

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
FACULTAD DE CIENCIAS



LICENCIATURA EN FÍSICA

RECONOCIMIENTO DE PATRONES EN IMÁGENES DIGITALES
USANDO LA TRANSFORMADA DE FOURIER NORMALIZADA Y LA
TRANSFORMADA ANALÍTICA DE FOURIER-MELLIN
NORMALIZADA

TESIS

que para cubrir los requisitos necesarios para obtener el grado de

LICENCIADO EN FÍSICA

Presenta:

RODRIGO VÉLEZ RÁBAGO

Ensenada, Baja California, México, **abril** de 2016

TESIS DEFENDIDA POR

Rodrigo Vélez Rábago

Y APROBADA POR EL SIGUIENTE COMITÉ



Dra. Selene Solorza Calderón

Director del Comité



Dr. Jesús R. Lerma Aragón

Miembro del Comité



M.C. Gloria E. Rubí Vazquez

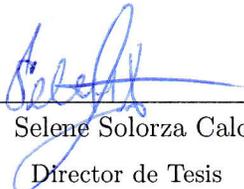
Miembro del Comité

Ensenada Baja California, México. Abril de 2016.

RESUMEN de la tesis de **RODRIGO VÉLEZ RÁBAGO**, presentada como requisito para la obtención del grado de LICENCIADO EN FÍSICA. Ensenada, Baja California, abril de 2016.

**RECONOCIMIENTO DE PATRONES EN IMÁGENES DIGITALES
USANDO LA TRANSFORMADA DE FOURIER NORMALIZADA Y LA
TRANSFORMADA ANALÍTICA DE FOURIER-MELLIN
NORMALIZADA**

Resumen aprobado por:



Dra. Selene Solorza Calderón

Director de Tesis

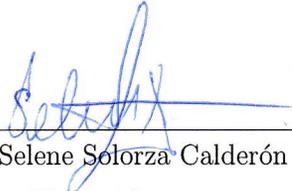
Este trabajo presenta un sistema de reconocimiento de patrones en imágenes digitales invariante a traslación, escala y rotación. El sistema utiliza la transformada de Fourier para lograr la invariancia a traslación y la transformada analítica de Fourier-Mellin normalizada para la invariancia a escala. A partir de la imagen resultante de la transformada analítica de Fourier-Mellin, se construyen dos firmas 1D: una obtenida por los marginales de los renglones (S_r) y la segunda generada por los marginales de las columnas (S_c). Ambas firmas son invariantes a traslación, escala y rotación. Para que el sistema de reconocimiento de patrones sea más robusto en la etapa de clasificación se necesita una tercera firma, en este trabajo se optó por usar la firma proveniente de la fase de la transformada de Fourier de S_c . Una vez que se asocian las tres firmas a la imagen, las características o atributos que la identificarán son las potencias de dichas firmas, con los que se construyen los ortoedros de confianza mediante la técnica estadística de diagramas de cajas. Con los ortoedros correspondientes a las imágenes de referencia en la base de datos, se genera el espacio de clasificación para el sistema de reconocimiento de patrones invariante a traslación, escala y rotación. De los espacios de clasificación obtenidos en las pruebas hechas al sistema de reconocimiento de patrones en imágenes digitales en blanco y negro, y en escala de grises, se llega a la conclusión de que tiene un nivel de confianza de al menos 95.4%.

Palabras Clave: Algoritmo de reconocimiento de patrones, transformada de Fourier, transformada de Fourier-Mellin, método de clasificación de imágenes digitales.

ABSTRACT of the thesis presented by **RODRIGO VÉLEZ RÁBAGO**, in fulfillment of the requirements of the degree of BACHELOR OF PHYSICS. Ensenada, Baja California, April 2016.

**DIGITAL IMAGE PATTERN RECOGNITION SYSTEM USING
NORMALIZED FOURIER TRANSFORM AND NORMALIZED
ANALYTICAL FOURIER-MELLIN TRANSFORM**

Approved by:



Dra. Selene Solorza Calderón
Thesis Advisor

This work presents a digital image pattern recognition system invariant to translation, scale and rotation. The system uses the Fourier transform to achieve the invariance to translation and the analytical Fourier-Mellin transform for the invariance to scale. The image obtained from the analytical Fourier-Mellin transform was used to construct two 1D signatures: one generated from the row marginals (S_r) and the second one obtained from the column marginals (S_c). Both signatures are invariant to translation, scale and rotation. In order to get a more robust pattern recognition system in the classification step, a third signature is needed, therefore this work uses the phase of the Fourier transform of S_c . Once the three signatures are associated with the image, the features or attributes that identify it are the powers of these signatures. The powers of the images are used in the confidence cuboids construction by the box-plot method. The classifier output space for the pattern recognition system invariant to translation, scale and rotation was generated with the cuboids corresponding to the reference images. From the cuboids classifier output space obtained using black and white digital images, and grayscale images, it is concluded that the pattern recognition system has at least a 95.4% confidence level.

Keywords: Pattern recognition algorithms, Fourier transform, Fourier-Mellin transform, digital image classification method.

*A mi madre, la fuente de
mi felicidad*

Agradecimientos

Agradezco sinceramente a mi asesora, Dra. Selene Solorza Calderón, por el apoyo brindado y por su infinita paciencia durante el desarrollo del presente trabajo. Su manera de trabajar es admirable y sus enseñanzas van más allá de lo académico.

A todos los profesores, estudiantes y personal de la Facultad de Ciencias de la Universidad Autónoma de Baja California de los que tanto he aprendido y que día a día han motivado mi crecimiento académico.

A mis padres, Carlos y Rosa Elia, por todos los sacrificios que hicieron para que hicieran esto posible, pero sobre todo por todo el apoyo y amor que me han proporcionado a lo largo de mi vida. A mis hermanos Rebeca, Sonia, Carlos y Ana por siempre estar ahí para mí.

A mis tíos Jesús, Tomás, Lucía y Luzma, a primo Gerardo y a mis amigos por su cariño y apoyo cuando más lo necesité.

A la Dra. Bogdanchikova, la Dra. Martynyuk y al CNyN por la gran huella que dejaron en mí y por su aportación a mi formación profesional.

Al CONACyT por la beca otorgada para la realización de esta 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.

Contenido

	Página
Resumen en español	i
Resumen en inglés	ii
Dedicatoria	iii
Agradecimientos	iv
Contenido	v
Lista de Figuras	vi
I. Introducción	1
II. Sistema de Reconocimiento de Patrones Invariante a Traslación, Escala y Rotación	3
II.1 Introducción	3
II.2 El sistema de reconocimiento de patrones	4
III. Nivel de Confianza del Sistema de Reconocimiento de Patrones	12
III.1 Imágenes en Blanco y Negro	12
III.2 Imágenes en escala de grises	14
IV. Conclusiones	16
REFERENCIAS	18

Lista de Figuras

Figura		Página
1	Ejemplos de escalas acromáticas http://4.bp.blogspot.com/-OGP7TV8IjaA/TdHrXmH-t9I/AAAAAAAAAM0/wOMQDFQZ_g8/s1600/13.jpg . . .	4
2	(a) $I_1(x, y)$. (b) $A_1(u, v)$. (c) $I_2(x, y)$. (d) $A_2(u, v)$	6
3	Ejemplos de espectros de amplitudes de la TAFM. (a) Imagen I_1 : La letra de tipo arial B sin transformaciones geométricas. (b) Imagen I_2 : La letra de tipo arial B rotada 315° y escalada -20% . (c) Imagen I_3 : letra arial Q sin transformaciones geométricas. (d) $A_1(u, v)$. (e) $A_2(u, v)$. (f) $A_3(u, v)$. (g) $A_1(e^\rho, \theta)e^{\rho\sigma}$. (h) $A_2(e^\rho, \theta)e^{\rho\sigma}$. (i) $A_1(e^\rho, \theta)e^{\rho\sigma}$. (j) $S_1(k, \omega)$. (k) $S_2(k, \omega)$. (l) $S_3(k, \omega)$	8
4	Las firmas de una imagen. (a) Imagen en blanco y negro conteniendo la letra arial B. (b) La firma S_r de la Fig. 4(a). (c) La firma S_c de Fig. 4(a). (d) La firma S_p de Fig. 4(a).	9
5	Las firmas S_r : en color negro la de la Fig. 3(a), en color rojo la de la Fig. 3(b) y en color azul la de la Fig. 3(c).	10
6	Las firmas S_c : en color negro la de la Fig. 3(a), en color rojo la de la Fig. 3(b) y en color azul la de la Fig. 3(c).	11
7	Las firmas S_p : en color negro la de la Fig. 3(a), en color rojo la de la Fig. 3(b) y en color azul la de la Fig. 3(c).	11
8	Base de datos de imagenes de referencia BW.	12
9	Espacio de clasificación de ortoedros de confianza para las imágenes de referencia de la Fig. 8.	13
10	Coil-20 object image library[1].	14
11	Espacio de clasificación de ortoedros de confianza para las imágenes en escalas de grises de la base de datos en Fig. 10.	15

Capítulo I

Introducción

El reconocimiento de patrones en imágenes digitales es empleado todos los días con distintos fines. Por citar algunos ejemplos: policías utilizan sistemas automatizados de reconocimiento de patrones para identificar sospechosos usando imágenes digitales de huellas dactilares; sistemas de seguridad permiten o niegan el acceso a una persona analizando los patrones de su iris. El trabajo de algunas áreas de las ciencias se apoya en estos sistemas de reconocimiento de patrones para clasificar objetos experimentales sin necesidad de hacerlo manualmente.

En la actualidad se realizan estudios para el desarrollo de paquetes de cómputo dedicados al reconocimiento de patrones invariantes a posición, orientación y escala [2-7]. Existen diversas maneras de obtener las invariancias deseadas, la transformada de Fourier es de las técnicas más utilizadas para lograrlo [2, 3, 4]. La transformada de Fourier acompañada de otras operaciones puede ayudar a crear algoritmos eficientes de reconocimiento de patrones. Su utilidad se ve reflejada en las relaciones que existen entre la transformada de Fourier de dos imágenes, donde una es una copia rotada o trasladada de la otra imagen. Sin embargo, la transformada de Fourier tiene sus limitaciones, ya que carece de una forma directa de relacionar una imagen y una copia escalada de dicha imagen. Este problema se puede solucionar haciendo uso de otra herramienta, la transformada analítica de Fourier-Mellin. En la década de los setenta la comunidad del área de óptica introdujo la transformada de Fourier-Mellin para el reconocimiento de patrones en imágenes ópticas [5]. Sin embargo, ésta presentaba al-

gunas dificultades para la estimación numérica de dicha transformada. Las dificultades desaparecieron con la introducción de la transformada analítica de Fourier-Mellin [5]. La principal ventaja de ésta radica en la relación que existe entre la transformada de dos imágenes, donde una imagen es una copia escalada o rotada de la otra. Esto implica la posibilidad de elaborar una metodología que utilice la transformada de Fourier y la transformada analítica de Fourier-Mellin para lograr las invariancias en rotación, escala y traslación.

La tesis comprende cuatro capítulos: en el primero se realiza una introducción general al reconocimiento de patrones en imágenes digitales. En el capítulo II se presenta la base teórica y la metodología para construir el sistema de reconocimiento de patrones. En el capítulo III se explica la manera de generar el espacio de clasificación de ortodros de confianza para imágenes en blanco y negro, y en escala de grises. Finalmente, en el capítulo IV se presentan las conclusiones.

Capítulo II

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

II.1 Introducción

Una imagen digital se define como una función bidimensional $I(x, y)$, donde x y y son las coordenadas espaciales y al entero $I(x, y)$ se le denomina la intensidad de la imagen en ese punto. Una imagen digital se representa por medio de una matriz $N \times M$, [10], dada por

$$I = \begin{bmatrix} I(1, 1) & I(1, 2) & \cdots & I(1, M) \\ I(2, 1) & I(2, 2) & \cdots & I(2, M) \\ \vdots & \vdots & \ddots & \vdots \\ I(N, 1) & I(N, 2) & \cdots & I(N, M) \end{bmatrix}. \quad (1)$$

Los posibles valores para $I(i, j)$ se determinan por el tipo de clase que se utilizará para procesarla: si se selecciona la clase *uint8*, entonces $I(i, j) = 0, \dots, 255$ y si es *uint16*, entonces $I(i, j) = 0, \dots, 65535$ [10]. En la Fig. 1 se observa gráficamente como es que la clase de la imagen rige el número de posibles valores que la intensidad tomará.

Una escala acromática es una modulación continua del blanco al negro, es decir una escala de grises. La escala de grises se utiliza para establecer comparativamente tanto el valor de la luminosidad de los colores puros como el grado de claridad de las correspondientes gradaciones de este color puro. En la Fig. 1 se aprecian algunos ejemplos de escalas acromáticas, las cuales difieren por el número de bits por píxel,

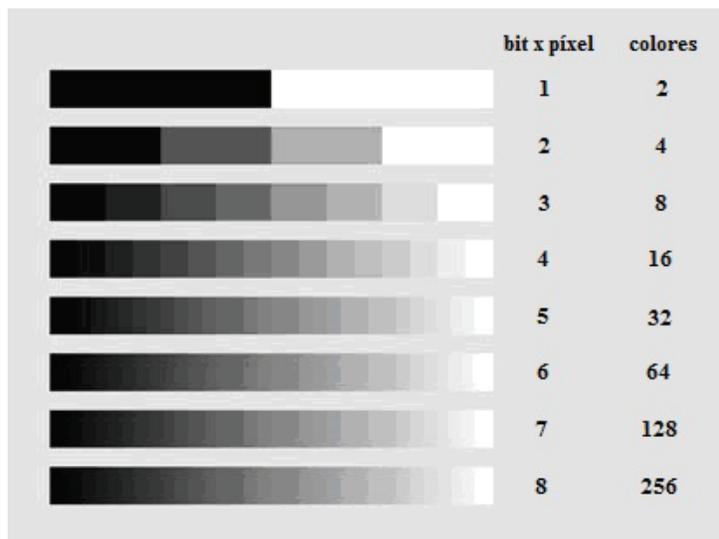


Figura 1: Ejemplos de escalas acromáticas http://4.bp.blogspot.com/-OGP7TV8IjaA/TdHrXmH-t9I/AAAAAAAAAM0/wOMQDFQZ_g8/s1600/13.jpg

es decir, el número de posibles valores que la intensidad puede tomar. Una escala monocromática es aquella en la que hay un sólo color y se forma con todas las variaciones de ese color, bien añadiendo blanco, negro o la mezcla de ambos, que es el color gris.

II.2 El sistema de reconocimiento de patrones

La transformada discreta de Fourier de la imagen $I(x, y)$, denotada por $F(u, v)$, está dada por

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) e^{-i2\pi(\frac{ux}{M} + \frac{vy}{N})}, \quad (2)$$

donde $u = 0, \dots, M-1$ y $v = 0, \dots, N-1$, y la correspondiente transformada discreta inversa de Fourier está dada por

$$I(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{i2\pi(\frac{ux}{M} + \frac{vy}{N})}, \quad (3)$$

para $x = 0, \dots, M - 1$ y $y = 0, \dots, N - 1$ [10]. Si la imagen $I_2(x, y)$ es una versión trasladada, rotada y escalada de la imagen en escala de grises $I_1(x, y)$, entonces

$$I_2(x, y) = I_1(\alpha x \cos \phi + \alpha y \sin \phi - \hat{x}, -\alpha x \sin \phi + \alpha y \cos \phi - \hat{y}), \quad (4)$$

donde \hat{x} y \hat{y} son los valores de traslación del objeto en los ejes horizontal y vertical, respectivamente y, ϕ y $\alpha \neq 0$ son los factores de rotación y escala del objeto, respectivamente [11]. Por lo tanto, la transformada de Fourier de las imágenes I_1 e I_2 se relacionan de la siguiente manera

$$F_2(u, v) = e^{-i[u(\frac{\hat{x}}{\alpha} \cos \phi - \frac{\hat{y}}{\alpha} \sin \phi) + v(\frac{\hat{y}}{\alpha} \cos \phi + \frac{\hat{x}}{\alpha} \sin \phi)]} F_1(\alpha u \cos \phi + \alpha v \sin \phi, -\alpha u \sin \phi + \alpha v \cos \phi), \quad (5)$$

donde F_1 y F_2 representan a las transformadas de Fourier de las imágenes I_1 e I_2 , respectivamente [10].

Como el espectro de amplitud de la transformada de Fourier de una imagen dada $I(x, y)$ se define por

$$A(u, v) = |F(u, v)| = \sqrt{R_F^2(u, v) + I_F^2(u, v)}, \quad (6)$$

donde R_F e I_F son la parte real e imaginaria de $F(u, v)$, respectivamente. Entonces, los espectros de amplitud de I_1 e I_2 están relacionados mediante

$$A_2(u, v) = A_1(\alpha u \cos \phi + \alpha v \sin \phi, -\alpha u \sin \phi + \alpha v \cos \phi). \quad (7)$$

En la Fig. 2(a) se muestra la imagen en blanco y negro (BW) de 369×369 píxeles de la letra B del alfabeto latino en fuente Arial y en la Fig. 2(b) su correspondiente espectro de amplitud. En la Fig. 2(c) se observa una versión trasladada de la letra B y su espectro de amplitud se visualiza en la Fig. 2(d). Para poder percibir fácilmente el contenido de las frecuencias de los espectros de amplitud, estos se grafican en escala

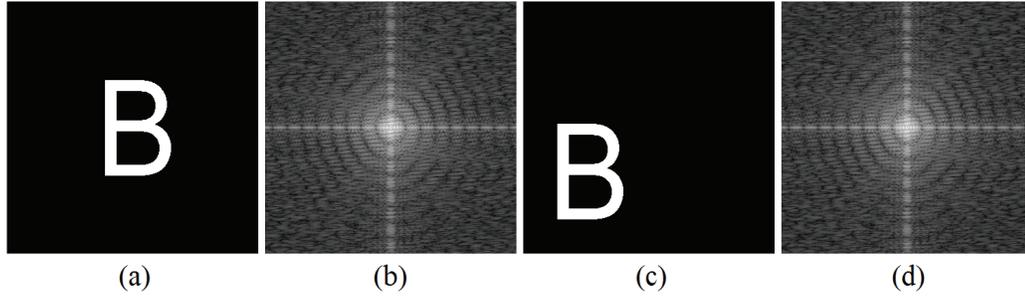


Figura 2: (a) $I_1(x, y)$. (b) $A_1(u, v)$. (c) $I_2(x, y)$. (d) $A_2(u, v)$.

logarítmica. Al comparar las imágenes de los espectros se observa que son iguales, esto quiere decir que el sistema es invariante a traslación. Analíticamente se deduce de la Ec.(7) donde $A_2(u, v)$ es una versión rotada y escalada de $A_1(u, v)$ pero \hat{x} y \hat{y} no son parte de la relación, por lo tanto se ha logrado la invariancia a traslación.

El espectro de amplitud $A(u, v)$ de la imagen dada $I(x, y)$ es el que se usa en la transformada analítica de Fourier-Mellin (TAFM) pero expresado en coordenadas logarítmico-polar, con el fin de calcular la TAFM vía la transformada de Fourier [9], es decir

$$M(k, \omega) = \mathcal{M}\{A(e^\rho, \theta)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_0^{2\pi} A(e^\rho, \theta) e^{\rho\sigma} e^{-i(k\theta + \rho\omega)} d\theta d\rho, \quad (8)$$

con $\rho = \ln(r)$ y $\sigma > 0$; en este trabajo $\sigma = 0.5$. Por ejemplo, en la Fig. 3(d) se observa el espectro de amplitud de la Fig. 3(a) y en la Fig. 3(g) el espectro de amplitud en coordenadas logarítmico-polar. La Fig. 3(b) es una versión trasladada, escalada y rotada de la Fig. 3(a) y su correspondiente espectro de amplitud en coordenadas cartesianas está dado en la Fig. 3(e). Comparando la Fig. 3(d) y la Fig. 3(e) se aprecia que el espectro de amplitud muestra el mismo ángulo de rotación y cómo se disminuyó la escala del objeto, el espectro presenta una dilatación con respecto al de la imagen original, si hubiera habido un incremento en la escala entonces el espectro presentaría una

contracción. Por otro lado, al comparar la Fig. 3(g) y la Fig. 3(h), donde se exhiben los espectros de amplitud de la Fig. 3(a) y la Fig. 3(b) en coordenadas logarítmico polar: en la Fig. 3(h) se observan dos picos brillantes largos, mientras que en la Fig. 3(g) se aprecia un pico en el centro y otro pico dividido en los extremos laterales de la imagen. Si en la Fig. 3(g) se realiza un desplazamiento angular horizontal de la misma magnitud que presenta el objeto de la Fig. 3(h), entonces los picos de ambos espectros van a coincidir. Por lo que el efecto de rotación genera un desplazamiento angular horizontal en los espectros de las imágenes. Además, el efecto de la escala se denota en la longitud de los picos ya mencionados, se observa que los de la Fig. 3(h) son más cortos que los de la Fig. 3(g). La Fig. 3(c) muestra la letra en fuente Arial Q, su espectro de amplitud en coordenadas cartesianas está dado en la Fig. 3(f) y en coordenadas logarítmico-polar en la Fig. 3(i), ambos muestran un patrón muy diferente al de las imágenes con la letra B, lo que permite caracterizar de manera única a las imágenes.

La Ec.(8) no es invariante a escala ni a rotación, pero normalizando la TAFM por su valor en el píxel central (c_x, c_y) , el espectro de amplitud de la TAFM normalizada ya es invariante a escala [9], esto es

$$S(k, \omega) = \left| \frac{M(k, \omega)}{M(c_x, c_y)} \right|. \quad (9)$$

En las Fig. 3(j), Fig. 3(k) y Fig. 3(l) se muestran $S_1(k, \omega)$, $S_2(k, \omega)$ y $S_3(k, \omega)$, la transformada analítica de Fourier-Mellin normalizada de los espectros de amplitud asociados a las Fig. 3(a), Fig. 3(b) y Fig. 3(c), respectivamente, de ellas se observa que las Fig. 3(j) y Fig. 3(k) tienen patrones muy similares, mientras que la Fig. 3(l) no tiene mucho en común con ellas.

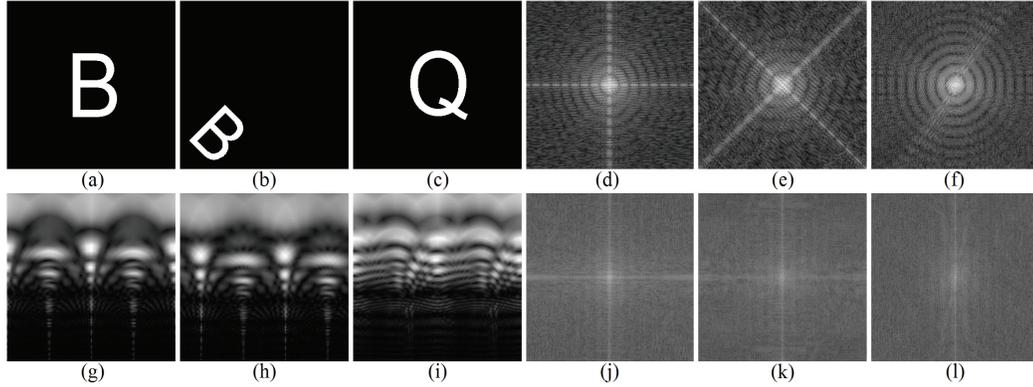
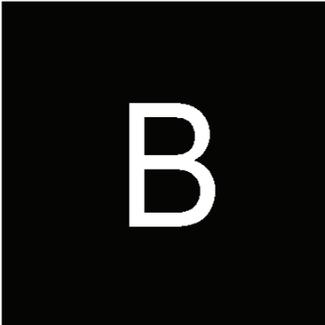


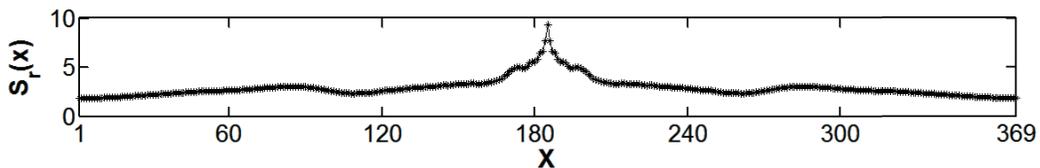
Figura 3: Ejemplos de espectros de amplitudes de la TAFM. (a) Imagen I_1 : La letra de tipo arial B sin transformaciones geométricas. (b) Imagen I_2 : La letra de tipo arial B rotada 315° y escalada -20% . (c) Imagen I_3 : letra arial Q sin transformaciones geométricas. (d) $A_1(u, v)$. (e) $A_2(u, v)$. (f) $A_3(u, v)$. (g) $A_1(e^\rho, \theta)e^{\rho\sigma}$. (h) $A_2(e^\rho, \theta)e^{\rho\sigma}$. (i) $A_3(e^\rho, \theta)e^{\rho\sigma}$. (j) $S_1(k, \omega)$. (k) $S_2(k, \omega)$. (l) $S_3(k, \omega)$.

Para obtener la invariancia a rotación y reducir la inversión en tiempo de cómputo, se construyen dos firmas 1D invariantes a traslación, escala y rotación mediante los marginales de renglón y columna de $S(k, \omega)$. A la firma generada a partir de los marginales de los renglones se le denomina S_r y a la firma basada en los marginales de las columnas se le denota por S_c . Una tercera firma invariante a traslación, escala y rotación se obtiene a partir de $S_p = \text{Arg}\{\mathcal{F}\{S_c\}\}$, donde \mathcal{F} es la transformada de Fourier de 1D y Arg es el argumento principal o la fase de un número complejo. En la Fig. 4 se observan las firmas correspondientes a la imagen en blanco y negro conteniendo la letra en fuente Arial B.

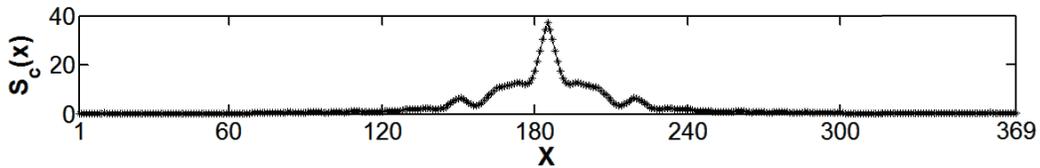
En las Fig. 5, Fig. 6 y Fig. 7 se observan las firmas S_r , S_c y S_p de tres imágenes distintas. Las firmas en color negro corresponden a la Fig. 3(a), en color rojo a la Fig. 3(b) y en color azul a la Fig. 3(c). Las tendencias de las curvas en negro y rojo para



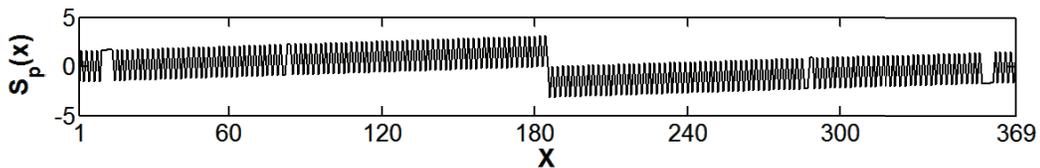
(a)



(b)



(c)



(d)

Figura 4: Las firmas de una imagen. (a) Imagen en blanco y negro conteniendo la letra arial B. (b) La firma S_r de la Fig. 4(a). (c) La firma S_c de Fig. 4(a). (d) La firma S_p de Fig. 4(a).

las tres firmas son muy parecidas, la diferencia se presenta por la rotación del objeto, sin embargo dichas diferencias no afectarán el rendimiento del sistema al momento de clasificar imágenes que presenten transformaciones geométricas de traslación, escala y rotación. Por otra parte, la tendencia de las curvas en azul difieren de las otras firmas, lo que permitirá que el sistema determine un patrón de otro de manera robusta.

Finalmente, para la construcción del espacio de clasificación (Capítulo III) el sistema de reconocimiento de patrones utiliza las potencias de las firmas [12], dadas por

$$P_q = \frac{\sum (S_q)^2}{N_q}, \quad (10)$$

donde $q = r, c, p$ y N_q es la longitud de la firma S_q .

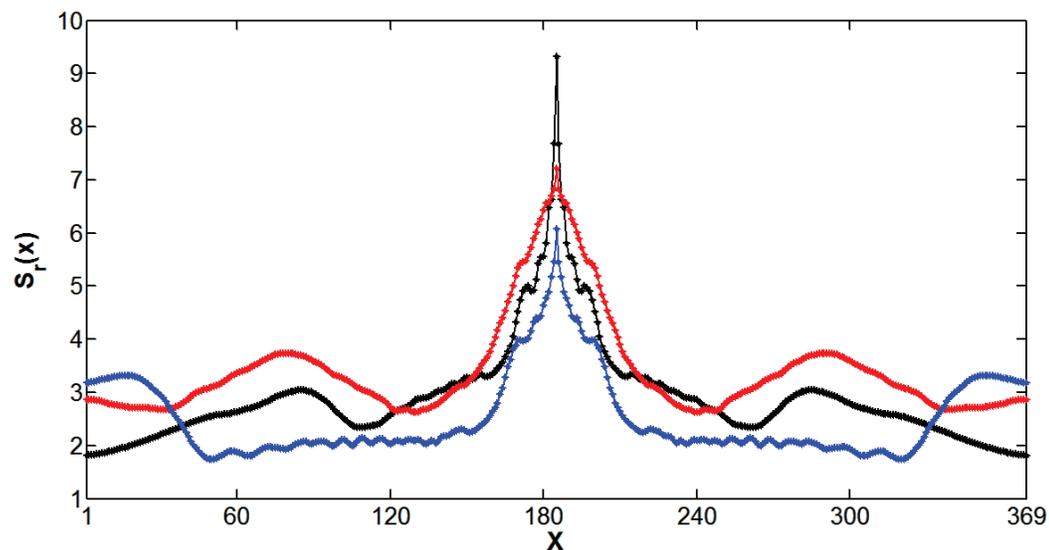


Figura 5: Las firmas S_r : en color negro la de la Fig. 3(a), en color rojo la de la Fig. 3(b) y en color azul la de la Fig. 3(c).

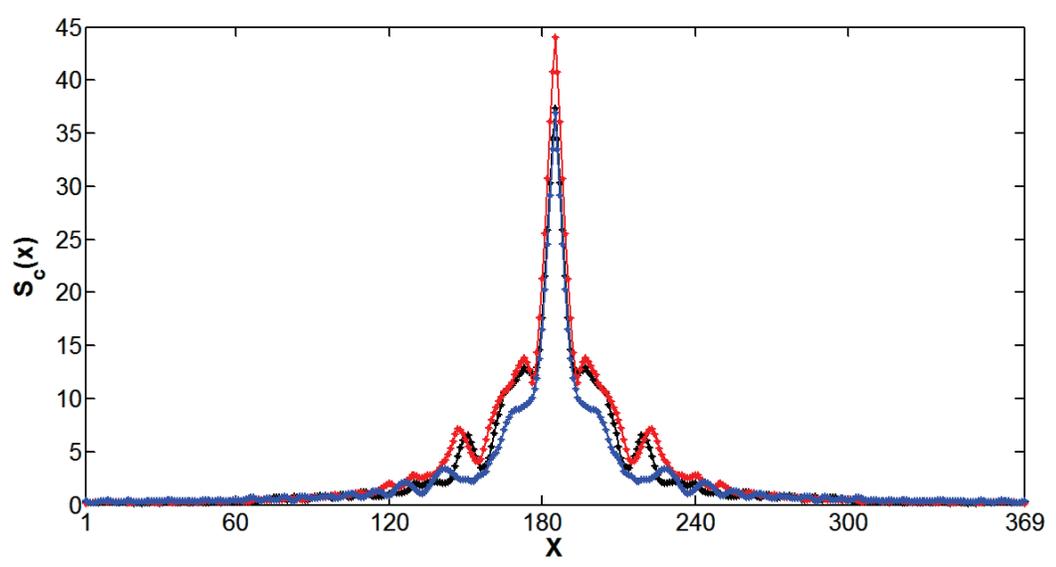


Figura 6: Las firmas S_c : en color negro la de la Fig. 3(a), en color rojo la de la Fig. 3(b) y en color azul la de la Fig. 3(c).

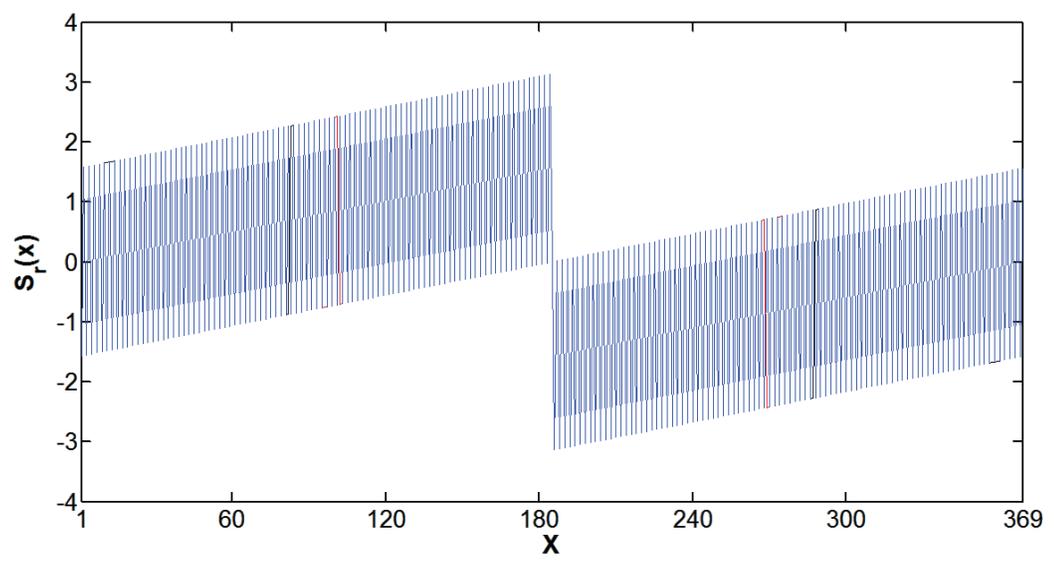


Figura 7: Las firmas S_p : en color negro la de la Fig. 3(a), en color rojo la de la Fig. 3(b) y en color azul la de la Fig. 3(c).

Capítulo III

Nivel de Confianza del Sistema de Reconocimiento de Patrones

Para probar el nivel de confianza del sistema de reconocimiento de patrones se utilizaron dos conjuntos de imágenes. El primer conjunto contiene imágenes en blanco y negro (BW) de 369×369 píxeles de las letras del alfabeto latino en formato Arial. El segundo conjunto está formado por imágenes de 318×318 píxeles en escala de grises seleccionadas de la librería de imágenes Coil-20[1].

III.1 Imágenes en Blanco y Negro

Para entrenar al sistema de reconocimiento de patrones, cada imagen en la base de datos de imágenes de referencia en la Fig. 8 fue rotada 360° con un paso de $\Delta\theta = 1^\circ$. Entonces por imagen se tienen 360 muestras rotadas, posteriormente cada una de ellas fue escalada $\pm 20\%$ con un paso de escala $\Delta h = 1\%$. Por lo tanto, se generaron $360 \times 41 = 14,760$ imágenes por cada imagen de referencia en la Fig. 8, obteniendo un total de $14,760 \times 11 = 162,360$ imágenes problema a clasificar.



Figura 8: Base de datos de imágenes de referencia BW.

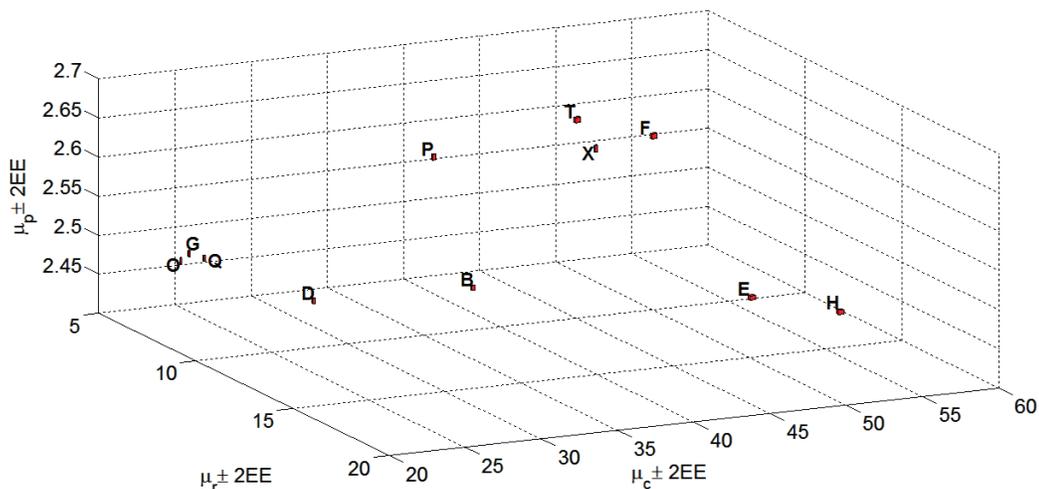


Figura 9: Espacio de clasificación de ortoedros de confianza para las imágenes de referencia de la Fig. 8.

Una vez generadas las muestras, se calcularon las firmas de cada una de ellas. Luego, se obtuvieron las correspondientes potencias de las firmas, para finalmente construir los intervalos de confianza de la siguiente manera: Sea R_k la k -ésima imagen de referencia en la base de datos, utilizando los 14,760 valores de las potencias de las firmas S_c^k se construyó el correspondiente intervalo de confianza del 95.4% usando el método estadístico de diagramas de cajas con $\mu_c \pm 2EE_c$, donde μ_c es la media de los valores de las potencias y EE_c es el error estándar. Los intervalos de confianza de S_r y S_p se construyeron de la misma manera que para S_r . Posteriormente, se formó un ortoedro cuyos lados están dados por los intervalos de confianza $\mu_c \pm 2EE_c$, $\mu_r \pm 2EE_r$, $\mu_p \pm 2EE_p$ y con vértices en las coordenadas: $(\mu_c - 2EE_c, \mu_r - 2EE_r, \mu_p - 2EE_p)$, $(\mu_c + 2EE_c, \mu_r - 2EE_r, \mu_p - 2EE_p)$, $(\mu_c + 2EE_c, \mu_r + 2EE_r, \mu_p - 2EE_p)$, $(\mu_c - 2EE_c, \mu_r + 2EE_r, \mu_p - 2EE_p)$, $(\mu_c - 2EE_c, \mu_r - 2EE_r, \mu_p + 2EE_p)$, $(\mu_c + 2EE_c, \mu_r - 2EE_r, \mu_p + 2EE_p)$, $(\mu_c - 2EE_c, \mu_r + 2EE_r, \mu_p + 2EE_p)$ and $(\mu_c + 2EE_c, \mu_r + 2EE_r, \mu_p + 2EE_p)$ [13]. En la Fig. 9 se muestra



Figura 10: Coil-20 object image library[1].

el espacio de clasificación de ortoedros de confianza para las imágenes de referencia de la base de datos en la Fig. 8. Dado que se asignó un espacio volumétrico a cada imagen y ninguno de ellos se traslapa, entonces el sistema de reconocimiento de patrones invariante a traslación, escala y rotación presenta un nivel de confianza de al menos 95.4%.

III.2 Imágenes en escala de grises

Para entrenar al sistema de reconocimiento de patrones con imágenes en escala de grises, cada imagen en la base de datos de imágenes de referencia en la Fig. 10 fue rotada 360° con un paso de $\Delta\theta = 10^\circ$. Entonces, por imagen se tienen 36 muestras rotadas, después cada una de ellas fue escalada $\pm 16\%$ con un paso de escala $\Delta h = 2\%$. Por lo tanto, se generaron $36 \times 17 = 612$ imágenes por cada imagen de referencia en la Fig. 10, obteniendo un total de $612 \times 20 = 12,240$ imágenes problema a clasificar. Una vez generadas las muestras, se calcularon las firmas de cada una de ellas. Luego, se obtuvieron las correspondientes potencias de las firmas, para finalmente construir los intervalos de confianza como fue descrito en la Sección III.1 para las imágenes en BW. En la Fig. 11 se muestra el espacio de clasificación de ortoedros de confianza [13] para las imágenes de referencia de la base de datos en la Fig. 10. Dado que se asignó un espacio volumétrico a cada imagen y ninguno de ellos se traslapa, entonces el sistema de reconocimiento de patrones invariante a traslación, escala y rotación presenta un

nivel de confianza de al menos 95.4% .

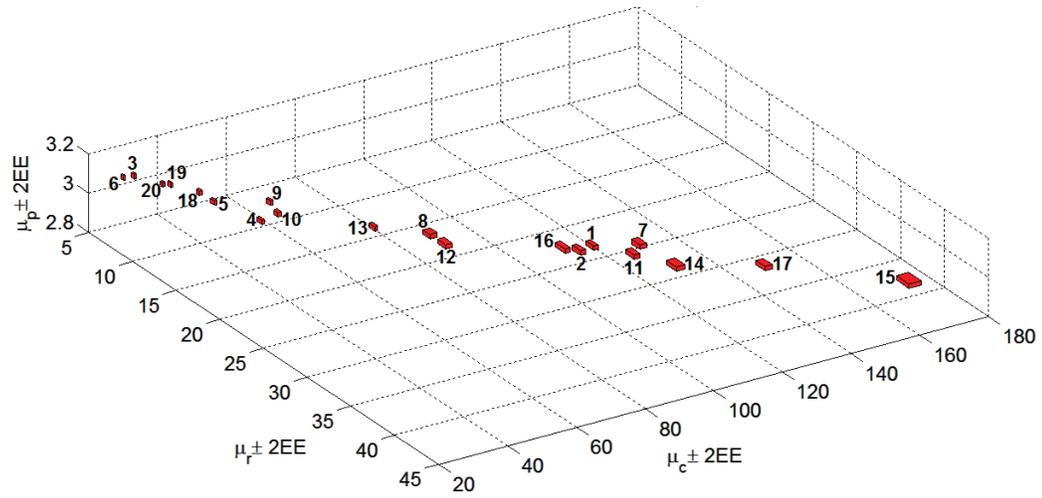


Figura 11: Espacio de clasificación de ortoedros de confianza para las imágenes en escalas de grises de la base de datos en Fig. 10.

Capítulo IV

Conclusiones

Basado en la teoría de la transformada de Fourier y la transformada de Fourier-Mellin, se desarrolló un sistema de reconocimiento de patrones que es invariante a traslación, escala y rotación. Dicho sistema fue diseñado para trabajar con imágenes en blanco y negro y en escala de grises.

Los sistemas de reconocimiento de patrones basados en la transformada analítica de Fourier-Mellin, por lo general, solamente poseen invariancia de escala y rotación, argumentando que la invariancia a traslación se puede lograr utilizando el centroide del objeto, en vez del centro real de la imagen. En este trabajo la invariancia a traslación se lleva a cabo de manera rápida y eficiente mediante el espectro de amplitud de las imágenes, dicho espectro es la imagen de entrada en la TAFM normalizada.

Para reducir el tiempo de cómputo considerablemente en la etapa de clasificación se generaron las firmas 1D invariantes a traslación, escala y rotación, a partir de la imagen obtenida del espectro de amplitud de la TAFM normalizada. Dichas firmas permiten reducir considerablemente el tiempo de cómputo en la etapa de clasificación. Además, en lugar de trabajar con múltiples planos de salida, como se hace en los sistemas de reconocimiento de patrones que utilizan funciones de correlación, aquí se genera solamente un espacio de clasificación de ortoedros, lo que permite reducir el número de operaciones de comparación y, por ende, el tiempo de cómputo.

De los espacios de clasificación de ortoedros obtenidos en las pruebas hechas al sistema de reconocimiento de patrones usando imágenes digitales en blanco y negro, y en escala de grises, se llega a la conclusión de que el sistema tiene un nivel de confianza de al menos el 95.4%.

Referencias

- [1] S.A. Nene, S. K. Nayar, and H. Murase, “Columbia Object Image Library (COIL-20),” *Technical Report CUCS-005-96*, 1996.
- [2] D. G. Lowe, “Object recognition from scale-invariant features,” in *The Proceedings of the Seventeenth IEEE International Conference on Computer Vision*, 1999, vol. 2, pp. 1150–1157.
- [3] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60(2), pp. 91–110, 2004.
- [4] H. Bay, T. Tuytelaars, and L. Vangool, “SURF: Speeded up robust features,” 2006, pp. 404–417, Springer.
- [5] J.R. Lerma Aragón and J. Álvarez Borrego, “Vectorial signatures for invariant recognition of position, rotation and scale pattern recognition,” *J. Mod. Opt.*, vol. 56(14), pp. 1598–1606, 2009.
- [6] J. Álvarez Borrego, S. Solorza, and M.A. Bueno-Ibarra, “Invariant correlation to position and rotation using binary mask applied to binary and gray images,” *Opt. Commun*, vol. 294, pp. 105–117, 2013.
- [7] S. Solorza and J. Álvarez Borrego, “Digital system of invariant correlation to position and rotation,” *Opt. Commun*, vol. 283(14), pp. 3613–3630, 2010.
- [8] S. Solorza-Calderón, “A position, rotation and scale invariant image descriptor based on rays and circular paths,” *Proceedings of SPIE, Applications of Digital Image Processing XXXVIII*, vol. 9599, pp. 95991V–1, 2015.
- [9] S. Derrode and F. Ghorbel, “Robust and efficient fourierâ“mellin transform approximations for gray-level image reconstruction and complete invariant description,” *Computer Vision and Image Understanding.*, vol. 83, pp. 57–78, 2001.
- [10] R.C. Gonzalez, R.E. Woods, and S.L. Eddins, *Digital Image Processing Using MATLAB*, Guestmark Publishing, 2009.
- [11] B. Srinivasa Reddy and B. N Chatterji, “An FFT-based technique for translation, rotation, and scale-invariant image registration,” *IEEE Transactions on Image Processing.*, vol. 5, pp. 1266–1271, 1996.
- [12] C. Barajas-García, S. Solorza-Calderón, and J. Álvarez-Borrego, “A digital image pattern recognition system invariant to rotation, scale and translation for color images,” *Opt. Commun.*, vol. 367, pp. 335–345, 2016.
- [13] J. Verdugo-Olachea, “Reconocimiento de patrones en imágenes digitales usando la transformada de Radon,” Tesis de Licenciatura, Universidad Autónoma de Baja California, 2015.