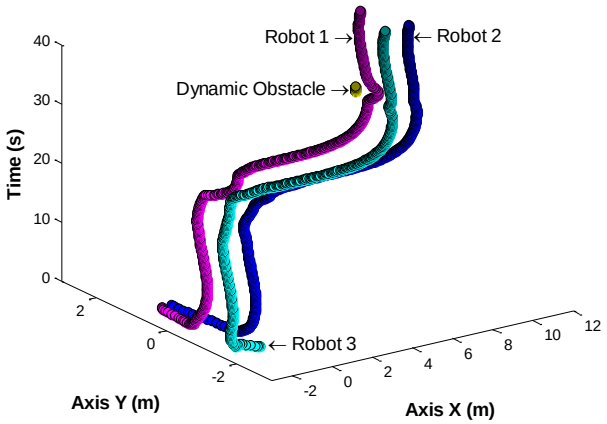


Figure 10. Evolution of each robots potentials functions

Figure 11 shows the evolution of time-parameterized trajectories of the robots and dynamic obstacle. The trajectories of the robots do not intersect with the obstacle, which implies that the formation of robots avoids collision with dynamic obstacle (obstacle 3). The collision avoidance for both static obstacles can be verified in Figure5. The R_1trajectory is of interest because it can collide with the dynamic obstacle.Figure 12 shows the difference in the trajectories for the two cases: when $\partial\phi_{t,x}/\partial t = 0$ and when $\partial\phi_{t,x}/\partial t \neq 0$.

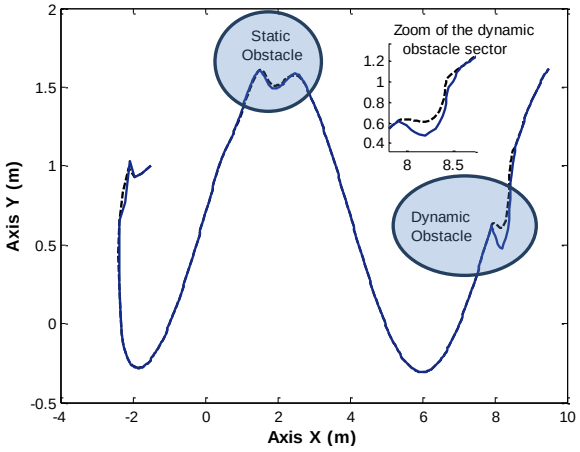


(a)

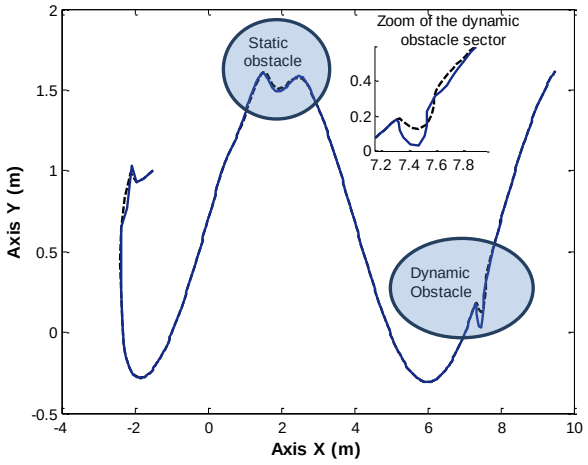


(b)

Figure 11. Time-Parameterized trajectories.a)Time-Parameterized trajectories [upper view]. b) Time-Parameterized trajectories [rotated view]



(a)



(b)

Figure 12. Trajectory-Evolution of the formation robots for R_1.12.a) low speeds.12.b) high velocities

Figure 12 shows the difference of the trajectory for R_1 when the variation of the potential field $\partial\phi_{t,x}/\partial t = 0$ (broken line) is not considered and when the variation of potential field $\partial\phi_{t,x}/\partial t \neq 0$ is considered (solid line).Figure 12(b) shows the trajectories of$R_1$when the dynamic obstacle speed has increased three times [3x]. In that event, a major difference in both cases is observable. When $\partial\phi_{t,x}/\partial t = 0$the robot R_1touches the dynamic obstacle, however when $\partial\phi_{t,x}/\partial t \neq 0$ the robot swerves to avoid the dynamic obstacle, which would implicate the importance of the variation of the potential field for dynamic obstacles.

5.2 Experiment 2

The difference is that in the first experiment only the formation of robots (the three robots) react to avoid collision, while in the second experiment all robots (six robots) react to avoid collisions between them, thus reducing control errors with respect to first experiment (see Figures 14, 15, 16 and 17). The form, posture, and trajectory parameters are the same as in the previous experiment. The initial conditions of the three new robots are $R_4 = (10,\ 2)$; $R_5 = (2.5,\ 2)$; $R_6 = (9, -1.5)$; $\psi_{r4} = -145^{\circ}$; $\psi_{r5} = 0^{\circ}$ and$\psi_{r6} = 180^{\circ}$; the speeds of R_4 andR_6 are $0.6\ (m/s)$ and$0.3\ (m/s)$ respectively.

Figure 13 reveals the evolution of the second experiment. Figures 14, 15, 16 and 17 show the form and posture errors, observing their increase when avoiding the collision with robots encountered navigating within the same environment.

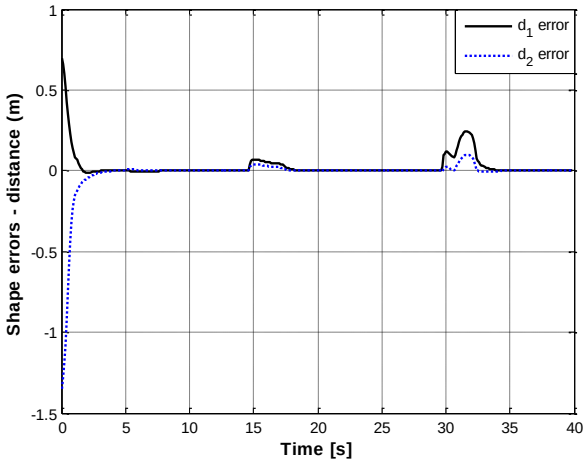


Figure 14.Evolution of shape error$\tilde{d}_1$and$\tilde{d}_2$

The control errors presented in $t = 14\ (s)$to$t = 17\ (s)$ are due to the possible collision with $R_4$and in $t = 29\ (s)$ to $t = 33\ (s)$are due to the possible collision with robots R_4 andR_6.

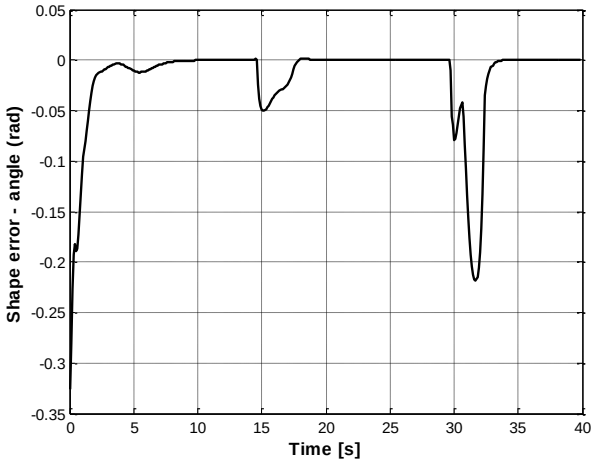


Figure 15.Evolution of shape error $\tilde{\beta}$

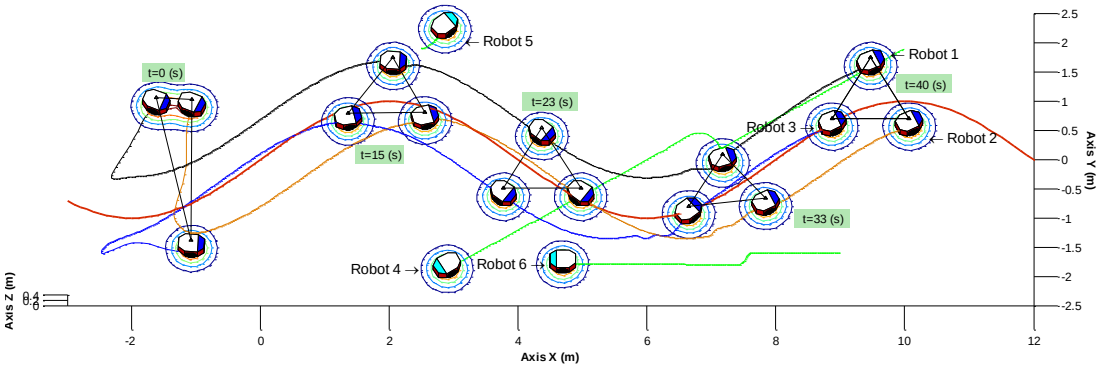


Figure 13.Positions and trajectories of obstacles and robots for experiment two

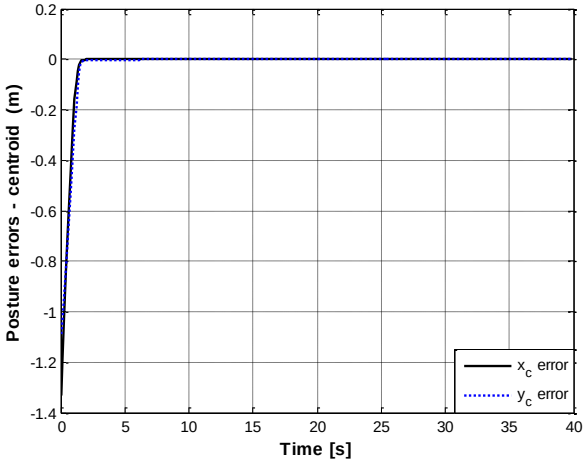


Figure 16.Evolution of posture errors$\tilde{x}_c$ and$\tilde{y}_c$

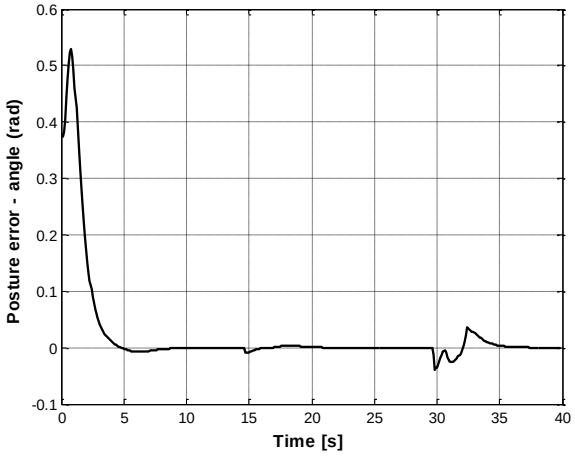
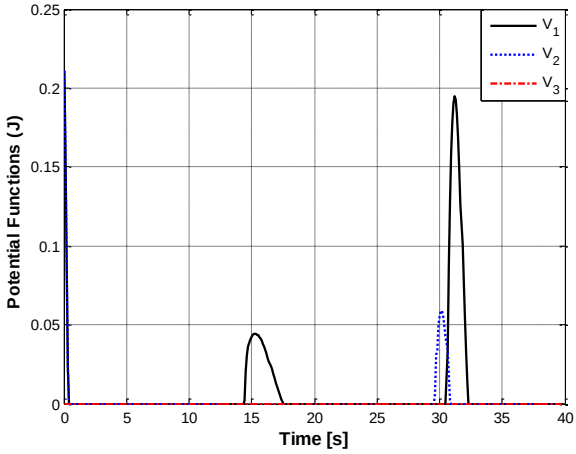
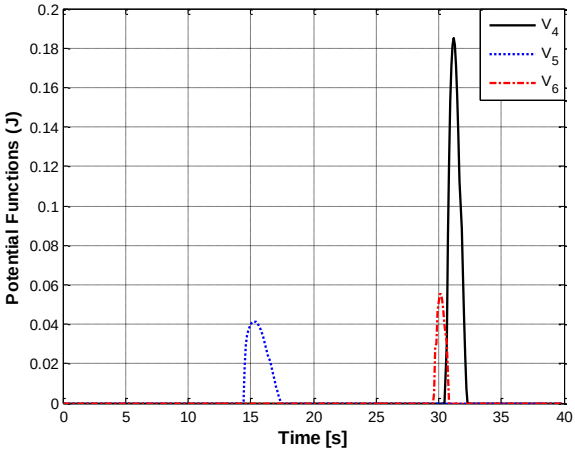


Figure 17.Evolution of posture errors$\tilde{\theta}$



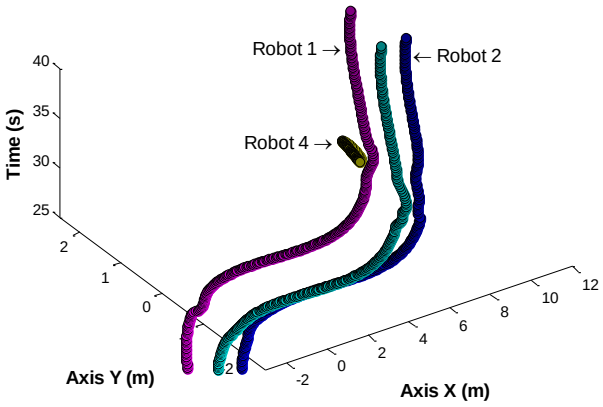
(a)



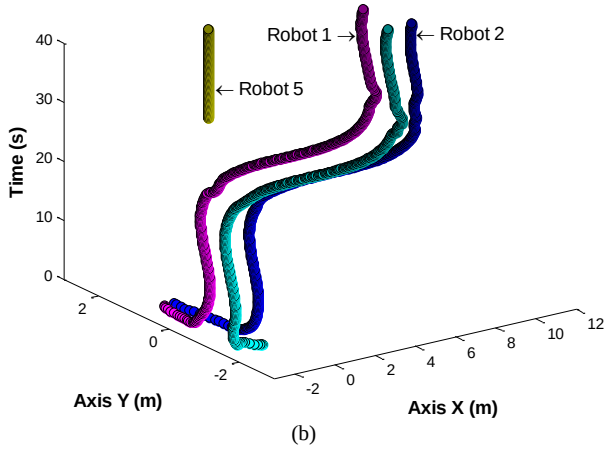
(b)

Figure 18.Evolution of the potentials functions for each robot. (a) Robots $R_1, R_2 and R_3$. (b) Robots $R_4, R_5 and R_6$.

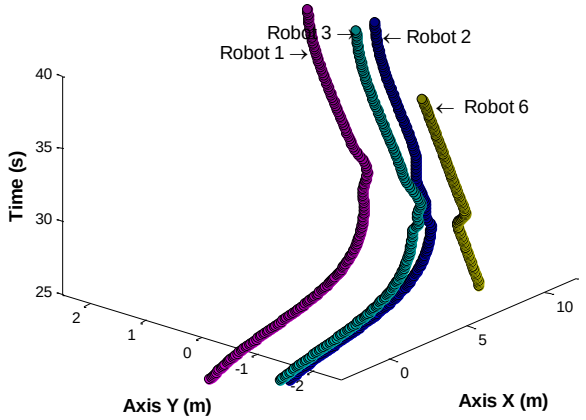
Figure 18 shows the evolution of the potential field for the six fictional robots. A reciprocity can be observed between potential fields. This allows observation of the interaction between robots in the event of possible collisions. Thus it can be seen that in $t = 14\ (s)$ to$t = 17\ (s)$ there is a possible collision between $R_1$and R_5.$t = 29\ (s)$to$t = 31\ (s)$there is a collision between R_2 and R_6. The same occurs in $t = 30\ (s)$to$t = 33\ (s)$where there is a collision between R_1andR_4, which corresponds to the observed trajectories in Fig 13.



(a)



(b)



(c)

Figure 19.Evolution of time-parameterized trajectories training for the three formation robots with the robots in the environment. (a) The three formation robots R_4 (b) The three formation robots with R_5 (c) The three formation robotswith R_6.

This verifies that the peak of the error control is lower in the second experiment than in first experiment, since R_4, R_5 and R_6robots contain obstacle avoidance algorithms and react to the presence of a possible collision with a formation. Figure 19 shows the evolution of the time-parameterized trajectories of the formation robots and robots in the environment. It can be seen that the formation robots avoid a collision with the environment robots. It leads to conclude that there are no collisions between the environment robots and the formation robots, or between the formation robots themselves. The required sensors for the implementation of the real experiments would be a laser sensor mounted on each robot to detect the position of the obstacles, and the positions of the robots are obtained using odometry.

6. **CONCLUSIONS**

This paper has developed a controller capable of working with multiple-control objectives using the definition of null space. The control objectives are such that the robot formation meets the objectives of shape and posture. (trajectory tracking, position and angle) and the avoidance of static and dynamic obstacles. Potential fields and temporal variations were used to model the dynamics of the obstacles to avoid collisions with the formation. The main contribution of the paper is to present a methodology considering multiple- control objectives using the null space of a Jacobian matrix associated with the definition of an artificial potential and its temporal variation linked to the primary target. The method does not present problems with the existence of local minimums and considers the dynamics of obstacles in motion. Another contribution is the use of the same controller for semi-structured environments with multiple robots having different control tasks. The simulations proved the efficacy of the proposed algorithms. As future work, experimental testing with real robots is expected to contribute to the validation of the proposed controllers.

REFERENCES

[1] G. Antonelli, F. Arrichiello, and S. Chiaverini The null-space-based behavioral control for mobile robots. In IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA),pages 15–20, 2005.

[2] G. Antonelli, F. Arrichiello, and S. Chiaverini. The null-space-based behavioral control for autonomous robotic systems. Intelligent Service Robotics, 1(1):27–39, 2008.

[3] G. Antonelli, F. Arrichiello, and S. Chiaverini. Experiments of formation control with multirobot systems using the null-space-based behavioral control. IEEE Transactions on Control Systems Technology, 17(5):1173–1182, 2009.

[4] G. Antonelli and S. Chiaverini. Kinematic control of platoons of autonomous vehicles. IEEE Transactions on Robotics, 22(6):1285–1292, 2006.

[5] T. Balch and R.C. Arkin. Behavior-based formation control for multirobot teams. IEEE Transactions on Robotics and Automation, 14(6):926–939, 1998.

[6] T.D. Barfoot and C.M. Clark. Motion planning for formation of mobile robots. Robotics and Autonomous Systems, 46(2):65˝U78, 2004.

[7] J. Chen, D. Sun, J. Yang, and H. Chen. Leader˝ Ufollower formation control of multiple non-holonomic mobile robots incorporatinga receding-horizon scheme. The International Journal of Robotics Research,29(6):727–747, 2010.

[8] L. Consolini, F. Morbidi, D. Prattichizzo, and M. Tosques. Leader-follower formation control of nonholonomic mobile robots with input constraints. Automatica, 40(5):1343–1349, 2008.

[9] N. Hogan. Impedance control: An approach to manipulation. ASME Journal of Dynamic Systems, Measurement, and Control, 107:1–23, March 1985.

[10] O. Khatib. Real time obstacle avoidance for manipulators and mobile robots. The International Journal of Robotics Research, 5(1):90–98, 1986.

[11] Y. Koren and J. Borenstein. Potential field methods and their inherent limitations for mobile robot navigation. In Proceedings of the IEEE conference on Robotics and Automation, pages 1398–1404, Sacramento, California, April 1991.

[12] J. Liu and J. Wu. Multiagent Robotic Systems. CRC Press, 2001.

[13] I. Mas, O. Petrovic, and C. Kitts. Cluster space specification and control of a 3-robot mobile system. In Proceedings of the 2008 IEEE International Conference on Robotics and Automation (ICRA'08), pages 3763–3768, 2008.

[14] F. Michaud, D. Letourneau, M. Guilbert, and J.M. Valin. Dynamic robot formations using directional visual perception. In IEEE/RSJ International Conference on Intelligent Robots and Systems, volume 3, pages 2740–2745, 2002.

[15] H. G. Tanner G. J. Pappas and V. Kumar. Leader-to-formation stability. IEEE Transactions on Robotics and Automation, 20(3):443 – 455, June 2004.

[16] V.T.L. Rampinelli, A.S. Brandao, F.N. Martins, M. Sarcinelli-Filho, and R. Carelli. A multi-layer control scheme for multirobot formations with obstacle avoidance. In International Conference on Advanced Robotics (ICAR), pages 1–6, 2009.

[17] T. Suzuki, T. Sekine, T. Fujii, H. Asama, and I. Endo. Cooperative formation among multiple mobile robot teleoperation in inspection task. In Proceedings of the 39th IEEE Conference on Decision and Control, 2000., volume 1, pages 358 – 363, 2000.

[18] C.R. Weisbin, J. Blitch, D. Lavery, E. Krotkov, C. Shoemaker, L. Matthies, and G. Rodriguez. Miniature robots for space and military issions. IEEE Robotics & Automation Magazine, pages 9 – 18, September 1999.

[19] C. Cruz, "Control de Formacion de Robots Móviles", Tesis doctoral, cap. 2, diciembre 2006.

Appendix A: FIRST APPENDIX

Now the stability of the position errors is analyzed. From (4) for *n*robots $\dot{X}_p = J_r\, U$ arises, assuming perfect velocity tracking $(U \equiv U_c)$ thus:

$$\dot{X}_p = J_f{}^{+}\left(\dot{q}_{fc} + \dot{q}_{obc}\right) + \left(I - J_f{}^{+}J_f\right)\dot{X}_{pc} \quad (30)$$

replacing (13) in (30):

$$J_p{}^{+}\dot{q}_p = J_f{}^{+}\left(\dot{q}_{fc} + \dot{q}_{obc}\right) + \left(I - J_f{}^{+}J_f\right)\dot{X}_{pc} \quad (31)$$

Multiplying both members of equality by $J_2 = I - J_f{}^{+}J_f$and developing

$$J_2 J_p{}^{+}\dot{q}_p = J_2 J_f{}^{+}\left(\dot{q}_{fc} + \dot{q}_{obc}\right) + J_2\left(I - J_f{}^{+}J_f\right)\dot{X}_{pc}$$

as$J_2 J_f{}^{+} = 0$

$$\begin{aligned} J_2 J_f{}^{+} &= \left(I - J_f{}^{+}J_f\right)J_f{}^{+} \\ &= J_f{}^{+} - J_f{}^{+}J_f J_f{}^{+} \\ &= J_f{}^{+} - J_f{}^{+}J_f J_f{}^{T}\left(J_f J_f{}^{T}\right)^{-1} \\ &= J_f{}^{+} - J_f{}^{+} = 0 \end{aligned}$$

then:

$$J_2 J_p{}^{+}\dot{q}_p = J_2\dot{X}_{pc}$$

replacing$\dot{X}_{pc}$ (24):

$$J_2 J_p{}^{+}\dot{q}_p = J_2 J_p{}^{+}\left(\dot{q}_{pd} + K_p \tanh\tilde{q}_p\right)$$

developing:

$$\boldsymbol{J_2 J_p{}^{+}\left(\dot{\tilde{q}}_p + K_p \tanh\tilde{q}_p\right) = 0}$$

Apuntamiento y Estabilización Automática de un Sistema Electro-Óptico ante Perturbaciones

Chaglla S. *; Terán D. *; Morales L. *; Rosales A. *; Pozo D. *; Rosero J. *
**Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica*
e-mail: jsanticabj@hotmail.com;daniloteranl@gmail.com;luis.moralesec@epn.edu.ec; andres.rosales@epn.edu.ec; david.pozo@epn.edu.ec;jorge.rosero@epn.edu.ec;

Resumen: En el presente documento se implementa un sistema electro-óptico, el mismo queposee un algoritmo que permite estabilizar una plataforma con dos grados de libertad ante perturbaciones y realizar apuntamiento hacia una dirección especificada mediante latitud y longitud, a través de "Google Earth". Este sistema, se encuentra constituido por una plataforma con movimientos en el eje vertical (cabeceo) y horizontal (guiñada). Además, dispone de una cámara de espectro visible, una cámara térmica y una cámara de medición láser para adquirir imágenes y tomar medidas de distancias.

Palabras clave:*Electro-óptico, cabeceo, guiñada, cámaras, estabilización, apuntamiento.*

Abstract: In this document, we implement an electro-optical system, which one has an algorithm to stabilize a platform with two degrees of freedom against disturbances and perform pointing in one direction, specified by latitude and longitude in "Google Earth". This system is constituted by a platform that moves in the vertical axis (pitch) and horizontal axis (yaw). It also has a visible spectrum camera, a thermal camera and laser rangefinder to acquire images and take measurements of distances.

Keywords:*Electro-optical, pitch, yaw, cameras, stabilization, pointing.*

1. INTRODUCCIÓN

La historia de los sistemas aéreos no tripulados es tan antigua como la aviación, ya que desde los finales del siglo XIX personajes como Nikola Tesla hablaron acerca de los aviones no tripulados, pero la paternidad de la misma es el norteamericano Elmer Sperry, Lawrence Sperry y Peter Cooper quienes establecieron las bases para esta tecnología, fundamentalmente en las aéreas de control y guiado, el primer avión no tripulado realizo su primer vuelo siendo en el año 1918. [1]

El cuidado del talento humano junto con los avances tecnológicos resultantes de la investigaciónmotiva y permiteel desarrollo de sistemas de navegación autónomos, dentro de los cuáles se encuentran tanto vehículos terrestres no tripulados (UGV, UnmannedGroundVehicles) como vehículos aéreos no tripulados (UAV, UnmannedAerialVehicle), los cuales pueden ingresar en sectores afectados para realizar inspecciones en situaciones de riesgo.Sin embargo, dichos vehículos no tripulados no serían de utilidad, sí no contaran entre sus componentes de un sistema de sensores que permita al operador reconocer el estado de la misión y la localización de dicho vehículo para poder salvaguardarlo.

Un conjunto amplio de aplicaciones coinciden en dotar al UAV de una cámara de vídeo y de los medios adecuados para capturar y/o transmitirla imagen a la estación base. Esta capacidad de adquirir la imagen en un espectro visible o infrarrojo, para su procesamiento manual o automático en línea o en una fase posterior, permite abordar acciones como la inspección, la vigilancia o la búsqueda, tareas de interés en múltiples campos.[2]

En el Ecuador, el Centro de Investigación y Desarrollo de la Fuerza Aérea Ecuatoriana (CIDFAE), conjuntamente con la EPN, se encuentra desarrollando el proyecto UAV, en el cual se pretende crear un Vehículo Aéreo no tripulado con capacidad para realizar misiones tanto para prevención de riesgos como para ayuda a la colectividad. Por esta razón, resulta de gran importancia la creación de un sistema electro – óptico que permita la adquisición y transmisión de imágenes a una estación a tierra.

En el presente documento se explican los conceptos y el modo de funcionamiento de un sistema electro-óptico, el mismo que además de adquirir imágenes, consta de un sistema de estabilización automática y de apuntamiento de objetivos.

De acuerdo con varios autores, se considera a un sensor o sistema electro – óptico (SEO), como una combinación de equipos que trabajan dentro del espectro óptico e incluyen sistemas de imágenes en el espectro visible, ultra violeta, láser, infrarrojo, y fotometría. [3]

Un SEO consta de dos categorías principalmente: la primera los dispositivos de visión e imágenesy la segunda los dispositivos de aplicaciones laser. [4]

2. ARQUITECTURA DE UN SISTEMA ELECTRO - ÓPTICO

El prototipo consta básicamente de una cámara de video en el espectro visible, una cámara de visión térmica, y una cámara de medición láser, las cuales se pretenden utilizar para misiones diurnas, nocturnas y medición de distancias respectivamente.

Además, se integra un sensor de medida inercial (IMU) y un receptor de sistema de posicionamiento global (GPS). La IMU permite conocer el comportamiento del sistema mediante los ángulos de navegación en los tres ejes: cabeceo

(pitch), guiñada (yaw) y alabeo (roll), utilizados para implementar la estabilización del mismo, es decir, recuperar su posición de equilibrio luego de haber sido afectado por una perturbación externa. Por su parte, el receptor GPS permite determinar la ubicación geográfica del sistema mediante sus coordenadas de latitud y longitud, característica aprovechada para realizar apuntamiento de objetivos. El apuntamiento consiste en dirigir la zona de visualización de las cámaras hacia un determinado objetivo del cual se conocen también sus coordenadas geográficas, ya que es seleccionado en un mapa geográfico. De esta manera, conocida la ubicación geográfica tanto del sistema como del objetivo, el computador realiza el cálculo del ángulo de giro necesario para visualizar al objetivo, el mismo que se debe encontrar en un radio máximo de 1 kilómetro de distancia del sistema. Para éste proyecto se usaun sistema de coordenadas Geodésicas el cual describe la posición de un punto sobre la superficie terrestre en términos de latitud, longitud y altura (LLA).[5]

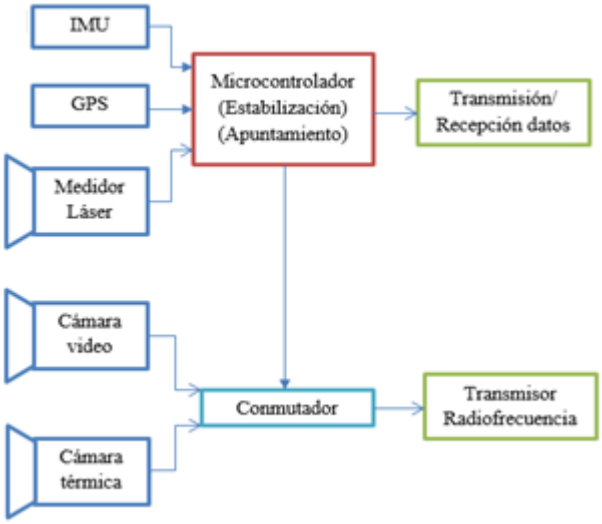


Figura 1. Diagrama de bloques de la plataforma

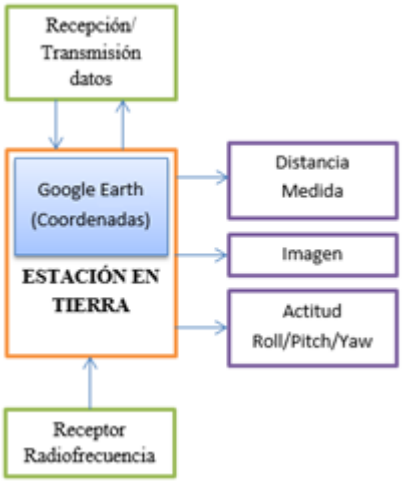


Figura 2. Diagrama de bloques Estación en Tierra

Este tipo de sistemas agregan funcionalidad, e incrementan el campo de aplicación a cualquier tipo de vehículo no tripulado, ya que al ser capaces de adquirir imágenes y enviarlas a un operador localizado a distancia, pueden ser utilizados para realizar tareas de vigilancia, monitoreo, seguimiento, entre otras, sin que sea necesario que un operador se encuentre en el vehículo, eliminando el riesgo para el mismo.

Un sistema electro-óptico debe mantener la misma actitud, para ello consta de movimientos con dos grados de libertad, uno en el eje de azimut o yaw, el cual permite realizar movimientos en el eje horizontal; y en el de elevacion o pitch, que permite realizar los movimientos en el eje vertical del sistema.

3. EQUIPOS UTILIZADOS

Los principales equipos que se utilizan para el sistema electro – óptico se encuentran los descritos en la Tabla 1.

Tabla 1. Equipos Utilizados

CANTIDAD	EQUIPO	DESCRIPCIÓN
1	Computador	Contiene la interfaz de usuario, envía comandos de ejecución, realiza cálculo de ángulo de giro.
1	Plataforma Robótica para cámaras Endurance Robotics PT3	Contiene a los equipos y sensores utilizados.
1	Cámara de video Sony HDR-PJ10	Permite adquirir imágenes en el espectro visible.
1	Cámara de medición láser LTI TruPulse 360	Permite realizar medicion de distancia hacia el objetivo visualizado.
1	Microcontrolador MBED	Realiza funciones de estabilizacion, comando de la plataforma, transmisión de datos al computador.
1	Sensor de medida inercial ArduIMU v3	Proporciona las medidas de los ángulos de navegación
1	GPS LS20031	Proporciona medidas de latitud,longitud y altura de la plataforma.
3	Servomotores Hitec HS-645MG	Permiten realizar el movimiento de la plataforma.
2	Módulos de comunicación Xbee pro S1	Realizan la transmisión de datos entre el computador y el microcontrolador.
1	Sistema de transmisión y recepción de radiofrecuencia	Transmiten las imágenes adquiridas por las cámaras térmica y de video.
2	Baterías LiPo 5000mAh	Alimentación eléctrica del sistema.

4. ADQUISICIÓN DE IMÁGENES

La primera función del sistema consiste en adquirir imágenes por medio de una cámara de video Sony HDR PJ10, la cual es capaz de grabar video con una resolución de 1920x1080 píxeles (Full HD) además de capturar imágenes a 3 Megapíxeles de resolucióny enviarlas a la estación del operador para que este pueda tomar decisiones respecto a la misión en curso. Esta cámara adquiere imágenes en el espectro visible, el mismo que corresponde a la región que el ojo humano es capaz de percibir, y que se encuentra comprendida en una longitud de onda de entre 370 nm y 750 nm aproximadamente. [3]
En la estación en tierra, donde se encuentra el computador y demás equipos de comunicación, el operador del sistema puede observar las imágenes enviadas por las cámaras en la interfaz de usuario desarrollada en Labview.

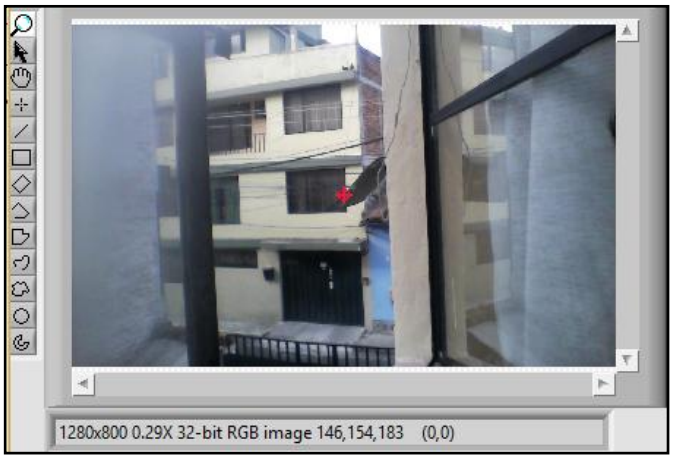


Figura 3. Panel de visualización de imágenes en el espectro visible

4.1Control de zoom

En varias ocasiones, los objetivos que se desean visualizar se encuentran tan alejados que no es posible distinguir detalles relevantes acerca delos mismos, razón por la cual, se implementa un control para el zoom motorizado que posee la cámara de video.

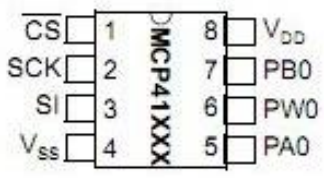


Figura 4. Diagrama de pines potenciómetro digital MCP41010

De fábrica, el control del zoom se realiza por medio de un potenciómetro analógico, el mismo permite ampliar las imágenes hasta 42 veces su tamaño original, sin embargo, en vista de que se requiere que el control sea totalmente digital, este se reemplaza por un potenciómetro digital MCP41010.
El circuito integrado MCP41010 fabricado por Microchip es un potenciómetro digital de 10kΩ entre sus terminales fijos. La variación de la resistencia del terminal variable se controla mediante la interfaz SPI (Serial Peripheral Interface) dicha comunicación se realiza mediante el microcontrolador y los comandos son enviados desde el computador. De esta forma, es posible variar la resistencia conectada a la cámara,permitiendoacercar o alejar imágenes.

4.2 Transmisión de video

La transmisión de imágenes se realiza desde el sistema electro – óptico hacia el computador mediante módulos de transmisión y recepción de radio frecuencia de 3W de potencia, y que operan a una frecuencia de 2.4GHz, para una distancia máxima de transmisión de imágenes de hasta 1000m.
Estas imágenes se visualizan en la interfaz del computador mediante un capturador de video USB. Este es un dispositivo externo que permite ingresar señales de audio y video al computador para poder manipularlo. Se conecta mediante un puerto USB, y posee conectores RCA y S-video para dispositivos externos como televisores, cámaras digitales, etc.

Figura 5. Transmisor y receptor de radiofrecuencia.

5. MEDICIÓN DE DISTANCIA

Otro aspecto de importancia para el sistema es el dispositivo láser para medición de distancia. Se utiliza una cámara TruPulse360 de LTI (Laser Technology Inc.), en la cual se enfoca el objetivo en la pantalla de visualización y la medición de la distancia se hace mediante la emisión de impulsos de energía infrarrojos que son invisibles e inofensivos para la vista. La unidad TruPulse determina la distancia midiendo el tiempo que tarda cada impulso en viajar de ida y vuelta entre el telémetro y el objetivo, los datos obtenidos por el sensor láser son enviados mediante comunicación serial hacia el microcontrolador a una velocidad de 9600 baudios.

Figura 6. Cámara láser para medición de distancia TruPulse360

Este instrumento posee alta sensibilidad y puede funcionar tanto con objetos reflectantes como no reflectantes. La distancia máxima de medición varía dependiendo de la calidad del objetivo y de las condiciones ambientales, de tal manera que para un objeto no reflectante, la distancia máxima de medición es de 1000m; mientras que, para un objeto reflectante, la distancia máxima de medición es de 2000m.Esta distancia puede variar en un bajo rango debido a factores como color y textura del objetivo, ángulo de disparo y condiciones de iluminación; por ejemplo, disparar perpendicularmente a un objetivo brinda un mejor alcance, de igual forma el alcance del medidor aumenta en un día nublado. [6]

6. ESTABILIZACIÓN

La estabilidad es la respuesta de un sistema o cuerpo al moverse de su posición de equilibrio. El equilibrio es el estado de un cuerpo o sistema cuando la resultante de sus fuerzas que actúan sobre él es nula.[7]

La estabilidad estática es la capacidad del cuerpo o sistema de recobrar su posición de equilibrio luego de haber sido afectado por alguna perturbación externa; el presente proyecto debe mantenerse en la posición inicial, es decir conservando su orientación angular con respecto al centro de gravedad del sistema.

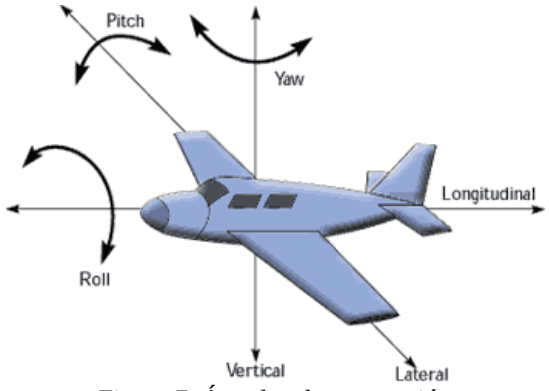


Figura 7. Ángulos de navegación

6.1 Ángulos de navegación

Los ángulos de navegación, conocidos también como ángulos de Euler, son un conjunto de coordenadas angulares usadas para describir la orientación de un objeto en tres dimensiones, es decir, determinar la actitud de un sistema en un momento dado.

Son tres ángulos, que permiten definir la orientación de un sistema de ejes ortogonales móvil respecto a otro sistema de ejes conocido como sistema de referencia. Éstos tres ángulos son guiñada (yaw), cabeceo (pitch) y alabeo (roll); y son equivalentes a tres rotaciones principales, llamadas igual que el eje sobre el cual se producen (roll, pitch y yaw), que permiten conocer la orientación del sistema.

- *Alabeo (Roll Φ):* Es la inclinación o rotación del sistema respecto al eje "x" o eje longitudinal.
- *Cabeceo (Pitch θ):* Es la rotación del sistema respecto al eje "y" o eje lateral o transversal.
- *Guiñada (Yaw Ψ):* Es la rotación del sistema o cuerpo respecto al eje "z" o eje vertical.

6.2 Unidad de Medida Inercial

La unidad de medida inercial(IMU) sirve para determinar la posición instantánea (ángulos de cabeceo, alabeo y guiñada) y la velocidad a la que se desplaza un cuerpo.La medida de estos parámetros se la hace mediante acelerómetros que son elementos que miden la aceleración lineal en términos de la gravedad de la tierray giroscopios que determinan la velocidad angular. Mediante estos datos se puede obtener la posición del sistema.

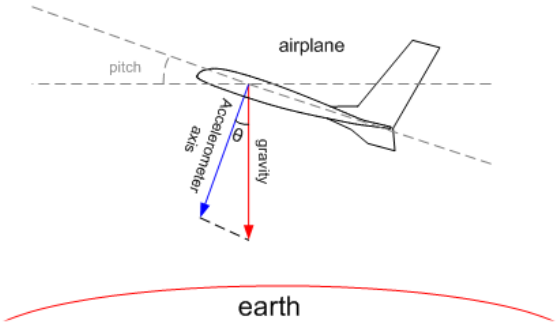


Figura 8. Actitud del sistema

6.3 Actitud

La actitud o AHRS (AttitudeHeading Reference System) es aquella que permite determinar la posición, la velocidad y la orientación de un cuerpo en función del tiempo con respecto a un sistema de referencia que es el eje "tierra" o eje horizontal. [9]

6.3.1 Sistema de navegación inercial

Son los datos que ingresan a un microcontrolador provenientes de los sensores de medida inercial para calcular constantemente la posición en la que se encuentra sin necesidad de poseer una referencia externa.

6.3.2 GPS

Es un sistema de posicionamiento global (Global PositioningSystem) conocido también como un sistema global de navegación por satélite (GNSS) que permite determinar en todo el mundo la posición de un objeto, una

persona o un vehículo mediante sus coordenadas geográficas (latitud y longitud). De forma general, un GPS funciona mediante una red de 24 satélites en órbitas circulares ubicadas a 12 horas sobre el planeta tierra con una inclinación de 55° respecto al ecuador, a 20.200 km; dichos satélites se desplazan a una velocidad de 4km/h, con trayectorias sincronizadas para cubrir toda la superficie de la Tierra.[10]

Para desarrollar el sistema electro – óptico, se utiliza un sensor GPS comercial LS20031, el cuál es un módulo receptor de antena inteligente, constituido por una antena integrada y circuitos del receptor GPS.

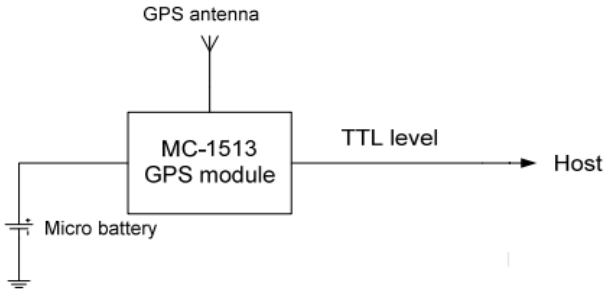


Figura 9. Diagrama de bloques sensor GPS LS20031

Esta unidad es capaz de entregar información acerca de las coordenadas de ubicación a una frecuencia de 5 veces por segundo (5Hz). Esta antena GPS puede hacer seguimiento de hasta 66 satélites a la vez mientras provee una rápida impresión de la primera localización, actualización de la navegación en un segundo, y bajo consumo de energía.

Este sensor se conecta con la unidad de medida inercial mediante comunicación serial, en un rango de velocidad de 57600 bits por segundo y es capaz de entregar salida de datos en 6 diferentes formatos de acuerdo a la NMEA (National Marine ElectronicsAssociation). Se utiliza el formato GLL que corresponde a posición geográfica (Latitud/Longitud) cuya trama de datos es: [11]

$GPGLL,2503.6319,N,12136.0099,E,053740.000,A,A*52

En la misma, se indica en primera instancia un encabezado de inicio de trama, la ubicación geográfica del sistema siendo el primer número correspondiente a latitud y expresado en ddmm.mmmm (d: grados, m: minutos), seguido de un indicador Norte o Sur; el siguiente número corresponde a longitud, expresadoendddmm.mmmm(d: grados, m: minutos), seguido por un indicador Este u Oeste, a continuación se encuentra el indicador de tiempo UTC (Tiempo Universal Coordinado) y finalmente, indicadores de fin de trama de datos.

Dicha trama es receptada, mediante comunicación serial, por la IMU, la cual la procesa y forma una sola trama en la cual se envían datos de Roll, Pitch, Yaw, Latitud, Longitud, y Altura hacia el microcontrolador y son aquellos que se utilizan para implementar los algoritmos de apuntamiento y estabilización.

7. ACTUADORES

Los actuadores son los elementos que permiten el movimiento del sistema. En el caso de los sistemas electro-ópticos es necesario tener dos grados de libertad. Para este fin, en el sistema planteado seutilizanservomecanismos.

Los servomecanismos (servomotores) son sistemas formados por partes mecánicas y partes electrónicas, estos mecanismos realizan un control de posición rotando el eje un ángulo de giro deseado.

Un servomecanismo realiza el control de posición en lazo cerrado, consta de un motor, un potenciómetro y un circuito de control con caja reductora.

Figura 10. Componentes de un servomotor. *[12]*

El potenciómetro se usa para el posicionamiento y con los trenes de pulsos el circuito de control revisa si el ángulo es correcto, en el caso de ser correcto el motor se encuentra apagado, mientras que en el otro caso el motor se direcciona a la nueva posición hasta llegar al ángulo adecuado. La caja reductora aumenta el torque del servomecanismo y reduce la velocidad del mismo

En el sistema electro-óptico desarrollado se utilizan servomotores que operan a una velocidad de 0.2 s/60°, que poseen un par de arranque de 9.6 kg.cm, y que pueden girar hasta 210° tanto en pitch como en yaw.

8. APUNTAMIENTO DE OBJETIVOS

El sistema tiene la capacidad de realizar el apuntamiento de objetivos, es decir, localizar en las cámaras un objetivo que el operador seleccione por medio del computador desde un mapa geográfico.

Para ello, el sistema debe conocer en primer lugar su posición en la superficie terrestre por medio de sus coordenadas geográficas, ésta posición es enviada por el sensor GPS incorporado en el módulo. Luego, se debe seleccionar un objetivo en un mapa geográfico, a partir de esto, se calculan las coordenadas de dicho punto y se calcula el ángulo existente entre el norte y el objetivo; dato que es enviado al microcontrolador para que calcule, respecto a la actitud del sistema, la dirección y el ángulo a girar.

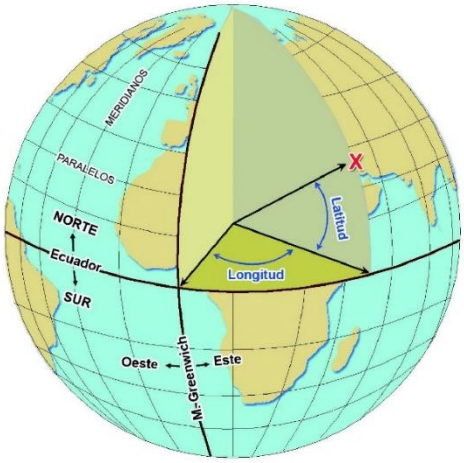


Figura 11. Coordenadas geográficas

La localización de un punto sobre la superficie terrestre se expresa en términos de:

Latitud: Es la distancia angular entre la línea ecuatorial y un punto sobre la superficie terrestre, a través del meridiano. Se mide en grados, puede ser norte o sur y entre un rango de 0° a 90°.

Longitud: Es la distancia angular entre el meridiano 0° (Meridiano de Greenwich) y un punto sobre la superficie terrestre con ayuda de los meridianos. Se mide en grados, pueden ser este u oeste y entre un rango de 0° a 180°.

Para obtener la ubicación geográfica del objetivo, se utiliza la interfaz de programación de aplicaciones (API) de Google Earth, el cual permite obtener coordenadas de latitud, longitud y altura de un determinado punto en la superficie terrestre directamente hacia LabView, a partir de su posición en pantalla.

En la Figura12, se muestra la posición del sistema electro – óptico (representado por una aeronave) y el objetivo seleccionado, para lo cual se utiliza programación de archivos KML (KeyholeMarkupLanguage) que es un lenguaje utilizado por Google Earth para representar datos geográficos en tres dimensiones.

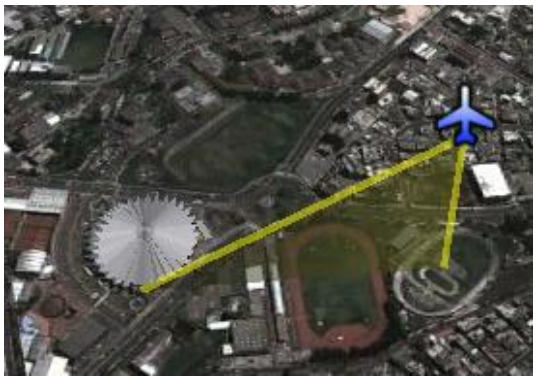


Figura 12. Ubicación de un objetivo seleccionado mediante un mapa geográfico

Dentro del API de Google Earth se encuentra una interfaz denominada **IApplicationGE**, que consta de una serie de funciones que permiten manejar el programa. Se utiliza la función **GetPointOnTerrainFromScreenCoords**, la misma que toma un punto de referencia del mapa y lo usa como centro para poder calcular las demás coordenadas de acuerdo con la posición punto en la pantalla. [13]

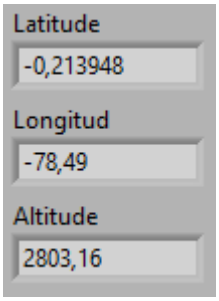


Figura 13. Posición del objetivo seleccionado expresada en coordenadas geográficas

Mediante el software implementado se obtiene la posición del objetivo (Coliseo Rumiñahui – Quito, Ecuador) en las coordenadas indicadas en la Figura 13.

Las coordenadas obtenidas se han comprobado utilizando "Google Maps".Posteriormente, se calcula el ángulo del objetivo respecto a la posición del sistema utilizando lassiguientesecuaciones empleadas ensistemas de aviación: [14]

$$ANG = a\tan 2(S, C) \quad (1)$$

$$S = \sin(\text{LON2-LON1})*\cos(\text{LAT2}) \quad (2)$$

$$C = \cos(\text{LAT1})*\sin(\text{LAT2})-\sin(\text{LAT1})*\cos(\text{LAT2})*\cos(\text{LON2-LON1}) \quad (3)$$

Tanto los datos que opera como el resultado se encuentran en radianes por lo que se debe transformar a grados antes de ser ingresados a la ecuación.

9. FUNCIONES DEL MICROCONTROLADOR

El microcontrolador es el elemento principal de control del sistema electro – óptico, ya que ordena todas las acciones que deben realizar los servomecanismos y demás elementos a controlarse, se utiliza un microcontrolador MBED LPC1768.

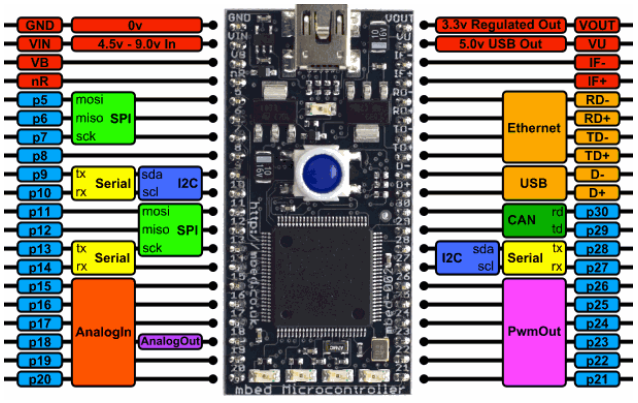


Figura 14. MbedLPC 1768.

Las características del microcontrolador se presentan en la siguiente Tabla 2:[15]

Tabla 2. Características del microcontrolador LPC 1768

PARÁMETRO	CARACTERÍSTICA
Velocidad	96 Mhz
Memoria Ram	32 KB
Memoria Flash	512KB
Alimentacion	5V USB – 4.5-9V
Comunicacion	SPI –I2C –UART
Pines	40 Entrada/Salida
PWM	6 pines
ADC	6 pines

Se utiliza comunicación serial para enlazar al computador y al microcontrolador a una velocidad de 115200 baudios, la cual es fundamental ya que envía y recibe datos y comandos que permiten el funcionamiento del sistema. La comunicación inalámbrica se realiza mediante módulos "Xbee pro s1"que tienen un alcance máximo de 1km a campo abierto.En pruebas realizadas en el Campus de la Escuela Politécnica Nacional con respecto al alcance de transmisión de datos de los módulos se tuvo medidas máximas de hasta 500 metros.

Una de las funcionesde esta comunicación es controlar el ancho de pulso de las señales PWM útiles parael movimiento de los servomotores, tanto en control manual, donde se ingresa un comando mediante teclado o joystick; como para su movimiento cuando se desea realizar estabilización o apuntamiento.

Debido a la existencia de un solo canal de transmisión de video, el microcontrolador debe realizar la multiplexación entre las imágenes de la cámara de espectro visible o la cámara térmica, para poderlas observar en la interfaz.

Otra de sus funciones es recibir los datos de guiñada y cabeceo obtenidos mediantela unidad de medida inercial, para poder conocer la orientación del sistema y de esta manera poder realizar un algoritmo de control para poder mantener una actitud deseada.

Además permite obtener los datos del sensor GPS de la ubicación geográfica de la plataforma, con dichos datos se realiza el direccionamiento de la plataforma hacia la zona de apuntamiento seleccionada por el usuario desde el computador.

El microcontroladorrecibe las tramas de datos enviadas en forma de caracteres, tanto por el sensor de medida láser como por el sensor de medida inercial, y los procesa de tal manera que los convierte en datos útiles para realizar cálculos matemáticos y utilizarlos en los algoritmos de estabilización y apuntamiento. El algoritmo de estabilización consiste en hacer corrección de la posición del sistema aumentando o disminuyendo ángulos en pitch o en yaw, la resolución de giro está ligada al máximo de los servomotores que es de 210 grados. Mientras que el de apuntamiento se realiza en su mayor parte en el computador, sin embargo, el microcontrolador se encarga de mover al sistema en el ángulo indicado por el computador.

10. CONCLUSIONES

Un sistema electro – óptico se encuentra conformado por elementos que trabajan dentro del espectro óptico de radiación electromagnética, y entre los cuales se encuentran imágenes en el espectro visible, infrarrojos, láser, etc., que permitan adquirir información que sea útil para ser procesada.

La medición de distancia utilizando tecnología láser permite obtener grandes alcances de medición incluso cuando se apunte hacia objetos no reflectantes. Para el medidor Láser TruPulse utilizado, el mayor alcance de medición utilizando objetos no reflectantes se encuentra cercano a 1000 metros, mientras que para un objeto reflectante, dicho alcance prácticamente se duplica, cercano a los 2000 metros. Para mejorar el alcance, se debe disparar perpendicularmente hacia el objetivo antes que dispararlo en ángulo cerrado, además de realizar mediciones en días nublados.

La distancia de transmisión de datos utilizando equipos de radiofrecuencia depende de la potencia de transmisión de dichos equipos junto con sus antenas. Por ejemplo, para los módulos X-bee, la potencia de transmisión es cercana a los 50mW por lo que su distancia máxima aproximada de transmisión es de 1 kilómetro; mientras que en los equipos para transmisión de video su potencia es de 3W por lo que también su distancia máxima de transmisión es de 1000 metros, esto debido a que transmiten audio y video, no solamente datos. Sin embargo, estos alcances máximos se ven afectados por la banda de frecuencia en la cual trabajan, ya que al utilizar la banda de frecuencia de 2.4 GHz se ven expuestas a grandes cantidades de interferencia, debido a que al ser una banda libre, varios equipos trabajan en la misma frecuencia, provocando interferencia y por tanto problemas en la comunicación.

Se realizó un análisis para la selección del microcontrolador a usarse en función de la cantidad de elementos a controlar, adquisición de datos y transmisión de datos, ya que el poseer recursos en exceso es innecesario para el sistema.Se utilizó los tres pórticos seriales que posee uno para la comunicación con el computador, el segundo puerto para la comunicación con la IMU, y el tercero para la cámara laser. Se utilizó la comunicación SPI para el control del zoom mediante el dispositivo MCP41010, el control PWM que ofrece este microcontrolador es bastante fino y de esta manera se puede manejar los servomotores que se encuentran en pitch y yaw, además de salidas digitales necesarias.

Se determinó los tiempos necesarios a los cuales se debe realizar la adquisición de datos de elementos como la unidad de medida inercial, los datos de GPS, los datos de cámara laser para poseer un correcto funcionamiento del sistema en conjunto.Al trabajar ciclos en labview bastante rápidos se producía un error de sobresaturación en el microcontrolador,

al momento de la recepción y transmisión que ejecutaba el microcontrolador, además la recepción que posee el microcontrolador de datos de IMU y laser se determinó que los ciclos de labview se deben realizar en cada 80 ms o tiempos mayores, y que los tiempos mínimos a los cuales se debe adquirir los datos de la IMU no deben ser menores a 200 ms y que los datos de la laser no deben ser menores a 20s ya que la laser por fabricación se apaga transcurrido ese tiempo.

Debido a que se trabaja con un microprocesador de 8bits los rangos de variación del zoom se realiza hasta en 256 pasos, esto se encuentra en una relación directa con el potenciómetro digital que controla el zoom de la cámara. La resistencia que se debe variar para bajar el zoom se encuentra bajo los 3.4Kohms y para aumentar el zoom es a partir de 7.8Kohms.

El software de programación gráfica LabView posee la capacidad de comunicarse e interactuar con programas externos cuya interfaz de programación de aplicaciones (API) se encuentre disponible, como por ejemplo Google Earth, cuyo API permite abrir un mapa geográfico y determinar, entre otras cosas, la posición geográfica de un punto dentro del mismo, en la ventana principal de LabView; capacidad aprovechada para implementar el algoritmo de apuntamiento de objetivos.

11. REFERENCIAS

[1] L. G. F. Bouillon, «Tecnologia Aeroespacial,» *perfilesIDS*, vol. 1, nº 1, p. 100, 2009.

[2] A. Barrientos, «Vehículos aéreos no tripulados para uso civil. Tecnología y,» Madrid.

[3] «Electro optics,» [En línea]. Available: www.tscm.com/eo-ir.pdf.

[4] K. Lane, «http://www.pghtech.org/,» [En línea]. Available: http://www.pghtech.org/Docs/pdf/News-And-Publications/Sub-Clusters/Electro-Optics.pdf. [Último acceso: 08 09 2013].

[5] M. I. E. Escalero, Control de sensores en aviones no tripulados y tratamiento de de imagenes utilizando maquinas de soporte de estados, Madrid: Escuela Politecnica Superior, 2010.

[6] Laser Technology Inc, TruPulse 360/360B Manual de usuario, 2009.

[7] C. A. Cabrera, «monografías.com,» [En línea]. Available: http://www.monografias.com/trabajos14/equilibriocuerp/equilibriocuerp.shtml. [Último acceso: 08 09 2009].

[8] Wikipedia, [En línea]. Available: http://es.wikipedia.org/wiki/%C3%81ngulos_de_navegaci%C3%B3n. [Último acceso: 07 06 2013].

[9] D. Pozo, Diseño y construccion de una plataforma didactica para medir angulos de inclinacion usando sensores inerciales como acelerometros y giroscopios, Quito, 2010.

[10] N. Chacón Mejía, «Sistema de posicionamiento Global,» [En línea]. Available: http://ocw.utpl.edu.ec/ingenieria-civil/topografia-aplicada/unidad-4-sistema-de-posicionamiento-global.pdf.

[11] LOCOSYS Technology Inc., «Spark Fun Electronics Inc.,» [En línea]. Available: https://www.sparkfun.com/products/8975. [Último acceso: 20 05 2013].

[12] UTSFM, «Servomotores,» Diciembre 2003. [En línea]. Available: http://www2.elo.utfsm.cl/~mineducagv/docs/ListaDetalladadeModulos/servos.pdf.

[13] Google, «Google Earth COM API Documentation,» [En línea]. Available: http://web.archive.org/web/20100408210316/http://earth.google.com/comapi/annotated.html.

[14] R. Bullock, 5 Junio 2007. [En línea]. Available: http://www.dtcenter.org/met/users/docs/write_ups/gc_simple.pdf.

[15] Mbed, «www.mbed.org,» [En línea]. Available: http://mbed.org/platforms/mbed-LPC1768/.

Diseño e Implementación de un Sistema de Localización y Mapeo Simultáneos (SLAM) para la Plataforma Robótica ROBOTINO®

Narváez V. *; Yandún F. *; Pozo D. *; Morales L. *; Rosero J. *; Rosales A. *; Auat F. **

**Escuela Politécnica Nacional, Facultad de Ingeniería Eléctrica y Electrónica, Quito, Ecuador*
e-mail: victor_narvaez_t@hotmail.com; fynsk8@yahoo.es; andres.rosales@epn.edu.ec;
david.pozo@epn.edu.ec;luis.moralesec@epn.edu.ec; jorge.rosero@epn.edu.ec;
*** Universidad Técnica Federico Santa María, Departamento de Electrónica, Valparaíso, Chile*
e-mail: fernando.auat@usm.cl

Resumen:*En este documento se presenta el desarrollo e implementación en MATLAB de un algoritmo de localización y mapeo simultáneo (SLAM)para un robot móvil comercial, operando en entornos estructurados estáticos, cuya localización está dada por la odometría del robot y el modelo del mismo. Para la realización del mapeo se utiliza el sensor de rango laser Hokuyo URG-04LX, una vez obtenido el mapa se implementa un algoritmo capaz de detectar esquinas dentro del entorno, con el fin de que puedan ser utilizadas en futuras aplicaciones de SLAM donde sea necesario la inclusión de "landmarks" o características de entorno. Para atenuar los errores de posicionamiento inherentes a la odometría propia del robot se implementa un Filtro Extendido de Kalman entre la odometría y el modelo omnidireccionaldel robot móvil, en base a lo cual se obtienenmejores resultados para lalocalización. La detección de esquinas se realiza usando la Transformada de Hough, algoritmo mediante el cual se identifica líneas presentes en el mapa y posteriormente se halla esquinas como un cruce entre dos líneas. Se presentan los resultados como un análisis de los mapas obtenidos contrastados con los entornos reales, además del análisis de errores sobre la localización del robot.*

Palabras Clave:*ROBOTINO, SLAM, Odometría, Transformada de Hough.*

Abstract:*This paper presents the development and implementation of a Simultaneous Localization And Mapping (SLAM) algorithm for a commercial mobile robot, operating in structured and static environments. The location is given by the odometry of the robot and its model. To perform the mapping a sensor laser range Hokuyo URG - 04LX is used.Later, an algorithm capable of detecting corners in the environment is implemented, they can be used in future applications of SLAM where being necessary the inclusion of landmarks or characteristics of environment. To mitigate the position errors inherent to odometryof the robot is implemented an Extended Kalman Filter between the odometry and omnidirectional mobile robot model. This technique improves the results of localization. Corner's detection is performed using the Hough transform algorithm, which identifies lines on the map and subsequently corners like a cross between two lines. We present the results as an analysis of the obtained maps that include contrasts with the real environments. In addition,we analyze the location error of the robot.*

Keywords: *ROBOTINO, SLAM, Odometry, Hough Transform.*

1. INTRODUCCIÓN

Uno de los principales objetivos de la robótica móvil es dotar a los robots de la autonomía necesaria para ser capaces de navegar a través de ambientes desconocidos brindando información acerca del entorno sin la ayuda del ser humano.Siendouno de los principales inconvenientes lalocalización del robot móvil cuando se desplaza por un entorno desconocido [1-3], el presente trabajo se enfoca en solucionar el problema de localización y mapeo mediante una reconstrucción del entorno en dos dimensionesutilizando la información entregada por un sensor laser [4], el cual entrega lecturas de distancias, que procesadas mediante un algoritmo se convierten en un mapa del entorno. En el presente trabajo se utiliza el sensor laser Hokuyo URG-04LX [5].

Adicionalmente, la plataforma robótica Robotino® [6]cuenta con encoders que entregan información de la posición del móvil con respecto a un sistema de coordenadas fijo. La estimación del desplazamiento de la plataforma móvil está

sujeta a un margen de error considerable debido principalmente a la resolución de los encoders y al deslizamiento de las llantas sobre la superficie.Por esta razón, es necesario el desarrollo de un algoritmo que permita obtener una posición fiable del móvil a lo largo del tiempo, dicho algoritmo es el Filtro Extendido de Kalman, el cual será estudiado en las siguientes secciones.

Para la implementación del Filtro Extendido de Kalman [7],[8], se ha considerado estimar la posición del robot móvil en base a su modelo cinemático omnidireccionaly como observador se tendrán los datos por odometría.

La detección de características (esquinas o *landmark*s) del entorno, es realizada tomando al mapa como una imagen binarizada y usando los diferentes métodos de detección de bordes que posee MATLAB y posteriormente aplicando la Transformada de Hough [9].

2. DESCRIPCIÓN DEL ROBOT MÓVIL

Robotino® ha sido diseñado por *FestoDidactic*, siendo una herramienta didáctica que permitemanejar un robot a nivel académico con una diversa gama de sensores quefacilitan la realización de diferentes tareas de control.

Figura 1. Robotino®, tomado de [6]

El software provisto por el fabricante para controlar al robot es ROBOTINOView, pero existe la posibilidad de controlarlo mediante los programas MATLAB o LABView con el uso de las respectivas librerías desarrolladas.Nuestro trabajo ha sido desarrollado utilizando MATLAB.

2.1 Accionamiento

Robotino® es una plataforma móvil con accionamiento omnidireccional, puede desplazarse en cualquier dirección sin necesidad de girar mediante el cambio de orientación.Cuenta con tres unidades de accionamiento omnidireccional permitiéndole movimientos en todas las direcciones: adelante, atrás, y lateralmente, considerando que además puede girar sobre su propio eje (manteniendo o no su orientación), ello significa que dispone de tres grados de libertad: dos de traslación y uno de rotación.

2.2 Modelo Cinemático

El esquema para el modelo omnidireccional del robot móvil se presenta en la Fig. 2.

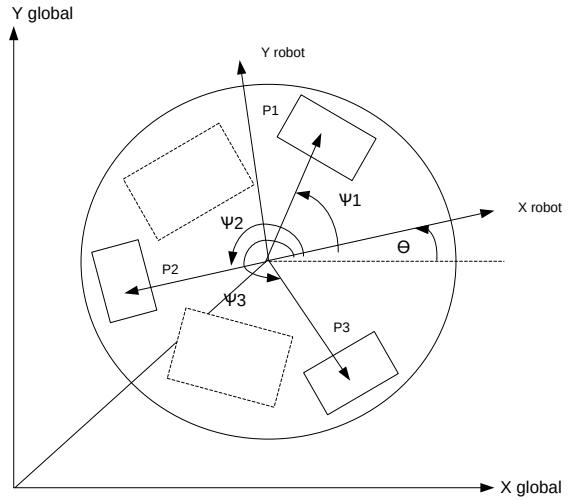


Figura 2. Representación gráfica de un robot con modelo omnidireccional

La expresión (1) representa la relación matemática entre las variables descritas en la Fig. 2 y el movimiento del robot.

$$\begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{\theta}\end{bmatrix} = r * \begin{bmatrix}-\sin(\theta+\psi 1) & \cos(\theta+\psi 1) & R\\ -\sin(\theta+\psi 2) & \cos(\theta+\psi 2) & R\\ -\sin(\theta+\psi 3) & \cos(\theta+\psi 3) & R\end{bmatrix}^{-1} * \begin{bmatrix}\omega 1\\ \omega 2\\ \omega 3\end{bmatrix} \quad (1)$$

En la ecuación (1), ω1, ω2, ω3 son las velocidades angulares de cada uno de los motores del robot, r es el radio de las ruedas, R es la distancia entre una rueda y el centro del robot, θ es el ángulo de orientación del robot respecto a los ejes de coordenadas globales, ψ1, ψ2, ψ3 corresponden a los ángulos en los cuales se disponen las ruedas del robot, finalmente ẋ, ẏ y$\dot{\theta}$ son las velocidades lineales y angular del robot. Todas las variables están referidas al sistema de coordenadas global. Además, la ecuación (1) describe el movimiento en tiempo continuo, aunque es necesario discretizar el modelo obtenido para realizar la implementación en base a un algoritmo de tipo recursivo, tal como se muestra en la ecuación (2).

$$\begin{bmatrix}x\\ y\\ \theta\end{bmatrix} = \begin{bmatrix}x_{k-1}\\ y_{k-1}\\ \theta_{k-1}\end{bmatrix} + \Delta T * r * F(\theta_{k-1}) * \begin{bmatrix}\omega 1_k\\ \omega 2_k\\ \omega 3_k\end{bmatrix} \quad (2)$$

Donde el subíndice k indica el tiempo, ΔTes el tiempo de muestreo y$F(\theta)$ está dado por la siguiente expresión:

$$F(\theta) = \begin{bmatrix} -\sin(\theta+\psi 1) & \cos(\theta+\psi 1) & R \\ -\sin(\theta+\psi 2) & \cos(\theta+\psi 2) & R \\ -\sin(\theta+\psi 3) & \cos(\theta+\psi 3) & R \end{bmatrix}^{-1} \quad (3)$$

Más detalles de la obtención del modelo pueden ser encontrados en [10].

3. ALGORITMO DE CONTROL

El algoritmo de control está constituido por tres secciones: elalgoritmo del Filtro Extendido de Kalman (EKF) para filtrar los datos de odometría provenientes del robot, el algoritmo que permite la construcción del mapa en base a los datos del sensor laser y un algoritmo basado en la Transformada de Hough para la detección de características (esquinas).

3.1 Implementación del Filtro de Kalman

El filtro toma como base para la etapa de predicción el modelo de un robot omnidireccional, el cual se describió previamente en (2), y como observador para la etapa de actualización los datos de la odometría (*x*, *y*, *θ*) brindados por el software del robot.

3.1.1 Etapa de Predicción

En esta etapa se usa el modelo cinemático del robot (2) para obtener una estimación previa de la posición (*x*, *y*, θ) del robot.

El cálculo de la matriz de covarianza estimada, considera el valor de la matriz de covarianzas en un instante anteriorP_{k-1}, además se tiene W_k que corresponde al Jacobiano del modelo con respecto al ruido, con lo cual se obtiene la siguiente expresión [8]:

$$P_k^- = A_{x,k} P_{k-1} A'_{x,k} + W_k Q_{k-1} W'_k \quad (4)$$

Donde $A_{x,k}$ es el Jacobiano del modelo con respecto al vector x. Esta ecuación es completamente válida, pero si se desea tomar en cuenta las perturbaciones en las entradas de control, esta expresión puede extenderse y añadir un término que asocie también la dinámica de tales incertidumbres, quedando la predicción de la matriz de covarianza como:

$$P_k^- = A_{x,k} P_{k-1} A'_{x,k} + A_{u,k} U_k A'_{u,k} \quad (5)$$

Donde $A_{u,k}$ es el Jacobiano del modelo con respecto a las entradas de control y U_k es la matriz de covarianzas asociada al proceso, cuyo valor inicial es determinante para que el filtro funcione adecuadamente.Este valor inicial fue encontrado experimentalmente y se encuentra expresado como se muestra en la expresión (6).

$$U_k = \begin{bmatrix} {\sigma_x}^2 & 0 & 0 \\ 0 & {\sigma_y}^2 & 0 \\ 0 & 0 & {\sigma_\theta}^2 \end{bmatrix} \quad (6)$$

donde σ_x^2, σ_y^2 y σ_θ^2 representan las varianzas correspondientes al proceso en los ejes *x*, *y*, *Θ*, respectivamente.

3.1.2 Etapa de Actualización

En primer lugar se calcula la ganancia de Kalman [8]:

$$K_k = P_k^- H'_k (H_k P_k^- H'_k + R_k)^{-1} \quad (7)$$

Donde H_k es la matriz identidad de 3x3 y se obtiene directamente de la posición adquirida por odometríaconsiderando al observador como un modelo de tipo lineal. Además, R_k es la matriz de covarianza asociada a la observación (odometría del robot), cuyo valor inicial es determinante para que el filtro funcione adecuadamente, este valor inicial fue encontrado experimentalmente.

$$R_k = \begin{bmatrix} {\sigma_{xr}}^2 & 0 & 0 \\ 0 & {\sigma_{yr}}^2 & 0 \\ 0 & 0 & {\sigma_{\theta r}}^2 \end{bmatrix} \quad (8)$$

donde σ_{xr}^2, σ_{yr}^2 y $\sigma_{\theta r}^2$ representan las varianzas correspondientes al observador (odometría del robot) en los ejes *x*, *y*, θ, respectivamente.

Posteriormente se calcula la posición corregida que servirá para realizar el mapeo. La expresión (9) muestra el cálculo de la posición corregida [8].

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - \hat{x}_k^-) \quad (9)$$

Finalmente, se calcula la nueva matriz de covarianza [8].

$$P_k = (I - K_k H_k)\, P_k^- \quad (10)$$

Al final de todo el proceso de filtrado, los valores de la matriz de covarianzas deberían ser bajos, puesto que el objetivo del filtro es disminuir los errores y con esto la covarianza de los elementos del vector de estados (diagonal principal de P). El vector de localización filtrado va a tener datos almacenados cada 0.5 segundos y con los cuales se va a realizar el mapeo del entorno.

3.2 Mapeo

Una vez que se han obtenido los datos por parte del sensor

láser y se han filtrado los datos de la odometría del robot, se aplica el algoritmo de mapeo para obtener la imagen del entorno en dos dimensiones.

Los datos que entrega el sensor láser están en coordenadas polares, es por ello que se los debe convertir a coordenadas rectangulares para posteriormente construir el mapa. Con esta transformación es preciso también establecer los ejes de coordenadas para el sistema de referencia local del robot, esto se lo realiza con las ecuaciones 11 y 12:

$$x = r * \cos\left(\theta - \frac{\pi}{2}\right) \quad (11)$$

$$y = r * \sin\left(\theta - \frac{\pi}{2}\right) \quad (12)$$

Quedando el sistema de referencia local del robot, como se muestra en la Fig. 3.

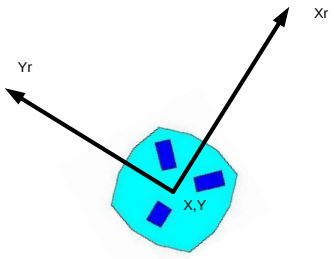


Figura 3. Sistema de referencia local al robot

Es así que los datos del sensor láser quedan relacionados al sistema de referencia local del robot, siendo necesario realizar transformaciones tanto de rotación como de traslación teniendo en cuenta el movimiento del robot, para que los datos obtenidos estén referidos al sistema global de coordenadas, como se muestra en las ecuaciones 13 y 14.

$$xmapa = xx * \cos(\theta_a) - yy * \sin(\theta_a) + x \quad (13)$$

$$ymapa = xx * \sin(\theta_a) + yy * \cos(\theta_a) + y \quad (14)$$

donde*xx* e *yy*son las medidas del láser en coordenadas cartesianas,*x*, *y* y $\theta_a$representan la posición actual del robot (luego de ser filtradas).

3.3 Detección de Características

Para esta etapa se utiliza la Transformada de Hough con el fin de detectar las líneas del mapa y posteriormente encontrar las esquinas como una intersección de dos líneas.

La Transformada de Hough[11]es una técnica que brinda la posibilidad de detectar figuras dentro de una imagen previamente binarizada, su uso se limitaba a la detección de rectas en una imagen, pero con el transcurrir del tiempo, se extendió hasta llegar a la identificación de cualquier Fig. que se pueda describir con parámetros, tales como circunferencias y elipses. En el presente artículo se hace énfasis en el estudio del reconocimiento de líneas.

3.3.1 Binarización de la imagen

En primera instancia se crea una cuadrícula con una separación de 6 cm entre cada línea (entre menor sea esta separación se obtiene mayor exactitud, pero el software trabajaría con un mayor número de datos). El objetivo es colocar los puntos del mapa obtenido en la intersección de líneas de la grilla más cercana a ellos, y posteriormente colocar unos (1L) en estas intersecciones y ceros (0L) en las cuales no lo hacen, teniendo así el mapa como una imagen binaria.

3.3.2 Detección de Bordes

Luego de obtener la imagen binarizada, se detectan los bordes de la misma, con este fin se usaron dos de los métodos que proporciona MATLAB para este efecto: Roberts y Prewitt [12]. Es importante destacar que estos métodos se aplicaron por separado a la misma imagen binarizada ya que los resultados obtenidos se complementaban entre sí.Se observó que ciertas líneas existentes en la imagen se detectaban por Roberts, y otras se detectaban con Prewitt, razón fundamental por la que se optó por someter a las imágenes a los dos métodos, recogiendo todos los datos de líneas al final del tratamiento de la imagen.

3.3.3 Detección de Líneas

Una vez detectados los bordes de la imagen, se aplica la Transformada de Hough, encontrando primero los parámetros de las posibles líneas, para posteriormente detectar los picos donde es más probable encontrar una línea en la imagen.

3.3.4 Localización de esquinas

Una vez detectadas las líneas, se promedian las que tengan unarango de distancia entre puntos extremos iniciales y finales, menor a un valor determinado.Posteriormente se encuentran las esquinas como un cruce entre líneas promediadas, teniendo en cuenta que una esquina va a estar en el cruce de dos rectas y el ángulo entre ellas debe ser aproximadamente 90°.

4. PRUEBAS Y RESULTADOS

Observamos la implementación en entornos reales[1], filtrando la odometría y variando diferentes parámetros del filtro o parámetros de las funciones que permiten detectar las características del entorno (esquinas). Para presentar las

[1]Las pruebas fueron realizadas en espacios físicos de libre acceso pertenecientes a la Facultad de Ingeniería en Electrónica y Control de la Escuela Politécnica Nacional (Quito- Ecuador)

pruebas y resultados se han tomado datos en dos ambientes diferentes, ambos estructurados y estáticos, tal como se muestra en la Fig. 4.

Figura 4. Pasillo

4.1 Mapeo

Se comparan los resultados del procesamiento usando únicamente los datos de la odometría, y los datos del procesamiento usando el Filtro de Kalman, se puede apreciar que la odometría entrega datos relativamente aceptables hasta que en cierto punto los errores acumulativos entorpecen el procesamiento y el mapa comienza a desviarse de su correcta construcción. Sin embargo, cuando se aplica el Filtro de Kalman el error disminuye y la elaboración del mapa mejora.

Además, se muestran los resultados al cambiar las constantes del Filtro de Kalman (Figuras 7 y 8) para la inicialización de las matrices de covarianza tanto de la observación (R) como en el proceso (Q). Estas variables contienen la ponderación de los errores presentes en cada caso.Los valores con los cuales se obtuvieron resultados más satisfactorios fueron encontrados de manera heurística, considerando el criterio de que al asignar valores pequeños, las perturbaciones se asumirán como pequeñas; mientras que si se inicializa con valores grandes, la perturbación será considerable.

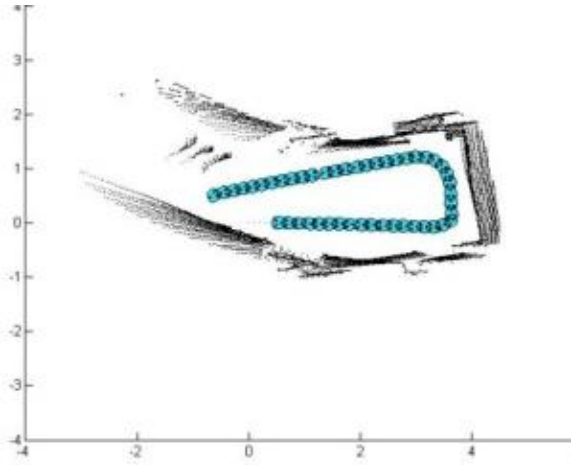


Figura 5. Mapa obtenido con odometría

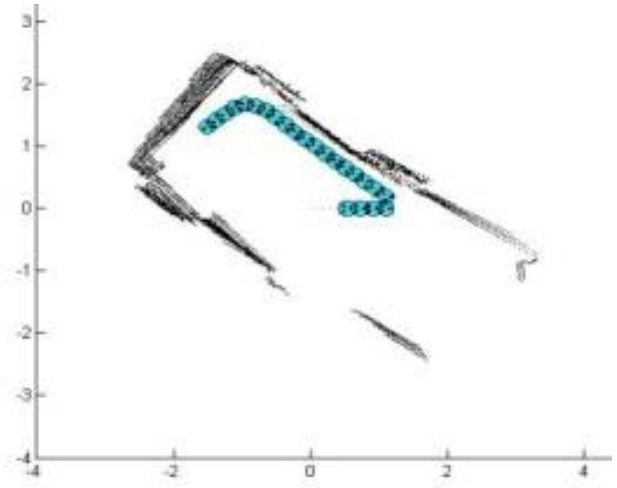


Figura 6. Mapa obtenido aplicando el Filtro Extendido de Kalman

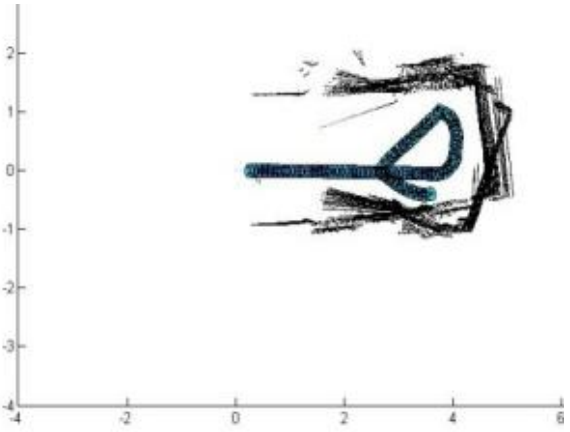


Figura 7. Mapa obtenido con constantes incorrectas en el Filtro Extendido de Kalman

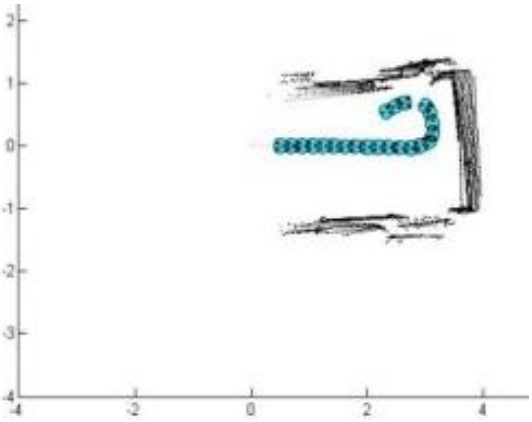


Figura 8. Mapa obtenido con constantes correctas para el Filtro Extendido de Kalman

Además de la experimentaciónen entornos estructurados y estáticos, se realizaron pruebas cuando algunas personas se encontraban circulando en el ambiente de prueba, esto se

puede notar en las marcas existentes dentro del mapa mostrado en la Fig. 9.La existencia de objetos móviles en el área de mapeo es algo crucial para el algoritmo, puesto que las personas pueden ser confundidas con características, lo que llevaría a un error en la detección de las mismas.Estos errores están descartados en esta aplicación puesto que el presente proyecto está orientado a entornos estáticos y estructurados.

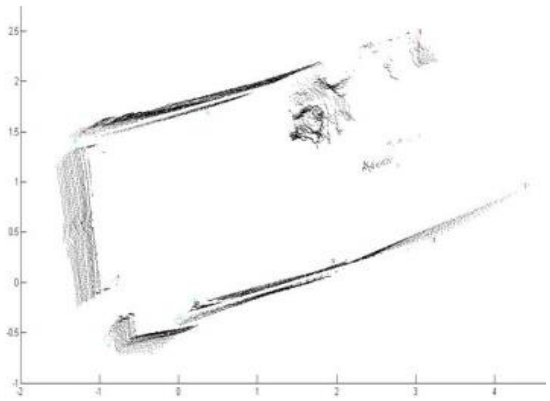


Figura9. Mapa obtenido con personas caminando en el entorno de prueba

A continuación, se muestra uno de los mapas finales, construido en el entorno de trabajo (Fig. 4)luego de afinar las constantes del EKF (*Extended KalmanFilter*) y realizar la detección de esquinas, las cuales se muestran como círculos en el gráfico.

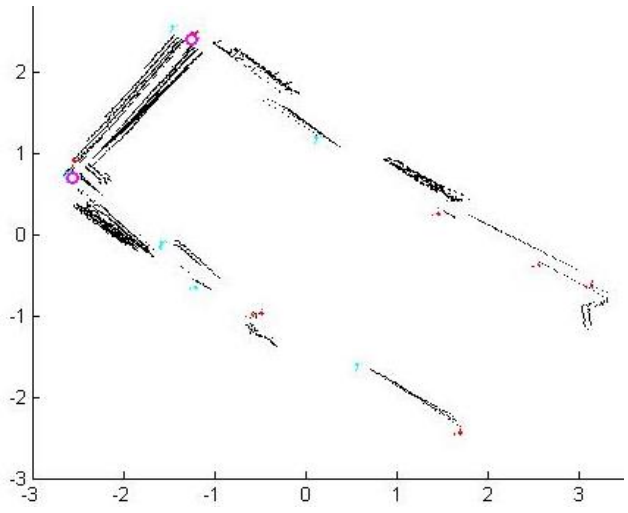


Figura 10. Mapa final

4.2 Localización

Se analiza la posición final del robot y la precisión en la detección de esquinas luego de un recorrido, también se compara la posición del robot en cada iteración del algoritmo durante un mismo recorrido, obteniendo los resultados que se muestran en la Tabla 1.

Tabla 1. Valores obtenidos en cuatro pruebas en entornos reales

Entornos	Variables	Odometría	Modelo	Kalman	Real
1	X(m)	2,887	2,45	2,648	2,66
	Y(m)	0,6214	0,449	0,611	0,61
	θ [°]	42,68	42,69	42,7	45
2	X(m)	2,978	2,572	2,748	2,78
	Y(m)	0,32	0,137	0,3	0,38
	θ [°]	45,97	46,32	45,980	45
3	X(m)	3,155	2,615	2,893	2,92
	Y(m)	-0,121	-0,089	-0,107	-0,2
	θ [°]	-3,85	-3,71	-3,83	-10
4	X(m)	3,162	2,616	2,9	2,88
	Y(m)	-0,095	-0,066	-0,081	-0,15
	θ [°]	-3,08	-3,29	-3,06	-5

En la Tabla 2 se muestran los errores obtenidos tanto con la odometría (A), con el modelo (B) así como con el Filtro de Kalman (C) y con los valores reales medidos directamente.Se puede apreciar que luego del filtrado, los datos tienden a aproximarse al valor real, a pesar que los datos de odometría difieren considerablemente de los valores reales.

Tabla 2. Cálculo de errores en la pose del robot

Entornos	Variables	% Error A	% Error B	% Error C
1	X(m)	8,53	7,89	0,45
	Y(m)	1,87	26,39	0,16
	θ [°]	5,16	5,13	5,11
2	X(m)	7,12	7,48	1,15
	Y(m)	15,79	63,95	21,05
	θ [°]	2,16	2,93	2,18
3	X(m)	8,05	10,45	0,92
	Y(m)	39,50	55,50	46,56
	θ [°]	61,50	62,90	61,70
4	X(m)	9,79	9,17	0,69
	Y(m)	36,67	56,00	46,00
	θ [°]	38,40	34,20	38,80

La Fig. 11.muestra la posición del robot en cada iteración del algoritmo (5 segundos), detallando tanto los valores de desplazamiento obtenidos directamente por la odometría, como por el Filtro de Kalman.Se grafican también los valores reales medidos in situ.

Se puede observar que si el número de iteraciones continúa creciendo, la línea de tendencia de la odometría difiere demasiado de la línea de tendencia de los valores reales, mientras que con el Filtro de Kalman se tiene un mejor resultado. Basados en la gráfica se puede afirmar que luego de procesar los valores dados por el Filtro de Kalman, el algoritmo implementado corrige la odometría y mejora la localización, pero debido a condiciones propias del robot estos resultados no son los ideales. Este comportamiento es justificado debido a que existen variables que no han

sidoconsideradas y que influyen en el sistema como por ejemplo: el deslizamiento de las ruedas, el tipo de piso, nivel de baterías del robot, deterioro de sensores del robot, etc.

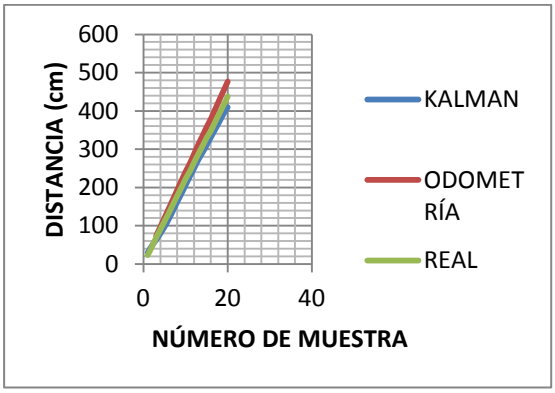


Figura 11. Comparación entre odometría, datos filtrados y datos reales; tomados cada 5 segundos

5. CONCLUSIONES

La calibración de las matrices de covarianza en el algoritmo del Filtro de Kalman es un factor que genera una sensibilidad muy alta en el comportamiento del Filtro, razón por la cual, al realizar un mínimo cambio en alguna de ellas los resultados se ven afectados drásticamente.

Si bien el sensor utilizado brinda buenos resultados, este presenta la limitación de verse afectado por la cantidad de luz del ambiente y los colores de los objetos circundantes.Para eliminar o reducir su grado de vulnerabilidad se pueden usar otros sensores como sonares, cámaras, etc.

La posibilidad de establecer una comunicación inalámbrica a través de la red *Wireless* propia de Robotino® es una gran ventaja ya que elimina el hecho de tener un medio físico tal como un cable.Sin embargo, esto incurre en una desventaja por el hecho de que esta comunicación se ve afectada por la contaminación WiFi, quesignifica que en presencia de otras redes inalámbricas el desempeño del robot no es óptimo, debido a que se puede ver afectada la emisión y recepción de datos.

Si se desea obtener datos de odometría más precisos es posibleañadir sensores exteroceptivos, como el sistema de navegación NorthStar [13].Sin embargo, se debe considerar que si bien los datos de la posición y orientación del robot pueden mejorar, los ambientes de trabajo se verían limitados debido a las características de funcionamiento del sensor, especialmente relacionado con el rango del alcance que es de 2.5m de altura y hasta 4m de distancia.

6. REFERENCIAS

[1] H. Durrant-Whyte, T. Bailey. “Simultanoeus Localization and Mapping (SLAM) Part I”, IEEE Robotics&Automation Magazine, September 2006.
[2] H. Durrant-Whyte, T. Bailey. “Simultanoeus Localization and Mapping (SLAM) Part II”, IEEE Robotics&Automation Magazine, September 2006.
[3] F. Auat. “ Localización y Reconstrucción Simultánea de Entornos por un Robot Móvil basada en la navegación Orientada a las Zonas de Máxima Incertidumbre”, Instituto de Automática, Universidad Nacional de San Juan, San Juan Argentina. Diciembre, 2008.
[4] F. Caramés, “ Técnicas de navegación de Robots basadas en Sistemas de medición por Láser”, Departamento de Informática y Automática, Universidad de Salamanca, Salamanca, España. Septiembre, 2007.
[5] Scanning Laser Range Finder URG-04LX-UG01 (Simple-URG) Specification, (2009, Agu). Available:http://www.hokuyo-aut.jp/02 sensor/07scanner/download/products/urg-04lx-ug01/data/URG-04LX_ UG01_spec_en.pdf.
[6] FESTO, © Festo Didactic GmbH & Co. KG • 544305 Manual, Denkendorf, 2007.
[7] Begoña F. Aguado, Janelcy A. Castaño, Eduardo Z. Casanova, Jaime G. García-Bermejo., “Diseño y Simulación de un Filtro Kalman para un Robot Móvil”, XXV Jornadas de Automática, Ciudad Real., Septiembre, 2004.
[8] G. Welch, G. Bishop., “An Introduction to the Kalman Filter”, Department of computer Science, University of North Carolina, Chapel Hill, EEUU., July, 2006.
[9] Universidad Nacional de Quilmes, “ Procesamiento Global empleando la Transformada de Hough”, Ingeniería en Automatización y Control Industrial, Septiembre, 2005.
[10] AallborgUniversity, “Modelo Cinemático Omnidireccional”, Denmark, 2012. Disponible: http://www.control.aau.dk/~tb/wiki/index.php/Kinematics
[11] Wikipedia, “Transformada de Hough, Teoría e Introducción”, 2013. Disponible: http://es.wikipedia.org/wiki/Transformada_de_Hough; teoría e introducción de la transformada de Hough.
[12] Universidad de Jaén, “Detección de bordes en una imagen”, Práctica 3, Departamento de Ingeniería, Telecomunicación y Automática.
[13] FESTO, “NorthstarSensor 2”. Manual,Denkendorf, 2009. Disponible: www.festo-didactic.com

Fuentes de Generación de Energía Eléctrica Convencional y Renovable a Nivel Mundial

J. Rosero*; L. Garza**; L. Minchala**; D. Pozo*; L. Morales*

*Escuela Politécnica Nacional, Departamento de Automatizacion y Control Industrial

e-mail: {jorge.rosero;david.pozo;luis.moralesec}@epn.edu.ec

**Instituto Tecnológico y de Estudios Superiores de Monterrey

e-mail: {legarza;li.minchala.phd.mty}@itesm.mx

Resumen: Actualmente la energía eléctrica es generada en su gran mayoría, por fuentes que consumen combustibles fósiles y/o energía nuclear, con los consecuentes riesgos y daños al medio ambiente que esto representa; esto sumado al aumento en los precios internacionales del petróleo, su inevitable agotamiento, la demanda de grandes cantidades de energía y el envejecimiento de las redes actuales, han hecho que en los últimos años se de un importante impulso y desarrollo a la generación de electricidad basada en fuentes renovables de energía que ha permitido incrementar de manera exponencial su aporte a nivel mundial. En este documento se presenta un análisis de las principales fuentes de generación de energía eléctrica, tanto convencionales como renovables, que se basa en una recopilación de datos históricos y estadísticos de producción, consumo, crecimiento y reservas, los mismos que permiten identificar a los países y regiones líderes en cada área.

Palabras claves: Combustible Fósil, Fuentes Renovables, Red Eléctrica.

Abstract: Currently, electric power is generated in its vast majority, for sources that consume fossil fuels and/or nuclear energy, with consequent risks and environmental damage that represents, this added to the increase in international oil prices, the inevitable exhaustion, the demand for large amounts of energy and aging current networks have made in the last years a significant development momentum and the generation of electricity from renewable sources of energy which has increased its contribution exponentially worldwide. This paper presents an analysis of the main sources of electric power generation, both conventional and renewable, which is based on a compilation of historical and statistical data on production, consumption, and reserves growth, enabling them to identify the countries and regions leaders in each area.

Keywords: Fossil Fuels, Renovable Sources, Electric Grid.

1. INTRODUCCIÓN

Un sistema eléctrico de potencia típico, está compuesto por tres grandes subsistemas: Sistema de Generación, Sistema de Distribución y Sistema de Consumidores; la red actual es fundamentalmente jerárquica y unidireccional, como se aprecia en la Figura 1, con bajos niveles de eficiencia (convierte únicamente una tercera parte del combustible consumido en electricidad) y grandes pérdidas en las líneas de transmisión (casi el 8 % de su generación), mientras que un alto porcentaje de esta potencia es usada durante pequeños intervalos de tiempo, solamente para cubrir los picos de demanda (20 % de su generación) [1].

El sistema de generación eléctrico convencional está basado en combustibles fósiles, energía nuclear, gas natural y plantas hidroeléctricas; en los últimos años una pequeña cantidad de energía renovable, proveniente de turbinas eólicas, paneles solares, plantas geotérmicas, biocombustibles, biomasa y otras, se han sumando al sistema de generación renovable. Según el reporte de la U.S. Energy Information Administration (EIA) [3], la generación eléctrica en Estados Unidos es 69 % basada

J. Rosero, D. Pozo y L. Morales están en el Departamento de Automatización y Control Industrial de la Escuela Politecnica Nacional en Quito Ecuador. L. Garza está en el Departamento de Mecatrónica y Automatización del Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Monterrey en México. L. Minchala es estudiante de Doctorado en el Instituto Tecnológico y de Estudios Superiores de Monterrey Campus Monterrey en México.

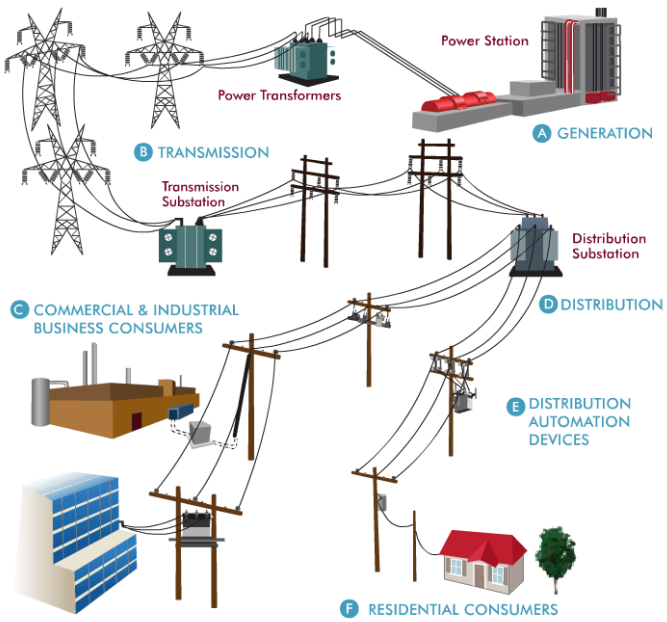


Figura 1. Sistema Eléctrico de Potencia en la actualidad [2]

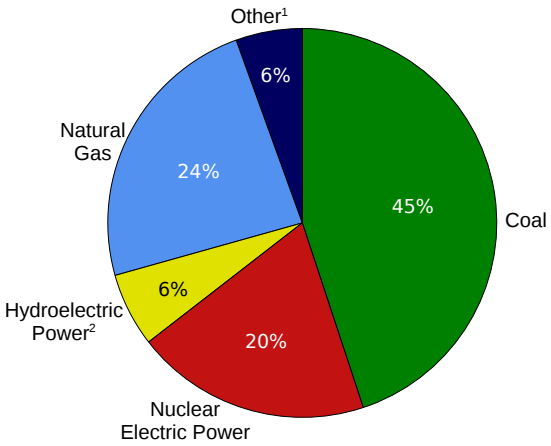


Figura 2. Generación Eléctrica de Estados Unidos en el 2010 [3].

en combustible fósil, 20 % por energía nuclear, 6 % por energía hidroeléctrica y apenas un 6 % por otras fuentes de energía, entre las que se encuentran las renovables, como se aprecia en la Figura 2. Por otro lado, de acuerdo al E.I.A. Country Analysis Brief [4], la generación eléctrica en China es aproximadamente 82 % basada en en combustible fósil, 16 % del sector hidroeléctrico y 2 % gracias a otras fuentes. Además, el E.I.A. International Energy Outlook 2011 (IEO2011) [5], proyecta que la generación de electricidad neta mundial se incrementará un 84 % en el 2035 con un uso predominante de carbón y gas natural y un importante desarrollo del sector hidroeléctrico y las energías renovables, como se aprecia en el Figura 3.

El combustible fósil es la principal fuente para generar electricidad, especialmente para China, que es el más grande productor y consumidor de carbón, mientras que Estados Unidos es el mayor productor y consumidor de petróleo [4], provocando que más del 50 % de las emisiones de CO_2 sean causadas por estos dos países y la Unión Europea [6], convirtiendo a la generación y transporte de electricidad como una de las mayores causas de contaminación ambiental.

Sin embargo, más del 82 % del incremento en generación de energía renovable ha sido en el sector hidroeléctrico y eólico, siendo China en el 2009, el mayor productor de energía hidroeléctrica y que actualmente construye la central hidroléctrica más grande del mundo, "Las tres gargantas"sobre el río Yangtsé, con una potencia instalada de 24GW, es decir, casi el doble que la central hidroeléctrica de Itaipú en Brasil (14 GW) [5]. Además, este país el próximo año contará con un nuevo record mundial, cuando entre en operación la tercera central hidroeléctrica más grande del mundo, Xiloudu con una potencia instalada de 13.86 GW [7].

Por otro lado, el Sistema de Distribución se realiza con ayuda de tres complejos subsistemas, como son: el de transmisión, el de subestaciones y el de despacho, los cuales fueron diseñados y construidos para *llevar* la electricidad desde las centrales de generación hacia los usuarios finales como se ilustra en la Figura 1, sin embargo éstos son sometidos a altos niveles de stress debido a sus condiciones de operación, como son: las variaciones bruscas de carga, presencia de armónicos y fenómenos naturales, capaces de causar graves daños sobre una red cada vez más deteriorada y sobrecargada, que la convierten en un sistema altamente vulnerable hoy en día.

El Sistema de Consumidores, según proyecciones hechas por la International Energy Agency en el Wolrd Energy Outlook 2012 [8], muestra que la población mundial crecerá de 6.8 billones de personas en el 2010 a 8.6 billones en el 2035, teniendo a futuro 1.7 billones de consumidores adicionales, principalmente en Asia y África, donde la población de la India superará la de China antes del año 2025, y para el 2035 alcanzará los 1.5 billones de habitantes, se conforma por tres grandes grupos claramente identificados: los consumidores residenciales, industriales y comerciales, los mismos que están cambiando, tanto en sus hábitos de consumo, como en el tipo de carga eléctrica que utilizan, siendo éstas en su gran mayoría formadas por dispositivos electrónicos de conmutación, los cuales inyectan grandes

cantidades de armónicos, que sumados causan graves perturbaciones en la red pública y que se traducen como pérdidas debido a los bajos factores de potencia.

Además, proyecciones hechas por la EIA en el IEO2011 [5], muestran que el consumo de energía a nivel mundial aumentará un 53 % para el 2035, donde China y la India serán los que encabecen dicho incremento, con un crecimiento promedio anual del 2.3 %. Hoy en día estos países lideran el crecimiento económico mundial y la demanda de energía, juntos consumieron alrededor del 10 % del total de la energía mundial en 1990 y un 21 % en 2008, mientras que China considera que su consumo ha sobrepasado al de Estados Unidos por primera vez en su historia. Se estima que estas dos naciones consumirán 38 % más energía que el resto de países en el 2020 y 67 % más en el 2035. En la Figura 4, las proyecciones muestran que el consumo de energía combinada de China e India serán el doble que la de Estados Unidos en el 2035, totalizando un 31 % a nivel mundial.

Finalmente, se proyecta que los actuales y futuros consumidores demandarán cada vez más una gran cantidad de energía, que muchas veces es mal utilizada o desperdiciada debido entre otras cosas, a la mala administración del consumo de energía, convirtiéndose en actores pasivos frente a lo que sucede en los niveles superiores de distribución, sin tener la posibilidad de contribuir a reducir los picos de consumo de energía aunque ellos son los responsables directos de los mismos [9]. Para satisfacer estos requerimientos y retos, muchos países, organizaciones no gubernamentales y el sector privado están trabajando incansablemente por desarrollar e implementar un nuevo concepto del sistema eléctrico actual, denominado redes inteligentes o smart grids, las cuales se convertirán en la siguiente revolución tecnológica de la nueva era.

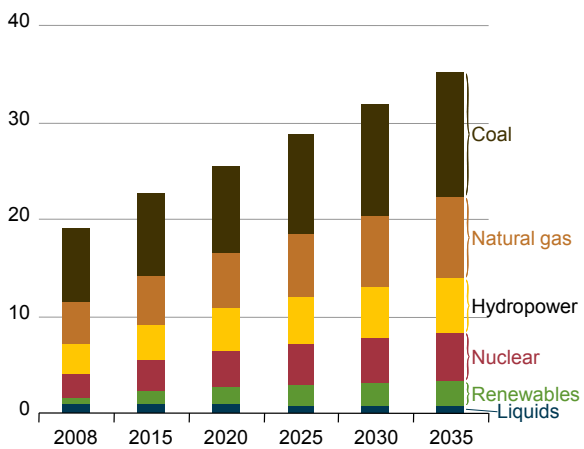


Figura 3. Generación Eléctrica Mundial Neta 2008-2035 (Trillones KWh) [5].

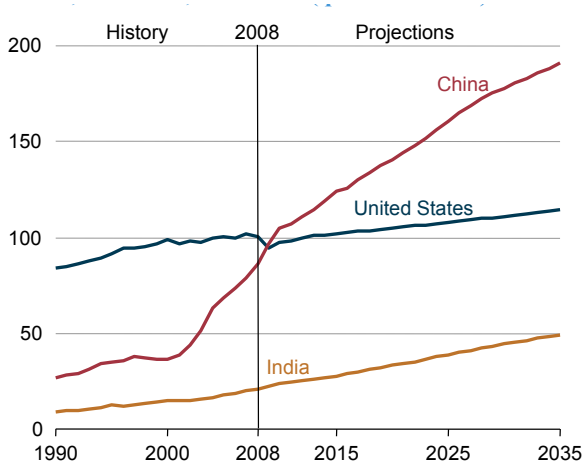


Figura 4. Consumo de Energía en EUA, China e India 1990-2035 (Quadrillones de BTU) [5].

2. FUENTES DE ENERGÍA

2.1 *Carbón*

De consistencia sólida, generalmente de color negro o café, con un 50 % en peso o 70 % en volumen de materia carbonosa, producida por la compactación y endurecimiento de residuos vegetales que vivieron y murieron hace aproximadamente 400 millones de años, el carbón es actualmente la fuente de energía más abundante sobre el planeta y sin lugar a duda, como el combustible fósil de mayor historia e importancia en el desarrollo de la humanidad, el mismo que ha sido usado desde épocas prehistóricas como fuente de calor y en la actualidad principalmente como combustible para la generación de energía eléctrica [10, 11].

La llegada de la revolución industrial jugó un papel muy importante en el desarrollo y consumo de éste combustible, en 1769 James Watt patentó la máquina a vapor, un revolucionario invento que usaba carbón para producir vapor y con él generar movimiento mecánico, que hasta esa fecha era realizado por animales y/o personas [12]. En 1882 Thomas Alba Edison construyó la primera central de generación termo-eléctrica usando carbón para proveer de electricidad a algunos residentes de la ciudad de Nueva York. En 1901, la *General Electric Company* construyó la primera planta termo-eléctrica de corriente alterna en Ehrenfeld, Pennsylvania, para la *Webster Coal and Coke Company*, eliminando los problemas de transmisión a larga distancia que en ese tiempo presentaba la corriente continua [13].

Entre 1970 y 1990 la producción de carbón en Estados Unidos prácticamente se duplicó, pasando de 520 millones a casi 1.1 billones de toneladas, cifra en la cual

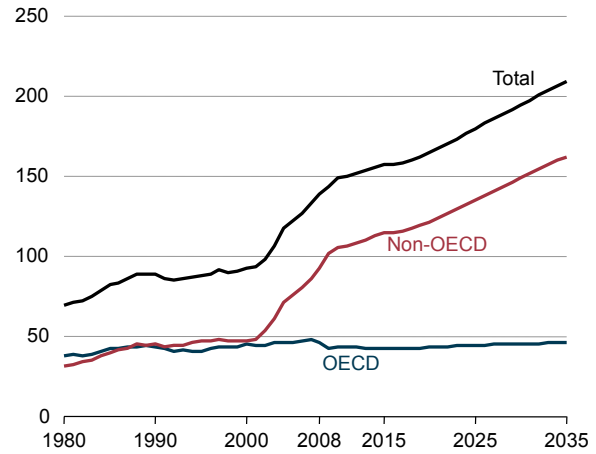


Figura 5. Consumo mundial de Carbón, 1980-2035 (QBTU) [5]

se mantuvo hasta la crisis mundial en 2009 donde se presento un ligero descenso [3]. De acuerdo al IEO2011, actualmente el productor y consumidor de carbón más grande del mundo es China, con una producción en 2010 que bordea los 70.5 quadrillones de BTU (QBTU) (British Thermal Unit), seguido de lejos por Estados Unidos con 22.6 QBTU; según sus predicciones el consumo mundial de carbón se incrementará en un 50 %, de 139 QBTU en 2008 a 209 QBTU en 2035, como se aprecia en la Figura 5, siendo China e India, países que no pertenecen al OECD (Organización para la Cooperacion y el Desarrollo Económico), quienes marquen dicho crecimiento [5].

Según el World Coal Association [14], el mercado más grande para el carbón actualmente es Asia, con una demanda que supera el 65 % del consumo mundial, especialmente en China donde el aumento en la demanda de energía del sector industrial y eléctrico es cubierto con este combustible; solo en 2009, este país consumió aproximadamente 3.5 billones de toneladas de este mineral, es decir, más del 46 % del consumo mundial, con un incremento de 180 % respecto al año 2000 [4]. Los mayores consumidores de carbón en el mundo son China, Estados Unidos, India, Rusia y Japón que en conjunto suman un 77 % de la producción mundial [5].

Según el Global Energy Statistical Yearbook 2011, el crecimiento en la producción de carbón en 2010 se mantuvo a una tasa de 5.2 %, donde China alcanzo un crecimiento de 8.9 % principalmente para cubrir su demanda interna. En el mismo año India creció al 3 %, más lentamente respecto al 2009 (8 %), mientras que Estados Unidos el segundo mayor productor de carbón, con un mercado mundial del 15 %, tuvo un ligero decremento de 0.3 %; Indonesia por su parte, creció a un ritmo de 7.3 % (10 % en 2009) y Rusia a un nivel de 5.4 % [15].

El carbón, a nivel mundial, es usado principalmente como combustible para la generación de energía eléctrica, éste representa el 40 % del combustible destinado a dicho propósito, sin embargo se estima que para el año 2035 su aporte disminuirá al 37 %. Otros usos del carbón están en el sector industrial, en la producción de acero, procesamiento de cemento, fabricación de papel, industria farmacéutica, entre otros [14]. Se estima que el uso de carbón para la generación eléctrica en China aumente de 28.7 QBTU en 2008 a 63.4 QBTU en 2035, esto significa a una tasa de crecimiento anual de 3.0 % en comparación con Estados Unidos que crecerá a un ritmo de 0.2 % anualmente en este sector [5]. China destina aproximadamente el 80 % de carbón para generar un 80 % de energía eléctrica, es decir 2756.8 Billones de Kilowatts-hora (BKW-h) [4]. Estados Unidos por su parte, destina cerca del 92 % de carbón para generar un 45 % de energía eléctrica, lo que representa 1850.7 BKW-h [3].

En cuanto a las reservas, se estima que mundialmente existe un total de 948 billones de toneladas, lo que representa 126.3 años de disponibilidad de este combustible, con Estados Unidos a la cabeza con 260.6 billones, seguido por Rusia con 173.1 y China con 126.2, es decir que estos tres países poseen el 59 % de las reservas mundiales de carbón [10][11]. En la actualidad el carbón representa la principal fuente de energía para el mundo, China tiene una dependencia a este combustible en un 71 % [4], mientras que India un 41 % [16], Estados Unidos un 21 % [3] y Rusia un 16 % [17].

2.2 *Gas Natural*

Hidrocarburo gaseoso altamente inflamable que está compuesto principalmente de metano y etano, incoloro e inodoro y que normalmente se encuentra disuelto en petróleo a muy alta presión denominándolo gas asociado o húmedo; otras veces se lo puede encontrar sólo, es decir, sin estar en contacto con ningún líquido, llamándolo gas no asociado o seco. El gas natural es un combustible fósil que se formó por la descomposición de plantas terrestres y restos marinos hace mas de 300 millones de años y actualmente representa la segunda fuente de energía más abundante del planeta y la de mayor crecimiento [18].

Fue en 1821 en Fredonia, Nueva York cuando William A. Hart taladró un pozo de 27 pies de profundidad en un esfuerzo por obtener un mayor flujo de gas proveniente de una filtración superficial; éste hecho es considerado como el primer pozo intencionalmente taladrado para obtener gas natural en América. Por mucho tiempo el

gas fue usado exclusivamente como combustible para iluminación, hasta que en el año 1885, Robert Bunsen inventó el *Quemador de Bunsen*, el cual combinaba gas con aire, permitiendo usarlo cómo calefacción doméstica y como fuente de calor para cocinar [19].

Una de las primeras grandes tuberías de gas fue construida en 1891, con una longitud de 120 millas (193Km), desde los pozos ubicados en Indiana hasta la ciudad de Chicago; sin embargo, no fue sino hasta después de la segunda guerra mundial cuando los avances en industria metalúrgica, nuevas técnicas de soldadura y materiales fueron desarrollados, permitiendo que se construyeran miles de millas de tuberías en Estados Unidos, llevando este combustible a nuevos mercados y convirtiendo hoy en día a esta red en una de las más grandes del mundo, a tal punto que se podría extender hasta la luna, ida y vuelta, dos veces [20].

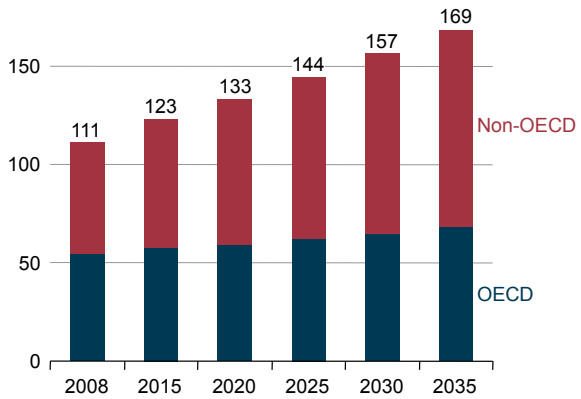


Figura 6. Consumo mundial de Gas Natural, 2008-2035 (trillones de pies cúbicos) [5]

El gas natural es el combustible fósil de más rápido crecimiento a nivel mundial, con un ritmo anual de consumo promedio del 1.6 % desde el 2008 hasta el 2035, como se aprecia en la Figura 6. La demanda en los países no-OECD (Organization for Economic Co-operation and Development) se incrementa casi tres veces más rápido que en los países OECD, con un crecimiento promedio del 2.3 % y 0.8 %, respectivamente, proyectado hasta el 2035. A nivel mundial los países no-OECD cuentan con alrededor del 76 % del incremento total de consumo de gas natural, mientras tanto los países OECD tienen una participación en el incremento mundial que va del 51 % en el 2008 al 59 % en el 2035.

El sector eléctrico e industrial juntos suman un incremento en la estimación de consumo de este combustible en un 87 % para el 2035. Irán y Qatar son los países que mayor producción de gas natural tienen actualmente, llegando a una quinta parte de la producción mundial.

El gas natural es el combustible fósil de combustión más limpio, el mismo que produce un 50 % menos CO2 que el carbón y un 30 % menos que el petróleo; es la fuente de energía primaria más eficiente y más barata del mercado, razón por la cual es usado principalmente como calefacción en hogares y oficinas, y como fuente de propulsión para maquinaria industrial y vehículos de transporte de carga; se estima que en pocos años también podrá ser usado en vehiculos de transporte de pasajeros, desplazando así al petróleo que hasta ahora es usado para dichas actividades [21]. Se estima que el gas natural para aplicaciones industriales contará con un crecimiento promedio del 1.7 % hasta el 2035, mientras que en el sector eléctrico crece a una tasa promedio anual del 2.0 %. Tanto el sector industrial como el electrico, juntos alcanzan el 87 % de consumo de gas natural proyectado.

Por otro lado, las reservas mundiales han crecido un 50 % en los últimos 20 años, superando las reservas de petróleo en el mismo período, con un incremento anual de 3.1 % desde 1980. La región que más crecimiento ha tenido en sus reservas ha sido el Medio Oriente, con Irán, Qatar, Arabia y Abu Dahbi a la cabeza. Tres países poseen más del 50 % Rusia con 25.2 %, Irán 15.7 %, Qatar 13.4 %, mientras que un total de 20 países poseen el 90.9 % de las reservas mundiales [5].

2.3 *Petróleo*

Compleja mezcla de hidrocarburos que se encuentran en forma líquida, gaseosa o sólida. El petróleo se refiere normalmente a la forma líquida o también llamado *crudo*, aunque técnicamente este formado por gas y otros elementos viscosos principalmente azufre. Las fases líquidas y gaseosas, actualmente constituyen la forma de energía fósil más importante del planeta. Su formación es igual que la del carbón y por ende también se lo considera como una fuente de energía no renovable. Los antiguos Sumerios, Asirianos y Babilonios, usaron éste hidrocarburo hace más de 5000 años A.C. en diversas aplicaciones, mientras que los Egipcios lo usaron con fines medicinales y los Árabes y Persas como armas de guerra [22].

En 1858 el canadiense Miller Williams (1818-1890) perforó el primer pozo petrolero en América y construyó una sencilla refinería para su procesamiento, a la cual llamo *J.M Williams Company* siendo la primera empresa petrolera en América y que posteriormente en 1860 se conocería como la *Canadian Oil Company* [23]. Sin embargo, un año más tarde, el 27 de Agosto de 1859 el estadounidense Edwin L. Drake (1819-1880) perforó

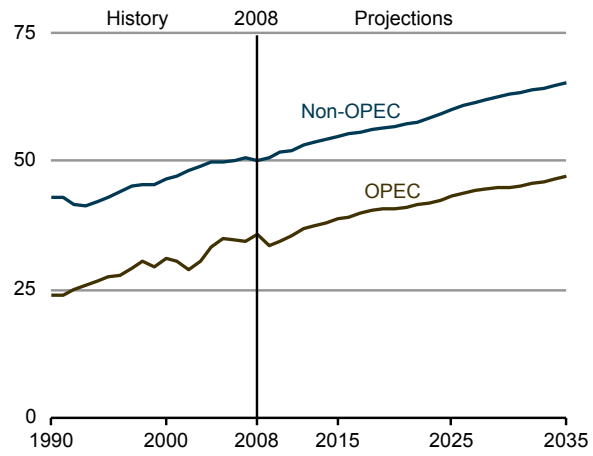


Figura 7. Producción mundial de combustible fósil líquido 1990-2035 (millones de barriles por día) [5]

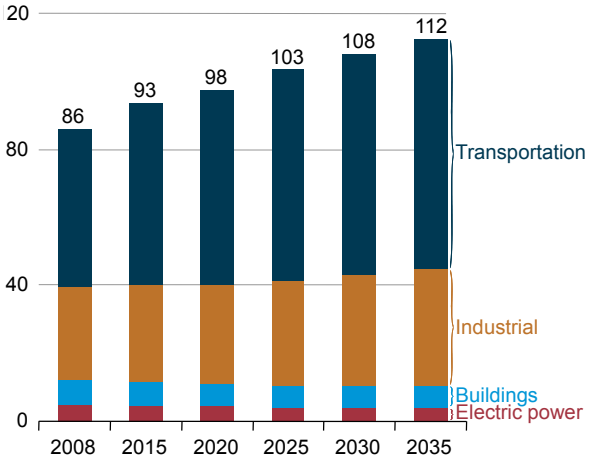


Figura 8. Consumo de combustible líquido por sector a nivel mundial (millones de barriles por día)[27]

un pozo petrolero de 69 pies de profundidad con una innovadora y revolucionaria técnica (que no la patentó) en Titusville, Pennsylvania siendo éste el punto de partida del *oro negro* en el continente Americano [24, 25].

En el 2010 la demanda de petróleo alcanzó los 86.7 millones de barriles por día y a este ritmo, para el 2035 la demanda será de 99 millones de barriles por día. Sin embargo, si se aplican políticas de reducción de CO2 es posible disminuir el consumo hasta 78 millones de barriles por día. En cuanto al precio, se estima que su valor aumente hasta 120 USD/barril en el 2035. Según el IEO2011 [5], más del 75 % del incremento en el consumo de combustible líquidos será destinado a los países Asiáticos y del Medio Oriente, gracias a su próspero crecimiento económico principalmente. Para cubrir esta demanda, la producción de combustibles líquidos deberá crecer a un ritmo de 26.6 billones de barriles por día, desde el 2008 hasta el 2035.

La cantidad de petróleo descubierto a nivel mundial en la década pasada bordea los 14 billones de barriles en promedio [26], mientras que la producción mundial de combustible fósil líquido estimada para el año 2035 excede con 26.6 millones de barriles por día a la del año 2008, cuyo aumento se refleja en ambos sectores, tanto OPEC (Organization of the Petroleum Exporting Countries) como no-OPEC. El primero se estima produzca 46.9 millones de barriles por día en el 2035, mientras que el segundo 65.3 millones de barriles por día, para el mismo año [5], como se aprecia en la Figura 7. Sin embargo, estas estimaciones pueden variar de manera drástica debido a las políticas de reducción de gases de efecto invernadero así como al incremento en los precios internacionales de petróleo.

Por otro lado, los sectores de mayor consumo de combustible líquido a nivel mundial, son el de la transportación y el industrial, como se aprecia en la Figura 8, siendo el primero el más importante con un 54 % del total de combustibles líquidos en el 2008, y se estima que para el 2035 incremente a 60 %. Sin embargo, la única fuente de energía para la generación de electricidad que no crece a nivel mundial es el petróleo, básicamente debido al incremento de su precio (se estima llegue a los 120 USD/barril en el 2035), es así que a nivel mundial la generación eléctrica basada en este combustible decrece un promedio de 0.9 % al año, de 1.0 trillones de Kilowatts-hora en el 2008 a 0.8 trillones de Kilowatts-hora en el 2035, siendo reemplazado por combustibles más económicos y menos contaminantes como el gas natural y la energía nuclear.

Finalmente, se estima que las reservas mundiales de petróleo están entre los 1470 [27] y 1652.6 billones de barriles [28], siendo Venezuela el país con mayor crecimiento en el 2011, mientras que la mayor cantidad de las reservas provadas se encuentran en el Medio Oriente. Aproximadamente, el 79 % de las reservas probadas de petróleo se encuentran en poder de 8 países, de los cuales solo Canada y Rusia no estan dentro del grupo OPEC (Organization of the Petroleum Exporting Countries). El país que mayor reservas posee es Arabia Saudita que cuenta con 260.1 billones de barriles, seguido de Venezuela con 211.2 billones de barriles y Canada con 175.2 billones de barriles, lo que a nivel mundial representa el 17.68 %, 14.35 % y 11.91 %, respectivamente. Ecuador aparece con una reserva probada de 6.5 billones de barriles lo que representa el 0.44 % a nivel mundial.

2.4 *Nuclear*

Las plantas nucleares son plantas de generación termo-eléctrica, las mismas que permiten hervir el agua hasta convertirla en vapor, que luego es usado para producir movimiento mecánico rotacional sobre las turbinas de generación eléctrica, del mismo modo que lo hacen las plantas de carbón, petróleo, geotérmicas, entre otras. La principal diferencia es que las plantas nucleares, en lugar de usar combustible fósil para este propósito, usan Uranio como principal fuente de energía. El uranio consiste en pequeñas pastillas de cerámica duras que son empaquetadas en largos tubos verticales e insertados dentro del reactor [29].

Esta tecnología basa su principio en el aprovechamiento de la energía liberada gracias a un proceso físico llamado fisión nuclear, que consiste en la división de los núcleos de átomos de Uranio en otros más pequeños y además en otras partículas como, neutrones, fotones (generalmente rayos gamma), rayos alfa y beta, los mismos que provocan que otros átomos se dividan, obteniendo un fenómeno llamado *reacción en cadena*. Éstas particulas contienen grandes cantidades de energía calorífica y cinética que calientan la materia circundante.

Ésta se desarrolló en 1940 en la Segunda Guerra Mundial, donde se usó con fines militares en la construcción de bombas de destrucción masiva, sin embargo luego se buscaron formas de aplicarla para la generación de energía eléctrica.

La primera planta de energía nuclear inició su operación a principios del año 1950 [30] y actualmente es una de las principales fuentes para la generación de energía eléctrica, proveyendo alrededor del 14 % a nivel mundial y un 21 % en los países OECD [29]. Los primeros países en usar esta nueva fuente de energía para la generación de electricidad fue la Ex-Unión Soviética en 1954, el Reino Unido en 1956, los Estados Unidos 1957 y Francia en 1963. Posteriormente otros países, como Bélgica, Canada, Alemania, Italia, Japón y Suecia les siguieron sus pasos.

En total 13 países son capaces de generar al menos la cuarta parte de la producción mundial, con datos de Agosto 2012 y como se ilustra en la Tabla 1, a nivel mundial 435 reactores para la generación de electricidad están operando, siendo Estados Unidos el país con mayor número de reactores en operación, que en total suman 104 y una capacidad de generación de 101 465 MW, que para este país representa alrededor del 20 % de su capacidad total. Además, el segundo país con mayor capacidad de generación es Francia, el mismo que cuenta con 58 reactores nucleares que entregan una potencia de 63 130 MW, que para este país representa el 74.8 % de su capacidad de generacón total [30].

Tabla 1. Capacidad de Generación Nuclear en Operación [29].

	No. de Reactores	Capacidad (MW)
Estados Unidos	104	101 465
Francia	58	63 130
Japón	50	44 215
Rusia	33	23 643
República de Corea	21	18 751
Ukrania	15	13 107
Canada	18	12 604
Alemania	9	12 068
China	16	11 816
Reino Unido	18	9 953
Suecia	10	9 326
España	8	7 567
Bélgica	7	5 927
Taipei	6	5 018
India	20	4 391
República Checa	6	3 766
Suiza	5	3 263
Finlandia	4	2 736
Bulgaria	2	1 906
Húngria	4	1 889
Brasil	2	1 884
Sudáfrica	2	1 830
República Eslovaca	4	1 816
México	2	1 300
Rumania	2	1 300
Argentina	2	935
Irán	1	915
Pakistan	3	725
Eslovenia	1	688
Holanda	1	482
Armenia	1	375
Total	435	368.791

Por otro lado, 66 nuevas plantas están bajo construcción en 14 países, como se ilustra en la Tabla 2, siendo China el país que mayor inversión realiza en esta área, buscando

Tabla 2. Capacidad de Generación Nuclear en Operación [29].

	No. de Reactores	Capacidad (MW)
China	26	26 620
Rusia	10	8 188
República de Corea	5	5 560
India	7	4 824
Japón	2	2 650
Taipei	2	2 600
Ukrania	2	1 900
Francia	1	1 600
Finlandia	1	1 600
Brasil	1	1 245
Estados Unidos	1	1 165
República Eslovaca	2	782
Argentina	1	692
Pakistan	2	630
Total	63	60 056

casi triplicar su capacidad de generación de 11 816 MW a 26 620 MW, con la construcción de 26 nuevos reactores nucleares. En el 2010, las plantas nucleares proveyeron el 13.5 % de la energía eléctrica mundial.

Entre otras de las aplicaciones que normalmente se dá a la energía nuclear se encuentran productos de uso cotidiano, como por ejemplo, relojes, lámparas fluorescentes, vidrios, alimentos, cerámicas, entre otros, los mismos que contienen pequeñas cantidades de materiales radioactivos que sirven para mejorar sus cualidades físicas o químicas. En agricultura, la radiación producida por los rayos gamma es usada para eliminar cierto tipo de bacterias y hongos, que provocan enfermedades e intoxicaciones. Mientras tanto en la industria, la mayor parte de los procesos relacionados con calibración e inspeción de materiales usa radiación para medir el espesor de partes, fallas en tuberías, fugas de líquidos, comportamiento de piezas en operación, entre otras. Por otro lado, en la medicina, el uso de pequeñas cantidades de radiación sobre el cuerpo humano es usado para realizar examenes de diagnóstico de enfermedades, incluso en algunos tipos de cáncer la exposición del paciente a terapias de radiación es muy común en la actualidad. [31]

2.5 *Renovable*

Con el término *energía renovable* se conocen a todas aquellas fuentes que pueden ser usadas para producir energía, cuyo ciclo y combustible es inagotable, las mismas que provocan una reducida contaminación y daño ambiental, entre las más comunes se encuentran, la solar, eólica, geotérmica, biocombustibles, biomasa y tidal [32].

Sin lugar a dudas, las energías renovables son una de las fuentes de generación de electricidad de mayor incremento en los últimos años, con una tasa de crecimiento promedio anual del 3.1 % y una producción, que en el 2008 era del 19 % y que en el 2035 se estima llegue al 23 %. La energía eólica muestra un salto significativo en su producción, pasando de 18 GW a finales del 2000 a 121 GW a finales del 2008 y con una tendencia creciente, estimando que en el 2035 represente el 27 % de toda la producción de energía renovable a nivel mundial [5]. Sin embargo, a pesar que este tipo de energía es amigable con el medio ambiente, aún no puede competir económicamente con los combustibles fósiles, debido a sus altos costos de implementación derivados de la tecnología de punta involucrada y su baja producción de energía promedio, siendo necesario que existan políticas gubernamentales que permitan subsidiar el costo real de generación a cambio de la reducción en la emisión de gases de efecto invernadero.

2.6 *Hídrica*

Este tipo de energía basa su principio de funcionamiento en el aprovechamiento de inmensas cantidades de agua, proveniente en la mayoría de los casos de ríos caudalosos, que son almacenadas de manera temporal en grandes reservorios llamados represas o diques, las mismas que se ubican en los lechos de los ríos aguas arriba. Esta agua almacenada es conducida por medio de tuneles o ductos aguas abajo, logrando formar una diferencia de altura bastante considerable que permite obtener energía potencial, la misma que será transformada en electricidad. El caudal que se vierte de manera directa sobre las paletas de las turbinas de generación, hace que éstas giren dando lugar al movimiento mecánico rotacional que se transforma en energía eléctrica al interior de los generadores eléctricos instalados para el efecto [33].

La primera planta de generación hidroeléctrica del mundo fue la construida a finales del Siglo XIX en las Cataratas del Niagara en la frontera entre Estados Unidos y Canada, por el padre de la energía eléctrica alterna el científico europeo Nicola Tesla [34] [35] . En la actualidad la energía hidroeléctrica es la fuente renovable de generación eléctrica más grande del mundo, donde China lidera su producción seguido de Brasil, Canada, Estados Unidos y Rusia. En la Tabla 3 se resumen a los países lideres a nivel mundial en el sector hidroeléctrico. Sin

embargo, entre el 2005 y el 2010, este tipo de generación se ubica en el segundo lugar después de la energía eólica, en cuanto a crecimiento [36] debido en gran medida a que una central de generación requiere mayor tiempo en ser construida que un parque eólico.

Tabla 3. Países líderes de generación de energía eléctrica [37].

País	Parte del Total en el 2012 (%)
China	23.4
Brasil	11.4
Canada	10.4
Estados Unidos	7.6
Rusia	4.5
Noruega	4.9
Japón	2.2
Venezuela	2.2
Suecia	2.1

Una de las grandes y principales ventajas que poseen las centrales hidroeléctricas es su velocidad y capacidad de respuesta ante fluctuaciones de carga, que permiten cubrir los picos de demanda de manera rápida y confiable, producto de sus grandes reservas de agua que son consumidas en instantes cortos de tiempo, actuando como una especie de colchón energético, a diferencia de la energía solar y eólica, las mismas que tienen un perfil de generación bastante impredecible e inestable debido principalmente a su naturaleza y a la carencia de dispositivos o métodos de almacenamiento seguros.

En el 2011, los países OECD contaron con una generación de energía hidroeléctrica de alrededor de 1,458 TWh o el 13.5 % de su producción total, mientras que los países no-OECD tuvieron 2,097 TWh o el 20 % de su producción total [38].

2.7 *Eólica*

El sol calienta las masas de aire en la superficie de la tierra de manera irregular, provocando diferencias de temperaturas que generan a su vez movimientos de grandes masas de aire. Este flujo ha sido aprovechado desde épocas milenarias como fuente de energía mecánica rotacional en molinos y actualmente es una de las mayores fuentes de generación eléctrica renovable después de las hidroeléctricas. Utilizando principios aerodinámicos los molinos de viento, capturan la energía cinética del viento para mover sus aspas generando movimiento mecánico rotacional que luego es transformado en energía eléctrica en el interior de las turbinas. Comercialmente existen varios modelos de ellas, pero las más comunes son las de tres aspas con eje horizontal equipados con una caja de engranes para multiplicar la velocidad del rotor del generador y alcanzar una frecuencia de generación cercana a los 50/60 Hz.

Los tamaños típicos de estas turbinas van desde los 200KW hata los 750KW, con determinados rangos de velocidad de viento para su generación en cada caso. El rápido desarrollo de esta tecnología ha permitido reducir los costos de instalación, de $2.2/W en 1980 a $1/W en la actualidad, del mismo modo los costos de energía se han visto reducidos de $0.40/KWh a menos de $0.04-$0.06/KWh en áreas con buenas condiciones de viento [40]. A nivel mundial en el año 2000, más de 3900MW adicionales de energía eólica fueron instalados, con ventas mayores a $ 3.9 billones y una tasa de crecimiento del 35 % anual, tendencia que ha ido en claro incremento exponencial hasta la fecha en los Estados Unidos, como se aprecia en la Figura 9.

En el año 2012 se añadieron 44.8 GW de potencia eólica, registrando un record histórico y completando un total de 282.5 GW a nivel mundial. Actualmente, son 100 países quienes logran cubrir un 3 % de la demanda global, estando el 85 % de la generación en manos de 21 de ellos [41], siendo China el país que a finales del 2010, alcanzó una capacidad instalada de energía eólica de 44.73 GW; el 64 % de ésta energía proviene de 127 plantas que superan los 100MW de capacidad, cubriendo sóla el 1.28 % del consumo total, superando así a los Estados Unidos y convirtiendose en el líder a nivel mundial [42]. En la Tabla 5 se listan los países que acumulan la mayor cantidad de potencia eólica instalada a nivel mundial.

Alrededor del mundo, grandes esfuerzos por reducir la emisión de gases de efecto invernadero se están realizando en el área eólica, es así que en el año anterior, en seis países se sumaron mas de 1GW de potencia a la red, como se aprecia en la Tabla 4, mientras que países como Canadá, Suiza y México sumaron cada uno 600 MW de nueva potencia proveniente del aire, siendo éste último el que mayor crecimiento muestra en este campo, con un incremento de 113 %.

Las espectativas de crecimiento son bastante ambiciosas, como por ejemplo, la Unión Europea prevee contar con un 20 % de energía eólica para el año 2020, al igual que Australia; mientras que China espera instalar 100 GW para el 2015. Por su parte, la República de Corea busca reemplazar un 11 % de su consumo con energía eólica. Mientras tanto en latinoamérica, México lidera el sector

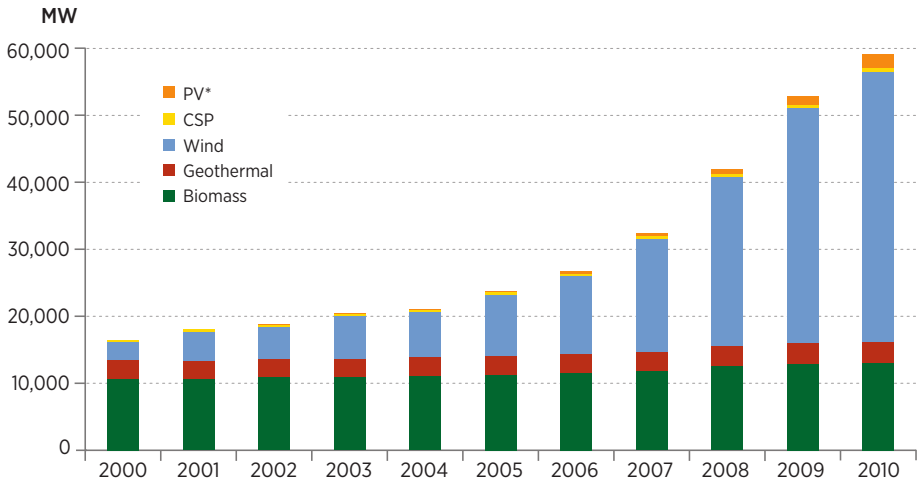


Figura 9. Capacidad de Generación Eléctrica Renovable de Estados Unidos [39].

Tabla 4. Incremento de Potencia Eólica [41].

País	Capacidad (GW)	Incremento (%)
Estados Unidos	13.13	28
China	12.96	21
Alemania	2.44	8
Inglaterra	1.82	28
Italia	1.27	18
España	1.11	5

eólico con un crecimiento estimado para el año 2020 de un 5 % del total, que representa 12GW.

2.8 *Solar*

La energía solar es sin duda, una de las formas de energía renovable más abundante sobre el planeta y es la causante de provocar muchas otras, como la eólica debido al calentamiento de las masas de aire. Se estima que la energía proveniente del sol está entre 1.5×10^{11} y 1.8×10^{11} MW, de los cuales la mitad alcanzan a ingresar a la superficie de la tierra y la otra mitad es reflejada por la atmósfera [40] [43]. Se calcula que logrando captar el 0.1 % de ésta energía y convirtiéndola en electricidad con un rendimiento del 10 %, se podría cubrir cuatro veces la demanda mundial de energía eléctrica; Incluso en 90 minutos, el sol es capaz de proveer toda la energía que el planeta requiere para un año [44]; por tal razón la energía solar es en magnitud, la fuente de energía con mayor futuro del planeta y la que mayor sustentabilidad puede ofrecer [45].

Existen dos clases principales de energía solar: Fotovoltáica y termosolar.

2.8.1 *Fotovoltáica*

Su funcionamiento se basa en que la luz solar esta compuesta por pequeñas partículas de energía llamadas fotones, los mismos que son absorbidos por los paneles solares. Los fotones excitan a los electrones que se encuentran en las celdas fotovoltaicas haciendo que estos empiecen a moverse y formen pequeños flujos de corriente.

Las celdas se colocan en arreglos para formar paneles solares de diferentes tamaños y por ende potencia eléctrica que luego es utilizada como una forma de energía. Estos paneles utilizan material semiconductor tales como el silicio para convertir la energía solar en energía eléctrica, los cuales no tienen partes móviles, no generan ruido, ni contaminación de ninguna clase, tienen un tiempo de vida útil de aproximadamente 20 años y requieren poco mantenimiento. Los costos relacionados a estos elementos han disminuido de \$50/W a inicios de 1980 a \$5/W hoy en día; del mismo modo el costo de la energía ha disminuido de \$0.90/KWh en 1980 a \$0.20/KWh en la actualidad. A pesar de su aún alto costo, en algunos países ésta tecnología es ampliamente usada para alimentar de energía a lugares remotos, equipos de telecomunicaciones, vehículos, entre otros [40].

En el período del 2000 al 2011, esta forma de energía

Tabla 5. Capacidad de Generación Eléctrica Eólica Instalada 2012 [41].

País	Capacidad (MW)
China	75 324
Estados Unidos	60 007
Alemania	31 315
España	22 785
India	18 421
Inglaterra	8 292
Italia	8 144
Francia	7 564
Canadá	6 201
Dinamarca	4 162
Suiza	3524
Japón	2 614
Australia	2 584
Brasil	2 508
Polonia	2 497
Holanda	2 431
Turquía	2 312
Rumania	1 905
Irlanda	1 827
Grecia	1 749
Austria	1 378
México	1 212

fue una de las de mayor crecimiento a nivel mundial, alcanzando a finales del 2011 una capacidad instalada de 65 GW, siendo Alemania e Italia los países que poseen la mitad de la capacidad mundial, seguidos por Japón, España, Estados Unidos y China [46].

2.8.2 *Termo-Solar*

Una nueva forma de producir energía derivada de la solar son los parques de generación termo-solar, que consisten en amplios campos dotados de paneles de alta reflectividad los cuales proyectan la energía capturada a una torre en la cual calientan un núcleo de sales fundidas a temperaturas del orden de los 900 °C y por medio de sistemas de transferencia de calor generan vapor el cual a su vez es usado para generación de energía eléctrica.

3. CONCLUSIONES

En este documento se presenta una descripción general de las diferentes fuentes de generación de energía eléctrica, donde se aprecia que los combustibles fósiles en la mayoria de los países son la principal fuente para la generación de electricidad y según diferentes proyecciones realizadas por organismos internacionales y empresas privadas, en los próximos años se mantendrá su uso aunque con una ligera tendencia a la baja, por lo que se puede preveer que el ritmo de contaminación ambiental no disminuirá en un futuro inmediato.

Por otro lado, se puede apreciar un claro esfuerzo internacional por apaliar en algo la cruda realidad del medio ambiente, por medio del uso de combustibles provenientes de fuentes renovables de energía, donde la eólica es la que mayor crecimiento a nivel mundial a tenido, seguido de la solar y la hidroeléctrica. Sin embargo, los números no son muy alentadores a nivel global, ya que menos del 10 % de la energía que necesita el planeta es suministrada por estas fuentes.

Muchos de los países analizados apuestan al uso de energías nucleares como principal fuente de generación de electricidad a pesar de los consecuentes riesgos que esto representa. Uno de éstos países es Francia cuyo porcentaje de generación está entre los más altos del mundo. Debido a los altos costos del petróleo y el carbón y a su inevitable agotamiento a mediano plazo, la energía nuclear bien utilizada podría convertirse en la principal fuente de generación, incluso por encima de las renovables que no tienen un terreno afianzado todavía.

Las energías renovables aún no logran superar a las fósiles y nucleares, principalmente por su alto costo de instalación y mantenimiento, a pesar de ello son las que mayor crecimiento están teniedo en los últimos años. Además, se espera que en el corto y mediano plazo puedan ser capaces de cubrir las necesidades básicas de consumo.

Finalmente, en el Ecuador el impulso a las fuentes de energía renovable en los ultimos años ha sido muy importante, tanto así que se estima que para el año 2017 el 100 % de la demanda de energía eléctrica será cubierta con este tipo de fuentes de generación, siendo las hidroeléctricas las que encabezaran dicho cambio. No obstante, en comparación a otros países su producción es aún incipiente y en algunos casos nula, como el uso de fuentes nucleares o geotérmicas. Es así, que se hace necesario un estudio completo y detallado de la matriz energética del

país, el mismo que será abordado en un próximo artículo de investigación.

REFERENCIAS

[1] H. Farhangi, "The path of the smart grid," *Power and Energy Magazine, IEEE*, vol. 8, no. 1, pp. 18–28, 2010.

[2] Smarth Texas Rethinking Energy. (2012, February 2012) Pathway to power. [Online]. Available: http://www.txuelectricdelivery.com/tech_reliable/pathwaytopower.aspx

[3] U.S. Energy Information Administration. (2011, March 20 2012) Annual Energy Review 2010. [Online]. Available: http://www.eia.gov/totalenergy/data/annual/pdf/aer.pdf

[4] Energy Information Administration. (2010, March 20 2012) Country Analysis Briefs, China. [Online]. Available: http://www.eia.gov/EMEU/cabs/China/pdf.pdf

[5] EIA. (2011, February 27 2012) International Energy Outlook 2011. [Online]. Available: http://www.eia.gov/FTPROOT/forecasting/0484(2011).pdf

[6] H. Rudnick, L. Barroso, D. Llarens, D. Watts, and R. Ferreira, "Flexible connections: Solutions and challenges for the integration of renewables in south america," *Power and Energy Magazine, IEEE*, vol. 10, no. 2, pp. 24–36, 2012.

[7] S. Patel. (2013, October 2013) China's second and third-largest mega-dams mark operational milestones. [Online]. Available: http://goo.gl/u4nKhw

[8] Maria van der Hoeven, "World Energy Outlook 2012," International Energy Agency, France, Tech. Rep., Jan. 2012.

[9] European-Commission. (2006, March 21 2012) Vision and strategy for europe's electricity networks of the future.

[10] U.S. Department of Energy. (2011, March 23 2012) A Brief Histoy of Coal Use. [Online]. Available: http://fossil.energy.gov/education/energylessons/coal/coal_history.html

[11] O. C. Kopp, "Coal," *Britanica Encyclopedia Online Academic Version*, March, 2012 2012. [Online]. Available: http://0-www.britannica.com.millenium.itesm.mx/EBchecked/topic/122863/coal

[12] U.S. Energy Information Administration. (2011, March 23 2012) Annual energy review 2010. [Online]. Available: http://www.eia.gov/kids/energy.cfm?page=tl_coal

[13] National Energy Technology Laboratory. (2010, March 23 2012) History of U.S. Coal Use. [Online]. Available: http://www.netl.doe.gov/keyissues/historyofcoaluse.html

[14] World Coal Association. (2012, April 02 2012) Coal, Uses of Coal. [Online]. Available: http://www.worldcoal.org/coal/uses-of-coal/

[15] Global Energy Statistical Yearbook 2011. (2010, April 02 2012) Coal and Lignite Production. [Online]. Available: http://yearbook.enerdata.net/2010-energy-consumption-data.html#/coal-and-lignite-production-in-2010.html

[16] U.S. Energy Information Administration. (2011, March 23 2012) Country Analysis Briefs, India. [Online]. Available: http://www.eia.gov/EMEU/cabs/India/pdf.pdf

[17] U.S. Energy Information. Administration. (2010, March 23 2012) Country Analysis Briefs, Russia. [Online]. Available: http://www.eia.gov/EMEU/cabs/Russia/pdf.pdf

[18] G. I. Atwater and J. P. J. Riva. (2012, March 29 2012) Natural Gas. [Online]. Available: http://0-www.britannica.com.millenium.itesm.mx/EBchecked/topic/406163/natural-gas

[19] U.S. Department of Energy. (2012, April 08 2012) The history of Natural Gas. [Online]. Available: http://fossil.energy.gov/education/energylessons/gas/gas_history.html

[20] Natural Gas Org. (2012, April 08 2012) History. [Online]. Available: http://www.naturalgas.org/overview/history.asp

[21] L. Borgard, D. McCurdy, "2012 Policy HandBook," American Gas Association, USA, Tech. Rep., Jan. 2013.

[22] G. I. Atwater and J. P. J. Riva. (2012, March 29 2012) Petroleum. [Online]. Available: http://0-www.britannica.com.millenium.itesm.mx/EBchecked/topic/454269/petroleum.

[23] F. Habashi, "The first oil well in the world," Quebec, Canada, 2000. [Online]. Available: http://www.scs.illinois.edu/~mainzv/HIST/bulletin_open_access/v25-1/v25-1%20p64-66.pdf

[24] R. McHenry, "The first oil well (150th anniversary of the birth of the oil industry)," August 26 2009. [Online]. Available: http://www.britannica.com/blogs/2009/08/the-first-oil-well/

[25] U. Davé, "Edwin drake and the oil well drill pipe)," Summer 2008. [Online]. Available: http://www.pabook.libraries.psu.edu/palitmap/DrakeOilWell.html

[26] Energy Information Administration. (2011, November) World Energy Outlook 2011. [Online]. Available: http://biblioteca.mty.itesm.mx/mty/bds/salta_bd.php?col_id=oecd

[27] Oil & Gas Journal, "Oil & gas reserves news," August 2012. [Online]. Available: http://www.ogj.com/oil-exploration-and-development/reserves.html

[28] British Petroleum, "Statistical of world energy 2012," June 2012. [Online]. Available: http://www.bp.com/sectionbodycopy.do?categoryId=7500&contentId=7068481

[29] Luis E. Echavarri, "Nuclear Energy Today 2012," International Energy Agency, France, Tech. Rep., Jan. 2012. [Online]. Available: http://www.oecd-nea.org/pub/nuclearenergytoday/6885-nuclear-energy-today.pdf

[30] Nuclear Energy Institute, "Nuclear energy around de world," 2012, visited on 04/04/2012,. [Online]. Available: http://www.nei.org/resourcesandstats/nuclear_statistics/worldstatistics/

[31] HPS. (2010, February) Consumer Products Containing Radioactive Materials. USA. [Online]. Available: http://www.hps.org/documents/consumerproducts.pdf

[32] N. Panwar, S. Kaushik, and S. Kothari, "Role of renewable energy sources in environmental protection: A review," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 3, pp. 1513 – 1524, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032110004065

[33] Britanica Encyclopedia Online. (2013, October 05 2013) Hydroelectric Power. [Online]. Available: http://0-www.britannica.com.millenium.itesm.mx/EBchecked/topic/278455/hydroelectric-power

[34] S. Markovic, T. Lazovic, L. Milovic, and B. Stojiljkovic, "The first hydroelectric power plant in the balkans built on the basis of tesla principles," in *Explorations in the History of Machines and Mechanisms*, ser. History of Mechanism and Machine Science, T. Koetsier and M. Ceccarelli, Eds. Springer Netherlands, 2012, vol. 15, pp. 395–406. [Online]. Available: http://dx.doi.org/10.1007/978-94-007-4132-4_27

[35] J. Cunningham, "Architect of power: Thomas e. murray amp; new york's electrical system [history]," *Power and Energy Magazine, IEEE*, vol. 10, no. 2, pp. 80–94, 2012.

[36] International Energy Agency. (2013, October 2013) Hydropower. [Online]. Available: http://www.iea.org/topics/hydropower/

[37] British Petroleum, "Bp statistical review of world energy june 2013," June 2013. [Online]. Available: http://www.bp.com/content/dam/bp/pdf/statistical-review/statistical_review_of_world_energy_2013.pdf

[38] IEA, "Electricity Information," International Energy Agency, USA, Tech. Rep., February 2012.

[39] U.S. Department of Energy. (2010, March 23 2012) 2010 Renewable Energy Data Book. [Online]. Available: www.nrel.gov/analysis/pdfs/51680.pdf

[40] S. Bull, "Renewable energy today and tomorrow," *Proceedings of the IEEE*, vol. 89, no. 8, pp. 1216 –1226, aug 2001.

[41] Hannele Holttinen and Jim Ahlgrimm, "Iea wind 2012 annual report," Tech. Rep., September 2013 2012. [Online]. Available: http://www.ieawind.org/annual_reports_PDF/2012/2012%20IEA%20Wind%20AR_smallPDF.pdf

[42] L. Jiang, Y. Chi, H. Qin, Z. Pei, Q. Li, M. Liu, J. Bai, W. Wang, S. Feng, W. Kong, and Q. Wang, "Wind energy in china," *Power and Energy Magazine, IEEE*, vol. 9, no. 6, pp. 36–46, 2011.

[43] B. Parida, S. Iniyan, and R. Goic, "A review of solar photovoltaic technologies," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 3, pp. 1625 – 1636, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032110004016

[44] IEA, "Solar Energy Perspectives," International Energy Agency, France, Tech. Rep., September 2011.

[45] M. Thirugnanasambandam, S. Iniyan, and R. Goic, "A review of solar thermal technologies," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 1, pp. 312 – 322, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032109001750

[46] International Energy Agency. (2013, September 2013) Solar (pv and csp). [Online]. Available: http://www.iea.org/topics/solarpvandcsp/

Variabilidad Climática y Caudales Mínimos en los Andes Ecuatorianos

Andrade L.*; Ríos C.**

** Escuela Politécnica Nacional, Facultad de Ingeniería Civil y Ambienta, Quito, Ecuador*
e-mail: laureano.andrade@epn.edu.ec
***Escuela Politécnica Nacional, Facultad de Ingeniería Civil y Ambiental, Quito, Ecuador*
e-mail:luis_riosc@hotmail.com

Resumen:*El estudio de caudales bajos o mínimos toma cada vez mayor importancia en relación con la gestión del agua. Se analizan 6 cuencas hidrográficas de la región Andina ecuatoriana, para las cuales se generaron series en valores mínimos de precipitación y de caudal, y de temperatura máxima, en base a registros históricos de 9 estaciones meteorológicas y 6 hidrológicas. Se analizan las características físico – geográficas de las cuencas, se establecen los balances hídricos respectivos, se evalúa la duración general de los caudales mediante formas gráficas adimensionales y se determinan los coeficientes de regulación natural de las cuencas. El análisis de tendencias entre las series refleja la disminución de las precipitaciones y de los correspondientes caudales de los períodos de aguas bajas, así también el incremento de la temperatura máxima mensual. El análisis correlacionante entre las series muestra la existencia de dependencias entre las series de precipitación y caudales mínimos.*

Palabras clave:*Caudales mínimos, duración general de caudales, precipitación, temperatura y estiajes.*

Abstract:*Low and minimum flows studies are more relevant nowdays in the relation with water resources management. Based in the historical data of 9 meteorological stations and 6 hydrological stations in the High Andes of Ecuador, 6 watersheds are analyzed for which time series of minimum precipitation, discharge, and maximum temperatures are generated. Physical and geographic characteristics of the basins are analyzed, water balance studies are done, general duration of flows by using dimensionless analysis is evaluated and basin natural regulation coefficients are determined. Trend analysis between the series reflects the decrease in rainfall and associated flows for low water periods, as well as the increase in the maximum temperature increase monthly. Correlation analysis between sets shows the existence of dependencies between sets of precipitation and minimum flows.*

***Key Words**: Low flows, flow duration curve, precipitation, temperature and droughts.*

1. INTRODUCCIÓN

El estudio de caudales mínimos o de estiaje se asocia generalmente con el diseño de proyectos de aprovechamiento de agua potable, riego, generación hidroeléctrica, navegación, recreación, manejo y preservación de cuencas, protección de ecosistemas acuáticos. Maspor la alta demanda de caudales y la notable concurrencia por el recurso, incluso con miles de usuarios en áreas relativamente pequeñas (**1**), la gestión del agua debería iniciarse desde los planes de desarrollo regionales y locales, tomando en cuenta la protección del recurso hídrico; luego enfocarse en el aprovechamiento integrado del agua y no solo sectorial; y por supuesto, considerar la administración y operación de los diferentes sistemas existentes de aprovechamiento o uso del agua. En este amplio contexto, la determinación y aseguramiento de los caudales mínimos (incluyendo los ecológicos) debe priorizarse, así también la protección de la calidad del agua.

El resumen para los tomadores de decisiones del Grupo de Trabajo 1 - IPCC (2007), evidencia la ocurrencia de sequías más intensas y prolongadas en amplias áreas desde 1970, particularmente en los trópicos y subtrópicos; y también, el incremento de los períodos secos, condicionados por las altas temperaturas y por la disminución de las precipitaciones, agravando la sequía.

En el contexto indicado se analiza:la incidencia de la variabilidad climática a través de los registros de precipitación y temperatura; y, se evalúa la ocurrencia y severidad de los caudales mínimos, tema que en gran parte corresponde al Proyecto de Titulación, Ríos (2010).

2. MARCO CONCEPTUAL

Según Vladimirov (1990), la escorrentía de los ríos en el período de aguas bajas, que es producto principalmente de la disminución o ausencia de precipitación, toma el nombre de caudales mínimos o de estiaje y al período durante el cual se presentan estos caudales, se denomina estiaje o estío.

Smathkin (2001) establece que los caudales bajos son un fenómeno estacional y un componente integral del régimen de caudales de cualquier río, y en cambio, la sequía es un evento resultante de una precipitación menor a la normal, durante un período relativamente largo.

Indica además que los caudales bajos son normalmente aportaciones de las descargas subterráneas o aportaciones superficiales de lagos, humedales (páramos) o de glaciales, y

anota que los caudales bajos ocurren generalmente en la misma estación del año.

En relación con el régimen climáticode la región Andina ecuatoriana, se destaca:

- La distribución estacional e interanual de la precipitación en el trópico de Sud América está condicionada por: la zona de convergencia intertropical (ZCIT);las masas de aire húmedas originadas en el Pacífico (oeste), en el Atlántico tropical (noreste) y en la cuenca Amazónica; y, por la cordillera de los Andes, que constituye una barrera física que incide en la ocurrencia y distribución espacial y temporal de la lluvias [9] y [15].
- En escala interanual el fenómeno océano – atmosférico de larga – escala y de baja frecuencia (El Niño) es el principal mecanismo de forzamiento de anomalías climáticas e hidrológicas en la zona[9] y [15].
- Las variaciones interanuales de la precipitación pueden ser considerables durante los fenómenos El Niño – Oscilación del Sur (ENSO), con lluvias torrenciales en la zona costera, que no llega a alturas superiores a 2 000 m s.n.m. [10] y [15].
- Los valles interandinos, localizados entre las Cordillera Occidental, y las Cordilleras Central y Oriental, soportan la influencia variada de masas de aire oceánicas y continentales, con dos períodos húmedos; febrero – mayo (muy lluvioso) y octubre – noviembre y dos períodos secos: junio – septiembre (más notable) y diciembre – enero [15].
- Las condiciones climáticas medias en Los Andes tropicales sufrieron cambios considerables durante el siglo XX; la temperatura aumentó en alrededor de 0.7 °C entre 1939 y 2006, en dependencia de la elevación y de la pendiente de las áreas; y, que este incremento permite entender las tasas de retroceso de los glaciares [16].

Varios autores referidos en [16] coinciden en que los ríos alimentados por páramos se caracterizan por tener un caudal base elevado, por la capacidad de retención de agua de los suelos y de la vegetación; y, que se enfrentan con la amenaza del cambio climático puesto que las temperaturas más altas ocasionarían mayores tasas de evapotranspiración y por consecuencia, una menor producción de agua de los páramos.

La ocurrencia junto con la variación espacial y temporal de los caudales mínimos pueden ser analizadas considerando los aspectos ya mencionados, en función de los escenarios propuestos por WMO (2009), que se presentan en forma de cubo Fig. 1.

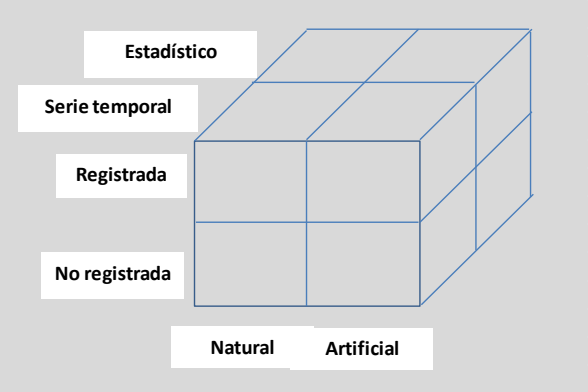


Figura 1. Esquema de cubo para analizar los caudales mínimos

En la gráfica anterior se definen los siguientes factores principales, en relación con las cuencas de estudio:

- La disponibilidad o no de estaciones hidrológicas con información representativa en la cuenca de interés (cuenca registrada o no);
- Los requerimientos operacionales de estudios hidrológicos, en función de la magnitud de las obras de aprovechamiento hídrico, que pueden demandar estudios estadísticos o análisis a mayor detalle mediante series temporales; y,
- La influencia antrópica en la cuenca, en función de los usos y derivaciones de agua, y también por el cambio en el uso del suelo, que puede demandar la conformación de series históricas de caudales, sumando los caudales derivados a los registros observados.

La combinación de los diferentes factores, define 8 escenarios potenciales, que condicionan el estudio de los caudales mínimos.

Por otra parte, los métodos de análisis de los caudales mínimos en relación con la determinación de sus características se sustentan en series de caudales históricos, y consisten básicamente en: curvas de duración general (CDG); análisis de frecuencias (AF) de eventos de caudales mínimos extremos, correspondientes a intervalos continuos de caudales bajos; y, separación de caudales base y caracterización de las curvas de recesión [5], [6], [8], [11], [12], [13] y [14].

Aplicaciones sobre AF a partir de series de caudales mínimos,para cuencas hidrográficas del sur ecuatoriano, se presentan en las referencias [1] y [2].

Entre los conceptos utilizados adelante se tienen los siguientes [12] y [14]:

- El coeficiente de regulación natural del caudal($\varphi 1$);
- El coeficiente modular (k) como relación entre el caudal y su valor medio (valor adimensional); y,
- La probabilidad (p) de que un caudal sea igualado o excedido (seguridad), expresada en fracción respecto a 1.

El coeficiente de regulación natural del caudal se determina mediante la siguiente relación (ver Fig. 2).

$$\varphi 1 = \int_0^1 pdk \qquad (1)$$

Hay que considerar que φ1 expresa el área bajo la CDG adimensional [12] y [14], en la cual la ordenada se representa con el coeficiente modular k, por lo que el coeficiente de regulación máximoserá 1, para cuencas con regulación hipotética total, si el coeficiente modular k = 1.

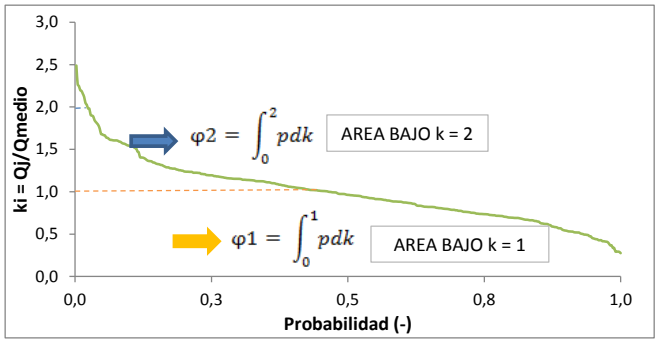


Figura 2. Áreas bajo la CDG adimensional

Con el fin de explicar la variación de los caudales en el tiempo, se determinó adicionalmente el coeficiente φ2, cuyo valor numérico representa el área bajo la CDG adimensional, considerando k = 2. La expresión de cálculo es por tanto:

$$\varphi 2 = \int_0^2 pdk \quad (2)$$

3. METODOLOGÍA APLICADA

Como procedimiento de análisis de los caudales mínimos se aplicó la siguiente secuencia:

- Selección de 6 cuencas hidrográficas con disponibilidad de registros históricos, localizadas en los Andes ecuatorianos: 2 de la vertiente occidental (ríos Alambi y Toachi) y 4 de la vertiente oriental (ríos Quijos, Yanahurco, Alao y Matadero);
- Conformación de la base de datos, consistente en: características físico – geográficas de las cuencas; series mensuales de precipitación y temperatura; y, series de caudales medios diarios, todo para el período 1971–2005 (35 años), con el siguiente detalle global: 11 estaciones con registros de temperatura media mensual, 18 estaciones con datos de precipitaciones medias mensuales y 6 estaciones hidrométricas con registros de caudales medios diarios; la ubicación de las estaciones constituye la sección de cierre de cada una de las cuencas hidrográficas consideradas;
- Relleno de datos faltantes mediante técnicas de regresión y correlación a nivel zonal con otras estaciones;
- Validación de los registros históricos aplicando pruebas no paramétricos, como son las curvas de masas y la prueba de rachas;
- Caracterización climática con la definición de los periodos de estiaje para cada una de las cuencas de drenaje;
- Determinación de series de precipitación acumulada, con valores mensuales mínimos anuales, considerando 1, 2, 3, 4, 5 y 6 meses consecutivos; y, generación de series de temperaturas máximas anuales para para intervalos de 1, 3 6 meses consecutivos;
- Obtención de series de caudales medios mensuales mínimos anuales, para intervalos de tiempo de 1, 2, 3, 4, 5 y 6 meses consecutivos;
- Obtención de los coeficientes de regulación natural φ1 y también φ2, para las cuencas hidrográficas;
- Análisis estadístico de las series generadas de caudales, evaluación de tendencias y de correlación entre series.

Para facilitar el análisis mediante la CDG, se generó una aplicación con la herramienta MACROS en Microsoft Excel, denominada CDG 1.05. El programa permiteobtener resúmenes de caudales característicos, dibujar la gráfica y determinar las características φ1 y φ2.

En base a la información meteorológica e hidrológica validada se establecieron balances hídricos anuales a nivel de cuenca hidrográfica; las láminas de evaporación real fueron estimadas mediante la relación de Turc.

Para el cálculo del balance hídrico anual se consideró a la cuenca como un sistema, con entradas y salidas de agua; y, se asumió despreciable la variación del volumen de almacenamiento de agua en la cuenca, en el período multianual 1971 – 2005, la ecuación utilizada es:

$$h_Q \cong h_p - h_e \quad (3)$$

Donde: h_Q: lámina de escurrimiento superficial media anual, (mm), h_p – lámina de precipitación anual media, (mm); y, h_e: lámina de evaporación anual media, (mm).

4. CUENCAS ANALIZADAS

En la selección de las cuencas se considerólos siguientes aspectos: la ubicación geográfica en la región Andina; el régimen climático; y, la disponibilidad de información climática e hidrológica.

La ubicación de las cuencas de estudio se muestra en la Fig.3,y sus características físico – geográficas principales se incluyen en la Tabla 1.

Las cuencas por localizarse en la región montañosa tienen suelos de origen volcánico, con capacidad de infiltración entre media y baja. Las áreas de drenaje varían entre 78.3 km^2 (cuenca del río Yanahurco) y 1526 km2 (cuenca del río Toachi); las pendientes medias de la cuencas fluctúan entre 25.3 % (cuenca del río Yanahurco) y 65.5 % (cuenca del río Alambi), que corresponden a áreas fuertemente accidentadas y escarpadas; y, la elevación media de las cuencas está, en todos los casos, por sobre los 2 200 m s.n.m.

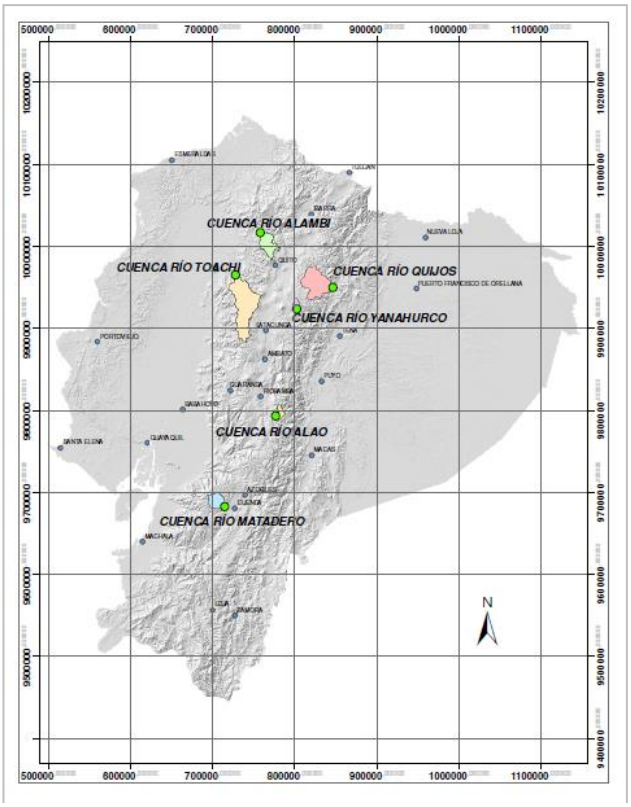


Figura 3. Ubicación de las cuencas hidrográficas

Tabla 1. Características físicas de las cuencas hidrográficas

Cuenca	X	Y	Área	Scuenca	Emedia	Vertiente
	(m)	(m)	(km²)	(%)	(m s.n.m.)	
Río Alambi	758166	10016593	442.0	65.5	2247.1	Occidental
Río Toachi	728005	9965099	1526.4	54.6	2845.4	Occidental
Río Alao	776892	9793144	108.0	59.4	3562.5	Oriental
Río Matadero	714912	9682700	304.0	55.1	3500.7	Oriental
Río Quijos	846638	9949710	932.5	31.2	3492.4	Oriental
Río Yanahurco	802637	9923470	78.3	25.3	3796.8	Oriental

En el cuadro anterior: X, Y: coordenadas de la sección de cierre (UTM WGS84 Z17S); Scuenca: pendiente media de la cuenca, (%) y, Emedia: elevación media de la cuenca, (m s.n.m).

En función de su ubicación las cuencas hidrográficas se identifican como:

- Cuencas con régimenclimático occidental: ríos Alambi y Toachi;
- Cuencas con régimen climático Andino, con influencia principalmente mixta: ríos Alao y Matadero; y,
- Cuencas con régimen climático oriental: río Quijos y Yanahurco.

El uso del suelo de las cuencas [7] se muestra en la Tabla 2.

Tabla 2.Cobertura vegetal en las cuencas hidrográficas

Cuenca	Páramo	Bosque	Cultivos y arbustos	Otros
	(%)	(%)	(%)	(%)
Río Alambi	2.0	47.7	50.2	0.1
Río Toachi	17.4	47.4	27.9	7.3
Río Alao	41.6	0.0	26.9	31.5
Río Matadero	85.5	3.2	9.7	1.5
Río Quijos	53.6	33.4	10.1	2.9
Río Yanahurco	100.0	0.0	0.0	0.0

El cuadro anterior muestra por una parte que la categoría páramo está presente, en porcentaje mayor al 41 % respecto al área, en las cuencas de la vertiente oriental (ríos Alao, Matadero, Quijos y Yanahurco), y, que la categoría Cultivos y Arbustos prevalecen en las cuencas de los ríos Alambi, Toachi y Alao (la categoría varios incluye principalmente áreas erosionadas, urbana y afloramientos rocosos).

5. ESTACIONES Y REGISTROS

Las estaciones hidrológicas principales consideradas en el análisis se muestran en la Tabla 3, en donde constan los siguientes códigos para las estaciones meteorológicas, de acuerdo al INAMHI: CP – climatológica principal, CO – climatológica ordinaria, PG – pluviográfica; y, PV – pluviométrica.

Tabla 3. Estaciones hidrometeorológicas principales

CUENCA	ESTACIÓN HIDROLÓGICA	ESTACIÓN METEOROLÓGICA
Río Alambi	Río Alambi en Churupamba (H136)	Nanegalito (M339) - PG La Perla (M213) - CP
Río Toachi	Río Toachi A.J. Pilatón (H161)	Las Pampas (M362) - CO Santo Domingo (M027) - CP
Río Alao	Río Alao en Hacienda Alao (H797)	Alao (M396)
Río Matadero	Río Matadero en Sayausí (H896)	El Labrado (M141) - CO Piscícola Chirimachay (M417) - PV
Río Quijos	Río Quijos en Baeza (H718)	Baeza (M215) - CP
Río Yanahurco	Río Yanahurco D.J. Valle (H722)	Papallacta (M188) - CO

Una vez conformada la base de datos en forma de series de precipitación, temperatura y de caudales medios diarios y mensuales para el período 1971-2005, se analizó la consistencia de la información mediante dobles masas y pruebas estadísticas de secuencias o rachas. Los resultados de las pruebas para las series de caudales medios anuales se indican en la Tabla 4.

Tabla 4. Prueba de secuencias para las series de caudales medios anuales

CUENCA	NA	NS	SERIE HOMOGÉNEA
Río Alambi	17	14	SI
Río Toachi	17	14	SI
Río Alao	17	16	SI
Río Matadero	17	16	SI
Río Quijos	17	15	SI
Río Yanahurco	17	14	SI

En la tabla anterior: NA - número de valores por sobre la mediana; y, NS - número de rachas en la serie.

6. CARACTERIZACIÓN CLIMÁTICA

La variación interanual y estacional de la precipitación y de la temperatura en el Ecuador se analizó de forma general en el numeral 2, sin embargo caben las siguientes puntualizaciones:

– El reconocimiento de que las cuencas hidrográficas tienen régimen climático occidental (del Pacífico), Oriental (del Atlántico tropical y Amazonía); y, régimen mixto, propio de las cuencas de la región montañosa de los Andes; y,

– La influencia del fenómeno ENSO, principalmente en las cuencas occidentales, en áreas con cotas menores a 2 000 m s.n.m.

La distribución de las lluvias en las cuencas - que se observa en la Fig.4 – presenta tres tendencias principales:

Clara presencia de masas de humedad occidentales en las cuencas de los ríos Alambi y Toachi, con un único período seco (junio – octubre);

Influencia prevaleciente de masas de humedad orientales en las cuencas de los ríos Alao, Quijos y Yanahurco; con un período de menores precipitaciones entre noviembre y febrero; y,

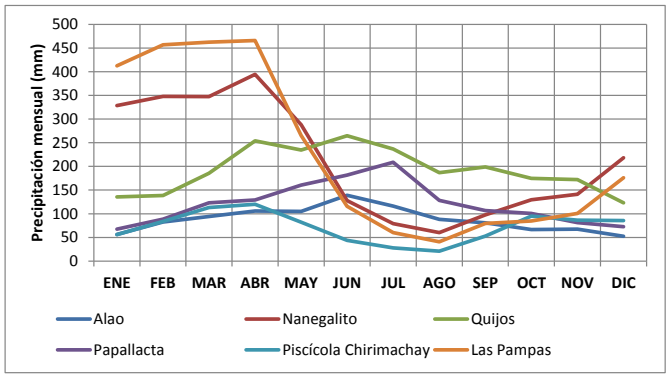


Figura 4. Variación mensual de la precipitación en las estaciones meteorológicas.

Régimen mixto de lluvias en la cuenca del río Matadero, con influencia occidental y oriental, con 2 períodos secos: junio septiembre (principal) y diciembre – enero.

La Tabla 5 muestra las características climáticas principales en función de los registros de las estaciones hidrometeorológicas, con lo cual se establece el período seco principal, y también se refiere la influencia del fenómeno ENSO con información de Rossel (1997).

Tabla 5. Características climáticas de las cuencas

Cuenca	Temperatura media anual	Precipitación anual media	Período seco principal	Influencia ENSO
	(°C)	(mm)		
Río Alambi	19.3	2552.4	Jun - Oct	Media
Río Toachi	23.5	2256.0	Jun – Oct	Media
Río Alao	7.9	1054.8	Nov - Feb	No
Río Matadero	16.3	867.6	Jun – Sep	No
Río Quijos	16.6	2326.8	Nov - Feb	No
Río Yanahurco	10.0	1448.4	Nov – Feb	No

Se propone el balance hídrico aproximado para cada una de las cuencas; ante la ausencia de información de evaporación real, se estimó esta característica mediante la relación de Turc, con datos de temperatura y precipitación. Las temperaturas y precipitaciones para las cuencas se evaluaron considerando la variación espacial y altitudinal de estas variables, con información de estaciones a nivel zonal. La lámina de escurrimiento se estableció en base al caudal medio anual de la cuenca; los valores obtenidos se muestran en la Tabla6.

Tabla 6. Balance hídrico para las cuencas hidrográficas.

CUENCA	Lámina de precipitación anual	Temperatura media anual	Lámina de escurrimiento anual	Lámina de evaporación anual	Lámina de evaporación anual Turc	Caudal medio anual
	(mm)	(mm)	(mm)	(mm)	(mm)	(m^3/s)
Río Alambi	1757	16	1238	519	822	17,3
Río Toachi	1414	14	955	459	692	46,2
Río Alao	2777	9	2320	457	562	7,9
Río Matadero	1127	11	725	402	573	7,0
Río Quijos	1962	12	1495	467	636	44,2
Río Yanahurco	1220	8	834	386	495	2,1

7. SERIES HIDROMETEOROLÓGICAS

Las variables utilizadas en el análisis estadístico-probabilístico para evaluar la relación entre variables climáticas e hidrológicas representativas de los caudales mínimos son precipitación,temperatura,y caudal, todo en valores mensuales. Se utilizó además registros diarios de caudal de las estaciones hidrológicas.

Las series históricas se disponen como una matriz $X_{i,j}$, en donde el subíndice i representa el año y el subíndice j los meses.

La Tabla 7indica a manera de ejemplo una serie de valores anuales característicos, definida como matriz $Z_{i,k}$; en donde i es el año al que corresponde el registro y k es el número de meses consecutivos considerados (entre 1 y 6).

Tabla 7. Matriz de valores críticos anuales

AÑO	NÚMERO DE MESES CONSECUTIVOS					
	1	2	3	4	5	6
1	$Z_{1,1}$	$Z_{1,2}$	$Z_{1,3}$	$Z_{1,4}$	$Z_{1,5}$	$Z_{1,6}$
2	$Z_{2,1}$	$Z_{2,2}$	$Z_{2,3}$	$Z_{2,4}$	$Z_{2,5}$	$Z_{2,6}$
…	…	…	…	…	…	…
n	$Z_{n,1}$	$Z_{n,2}$	$Z_{n,3}$	$Z_{n,4}$	$Z_{n,5}$	$Z_{n,6}$

Las relaciones para la determinación de los valores críticos de las series de precipitación, caudal y temperatura se incluyen como Apéndice A.

8. DURACIÓN GENERAL DE CAUDALES

A través de la duración general o permanencia de caudales se determinan los gastos correspondientes al rango de probabilidades de excedencia entre 50% y 90%, tomando en cuenta la práctica hidrológica nacional. Esta experiencia considera como caudales limitantes del uso del recurso hídrico (caudales mínimos) a los valores correspondientes a

las frecuencias 80% (para riego), 90% o 95% (para abastecimiento humano).

Las ventajas y desventajas que se presentan al determinar los caudales mínimos en base a la duración general son [2]:

- Se considera solo datos históricos y la distribución de probabilidades utilizada es en general empírica;
- Los resultados dependen en mucho del tamaño de la muestra (representatividad temporal); y,
- No toma en cuenta la estacionalidad de los caudales, sino implícitamente;

Los caudales medios diarios y mensuales fueron procesados para todo el período (1971 – 2005) y también para cada uno de los años, mediante curvas de duración general, con lo que se obtuvo los caudales mínimos con probabilidad de excedencia. En la Tabla 8 se presenta de forma esquemática los valores de la correspondiente matriz Qi,p, en donde i es el año, y p, la probabilidad de excedencia en %.

Tabla 8. Matriz de caudales medios y mínimos anuales de la CDG

AÑO	CAUDALES MEDIOS Y MÍNIMOS ANUALES					
	Q_{90}	Q_{80}	Q_{70}	Q_{60}	Q_{50}	Qm
1	$Q_{1,90}$	$Q_{1,80}$	$Q_{1,70}$	$Q_{1,60}$	$Q_{1,50}$	$Q_{1,m}$
2	$Q_{2,90}$	$Q_{2,80}$	$Q_{2,70}$	$Q_{2,60}$	$Q_{2,50}$	$Q_{2,m}$
…	…	…	…	…	…	…
N	$Q_{n,90}$	$Q_{n,80}$	$Q_{n,70}$	$Q_{n,60}$	$Q_{n,50}$	$Q_{n,m}$

En la tabla anterior: $Qi_{,90}$: caudal diario con probabilidad de excedencia del 90%, del año i, (m^3/s); $Qi_{,80}$: caudal diario con probabilidad de excedencia del 80%, del año i, (m^3/s).

Los caudales característicos de la CDG en valores diarios del período (1971 – 2005), para las cuencas hidrográficas, se muestran en el Apéndice B, en tanto que las formas gráficas correspondientes se presentan en valores específicos en la Fig.5.

La CDG en caudales específicos posibilita conocer la producción de caudales de las cuencas, sin embargo su forma dependerá de la mayor o menor humedad de la cuenca, por lo que se analiza las CDG en valores adimensionales de caudal, con ayuda del coeficiente modular (k), como se muestra en la Fig. 6 y en la Fig. 7, en donde se incluye la tendencia medias para las cuencas occidentales y orientales.

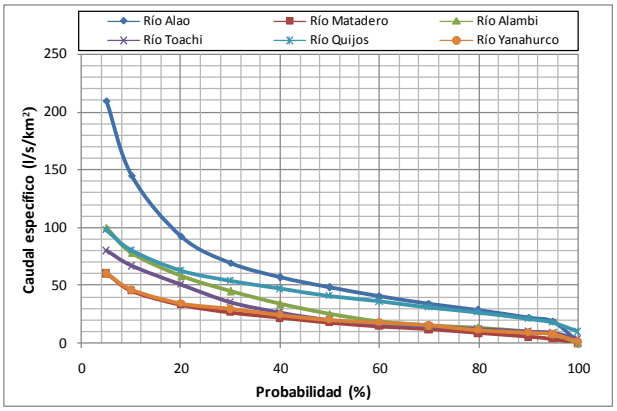


Figura 5. CDG de caudales específicos paras las cuencas hidrográficas

De la gráfica anterior se desprende:

- La superposición de las curvas correspondientes a las cuencas de los ríos Matadero y Yanahurco, en todo el rango de probabilidades;
- La coincidencia en las curvas correspondientes a las cuencas de los ríos Toachi, Matadero, Alambi y Yanahurco, en el tramo con probabilidad de excedencia mayor a 60 %; y,
- La coincidencia en las curvas de las cuencas de los ríos Quijos y Alao, en el rango mayor al 70 %.

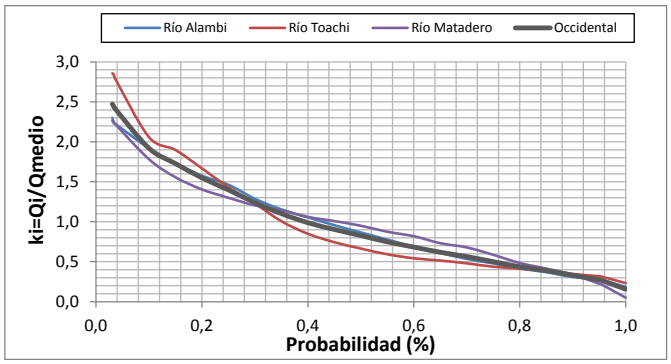


Figura 6. CDG adimensionales para las cuencas occidentales

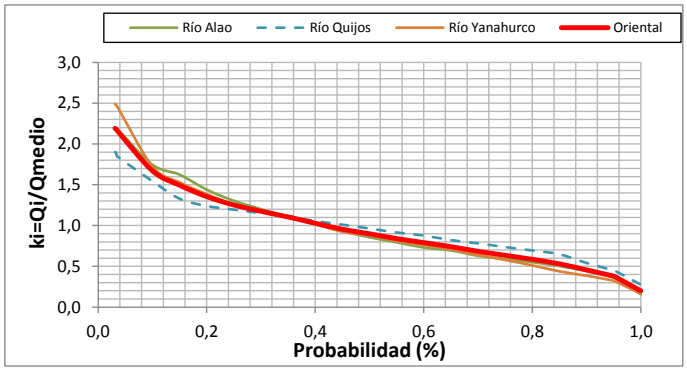


Figura 7. CDG adimensionales para las cuencas orientales.

Las coordenadas de las dos CDG adimensionales muestra la mayor humedad para las cuencas orientales (línea en rojo), respecto a las cuencas occidentales (línea negra), considerando el área φ1, bajo la ordenada k =1 (ver Fig. 8). La información de soporte consta como Apéndice C.

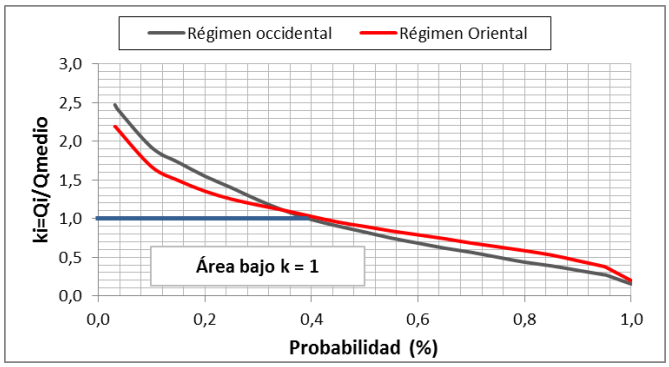


Figura 8. CDG adimensionales para las cuencas occidentales y orientales

El valor φ1 medio para las cuencas con influencia de las masas de humedad occidentales, es del orden de 0.73; y, para las cuencas con influencia oriental, alcanza 0,76. Esta condición se podría asociar con las mayores precipitaciones en el período con estiaje, que presentan generalmente las cuencas bajo influencia oriental de lluvias. Este rango de valores corresponde a regiones boscosas con presencia de lagos entre el 5 y 10 % del área de drenaje [12]. El valor φ2, toma valores del orden de 0.93 para todas las cuencas.

9. TENDENCIA DE LAS SERIES ANUALES

En base a las series generadas de precipitación, caudal mínimo y temperatura máxima, se analizó la variación temporal, para el periodo 1971-2005. Como metodología se utilizó la tendencia lineal para las series, como se muestra en la Fig. 9, para la estación Las Pampas, y en la Fig. 10, para los caudales mínimos de la estación hidrométrica Toachi A.J. Pilatón (en los dos casos para la cuenca del río Toachi).

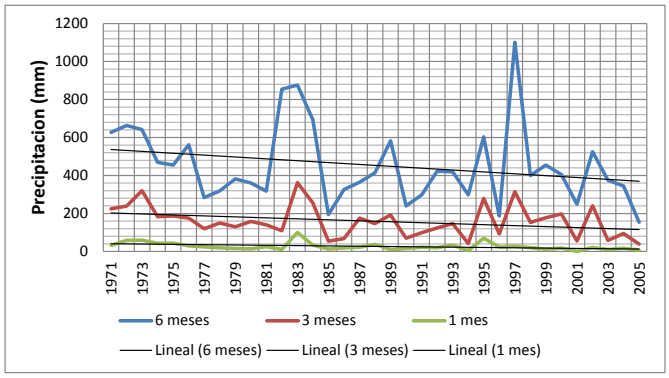


Figura 9. Decrecimiento de las precipitaciones de la estación Las Pampas (cuenca del río Toachi).

En la Tabla 9 se incluye en resumen, la tendencia para las variables hidrometeorológicas analizadas, negativa (disminución) y positiva (aumento).

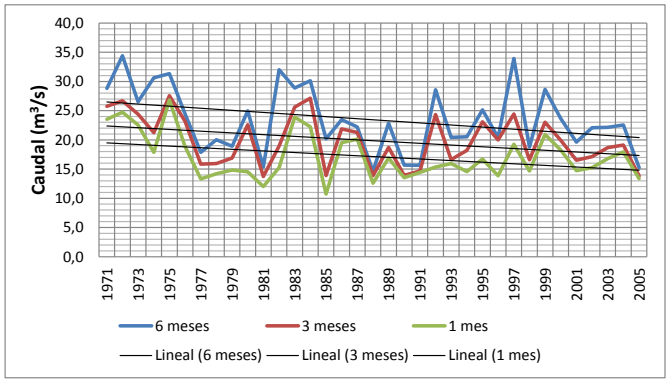


Figura 10. Decrecimiento de los caudales mínimos – río Toachi

Tabla 9. Variación (%) de las variables hidrometeorológicas consideradas

CUENCA	PRECIPITACIONES MÍNIMAS			CAUDALES MINIMOS			TEMPERATURAS MAXIMAS		
	1 MES	3 MESES	6 MESES	1 MES	3 MESES	6 MESES	1 MES	3 MESES	6 MESES
Alambi	-83	-60	-40	-51	-36	-24	3	3	6
Toachi	-75	-40	-30	-22	-22	-24	0	1	2
Alao	-36	-44	-27	-18	-45	-32	6	4	1
Matadero	-100	-40	-125	-79	-54	-42	8	5	4
Quijos	9	0	-11	-16	-15	-18	5	5	5
Yanahurco	0	-27	-29	-45	-33	-6	14	14	12

El Apéndice D muestra la información correspondiente, de donde se concluye lo siguiente:

- De manera general las series de precipitación observan un disminución de los valores, siendo más notable para el intervalo de 6 meses consecutivos de precipitaciones mínimas; las series de las estaciones representativas de las cuencas de los ríos Quijos y Yanahurco, muestran una tendencia contraria para intervalos de tiempo de 1 mes y de 3 meses consecutivos (zona con clara influencia oriental de lluvias);
- Las series de caudales mínimos en todos los casos presentan disminución de los valores (para intervalos de 1, 3 y 6 meses consecutivos);
- Las series de temperaturas máximas (17 de las 18 analizadas), observan el incremento de la temperatura, en valores entre 0,1 y 1,4 °C en el período.

Con el fin de contrastar la información anterior se aplicó a las series la prueba estadística para medias, considerando la hipótesis nula, de que la diferencia entre los valores medios de las muestras es igual a 0. Los resultados se muestran en la Tabla 10. Para el efecto se dividió a las series en dos,la primera correspondiente al período 1971 – 1988, y la segunda con los datos del período 1989 – 2005.

Tabla 10. Resultados de la prueba de diferencia de medias (si – se rechaza la hipótesis nula ; no – no se rechaza la hipótesis nula).

CUENCA	PRECIPITACIÓN ACUMULADA			CAUDALES MÍNIMOS			TEMPERATURA
	1 MES	3 MESES	6 MESES	1 MES	3 MESES	6 MESES	6 MESES
Alambi	si	si	si	si	no	no	si
Toachi	si	no	no	no	no	no	no
Alao	no	no	si	no	no	si	si
Matadero	si	si	si	si	si	si	si
Quijos	no	no	no	no	no	si	si
Yanahurco	no	no	no	si	no	no	si

10. RELACIONES CAUDALES MÍNIMOS – PRECIPITACIONES MÍNIMAS MENSUALES

Si bien los caudales de estiaje y los caudales mínimos, dependen del afloramiento de aguas subterráneas, y de los aportes de lagos, humedales y glaciales (en función del grado de cobertura en la cuenca), mas es claro que la distribución estacional principalmente de las precipitaciones en la zona Andina condiciona los períodos de ocurrencia de aguas bajas, e incluso la magnitud de estos caudales.

Para evaluar este condicionamiento se analizan conjuntamente las series de precipitación mínima y de caudales mínimos, a través de matricescorrelacionantes. En la Tabla 11 se muestran los resultados, como ejemplo para las cuencas de los ríos Alambi, Matadero y Yanahurco: en azul claro se presentan los coeficientes de correlación mayores a 0.6; y, en color gris constan los coeficientes de correlación entre 0.3 y 0.6.

CUENCA	VARIABLE		PRECIPITACIÓN MÍNIMA			CAUDAL MÍNIMO		
			1 mes	3 meses	6 meses	1 mes	3 meses	6 meses
Alambi	Precipitación mínima	1 mes	1,00					
		3 meses	0,83	1,00				
		6 meses	0,53	0,78	1,00			
	Caudal mínimo	1 mes	0,80	0,79	0,68	1,00		
		3 meses	0,63	0,81	0,69	0,82	1,00	
		6 meses	0,39	0,63	0,81	0,67	0,84	1,00
Matadero	Precipitación mínima	1 mes	1,00					
		3 meses	0,66	1,00				
		6 meses	0,65	0,83	1,00			
	Caudal mínimo	1 mes	0,55	0,65	0,69	1,00		
		3 meses	0,50	0,67	0,64	0,89	1,00	
		6 meses	0,58	0,62	0,69	0,85	0,88	1,00
Yanahurco	Precipitación mínima	1 mes	1,00					
		3 meses	0,61	1,00				
		6 meses	0,26	0,62	1,00			
	Caudal mínimo	1 mes	0,25	0,37	0,32	1,00		
		3 meses	0,32	0,48	0,24	0,80	1,00	
		6 meses	0,40	0,52	0,41	0,66	0,73	1,00

Tabla 11. Matriz correlacionanteentre series de precipitaciones mínimas y de caudales mínimos anuales.

En base a los datos de la tabla anterior se concluye la existencia de dependencia entre las series de precipitación, entre las series de precipitación y caudal, y, entre las propias series de caudales mínimos.

En general, las correlaciones son mejores para las cuencas con influencia climática occidental (del Pacífico), puesto que el período de estiaje es más largo y severo (5 meses, entre junio y octubre); en tanto, en las cuencas con influencia oriental, el estiaje es más corto (4 meses, entre noviembre y febrero), y generalmente menos pronunciado.

11. CONCLUSIONES Y RECOMENDACIONES

El régimen climático estacional e interanual de los Andes ecuatorianos ha sido estudiado por varios autores, quienes destacan:

- Que la distribución estacional e interanual de la precipitación está condicionada por fenómenos de escala regional, como son: la ZCIT; la circulación del aire que transporta humedad del Pacífico, del Atlántico Norte y de la Amazonía; y, el ENSO. Todo condicionado por la presencia de la cordillera de los Andes. Como consecuencia, los valles interandinos presentan por lo general dos períodos secos: junio – septiembre (más pronunciado) y diciembre – enero.
- Los métodos de análisis de los caudales mínimos se sustentan en series históricas de caudales y consisten en la aplicación de CDG, AF de series de caudales mínimos, y curvas de recesión. Se aplican además las CDG adimensionales, en las cuales en lugar del caudal absoluto o específico, se considera el coeficiente modular k, que es la relación entre el caudal y su valor medio anual.

Las cuencas estudiadas son 6, y se localizan en la región intermontañosa y en las laderas exteriores de los Andes, con régimen climático occidental, mixto y oriental. Los suelos de las cuencas son por lo general de origen volcánico con capacidad de infiltración entre media y baja. El uso del suelo corresponde a páramo, bosques y cultivos, con gran variación de cobertura de una cuenca a otra. Las áreas de drenaje varían entre 78.3 km^2 y 1 526 km^2, y, sus pendientes son fuertemente accidentadas y escarpadas. La elevación media está por sobre los 2 200 m s.n.m

La información básica hidrometeorológica, corresponde principalmente a 9 estaciones meteorológicas; y, 6 estaciones hidrológicas, las ubicaciones de estas últimas constituyen las secciones de cierre de las cuencas estudiadas.

El balance hídrico anual generado para las cuencas muestra que la precipitación varía entre 1 127 mm (río Matadero) y 2 777 mm (río Alao); la lámina de evaporación estimada mediante la ecuación de Turc, está entre 369 (río Yanahurco) y 525 mm (río Alambi); y, el escurrimiento anual determinado a partir de los caudales medios anuales, varía entre 721 mm (río Matadero) y 2 281 mm (río Alao).

Se conformaron series mínimas anuales de: precipitación acumulada para períodos de 1 - 6 meses; y, de caudales medios para períodos de 1 – 6 meses consecutivos. Se generaron además series de temperaturas máximas, a partir de los valores medios de 1, 3 y 6 meses consecutivos.

La duración general de caudales en valores adimensionales, permite agrupar las formas de las curvas en dos, que corresponden a las cuencas con influencia principal de masas húmedas occidentales y orientales.

El coeficiente de regulación natural $\varphi 1$ es del orden de 0.73 para las cuencas occidentales y 0.76 para las orientales, lo cual indica que las cuencas con régimen de lluvias oriental (bajo influencia de las masas de aire originadas en el Atlántico tropical y en la Amazonía), presentan un caudal base mejor distribuido en el tiempo.

El análisis de tendencia de las series anuales de precipitaciones y caudales mínimos y de temperaturas máximas, reflejan la disminución de los valores en el período 1971 – 2005. Las series de temperaturas máximas (17 de las 18) presentan aumentos de entre 0.1 y 1.4 °C.

Se aplicó a las series la prueba estadística de diferencia de los valores medios, y se determinó que en el 48 % de los casos (20 de 42 series analizadas) se rechaza la hipótesis nula: diferencia de medias = 0.

El análisis correlacionante entre series de precipitaciones mínimas y caudales mínimos, muestra la existencia de dependencias entre las variables analizadas. En general, se obtienen mejores coeficientes de correlación para las cuencas con influencia occidental de lluvias, aspecto que podría explicarse por observar períodos de estiaje más pronunciado (largo y severo) en relación con el período de caudales bajos de las cuencas con influencia oriental de lluvias.

Se recomienda ampliar el análisis desarrollado considerando otras cuencas hidrográficas en la región Andina. El principal limitante del estudio es la insuficiencia de estaciones meteorológicas por sobre los 3 000 m de altitud, aspecto que redunda en la obtención de series climáticas que representen de mejor manera el clima estacional e interanual de las partes altas de las cuencas Andinas.

12. DISCUSIÓN DE RESULTADOS

Si bien el presente análisis está limitado por la insuficiente representatividad de las estaciones meteorológicas consideradas, que redunda en la evaluación del régimen estacional e interanual de la precipitación y temperatura de las cuencas, sin embargo, el objetivo principal del trabajo se ha alcanzado, esto es mostrar la gran dependencia entre la

variabilidad climática y los caudales mínimos en las cuencas estudiadas, ubicadas en los Andes ecuatorianos.

La bibliografía analizada en general establece que los factores que inciden en el régimen de caudales mínimos son principalmente el afloramiento de las aguas subterráneas, y la retención y aportaciones de lagos y humedales (por su capacidad reguladora), más la dependencia analizada entre precipitaciones y caudales en los períodos de estiaje, muestra justamente, que la ocurrencia y distribución de las lluvias estacionales (en el período de estiaje) incide directamente en la presencia y variabilidad de los caudales mínimos de los ríos, a través del escurrimiento superficial.

13. REFERENCIAS

[1] L. S. Andrade, "*El proyecto hidroeléctrico Paute y su problema hidrometeorológico*". Revista Politécnica, vol. 17, 1992a.

[2] L. S. Andrade, "*Método de cálculo de los caudales mínimos para el sur ecuatoriano*". Revista Politécnca, vol. 17, 1992b.

[3] B. De Viebre y X. Coello, "*Caracterización de la demanda hídrica", Proyecto manejo integrado de los recursos hídricos de la hoya de Quito, UICN-SUR, 2008.*

[4] IPCC, (2007). "*Summary for Poliycymarkers. In: Climate Change 2007: The Physical Scince Basis*". Contribution of working Group I to the Fourth Assesment Report of the Intergovernmental on Climate Change. Cambridge University Press, Cmbridge, United Kingdom and New York, NY, USA.

[5] Guía de Practicas Hidrológicas, OMM, No. 168, pp. 413 – 417; 407 – 503.

[6] *Manual on Low – Flow, Estimating and Prediction, Operational* Hydrolgy, Report No. 50, WMO – No. 1029, IHP/HWRP.

[7] L. E. Ríos, "*Incidencia de la variabilidad climática en los caudales mínimos del Ecuador",* Proyecto de Titulación, *EPN, 2010.*

[8] G., Poveda et al., "*Estimación de caudales mínimos para Colombia mediante regionalización y aplicación de las curva de recesión de caudales*", Meteorología Colombiana, No. 6, pp. 73 – 80, 2002.

[9] G., Poveda and J. Mesa, "*Feedbacks between hydrological processes in tropical South America and large – scale ocean – atmospheric phenomena*", Journal of Climate, vol. 10, pp. 2690 – 2702, 1996.

[10] F. Rossel, "*Influencia de El Niño sobre los regímenes hidro-pluviométricos del Ecuador*".*INAMHI, ORSTOM, pp.203, 1997.*

[11] V. U. Smakhtin, "*Low flow hydrology: a review*", Journal of Hydrology, vol. 240, pp. 147 – 186, 2001.

[12] D. L. Sokolovski,"*Escurrimiento fluvial"*. Ed. Gidrometeoizdat, pp. 249 - 253, 1968 (en ruso).

[13] W. J. Viessman and L. L. Lewis, "*Introduction to hydrology*", Ed. Pearson – Prentice Hall, 2012, pp. 68 – 70.

[14] A. M. Vladimirov, "*Cálculos hidrológicos"*. Ed. Gidrometeoizdat, pp. 201 - 241, 1990 (en ruso).

[15] M. Vuille, et al., "*Climate variability in the Andes of Ecuador and its relation to tropical Pacific and Atlantic sea surface temperature anomalies*", Journal of Climate, vol. 13, pp. 2520 – 2535, 2000.

[16] M. Vuille, "*El cambio climático y los recursos hídricos en los Andes tropicales"*. BID, IDB Technical Note; 517, pp. 29, 2013.

Apéndice A.DETERMINACIÓN DE LOS VALORES DE LAS SERIES.

Apéndice B. CAUDALES MEDIOS Y MÍNIMOS, TOMADOS DE LAS CDG DE LAS CUENCAS HIDROGRÁFICAS, PERÍODO (1971 – 2005).

Apéndice C. COORDENADAS DE LAS CDG ADIMENSIONALES (k) – PERÍODO (1971 – 2005).

Apéndice D. TENDENCIAS DE LAS SERIES ANUALES – PERÍODO (1971 – 2005).

Apéndice A
DETERMINACIÓN DE LOS VALORES DE LAS SERIES

a) Los valores mínimos de **precipitación** se determinan con las siguientes relaciones:

Valor mínimo de 1 mes consecutivo:

$$Z_{1,1} = mín(X_{1,1},\ X_{1,2}, \ldots., X_{1,12}) (1)$$

Valor mínimo de 2 meses consecutivos:

$$Z_{1,2} = mín(X_{1,1} + X_{1,2},\quad X_{1,2}, +X_{1,3}, \ldots, X_{1,11} + X_{1,12}) \qquad (2)$$

Valor mínimo de 3 meses consecutivos:

$$Z_{1,3} = mín\ (X_{1,1} + X_{1,2} + X_{1,3},\quad X_{1,2}, +X_{1,3} + X_{1,4}, \ldots., X_{1,10} + X_{1,11} + X_{1,12}) \qquad (3)$$

Valor mínimo de 4 meses consecutivos:

$$Z_{1,4} = mín(X_{1,1} + X_{1,2} + X_{1,3} + X_{1,4},\quad X_{1,2}, +X_{1,3} + X_{1,4} + X_{1,5}, \ldots., X_{1,9} + X_{1,10} + X_{1,11} + X_{1,12}) \qquad (4)$$

Valor mínimo de 5 meses consecutivos:

$$Z_{1,5} = mín(X_{1,1} + X_{1,2} + X_{1,3} + X_{1,4} + X_{1,5},\quad X_{1,2}, +X_{1,3} + X_{1,4} + X_{1,5} + X_{1,6}, \ldots\ldots, X_{1,8} + X_{1,9} + X_{1,10} + X_{1,11} + X_{1,12}) \qquad (5)$$

Valor mínimo de 6 meses consecutivos:

$$Z_{1,6} = mín(X_{1,1} + X_{1,2} + X_{1,3} + X_{1,4} + X_{1,5} + X_{1,6},\quad X_{1,2}, +X_{1,3} + X_{1,4} + X_{1,5} + X_{1,6} + X_{1,7}\ , \ldots\ldots, X_{1,7} + X_{1,8} + X_{1,9} + X_{1,10} + X_{1,11} + X_{1,12}) \qquad (6)$$

Donde i: año del registro; j: mes del registro; y, $X_{i,j}$: valor de la variable meteorológica correspondiente al año i y mes j.

b) Los **caudales** mensuales consecutivos mínimos se estiman por ejemplo, para el caso de tres meses consecutivos, con la siguiente relación:

$$Z_{1,3} = \text{mín}\left(\frac{X_{1,1} + X_{1,2} + X_{1,3}}{3}, \frac{X_{1,2}, +X_{1,3} + X_{1,4}}{3}, \ldots, \frac{X_{1,10} + X_{1,11} + X_{1,12}}{3}\right) (10)$$

c) En tanto que para la determinación de los valores máximos consecutivos de **temperatura**, se considera la siguiente relación, para el caso de 3 meses consecutivos.

$$Z_{1,3} = máx\left(\frac{X_{1,1} + X_{1,2} + X_{1,3}}{3}, \frac{X_{1,2}, +X_{1,3} + X_{1,4}}{3}, \ldots, \frac{X_{1,10} + X_{1,11} + X_{1,12}}{3}\right) (11)$$

Apéndice B
CAUDALES MEDIOS Y MÍNIMOS, TOMADOS DE LAS CDG DE LAS CUENCAS HIDROGRÁFICAS PERÍODO (1971 – 2005)

CUENCA	CAUDALES ABSOLUTOS (m^3/s)						CAUDALES ESPECÍFICOS ($l/s/km^2$)					
	90%	80%	70%	60%	50%	Medio	90%	80%	70%	60%	50%	Medio
Río Alambi en Churupamba	4,70	5,70	6,80	8,40	11,3	16,1	10,6	12,9	15,4	19,0	25,6	36,5
Río Toachi A.J. Pilatón	16,1	19,2	22,6	25,9	31,0	47,3	10,5	12,6	14,8	17,0	20,3	31,0
Río Alao en Hacienda Alao	2,42	3,08	3,68	4,40	5,20	7,81	22,4	28,5	34,1	40,7	48,2	72,3
Río Matadero en Sayausí	1,62	2,66	3,64	4,55	5,54	6,95	5,33	8,75	12,0	15,0	18,2	22,9
Río Quijos en Baeza	19,8	25,1	29,3	33,8	38,5	44,2	21,2	26,9	31,4	36,2	41,3	47,4
Río Yanahurco D.J. Valle	0,70	0,90	1,20	1,40	1,60	2,07	8,94	11,5	15,3	17,9	20,4	26,4

Apéndice C

COORDENADAS DE LAS CURVAS DE DURACIÓN GENERAL ADIMENSIOANLES (k) - PERÍODO (1971 – 2005)

PROBABILIDAD		CUENCAS							
(%)	(…)	ALAMBI	TOACHI	ALAO	MATADERO	QUIJOS	YANAHURCO	OCCIDENTE	ORIENTE
3,1	0,03	2,30	2,86	2,18	2,26	1,90	2,49	2,47	2,19
3,3	0,03	2,23	2,85	2,18	2,26	1,89	2,47	2,45	2,18
3,6	0,04	2,23	2,80	2,18	2,23	1,84	2,45	2,42	2,16
10,0	0,10	1,92	2,06	1,74	1,78	1,54	1,72	1,92	1,67
15,0	0,15	1,73	1,90	1,62	1,56	1,33	1,53	1,73	1,49
20,0	0,20	1,58	1,67	1,44	1,40	1,24	1,38	1,55	1,35
25,0	0,25	1,46	1,43	1,30	1,30	1,19	1,26	1,40	1,25
30,0	0,30	1,29	1,23	1,20	1,20	1,15	1,17	1,24	1,17
35,0	0,35	1,15	1,01	1,10	1,14	1,11	1,10	1,10	1,11
40,0	0,40	1,05	0,85	1,01	1,06	1,05	1,02	0,99	1,03
45,0	0,45	0,96	0,74	0,93	1,01	1,01	0,92	0,90	0,95
50,0	0,50	0,87	0,66	0,86	0,95	0,96	0,88	0,83	0,90
55,0	0,55	0,78	0,59	0,79	0,87	0,91	0,81	0,75	0,84
60,0	0,60	0,69	0,54	0,73	0,82	0,88	0,76	0,68	0,79
65,0	0,65	0,61	0,51	0,69	0,73	0,82	0,71	0,62	0,74
70,0	0,70	0,53	0,48	0,63	0,68	0,78	0,64	0,56	0,68
75,0	0,75	0,48	0,44	0,59	0,59	0,74	0,58	0,50	0,63
80,0	0,80	0,41	0,41	0,55	0,48	0,69	0,51	0,44	0,59
85,0	0,85	0,37	0,39	0,50	0,41	0,65	0,44	0,39	0,53
90,0	0,90	0,31	0,35	0,44	0,34	0,54	0,38	0,33	0,46
95,0	0,95	0,28	0,32	0,36	0,22	0,45	0,32	0,27	0,38
95,2	0,95	0,28	0,32	0,36	0,22	0,44	0,32	0,27	0,37
100,0	1,00	0,19	0,23	0,16	0,05	0,28	0,17	0,16	0,20

Apéndice D

TENDENCIAS DE LAS SERIES ANUALES - PERÍODO (1971 – 2005)

CUENCA	REFERENCIA	PRECIPITACIONES MÍNIMAS (mm)			CAUDALES MINIMOS (m^3/s)			TEMPERATURAS MAXIMAS (°C)		
		1 MES	3 MESES	6 MESES	1 MES	3 MESES	6 MESES	1 MES	3 MESES	6 MESES
Alambi	1971	60	300	810	7,3	8,8	10,5	19,3	19,5	19,6
	2005	10	120	490	3,6	5,6	8,0	19,8	20,0	20,8
	Dif. Período	*-50*	*-180*	*-320*	*-3,7*	*-3,2*	*-2,5*	*0,5*	*0,5*	*1,2*
	Variación anual	-2,0	-7,2	-12,8	-0,15	-0,13	-0,10	0,02	0,02	0,05
Toachi	1971	40	200	530	19,2	22,3	26,7	24,2	24,5	24,7
	2005	10	120	370	15,0	17,3	20,3	24,3	24,8	25,2
	Dif. Período	-30	-80	-160	-4,2	-5,0	-6,4	0,1	0,3	0,5
	Variación anual	-1,2	-3,2	-6,4	-0,17	-0,20	-0,26	0,00	0,01	0,02
Alao	1971	55	250	600	3,3	4,9	8,2	8,0	8,2	8,6
	2005	35	140	440	2,7	2,7	5,6	8,5	8,5	8,7
	Dif. Período	-20	-110	-160	-0,6	-2,2	-2,6	0,5	0,3	0,1
	Variación anual	-0,8	-4,4	-6,4	-0,02	-0,09	-0,10	0,02	0,01	0,00
Matedero	1971	10	100	160	3,8	5,6	7,2	8,5	9,2	9,5
	2005	0	60	-40	0,8	2,6	4,2	9,2	9,6	9,8
	Dif. Período	-10	-40	-200	-3,0	-3,0	-3,0	0,7	0,4	0,4
	Variación anual	-0,4	-1,6	-8,0	-0,12	-0,12	-0,12	0,03	0,02	0,01
Quijos	1971	110	440	1120	25,0	33,0	45,0	16,5	16,9	17,2
	2005	120	440	1000	21,0	28,0	37,0	17,4	17,7	18,0
	Dif. Período	10	0	-120	-4,0	-5,0	-8,0	0,9	0,8	0,8
	Variación anual	0,4	0,0	-4,8	-0,16	-0,20	-0,32	0,04	0,03	0,03
Yanahurco	1971	40	220	720	1,1	1,4	1,8	9,5	9,8	10,2
	2005	40	160	510	0,6	0,9	1,7	10,8	11,2	11,4
	Dif. Período	0	-60	-210	-0,5	-0,5	-0,1	1,3	1,4	1,2
	Variación anual	0,0	-2,4	-8,4	-0,02	-0,02	0,00	0,05	0,06	0,05

Fiabilidad del Modelamiento Sísmico, en el Rango No Lineal, en Función de la Longitud de las Zonas Plásticas Asumidas Para Vigas

Celi C.*; Arellano B.*

** Escuela Politécnica Nacional, Facultad de Ingeniería Civil y Ambiental, Quito, Ecuador*
e-mail: carlos.celi@est.epn.edu.ec; roarebue@andinanet.net

Resumen: *Este artículo, presenta los resultados del análisis matemático de un pórtico de 4 vanos y 5 pisos, sometido a una carga monotónica, con una distribución de fuerzas laterales que corresponden a una combinación del primero y segundo modo de vibración de la estructura; dichas fuerzas se aplican en ciclos únicamente de carga de forma incremental, hasta alcanzar un desplazamiento de control asumido, es decir se realiza un Análisis Estático No Lineal (NSP), se proveen mecanismos de disipación de energía, en función de tres distintas longitudes de zonas plásticas para las vigas del pórtico y una única longitud de zonas plásticas para todas las columnas.*

Se realizaron 200 combinaciones de estos probables mecanismos de rotulación, comparando los resultados entre sí, con el empleó de programas computacionales como son Sap2000 y Matlab.

El presente artículo, representa la primera parte de una investigación denominada Análisis Paramétrico de Fiabilidad del Modelamiento Sísmico en el Rango No Lineal, en Función del Tipo de Modelamiento Empleado, en el cual se intenta identificar la eficiencia de dos diferentes procesos matemáticos, de simulación del mecanismo de disipación de energía como son, el Método de las Fibras y el Método de Zonas Plásticas Asumidas.

Palabras clave: *Monotónica, Zona Plástica, Análisis Estático No Lineal, Disipación de Energía.*

Abstract: *This article presents the results of mathematical analysis of one frame of four bays and five stories, subjected to monotonic loading, with lateral forces distribution corresponding to a combination of the first and second mode of vibration of the structure; these forces are applied in cycles of only incrementally load up to reach a displacement assumed control, ie, will conduct a Nonlinear Static Analysis (NSP), will be provide energy dissipation mechanisms, based on three different lengths of plastic zones for frame beams, and a single plastic zone length for all columns.*

We performed 200 combinations of these potential mechanisms, comparing the results among themselves, using computer programs such as Sap2000 and Matlab.

This paper represents the first part of a research called Parametric Reliability Analysis of Seismic Modeling the Nonlinear Range in Function Modeling Type Used, in which you try clarified the efficiency of two different processes of mathematical simulation, for energy dissipation mechanism such as the method of Fibres and Plastic Zone Assumed Method.

Keywords: *Monotonic, Plastic Zone, Nonlinear Static Analysis, Energy Dissipation.*

1. INTRODUCCIÓN

El procedimiento de mayor aceptación en nuestro medio, es realizar un análisis lineal, para una estructura sometida a una distribución de fuerzas laterales, que provienen de un Análisis Estático o Dinámico, para finalmente comprobar entre otros parámetros, el cumplimiento de una deriva máxima permitida, que es función del Factor R (Factor de Reducción de Resistencia Sísmica), que a su vez está en función de la ductilidad del sistema Estructural, es decir se obtiene un desplazamiento relativo inelástico.

Este tipo de procedimiento, aunque de buenos resultados, presenta limitaciones al preveer si la estructura controlada por ese desplazamiento inelástico, se encuentra en capacidad de resistir mayores deformaciones dentro de ciertos parámetros de seguridad, puesto que las fuerzas para las cuales se controló el desplazamiento inelástico relativo, no necesariamente son las fuerzas laterales que pueden presentarse durante la vida útil de la estructura.

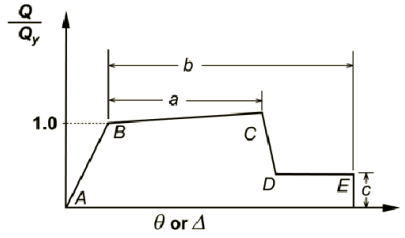
Se compensa esta deficiencia realizando cálculos adicionales como son el diseño de la conexión viga – columna, que

supone diseñar un nudo fuerte que soporte las acciones del evento sísmico que supere al sismo de diseño; esto supone diseñar las conexiones para una combinación extrema de carga lateral; es por esta razón que el Diseño por Desempeño, provee una buena alternativa, asumiendo niveles de daño que se aceptan para una amplia variación de una carga lateral incremental.

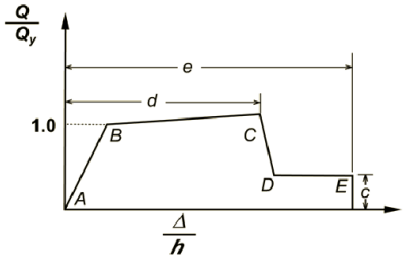
Para realizar este tipo de control, se empleó un Análisis Estático No Lineal (NSP), un paso intermedio entre el análisis lineal elástico y el análisis no lineal dinámico [6]. Dentro del Análisis No lineal Estático, la técnica de Push Over es la más empleada, para determinar la curva de capacidad resistente de las estructuras ante acciones sísmicas, curva que relaciona el cortante V con el desplazamiento lateral máximo en el tope de la estructura (Dt).

La relación carga-deformación deberá seguir lo dispuesto en 6.4.1.2 del documento FEMA-356 [6]; dicho numeral, presenta el **Procedimiento No Lineal**, que en su parte pertinente dice: cuando se emplee un procedimiento no lineal para determinar las acciones en una estructura, la componente de respuesta de carga-deformación deberá ser representada por una relación no lineal de carga-deformación, pero se permitirá una relación lineal cuando la componente no ocurra donde se dé una respuesta no lineal [6].

EL procedimiento Estático No Lineal (NSP), usa generalmente la relación carga-deformación como se muestra en la Fig. 1; también se permiten otras curvas que definan el comportamiento de la estructura bajo cargas monotónicas.



(a) Deformation



(b) Deformation ratio

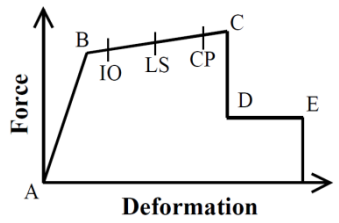


Figura1. Generalización de la relación fuerza-deformación para elementos de concreto. [6]

La Fig. 1, describe una relación lineal desde A (componente sin carga) a un rendimiento efectivo B, luego una reducción de rigidez lineal entre el punto B y C, seguida de una repentina reducción de resistencia a la carga lateral en el punto D, continuando hasta una reducida resistencia lateral en el punto E y una pérdida total de resistencia a partir de entonces [6].
Los puntos IO,LS,CP son usados como criterios aceptables de formación de rótulas plásticas, donde :
LS: Life Safety (seguridad para la vida)
CP: Collapse Prevention (prevención de colapso)
IO: Stand for Immediate Occupancy (de inmediata ocupación)

2. DESCRIPCIÓN Y PROCEDIMIENTO

Se procede con el análisis y diseño de un pórtico, obedeciendo a un procedimiento lineal, la configuración de la estructura como se muestra en la Fig. 2, que estará sometida a las cargas de servicio de la Tabla 1. Para determinar la distribución de cargas laterales que actúan en la estructura, se realizó un análisis Estático y Dinámico, descrito en CPE INEN 5:2001 [5] mediante programación en Matlab.
El Análisis Dinámico es el que presenta mayores fuerzas laterales hacia la estructura una vez realizada la superposición modal.

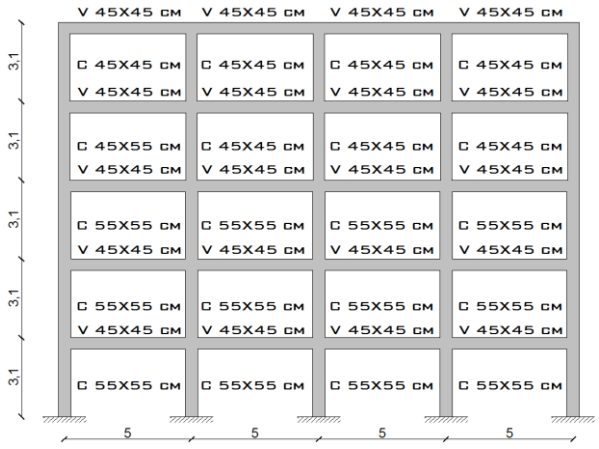


Figura 2. Geometría y Secciones

Los criterios empleados en el Análisis Estático y Dinámico, y los resultados de fuerzas laterales se presentan en la Tabla 1.

La referencia [2] se empleó para el diseño del pórtico, cuyas combinaciones de carga son de concordancia con el análisis proveniente de CPE INEN 5:2001 [5].

Tabla 1. Cargas empleadas y propiedades del hormigón.

	Carga por ml			Fuerza Lateral
Nivel	WD (Ton/ml)	WL (Ton/ml)	WP (Ton/ml)	F (Ton)
3.10	0.333	0.333	0.939	19.734
6.20	0.333	0.333	0.939	34.615
9.30	0.333	0.333	0.897	40.698
12.40	0.333	0.333	0.861	47.997
15.50	0.083	0.333	0.799	49.065

E = 200000 kg/cm2
f'c = 210 kg/cm2

R:	10.00
Z:	0.40
I:	1.00
g:	980.00
ΦEx :	0.90
ΦEy :	0.90
ΦP :	1.00
S :	1.20

El pórtico a analizar posee 90 probables mecanismos de formación de rótulas plásticas, que corresponden a los 2 extremos de cada una de las vigas y en cabeza y pie de cada una de las columnas que conforman el pórtico; con este precedente se decidió emplear 3 tipos diferentes de longitudes de zonas plásticas asumidas, estos son, 0 %, 2.5 % y 5 %; se realizaron 200 modelos basados en las distintas combinaciones de los tres parámetros antes mencionados; en la Fig. 3. se muestran 2 de las 200 combinaciones empleadas.

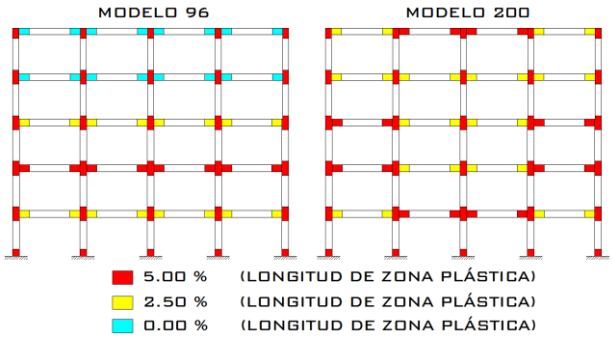


Figura 3. Distribución de Zonas Plásticas, Modelos 96 y 200.

Antes de realizar el Análisis Estático No Lineal (NSP), empleando el programa Sap 2000, se realizó una comprobación manual, en base al ATC 40 [4]; se describe un procedimiento simplificado para la obtención de la curva de capacidad:

1. Aplicar una carga concentrada horizontal en el tope de la estructura
2. Aplicar fuerzas laterales en cada piso de acuerdo a CPE INEN 5:2001[5], sin considerar la fuerza concentrada F1 del tope de la estructura

$$Fx = \left[\frac{Wi * hi}{\sum Wi * hi}\right] * V \quad (1)$$

Wi = peso del piso "i"
hi = altura del piso "i"
V = cortante basal

3. Aplicar las fuerzas laterales en proporción al producto de las masas de los pisos y el primer modo de vibración del modelo de la estructura elástica.

$$Fx = \left[\frac{Wi * \phi i}{\sum Wi * \phi i}\right] * V \quad (2)$$

Wi = peso del piso "i"
Φi = primer modo de vibración del piso "i"
V = cortante basal

4. Se procederá igual que en el paso 3 hasta alcanzar la primera fluencia, luego para cada incremento que sobrepase dicha primera fluencia, se deberán ajustar las fuerzas de manera consistente con el cambio en la forma de la deflexión.
5. Similar a los pasos 3 y 4, se incluirán los efectos de los modos superiores de vibración en la determinación de la fluencia de los elementos individuales de la estructura, mientras se plotea la curva de capacidad del edificio en términos de las fuerzas laterales del primer modo de vibración y los desplazamientos que produzca.
6. Plotear el cortante en la base y el desplazamiento en el tope.
7. Realizar un modelo usando una cantidad muy pequeña de rigidez en los elementos que se espera que fluyan primero.
8. Aplicar un nuevo incremento de cargas laterales a la estructura modelada anteriormente.
9. Añadir el nuevo incremento de cargas laterales que produce el desplazamiento del tope, al sub total previo para obtener el cortante total en la base y el desplazamiento total del tope.
10. Repetir los pasos 7, 8 y 9 hasta que la estructura alcance su límite último, como es: inestabilidad por efectos P-Δ (considerables distorsiones más allá del desempeño esperado).
    Cuando se alcanza el nivel de deformación lateral se produce una pérdida en la capacidad de soportar las cargas gravitacionales como se muestra en la Fig. 4.

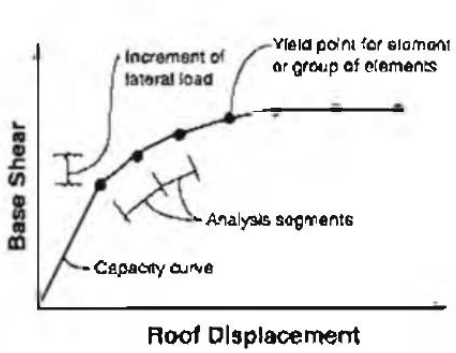


Figura 4. Curva de Capacidad. [4]

11. El modelo explícito de la degradación global de la resistencia, se da cuando la carga incremental se detiene en el paso 10 descrito anteriormente, como resultado de alcanzar un nivel de deformación lateral total o significativa, en el que los elementos (o grupo de elementos) no puedan resistir más carga. Esto es cuando la resistencia que pueden absorber los elementos estructurales se ha degradado significativamente, y la rigidez ha sido reducida o eliminada.
    Deberá realizarse una nueva curva empezando desde el paso 3, y crear tantas curvas pushover como sean necesarias para adecuadamente describir la pérdida total de resistencia de la estructura.
    Finalmente se trazará una curva que inicia por el trazado de la primera curva, luego se realiza una transición a la segunda curva, correspondiente al desplazamiento donde ocurrió la primera degradación significativa de la resistencia, y así para los demás puntos de degradación.
    Esta curva toma una forma **sawtooth** (cierra dentada), como se aprecia en la Fig. 5.

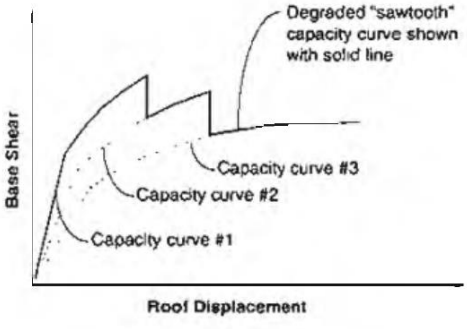


Figura 5. Curva de Capacidad, con el modelo de Degradación Global. [4]

Se procede a realizar la curva de capacidad empleando programación en Matlab, siguiendo los pasos anteriormente mencionados, con base en los procesos de cálculo presentes en FEMA 273 [7], ATC40[4] y ASPD [3], para lo cual previamente se debe establecer la degradación de la inercia de los elementos estructurales, a través de un diagrama de momento-curvatura [3].

Para el modelo de hormigón se emplea el Modelo de PAPOVIC [10], en el cual se asume un valor de k = 1.2 (factor que considera el refuerzo de confinamiento del elemento estructural), el valor de **k** se obtuvo realizando el proceso descrito en MANDER [8].

La curva del modelo de hormigón calculado para el presente artículo se observa en la Fig. 6. Como se trata de una comprobación de los resultados que se realiza en Sap 2000, para la programación en Matlab, no se considera el aporte de adherencia ni cortante, además de que la estructura será representada como un péndulo invertido con masas concentradas en cada uno de sus pisos.

El momento-curvatura se realizará de una columna intermedia del pórtico a analizar, empleando el Método de Dovelas [3] para la primera parte de la gráfica (M-φ), y el cálculo convencional de momento Curvatura [8] para el resto de la gráfica (M-φ); además se realiza el arreglo bi-lineal que iguale el área debajo de la curva del diagrama momento-curvatura (M-φ) y cumpla con la condición que la gráfica real y la curva idealizada se intercepten en el 75% del momento de fluencia según las recomendaciones de Paulay y Priestley [9], como se muestra en la Fig. 7.

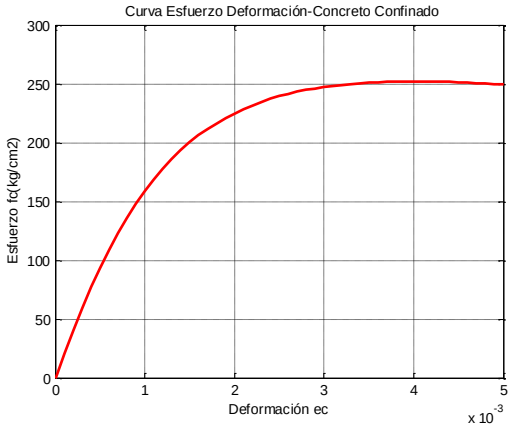


Figura 6. Curva Esfuerzo Deformación, Concreto – Confinado, [9]

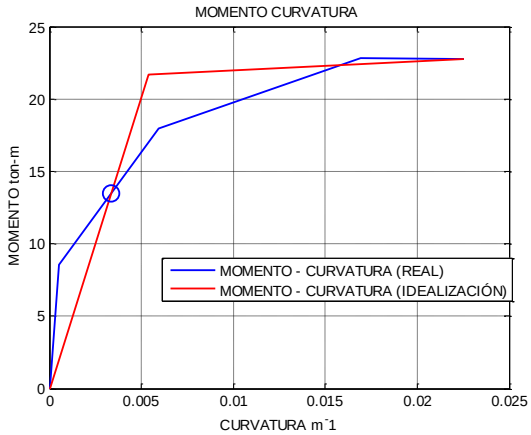


Figura 7. Momento-Curvatura (M-φ), Columna Central

Como se mencionó, a partir del diagrama momento – curvatura (M-φ) se obtiene la degradación de la inercia del elemento, ya que el módulo de rigidez (EI) es igual al momento aplicado dividido para la curvatura en ese punto. Para el presente artículo se considera que el módulo de elasticidad es constante, con el fin de simplificar el cálculo de la degradación del módulo de rigidez.

Se emplean secantes para el cálculo de la degradación, tomando en consideración que todas las secantes tienen como origen el inicio del diagrama momento – curvatura, y no se toma en consideración un desplazamiento plástico irrecuperable [8], para el inicio de la próxima secante, con el fin de simplificar el cálculo.

$$E * Ief = \left[\frac{M}{\phi}\right] \quad (3)$$

$$Icr = \left[\frac{E * Ief}{E * Ig}\right] \quad (4)$$

Φ = curvatura
Ief = inercia efectiva de la sección
Ig = inercia gruesa de la sección
Icr = relación de inercias

En la Fig. 8, representa la degradación de la inercia del elemento en función del momento aplicado.

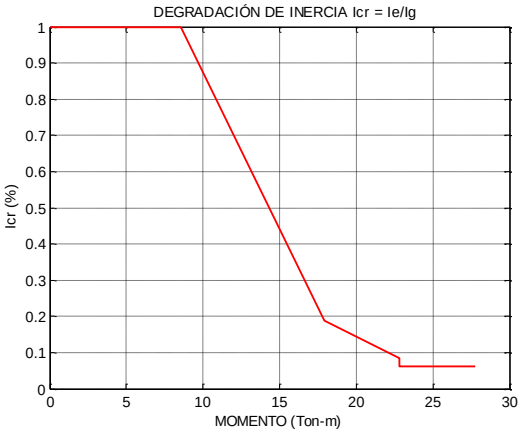


Figura 8. Degradación de Inercia, Columna central

Con fin ilustrativo se calculó la degradación del módulo de elasticidad, Fig. 9, a partir de la curva esfuerzo-deformación, previamente calculada, observándose que para el rango de variación de las deformaciones del hormigón, el módulo de elasticidad no varía significativamente.

Se realiza un cálculo manual lineal, para obtener los desplazamientos en el tope de la estructura, usando el modelo de péndulo invertido antes mencionado sometido a cargas transversales, con incrementos de carga de 0.50 t, obteniendo una curva de capacidad lineal.

Se observan las acciones (momentos) correspondientes a cada ciclo de carga, y se los correlaciona con la Fig. 8, la cual presenta la degradación de la inercia en función del momento aplicado, de tal manera de realizar un cálculo interactivo, agrietando cada vez más las secciones en función del momento resultante. Una vez realizadas las interacciones, se obtiene una curva de capacidad en función del incremento de las fuerzas laterales, obtenidas de la superposición modal del primero y segundo modo de vibración de la estructura.

Para simplificar el cálculo de la curva de capacidad, se realiza una bilinealización, para lo cual es mandatorio obtener la rigidez antes de la primera fluencia, y la rigidez post-fluencia, para lo cual la Referencia [7], presenta algunos criterios de cálculo; en la presente investigación se determinó:

Ke = 115 ton (rigidez equivalente antes de la primera fluencia)

Kp = 4.6 ton (rigidez post-fluencia)

Donde, **Ke** es la rigidez secante con origen en el centro de coordenadas de la curva de capacidad hasta la primera fluencia, y **Kp** es la pendiente entre la primera fluencia y el punto de colapso teórico de la curva de capacidad.

El colapso teórico se considera cuando el desplazamiento lateral en el tope es 5% de la altura del edificio, según ASDP [3].

Una vez establecidos Ke, Kp, y por tanto los desplazamientos correspondientes a ellos, se calcula la curva de capacidad bilinealizada según las ecuaciones presentes en ASPD [3], a partir de las matrices de compatibilidad, vector de cargas, cortante basal, entre otros parámetros propios de la estructura a analizar, obteniendo la curva de capacidad que se muestra en la Fig.10.

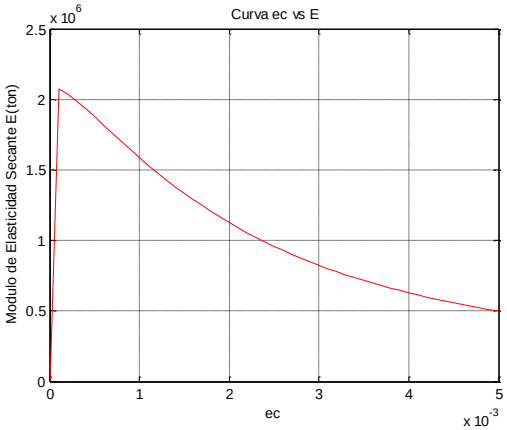


Figura 9. Degradación Modulo de Elasticidad

Antes de obtener el punto de desempeño, primero se obtiene el espectro de capacidad (Sa-Sd) a partir de la curva de capacidad.

La curva de capacidad previamente calculada es una bilinelización, por lo que si se desea, se puede obtener el espectro de capacidad con una desviación estadística, para absorber las posibles fallas en la estimación de Ke y Kp [3]. El espectro de capacidad más la desviación se muestra en la Fig. 11.

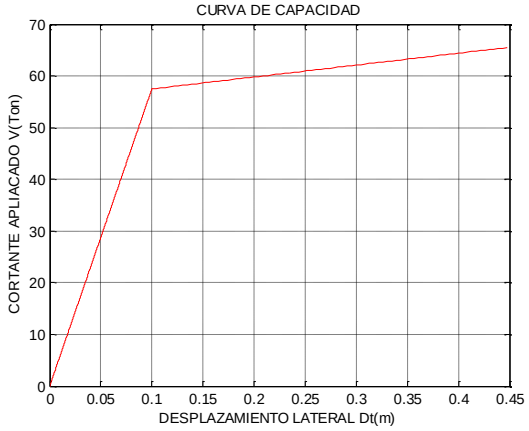


Figura 10. Curva de Capacidad del Pórtico Analizado

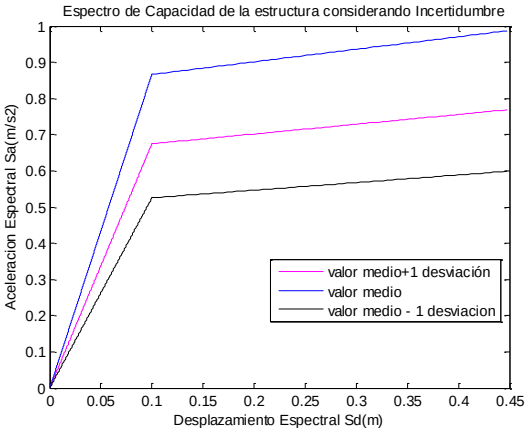


Figura 11. Espectro de Capacidad, Considerando Incertidumbre.

Como se mencionó en la Tabla 1, se emplea el espectro inelástico para un suelo S2 de acuerdo a CPE INEN 5:2001 [5], transformándolo en un espectro inelástico en función de (Sa -Sd), con base en las ecuaciones propuestas en ASPD [3].
Realizada esta conversión, se procede a calcular la reducción del espectro inelástico de aceleración en función de (Sa - Sd), siguiendo la recomendación del método 1, dispuesta en la referencia [4], que es un proceso interactivo de reducción del espectro de aceleración inelástico en función de una probable intersección con el espectro de capacidad de la estructura.
Este proceso se realizó tantas veces hasta que la respuesta convergió, es decir se obtuvo el **Punto de Desempeño de la Estructura**, como se muestra en la Fig.12.
Las coordenadas resultantes del punto de desempeño son: Sa = 0.68 m/s2 y Sd = 0.105 m; realizando la conversión pertinente, se tiene que el punto de desempeño en coordenadas de la curva de capacidad se localiza en: V = 60.88 Ton y Dt = 0.14m, siendo este resultado muy próximo a los que se obtuvieron con el Programa Sap 2000 cuando se consideró plasticidad de 0.05 en los 90 probables mecanismos de formación de rótulas plásticas.
Una vez que se comprobaron los resultados que entrega el Programa Sap 2000, se procede a la simulación de las 200 diferentes combinaciones de mecanismos de formación de rótulas plásticas que se mencionó anteriormente, empleando las recomendaciones de la referencia [6].

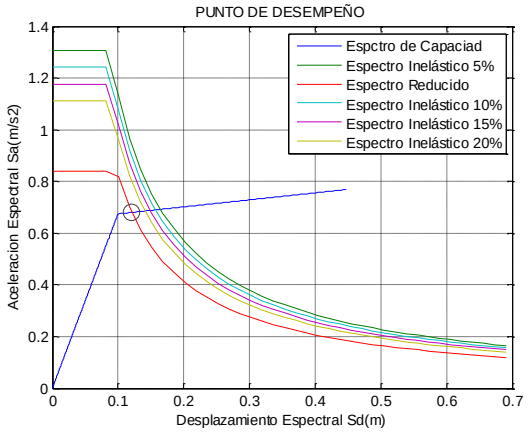


Figura 12. Punto de Desempeño

Se presentan algunas curvas de capacidad calculadas con el programa Sap 2000 y filtradas-ordenadas mediante programación en Matlab, en las Fig. 13, Fig. 14. y Fig. 15.

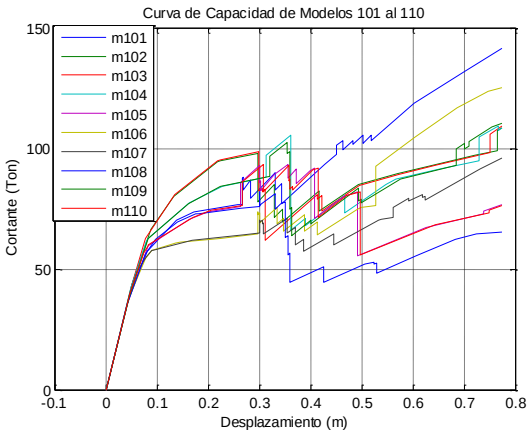


Figura 13. Curvas de Capacidad, Modelos 101 – 110

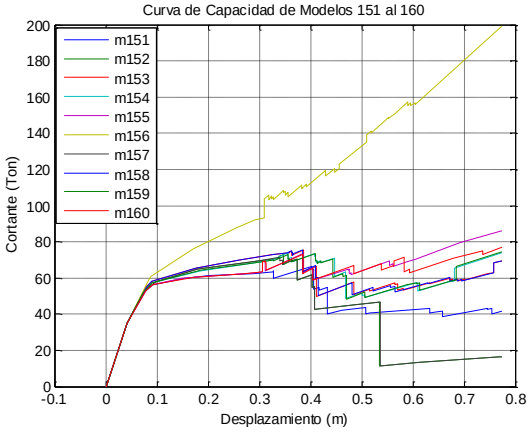


Figura 14. Curvas de Capacidad, Modelos 151 – 160

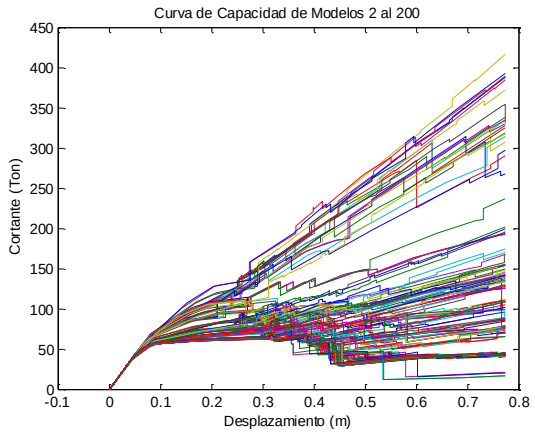


Figura 15. Curvas de Capacidad, Modelos 2 – 200

La Fig. 16, muestra la variación del cortante del punto de desempeño, en cada uno de los 200 modelos realizados.

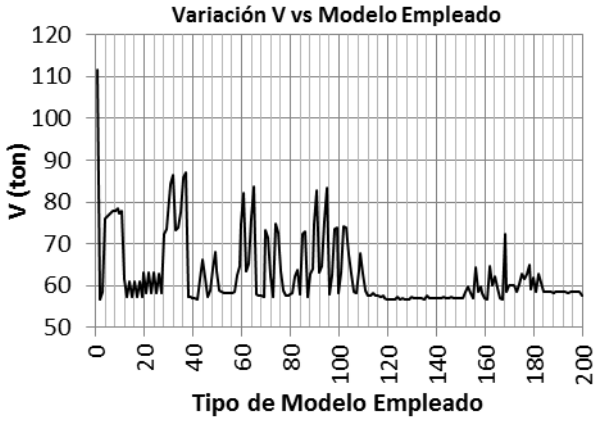


Fig 16. Variación del Cortante (V) ,en Función del Modelo Empleado

La Fig. 17, muestra la variación del desplazamiento en el tope de la estructura para el punto de desempeño, en cada uno de los 200 modelos realizados.

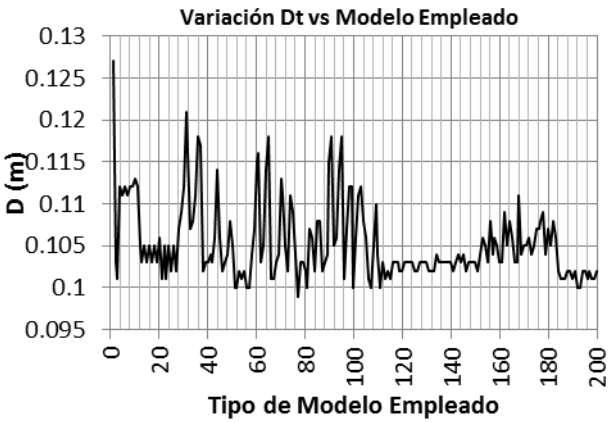


Figura 17. Variación del Desplazamiento (Dt), en Función del Modelo Empleado

La Fig. 18, muestra las coordenadas del punto de desempeño, en función del cortante en el tope de la estructura y su desplazamiento respectivo, en cada uno de los 200 modelos realizados.

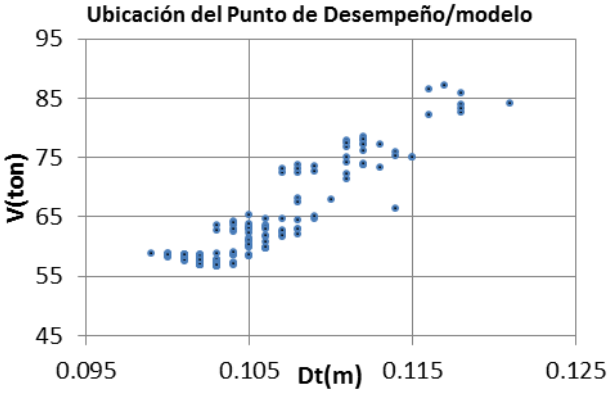


Figura 18. Coordenadas del Punto de Desempeño, en Función del Modelo Empleado

La Fig. 19, muestra la variación entre el cortante (V) cuando se alcanzó el mayor número de rótulas plásticas para el nivel de desempeño CP (prevención de colapso) y el cortante (V) cuando se alcanzó el punto de desempeño, en cada uno de los 200 modelos realizados.

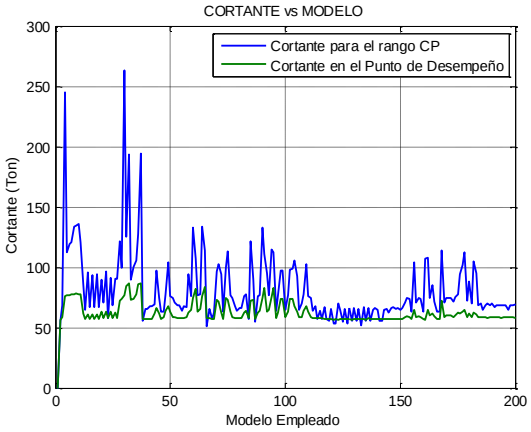


Figura 19. Variación del Cortante (V) entre Niveles de Desempeño.

La Fig. 20, muestra la variación entre el desplazamiento en el tope de la estructura (Dt) cuando se alcanzó el mayor número de rótulas plásticas para el nivel de desempeño CP (prevención de colapso) y el desplazamiento en el tope de la estructura (Dt), cuando se alcanzó el punto de desempeño, en cada uno de los 200 modelos realizados.

La Fig 21, muestra el **Porcentaje de Ciclos de Carga** necesarios para obtener el mayor número de formación de rotulas plásticas en el nivel de desempeño CP (prevención de colapso), en cada uno de los 200 modelos realizados.

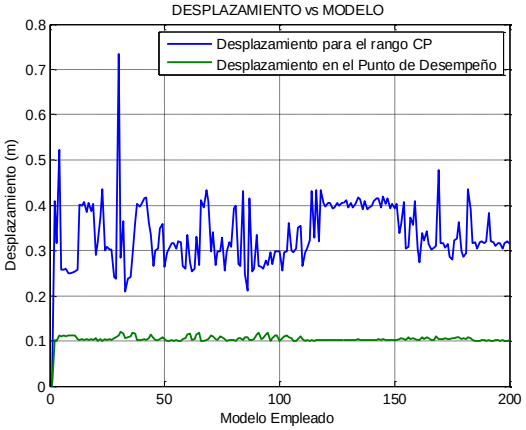


Figura 20. Variación del Desplazamiento en el Tope (Dt) entre Niveles de Desempeño.

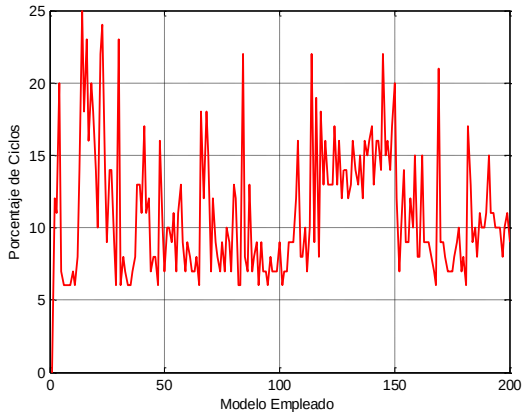


Figura 21.

3. CONCLUSIONES

De las Fig. 16 y Fig. 17, se observa una normalización de resultados, independientemente de la combinación de zonas plásticas asumidas, no con esto excluyendo que exista una variación de resultados entre las últimas modelaciones, sino que la variación no es de gran magnitud, en comparación con los primeros modelos, los cuales poseen un gran número de mecanismos de rotulación plástica prácticamente lineales.
Esta conclusión se da, puesto que la variación del cortante en el punto de desempeño entre los modelos 2 al 120, es del orden de 34 ton, mientras que del modelo 120 al 200 es de alrededor de 7 ton, mostrando una clara tendencia de los resultados para las próximas combinaciones, tomando en cuenta que el valor del cortante promedio es de alrededor de 58 ton.
No obstante en la Fig. 17, la variación del desplazamiento en el tope de la estructura para el punto de desempeño, es de baja magnitud en relación a su valor nominal, es decir que dicho desplazamiento es de alrededor de 0.105 m, y la variación entre los modelos 2 al 200 es de alrededor de 0.021 m.

En primera instancia se planificó realizar 1000 modelos basados en la combinación de zonas plásticas asumidas, pero ya que en las primeras 200 modelaciones, se observó una clara tendencia de normalización de los resultados, se concluye que las siguientes 800 permutaciones, corresponderán a resultados similares a la tendencia previamente observada, ya que estas 800 permutaciones son la combinación de zonas plásticas asumidas de 0.025 y 0.05, las cuales representan un comportamiento altamente no lineal.

De la Fig. 19. se observa que existe una variación del orden de 10 ton entre los cortantes (V) para el punto de desempeño en comparación cuando se alcanza el mayor número de rótulas plásticas en rango CP (prevención de colapso), entre los modelos que presentan una tendencia de resultados, es decir que la estructura se encuentra en capacidad de absorber hasta un 18 % más de fuerza lateral, antes de incursionar en rangos de desempeño peligrosos.

Si se correlacionan los desplazamientos en función del rango de desempeño CP (prevención de colapso) Fig. 20, versus el desplazamiento relativo inelástico con el que se controla la estructura calculada linealmente, se observa que.

Desplazamiento Promedio Rango CP = 0.36 m
Desplazamiento Relativo Inelástico = 0.31 m

Es decir, que el chequeo que se realiza para la estructura calculada linealmente, es bastante cercano al calculado mediante un Análisis Estático No Lineal (NSP), con la diferencia que el chequeo inelástico no permite determinar si la estructura, una vez que alcanzó el desplazamiento inelástico, presente graves daños en el arranque y fin de las vigas de todos los pisos, y que dichas rotulas ya no serán capaces de seguir disipando energía de manera segura, puesto que se encuentran en el nivel de desempeño CP (prevención de colapso).
Esto demuestra que aunque el proceso de cálculo lineal, controlado por desplazamientos inelásticos relativos, ofrece un control para la seguridad de la estructura, no permite visualizar el grado de daño que la misma afrontará una vez alcanzado el desplazamiento inelástico, y tampoco poder asegurar si la estructura podrá seguir disipando energía de manera segura.
La comparación entre los resultados obtenidos del desplazamiento relativo Lineal Inelástico y el desplazamiento promedio en el rango CP, no necesariamente es parámetro de validación el Análisis Estático No Lineal con uso de Zonas Plásticas Asumidas, puesto que esta investigación pretende mostrar que la longitud de la zona plástica asumida para vigas es altamente influyente en la curva de capacidad de la estructura, dependiendo de principalmente de la posición más que el rango de la Longitud de zona plástica. No obstante esta es la primera parte de una investigación de mayor profundidad en la cual se pretende mostrar que tan fiable es el uso de zonas plásticas asumidas vs un análisis más detallado

empleando fibras con propiedades de corte, flexión y adherencia.

Una vez realizadas las 200 modelaciones, se concluye que la fiabilidad del modelamiento depende en gran medida de la linealidad que se asume en la zona plástica, y de la ubicación de estas zonas en el modelo, es decir, aunque se empleen zonas plásticas de una alta disipación de energía, como son valores de 0.025 y 0.05, presentarán resultados lineales si se colocan zonas plásticas de baja de disipación de energía en los primeros pisos.
La fiabilidad del modelamiento no se ve significativamente afectada si se emplean o no zonas plásticas para los pisos superiores, o si se combinan zonas plásticas de alta disipación de energía en la estructura, en los primeros pisos.

REFERENCIAS

[1] A. Habibullah, "Practical Three Dimencional Nonlinear Static Pushover Analysis", Structural Magazine, 1998.
[2] ACI 318 – 99 Código de Diseño de Hormigón Estructural, American Concrete Institute, Washington, D.C., 1999.
[3] Aguiar R., Análisis Sísmico por Desempeño, Centro de Investigaciones Científicas Escuela Politécnica del Ejercito, Quito, 2003.
[4] ATC 40, Seismic Evaluation and Retrofit of Concrete Buildings, California Seismic Safety Commission, Volume 1, Report No SSC 96 - 01, California, 1996.
[5] CPE INEN 5:2001, Código Ecuatoriano de la Construcción, Instituto Ecuatoriano de Normalización, Quito – Ecuador, 2001.
[6] FEMA 356, American Society of Civil Engenieeres. Prestandard and Comentary for the Seismic Rehabilitation of Buldings, Federal Emergency Management Agency, Washington.D.C, 2000.
[7] FEMA 273, NEHRP Guidelines for the Seismic Rehabilitation of Buldings, Federal Emergency Management Agency, Washington.D.C, 1997.
[8] J.B. Mander, Theoritical Stress – Strain Model for Confined Concrete, Journal of Structural Engineering, ASCE,Vol 114, No 8, pp 22686, 1988.
[9] Paulay, T., Pristley, M., Seismic Design of Reinforced Concrete and Masonry Structures, John Wiley and Sons, Londres, U.K., 1992.
[10] S. Popovics, A numerical approach to the complete stress strain curve for concrete. Cement and concrete research, 3(5), 583-599, 1973

Data Mart Para El Sistema De Servicios Sociales Del Conadis

Mosquera L.*; Hallo M.**

**Escuela Politécnica Nacional, Facultad de Ingeniería de Sistemas*
Quito, Ecuador; e-mail: eduardo.mosquera; maria.hallo@epn.edu.ec

Resumen: *El Consejo Nacional de Discapacidades Conadis, brinda servicios a nivel nacional y requiere conocer indicadores que demuestren el alcance de los servicios y permitan tomar decisiones para el futuro. Para apoyar este requerimiento se desarrolló un prototipo de data mart, para apoyar al proceso de toma de decisiones. Para el desarrollo del sistema, se utilizó como metodología el ciclo de vida de Ralph Kimball, el proceso cubrió las etapas de planificación, recolección de requerimientos, definición de la arquitectura y selección de herramientas, diseño lógico y físico del data mart, diseño y construcción del proceso ETL, diseño y construcción de aplicaciones de usuario final, despliegue del data mart. Todo el software utilizado para la construcción de los diferentes componentes del data mart son de libre difusión; esto incluye la base de datos, el servidor OLAP, la herramienta para ETL y las aplicaciones de usuario final .Entre las aplicaciones de usuario se destaca la aplicación Mapsic, la cual corresponde a una aplicación GIS en web desarrollada con la finalidad de mejorar la interacción de los usuarios con la información proporcionada por el data mart. La metodología fue útil y el sistema desarrollado resolvió los requerimientos urgentes detectados..*

Palabras clave: *Data mart, Conadis, Data Warehouse, Mapsic*

Abstract*: The National Council for disabilities Conadis, offers services at national level and require to know indicators showing the range of services to take decisions for the future. To solve this requirement a data mart prototype was developed to help in decision making process. The system was developed using the Ralph Kimball Life cycle methodology. The process include planning, requirements analyse, architecture definition, tools selection, logical and physical data mart design, ETL design, end user applications design and implementation, data mart deployment. The different data mart components are developed using free software such us database, OLAP server, ETL tools and end user applications.One of the end user application developed is Mapsi, a Web GIS application to improve the interaction between the user and the data mart information. The methodology was useful and the developed system solved the urgent requirements detected*

Keywords: ***Data mart, Conadis, Data Warehouse, Mapsic****).*

1. INTRODUCCIÓN

Un *data mart* es un subconjunto de datos orientado al análisis de un área del negocio.

La propuesta de un *data mart* para el sistema de servicios sociales del CONADIS(Consejo Nacional de Discapacidades), nace de la necesidad de contar con información oportuna y precisa que respalde los procesos de la organización y ayude a la toma de decisiones en el área de las discapacidades, lo cual fue detectado en entrevistas a personeros del Conadis.

En el mundo del *data warehouse* son múltiples las soluciones orientadas al campo empresarial, negocios y ventas; sin embargo los campos de aplicación para esta tecnología son variados, para el caso de este proyecto, el campo de aplicación es el campo social, buscando una mejor comprensión y comunicación entre la organización y la comunidad.

Existen casos de éxito sobre la aplicación de las tecnologías de *data warehouse* en campos de ámbito social, a continuación se mencionan algunos de estos casos:

- El sistema de información demográfica (SID) desarrollado por el Instituto Nacional de Estadísticas de España, es un *data warehouse* que permite la tabulación detallada de variables demográficas hasta el nivel municipal con datos de estadísticas de variaciones Residenciales, sección censal y unidad poblacional[1].

- El *data mart* del *Health and Hospital Corporation of Marion Country*en Oregón - Estados Unidos, que a través de Internet presenta resultados instantáneos a solicitudes personalizadas de información de salud en el condado de Marion. Proporciona todas las estadísticas locales de salud, tales como

nacimientos, defunciones, enfermedades transmisibles, cáncer e información hospitalaria a través de Internet [2].

- El data warehouse del Instituto Nacional de Empleo en Madrid – España, proyecto desarrollado con el propósito de estudiar grandes volúmenes de información desde distintas perspectivas del mercado de trabajo para la toma de decisiones a nivel táctico y estratégico[3].

En este documento se muestran los aspectos y tareas más relevantes que se llevaron a cabo para la construcción de la propuesta de *data mart* para el sistema de servicios sociales del CONADIS.

2. METODOLOGIA

Para el desarrollo del *data mart* se utilizó la metodología de Ralph Kimball[4] con la siguientes etapas:

- Planificación del Proyecto, Gestión del proyecto
- Definición de Requerimientos
- Arquitectura Técnica, selección e instalación de productos.
- Modelo Dimensional, Diseño físico, Diseño ETL
- Especificaciones de Usuario Final, Desarrollo de aplicaciones
- Despliegue, Implantación, mantenimiento

2.1 Planificación del proyecto

El data mart apoyará a los procesos de la Dirección técnica y del Centro de información del CONADIS, procesos relacionados al sistema web SIC, tales como: registro nacional de discapacidades, solicitud y entrega de ayudas técnicas, solicitud y entrega de medicamentos o insumos, solicitud y autorización de vehículos para personas discapacitadas.

Las aplicaciones desarrolladas permiten realizar consultas interactivas y distribuir en el un mapa indicadores tales como, el número de personas con discapacidades, de tal forma que se puedan realizar programas convenientes de apoyo social de acuerdo a diversos factores.

Para el desarrollo del *data mart* se usó únicamente productos de software libre, manteniendo así los lineamientos gubernamentales para las administraciones públicas del Ecuador[5] .

2.2 Definición de requerimientos

En esta etapa se definen los requerimientos que soportan todas las demás etapas del ciclo de vida del data mart, así: se definen requisitos para la arquitectura y selección de productos, requerimientos de análisis e información para el modelado dimensional, también se determinan especificaciones para las aplicaciones de usuario final.

Para llevar acabo la definición de requerimientos, se recolectó información mediante las siguientes técnicas:

- Entrevistas al personal de la organización
- Análisis de documentos y sistemas fuente

La información obtenida incluye el ambiente de hardware y software bajo el cual trabaja la organización y estructura de los sistemas de información que manejan.

Se analizó una serie de consultas SQL (Structured Query Language) que normalmente se utilizan para obtener información de la base de datos del sistema web SIC, estas consultas generan reportes totalizados de personas discapacitadas, ayudas técnicas, medicamentos.

En base a estas consultas, se detectaron como indicadores claves:

- Número de personas discapacitadas
- Cantidad de ayudas técnicas solicitadas y entregadas
- Cantidad de medicamentos solicitados y entregados
- Cantidad de vehículos autorizados

Se detecto además datos descriptivos bajo los cuales se categoriza y totaliza la información, datos como: provincia, cantón, género, discapacidad, edad, tipo de ayuda técnica, tipo de medicamento, entre otros.

Se analizó la base de datos del sistema web SIC para identificar el origen de la información anteriormente descrita y buscar otros posibles indicadores y datos descriptivos; como resultado se obtuvo los siguientes indicadores adicionales:

- Porcentaje de discapacidad de una persona
- Edad de la discapacidad de una persona
- Monto y subvención de ayudas técnicas
- Monto y subvención de medicamentos

- Cantidad de tiempo que lleva autorizar un vehículo

2.3 Arquitectura técnica del *data mart.*

En esta etapa del ciclo de vida se busca consolidar tres factores:

- Los requerimientos de la organización.
- El ambiente actual de hardware y software.
- Los componentes requeridos para la implementación del *data mart*

Para esta tarea se consideraron dos escenarios:

Back Room.- Encapsula todos los procedimientos y recursos necesarios para llevar a cabo el tratamiento de los datos, básicamente los procesos ETL(Extract, Transform and Load).

Front Room.- Encapsula todos los procedimientos y recursos necesarios para presentar los datos ya tratados al usuario final.

En medio de estos dos escenarios se ubica el servidor de presentación y los metadatos; la figura 2 muestra la ubicación y relación que existe entre los componentes que conforman la arquitectura del *data mart*.

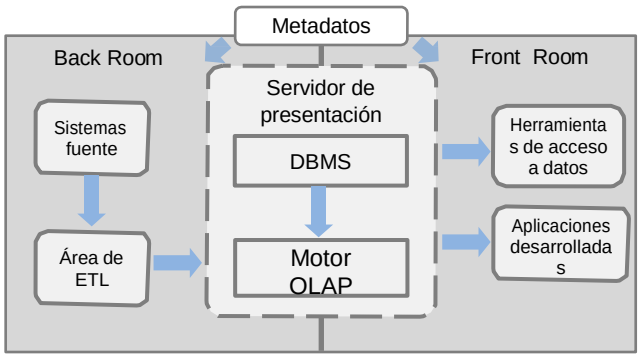


Figura 1. Arquitectura de un *data mart*. Extraído de [6]

La arquitectura presenta dos tipos de componentes fundamentales, servicios y repositorios de datos.

Servicios.- Funciones necesarias para llevar a cabo las tareas requeridas en el *data mart*.

Repositorios de datos.- Lugares temporales o permanentes en los cuales se almacenan los datos.

Los servicios para el *Back Room* son los siguientes:

Servicio de extracción de datos.- Es el encargado de obtener los datos de los sistemas fuente.

Servicio de transformación de datos.- Cambios que sufren los datos necesarios para el rendimiento del *data mart*, entre estas actividades están: generación de claves subrogadas, mapeo de claves de un sistema a otro, desnormalización de tablas, conversión de los tipos de datos, cálculo de valores.

Servicio de carga de datos.- Es el encargado de almacenar los datos en los sistemas destino.

Servicio de control de tareas.- Ayuda a controlar los errores y a manejar las excepciones mediante la definición y programación de tareas, monitoreo y registro de eventos.

Los repositorios de datos para el *Back Room* son los siguientes:

Sistemas fuente.- Fuente de datos que alimenta el *data mart*, sobre este repositorio de datos ocurre el servicio de extracción de datos.

Área de ETL.- Repositorio temporal de datos, aquí tiene lugar el proceso de transformación de datos y el servicio de control de tareas, el cual consiste en la calendarización y monitoreo de los procesos ETL.

DBMS.- Sistema de administración de bases de datos que administra la base de datos del *data mart.*

Los servicios para el *Front Room* son los siguientes:

Servicio de navegación sobre el data mart.- El proceso de navegación a través de la información es uno de los proceso más importantes del *Front Room*, permite al usuario mediante una herramienta del sistema buscar algún dato que le sea importante o necesite saber.

Servicio de acceso y seguridad.- Estos servicios también pueden ser llamados de autorización y autenticación, y ambos se refieren a las conexiones que los usuarios necesitan para acceder a las bases de datos. La autenticación se refiere a alguna técnica para verificar que la persona es quien dice ser, este servicio puede ser controlado por otro sistema.

Servicio de administración de consultas.- La administración de consultas le permite al usuario una interacción con las bases de datos.

Servicio de reportes estándar.- El servicio de reportes estándar le permite al usuario generar reportes con formatos estándar.

Los repositorios de datos para el *Front Room* son los siguientes:

Motor OLAP.- Genera los cubos y dimensiones con la información almacenada en la base de datos del *data mart* y procesa las consultas analíticas, responde al

Servicio de administración de consultas MDX (Multidimensional Expressions).

Herramientas de acceso a datos.- Herramientas que permiten a los usuarios interactuar con los datos almacenados en el *data mart*, mediante consultas MDX y SQL, además deben proporcionar los servicios de navegación sobre el *data mart* y servicio de reportes estándar.

Aplicaciones desarrolladas.- Para este caso se planteó la construcción de una aplicación web que permita al usuario interactuar con los datos almacenados en el *data mart* mediante consultas SQL, valiéndose de un componente GIS (Geographic Information System) para representar la información de manera gráfica en un mapa.

Los metadatos del *Back Room* están relacionados a los procesos ETL, guían la extracción, limpieza y carga de datos; por lo tanto entre estos metadatos esta la información acerca de los sistemas fuente y el área ETL.

Los metadatos del *Front Room* son más descriptivos, ayudan a las herramientas de consulta, nos permite entender de donde vienen los datos; estos metadatos cubren información acerca del *data mart* y herramientas de acceso a datos.

Selección de productos

En esta etapa del ciclo de vida, se busca elegir herramientas y aplicaciones de software apropiadas para la implementación de la arquitectura propuesta.

Los productos seleccionados son de software libre, para seguir los lineamientos de la organización y además satisfacen los servicios tanto para el *Back Room* como para el *Front Room*.

La tabla 1 resume los productos de software elegidos para cada componente de la arquitectura del Data mart.

Tabla 1. Productos seleccionados para el Data mart	
Back Room	
DBMS	PostgreSQL 8.4
Área ETL	Pentaho data intregation – Kettle 3.2
Front Room	
Motor OLAP	Mondrian 3.1.6
Acceso a Datos	Openi 2.0 RC2
Otros productos necesarios	
Servidor de aplicaciones java	Tomcat 6.0
Servidor web	Apache 2.2
Servidor GIS	Mapserver 3.0

2.4 Modelado dimensional

Cada modelo dimensional esta compuesto por una tabla cuya clave es compuesta, llamada tabla de hechos, y un grupo de tablas pequeñas llamadas tablas de dimensiones.

Cada tabla de dimensión tiene una clave primaria singular que corresponde exactamente con uno de los componentes de la clave compuesta de la tabla de hechos, esta estructura adopta el nombre de modelo en estrella.

En base a la información descriptiva de las consultas SQL analizadas en la etapa de requerimientos, se determinaron las dimensiones agrupando los atributos descriptivos que están relacionados[7]. Se definieron un total de 13 dimensiones.

Para la conformación de las tablas de hechos se agruparon las medidas asociadas a los indicadores claves detectados en la etapa de requerimientos, la agrupación se realizó en función del proceso asociado al sistema web SIC. Se determinó las siguientes tablas de hechos:

- Tabla de hechos de personas discapacitadas
- Tabla de hechos de ayudas (agrupa ayudas técnicas y medicamentos)
- Tabla de hechos de autorización de vehículos

Para cada tabla de hechos se determinó la granularidad, se seleccionaron las dimensiones y medidas. Las figuras 3, 4 y 5 muestran los diagramas de las Tablas de hechos conformadas[8].

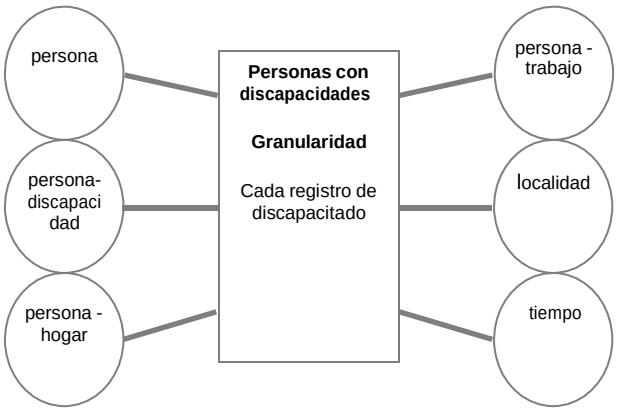


Figura 2. Diagrama de la tabla de hechos de personas con discapacidades.

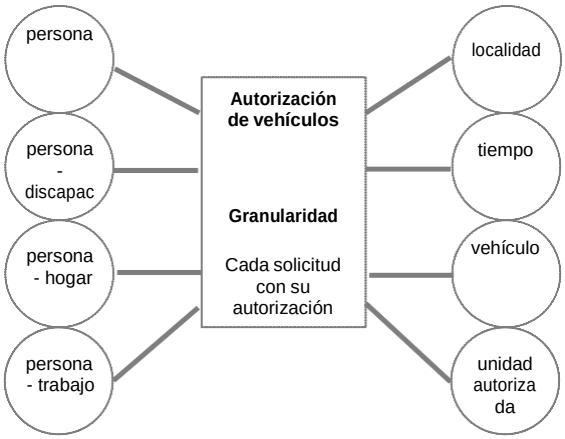


Figura 3. Diagrama de la tabla de hechos de autorización de vehículos.

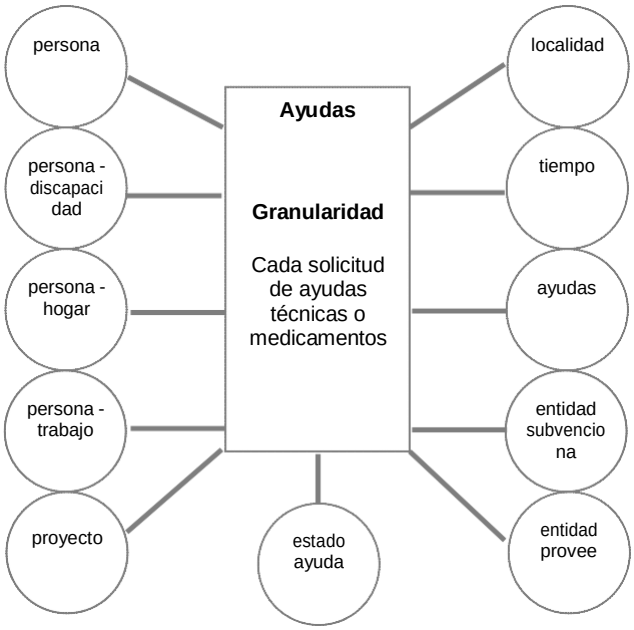


Figura 4. Diagrama de la tabla de hechos de ayudas.

2.5 Diseño físico

En esta etapa del ciclo de vida se contemplan los pasos para convertir el diseño lógico en una base de datos física, las dimensiones y tablas de hechos se definen como tablas de la base de datos.

Como primera tarea se elaboró estándares para nombrar los objetos de la base de datos y del ETL.

Mediante el uso de una herramienta de modelado de datos, se paso el modelo lógico a un modelo físico respetando los estándares definidos para los objetos de la base de datos, tipos de datos, etc.

La herramienta permitió generar el código necesario para crear la base de datos sobre PostgreSQL.

Como segunda tarea se elaboró el plan de índices, este plan consideró los siguientes índices:

Índices para las Tablas de hechos.- Corresponde a un índice compuesto por todas las claves dimensionales de la Tabla de hechos, también se definieron índices para cada clave dimensional por separado. Estos índices tienen como propósito mejorar el rendimiento de las consultas, el planificador de consultas del DBMS es el encargado de utilizar convenientemente estos índices.

Índices para las tablas de Dimensiones.- Se creo un índice para la clave subrogada de la tabla, índices para los atributos e índices compuestos para las jerarquías entre atributos utilizando las claves primarias asociadas tomadas de la base de datos fuente. Estos índices tienen como propósito mejorar el rendimiento de las consultas, especialmente cuando se declaran filtros y ordenamientos.

Índices para el proceso ETL.- Para optimizar la carga y actualización de datos en las tablas de Dimensiones y especialmente en las Tablas de hechos, se definieron índices simples y compuestos por las claves operacionales en las tablas de Dimensiones para mejorar el proceso de obtención de claves subrogadas.

Para asegurar el correcto funcionamiento del planificador de consultas, los índices y estadísticas deben estar actualizados, esta operación se definió posteriormente como parte del proceso ETL.

Como tercera y última tarea se construyeron los cubos OLAP (Online Analytical Processing), Como fue determinado en la selección de productos, el motor OLAP a ser utilizado es Mondrian, por lo tanto la construcción de cubos se la realizó mediante el esquema de etiquetas XML propuesto por Pentaho – Mondrian. La figura 6 muestra la estructura básica del esquema[9].

```
<Schema>
    <Cube>
        <Dimension>
            <Hierarchy>
                <Level/>
            </Hierarchy>
        </Dimensión>
        <DimensionUsage/>
        <Measure />
    </Cube>
</Schema>
```

Figura 5. Estructura básica del esquema Mondrian

2.6 Diseño y construcción de los ETLs

En esta etapa se desarrolló el *Back Room*, especialmente en el área ETL donde se definen los procesos de extracción, limpieza, transformación y carga de datos desde la base de datos fuente hacia la base de datos del *data mart*.

Como primera tarea se elaboró el plan ETL[10], este plan detalla la forma como se realizó la carga de datos desde los sistemas fuente hacia el *data mart*. Se plantearon tres procesos y son los siguientes:

Carga inicial de datos.- Este proceso carga datos sobre el *data mart* por primera vez, el primer trabajo consiste en la creación de la base de datos, el segundo trabajo consiste en llenar la dimensión tiempo, como tercer paso se ejecutan los trabajos para llenar el resto de dimensiones y finalmente se ejecutan los trabajos para llenar las tablas de hechos.

Debido a que se trata de la carga inicial, los registros que se cargan a las tablas de hechos y dimensiones comprenden desde los más antiguos hasta los ingresados en el sistema operacional la fecha en que se ejecuta la carga inicial.

Carga incremental de datos.- Este proceso se ejecuta periódicamente y permite mantener actualizadas las dimensiones cargando nuevos registros, así como actualizando los datos ya registrados de acuerdo a la política de cambio que se planteo para cada atributo.

En el caso de las tablas de hechos, se cargan los nuevos registros comprendidos entre la última ejecución de la carga incremental o inicial hasta la fecha anterior a la nueva ejecución.

Actualización de datos.- Este proceso permite actualizar las dimensiones cargando nuevos registros así como actualizando los antiguos.

En el caso de las tablas de hechos, aunque es poco común realizar actualizaciones sobre los registros, este proceso permite actualizar los registros comprendidos entre dos fechas, una inicial y una final suministradas por el usuario.

Como siguiente tarea se determinó el flujo de datos base tanto para dimensiones como tablas de hechos, este flujo muestra las etapas por las cuales pasan los datos desde que son extraídos del sistema operacional hasta que son cargados al *data mart*. También se identificaron las transformaciones y cálculo de datos, muchas de estas operaciones se realizan como parte de la extracción de datos.

El flujo de datos base para la carga y actualización de las dimensiones[11]consta de los siguientes pasos:

1) *Extracción de datos de las tablas origen:* Mediante código SQL se seleccionan las columnas que proveen los datos, los nombres de las columnas se las redefine conforme al nombre de las columnas de la tabla de dimensión. En este paso se lleva acabo una primera parte de las trasformaciones.

2) *Transformaciones y cálculos:* Los registros entran a este paso y se llevaron a cabo las transformaciones y cálculos definidos, de ser necesario este paso añade columnas a los registros.

3) *Asignación de la clave subrogada:* En base a las claves operacionales se asigna un número secuencial único al registro, esta es la clave subrogada.

4) *Carga y actualización de datos:* Los registros del paso anterior son cargados a la tabla de dimensión, los registros nuevos son insertados mientras que los registros alterados son actualizados.

El flujo de datos base para la carga y actualización de las tablas de hechos consta de los siguientes pasos:

1) *Extracción de datos de las tablas origen:* Mediante código SQL se seleccionan las claves operacionales correspondientes a las dimensiones asociadas a la tabla de hechos, se seleccionan las columnas que proveen los datos para las medidas, los nombres de estas columnas se las redefine para que coincidan con las medidas de la tabla de hechos. En este paso se lleva acabo una primera parte de las trasformaciones esencialmente el cálculo de algunas medidas. Los registros seleccionados están dentro de un rango de fechas, las cuales dependen del tipo de proceso ETL.

2) *Transformaciones y cálculos:* Se llevaron a cabo las transformaciones y cálculos definidos, de ser necesario este paso añade columnas a los registros y realiza validaciones para comprobar la integridad de los cálculos, en caso de existir datos incoherentes estos son exportados a un archivo para su posterior análisis.

3) *Obtención de claves dimensionales:* Se busca remplazar las claves operacionales con sus respectivas claves dimensionales, las cuales se añaden como nuevas columnas a los registros. Las claves operacionales son removidas de los registros.

4) *Validación de claves:* Las claves obtenidas en el paso anterior son sometidas a una verificación para asegurar que no hayan datos nulos, en caso de hallarse registros con claves nulas estos son exportados a un archivo para su posterior análisis. Los registros que pasan la verificación son ordenados en base a las claves dimensionales.

5) *Extracción de datos de la tabla de hechos:* Mediante código SQL se seleccionan todas las columnas de la tabla y se las ordena por las claves dimensionales, los registros seleccionados están dentro de un rango de fechas, las cuales dependen del tipo de proceso ETL.

6) *Comparación y Marcado de filas:* Los registros de los pasos 4 y 5 son comparados en base a las claves dimensionales y dependiendo del resultado de la comparación se marca cada registro como: idéntico, alterado, nuevo, borrado; esta información se almacena en una nueva columna la cual se añade a los registros.

7) *Actualización de Registros:* Los registros del paso anterior son conmutados dependiendo de la marca, los nuevos son insertados, los alterados son actualizados y los borrados son eliminados de la tabla de hechos.

2.7 Construcción de las aplicaciones de usuario

Al iniciar esta etapa del ciclo de vida, nos concentraremos en el *Front Room*.

La principal actividad de desarrollo que se lleva acabo en el *Front Room* es la creación de un primer grupo de reportes, los cuales adoptan el nombre de plantillas de la aplicación de usuario [10], un segundo esfuerzo esta enfocado en la creación de la aplicación Mapsic.

El usuario interactúa con el Data mart mediante los elementos del *Front Room*, es decir mediante herramientas de acceso a datos y aplicaciones desarrolladas sobre el Data mart. Como fue determinado en la selección de productos para el *Front Room*, la herramienta de acceso a datos es el sistema de reportes web Openi, adicionalmente se planteo el desarrollo de una pequeña aplicación GIS para realizar consultas interactivas.

Como primera tarea se elaboraron las plantillas de la aplicación de usuario, estos reportes fueron construidos tomando como referencia los reportes observados en la sección de estadísticas del sitio web del CONADIS, se definieron los siguientes reportes como plantillas iniciales:

Análisis cruzado

- Causa de deficiencia por edades
- Distribución local de PCD (personas con discapacidades) por situación laboral
- Distribución local de ayudas técnicas

Distribución por categoría

- Distribución de PCD por deficiencia principal
- Tiempo promedio para autorizar un vehículo

Promedio por categoría

- Porcentaje de PCD por género

Tendencias por categoría

- PCD registradas por año
- Monto total en ayudas entregadas al año
- Vehículos autorizados al año

Como segunda tarea se desarrolló la aplicación Mapsic, esta aplicación web tiene como propósito permitir a los usuarios realizar consultas interactivas, obtener información sobre el número y porcentaje de personas discapacitadas y su distribución provincial y cantonal en el Ecuador.

Para el desarrollo se siguieron las etapas de requerimientos, análisis, diseño, construcción y pruebas, de entre estas etapas se resaltó el diseño de la arquitectura, que corresponde a un diseño en cuatro capas. La tabla 2 muestra en detalle en que consiste cada capa.

Tabla 2. Capas de la aplicación - Mapsic	
Capa	Descripción
Presentación	Interfaz de usuario estilo web, utilizando xhtml, css y javascript. Presenta ayudas visuales tales como: resaltado de filas y columnas, ordenamiento por columnas, en los reportes. Paneles expandibles e indicadores de carga.
Comunicación	Las llamadas al servidor son asincrónicas, el refrescamiento del contenido de la pagina web es mediante Ajax.
Lógica	Procedimientos para la generación de reportes, procedimientos para la generación de mapas (GIS). Definición de constantes y variables de sesión.
Datos	Procedimientos para acceder y realizar consultas a bases de datos. Procedimientos para acceder a shape files (mapas vectoriales). Definición de metadatos

2.8 Despliegue

El despliegue es la convergencia de tecnología, datos y aplicaciones; es la integración de los componentes que se han ido desarrollando en el transcurso del ciclo de vida del *data mart*, con el fin de que los usuario puedan empezar a sacar provecho.

Esta etapa también contempla capacitación para los usuarios y la preparación del ambiente para poder integrar cada uno de los componentes del *data mart*.

Entre las tareas llevadas a cabo en esta etapa final está la elaboración y ejecución de pruebas para validar el funcionamiento de los componentes del *data mart*, como son los procesos ETL, los reportes iniciales y la aplicación Mapsic.

3. CONCLUSIONES

El ciclo de vida propuesto por Kimball facilita las tareas de desarrollo de un data warehouse, en este caso las del *data mart*, define los productos que se deben obtener tras finalizar cada etapa, además de presentar plantillas, formatos y diagramas para la documentación.

El *data mart* mejora el trabajo de los usuarios poniendo a su disposición la información que necesitan cuando la necesitan, por lo tanto es importante seleccionar herramientas apropiadas que se ajusten a las necesidades de la organización.

Una de las actividades más importantes dentro del desarrollo del *data mart* es la definición del proceso ETL, especialmente el aseguramiento de la calidad de los datos, por lo tanto es importante contar con sistemas fuentes confiables que validen y controlen el registro de datos, estos mecanismos facilitan en gran medida el tratamiento de los datos.

El esfuerzo aplicado en estructurar consultas SQL para el análisis de datos es alto, especialmente en bases de datos transaccionales esta tarea es compleja; ahora mediante el *data mart* el cual presenta una base de datos apropiada para realizar análisis y consultar grandes cantidades de datos, el estructurar consultas en SQL involucra menos esfuerzo, sin embargo es más eficaz utilizar MDX mediante los cubos OLAP.

El trabajo ha permitido elaborar y validar una solución de apoyo a la toma de decisiones para el CONADIS utilizando software libre.

REFERENCIAS

[1] Plan de actuación del Instituto Nacional de Estadísticas para el año 2009,http://www.mzes.uni-mannheim.de/eda/e/series/2242129-4/2009.pdf (accesado noviembre 2013).

[2] Missouri Department of Health,World Access to Indy's Health, http://www.hhcdatamart.com/ (accesado julio 2013).

[3] INEM,Data warehouse del Instituto Nacional de Empleo en Madrid, http://foil.oit.or.cr/documentos/index.php?option=com_k2&view=foildoc&id=224:sistema-de-informacion-data-warehouse&Itemid=55 (accesado julio 2013)

[4] KIMBALL Ralph, *et al*. The Data Warehouse Lifecycle Toolkit. 1a ed. New York: Wiley. 1998. Capítulo 2, p. 33 – 36.

[5] CORREA Rafael, Decreto 1014, Registro Oficial 322, Quito: Editora Nacional,2008.

[6] KIMBALL Ralph, *et al*. The Data Warehouse Lifecycle Toolkit. 1a ed. New York: Wiley. 2001. Capítulo 8, p. 327 – 330, Capítulo 9, p. 336 – 364, Capítulo 10, p. 375 – 386.

[7] KIMBALL Ralph, ROSS Margy. The Data Warehouse Toolkit. 2a ed. New York: Wiley. 2002. Capítulo 6, p. 154-157.

[8] KIMBALL Ralph, *et al*. The Data Warehouse Lifecycle Toolkit. 1a ed. New York: Wiley. 1998. Capítulo 5, p. 194 – 197.

[9] Mondrian Documentation, How to Design a Mondrian Schema. http://mondrian.pentaho.org/documentation/schema.php. 2011-06-24.

[10] KIMBALL Ralph, *et al*. The Data Warehouse Lifecycle Toolkit. 1a ed. New York: Wiley. 1998. Capítulo 16, p. 612 – 615.

[11] KIMBALL Ralph, CASERTA Joe. The Data Warehouse ETL Toolkit. 1a ed. New York: Wiley. 2004. Capítulo 5.

Accesibilidad web en las Universidades del Ecuador. Análisis preliminar

Navarrete R. *; Luján S.**

**Escuela Politécnica Nacional, Facultad de Ingeniería de Sistemas, Departamento de Informática y Ciencias de la Computación*
Quito, Ecuador (Tel: 593-2-2507-144; e-mail: rosa.navarrete@epn.edu.ec)
***Universidad de Alicante, Departamento de Lenguajes y Sistemas Informáticos*
Alicante, España (e-mail: sergio.lujan@ua.es)

Resumen: *El interés de la accesibilidad web en Ecuador se enmarca en la declaración del Plan Nacional para el Buen Vivir, 2013- 2017, que en el segundo objetivo promueve la igualdad, inclusión y equidad social, además del interés del Estado en procurar un país sin barreras, explicitado en la Ley Orgánica de Discapacidades. Para alinearse con esta estrategia estatal, las universidades públicas deben procurar el cumplimiento de las normativas sobre accesibilidad web en sus portales web institucionales. En este artículo se verifica, a través de herramientas de software, el nivel de cumplimiento de las directrices sobre accesibilidad web y la observancia de la sintaxis de lenguajes de marcado y hojas de estilo, en los portales web de las principales universidades del Ecuador. A partir de los resultados del análisis se podrá establecersi es necesario mejorar el diseño de los portales web de las universidades.*

Palabras clave: *accesibilidad web; portales web universitarios; WCAG; TAW.*

Abstract: *The interest about web accessibility in Ecuador is part of the declaration of the National Plan for Good Living, 2013 - 2017, the second objective that promotes equality, inclusion and social equity, and the state's interest in procure a country without barriers, explained in the Disabilities Act. To align with this strategy state public universities must ensure compliance with web accessibility standards on their institutional web portals. In this paper we verify, through software tools, compliance with web accessibility guidelines and adherence to the syntax of markup languages and style sheets, web sites of the major universities of Ecuador.From the analysis results will be established if necessary to improve the design of university web sites.*

Keywords: *web accessibility; university web portals; university; WCAG; TAW*

1. INTRODUCCION

En el Ecuador, la Ley Orgánica de Discapacidades[1], recoge como uno de sus principios rectores, "Accesibilidad: se garantiza el acceso de las personas con discapacidad al entorno físico, al transporte, la información y las comunicaciones, incluidos los sistemas y las tecnologías de información y las comunicaciones, y a otros servicios abiertos al público o de uso público*"*.

Adicionalmente, la ley mencionada establece en el "**Art. 65.-** Atención prioritaria en portales web.- Las instituciones públicas y privadas que prestan servicios públicos, incluirán en sus portales web, un enlace de acceso para las personas con discapacidad, de manera que accedan a información y atención especializada y prioritaria, en los términos que establezca el reglamento".

En el mismo contexto, el "**Art. 33**, Accesibilidad a la educación.- La autoridad educativa nacional en el marco de su competencia, vigilará y supervisará, que las instituciones educativas escolarizadas y no escolarizadas, especial y de educación superior, públicas y privadas, cuenten con infraestructura, diseño universal, adaptaciones físicas, ayudas técnicas y tecnológicas para las personas con discapacidad".

El Plan Nacional para elBuen vivir 2013-2017 [9], recoge como objetivo, "Auspiciar la igualdad, la cohesión, la inclusión y la equidad social y territorial, en la diversidad"; tales enunciados se promueven también con la accesibilidad web.

En base al marco legal vigente se podría interpretar que las universidades públicas, cuyo presupuesto lo financia el Estado, al ofrecer servicios de educación superior, están obligadas a garantizar la accesibilidad web en sus portales web institucionales.

Sin embargo, la existencia de un marco legal no garantiza que se cumplan los principios de accesibilidad web. No se conocen estudios que refieran el estado de cumplimiento de la legislación sobre accesibilidad web en los portales web de universidades públicas del país. La aportación de este trabajo, es ofrecer un análisis de la accesibilidad de los portales web de las universidades públicas de las categorías A y Bestablecidaspor el Consejo Nacional de Evaluación y Acreditación de la Educación Superior de Ecuador. Los resultados de estudio podrán ayudar a las instancias de decisión a tomar las medidas correctivas necesarias para el cumplimiento de los objetivos señalados en el Plan Nacional para elBuen Vivir y en la Ley Orgánica de discapacidades.

El artículo concluye con algunas sugerencias para que se logren esos objetivos.

2. ACCESIBILIDAD WEB

La accesibilidad web en su sentido más amplio, se concibe como un acceso universal a la Web, independientemente del hardware, software, lenguaje, cultura, localización geográfica o de las capacidades físicas o mentales de los usuarios [11].
La meta de la accesibilidad web es asegurar que la información o servicios entregados a través de sitos web estén disponibles y puedan ser utilizados por la mayor audiencia posible.
El organismo Web AccessibilityInitiative, WAI[17], perteneciente al World Wide Web Consortium W3C, a través del Grupo de trabajo sobre Guías para el Contenido web accesible, (Web Content AccesibilityGuidelinesWorkingGroup, WCAG WG), ha propuesto un conjunto de normas llamadas Pautas de Accesibilidad al Contenido en la Web (Web Content AccessibilityGuidelines, WCAG). El marco legal vigente en Ecuador no determina los requisitos técnicos de accesibilidad web de cumplimiento obligatorio, pero dado que las WCAG se han adoptado como estándares de facto y han sido recogidas en la legislación de un número importante de países[7], serán las que se utilicen como marco de referencia para este trabajo.
El reto de desarrollar contenido web accesible ha motivado la evolución de un conjunto de técnicas para abordar los problemas de la accesibilidad web. A continuación se citan brevemente las guías desarrolladas por la W3C. La primera, WCAG 1.0 de 1999 y la segunda, WCAG 2.0, de 2008.

2.1 Guía WCAG 1.0

Es una recomendación del 5 de mayo de 1999. Presenta 14 pautas que resumen los principios fundamentales del diseño web accesible[19]:

1. Proporcione alternativas equivalentes para el contenido visual y auditivo.
2. No se base sólo en el color.
3. Utilice marcadores y hojas de estilo y hágalo apropiadamente.
4. Identifique el idioma usado.
5. Cree tablas que se transformen correctamente.
6. Asegúrese de que las páginas que incorporen nuevas tecnologías se transformen correctamente.
7. Asegure al usuario el control sobre los cambios de los contenidos tempo-dependientes.
8. Asegure la accesibilidad directa de las interfaces incrustadas.
9. Diseñe para la independencia del dispositivo.
10. Utilice soluciones provisionales.
11. Utilice las tecnologías y pautas W3C.
12. Proporcione información de contexto y orientación.
13. Proporcione mecanismos claros de navegación.
14. Asegúrese de que los documentos sean claros y simples.

La accesibilidad no se puede conseguir sin el dominio de las pautas, sin saber cómo utilizan la web los usuarios con discapacidad y, fundamentalmente, sin un manejo de los lenguajes y estándares que hacen funcionar la web[2].
Las guías presentadas han sido categorizadas en tres niveles de prioridad basadas en su impacto en la accesibilidad.

Prioridad 1

Un desarrollador de contenidos Web debe satisfacer estenivel de verificación, en caso contrario la información no será accesible a uno o más grupos de usuarios.

Prioridad 2

Un desarrollador de contenidos Web debería satisfacer estenivel de verificación, en caso contrario, uno o más grupos de usuario encontrarán dificultad en acceder a la información de modo satisfactorio.

Prioridad 3

Un desarrollador de contenidos Web podría satisfacer este nivelde verificación para mejorar el acceso de uno o más grupos de usuario.

En concordancia con estas prioridades, se tienen tres niveles de conformidad en la evaluación de accesibilidad:

Nivel A. Todas las verificaciones de prioridad 1, han sido satisfechas.

Nivel AA. Todas las verificaciones de prioridad 1 y de prioridad 2, han sido satisfechas.

Nivel AAA. Todas las verificaciones de prioridad 1, prioridad 2 y prioridad 3 han sido satisfechas.

Una descripción detallada de la verificación de las pautas de accesibilidad puede consultarse en el documento oficial de la W3C, sobre la guía WCAG 1.0[12].

2.2 Guía WCAG 2.0

Recomendación publicada 11 de diciembre de 2008. Éstas fueron aprobadas como estándar internacional ISO/IEC 40500:2012, el 15 de octubre de 2012[5].
Las WCAG 2.0 mejoran el estándar inicial creado por el W3C, aplicándose a tecnologías más avanzadas, siendo más flexibles, detallando de forma más precisa cuáles son las comprobaciones a realizar para probar la calidad de su aplicación e incorporando una mejor documentación para facilitar su comprensión.
A diferencia de las WCAG 1.0 desarrolladas exclusivamente para las tecnologías web que existían en el año 1999 (HTML, CSS y JavaScript), las WCAG 2.0 se desarrollaron de un modo tecnológicamente neutro. Sin embargo, también se desarrollaron las técnicas suficientes que proporcionan orientación y ejemplos para cumplir las pautas usando tecnologías específicas. Se compone de 4 principios, 12 directrices y 61 criterios de cumplimiento, además de un conjunto de técnicas suficientes y técnicas de asesoramiento [18]. Los principios son los siguientes:

Principio 1: Perceptibilidad

La información y los componentes de la interfaz de usuario deben presentarse a los usuarios de la manera en que puedan

percibirlos. Se compone de 4 directrices y 22 criterios de cumplimiento.

Principio 2: Operabilidad

Los componentes de la interfaz de usuario y la navegación deben ser operables. Se compone de 4 directrices y 20 criterios de cumplimiento.

Principio 3: Comprensibilidad

La información y el manejo de la interfaz de usuario deben ser comprensibles. Se compone de 3 directrices y 17 criterios de cumplimiento.

Principio 3: Robustez

El contenido debe ser lo suficientemente robusto como para confiarse en su interpretación por parte de una amplia variedad de agentes de usuario, incluidas las tecnologías de asistencia. Se compone de 1 directriz y 2 criterios de cumplimiento.

Una descripción detallada de la aplicación de las directrices y criterios de cumplimiento de cada principio puede verse en el sitio oficial de la W3C, sobre la guía WCAG 2.0[16].

3. OBJETIVOS DE LA INVESTIGACIÓN

A. Objetivo general

Establecer el nivel de accesibilidad web en los portales web institucionales de las Universidades de Categoría A y B del Ecuador, a través del análisis de cumplimiento de los estándares de facto sobre accesibilidad web, WCAG 1.0 y WCAG 2.0.

B. Objetivos específicos

Verificar en los portales web analizados, el nivel de conformidad de la accesibilidad, utilizando las guías WCAG 1.0 y WCAG 2.0.
Verificar el cumplimiento de la sintaxis del lenguaje de marcado y de las hojas de estilo.
Identificar los problemas de accesibilidad más comunes en los portales web analizados.

4. PROCEDIMIENTO DE LA INVESTIGACIÓN

A. Selección de las Universidades para el análisis

En Ecuador existen 71 universidades reconocidas por el Consejo de Educación Superior, CES: 29 universidades públicas, 9 universidades particulares cofinanciadas y 33 universidades particulares[3]. Las universidades se encuentran calificadas en 4 categorías, de acuerdo a la última evaluación realizada por el Consejo Nacional de Evaluación y Acreditación de la Educación Superior, CONEA entre el 2010-2011. La categoría A corresponde a las universidades más destacadas, con un total de 14 universidades, de las cuales 9 son públicas; de éstas, 6 ofrecen carreras de pregrado y postgrado, y 3 ofertan exclusivamente programas de postgrado. La categoría B corresponde a las universidades con la subsiguiente evaluación, con un total de 10 universidades, de las cuales 6 son públicas y ofrecen carreras de pregrado y postgrado.

En razón de que la accesibilidad es una política del Estado, se ha considerado para esta investigación a las universidades públicas, cuyo presupuesto lo financia el Estado, de las primeras categorías: A y B, que no ofertan únicamente programas de postgrado; a fin de tener la misma base de comparación.

Las universidades seleccionadas para el análisis se muestran en la Tabla 1.

B. Definición de las herramientas de análisis

Las herramientas automáticas de evaluación de la accesibilidad web son programas de software o servicios en línea, que realizan una inspección automática para ayudar a determinar si un portal web satisface las guías de accesibilidad web.
Existen numerosas herramientas automáticas para la evaluación de la accesibilidad web [13], pero en consideración a algunos criterios de selección[4], se ha optado por la herramienta automatizada TAW [10]. Esta herramienta automática de evaluación de la accesibilidad web desarrollada por la Fundación CTIC (Centro Tecnológico de la Información y la Comunicación), sede del W3C en España, está en línea desde abril de 2001 y fue el primer verificador de la accesibilidad web en español.
TAW evalúa de forma automática el nivel de conformidad de las pautas de la WCAG 1.0 y WCAG 2.0. Genera un reporte de salida en el que aparecen los resultados del análisis, indicando el número de problemas detectados, en relación al nivel de prioridad del criterio de verificación correspondiente a la guía con la que se evalúa. Estos informes de evaluación no son definitivos, pero constituyen una primera aproximación para efectos de este trabajo. Solamente un análisis exhaustivo por parte de un experto puede ofrecer un análisis completo y fiable de la accesibilidad web de una página[6].
Para complementar el análisis se utilizarán herramientas de la W3C, que verifican la correcta sintaxis del lenguaje utilizado en las páginas web: W3C MarkupValidationService[15] para el lenguaje de marcado y W3C CSS ValidationService[14], para las hojas de estilo.

C. Aplicación de las herramientas de análisis

De cada una de las universidades seleccionadas se evalúan 6 páginas comunes a todos los portales web y que corresponden a las siguientes:

1. *Home Page, Presentación, Inicio*. Permite el acceso al resto de la información contenida en el portal web. Si la primera página no es accesible, es difícil acceder al resto de páginas del portal web.
2. *Oferta Académica, Carreras.* Información sobre las carreras de pregrado y postgrado ofertadas por la Universidad.
3. *Carrera de pregrado.* Una página descriptiva de alguna carrera.
4. *Programa de postgrado.* Una página descriptiva de algún programa o maestría.

5. *Investigación*. Información sobre las investigaciones que desarrolla la Universidad. Una página representativa de investigación.
6. *Vinculación con la colectividad*. Servicios que la Universidad presta a la colectividad. Una página representativa de algún proyecto o servicio de relación con la colectividad.

Se utiliza el TAW en línea, para la validación del cumplimiento de WCAG 1.0, con el URL de cada una de las respectivas páginas de las universidades en análisis. Se configura el análisis para los niveles de conformidad A y AA, dado que son los exigibles; el nivel AAA es deseable, por lo que no se considera.

A continuación se ha aplicado el TAW en línea, para la validación del cumplimiento de WCAG 2.0, en las mismas páginas a las que se aplicó la validación de WCAG 1.0, de cada universidad. El resultado reporta el número de errores, advertencias y elementos no verificados, para cada uno de los principios: Perceptible, Operable, Comprensible y Robusto.

Por último, se realiza la verificación de sintaxis, a través de herramientas en línea: W3C MarkupValidationService para el lenguaje de marcado y W3C CSS ValidationService, para las hojas de estilo.

Estas herramientas reportan un número de errores y validan el cumplimiento con respecto de la sintaxis de los lenguaje de marcado y hojas de estilo.

5. RESULTADOS Y DISCUSIÓN

5.1 Universidades analizadas

Las universidades elegidas para el análisis se muestran en la Tabla 1 y Tabla 2.

Tabla 1. Universidades Categoría A

UNIVERSIDAD		URL PORTAL WEB
EPN	ESCUELA POLITECNICA NACIONAL	http://www.epn.edu.ec
ESPE	ESCUELA POLITÉCNICA DEL EJÉRCITO	http://www.espe.edu.ec
ESPOL	ESCUELA POLITÉCNICA DEL LITORAL	http://www.espol.edu.ec
ESPOCH	ESCUELA POLITÉCNICA DEL CHIMBORAZO	http://www.espoch.edu.ec
UCE	UNIVERSIDAD CENTRAL DEL ECUADOR	http://www.uce.edu.ec
UC	UNIVERSIDAD DE CUENCA	http://www.ucuenca.edu.ec

Tabla 2. Universidades Categoría B

UNIVERSIDAD		URL PORTAL WEB
UAE	UNIVERSIDAD AGRARIA DEL ECUADOR	http://www.uagraria.edu.ec/
UEB	UNIVERSIDAD ESTATAL DE BOLÍVAR	http://www.ueb.edu.ec/
UNACH	UNIVERSIDAD NACIONAL DE CHIMBORAZO	http://www.unach.edu.ec/
UNL	UNIVERSIDAD NACIONAL DE LOJA	http://www.unl.edu.ec/
UTN	UNIVERSIDAD TÉCNICA DEL NORTE	http://www.utn.edu.ec/
UG	UNIVERSIDAD DE GUAYAQUIL	http://www.ug.edu.ec/

5.2 Verificación de la Accesibilidad con WCAG 1.0

Se analiza el cumplimiento del nivel de conformidad AA, que es exigible. Además, es el que ha adoptado enla mayoría de los países,en la legislación pertinente a accesibilidad web.

Resultados de análisis WCAG 1.0 en todas las páginas

En la Tabla 3 y Tabla 4, se muestran los resultados de la aplicación de la herramienta TAW, con relación a la WCAG 1.0, detección de conformidad de nivel AA.

Tabla 3. Resultados de análisis WCAG 1.0 Universidades Categoría A

REFERENCIA	EPN	ESPE	ESPOL	ESPOCH	UCE	UC	TOTAL	%
Total de páginas analizadas en el portal web	6	6	6	6	6	6	36	100
Número de páginas que cumplen con el nivel A	4	1	2	1	4	0	12	33.33
Número de páginas que cumplen con el nivel AA	0	0	0	0	0	0	0	0

Tabla 4. Resultados de análisis WCAG 1.0 Universidades Categoría B

REFERENCIA	UAE	UEB	UNACH	UNL	UTN	UG	TOTAL	%
Total de páginas analizadas en el portal web	6	(*)5	(*)5	6	6	6	33	100
Número de páginas que cumplen con el nivel A	1	5	1	0	5	2	14	42.42
Número de páginas que cumplen con el nivel AA	0	0	0	0	0	0	0	0

(*) En las universidades señaladas, no se encontraron páginas web informativas sobre las carreras, únicamente archivos descargables.

Los resultados de la Tabla 3 y Tabla 4 no muestran una diferencia significativa en los resultados obtenidos en los dos grupos de universidades. De las páginas analizadas, ninguno de los portales web, cumple con el nivel de conformidad A, mínimo exigible. Es decir hay incumplimiento en puntos de control de prioridad 1. Los errores más frecuentes son: no se utilizan atributos para proporcionar un texto equivalente para elementos no textuales; no se incluyen encabezamientos en tablas; no se organiza el documento HTML para que pueda interpretarse sin asociarlo a la hoja de estilo.

No se observa conformidad de nivel AA, que verifica los puntos de control de prioridad 2. Los errores más frecuentes detectados se relacionan a fallas en el cumplimiento de

gramáticas formales, como en el lenguaje de marcado y hojas de estilo. Algunos portales web utilizan tablas para maquetación.

Análisis WCAG 1.0 en la página Home

Es importante verificar el cumplimiento de las directrices en la página *Home* o de *Inicio*, dado que es la página de ingreso al portal web de la Universidad, En las Tablas 5 y Tabla 6, se muestra el número de errores detectados en esa página, con el análisis del WCAG 1.0, nivel AA, para cada categoría de universidad.

Tabla 5. Análisis WCAG 1.0 en la página HomeUniversidades Categoría A

Universidad	Número de errores	
	Prioridad 1	Prioridad 2
EPN	8	17
ESPE	61	195
ESPOL	0	3
ESPOCH	56	222
UCE	0	88
UC	8	29
Promedio	22,16	92,33

Tabla 6. Análisis WCAG 1.0 en la página HomeUniversidades Categoría B

Universidad	Número de errores	
	Prioridad 1	Prioridad 2
UAE	9	124
UEB	0	10
UNACH	1	21
UNL	9	62
UTN	0	86
UG	3	45
Promedio	3,66	58

Los resultados de la Tabla 5 y Tabla 6 muestran que dos universidades cumplen el nivel de conformidad A, en la página Home, en cada categoría. El promedio del número de errores encontrados es mayor en universidades de categoría A, para los niveles de conformidad A y AA. El número promedio de errores encontrados para puntos de verificación de prioridad 2 es significativo y concierne principalmente a fallas en la maquetación general de la página.

5.3 Verificación de la accesibilidad con WCAG 2.0

Resultados de análisis WCAG 2.0 en todas las páginas

En la Tabla 7 y Tabla 8, se muestra el número de errores detectados en la aplicación de la herramienta TAW, con relación a la WCAG 2.0, nivel AA, en las páginas analizadas para cada universidad en cada una de las categorías, bajo los criterios: Perceptible (P), Operable (O), Comprensible (C), Robusto (R).

Tabla 7. Análisis WCAG 2.0 Universidades Categoría A

Universidad	Páginas analizadas	Número de errores por criterio			
		P	O	C	R
EPN	6	214	65	13	163
ESPE	6	555	36	5	131
ESPOL	6	103	34	7	19
ESPOCH	6	1865	579	12	392
UCE	6	101	11	2	48
UC	6	442	28	1	27
Total	36	3280	753	40	780
Promedio errores/página		91,11	20,92	1,11	21,67

Tabla 8. Análisis WCAG 2.0 Universidades Categoría B

Universidad	Páginas analizadas	Número de errores por criterio			
		P	O	C	R
UAE	6	84	9	8	32
UEB	(*) 5	265	12	10	11
UNACH	(*) 5	72	7	2	3
UNL	6	131	42	3	42
UTN	6	92	7	18	14
UG	6	128	43	8	19
Total	34	772	120	49	121
Promedio errores/página		22,71	3,53	1,44	3,56

Como puede observarse, en la Tabla 7 y Tabla 8, para las dos categorías hay un número alto de errores del principio de Perceptibilidad. Hay dos universidades categoría A, con un alto número de errores en este principio. Los errores más comunes detectados son: no proporcionar texto alternativo para contenido no textual, el contenido multimedia es dependiente del contexto, no se tiene una clara distinción de lo importante en las páginas.

Análisis WCAG 2.0 en la página Home

Es importante verificar el cumplimiento de las directrices en la página *Home* o de *Inicio*, como ingreso al portal web. En la Tablas 9 y Tabla 10, se muestran los errores detectados en esa página, en cada grupo de universidades, con el análisis del WCAG 2.0, nivel AA, bajo los criterios: Perceptible (P), Operable (O), Comprensible (C), Robusto (R).

Tabla 9. Análisis WCAG 2.0 página HomeUniversidades Categoría A

Universidad	Número de errores por criterio			
	P	O	C	R
EPN	17	5	3	9
ESPE	140	8	1	4
ESPOL	1	1	1	0
ESPOCH	90	17	2	33
UCE	6	0	0	4
UC	98	4	0	7
Promedio	58,67	5,83	1,17	9,50

Tabla 10. Análisis WCAG 2.0 página HomeUniversidades Categoría B

Universidad	Número de errores por criterio			
	P	O	C	R
UAE	24	2	1	13
UEB	113	8	2	3
UNACH	18	3	0	1
UNL	45	4	0	25
UTN	14	2	6	5
UG	15	23	1	6
Promedio	38,17	7,00	1,67	8,83

En la Tabla 9 y Tabla 10 se observa que el principio de Perceptibilidad muestra mayor número de errores y que el promedio de errores es mayor en las universidades tipo A es mayor.Los usuarios tienen dificultad en percibir los componentes de la interfaz de ingreso.

5.4 Verificación de sintaxis de lenguaje de marcado

Verificación del lenguaje de marcado en todas las páginas

En la Tabla 11 y Tabla 12, se presentan los errores de sintaxis detectados en el lenguaje de marcado de las páginas de las universidades, a través del servicio en línea, W3C MarkupValidationService.

Tabla 11. W3C MarkupValidationService Universidades Categoría A

Universidad	Páginas analizadas	Número de errores	Páginas que cumplen validación
EPN	6	68	0
ESPE	6	768	0
ESPOL	6	34	2
ESPOCH	6	3107	0
UCE	6	434	0
UC	6	120	0
Total	36	4531	2
Errores/página		125,86	
% páginas validadas			5,56%

Tabla 12. W3C MarkupValidationServiceUniversidades Categoría B

Universidad	Páginas analizadas	Número de errores	Páginas que cumplen validación
UAE	6	179	0
UEB	5	49	0
UNACH	5	84	0
UNL	6	72	0
UTN	6	25	3
UG	6	344	0
Total	34	753	3
Errores/página		22.15	
% páginas validadas			8,82%

Verificación del lenguaje de marcado en la página Home

En la Tabla 13 y Tabla 14, se muestran los errores de sintaxis detectados en el lenguaje de marcado en la página Home o Inicio, de las universidades de cada categoría, a través del servicio en línea, W3C MarkupValidationService.

Tabla 13. W3C MarkupValidationService página Home Universidades Categoría A

Universidad	Número de errores	Cumple validación
EPN	9	No
ESPE	126	No
ESPOL	0	Si
ESPOCH	231	No
UCE	71	No
UC	37	No
Total	474	

Tabla 14. W3C MarkupValidationService, página Home Universidades Categoría B

Universidad	Número de errores	Cumple validación
UAE	30	No
UEB	13	No
UNACH	18	No
UNL	38	No
UTN	0	Si
UG	112	No
Total	211	

En la Tabla 13 y Tabla 14 se observa que, únicamente una universidad en cada grupo, pasa la validación de la página Home. Una universidad del grupo A, tiene un número importante de errores; y las universidades de categoría A tienen el mayor número de errores en la página Home de sus portales web.

5.5 Verificación de sintaxis de CSS

Verificación de hojas de estilo en todas las páginas

En la Tabla 15 y Tabla 16, se presentan los errores de sintaxis detectados en las hojas de estilo, utilizando el servicio en línea,W3C CSS ValidationService.

Tabla 15. W3C CSS ValidationServiceUniversidades Categoría A

Universidad	Páginas analizadas	Número de errores	Páginas que cumplen validación
EPN	6	275	0
ESPE	6	389	0
ESPOL	6	8	5
ESPOCH	6	599	0
UCE	6	582	0
UC	6	138	0
Total	36	1991	5
Errores/página		55,31	
% páginas validadas			13,89%

Tabla 16. W3C CSS ValidationService Universidades Categoría A

Universidad	Páginas analizadas	Número de errores	Páginas que cumplen validación
UAE	6	1	5
UEB	5	110	0
UNACH	5	260	0
UNL	6	84	3
UTN	6	3	3
UG	6	244	1
Total	34	702	12
Errores/página		20,65	
% páginas validadas			35,29

En la Tabla 15 y Tabla 16, el escenario es igual a la validación de lenguaje de marcado. El número de errores en universidades A, es significativamente superior al de universidades B, por tanto lo es el promedio de errores / página. Se observa que dos universidades A, reportan un número considerable de errores. El número de páginas en universidades del tipo B, que cumple la validación, es más que el doble, que para las universidades tipo A. El porcentaje de páginas que cumple la validación es significativamente superior para universidades categoría B, comparadas con universidades categoría A.

Verificación de hojas de estilo en la página Home

En la Tabla 17 y Tabla 18, se muestran los errores de sintaxis detectados en las hojas de estilo, en la página Home o Inicio, de las universidades de cada categoría, utilizando el servicio en línea,W3C CSS ValidationService. Este servicio está disponible ingresando el URL a analizar.

En la Tabla 17 y Tabla 18 se observa que ninguna universidad de categoría A pasa la validación de hoja de estilo.

Una universidad de categoría B, pasa la validación. El número de errores en la página Home de universidades categoría A, es superior al de universidades categoría B.

Tabla 17. W3C CSS ValidationService página Home Universidades Categoría A

Universidad	Número de errores	Cumple validación
EPN	10	No
ESPE	8	No
ESPOL	8	No
ESPOCH	99	No
UCE	86	No
UC	11	No
Total	222	

Tabla 18. AnálisisW3C CSS ValidationService, página Home Universidades Categoría B

Universidad	Número de errores	Cumple validación
UAE	0	Si
UEB	22	No
UNACH	9	No
UNL	53	No
UTN	1	No
UG	69	No
Total	154	

6. CONCLUSIONES

La evaluación automática permite realizar una inspección rápida, ayuda a tener una primera impresión de la accesibilidad de una página web, pero no proporciona un análisis definitivo y fiable, ya que puede no detectar errores importantes o puede señalar errores que realmente no existen (falsos positivos). Únicamente un análisis manual por parte de un experto puede ofrecer un análisis completo y fiable de la accesibilidad web de una página. Por tanto, los resultados que se muestran en este artículo deben ser considerados una referencia preliminar del estado de la accesibilidad web en los portales web analizados.

En términos generales, los portales web de las universidades analizadas no cumplen las directrices de la WAI, definidas en la WCAG 1.0 y WCAG 2.0.

Comparando los grupos de universidades, en conjunto hay mejores resultados en las universidades de categoría B. Esto puede deberse a que, los diseños de los portales en ciertas universidades categoría A, incluyen muchos componentes multimedia, que requieren un tratamiento especial para que sean accesibles.

Se requiere corregir errores en todos los portales web de las Universidades analizadas. Por el nivel y el número de errores reportados, algunas universidades deben realizar rediseños importantes en sus portales web, a fin de cumplir con las guías de accesibilidad web y buenas prácticas en el uso de la sintaxis de los lenguajes de marcado y de hojas de estilo. Es posible que algunas universidades puedan mejorar el cumplimiento de las pautas de accesibilidad web, con un esfuerzo razonable de revisión y corrección del código de sus portales web.

Si bien en Ecuador el marco legal existente hace referencia a la accesibilidad web en los portales institucionales de las universidades, este marco es genérico y no define requisitos técnicos, como por ejemplo, el nivel mínimo de cumplimiento de la accesibilidad. Para facilitar la implantación de la accesibilidad web en el Ecuador, el marco legal debería definir los requisitos técnicos y lo recomendable es que se basen en los estándares internacionales, a fin de utilizar herramientas, bibliografía, experiencias, desarrolladas en torno a esos estándares.

Finalmente, sería recomendable que las universidades,lideraran este tema, dada su importancia como referentes de la sociedad. Esto les permitirá alinearse con la

estrategia estatal expresada en el Plan Nacional para el Buen Vivir 2013-2017 y en la Ley Orgánica de Discapacidades.

REFERENCIAS

[1] Asamblea Nacional, "Ley Orgánica de Discapacidades," Registro Oficial, Quito, Ley Orgánica Año IV No.796, Aug. 2012. [Online]. http://www.conadis.gob.ec/index.php?option=com_content

[2] Benavidez, C. (2013, July) Libro Blanco eXaminator. [Online]. http://examinator.ws/info/libro_blanco_examinator.pdf

[3] Consejo de Educación Superior, CES. (2013, July) IES, Universidades y Escuelas politécnicas. [Online]. http://www.ces.gob.ec/ies/universidades-y-escuelas-politecnicas/por-categoria/categoria-a

[4] Grooves, K., "Choosing an Automated Accessibility Testing Tool: 13 Questions you should ask," in *International Technology and Persons with Disabilities Conference*, San Diego, CA, 2013.

[5] ISO. (2013, July) ISO/IEC 40500:2012, Information Technology – W3C Content Accesibility Guidelines (WCAG) 2.0. [Online]. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=58625

[6] Luján- Mora, S. (2013, Aug.) Universidad de Alicante - Accesibilidad Web - Evaluación de la accesibilidad web. [Online]. http://accesibilidadweb.dlsi.ua.es/?menu=evaluacion

[7] Sergio Luján-Mora. Accesibilidad Web: Legislación. [Online]. http://accesibilidadweb.dlsi.ua.es/?menu=legislacion

[8] Luján-Mora, S. (2013, July) Universidad de Alicante - Accesibilidad Web - Tabla de puntos de verificación de WCAG 1.0. [Online]. http://accesibilidadweb.dlsi.ua.es/?menu=puntos-1.0

[9] Secretaría Nacional de Planificación y Desarrollo. (2013, Aug.) Plan Nacional para el Buen Vivir 2013-2017. [Online]. http://www.buenvivir.gob.ec

[10] TAW. (2013, Aug.) CTIC Centro Tecnológico. [Online]. http://www.tawdis.net/index.html?lang=es

[11] W3C. (2013, Aug.) Accessibility. [Online]. http://www.w3.org/standards/webdesign/accessibility

[12] W3C. (2013, July) Checklist of Checkpoints for Web Content Accessibility Guidelines 1.0. [Online]. http://www.w3.org/TR/1999/WAI-WEBCONTENT-19990505/full-checklist

[13] W3C. (2013, July) Complete List of Web Accesibility Evaluation Tools. [Online]. http://www.w3.org/WAI/ER/tools/complete

[14] W3C. (2013, Aug.) CSS Validation Service. [Online]. http://jigsaw.w3.org/css-validator/

[15] W3C. (2013, Aug.) Markup Validation Service. [Online]. http://validator.w3.org

[16] W3C. (2013, Aug.) Understanding WCAG 2.0. [Online]. http://www.w3.org/TR/UNDERSTANDING-WCAG20/

[17] W3C. (2013, Aug.) Web Accesibility Initiative. [Online]. http://www.w3.org/WAI/

[18] W3C. (2013, Aug.) Web Accessibility Initiative, Comparison of WCAG 1.0 Checkpoints to WCAG 2.0, in Numerical Order. [Online]. http://www.w3.org/WAI/WCAG20/from10/comparison/

[19] W3C. (2013, July) Web Content Accessibility Guidelines 1.0 , Recommendation 5-May-1999. [Online]. http://www.w3.org/TR/WCAG10/

Análisis de Sentimientos Aplicado a Referencias Bibliográficas

Hernández M.*; Gómez J.**

***Escuela Politécnica Nacional, Facultad de Ingeniería de Sistemas*
e-mail: myriam.hernandez@epn.edu.ec
*** Universidad de Alicante, Departamento de Lenguajes y Sistemas Informáticos*
e-mail: jmgomez@ua.es

Resumen: *Actualmente índices como el hindex calculan el impacto de un artículo científico considerando solamente el número y la frecuencia con que una cita bibliográfica es mencionada, sin tomar en cuenta el propósito ni la polaridad con que fue referenciada; éste método de cómputo puede llevar a conclusiones equivocadas a la hora de valorar la contribución científica de algunos investigadores o el impacto de las revistas en las que publican. En este trabajo se propone complementar las medidas cuantitativas de impacto con criterios subjetivos usando Análisis de Sentimientos para detectarlos en el contexto de citas bibliográficas dentro del texto de los artículos. Se presenta una visión general de lo que involucra el Análisis de Sentimientos basado en los enfoques de reconocimiento de patrones y aprendizaje automático; se plantea una propuesta de diseño que combine estas técnicas, las aplique en referencias bibliográficas y obtenga nuevos elementos para una visión más completa de la influencia de un artículo en la literatura científica.*

Palabras clave: *Análisis de sentimientos, bibliometría, polaridad, reglas, aprendizaje automático.*

Abstract: *Currently indexes like hindex, calculate the impact of a scientific article considering only the number and the frequency of citation mentions, regardless of the purpose or polarity that authors intended; this method of calculation may be misleading when evaluating the scientific contribution of some researchers or the impact of the journals in which they publish. This paper aims to complement quantitative measures of impact analysis using subjective criteria detected with Sentiment Analysis in the context of citations in the text of the articles. We present an overview of what is involved in Sentiment Analysis using different approaches: pattern recognition and machine learning based technologies; we present a design proposal that combines these techniques applied in the context of bibliographic references to obtain a more complete view of the influence of an article in the scientific literature.*

Keywords: *Sentiment analysis, bibliometrics, polarity, rules, machine learning.*

1. INTRODUCCION

Una cita bibliográfica es un texto que aparece en un artículo científico y referencia un trabajo previo con distintos propósitos como, por ejemplo, para comparar resultados o reconocer las fuentes usadas. Las citas en un texto científico normalmente son numerosas. Conectan piezas de trabajos de investigación a través de distintos periodos de tiempo y diferentes comunidades en la literatura científica. La clasificación subjetiva de estas referencias proporciona información valiosa que facilita la investigación. Este tipo de citas categorizadas se aplican para generar resúmenes y revisiones automáticas.

Los índices de impacto como el hindex [7] se relacionan con el número de veces que un artículo ha sido citado. En el cálculo de este tipo de índices, una cita se interpreta como que un autor A ha sido influenciado por el trabajo de un autor B, sin especificar el tipo de influencia [11]. Tratar a todas las referencias de igual manera puede llevar a conclusiones equivocadas en cuanto al verdadero impacto de una cita bibliográfica [14]. Por ello, se considera necesario analizar otros indicadores que muestren la intención del autor al realizar la referencia. El análisis automático de criterios subjetivos presentes en un texto se conoce como Análisis de Sentimientos (AS) y podría servir para complementar la estimación del impacto de una publicación.

El AS, es un tema actual de investigación en el área de procesamiento del lenguaje natural en el campo de extracción de la información; su campo de aplicación es muy amplio en el monitoreo de emociones en disciplinas tan diversas como el marketing, ciencias políticas y economía.

El AS se puede aplicar en el estudio de citas bibliográficas, usando el contexto de las citas para detectar la intención y sentimientos de un autor al hacer una referencia o cita bibliográfica dentro de un artículo científico. Con esta información se puede mejorar el análisis bibliométrico de las referencias [1]. El análisis sintáctico y semántico del contenido de las oraciones alrededor de las citas se obtiene de las teorías existentes sobre ellas como se presenta en la Tabla 1 que ha sido adaptada a partir de [11].

Para realizar avances en el análisis de las citas es necesario que se combinen las dos orientaciones principales que se

manejan en estos momentos: descripción cuantitativa de las referencias bibliográficas y la descripción cualitativa de las mismas [11].

Tabla 1: Bases teóricas para entender las principales características de una cita bibliográfica

Características y Orientación	Orígen conceptual	Lo que asume	Métodos analíticos
Numérica - Explícita	La ciencia necesita medir los distintos fenómenos	El número de citas puede ser considerado un indicador de impacto	Conteo de citas, frecuencia de citas
Literal - Explícita / Implícita	La naturaleza simbólica de las palabras	Las palabras y elecciones lingüísticas en el texto en que se realizan las citas pueden indicar la función y sentimiento de la cita	Procesa-miento de lenguaje natural, aprendizaje automático
Socio cultural - Implícita	La naturaleza individual y social de las selecciones	El contenido semántico del contexto de las citas sugiere las motivaciones de las citas	Análisis de contenido

2. ANÁLISIS DE SENTIMIENTOS

2.1 Visión general

El AS se refiere en general a la extracción de entidades que indican emociones en un texto para clasificarlas como positivas, negativas o neutrales. Hay distintos enfoques para el análisis de sentimientos que a la final termina siendo una categorización de palabras que se puede realizar con técnicas basadas en reconocimientos de patrones, o usando aprendizaje automático ya sea supervisado, no supervisado o híbrido [13].

La extracción de unidades que expresan sentimientos en un texto se puede hacer usando distintos métodos como el planteado por [6], donde una unidad de sentimiento está conformada por un sentimiento, un predicado y su argumento. Por ejemplo, las siguientes oraciones en un comentario de un cliente de una cámara digital, contiene tres unidades de sentimientos (1a), (1b) y (1c). Aparentemente estas unidades indican que la cámara tiene buenas características en sus lentes y cargador y una mala característica en su precio.

La cámara tiene excelentes lentes,
pero el precio es muy alto.
Yo no creo que la calidad del cargador
presente problemas. (1)

[favorable] excelentes (lentes) (1a)
[no favorable] alto (precio) (1b)
[favorable] problemático $^{+neg}$ (cargador) (1c)

La extracción de estas unidades de sentimientos se llama traducción de texto a unidades de sentimiento. Esta no es una tarea trivial porque se requieren muchas operaciones de tipo sintáctico y semántico, que incluyen extracción de los argumentos de la oración, manejo de información expresada como negación, desambiguación de sentimientos para detectar que "alto" en este caso es negativo al aplicarse a precio, a pesar de que podría ser un adjetivo positivo si se aplicara a otro parámetro, como por ejemplo, calidad. En esta tarea se pueden reusar herramientas de traducción automática [8].

El AS tiene como objetivo 3 tareas esenciales, de acuerdo a [4]: la extracción de la polaridad objetivo/subjetiva del texto (polaridad SN) [18], la determinación de la polaridad positivo / negativa (polaridad PN) del texto y la intensidad de la polaridad PN. Recursos léxicos de acceso libre como SentiWordNet [4] aplican este análisis y tienen unidades asociadas a tres valores numéricos, cada uno de los cuales describen cuán objetivos, positivos y negativos, respectivamente, son los términos contenidos en ellas.

El AS enfrenta problemas debido a la naturaleza misma del lenguaje humano, cuya estructura presenta dificultades para la detección automática. Para lograr una mayor precisión en la determinación de la polaridad, se sugiere aplicar una combinación de técnicas que permitan mejorar los resultados. En este contexto, una técnica adicional que se puede aplicar es el análisis de discurso.

En [21] se muestra que métodos híbridos de clasificación que combinan distintos algoritmos y enfoques, pueden mejorar la eficiencia de este análisis en términos de micro y macro F1 promedio. Siendo F1 una medida que toma la precisión y la recuperación como medidas de la eficiencia del clasificador. Otra forma de mejorar la eficiencia es el uso consecutivo de clasificadores de distinto tipo para lograr una sinergia en su funcionamiento y una mejor eficiencia.

La Tabla 2 (adaptada de [21]), muestra la forma como se obtendrían estas métricas para determinar la eficiencia de los algoritmos en AS, la tabla se refiere a la comparación entre los sentimientos evaluados por un humano versus los clasificados por una máquina.

Hay que tomar en cuenta la dificultad que presenta comparar objetivamente la eficiencia de distintos enfoques para AS debido a la aplicación de diversos métodos de evaluación y de datos de prueba.

Tabla 2: Aciertos

	La máquina dice SI	**La máquina dice NO**
El humano dice SI	True positive = TP	Fail negative= FN
El humano dice NO	Fail positive = FP	True negative = TN

Con estos valores se pueden calcular las variables precisión, recuperación y F1 para medir la eficiencia de los algoritmos para AS.

$$\text{Precisión (P)} = TP / (TP + FP)$$
$$\text{Recuperación (R)} = TP / (TP + FN) \quad (2)$$
$$F1 = 2.P.R / (P + R)$$

2.2 Técnicas

Se puede entender al AS como una categorización de palabras de acuerdo a características predefinidas que son palabras o conjuntos de palabras que denotan sentimientos positivos, negativos o neutrales.

En trabajos existentes, para automatizar el análisis de sentimientos se han aplicado diferentes enfoques, algunos relacionados con procesamiento de lenguaje natural (PLN) y basados en reconocimiento de patrones [8], otros usan algoritmos de aprendizaje automático y también aprendizaje no supervisado.

Los enfoques basados en PLN y reconocimiento de patrones se enfocan en el uso de herramientas de procesamiento como etiquetadores de las distintas partes de las oraciones y sentencias (etiquetadores de Part-of-Speech POS) y analizadores sintácticos. Las etiquetas PoS permiten clasificar el componente como adverbio, verbo, adjetivo, sustantivo, etc. como parte del análisis sintáctico; lo que permite pasar a la siguiente etapa o análisis semántico y posteriormente a la extracción de unidades de sentimientos. Los resultados generados por las herramientas se procesan para generar una colección de patrones. Cada patrón tiene asignado un sentimiento positivo o negativo y por lo tanto tienen definida su polaridad. Este tipo de tareas las realiza, por ejemplo, el analizador sintáctico Montylingua [10].

Las características pueden ser ngramas, donde n corresponde al número de palabras que componen la frase que va a ser analizada. Por ejemplo "excelente" es un unigrama, "terriblemente complicado" es un bigrama, "un gran trabajo" es un trigrama. Los unigramas normalmente no son suficientemente específicos y no se usan para extracción de unidades de sentimiento, los bigramas son los más utilizados y los trigramas son muy específicos lo que les hace perder frases que pueden ser importantes. Normalmente no se usan combinaciones de más de tres palabras, pero podrían ser de utilidad si es que estamos en una búsqueda definida, por ejemplo de un título de un artículo que se está analizando. Este tipo de búsqueda es computacionalmente simple y fácil de entender pero puede requerir de una larga lista de palabras. Las palabras de los ngramas se ubican en una colección conocida como "Bag-of-Words".

El aprendizaje no supervisado se enfoca en explotar corpus de prueba para determinar el sentimiento de una expresión, se puede inferir la polaridad de las palabras con su asociación estadística usando un conjunto de palabras paradigma que sirva como ejemplo para mostrar sus distintas posibles acepciones tanto positivas como negativas como se presenta en [20].

En el aprendizaje automático algoritmos muy utilizados son Näive Bayes y Support Vector Machine, entre otros [15].

Los métodos híbridos combinan distintos tipos de clasificadores en distinto orden y secuencia hasta lograr una mejora de la eficiencia. Este resultado puede variar de acuerdo a los campos de aplicación y se requeriría experimentación para determinar la mejor combinación y secuencia [21]. Por último, existen sistemas que utilizan recursos externos como léxicos especializados en AS, por ejemplo, SentiWordNet, para introducir esta información como características a los modelos de entrenamiento [5]. De esta forma no sólo se enriquece el modelo con información léxica y sintáctica, sino que se incluye también información semántica.

3. CLASIFICADORES BASADOS EN RECONOCIMIENTO DE PATRONES VS CLASIFICADORES CON APRENDIZAJE AUTOMÁTICO

Los clasificadores basados en reconocimiento de patrones son diferentes a los que usan aprendizaje automático porque parten de un mecanismo subyacente distinto. Se definen un grupo de reglas que se extraen de corpus de prueba como un grupo de patrones, para ello se usan plantillas o métodos heurísticos. De acuerdo a [21], el mapeo entre cada patrón y una categoría lleva a la construcción y la optimización de un conjunto de reglas.

Una regla consiste de un antecedente y un consecuente. Un antecedente se define como una condición que consiste de un token o secuencia de tokens concatenados por el operador ^. Un token puede ser una palabra "?" o un término objetivo "#" que varía de acuerdo al contexto del análisis, puede ser por ejemplo el nombre de un autor de un artículo. El consecuente representa un sentimiento positivo o negativo.

$$\{ [\text{token 1}] \text{ ^ } [\text{token 2}] \text{ ^ } \ldots \text{ ^ } [\text{token n}]\} \rightarrow \{+/-\}$$
$$\{\text{excelente}\} \rightarrow \{+\} \quad (3)$$
$$\{\text{absurdo}\} \rightarrow \{-\}$$

Se asumen dos sentencias:

1. Laptop-A es más costoso que Laptop-B
2. Laptop-A es más costoso que Laptop-C (4)

La palabra objetivo es Laptop-A. La regla derivada de estas sentencias es:

{ # ^ más ^ costoso ^que ^ ? } → {-} (5)

La interpretación de esta regla es que Laptop-A tiene un sentimiento más negativo que las otras dos debido a su precio.
Para la extracción de reglas se toma en cuenta también la presencia de negación o la proximidad de otras oraciones. Dado un conjunto de entrenamiento el patrón asociado se enfoca en extender la regla existente o manejar un nuevo conjunto de ejemplo u optimizar las reglas existentes para habilitar una concordancia eficiente (inducción de reglas).
En contraste un clasificador con aprendizaje automático considera un conjunto de características significativas, cada una de las cuales tiene asignado un peso. Por lo tanto, dado un conjunto de entrenamiento, el modelo asociado se enfoca en optimizar los pesos y otros parámetros de tal manera que el modelo puede lograr un alto nivel de eficiencia en un conjunto de ejemplo no visto anteriormente. Se asume la independencia de características lo que no es siempre cierto. Adicionalmente, se tiene que hacer a veces un compromiso entre la eficiencia del modelo y el tiempo que se necesita para entrenarlo. Un algoritmo de aprendizaje automático eficiente se requiere para entrenar el modelo. Posteriormente, un clasificador hará uso del modelo que contiene los pesos aprendidos y los parámetros para clasificar un ejemplo no visto.

3.1 Técnicas de clasificación usando aprendizaje automático

Debido a que son ampliamente usados en el aprendizaje automático para determinar la categoría de los sentimientos en un texto, en el presente trabajo se van a mencionar dos algoritmos: Naïve Bayes [15], y Support Vector Machines (SVM) [12].

3.2 Naïve Bayes

Aplica un marco probabilístico y usa el teorema de Bayes para calcular la probabilidad de una clase (positiva, negativa). Por ejemplo, dado un texto X cuya clase no se conoce, y una clase particular C, la probabilidad de una clase C de un texto se calcula con el Teorema de Bayes:

$$P(C|X) = P(X|C)P(C)/P(X) \quad (6)$$

Donde P(C|X) es la probabilidad a posteriori una clase C basada en el texto X. P(X|C) es la probabilidad posterior de un texto dada una clase. P(C) es la probabilidad de una clase. P(X) es la probabilidad de un texto. Se asume que las probabilidades de las características son independientes de la clase.

3.3 Support Vector Machine

El algoritmo SVM explota el uso de vectores que modelan la distribución de características en un texto.

Como se presenta en la Fig. 1, cada vector es un punto de un espacio n-dimensional, donde n es el número de características que contienen un valor booleano que indica si la característica existe en el texto o un número que representa la frecuencia de ocurrencia de la característica. El objetivo de SVM es modelar un hiperplano que divide grupos de vectores en categorías separadas maximizando la separación entre los datos de entrenamiento. En su forma más simple SVM diferencia entre dos categorías. Los vectores soporte son los más cercanos a la línea. La tarea es determinar los vectores que mejor describen la división entre categorías que tengan la mayor distancia entre clases (margen).
En el caso de datos que no son linealmente separables Fig. 2 se utilizan Kernels que mapean un espacio a otro con más dimensiones en el cual el espacio de características es separable para el set de entrenamiento Fig. 3.

Por ejemplo:

Un texto (cadena) se representa como el conteo de todas las subcadenas de k caracteres.

$$\Theta = (T) = (\#\text{subcadenastipo}_1, \ldots, \#\text{subcadenastipo}_n) \quad (7)$$

La función tiene n dimensiones.
Función Kernel: Conteo de todas las subcadenas que coinciden.

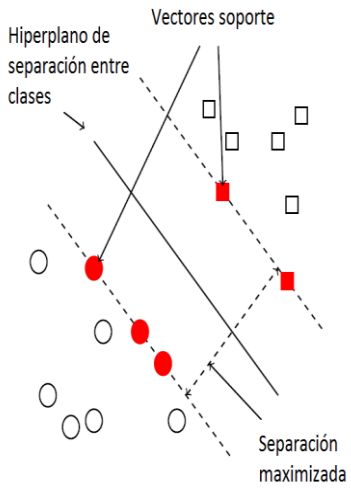


Figura 1: Algoritmo SVM

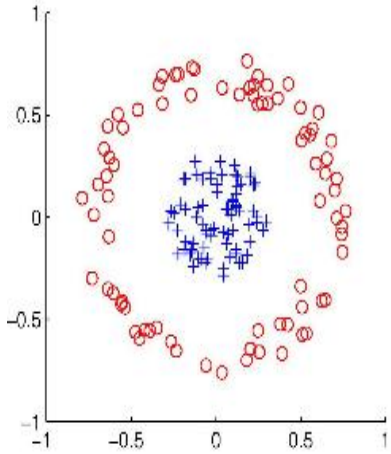


Figura 2: clases no linealmente separables en un espacio de dos dimensiones

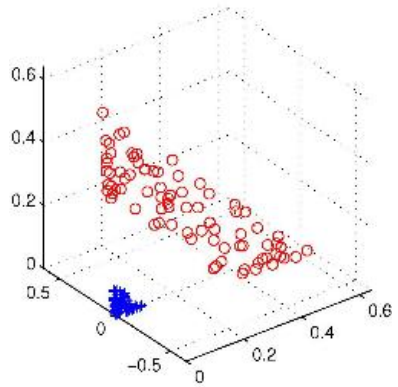


Figura 3: Clases linealmente separables en un espacio de tres dimensiones

En un enfoque de Bag-of-Words en un Kernel de cadena se tienen subcadenas separadas con espacios en blanco:

$\vec{b}$ = (#Yo, #estoy, #realmente, #muy, #feliz, #infeliz, #con, #esta, #cámara)
x = Yo estoy realmente feliz con esta cámara.
$\vec{x}$ = (1,1,1,0,1,0,1,1,1) (8)
z = Yo estoy muy infeliz con esta cámara.
$\vec{z}$ = (1,1,0,1,0,1,1,1,1)

La comparación se hace fácilmente en el Kernel y no con el espacio inicial. Este método no requiere que conozcamos el separador de las clases en el espacio inicial.

$$K(\vec{x}, \vec{z}) = \sum (x_i \cdot z_i) = 4 \quad (9)$$

K nos da el valor del margen de separación entre las clases.
Muchas aplicaciones usan árboles para análisis sintáctico. En este caso también se puede usar SVM cambiando el árbol a un vector de características enumerando todas las ramas de dos nodos y siguiendo un procedimiento similar al descrito, contando el número de ramas comunes.

Se pueden poblar los vectores de características para el SVM usando recursos léxicos y semánticos como WordNet [20], SentiWordNet [4] o WordNetAffect [17]. También se pueden utilizar herramientas de análisis de textos para obtener información léxica o sintáctica como el Freeling [3], Stanford NLP [16] o MiniPAR [9]. El WordNet es una base de datos léxica en inglés aunque tiene sus correspondientes versiones en otros idiomas como el EuroWordNet [19]. Los sustantivos, verbos, adjetivos y adverbios se agrupan en colecciones de sinónimos cognitivos (synsets), cada uno de los cuales expresa un concepto distinto. Los synsets se interrelacionan lexical y semánticamente. Los recursos como el SentiWordNet y el WordNetAffect relacionan algunos de los synset de WordNet que normalmente expresan sentimientos con su polaridad e intensidad. Diversas combinaciones de todos estos recursos han sido utilizadas en sistemas de aprendizaje automático con SVM [5], [2].

3.4 Análisis de Contexto

Según [11] y [12] se asume que las palabras alrededor de la cita objetivo proporcionan evidencia semántica para inferir la motivación del autor que la referencia. Si la cita contribuye al tema central, en el análisis bibliométrico se debería acreditar un mayor valor a su impacto. La ubicación de la entrada bibliográfica dentro del artículo científico proporciona información sobre la función de la cita. En [6] se establece que citas múltiples a una referencia indica una relación cercana y útil de esa referencia al artículo que la menciona. Una publicación citada en la introducción o en la sección de revisión de la literatura y mencionada nuevamente en la metodología o en la sección de discusión, ofrece una mayor contribución a la publicación que otras que han sido mencionadas una sola vez.
Por otro lado, la ventana que rodea a la cita objetivo {-n palabras, +n palabras} se usa para inferir el sentimiento y propósito del autor al hacer la referencia. [11] sugiere un n de 150, pero aclara que deben realizarse pruebas para hallar un valor óptimo para el tamaño de la ventana.

4. PLANTEAMIENTO DE UN SISTEMA PARA ANÁLISIS DE SENTIMIENTOS APLICADO A REFERENCIAS BIBLIOGRÁFICAS

Se plantea desarrollar un sistema analizador de polaridad de sentimientos de referencias bibliográficas en artículos cientificos, soportado por métodos híbridos que combinen varios enfoques y tecnologías con el fin de producir resultados confiables por su mayor eficiencia en cuanto a recuperación y precisión; los resultados se compararán con una línea base del estado del arte actual. Se tiene la visión de realizar un análisis fundamentado en el contexto próximo de las citas. El sistema tendrá las características generales que se describen en el diagrama lógico y los módulos propuestos, que se refieren a continuación.

4.1 Diagrama Lógico

El diagrama lógico se presenta en la Fig. 4. La arquitectura se define por capas para permitir reutilización, estandarización, flexibilidad y facilidad en el manejo de errores.

La capa de interfaz permite la comunicación con sistemas externos y cómo se presenta los resultados al usuario. La capa de aplicación maneja la lógica de los algoritmos y técnologías para el rastreador, extractor y analizador de sentimientos. La capa de negocio gestiona la comunicación y el flujo de trabajo entre la interfaz web y las aplicaciones. La capa de persistencia maneja el almacenamiento de las entradas y salidas del sistema.

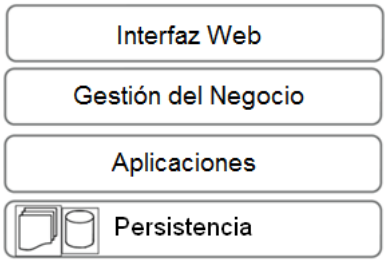


Figura 4: Arquitectura

4.2 Módulos principales

El sistema que se plantea desarrollar, se muestra en la Fig. 5 con los siguientes módulos principales:

- **Interfaz web:** Gestiona las entradas y salidas del sistema y proporciona un método de comunicación entre el usuario final y el sistema.
- **Rastreador de referencias bibliográficas**: Navega artículos científicos y obtiene las referencias bibliográficas.
- **Extracción de información**: Sistema que, a partir de una referencia bibliográfica obtenida por el rastreador, es capaz de extraer cada uno de sus campos (título, autores, revista, año publicación, editor, etc.) junto con información de contexto en la referencia. Obtienen la información objetivo dentro de una ventana con un tamaño escogido heurísticamente con el fin de agregar contexto suficiente y se calculan indicadores de ubicación y frecuencia de las entradas objetivo.
- **Analizador de textos**: Dividen las oraciones contenidas en las ventanas alrededor de las citas en sus partes sintácticas: verbos, sustantivos, adjetivos, adverbios, etc. (POS Tagging).
- **Analizador de sentimientos**: Definen unidades de sentimientos. Aplica en los datos obtenidos combinaciones de las tecnologías mencionadas en este artículo y otras que sean adecuadas hasta lograr resultados aceptables en la determinación del sentimiento de los ngramas para los autores definidos dentro de las palabras objetivos. Además trabajan en el establecimiento del propósito del autor al realizar la referencia.
- **Repositorio**: Almacena información de entrada y salida del sistema.

5. CONCLUSIONES

En este trabajo se ha planteado el desarrollo de un sistema de análisis de polaridad de sentimientos aplicado a referencias bibliográficas. Se propone usar una combinación de enfoques y tecnologías que permita una sinergia entre ellos hacia el logro de resultados confiables.

La tarea básica del AS es la definición de polaridad de sentimientos, lo que en el análisis de citas bibliográficas dentro de un artículo técnico está relacionado con la determinación de si la entrada fue referenciada en forma positiva o negativa. De la polaridad puede inferirse el tipo de impacto, función y contribución de la cita en ese artículo.

Debido a la naturaleza del lenguaje natural, existen retos que aún no han sido resueltos en la detección automática de la polaridad en un texto. Aún se requiere un mejor entendimiento del puente que existe entre las estructuras sintácticas y semánticas del lenguaje natural y la información estructurada, procesable por una computadora, que pueda analizar las oraciones cercanas a la entrada bibliográfica y obtener elementos subjetivos que aporten a la determinación de su verdadera influencia.

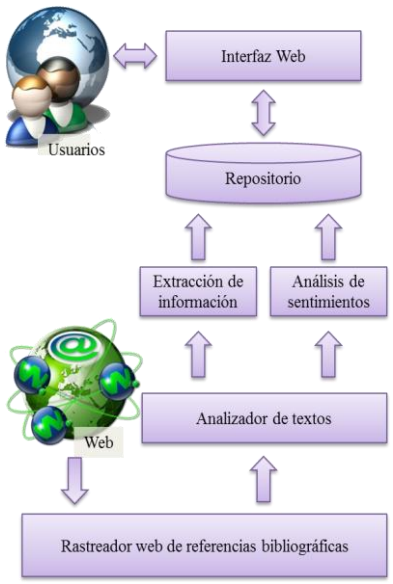


Figura 5: Módulos del sistema

REFERENCIAS

[1] A. Abu-Jbara and D. R. Radev, (2012, June). Reference scope identification in citing sentences. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 80-90). Association for Computational Linguistics.

[2] E. Boldrini, J. Fernández, J.M. Gómez and P. Martínez Barco. (2011). EmotiBlog: towards a finer-grained sentiment analysis and its application to opinion mining. Actas de las IV Jornadas TIMM.

[3] X. Carreras, I. Chao, L. Padró and M. Padró. FreeLing: An Open-Source Suite of Language Analyzers Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC'04), 2004.

[4] A. Esuli and F. Sebastiani (2006). SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. Proceedings of LREC 2006 – 5th Conference on Language Resources and Evaluation, 22-28/5/2006, Genova (IT).

[5] J. Fernández, E. Boldrini, J. M. Gómez, and P. Martínez-Barco, (2011). Evaluating EmotiBlog robustness for sentiment analysis tasks. In *Natural Language Processing and Information Systems* (pp. 290-294). Springer Berlin Heidelberg.

[6] G. Herlach, (1978). Can retrieval of information from citation indexes be simplified? Multiple mention of a reference as a characteristic of the link between cited and citing article. *Journal of the American Society for Information Science*, *29*(6), 308-310.

[7] J. E. Hirsch, (15 November 2005). An index to quantify an individual's scientific research output. *PNAS 102* (46): 16569–16572.

[8] K. Hiroshi, N. Tetsuya and W. Hideo, (2004, August). Deeper sentiment analysis using machine translation technology. In *Proceedings of the 20th international conference on Computational Linguistics* (p. 494). Association for Computational Linguistics.

[9] D. Lin (2003). Dependency-based evaluation of MINIPAR. In Treebanks (pp. 317-329). Springer Netherlands.

[10] H. Liu, (2004). MontyLingua: An end-to-end natural language processor with common sense. Available at http://web.media.mit.edu/hugo/montylingua

[11] X. Liu, J. Zhang and C. Guo, (2012, October). Full-text citation analysis: enhancing bibliometric and scientific publication ranking. In *Proceedings of the 21st ACM international conference on Information and knowledge management* (pp. 1975-1979). ACM.

[12] M. D. Lytras, and P. Ordonez de Pablos, (2009). *Social Web Evolution: Integrating Semantic Applications and Web 2.0 Technologies*. Information Science Reference-Imprint of: IGI Publishing.

[13] R. Prabowo, and M. Thelwall, (2009). Sentiment analysis: A combined approach. *Journal of Informetrics*, *3*(2), 143-157.

[14] F. Radicchi, (2012). In science "there is no bad publicity": Papers criticized in comments have high scientific impact. Scientific reports, 2.

[15] F. Sebastiani (2002). Machine Learning in Automated Text Categorization. ACM Computing Surveys, 34(1), 1-47.

[16] Stanford NLP. http://nlp.stanford.edu/software/index. Shtml. Verificado por última vez: 04 October, 2013.

[17] C. Strapparava and A. Valitutti, (2004). WordNet Affect: an Affective Extension of WordNet. In LREC (Vol. 4, pp. 1083-1086).

[18] P. D. Turney, and M. L. Littman, (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS)*, *21*(4), 315-346.

[19] P. Vossen, (1998). EuroWordNet: a multilingual database with lexical semantic networks. Boston: Kluwer Academic.

[20] WordNet. http://wordnet.princeton.edu/. Verificado por última vez: 02 October, 2013.

[21] G. Zhang, Y. Ding and S. Milojević, (2013). Citation content analysis (cca): A framework for syntactic and semantic analysis of citation content. *Journal of the American Society for Information Science and Technology*.