

Reconocimiento de Patrones en Imágenes Digitales a Color usando el Descriptor RFM

Verdugo-Olachea J.,^{1*} Solorza-Calderón S.¹

¹Facultad de Ciencias, Matemáticas Aplicadas, UABC.

* Correo electrónico: jonathan.verdugo@uabc.edu.mx

Resumen

En este trabajo se propone un sistema digital de reconocimiento de patrones para imágenes a color basado en firmas 1D invariantes a traslación, escala y rotación (RST). La invariancia a rotación se obtiene por medio de la transformada de Radon. Para la invariancia a escala se utiliza la transformada de Fourier-Mellin normalizada. La invariancia a traslación se consigue a través del espectro de amplitud de la transformada de Fourier de la imagen. Al trabajar en el espacio de color RGB la imagen se separa en tres imágenes monocromáticas, las cuales corresponden al canal rojo (R), verde (G) y azul (B). Al aplicar las transformadas integrales a cada una de las imágenes monocromáticas se generan tres imágenes, denominadas Radon-Fourier-Mellin (RFM) que son invariantes a traslación, escala y rotación, por lo que para una imagen a color se generan tres imágenes Radon-Fourier-Mellin. Para cada una de las imágenes Radon-Fourier-Mellin (señal 2D) se construye una firma 1D invariante a traslación, escala y rotación, de la cual se calcula su potencia. Como para la imagen se tienen tres firmas 1D (una para cada canal), entonces se tienen tres potencias. Las tres potencias son los atributos que se le asignan a la imagen a color para su clasificación y se utilizan para generar un espacio de clasificación 3D que tiene un nivel de confianza de al menos el 95.4%. Para mostrar la eficiencia del sistema se emplea una base de datos de 18 imágenes de referencia a color que contienen mariposas, dicho conjunto fue seleccionado por la similitud que presentan en su morfología y gama de colores.

Palabras clave: Reconocimiento de Patrones; Transformadas Integrales; Sistema RFM.

1 Introducción

Por medio de la ciencia y la tecnología, se ha tratado de imitar el proceso cognitivo que realiza el cerebro en la toma de decisiones. Un ejemplo de ello, es el campo denominado reconocimiento de patrones, que es un área del procesamiento de imágenes digitales en donde se desarrollan técnicas para: detección, segmentación, localización y reconocimiento de objetos en imágenes digitales; identificación de un objeto en diversas escenas; seguimiento de un objeto en un vídeo (secuencia de imágenes) y clasificación de imágenes digitales por su contenido [10]. Una de las metas en este campo es reconocer objetos de forma automática con un alto nivel de confianza y bajo

costo computacional; no importa si se hace rotar, escalar o desplazar, con diferentes tipos de ruido, iluminación o incluso si sólo se cuenta con un fragmento del objeto [25].

Los sistemas digitales de reconocimiento de patrones basados en funciones de correlación han sido estudio de interés por muchos años [4, 5, 23, 25, 26], estos suponen que las imágenes son invariantes bajo un conjunto de transformaciones (traslación, escala y rotación), por lo que son una herramienta muy útil en la identificación de objetos, ya que lo reconocerán independientemente de la posición, orientación y tamaño en el que se presente [23]. Dichas técnicas son muy efectivas, sin embargo, por lo regular son muy costosas

operacionalmente en tiempo de cómputo.

Con el objetivo de reducir el tiempo de cómputo, en este trabajo se propone un único espacio de clasificación, en lugar de tener múltiples planos de salida como se hace en los sistemas de reconocimiento de patrones que utilizan funciones de correlación [23, 25] o métricas [4, 5, 19, 28]. El documento se organiza como sigue: en la Sección 2 se explica cómo se crea el descriptor RFM. En la Sección 3 se construyen los espacios de clasificación de al menos el 95.4% de nivel de confianza. La Sección 4 muestra los resultados. Por último, en la Sección 5 se presentan las conclusiones.

2 Sistema de Reconocimiento de Patrones RFM Invariante a Traslación, Escala y Rotación

En la Fig. 1a, 1b y 1c se muestran tres imágenes en blanco y negro con cuadrados que presentan diferente factor de traslación, escala y rotación. A estas imágenes se les llamó K_1 , K_2 y K_3 . K_1 contiene el cuadrado que está trasladado con respecto al centro de la imagen, escalado un 10% con respecto a K_2 y sin rotación. K_2 no presenta traslación ni escalamiento y está rotado 22° . K_3 es la imagen que contiene el cuadrado que está trasladado con respecto al centro de la imagen, escalado -10% con respecto a la imagen K_2 y rotado 45° .

En el sistema de reconocimiento de patrones en imágenes digitales a color Radon-Fourier-Mellin (RFM), el primer paso es obtener la invarianza a traslación. Para esto se utiliza el módulo de la transformada de Fourier [26], esto es

$$A_i = \left| \mathcal{F}\{K_i\} \right|, \quad i = 1, 2, 3. \quad (1)$$

Solamente por cuestiones de visualización, en la Fig. 1b, 1d y 1f se muestra el módulo de la transformada de Fourier en escala logarítmica de las imágenes K_1 , K_2 y K_3 , respectivamente. Se observa que A_{K_1} , A_{K_2} y A_{K_3} presentan una deformación debido a las diferentes escalas de las

imágenes de los cuadrados. También se observa que A_{K_2} y A_{K_3} presentan una rotación con el mismo ángulo que K_2 y K_3 .

El siguiente paso en el sistema RFM es incorporar la invarianza a escala. Esto se hará por medio de la transformada analítica de Fourier-Mellin (TAFM) que se calculará vía la transformada de Fourier [4, 19, 27], mediante

$$M(k, c) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_0^{2\pi} A(e^\rho, \theta) e^{\rho\sigma} e^{-i(k\theta + c\rho)} d\theta d\rho, \quad (2)$$

donde A es la imagen obtenida en la ecuación 1, $\rho = \ln(r)$ y $\sigma \in \mathbb{R}^+$.

Sin embargo, la ecuación 2 no es invariante a escala, propiedad que sí posee la transformada de Fourier-Mellin. Para recuperar la invariancia a escala, se normaliza el módulo de la transformada analítica de Fourier-Mellin por su valor en el centro de coordenadas [4, 24, 27], es decir

$$S_M(k, c) = \left| \frac{M(k, c)}{M(c_x, c_y)} \right|, \quad (3)$$

donde (c_x, c_y) es el píxel central de la imagen.

La Fig. 2a, 2c y 2e muestra las transformada analítica de Fourier-Mellin de A_{K_1} , A_{K_2} y A_{K_3} , respectivamente. Se observa que las imágenes no presentan invariancia a escala y que la rotación se manifiesta por medio de una traslación circular en la variable angular. La Fig. 2b, 2d, 2f presenta la transformada analítica de Fourier-Mellin normalizada de la Fig. 2a, 2c y 2e, respectivamente, donde se observa la respuesta del sistema de reconocimiento de patrones invariante a traslación y escala.

Por último, la invarianza a rotación se obtiene por medio de la transformada de Radon,

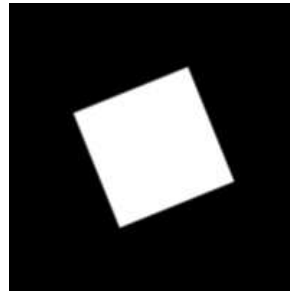
$$\mathcal{R}(p, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_M(k, c) \delta(p - k \cos \theta - c \sin \theta) dk dc, \quad (4)$$

donde $S_M(k, c)$ es la imagen obtenida en la ecuación 2, y δ es la función delta de Dirac, $-\infty < p < \infty$ y $0 \leq \theta < \pi$.

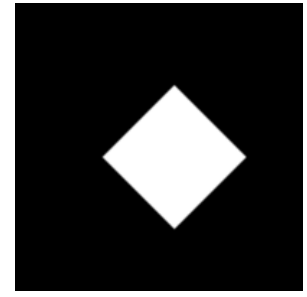
En la Fig. 3 se muestra la respuesta del descriptor RFM invariante a traslación, escala y rotación de las imágenes K_1 , K_2 y K_3 , respectivamente.



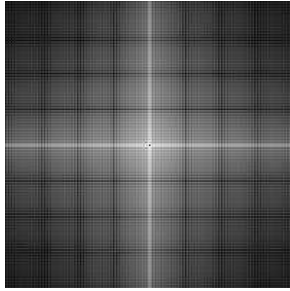
(a) K_1 .



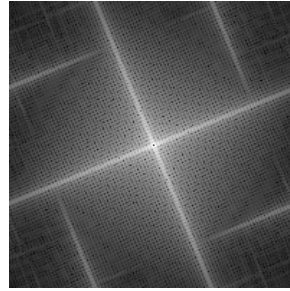
(c) K_2 .



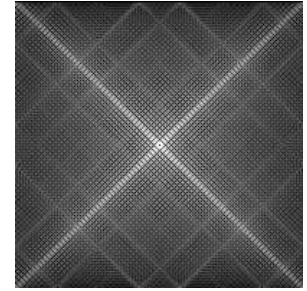
(e) K_3 .



(b) $A_{K_1} = |\mathcal{F}\{K_1\}|$.

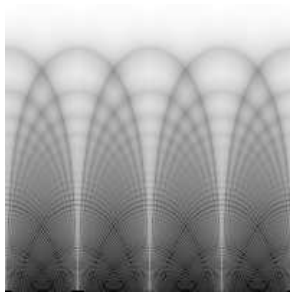


(d) $A_{K_2} = |\mathcal{F}\{K_2\}|$.

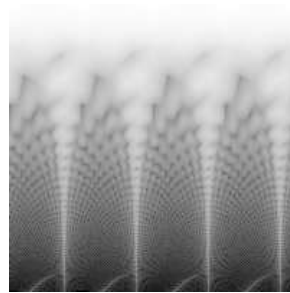


(f) $A_{K_3} = |\mathcal{F}\{K_3\}|$.

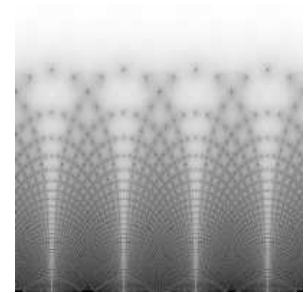
Figura 1: El módulo de la transformada de Fourier de imágenes digitales (señales 2D). Por cuestiones de visualización, el módulo de la transformada de Fourier de las imágenes se muestra en escala logarítmica.



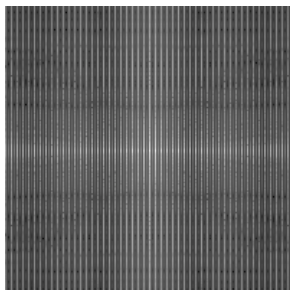
(a) $K_1 = A_{K_1}(e^\rho, \theta)e^{\rho\sigma}$.



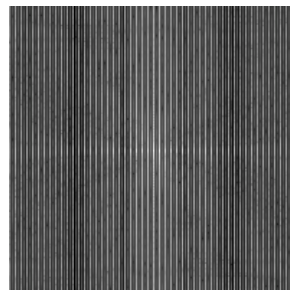
(c) $K_2 = A_{K_2}(e^\rho, \theta)e^{\rho\sigma}$.



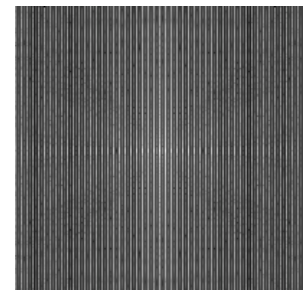
(e) $K_3 = A_{K_3}(e^\rho, \theta)e^{\rho\sigma}$.



(b) S_{K_1} .



(d) S_{K_2} .



(f) S_{K_3} .

Figura 2: El módulo de la transformada analítica de Fourier-Mellin normalizada usando $\sigma = 1$. Por cuestiones de visualización, el módulo de la transformada analítica de Fourier-Mellin se muestra en escala logarítmica.

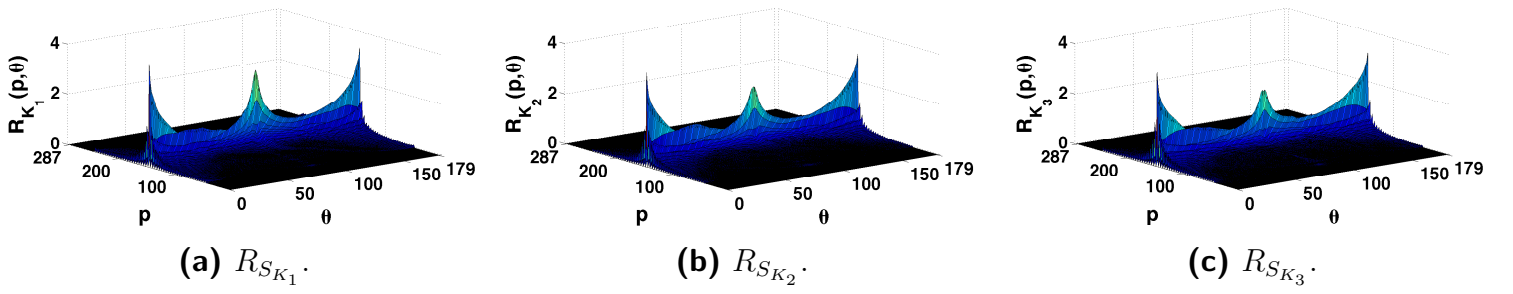


Figura 3: Ejemplos de la respuesta del sistema de reconocimiento de patrones RFM.

Para reducir el número de operaciones y por ende el tiempo de cómputo del sistema de reconocimiento de patrones en la etapa de la clasificación, se construyó la firma 1D invariante a traslación, escala y rotación, mediante

$$V(p) = \sum_{\theta=0}^{179} R_A(p, \theta), \quad (5)$$

donde R_A es la imagen obtenida por medio de la ecuación 4 y $p = 0, \dots, N$, con N representando el número de renglones de la imagen R_A . La ecuación 5 no es otra cosa más que la suma de los valores de intensidad por renglón de la imagen R_A (valores marginales por renglón).

3 El espacio de Clasificación

La base de datos de imágenes de referencia a color utilizada para mostrar la eficiencia del sistema de reconocimiento de patrones RFM consta de 18 imágenes digitales de mariposas, Fig. 4. Cada imagen de la Fig. 4 fue rotada 360° con $\Delta\theta = 1^\circ$ como tamaño de paso, cada una de esas 360 imágenes se escalaron $\pm 15\%$, variando la escala con un tamaño de paso $\Delta k = 1\%$. En total se obtuvieron 11,160 diferentes imágenes a color de una mariposa. Al trabajar en el espacio de color RGB, las imágenes se representan a través de tres imágenes monocromáticas, las cuales corresponden a los canales rojo, verde y azul, respectivamente. De esa manera, se tendrá un total de 33,480 imágenes en escala de grises para cada imagen de referencia de la Fig. 4.

Para construir un espacio de clasificación, se trabajará con la potencia de la firma 1D, ecua-

ción 5. La potencia se define como

$$P_{\theta,k}^C = \frac{1}{N} \sum_{p=1}^N \left(V_{\theta,k}^C(p) \right)^2, \quad (6)$$

donde $C = R, G$ o B (el canal de color) y N es el tamaño de la firma.

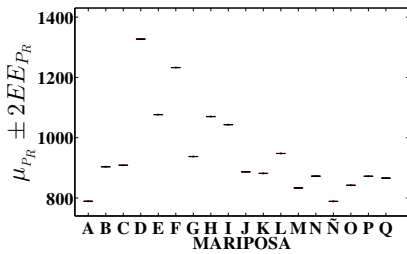
Para cada canal de cada imagen se generaron 11,160 imágenes monocromáticas, por lo que por cada imagen se generan 11,160 elementos escalares correspondientes a las potencias de las firmas 1D invariantes a traslación, escala y rotación. Por el teorema del límite central podemos suponer que los datos del conjunto tienen un comportamiento aproximado al de una distribución normal con media muestral $\mu_{P_C} \sim \mu$ y con un error estándar $\frac{\sigma_{P_C}}{\sqrt{n}}$ ($1EE_{P_C}$), donde n es el tamaño de la muestra, para este ejemplo $n = 11,160$. Por medio de esto, se generaron intervalos de confianza del 95.4% de la forma $[\mu_{P_C} - 2EE_{P_C}, \mu_{P_C} + 2EE_{P_C}]$.

En la Fig. 5 se muestra el diagrama de cajas de los intervalos de confianza del 95.4% construidos a partir de las potencias de cada una de las imágenes. En el eje horizontal se indica la imagen monocromática de referencia y en el eje vertical el valor de la media de los datos P_C con $\pm 2EE_{P_C}$. El punto central en color negro representa el valor de la media muestral μ_{P_C} , el rectángulo en negro corresponde a los valores de la media $\mu_{P_C} \pm 1EE_{P_C}$ y los bigotes en rojo a los valores de la media $\mu_{P_C} \pm 2EE_{P_C}$.

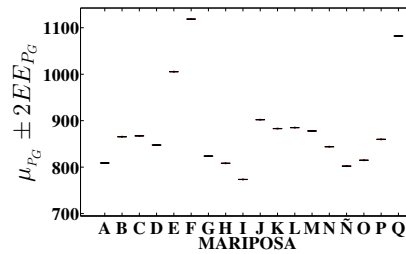
Para las imágenes de referencia de la Fig. 4, el espacio de clasificación se construyó mediante los intervalos mostrados en la Fig. 5. Se generó un ortoedro para cada imagen cuyas aris-



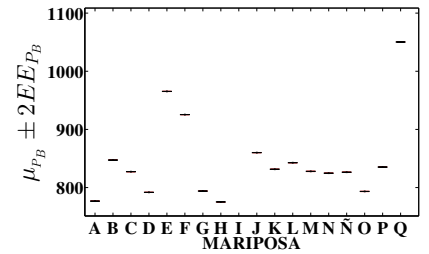
Figura 4: Imágenes a color de mariposas.



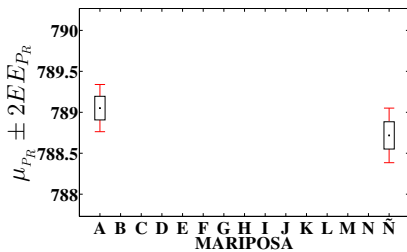
(a)



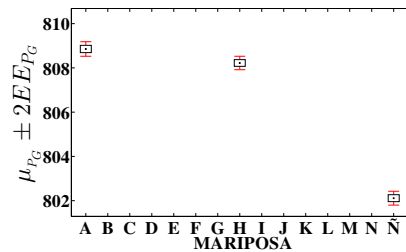
(c)



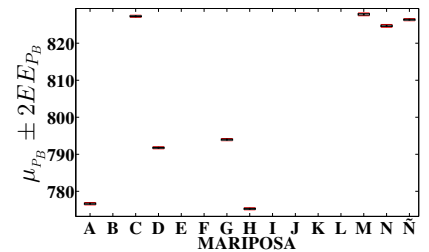
(e)



(b)



(d)



(f)

Figura 5: Intervalos de confianza. (a), (c) y (e) intervalos de confianza del 95.4% generados a partir de los conjuntos P_R , P_G , P_B , respectivamente. (b), (d) y (f) ampliación en la zona de A-Ñ de las Fig. 5a, 5c y 5e, respectivamente.

tas corresponden a los intervalos de confianza del 95.4% de los canales R , G y B . En la Fig. 6a se muestran dichos espacios de clasificación. Para mostrar con claridad que el sistema de reconocimiento de patrones RFM clasifica eficientemente cada una de las imágenes, se presenta una ampliación en la Fig. 6b donde se observa que ninguno de los ortoedros se traslapa. Logrando de esa manera tener un único espacio de clasificación y así reducir el tiempo de cómputo en la

etapa de clasificación de las imágenes digitales.

4 Conclusiones

Basado en la teoría de la transformada de Fourier, la transformada de Fourier-Mellin y la transformada de Radon, se desarrolló el sistema de reconocimiento de patrones Radon-Fourier-Mellin para imágenes digitales a color que es invariante a traslación, escala y rotación con un

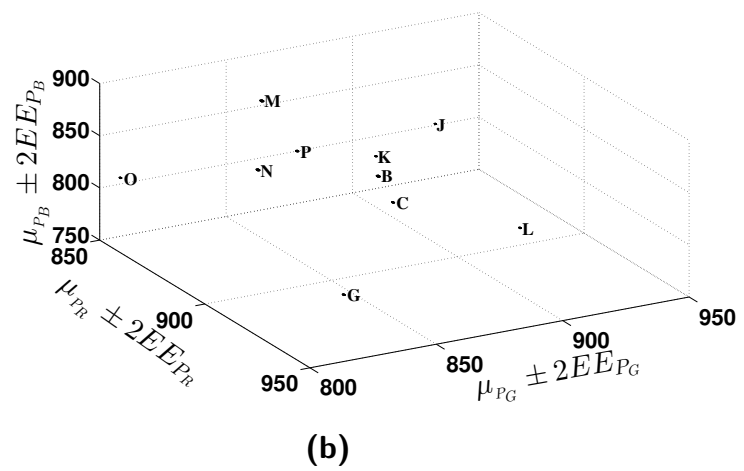
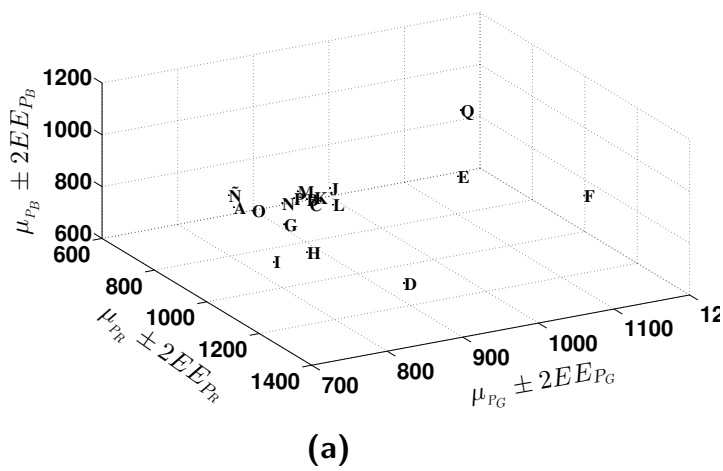


Figura 6: Espacio de clasificación. (a) Ortoedros de confianza del 95.4 % para la base de datos de la Fig. 4. (b) Amplificación de la zona más concurrida de la Fig. 6a.

nivel de confianza de al menos el 95.4 %.

Los sistemas de reconocimiento de patrones RFM existentes en la literatura utilizan la función de correlación 2D en la etapa de clasificación, esto significa que trabajan con imágenes (matrices) invariantes a traslación, escala y rotación. En este trabajo se muestra una técnica de clasificación que permite reducir el tiempo de cómputo, por medio de firmas 1D invariantes a traslación, escala y rotación, puesto que ahora sólo se requiere usar la correlación 1D. Más aún, en lugar de trabajar con múltiples planos de salida, como se hace en los sistemas de reconocimiento de patrones que utilizan funciones de correlación, se generó un único espacio de clasificación utilizando las potencias de las firmas, esto permite reducir aún más el número de operaciones y por ende, el tiempo de cómputo.

Agradecimientos. A CONACYT por la beca otorgada para la realización de la tesis bajo el proyecto de Ciencia Básica SEP-CONACYT 2011 con número 169174 y titulado reconocimiento de patrones en imágenes digitales mediante máscaras binarias de anillos concéntricos y firmas unidimensionales invariantes a posición, rotación y escalamiento.

Referencias

- [1] Arodz, T. Invariant object recognition using Radon-Based transform. *Computing and Informatics*, 24: 183–199, 2012.
- [2] Bracewell R. N. The Fourier Transform and its Applications. *McGraw-Hill Higher Education*, 2000.
- [3] Cristancho A. A., Rojas C. L. Algoritmo de tomografía local basado en la transformada discreta wavelet. *Departamento de Ciencias Básicas, Universidad EAFIT, Medellín*, 2011.
- [4] Derrode S., Ghorbel F. Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description. *Computer vision and image understanding*, 83: 57–78, 2010.
- [5] Derrode S. Représentation de formes planes à niveaux de gris par différentes approximations de Fourier-Mellin analytique en vue d’indexation de bases d’images. *Tesis Doctoral, Université de Rennes I*, 1999.
- [6] Epstein, C. L. Introduction to the mathematics of medical imaging. *Siam*, 2008.
- [7] ELouedi I., Fournier R., Nait-Ali A., Hamouda A. The polynomial discrete Radon transform. *Signal, Image and Video Processing*, 1–10, 2014.
- [8] Essaidani D., Seddik H., Ben B., E. Asynchronous invariant digital image watermarking in Radon field for resistant encrypted watermark. *International Journal of Network Security*, 2014.
- [9] Feldman J. Discrete-time Fourier Series and Fourier Transforms. 2007.
- [10] González R. C., Woods R. E. Digital Image Processing. *Prentice Hall*, New York 2002.

- [11] González R. C., Woods R. E., Eddins S. L. Digital Image Processing Using MATLAB. *Tata McGraw Hill Education Private Limited*, New York 2010.
- [12] HSU P. H. Análisis de Fourier. *Prentice Hall*, New York 1970.
- [13] Kailey B. Mathematic of medical imaging inverting the radon transform.
- [14] Lokenath, D., Dambaru, B. Integral Transforms and Their Applications. *Chapman & Hall/CRC*, EE.UU. 2007.
- [15] Lokenath, D., Dambaru, B. Integral Transforms and Their Applications. *Chapman & Hall/CRC*, EE.UU. 2015.
- [16] Miciak M. Radon transformation and principal component analysis method applied in postal address recognition task. *IJCSA*, 7.3: 33–44, 2010.
- [17] Osgood B. The Fourier Transform and its Applications. *Notas de clase, Electrical Engineering Department*, Stanford University 2007.
- [18] Butzar P. L., Jansche S. Mellin-Fourier Series and the Classical Mellin Transform. *Computers & Mathematics with Applications*, 40: 49–62, 2000.
- [19] Quintero V. L. C. Descriptor invariante a posición y escala para imágenes digitales usando la transformada analítica de Fourier-Mellin. *Tesis de Licenciatura*, Universidad Nacional Autónoma de Nicaragua UNAN-LEÓN, 2015.
- [20] Rodríguez M. R., Sossa A. J. H. Procesamiento y Análisis Digital de imágenes. *Alfaomega*, México, 2012.
- [21] Ross M. S. Introduction to probability and statistics for engineers and scientists. *Elsevier Academic Press*, EE.UU., 2004.
- [22] Shih F. Y. Image processing and pattern recognition. *Wiley*, New Jersey, 2010.
- [23] Solís V. A., Álvarez B. J., Solorza S. Adaptive non-linear correlation with a binary mask invariant to rotation and scale. *Optics Communications*, 339: 185–193, 2015.
- [24] Solorza-Calderón, S., Verdugo-Olachea J. A RFM Pattern Recognition System Invariant to Rotation, Scale and Translation. *Iberoamerican Congress on Pattern Recognition*, Springer International Publishing, 2015.
- [25] Solorza S., Álvarez B. J. Position and rotation-invariant pattern recognition system by binary rings masks. *Journal of Modern Optics*, 62: 851–864, 2015.
- [26] Solorza S., Álvarez B. J. Reconocimiento de patrones en imágenes digitales de fósiles de diatomeas. *Notas de clase*.
- [27] Verdugo-Olachea, J. Reconocimiento de Patrones en Imágenes Digitales usando la Transforma de Radon. *Tesis de Licenciatura*, Universidad Autónoma de Baja California, Facultad de Ciencias, 2015.
- [28] Villegas S. M. Determinacion de la posición de objetos en imágenes usando la Transformada de Fourier-Mellin Analítica. *Proyecto Final de Carrera*, Universidad Politécnica de Valencia 2004.
- [29] Zhang, Y., Wu, L. A rotation invariant image descriptor based on Radon transform. *International Journal of Digital Content Technology and its Applications* 5: 209–217, 2011.
- [30] Zwicke P. E., Kiss Jr, I. A new implementation of the Mellin transform and its application to radar classification of ships. *Pattern Analysis and Machine Intelligence*, 2: 191–199, 1983.