

Tesis defendida por
Alfredo Solís Ventura
y aprobada por el siguiente comité:

Dr. Josué Álvarez Borrego
Director del Comité

Dra. Rosa Reyna Mouriño Pérez
Miembro del Comité

Dra. Diana Tentori Santa Cruz
Miembro del Comité

Dr. Israel Rocha Mendoza
Miembro del Comité

Dr. Pedro Negrete Regagnon
Coordinador del Programa de Posgrado
en Ciencias en Óptica

Dr. David Hilario Covarrubias Rosales
Director de la Dirección de Estudios de
Posgrado

16 de agosto del 2012

**CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE
EDUCACIÓN SUPERIOR DE ENSENADA**



**Programa de Posgrado en Ciencias
en Óptica con orientación en Óptica Física**

**Correlación no lineal invariante a posición y rotación utilizando máscaras
adaptativas binarias**

Tesis

que para cubrir parcialmente los requisitos necesarios para obtener el grado de
Maestro en Ciencias

Presenta:

Alfredo Solís Ventura

Ensenada, Baja California, México,
2012.

Resumen de la tesis de **Alfredo Solís Ventura**, presentada como requisito parcial para la obtención del grado de Maestro en Ciencias en Óptica con orientación en Óptica Física.

Correlación no lineal invariante a posición y rotación utilizando máscaras adaptativas binarias.

Resumen aprobado por:

Dr. Josué Álvarez Borrego
Director de Tesis

En el reconocimiento de imágenes el desplazamiento o rotación de un objeto, dificulta o imposibilita su detección. En este trabajo se tiene como objetivo desarrollar un algoritmo digital que se ejecute con la mayor rapidez posible para el reconocimiento de imágenes, aun cuando estas presenten rotación, desplazamiento y contaminación por ruido, la transformada de Fourier de una imagen siempre estará centrada en el plano de frecuencias y es de esta manera que se logra obtener la invariancia a posición, para lograr la invariancia a rotación se proponen dos soluciones: la primera solución desarrollada, consta de una metodología iterativa, que cumple con invariancia a posición y rotación pero el tiempo de cómputo que utiliza es mucho y limita su aplicación práctica. Debido a esto se desarrolla una segunda metodología, que con la ayuda de una correlación no lineal adaptativa, logra la identificación de la imagen objetivo con un nivel de confianza del 100% y en un corto tiempo de cómputo utilizando una fracción de segundo por correlación. La metodología adaptativa se inicia tomando como referencia la parte real y positiva de la transformada de Fourier de una imagen objetivo, a partir de esta referencia se generan máscaras adaptativas binarias que son una serie de anillos concéntricos. Al realizar una multiplicación punto por punto de la máscara de anillos y el módulo de la transformada de Fourier de la imagen problema se obtienen una serie de anillos que contienen valores específicos del módulo de la transformada de Fourier. Estos anillos tienen la propiedad de ser invariantes a rotación. Al sumar los valores de cada anillo se obtiene una firma vectorial unidimensional capaz de representar a la imagen que se desea identificar, a pesar de estar desplazada, rotada e inmersa en ruido. Como método de comparación entre firmas vectoriales se utiliza una correlación no lineal, a la cual se le agrega la característica de ser adaptativa mediante un índice invariante a rotación, que se obtiene con una relación entre los promedios de la imagen problema y de la imagen objetivo, logrando así, la identificación automática de la imagen problema. Se demuestra la efectividad del método utilizando imágenes en blanco y negro, y en escala de grises; en donde cada una de ellas se gira grado por grado desde 0 hasta 359°. Fue posible identificar con éxito las 26 letras del abecedario en un formato blanco y negro y 19 especies de diatomeas fósiles en escala de grises, la identificación de diatomeas fósiles es de gran importancia para determinar las condiciones ambientales tanto del presente como del pasado y son útiles en estudios sobre la calidad del agua y el cambio climático. Se muestran resultados en donde se logra identificar con un nivel de confianza del 100% cada una de estas imágenes.

Palabras clave: **correlación no lineal, máscara binaria, firmas unidimensionales.**

Abstract of the thesis presented by **Alfredo Solís Ventura**, as a partial requirement to obtain the Master of Science degree in Optics with orientation in Physical Optics.

**Nonlinear correlation invariant to position and rotation using
binary adaptive masks**

Approved by:

Dr. Josué Álvarez Borrego
Thesis Advisor

Modifications due to displacement or rotation of an object to identify, adds difficulty or makes impossible the identification of the object. The objective is to develop a low computational cost digital algorithm to recognize images even when these images appear rotated, displaced and immersed in noise. The Fourier transform of an image, always will be centered on the frequency domain, and in this way, the invariant position is obtained. The first methodology developed consist of an iterative algorithm. This algorithm gets position and rotation invariance but it uses a lot of computational time. To avoid it; a second methodology was developed with the aid of an adaptive nonlinear correlation, achieving excellent results. The adaptive methodology begins taking as reference, the real and positive part of the target image Fourier transform. Starting in this reference, a binary adaptive mask was generated. This mask is a set of concentric rings. Calculating the point to point multiplication of the binary ring mask with the modulus of the Fourier transform of the problem image, we obtained a set of concentric rings, with the specific sampled values of the modulus of the problem image Fourier transform. These rings give us the property of rotation invariance. Making the sum of the values in each ring, we get a one-dimensional vectorial signature. This signature can represent a target image with invariance to rotation and position even in presence of noise. As comparison method between vectorial signatures, a nonlinear correlation was used. An adaptive characteristic was added to this nonlinear correlation with the aid of a rotational invariant index. This was obtained from a ratio between the mean values of the problem image and the target image, thus, the automatic identification of the problem image was achieved. The effectiveness of the method was demonstrated using black and white images and gray scale images. Each of these were rotated degree by degree from 0 to 359°. It was possible to identify with success the 26 letters of the alphabet in black and white format and 19 different fossil diatoms in gray scale format, the diatoms are a kind of microscopic algae and they are one of the most common types of phytoplankton. They live in fresh or marine water and are producers within the food chain. Hence, the diatom communities are used in the studies of water quality; also they are a useful tool for monitoring environmental conditions, past and present. Results with 100% of confidence level in the identification of these images are shown.

Key words: nonlinear correlation, binary mask, one-dimensional signatures.

Dedicatoria

*A mi esposa Yaneth, a mi hermosa hija Vanessa
y a mi bebé, el que mi esposa cuida y lleva en su vientre.*

*Porque además de llenar de amor y alegría todos los días de mi vida,
ustedes son la fuente de mi inspiración y el principal motivo de mi superación.*

Agradecimientos

Al CICESE, por que además de tener una excelente infraestructura; durante mis estudios de maestría me abrió sus puertas y me permitió utilizar sus instalaciones las 24 horas del día y todos los días del año, haciéndome sentir como en casa.

Al Dr. Josué Álvarez Borrego, por tenerme paciencia, por sus consejos y por ser ese guía indispensable, para concluir exitosamente este trabajo de tesis.

A los miembros del comité de tesis: Dra. Diana Tentori Santa Cruz, Dra. Rosa Reyna Mouriño Pérez y Dr. Israel Rocha Mendoza, por sus valiosos comentarios y sugerencias.

A mi esposa Yaneth y a mi hija Vanessa, quienes además de brindarme su amor incondicional y apoyo para llegar a concluir mis estudios de maestría; también sacrificaron apreciados momentos familiares.

A mis padres, que fueron los que me inculcaron estudiar para superarme.

Al CONACYT por su apoyo económico al otorgarme la beca para realizar mis estudios de maestría, así como también a su financiamiento del proyecto de ciencia básica: “Correlación invariante a posición, rotación, escala, ruido e iluminación para identificación de organismos y estructuras microscópicas y macroscópicas”, número 102007 a cargo del Dr. Josué Álvarez Borrego.

Contenido

	Página
Resumen	i
Abstract	ii
Dedicatoria	iii
Agradecimientos	iv
Contenido	v
Lista de figuras	vii
Capítulo I. Introducción	1
I.1 Antecedentes.....	1
I.1.1 Máscaras binarias	2
I.1.2 Filtros de correlación.....	6
I.2 La Transformada de Fourier	8
I.2.1 Transformada de Fourier	8
I.2.2 Transformada discreta de Fourier.....	9
I.2.3 Propiedades de la transformada de Fourier	10
I.3 Métricas de desempeño	14
I.3.1 Capacidad de discriminación.....	14
I.3.2 Energía pico de correlación (PCE)	14
I.4 Ruido en la imagen	15
I.4.1 Ruido gaussiano.....	15
I.4.2 Ruido impulsivo	16
I.5 Filtros de correlación.....	18
I.5.1 Filtro clásico acoplado.....	18
I.5.2 Filtro sólo de fase	19
I.5.3 Filtro no lineal	19
I.6 Objetivos.....	20
I.6.1 Objetivo general	20
I.6.2 Objetivos particulares	20
Capítulo II. Metodología	21
II.1 Configuración de: imágenes, programas y equipos utilizados.....	21
II.2 Metodología iterativa de correlación no lineal invariante a posición y rotación	22
II.2.1 Identificación del mejor transecto	24
II.2.2 Máscara adaptativa binaria de anillos concéntricos	26
II.2.3 La firma de la imagen	28
II.2.4. Correlación no lineal	30
II.3 Metodología adaptativa de correlación no lineal invariante a posición y rotación.....	30
II.3.1 Máscara de anillos concéntricos, utilizada para la metodología adaptativa.....	31
II.3.2 Índice invariante a rotación	33
II.3.3 La firma utilizada en la metodología adaptativa	33
II.3.4 Correlación no lineal adaptativa.....	34

Capítulo III. Resultados	37
III.1 Reconocimiento de imágenes utilizando la metodología iterativa	37
III.1.1 Aplicación de la metodología iterativa a imágenes binarias.....	37
III.1.2 Aplicación de la metodología iterativa a imágenes reales	40
III.2 Reconocimiento de imágenes utilizando la metodología adaptativa	42
III.2.1 Aplicación de la metodología adaptativa a imágenes binarias	42
III.2.2 Aplicación de la metodología adaptativa a imágenes reales.....	50
III.3 Valor óptimo del valor de no linealidad k	56
III.4 Reconocimiento de imágenes inmersas en ruido	57
Capítulo IV. Discusión	60
Conclusiones	62
Referencias bibliográficas	63

Lista de figuras

Figura	Página
1	Primeras máscaras binaras propuestas para generar firmas vectoriales a partir de una imagen real (Rosen <i>et al.</i> , 1963).....2
2	Ilustración de algunas máscaras propuestas por Lendaris y Stanley, 1965. 3
3	Máscaras binarias: a) máscara de anillos concéntricos, b) máscara de transectos de circunferencia separados 9° entre sí, c) máscara de líneas paralelas. 3
4	Ilustración de un detector RWD (<i>Ring-Wedge Detector</i>), (Asselin et al. 1994). 4
5	Ejemplos de máscaras adaptativas. (a) Letra T Arial. (b) Máscara adaptativa correspondiente a la letra T. (c) Letra X Arial. (d) Máscara adaptativa correspondiente a la letra X. 5
6	Distribución gaussiana graficada con diferentes parámetros de μ y σ^2 15
7	Imagen inmersa en ruido gaussiano con $\mu = 0$: (a) Sin ruido, (b) $\sigma^2=0.1$ y (c) $\sigma^2=0.5$ 16
8	Imagen inmersa en ruido sal y pimienta: (a) Sin ruido, (b) Densidad= 0.1 y (c) Densidad= 0.5..... 17
9	Metodología iterativa de correlación no lineal invariante a posición y rotación. 23
10	Diagrama de flujo para obtener el mejor transecto de una imagen de 256×256 pixeles 24
11	(a) Imagen formada por 256×256 pixeles con la letra D Arial centrada. (b) La parte real de la transformada de Fourier de la imagen. 26
12	(a) Ejemplo del promedio de los 180 mejores transectos. (b) Binarización del promedio de los 180 mejores transectos. 27
13	(a) Imagen objetivo. (b) Máscara de anillos concéntricos asociada a la letra D Arial. 27
14	Procedimiento para obtener la firma correspondiente a la imagen binaria de una letra B Arial centrada de 128×128 pixeles; donde el símbolo * especifica una multiplicación punto a punto..... 28
15	Procedimiento para obtener la firma correspondiente a la imagen binaria de una letra X Arial centrada de 128×128 pixeles; donde el símbolo * especifica una multiplicación punto a punto..... 29
16	Diagrama de flujo a seguir para construir la máscara binaria de anillos concéntricos utilizada en la metodología adaptativa. 31
17	Metodología adaptativa de correlación no lineal invariante a posición y rotación. 32
18	Valores de correlación para el filtro de la letra E obtenidos con la metodología adaptativa, (a) utilizando la correlación no lineal C_{NL1} , (b) utilizando la correlación no lineal adaptativa C_{NL2} 34

19	Imágenes binarias de las letras B, D, E y F tipo Arial de 256×256 pixeles	38
20	Correlaciones obtenidas utilizando el método iterativo para: (a) filtro letra B, (b) filtro letra D, (c) filtro letra E y (d) filtro letra F	39
21	Imágenes de diatomeas utilizadas para probar la metodología iterativa: (A) <i>Actinocyclus ingens-Rattray</i> , (B) <i>Thalassiosira oestruppii var 1</i> , (C) <i>Actinocyclus ellipticus-Grunow in van Heurck</i> , (D) <i>Nitzchia praereinholdii-Schrader</i> , (E) <i>Bogorovia praepaleacea (Schrader) Jouse</i>	40
22	Correlaciones obtenidas utilizando el método iterativo para: (a) filtro diatomea A, (b) filtro diatomea B, (c) filtro diatomea C, (d) filtro diatomea D y (e) filtro diatomea (E)	41
23	Gráficas de cajas del valor relativo de correlación de las letras del abecedario desde la (A) a la (Z), con respecto al resto, en rotación desde 0 hasta 359° con incrementos de 1°, obtenidas utilizando la metodología adaptativa, donde las cajas representan un error estándar ($\pm EE$) y las barras de las cajas representan dos errores estándar ($\pm 2*EE$) para el pico de correlación.	46
24	Acercamiento a la escala correspondiente al valor relativo de correlación (Metodología adaptativa), para: (a) Filtro Letra D, (b) Filtro Letra E, (c) Filtro Letra H, (d) Filtro Letra K, (e) Filtro Letra L, (f) Filtro Letra M, (g) Filtro Letra N, (h) Filtro Letra P, (i) Filtro Letra R, (j) Filtro Letra S, (k) Filtro Letra X y (l) Filtro Letra Z.....	49
25	Imágenes de diatomeas utilizadas para probar la metodología adaptativa.	50
26	Gráficas de cajas del valor relativo de correlación de las diatomeas desde la (A) a la (S), con respecto al resto, en rotación desde 0 hasta 359° obtenidas utilizando la metodología adaptativa, con incrementos de 1°, donde las cajas representan un error estándar ($\pm EE$) y las barras de las cajas representan dos errores estándar ($\pm 2*EE$) para el pico de correlación.	54
27	Acercamiento a la escala correspondiente al valor relativo de correlación (Metodología adaptativa) para: (a) Filtro diatomea G, (b) Filtro diatomea J, (c) Filtro diatomea N y (d) Filtro diatomea P.	55
28	Gráfica de cajas para el pico de correlación energía PCE, contra los valores de no linealidad k, utilizando la metodología adaptativa y la letra B Arial como filtro.....	56
29	Proceso para obtener un valor de coeficiente de discriminación (CD).....	57
30	Letra B Arial afectada por ruido sal y pimienta.....	57
31	Rendimiento del filtro en presencia de ruido sal y pimienta.	58
32	Letra B Arial afectada por gaussiano aditivo de media cero.	58
33	Rendimiento del filtro en presencia de ruido gaussiano de media cero.....	58
34	Coficiente de discriminación del filtro en presencia de ruido gaussiano y ruido impulsivo.	59
35	Coficiente de discriminación del filtro en presencia de ruido gaussiano ($\sigma^2=0$, $\sigma^2=0.6$ y $\sigma^2=1.2$) y ruido impulsivo.	59

Capítulo I. Introducción

I.1 Antecedentes

En la captura de una imagen intervienen diversos procesos físicos y electrónicos que deben ser comprendidos para la correcta interpretación de la imagen y los resultados que eventualmente se obtengan a partir de su tratamiento. La imagen digital capturada debe ser procesada por una computadora para extraer las magnitudes de interés. En estos procesos están involucradas muchas disciplinas; tales como la física del estado sólido, la óptica, la electrónica digital, la electrónica analógica, la programación, el diseño de algoritmos, el análisis matemático y el álgebra.

Para la ciencia, las ventajas del procesamiento digital de imágenes resultan muy atractivas, principalmente como herramienta de medición. La posibilidad de realizar observaciones a intervalos regulares con alta velocidad y de forma automática o la habilidad de manejar grandes volúmenes de datos, son sólo algunas de sus ventajas.

En estos últimos años, los avances de las tecnologías de adquisición de imágenes y de los dispositivos semiconductores en general, han hecho que estos sistemas estén al alcance de cualquier laboratorio. Hoy es posible realizar mediciones de gran calidad con sistemas de adquisición y computadoras de propósito general.

A continuación se presenta cómo han evolucionado las máscaras binarias y el desarrollo de filtros de correlación en los últimos años, ya que la metodología desarrollada para este trabajo se basa principalmente en proponer un nuevo filtro de correlación no lineal con firmas vectoriales generadas a partir de máscaras adaptativas binarias, para llevar a cabo el reconocimiento confiable de objetos distorsionados y degradados por cambios de rotación y ruido.

I.1.1 Máscaras binarias

El propósito de utilizar una máscara binaria es tomar muestras del patrón de difracción con el fin de generar una firma vectorial unidimensional que represente de manera efectiva a una imagen, facilitando y agilizando cualquier tipo de procesamiento posterior.

Con el propósito de reducir la cantidad de información a procesar al momento de realizar comparaciones entre diferentes patrones, investigadores del instituto de investigación de Standford, trabajando en un proyecto con fines militares, fueron los primeros en obtener la firma de una imagen utilizando máscaras binarias para adquirir muestras de una imagen real. Esta firma la obtuvieron a través de un sensor que consta de una matriz de 8×8 transductores, se colocan cinco máscaras frente a los sensores formando barras horizontales y otras cinco formando barras verticales como lo muestra la figura 1, de tal forma que cada una de estas máscaras obtendrá una señal eléctrica de diferente intensidad, generando así, una firma vectorial de diez índices diferentes para cada imagen, (Rosen *et al.*, 1963).

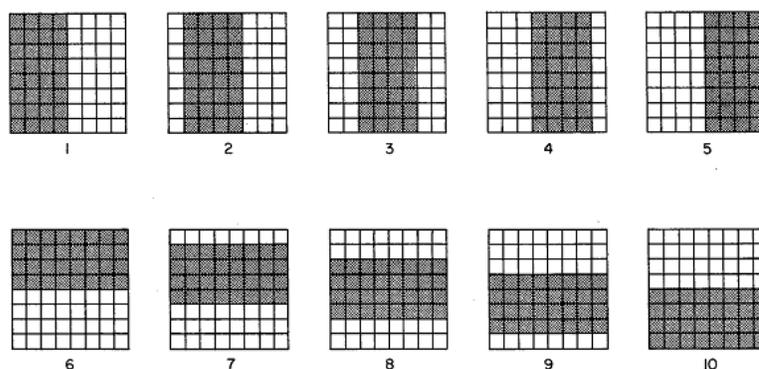


Figura 1. Primeras máscaras binarias propuestas para generar firmas vectoriales a partir de una imagen real (Rosen *et al.*, 1963).

Años más tarde en los laboratorios de investigación de *General Motors*, trabajando para el departamento de defensa de los Estados Unidos, describieron por primera vez el uso de máscaras binarias para obtener muestras de un patrón de difracción (Lendaris y Stanley, 1965) y aprovechar la invariancia a posición; ya que el patrón de difracción de una imagen problema desplazada en posición siempre estará centrada en el eje óptico. Como máscaras

binarias ellos propusieron utilizar líneas simples que podían ser horizontales, verticales o con una inclinación de 45° , también propusieron utilizar anillos concéntricos (Figura 2).



Figura 2. Ilustración de algunas máscaras propuestas por Lendaris y Stanley, 1965.

En 1970, con el propósito de identificar imágenes en fotografías de tomas aéreas, se utilizó una máscara de anillos concéntricos en combinación con una máscara que contiene transectos de circunferencia cada 9° (Lendaris y Stanley, 1970). Cada una de estas máscaras, ilustradas en la figura 3, toman muestras del patrón de difracción que son de gran utilidad. El muestreo con la máscara de anillos concéntricos logró muestras de un rango de frecuencias espaciales específico del patrón de difracción en dos dimensiones y la firma obtenida tiene la característica de ser invariante a rotación, con las muestras obtenidas a partir de los transectos de circunferencia separados 9° entre sí, se obtuvo una invariancia a escala y se logró obtener la aproximación de la posición del mejor transecto, lo cual da una referencia en grados de la rotación de la imagen problema.

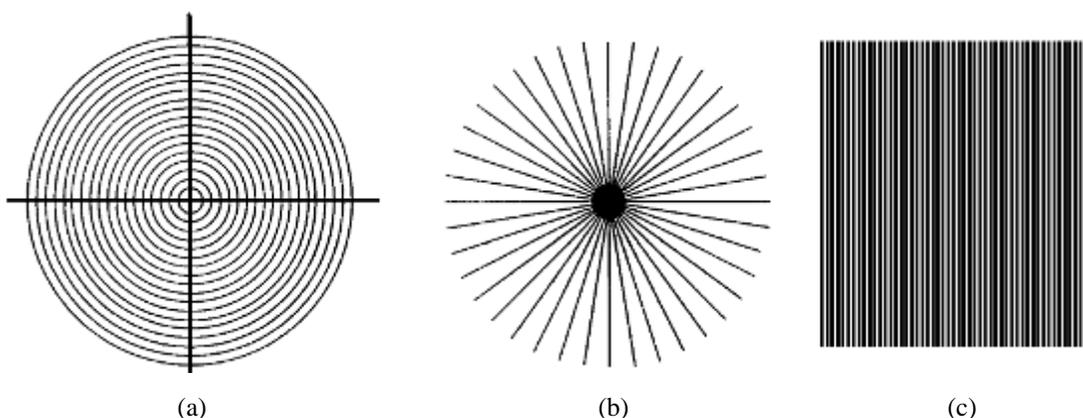


Figura 3. Máscaras binarias: a) máscara de anillos concéntricos, b) máscara de transectos de circunferencia separados 9° entre sí, c) máscara de líneas paralelas.

Otra máscara binaria, que tiene como propósito principal tomar muestras de un patrón de difracción que es invariante a rotación y escala, se ilustra en la figura 4 y esta se conoce

como detector RWD (*Ring-Wedge Detector*). Consiste en dos regiones de medio círculo, un semicírculo está compuesto por una serie de anillos concéntricos y el otro semicírculo se forma con transectos de circunferencia separados con ángulos equidistantes (George *et al.*, 1972).

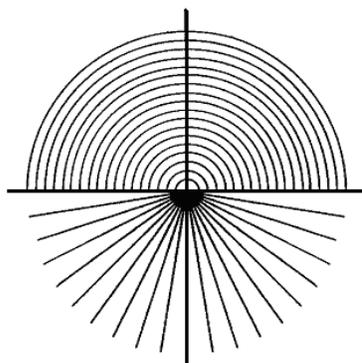


Figura 4. Ilustración de un detector RWD (*Ring-Wedge Detector*), (Asselin *et al.* 1994).

A principios de la década de los 80's se utilizaron máscaras binarias de 32 anillos concéntricos generadas por computadora, aplicadas a sistemas óptico-digitales. En ellas se aprovechó la velocidad de respuesta de un sistema óptico para generar un patrón de difracción y obtener así la transformada de Fourier con alta resolución, a la velocidad de la luz, independientemente del tamaño de la imagen problema. La imagen del patrón de difracción se capturó con una cámara de televisión para ser digitalizada y aplicarle la máscara binaria dejando sólo una tarea mínima de procesamiento digital (O'Toole y Stark, 1980).

Una máscara tipo RWD tiene la ventaja de ser invariante a posición y escala, pero cuando distorsiones generadas por escalamiento y rotación están presentes al mismo tiempo, interpretar los resultados de muestreo resulta complicado. Esto debido a que si el patrón de difracción sufre cambios de rotación, las muestras recopiladas por los anillos permanecen constantes, pero los datos de los transectos circulares separados formando ángulos equidistantes será muy variable. Ahora, si suponemos cambios de escala, las muestras obtenidas de los transectos circulares permanecerán constantes pero las muestras recopiladas por los anillos será variable. Por este motivo si sólo se requiere una invariancia a rotación, se recomienda utilizar máscaras de anillos concéntricos (Asselin *et al.* 1994).

La implementación de máscaras binarias de anillos circulares concéntricos ofrece invariancia a rotación, pero si los objetos a reconocer son similares entre sí, tendrán patrones de difracción similares y es muy probable que las firmas generadas también sean similares entre sí. Entonces, la discriminación entre objetos similares no es muy efectiva. Utilizar los valores de la fase de la transformada de Fourier a cambio de los valores de su intensidad, es una variante que mejora la discriminación entre objetos similares, la firma se obtiene, entonces, a partir de una multiplicación punto por punto de la máscara de anillos con los valores de la fase de la transformada de Fourier (Fares *et al.*, 2000). Al tomar como referencia los valores de la fase se pierde la invariancia a posición, por lo que es de interés el generar firmas que toman muestras de la intensidad del patrón de difracción y así conservar la invariancia a posición.

En investigaciones más recientes se han utilizado máscaras de anillos que se adaptan a las frecuencias más representativas del patrón de difracción de una imagen (Solorza Calderón y Álvarez Borrego, 2010), logrando así una invariancia a rotación con la propiedad de discriminación entre objetos similares. Se ha comprobado que las máscaras adaptativas pueden ser obtenidas a partir, tanto de la parte real, como de la parte imaginaria de la transformada de Fourier y que las máscaras de anillos pueden ser utilizadas de forma invertida (Chaparro Magallanez *et al.*, 2011).

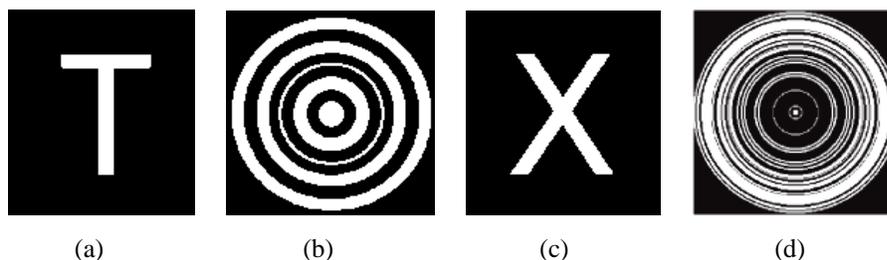


Figura 5. Ejemplos de máscaras adaptativas. (a) Letra T Arial. (b) Máscara adaptativa correspondiente a la letra T. (c) Letra X Arial. (d) Máscara adaptativa correspondiente a la letra X. (Solorza Calderón y Álvarez Borrego, 2010)

Con metodologías que utilizan máscaras adaptativas de anillos se ha logrado además de la invariancia a posición y rotación la invariancia a escalamiento. Esto se logra agregando el uso de filtros compuestos (Solorza Calderón y Álvarez Borrego, 2011).

I.1.2 Filtros de correlación

Desde la aparición del filtro acoplado clásico (Vander Lugt, 1964), empezó un interés para desarrollar filtros cada vez más eficientes. El algoritmo de la transformada rápida de Fourier (Cooley y Tukey, 1965) ha hecho posible el avance de métodos más eficientes de correlación de imágenes digitales, debido a su alta velocidad de procesamiento. La correlación puede ser un método eficaz para reconocer objetos, pero un simple cambio en la rotación puede hacer que el valor de correlación disminuya tanto que no se puedan reconocer a pesar de ser iguales. Una solución a este problema es el uso de filtros armónicos circulares que tienen invariancia a rotación por lo que el valor de correlación no se ve afectado por un cambio en la orientación del objeto; este filtro fue diseñado con fines militares y es generado a partir del centro propio o geométrico de una imagen, a un orden dado (Hsu *et al.*, 1982). El primer trabajo de filtros armónicos circulares con fines no militares fue presentado por nuestro grupo de investigación donde estos filtros invariantes a rotación se aplicaron en el reconocimiento de microorganismos planctónicos (Zavala Hamz y Álvarez Borrego, 1997). Basándose en la transformada de Fourier o en el patrón de difracción, el grupo de procesamiento de imágenes del departamento de óptica del CICESE ha desarrollado diferentes sistemas de correlación digital para el reconocimiento de patrones utilizando filtros simples y compuestos, lineales y no lineales invariantes a posición rotación y escalamiento. Estos filtros se han perfeccionado gracias al constante trabajo de diferentes colaboradores del grupo: Pech Pacheco *et al.* (2003) realizó la identificación de objetos basándose en la correlación de fase en el dominio de la transformada de escala para el reconocimiento automático de caracteres logrando invariancia a posición, rotación y escala. En otra investigación Lerma Aragón y Álvarez Borrego (2009) utilizaron dos firmas vectoriales por imagen: una firma de rotación y otra firma de escala, basadas en las propiedades de las transformadas de escala y de Fourier, como técnica de comparación utilizan la distancia euclidiana entre las firmas de la imagen objetivo y la imagen problema para lograr el reconocimiento de patrones invariante a posición, rotación y escala. Coronel Beltrán y Álvarez Borrego (2010) desarrollaron sistemas digitales de correlación invariante a posición, rotación y escala usando filtros no

lineales, en donde la correlación no lineal se realiza mediante las operaciones de la transformada de Fourier, el mapeo de coordenadas cartesianas a polares y la transformada de escala. Estos filtros fueron aplicados a las letras del alfabeto completo tipo Arial. Concluyeron que los resultados de los filtros no lineales muestran un mejor desempeño cuando son comparados con filtros sólo de fase.

Los filtros no lineales al realizar las operaciones de correlación digital dan información de la similitud de las imágenes entre los diferentes objetos. Esta clase de filtros tienen ventajas comparados con los filtros clásicos, debido a que tienen mayor capacidad para discriminar objetos, ya que el máximo valor del pico de correlación se encuentra bien localizado y con menor ruido.

Por lo anterior, en trabajos recientes se ha utilizado la correlación digital no lineal (ley-k) entre firmas unidimensionales creadas a partir de máscaras binarias adaptativas de círculos concéntricos, incrementando la capacidad de discriminación y conservando la invariancia a posición y rotación. Estos novedosos sistemas han logrado la identificación exitosa de la imagen problema aun cuando esta imagen se encuentra inmersa en ruido (Solorza Calderón y Álvarez Borrego, 2010; Chaparro Magallanez *et al.*, 2011; Solorza Calderón *et al.*, 2012).

En este trabajo se diseñaron dos metodologías que utilizan como herramientas: la transformada rápida de Fourier para obtener el patrón de difracción de las imágenes, las máscaras binarias de anillos concéntricos, las firmas vectoriales unidimensionales y la correlación no lineal como método de comparación. Como aportación se proponen: la metodología iterativa, que se basa en identificar el mejor transecto de la parte real del patrón de difracción de la imagen y en generar distorsiones por rotación grado por grado en cada imagen y la metodología adaptativa, que utiliza una correlación no lineal adaptativa en combinación con el valor promedio de la imagen como índice invariante a rotación. Estas dos metodologías logran mejorar la invariancia a rotación presentada por Solorza Calderón y Álvarez Borrego (2010).

I.2 La Transformada de Fourier

“El desarrollo matemático de la transformada de Fourier fue explicado por Jean Baptiste Joseph Fourier, en su libro la Teoría Analítica del Calor, publicado en 1822” (Cañada, 2006). Posteriormente, en 1965 Cooley y Tukey publicaron su artículo titulado: Un algoritmo para calcular las Series de Fourier Complejas, conocido como algoritmo FFT (*Fast Fourier Transform*), el cual reduce de forma considerable el número de multiplicaciones requeridas para el cálculo. La Transformada de Fourier convierte una imagen desde el dominio espacial al dominio de la frecuencia. Cada punto de la imagen en el dominio de Fourier representa una frecuencia particular contenida en la imagen en el dominio del espacio. La transformada de Fourier se utiliza en un amplio rango de aplicaciones, tales como filtrado, reconstrucción y compresión de imágenes. Esta transformación nos permite una manipulación más eficiente de la imagen debido a que muchas operaciones son más sencillas. Particularmente aquellas relacionadas con el filtrado.

I.2.1 Transformada de Fourier

“En el caso de una función de dos dimensiones $f(x, y)$, como es el caso de una imagen, el par de transformación de Fourier es el siguiente:

$$F(u, v) = \iint_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux + vy)} dx dy, \quad (1)$$

$$f(x, y) = \iint_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux + vy)} du dv, \quad (2)$$

donde (x, y) son las coordenadas de un pixel en la imagen, j es el número imaginario $\sqrt{-1}$ y (u, v) son las frecuencias espaciales de la imagen” (González y Woods, 2002, p. 150).

I.2.2 Transformada discreta de Fourier

El mundo real es de naturaleza continua en tiempo y espacio. Sin embargo, en diversas ocasiones es necesario trabajar con señales discretas. Esto se debe a que un gran número de sensores operan con este tipo de señales. Un ejemplo de esto se da en el procesamiento digital de imágenes, donde las imágenes de entrada son discretas tanto en espacio como en intensidad. “Para esto se supone que $f(x, y)$ es una función discreta, tomando N muestras separadas Δx unidades y M muestras separadas Δy unidades. Entonces el par de transformación discreta de Fourier es:

$$F(u, v) = \left(\frac{1}{MN} \right) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)}, \quad (3)$$

para valores de $u=0, 1, 2, \dots, M-1$ y $v=0, 1, 2, \dots, N-1$;

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M + vy/N)}, \quad (4)$$

para valores de $x=0, 1, 2, \dots, M-1$ y $y=0, 1, 2, \dots, N-1$.” (González y Woods, 2002, p. 154)

Los componentes de la transformada de Fourier son valores complejos, por consiguiente es común al trabajar con números complejos expresarlos en coordenadas polares, por lo tanto la magnitud y fase son:

$$F(u, v) = |F(u, v)| e^{-j\phi(u, v)}, \quad (5)$$

$$|F(u, v)| = \sqrt{\text{Re}(F(u, v))^2 + \text{Im}(F(u, v))^2}, \quad (6)$$

$$\phi(u, v) = \text{ArcTan} \left[\left| \frac{\text{Im}(F(u, v))}{\text{Re}(F(u, v))} \right| \right]. \quad (7)$$

I.2.3 Propiedades de la transformada de Fourier

Las propiedades más importantes de la transformada de Fourier en el procesamiento de imágenes digitales son las siguientes (González y Woods, 2002, p. 194):

1) Linealidad. La transformada de la suma de dos funciones es la suma de sus transformadas individuales.

$$F(f_1(x, y) + f_2(x, y)) = F(f_1(x, y)) + F(f_2(x, y)). \quad (8)$$

La transformada de Fourier es distributiva respecto a la suma pero no al producto,

$$F(f_1(x, y)f_2(x, y)) \neq F(f_1(x, y))F(f_2(x, y)). \quad (9)$$

2) Separabilidad. Se puede calcular la transformada de Fourier bidimensional aplicando sucesivamente la transformada unidimensional. Por lo tanto, el par de transformadas discretas de Fourier puede expresarse en la forma separable como,

$$F(u, v) = \left(\frac{1}{M} \right) \sum_{x=0}^{M-1} f(x) e^{-j2\pi(ux/M)} \left(\frac{1}{N} \right) \sum_{y=0}^{N-1} f(y) e^{-j2\pi(vy/N)}, \quad (10)$$

$$f(x, y) = \sum_{u=0}^{M-1} F(u) e^{j2\pi(ux/M)} \sum_{v=0}^{N-1} F(v) e^{j2\pi(vy/N)}. \quad (11)$$

3) Traslación. La multiplicación por un exponencial en el dominio espacial corresponde a traslación en frecuencia,

$$f(x, y) e^{j2\pi(u_0x + v_0y)/N} \Leftrightarrow F(u - u_0, v - v_0). \quad (12)$$

Una traslación de la función en el dominio espacial introduce un corrimiento en fase en el dominio de la frecuencia,

$$f(x - x_0, y - y_0) \Leftrightarrow F(u, v) e^{-j2\pi(x_0u + y_0v)/N}. \quad (13)$$

4) Periodicidad. La transformada de Fourier y su inversa son funciones periódicas de periodo N , es decir,

$$F(u, v) = F(u + N, v) = F(u, v + N) = F(u + N, v + N). \quad (14)$$

5) Simetría conjugada. Si $f(x, y)$ es una función real, la transformada de Fourier presenta también simetría conjugada,

$$F(u, v) = F^*(-u, -v). \quad (15)$$

6) Rotación. Si introducimos coordenadas polares en la transformada de Fourier, $x = r \cos \theta$; $y = r \sin \theta$; $u = \omega \cos \phi$; $v = \omega \sin \phi$, entonces $f(x, y)$ y $F(u, v)$ se convierten en $f(r, \theta)$ y $F(\omega, \phi)$. Además se cumple,

$$f(r, \theta + \theta_0) = F(\omega, \phi + \theta_0), \quad (16)$$

al girar $f(x, y)$ a un ángulo θ_0 , produce un giro con el mismo ángulo en $F(u, v)$.

7) Escalamiento. Para dos escalares a y b se tiene,

$$af(x, y) \Leftrightarrow aF(u, v), \quad (17)$$

$$f(ax, by) \Leftrightarrow \frac{1}{|ab|} F\left(\frac{u}{a}, \frac{v}{b}\right), \quad (18)$$

una expansión en el dominio espacial (x, y) resulta en una contracción en el dominio de la frecuencia (u, v) y viceversa.

8) Teorema de Parseval. Este es: “la integral del cuadrado del valor absoluto de una función es igual a la integral del cuadrado del valor absoluto de su transformada de Fourier. Este teorema se interpreta como un enunciado de la conservación de la energía, donde el área bajo la curva de esta integral es la variancia total de la función” (Hsu, 1984, p. 115):

$$\iint_{-\infty}^{\infty} |f(x, y)|^2 dx dy = \iint_{-\infty}^{\infty} |F(u, v)|^2 du dv. \quad (19)$$

9) Teorema de convolución. “La convolución de dos funciones $f(x, y)$ y $g(x, y)$, es otra función $h(x, y)$ definida por la doble integral,

$$h(x, y) = f(x, y) * g(x, y) = \iint_{-\infty}^{\infty} f(x', y') g(x - x', y - y') dx' dy', \quad (20)$$

donde $*$ denota convolución.

Si $f(x, y)$ tiene una transformada de Fourier $F(u, v)$ y $g(x, y)$ tiene una transformada de Fourier $G(u, v)$, entonces $f(x, y) * g(x, y)$ tiene la transformada de Fourier $F(u, v)G(u, v)$; es decir, la convolución de dos funciones en el dominio espacial, significa la multiplicación de la transformada de esas dos funciones” (Hsu, 1984, p. 117).

Usando $\mathfrak{T}\{ \}$ para denotar una transformada de Fourier, podemos expresar el teorema de convolución con las siguientes ecuaciones:

$$\mathfrak{T}\{f(x, y) * g(x, y)\} = F(u, v)G(u, v), \quad (21)$$

$$\mathfrak{T}\{f(x, y)g(x, y)\} = F(u, v) * G(u, v), \quad (22)$$

por lo tanto es posible calcular la convolución de dos señales haciendo uso de la transformada de Fourier:

$$f(x, y) * g(x, y) = \mathfrak{T}^{-1}\{F(u, v)G(u, v)\}. \quad (23)$$

La convolución para señales continuas está descrita por la Ec. (23); sin embargo, “puede ser fácilmente expresada en términos de funciones discretas de la siguiente manera:

$$h(x, y) = f(x, y) * g(x, y) = \sum_{x'=0}^{M-1} \sum_{y'=0}^{N-1} f(x', y') g(x - x', y - y'), \quad (24)$$

donde M es el número total de valores en el eje x , N es número total de valores en el eje y donde $*$ denota una operación de convolución.”, (González y Woods, 2002, p. 205).

10) Teorema de correlación. “La correlación de dos funciones reales $f(x, y)$ y $g(x, y)$ es otra función $h(x, y)$ y está definida por la integral,

$$h(x, y) = f(x, y) \otimes g(x, y) = \iint_{-\infty}^{\infty} f(x', y') g(x + x', y + y') dx' dy', \quad (25)$$

donde el símbolo \otimes denota una correlación.

En el caso de que las funciones sean complejas, se debe utilizar el complejo conjugado de una de las dos funciones:

$$h(x, y) = f(x, y) \otimes g(x, y) = \iint_{-\infty}^{\infty} f(x', y') g^*(x + x', y + y') dx' dy', \quad (26)$$

donde el signo * indica que es el complejo conjugado” (Hsu, 1984, p. 120).

Usando $\mathfrak{F}\{ \}$ para denotar una transformada de Fourier, podemos resumir lo establecido por el teorema de correlación a las siguientes ecuaciones:

$$\mathfrak{F}\{f(x, y) \otimes g(x, y)\} = F(u, v) G^*(u, v), \quad (27)$$

$$\mathfrak{F}\{f(x, y) g^*(x, y)\} = F(u, v) \otimes G(u, v), \quad (28)$$

por lo tanto es posible calcular la correlación de dos señales haciendo uso de la transformada de Fourier de la siguiente manera:

$$f(x, y) \otimes g(x, y) = \mathfrak{F}^{-1}\{F(u, v) G^*(u, v)\}, \quad (29)$$

donde * indica el complejo conjugado.

Así también, la correlación para señales continuas está descrita por la Ec. (25), sin embargo puede ser fácilmente expresada en términos de funciones discretas de la siguiente manera:

$$h(x, y) = f(x, y) \otimes g(x, y) = \sum_{x'=0}^{M-1} \sum_{y'=0}^{N-1} f(x', y') g(x + x', y + y'), \quad (30)$$

donde M es el número total de valores en el eje x, N es número total de valores en el eje y \otimes denota una operación de correlación (González y Woods, 2002, p. 206).

I.3 Métricas de desempeño

Una métrica de desempeño para filtros de correlación cuantifica qué tan bien puede definirse ese pico de correlación. En la actualidad existe un gran número de métricas que permiten evaluar el desempeño de los filtros de correlación (Kumar y Hassebrook, 1990). Entre las métricas más utilizadas podemos mencionar la relación señal a ruido, la capacidad de discriminación, la energía pico de correlación y la eficiencia de luz, entre otras. En este trabajo se utilizan: la capacidad de discriminación y la energía pico de correlación.

I.3.1 Capacidad de discriminación

El coeficiente de discriminación (CD) evalúa la habilidad de un filtro para detectar un objeto deseado y rechazar cualquier otro objeto; por ejemplo, el fondo. Este coeficiente está dado por (Yaroslavsky, 1993),

$$CD = 1 - \frac{|C_B(0,0)|^2}{|C_T(0,0)|^2}, \quad (31)$$

donde $C_B(0,0)$ es el valor máximo en el plano de correlación sobre el área del objeto a rechazar y $C_T(0,0)$ es el valor máximo en el plano de correlación sobre el área de la posición del objeto de referencia. El valor máximo de esta medida es la unidad, mientras que valores negativos indican que el objeto no puede ser reconocido por el filtro.

I.3.2 Energía pico de correlación (PCE)

La energía PCE (*Peak to correlation energy*) por sus siglas en inglés, es una relación de la energía en el origen, que compara la energía total del plano de correlación con respecto al valor del pico de correlación,

$$PCE = \frac{|E\{c(0,0)\}|^2}{E\{|c(x,y)|^2\}}, \quad (32)$$

donde el numerador es el valor esperado de la intensidad del pico de correlación y el denominador es el valor esperado de la energía promedio en el plano de correlación (Javidi, 2002, p. 299).

I.4 Ruido en la imagen

Si la imagen de entrada se obtiene mediante la captura de una escena del mundo real, es de esperar que esta imagen esté distorsionada por la acción de al menos una fuente de ruido. El ruido es una mezcla de señales aleatorias no deseadas, que están presentes en la imagen pero que son ajenas a la imagen original. En este trabajo, se agrega ruido a las imágenes intencionalmente, con el fin de evaluar el desempeño de los filtros, para los casos en los que la imagen problema está inmersa en ruido gaussiano, ruido impulsivo y la combinación de ambos.

I.4.1 Ruido gaussiano

Suele ser provocado por los componentes electrónicos tales como sensores y digitalizadores. En el ruido gaussiano, la intensidad de cada pixel de la imagen se ve alterada en cierta medida, con respecto a la intensidad del pixel en la imagen original y tiene la siguiente función de densidad de probabilidad (González y Woods, 2002, p. 223):

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (33)$$

donde x es la variable aleatoria y representa el nivel de gris, μ es la media del valor promedio de x y σ^2 es la variancia de x .

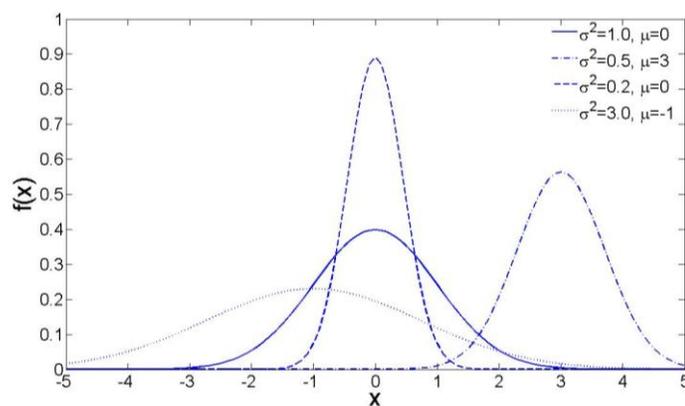


Figura 6. Distribución gaussiana graficada con diferentes parámetros de μ y σ^2 .

Cuanto mayor es el valor de la variancia, más se dispersan los datos en torno a la media. Un valor pequeño de este parámetro indica una gran probabilidad de obtener datos cercanos al

valor medio de la distribución. La media indica la posición de la campana, de modo que para diferentes valores de μ , la gráfica se desplaza a lo largo del eje x (Figura 6). La figura 7 muestra el efecto del ruido gaussiano en una imagen de 225×180 píxeles en escala de grises.

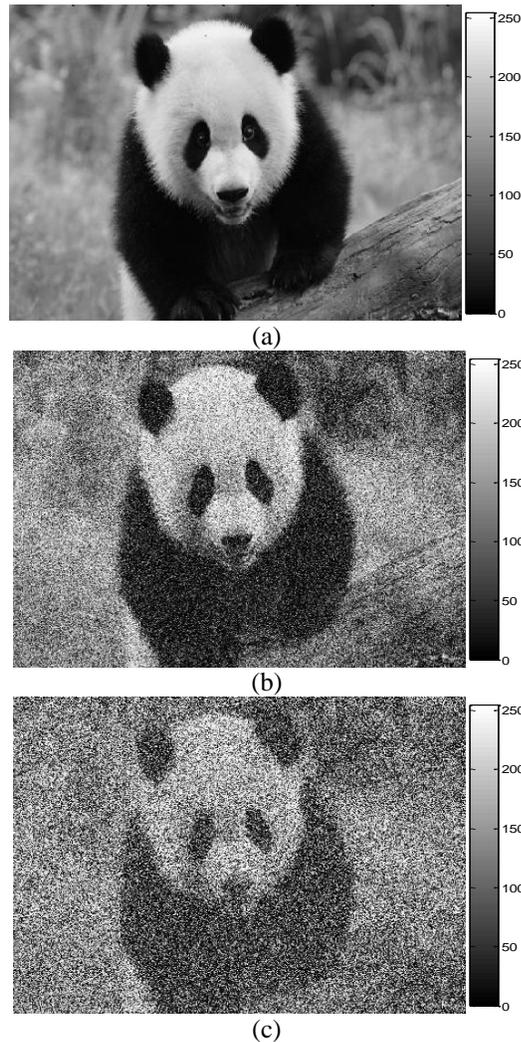


Figura 7. Imagen inmersa en ruido gaussiano con $\mu = 0$: (a) Sin ruido, (b) $\sigma^2=0.1$ y (c) $\sigma^2=0.5$.
(Imagen original: Mak *et al.*, 2010)

I.4.2 Ruido impulsivo

En el ruido impulsivo (ruido sal y pimienta) el valor que toma el píxel no tiene relación con el valor de la imagen original, sino que toma valores muy altos o muy bajos. “Toma el valor máximo (sal) o el mínimo (pimienta), para una imagen de 256 niveles de gris el valor

típico de la pimienta es 0 y de la sal 255. Defectos que contribuyen a este tipo de ruido son las manchas de polvo dentro de las componentes ópticas de la cámara, o bien un sensor CCD defectuoso, que realizará una captura errónea” (González y Woods, 2002, p. 226). La medida del ruido está determinada por un índice de densidad “d” que nos proporciona la información de pixeles afectados de la imagen que va de 0 a 1, cuando la densidad es 1 la totalidad de los pixeles están dañados. Una densidad de ruido impulsivo de 0.5, equivale al caso en donde la mitad de los pixeles están siendo afectados. En la figura 8 se muestra el efecto del ruido impulsivo en una imagen de 225×180 pixeles en escala de grises.

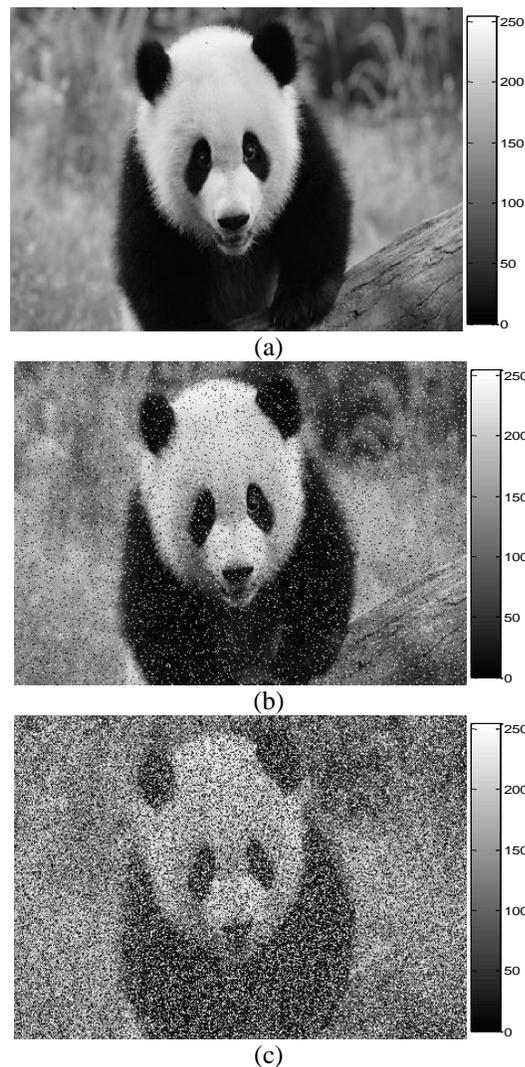


Figura 8. Imagen inmersa en ruido sal y pimienta: (a) Sin ruido, (b) Densidad= 0.1 y (c) Densidad= 0.5.
(Imagen original: Mak *et al.*, 2010)

I.5 Filtros de correlación

I.5.1 Filtro clásico acoplado

Este filtro también es llamado CMF (*Classical Matched Filter*) por sus siglas en inglés, fue propuesto por Vander Lugt (1964) y tiene la ventaja de que optimiza la relación señal a ruido a su salida, cuando la imagen problema se encuentra degradada por ruido blanco aditivo. Algunas limitaciones de este filtro clásico son producir: picos de correlación anchos, lóbulos laterales grandes, baja discriminación relativa y poca eficiencia de luz en su implementación óptica. La respuesta en frecuencia del filtro de acoplamiento clásico está dada por:

$$H(u, v) = \beta \frac{T^*(u, v)}{P_n(u, v)} \equiv H_{\text{CMF}}(u, v), \quad (34)$$

donde $T^*(u, v)$ y β son el complejo conjugado de la transformada de Fourier del objeto a reconocer y una constante arbitraria, respectivamente. $P_n(u, v)$ representa la densidad espectral de potencia del ruido.

Para el caso en que $P_n(u, v)$ y β son igual a uno, tenemos que $H_{\text{CMF}}(u, v)$ se puede expresar simplemente como,

$$H_{\text{CMF}}(u, v) = T^*(u, v), \quad (35)$$

como $H_{\text{CMF}}(u, v)$ es una función compleja, se puede expresar en su forma polar:

$$H_{\text{CMF}}(u, v) = |H(u, v)| e^{j\phi(u, v)}, \quad (36)$$

donde $|H(u, v)|$, es el módulo de la transformada de Fourier del objeto y $e^{j\phi(u, v)}$ es la respuesta de fase. Como los filtros espaciales son elementos pasivos; es decir, no amplifican la distribución de luz, $|H(u, v)| \leq 1$. En consecuencia, los valores posibles del filtro complejo deben estar en o dentro de una circunferencia unitaria centrada en el origen del plano complejo.

I.5.2 Filtro sólo de fase

Este filtro no considera pérdidas en la eficacia de la luz a través del sistema, es decir la respuesta en amplitud es unitaria. Entonces la función de un filtro sólo de fase está dada por (Horner y Gianino, 1984):

$$H_{\text{POF}}(u, v) = e^{j\phi(u, v)}, \quad (37)$$

es decir, sólo toma valores en el círculo unitario del plano complejo. Con el POF se obtienen picos de correlación más estrechos y es más selectivo que el CMF.

I.5.3 Filtro no lineal

Un filtro no lineal es un filtro cuya salida es una función no lineal de la entrada. La expresión para el filtro no lineal (ley k) está dada por (Vijaya Kumar y Hassebrook, 1990),

$$H_{\text{NLF}}(u, v) = |T(u, v)|^k e^{j\phi(u, v)}, \quad (38)$$

donde, $|T(u, v)|$, es el módulo de la transformada de Fourier del objeto que deseamos reconocer, k es el factor de no linealidad que toma valores $0 < k < 1$. Al variar los valores de k en este intervalo, podemos manipular las capacidades discriminantes del filtro y determinar el mejor valor k. Cuando el operador no lineal k modifica la transformada de Fourier tanto de la imagen problema como la de la imagen objetivo, consideramos que tenemos un procesador no lineal. El filtro no lineal tiene ventajas comparado con los filtros clásicos acoplados, los filtros sólo de fase y otros filtros lineales; debido a su gran capacidad para discriminar objetos, su máximo pico de correlación está bien localizado y el plano de salida es menos ruidoso.

I.6 Objetivos

I.6.1 Objetivo general

Identificación de objetos mediante una firma unidimensional realizada en una máscara adaptativa binaria, independientemente de que el objeto esté rotado o en diferente posición en la imagen.

I.6.2 Objetivos particulares

- 1) Obtener una firma unidimensional a partir de la transformada de Fourier de la imagen de referencia.
- 2) Encontrar la firma unidimensional de la imagen problema cuando ésta esté rotada a cualquier ángulo.
 - i. Encontrar el rango de ángulos en los cuales puede ser obtenida la firma unidimensional con mínima distorsión.
 - ii. Desarrollar una metodología que permita obtener la firma unidimensional aun cuando el objeto esté en la región de alta distorsión.
- 3) Utilizar una correlación no lineal (ley k) para correlacionar la firma de la imagen de referencia con la firma de la imagen problema.
- 4) Probar la efectividad del filtro en presencia de diferentes tipos de ruido (gaussiano, impulsivo y ambos).
- 5) Probar el filtro utilizando imágenes reales en escala de grises.

Capítulo II. Metodología

La necesidad de procesar imágenes en tiempo real agrega dificultad al reconocimiento de imágenes, ya que los objetos a reconocer pueden presentarse desplazados y además tener distorsiones debidas a rotación y diferentes tipos de ruido. Por lo tanto, es necesario el desarrollo de técnicas específicas para reconocer un mismo objeto de manera invariante a una alteración concreta o a un conjunto de ellas. Una modificación por desplazamiento o rotación de un objeto que se desea identificar, dificulta o imposibilita su detección. Por esta razón, se tiene como objetivo desarrollar un algoritmo digital para reconocer imágenes, aun cuando estas imágenes se presenten giradas, desplazadas e inmersas en ruido.

En este capítulo se muestra la metodología desarrollada para lograr una correlación no lineal invariante a posición y rotación utilizando máscaras adaptativas binarias. Estas máscaras están construidas por anillos concéntricos que se obtienen a partir de la parte real y positiva de la transformada de Fourier de la imagen problema. Al realizar una multiplicación punto por punto de la máscara de anillos y el módulo de la transformada de Fourier de la imagen problema, se estará obteniendo una serie de anillos que ahora contienen valores específicos del módulo de la transformada de Fourier. Estos anillos tienen la propiedad de ser invariantes a rotación. Al realizar la sumatoria de los valores de intensidad de cada anillo se obtiene una firma vectorial unidimensional capaz de representar a la imagen que se desea identificar, a pesar de estar desplazada, rotada e inmersa en ruido. Como método de comparación entre firmas vectoriales, se utiliza una correlación no lineal (ley k), que logra la identificación automática de la imagen objetivo.

II.1 Configuración de: imágenes, programas y equipos utilizados

Las imágenes binarias utilizadas en los experimentos son letras tipo Arial de la A a la Z con un tamaño de 256×256 píxeles y las imágenes reales son 19 especies de diatomeas fósiles en escala de grises (Figura 25) con un tamaño de 320×320 píxeles. “La identificación de estos organismos es de gran importancia para determinar las condiciones ambientales tanto

del presente como del pasado y son útiles en estudios sobre la calidad del agua y el cambio climático”, (Solorza Calderón y Álvarez Borrego, 2010). Los resultados son representados en diagramas de caja, donde el promedio de estos valores está representado por el cuadrado chico (\square). A un error estándar tenemos un valor de confianza de 68.3% y a dos errores estándar el nivel de confianza es del 95.4%. De esta forma, si las correlaciones no se traslapan, entonces se tiene un nivel de confianza del 100%. Las gráficas de caja fueron realizadas en el programa STATISTICA 6, la programación fue hecha en MATLAB versión R2011b, en una computadora con procesador AMD Athlon64, con una velocidad de reloj de 2.20 GHz, 533 MHz de velocidad de bus y 1GB DDR de memoria RAM.

II.2 Metodología iterativa de correlación no lineal invariante a posición y rotación

En la figura 9 se muestra un diagrama de flujo seccionado en pasos identificados con las letras de la (a) a la (l), este diagrama corresponde a la metodología iterativa que es utilizada para obtener una correlación no lineal invariante a posición y rotación. Se tienen como entrada una imagen problema (IP) y una imagen objetivo (IO) (pasos a1 y a2). Después ambas imágenes se rotan θ grados (pasos b1 y b2) y se obtiene la parte real de la transformada de Fourier de estas dos imágenes (pasos c1 y c2). La siguiente acción es encontrar el mejor transecto de la parte real de la transformada de Fourier y guardarlo en una variable T_p para la imagen problema y en T_o para la imagen objetivo (pasos d1 y d2). Hasta este momento se ha localizado el primer mejor transecto de cada imagen a cero grados. Enseguida se incrementa el valor de θ adicionándole 1° (pasos e1 y e2), de modo que ahora se inician los pasos (b1) y (b2) de la figura 9 con el valor de $\theta = 1^\circ$, provocando una rotación en las imágenes IP e IO, desde 0 hasta 179° grados (pasos f1 y f2), haciendo 180 iteraciones tanto para IP (paso, f1), como para IO (paso f2). En cada una de estas iteraciones se obtiene la parte real de la transformada de Fourier y se selecciona el mejor transecto. Promediando los 180 mejores transectos, se obtiene el mejor transecto promedio (pasos h1 y h2), se binariza este transecto promedio y a partir de este se obtiene una máscara binaria de anillos concéntricos (pasos i1 e i2) que será multiplicada punto a punto por el módulo de la transformada de Fourier (pasos g1 y g2). Se obtienen así una serie de

anillos concéntricos; cada uno de los cuales contiene los valores del módulo de la transformada de Fourier. Finalmente la firma vectorial de las imágenes se obtiene al sumar los valores del módulo de la transformada de Fourier de cada anillo (pasos j1 y j2). Se realiza una correlación no lineal entre estas dos firmas (paso k) y se selecciona el máximo valor de correlación (paso l).

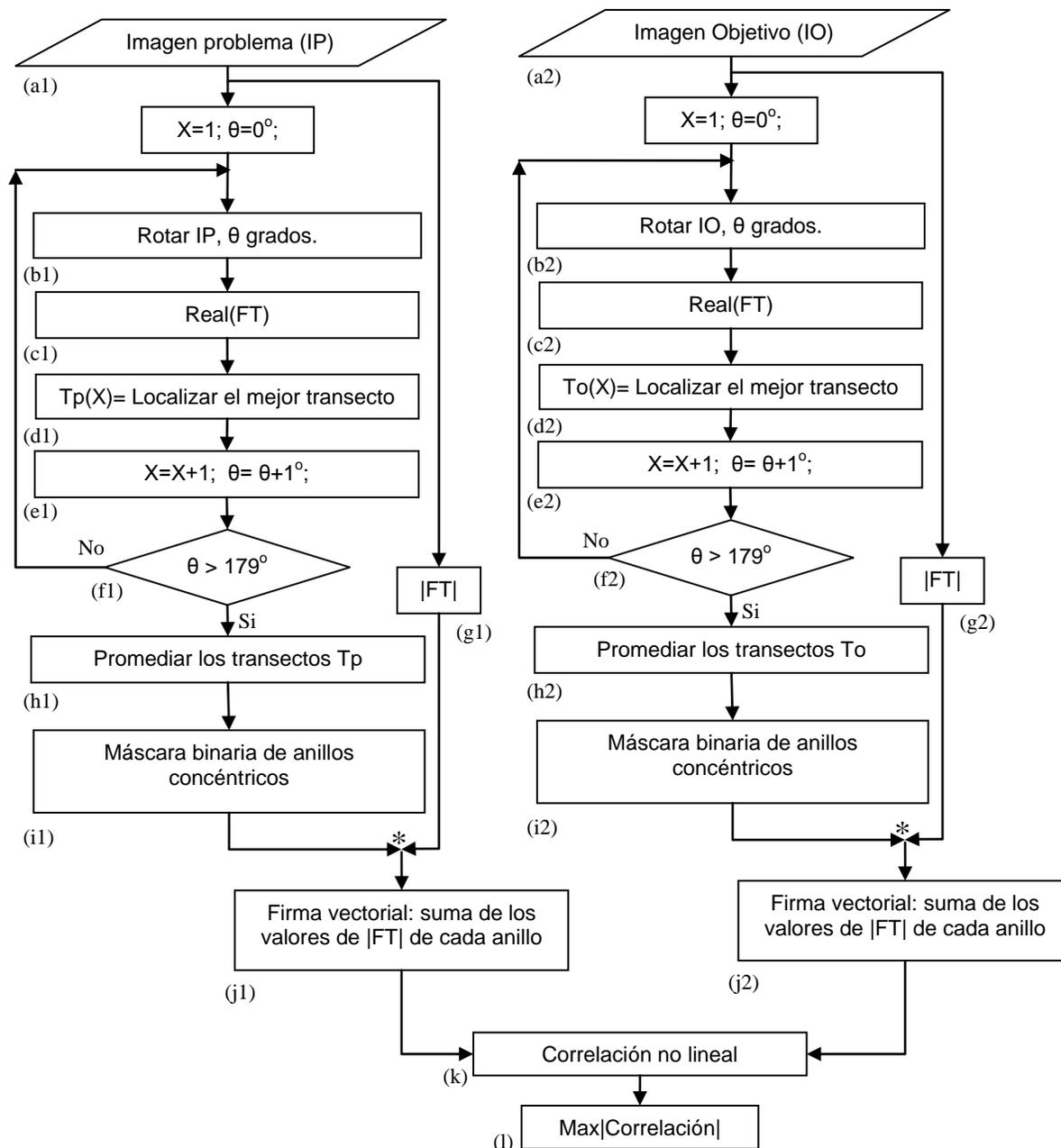


Figura 9. Metodología iterativa de correlación no lineal invariante a posición y rotación.

II.2.1 Identificación del mejor transecto

Existe un gran problema cuando tenemos imágenes del mundo real. No sabemos la rotación que presentan las imágenes en el momento exacto que estas son analizadas y encontrar el mejor transecto de la parte real de la transformada de Fourier nos puede dar la solución a este inconveniente.

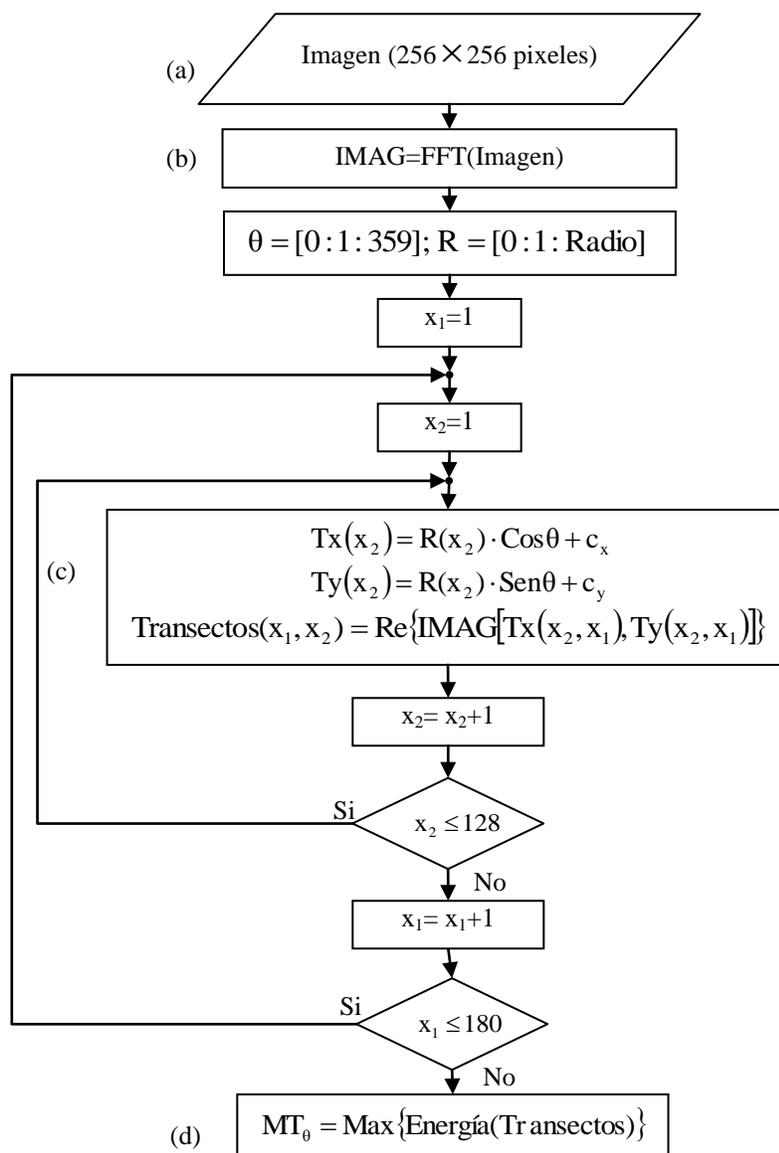


Figura 10. Diagrama de flujo para obtener el mejor transecto de una imagen de 256 × 256 píxeles.

Entonces, de la imagen objetivo (IO) escogemos el mejor transecto el cual será representado por el perfil con mayor energía de la parte real de la transformada de Fourier, para calcular la máscara binaria y su respectiva firma. De la misma manera para cada imagen problema (IP), seleccionamos en el plano de Fourier, el perfil con mayor energía. De esta manera, si la imagen problema es la misma que la imagen de objetivo, el perfil será el mismo, sin importar el ángulo de rotación de las imágenes.

El conjunto de coordenadas que contienen el mejor transecto de la parte real de los valores espectrales son calculadas con las siguientes ecuaciones:

$$Tx = R \cdot \text{Cos}\theta + c_x, \quad (39)$$

$$Ty = R \cdot \text{Sen}\theta + c_y, \quad (40)$$

donde las coordenadas (c_x, c_y) corresponden al centro de la imagen, R es el radio en pixeles, θ es un vector con valores desde 0 hasta 359° con incrementos $\Delta\theta=1^\circ$, Tx contiene el conjunto de coordenadas en x y Ty el conjunto de coordenadas en y .

La figura 10 es un diagrama de flujo que muestra los pasos para la localización del mejor transecto de una imagen (MT_0) teniendo como entrada una imagen de 256×256 pixeles (Figura 10a). Se inicia obteniendo la transformada de Fourier de la imagen de entrada (Figura 10b), utilizando las coordenadas obtenidas con las Ecs. 39 y 40 se obtienen 180 transectos para cada imagen (Figura 10c), los transectos tienen una separación igual a 1° además la parte real del contenido espectral es simétrica de 0 a 179° y de 180 a 359° , es por eso que sólo se utilizan 180 transectos de donde finalmente el mejor transecto (MT_0) es el transecto con mayor energía del conjunto de 180 transectos (Figura 10d).

II.2.2 Máscara adaptativa binaria de anillos concéntricos

La máscara binaria de anillos concéntricos se construye a partir de la parte real de la transformada de Fourier en dos dimensiones de una imagen.

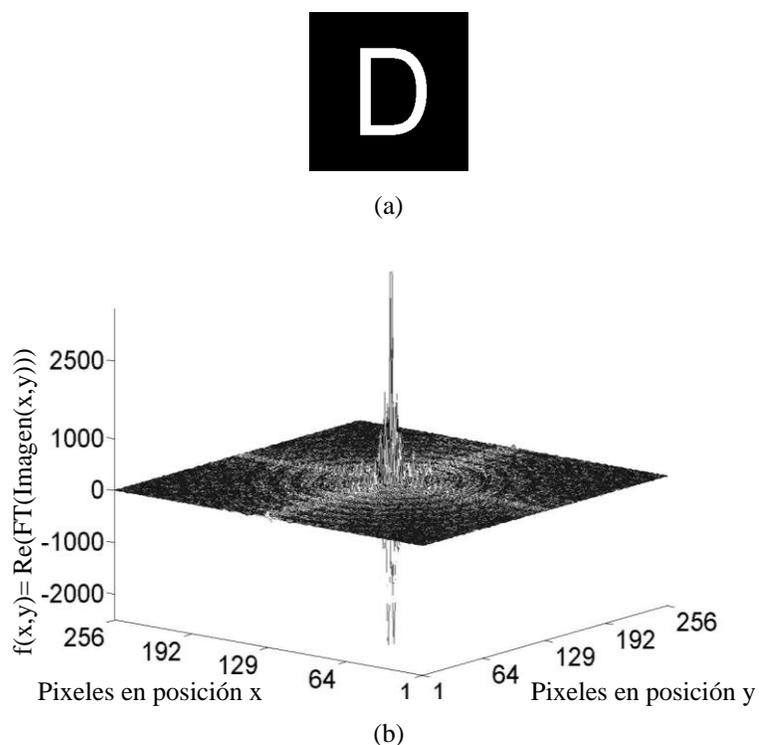


Figura 11. (a) Imagen formada por 256×256 píxeles con la letra D Arial centrada. (b) La parte real de la transformada de Fourier de la imagen.

En la figura 11 se muestra un ejemplo de este procedimiento al obtener la parte real de la transformada de Fourier de una imagen en blanco y negro de una letra D.

Si tomamos el mejor transecto MT_0 de la figura 11b y siguiendo el diagrama de la figura 9, después de obtener el mejor transecto tanto para IP como para IO, lo guardamos en T_p y T_o , estas variables guardarán 180 mejores transectos diferentes cada una (Figura 9d); que corresponden a la imagen problema y a la imagen objetivo, respectivamente. Estos 180 mejores transectos se obtienen debido a que la imagen de entrada se gira (Figura 9b) desde 0 hasta 179° (Figura 9f) con intervalos $\Delta\theta = 1^\circ$ (figura 9e) con el fin de generar las posibles

distorsiones causadas por rotación que puede presentar la imagen. Al cumplirse la condición $\theta > 179^\circ$ se promedian los 180 mejores transectos. Como lo muestra la figura 12 el resultado de este promedio se binariza para formar la función de una variable $Z_p(y)$ y $Z_o(y)$ (Figura 12b),

$$Z_p(y) = \begin{cases} 1, & \text{si promedio}(T_p(y)) > 0 \\ 0, & \text{en cualquier otro caso} \end{cases}, \quad (41)$$

$$Z_o(y) = \begin{cases} 1, & \text{si promedio}(T_o(y)) > 0 \\ 0, & \text{en cualquier otro caso} \end{cases}. \quad (42)$$

Después, tomando el eje vertical como eje de rotación, las gráficas de Z_p y Z_o son rotadas 360° para obtener anillos concéntricos de diferentes anchos. Finalmente, mapeando estos anillos en dos dimensiones se construye la máscara binaria de anillos asociada a las imágenes IP e IO. La figura 13 muestra la máscara de anillos asociada a la imagen de la letra D Arial.

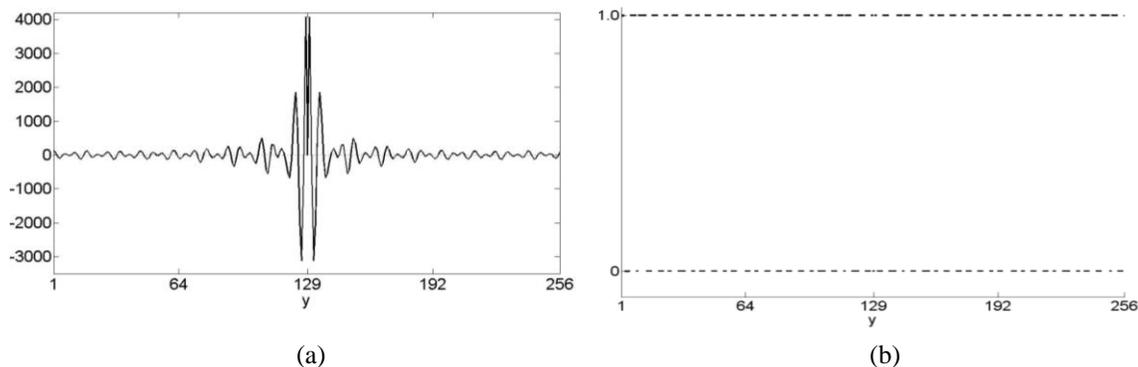


Figura 12. (a) Ejemplo del promedio de los 180 mejores transectos. (b) Binarización del promedio de los 180 mejores transectos.



Figura 13. (a) Imagen objetivo. (b) Máscara de anillos concéntricos asociada a la letra D Arial.

II.2.3 La firma de la imagen

El objetivo principal de este trabajo es identificar una imagen, sin importar el ángulo de rotación que presente en el eje vertical. Para obtener la invariancia a posición se obtiene el módulo de la transformada de Fourier del objeto. La invariancia a rotación se logra cuando la máscara binaria se multiplica punto a punto con el módulo de la transformada de Fourier, finalmente los valores del módulo de la transformada de Fourier de cada anillo se suman para asignarlos al índice del anillo correspondiente y así obtener una firma (figura 14).

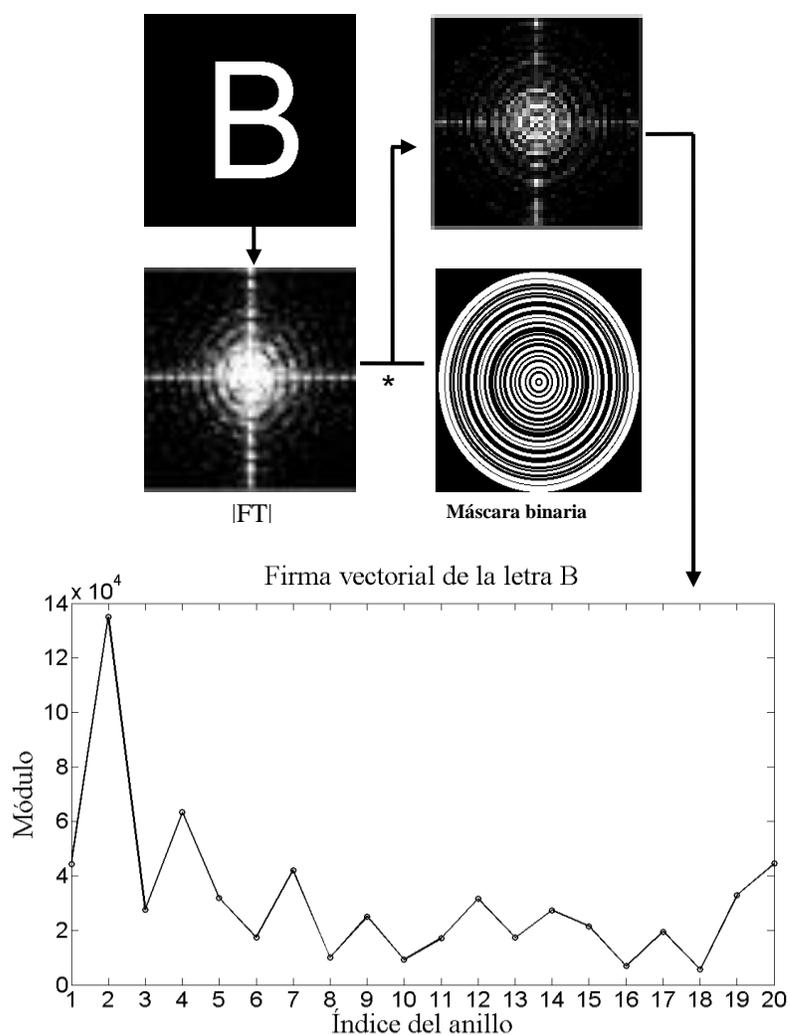


Figura 14. Procedimiento para obtener la firma correspondiente a la imagen binaria de una letra B Arial centrada de 128×128 píxeles; donde el símbolo * especifica una multiplicación punto a punto.

El mismo procedimiento se muestra en la figura 15. De las figuras 14 y 15, podemos observar que en imágenes distintas se producen firmas diferentes.

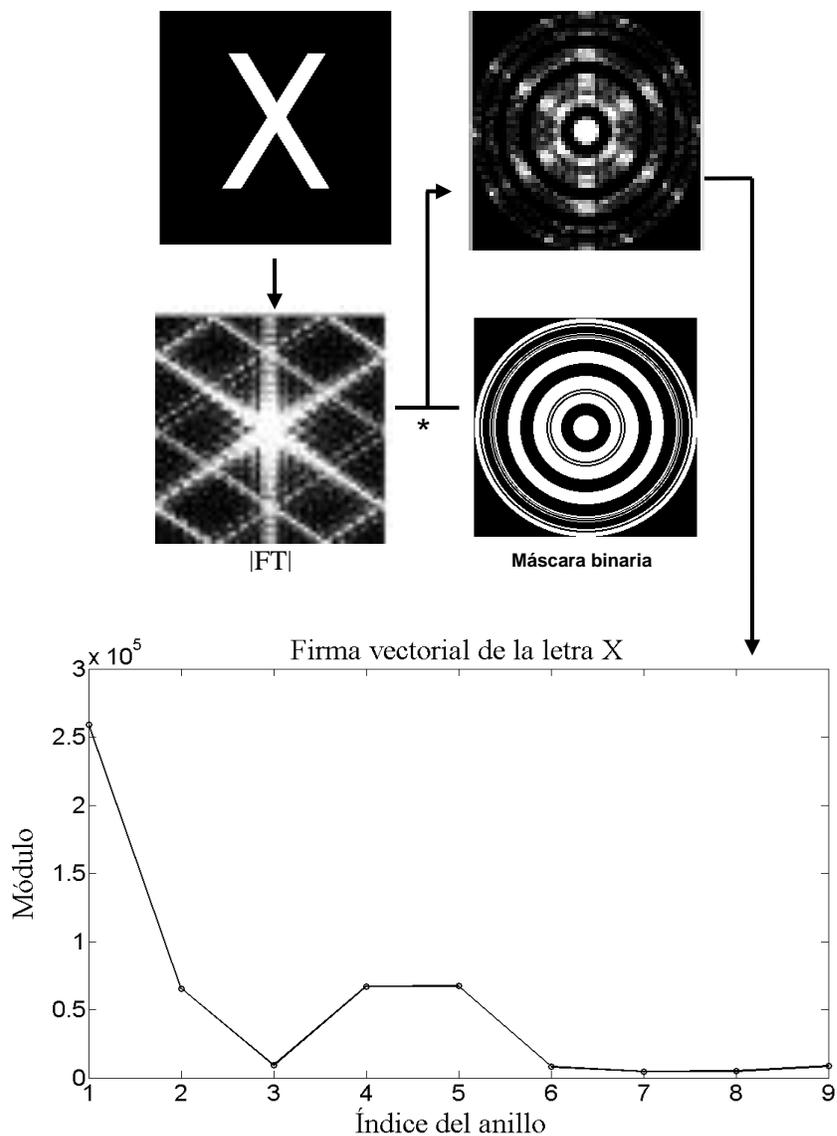


Figura 15. Procedimiento para obtener la firma correspondiente a la imagen binaria de una letra X Arial centrada de 128×128 píxeles; donde el símbolo * especifica una multiplicación punto a punto.

II.2.4. Correlación no lineal

Las correlaciones entre las firmas de la imagen objetivo y la imagen problema se calculan de la siguiente manera:

$$C_{NLI}(F_IP, F_IO) = \text{FFT}^{-1} \left(\left| \text{FFT}(F_IP) \right|^k e^{i\phi_{F_IP}} \cdot \left| \text{FFT}(F_IO) \right|^k e^{-i\phi_{F_IO}} \right), \quad (43)$$

donde FFT^{-1} representa la transformada inversa de Fourier. $|\text{FFT}(F_IP)|$ y $|\text{FFT}(F_IO)|$ son la magnitud de la transformada de Fourier de la firma de la imagen problema y de la imagen objetivo, respectivamente; k es el valor de no linealidad, ϕ_{IP} y ϕ_{IO} son la fase de la transformada de Fourier de la firma de la imagen problema y de la imagen objetivo, respectivamente.

II.3 Metodología adaptativa de correlación no lineal invariante a posición y rotación

Con el fin de producir máscaras binarias que representen a las imágenes de entrada con todo y sus distorsiones causadas por la rotación de la imagen, la metodología iterativa mostrada en la figura 9, en resumen, realiza la FFT-2D, gira la imagen con $\Delta\theta = 1^\circ$ y localiza el mejor transecto (MT_0). Esto se hace 180 veces y se realiza tanto para la imagen problema (IP) como para la imagen objetivo (IO). Debido a esto, al realizar sólo una correlación se estarán realizando 360 transformadas FFT-2D, 360 giros y la localización de 360 mejores transectos; sólo para poder construir dos máscaras binarias de anillos concéntricos, una correspondiente a IP y otra a IO. Con la metodología iterativa se obtuvieron buenos resultados, pero generando un tiempo de cómputo muy grande de aproximadamente 39 segundos y este detalle es una situación no deseada. Para la metodología iterativa, la idea original fue girar las imágenes de entrada desde 0 hasta 359° , pero debido a que las distorsiones generadas por rotación desde 0 hasta 179° , son simétricas a las de 180 hasta 359° , se logró reducir la cantidad de iteraciones de 360 a 180, conservando la efectividad del método. Otro intento por mejorar el tiempo de cómputo fue aumentar el valor $\Delta\theta$ y realizar menos iteraciones, pero esta acción restó efectividad de

discriminación al método. Con el fin de eliminar las iteraciones se desarrolló una nueva metodología que utiliza la misma máscara binaria de anillos concéntricos tanto para la imagen objetivo como para la imagen problema y una nueva correlación no lineal, eliminando así las iteraciones. Esta nueva metodología es llamada metodología adaptativa.

II.3.1 Máscara de anillos concéntricos, utilizada para la metodología adaptativa

A diferencia de la metodología iterativa de la figura 9, donde se obtenía una máscara a partir de la imagen problema (IP) y otra máscara a partir de la imagen objetivo (IO), en esta metodología se va a utilizar una misma máscara de anillos concéntricos para las dos imágenes, la cual debe ser generada a partir de la imagen objetivo (IO). La máscara binaria de anillos concéntricos (MBAC) para cada imagen objetivo, se construye siguiendo el diagrama de la figura 16.

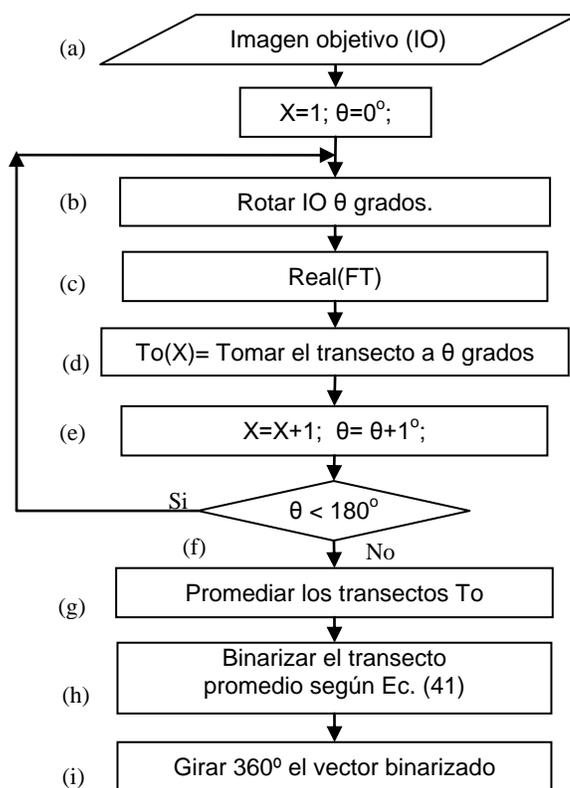


Figura 16. Diagrama de flujo a seguir para construir la máscara binaria de anillos concéntricos utilizada en la metodología adaptativa.

La máscara binaria de anillos concéntricos (MBAC) se genera a partir de IO (Figura 16a). Se elimina la búsqueda del mejor transecto y en cambio se utiliza el transecto central a θ grados de la parte real de la transformada Fourier (Figura 16d). A partir de las iteraciones de los pasos b, c, d y e de la figura 16, se obtendrán 180 transectos T_o , los cuales serán promediados (Figura 16g) y al binarizar este transecto promedio como lo indica la Ec. (41) (Figura 16h), lograremos obtener la única máscara que se va a utilizar en el paso f de la figura 17.

La solución para la reducción del tiempo de cómputo se logró siguiendo el diagrama de flujo de la figura 17. En este diagrama se han eliminando las iteraciones y se ha agregando una variante a la Ec. (43), para lograr una correlación no lineal adaptativa. Esta nueva metodología, llamada metodología adaptativa, logró reducir el tiempo de cómputo de 39 segundos a 0.19 segundos.

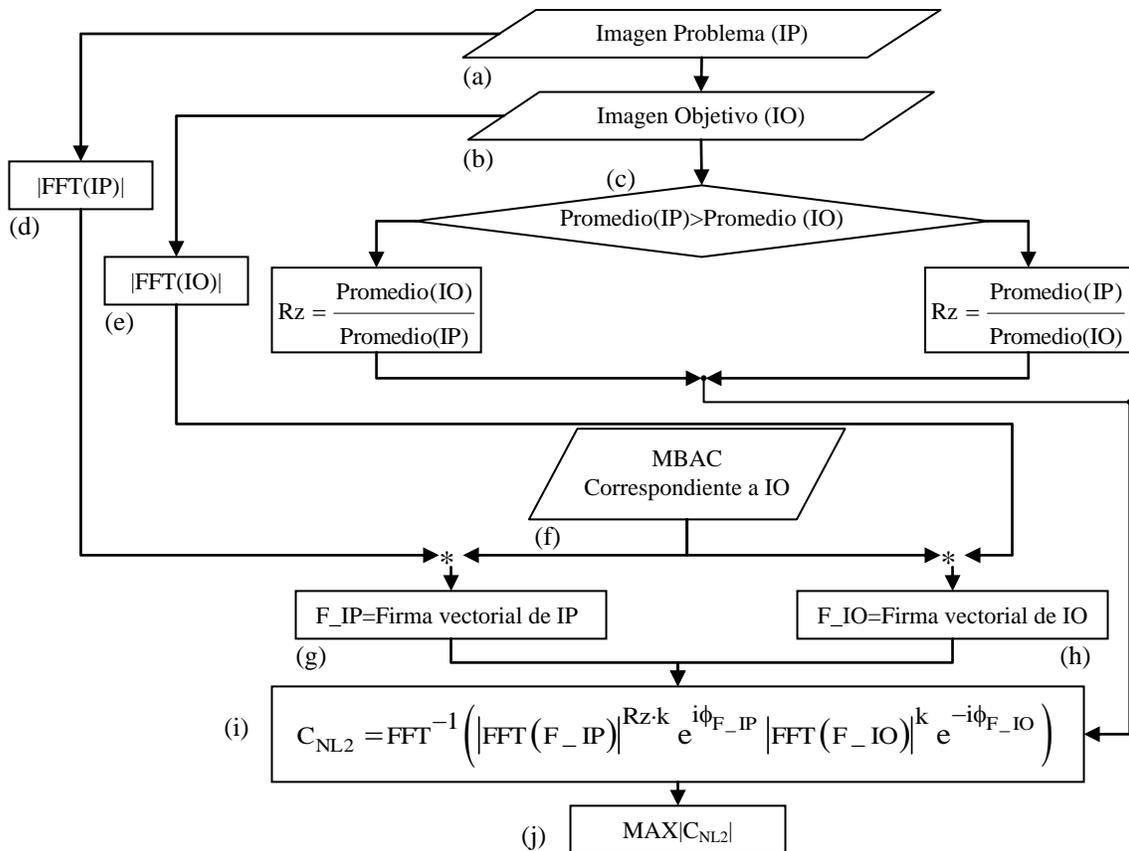


Figura 17. Metodología adaptativa de correlación no lineal invariante a posición y rotación.

II.3.2 Índice invariante a rotación

El valor del área de una imagen, puede ser utilizado como un índice invariante a rotación. Esta invariancia se debe a que el valor del área de la imagen es el mismo, sin importar la rotación que la imagen presente (Fimbres Castro *et al.*, 2012). De la misma manera, el promedio de la imagen puede ser utilizado como índice invariante a rotación, ya que el valor promedio de la imagen siempre será el mismo sin importar el ángulo de rotación de la imagen problema.

El índice invariante a rotación (R_z) se introduce en la Ec. (43) para formar una correlación no lineal adaptativa mostrada en la Ec. (45). Este índice se obtiene a través de la razón entre los promedios de la imagen objetivo y la imagen problema (Figura 17c).

$$\begin{aligned} & \text{SI Promedio(IP)} > \text{Promedio(IO)} \text{ ENTONCES,} \\ R_z &= \text{Promedio(IO)} / \text{Promedio(IP)} \text{ SI NO, } R_z = \text{Promedio(IP)} / \text{Promedio(IO)} \end{aligned} \quad (44)$$

Entonces, nos podemos dar cuenta que el índice R_z , siempre será menor a la unidad y cuando la imagen problema corresponda a la imagen objetivo, este índice será igual o aproximadamente igual a uno.

II.3.3 La firma utilizada en la metodología adaptativa

La máscara binaria de anillos concéntricos (MBAC) correspondiente a la imagen objetivo (figura 17f) se obtiene siguiendo todos los pasos del diagrama mostrado en la figura 16. En seguida, se realiza una multiplicación punto por punto entre el módulo de la transformada de Fourier de cada imagen (Figuras 17d y 17e) y la máscara binaria (Figura 17f), para formar dos firmas vectoriales unidimensionales: F_{IP} (Figura 17g) y F_{IO} (Figura 17h); las cuales representan a la imagen problema y a la imagen objetivo, respectivamente.

II.3.4 Correlación no lineal adaptativa

Si se utiliza la misma máscara de anillos para dos imágenes diferentes, a pesar de que estas imágenes generen patrones de difracción distintos entonces tendremos dos firmas muy parecidas entre sí, las gráficas de la figura 18 muestran valores de correlación para el filtro letra E que se generó utilizando la metodología adaptativa de la figura 17.

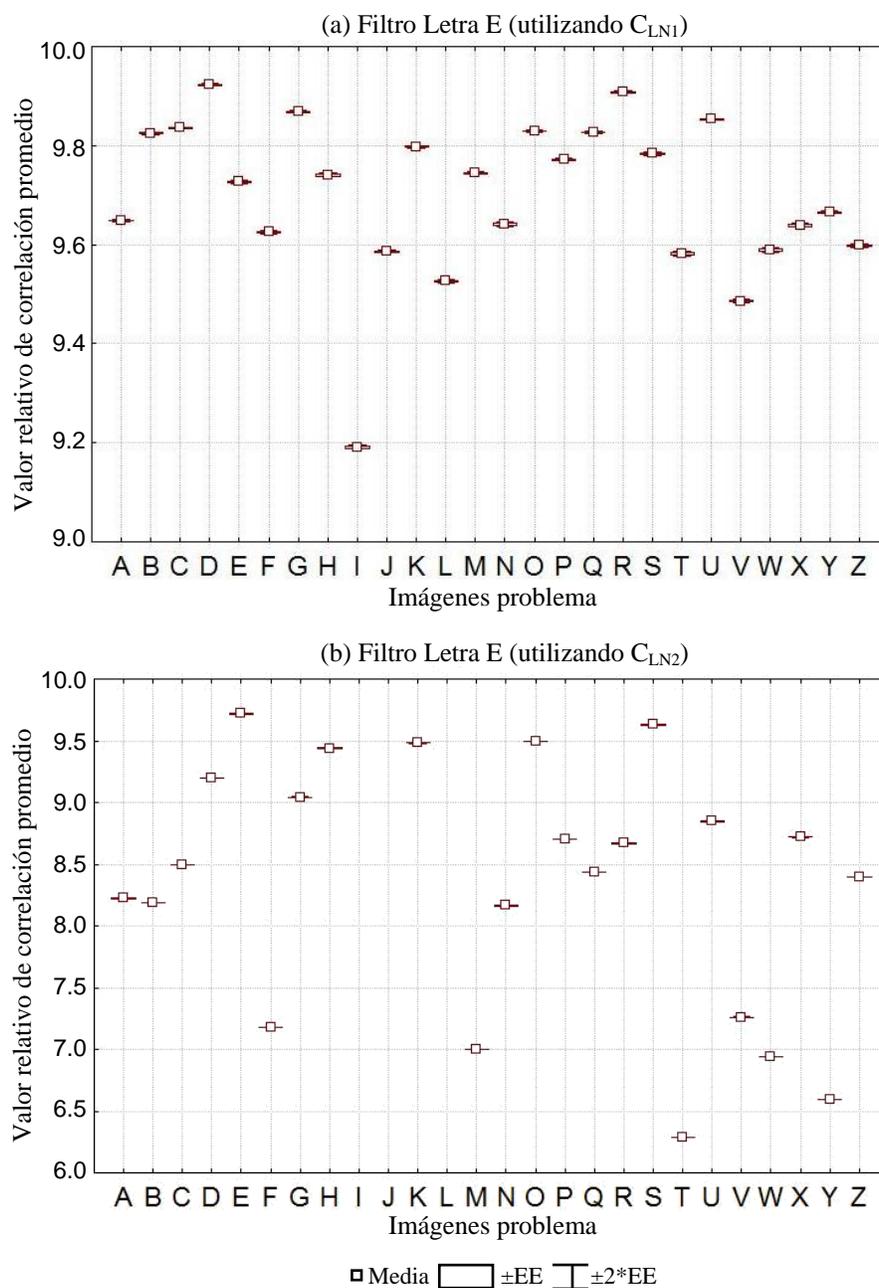


Figura 18. Valores de correlación para el filtro de la letra E obtenidos con la metodología adaptativa, (a) utilizando la correlación no lineal C_{NL1} , (b) utilizando la correlación no lineal adaptativa C_{NL2} .

En la figura 18a se muestran los valores de correlación obtenidos al utilizar la Ec. (43) como medio de comparación en el paso i de la figura 17, cuando se utiliza una misma máscara para todas las correlaciones se provoca que imágenes diferentes correlacionen con valores similares. Esto quiere decir que la correlación entre dos firmas será alta aun cuando la imagen problema no corresponde a la imagen objetivo. Para solucionar este problema se necesitó introducir un índice invariante a rotación adicional dado por la Ec. (44). Este índice multiplica al exponente k de la componente de magnitud correspondiente a la firma de la imagen problema en la Ec. (43), para formar una nueva correlación no lineal adaptativa C_{NL2} , tal como se muestra a continuación:

$$C_{NL2}(F_IP, F_IO) = \text{FFT}^{-1} \left(| \text{FFT}(F_IP) |^{Rz \cdot k} e^{i\phi_{F_IP}} | \text{FFT}(F_IO) |^k e^{-i\phi_{F_IO}} \right); \quad (45)$$

donde Rz representa la razón entre los promedios de la imagen problema y la imagen objetivo según la Ec. (44), FFT^{-1} representa la transformada inversa de Fourier, $| \text{FFT}(F_IP) |$ y $| \text{FFT}(F_IO) |$ son la magnitud de la transformada de Fourier de la firma de la imagen problema y de la imagen objetivo, respectivamente, k es el valor de no linealidad, ϕ_{IP} y ϕ_{IO} son la fase de la transformada de Fourier de la firma de la imagen problema y de la imagen objetivo, respectivamente.

En la metodología adaptativa de la figura 17 se utiliza la Ec. (45) como método de comparación. Esta es una correlación no lineal adaptativa entre dos firmas vectoriales (Figura 17i). En la figura 18 se muestra una comparación de los valores de correlación para el filtro de la letra E utilizando la correlación no lineal dada por la Ec. (43) (figura 18a) y la correlación no lineal adaptativa dada por la Ec. (45) (figura 18b). Se aprecia claramente que la correlación no lineal adaptativa (C_{NL2}) discrimina con éxito las letras utilizando la misma máscara binaria, aun cuando estas generan patrones de difracción similares, como es el caso del patrón de difracción de la letra E y la letra F (Asselin *et al.*, 1994). La excelente discriminación entre la letra E y la letra F mostrada en la figura 18b, se obtiene debido a

que el valor promedio de la letra E es mayor al valor promedio de la letra F y la Ec. (45) que utiliza la razón entre los promedios de las imágenes, genera un valor de Rz menor a uno.

Para la figura 18b se utilizó una escala de 6.0 a 10.0 con el fin de tener una visualización más clara de los valores de correlación promedio, los valores de correlación promedio para las letras I, J, y L no se alcanzan a visualizar en la figura 18b debido a que estas letras generaron valores de correlación promedio menores que 6.0.

Capítulo III. Resultados

III.1 Reconocimiento de imágenes utilizando la metodología iterativa

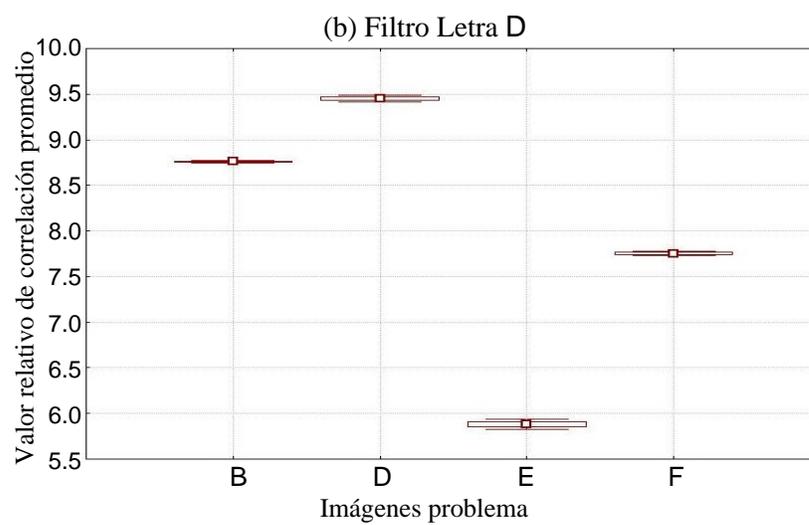
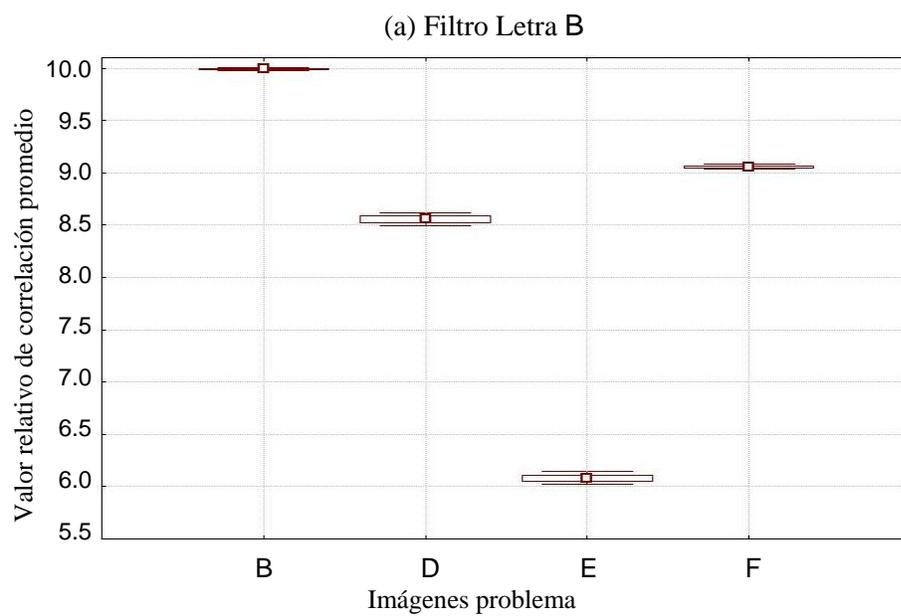
Los resultados obtenidos con la metodología iterativa permiten cumplir los objetivos de este trabajo, con el inconveniente de que cada iteración genera tiempo de cómputo. Esto significa que la identificación de la imagen problema es muy lenta (aproximadamente 39 segundos por correlación), lo cual hace complicada su aplicación en el caso de querer identificar imágenes dentro de una base de datos con miles de imágenes. Además, el elevado tiempo de cómputo de este método hace imposible aplicarlo a la identificación de imágenes en tiempo real. Con la finalidad de eliminar el excesivo tiempo de cómputo, se propuso una solución que elimina las iteraciones de esta metodología. El resultado, la llamada metodología adaptativa, es mucho más rápida (aproximadamente 0.19 segundos por correlación). Primeramente se muestran los resultados obtenidos con el método iterativo utilizando letras del abecedario como imágenes binarias y diatomeas fósiles como imágenes en escala de grises.

III.1.1 Aplicación de la metodología iterativa a imágenes binarias

Con la metodología iterativa se lograron resultados exitosos en cuanto a la identificación de imágenes binarias con invariancia a rotación, debido a que por cada letra tenemos 360 imágenes diferentes representando giros grado por grado se requiere de un tiempo de cómputo aproximado de 3.9 horas ($360 \text{ imágenes} \times 39 \text{ segundos}$) para cada letra del abecedario, debido a esto sólo se seleccionaron las letras B, D, E y F para comprobar la efectividad del método (Figura 19). En la base de datos formada por estas cuatro letras, podemos encontrar a cada una de ellas giradas desde 0 hasta 359° formando un total de 1,440 imágenes diferentes. Se seleccionaron estas letras debido a la similitud que existe entre ellas (Fares et al., 2000; Coronel Beltrán y Álvarez Borrego, 2009; Solorza Calderón y Álvarez Borrego, 2010).



Figura 19. Imágenes binarias de las letras B, D, E y F tipo Arial de 256×256 píxeles



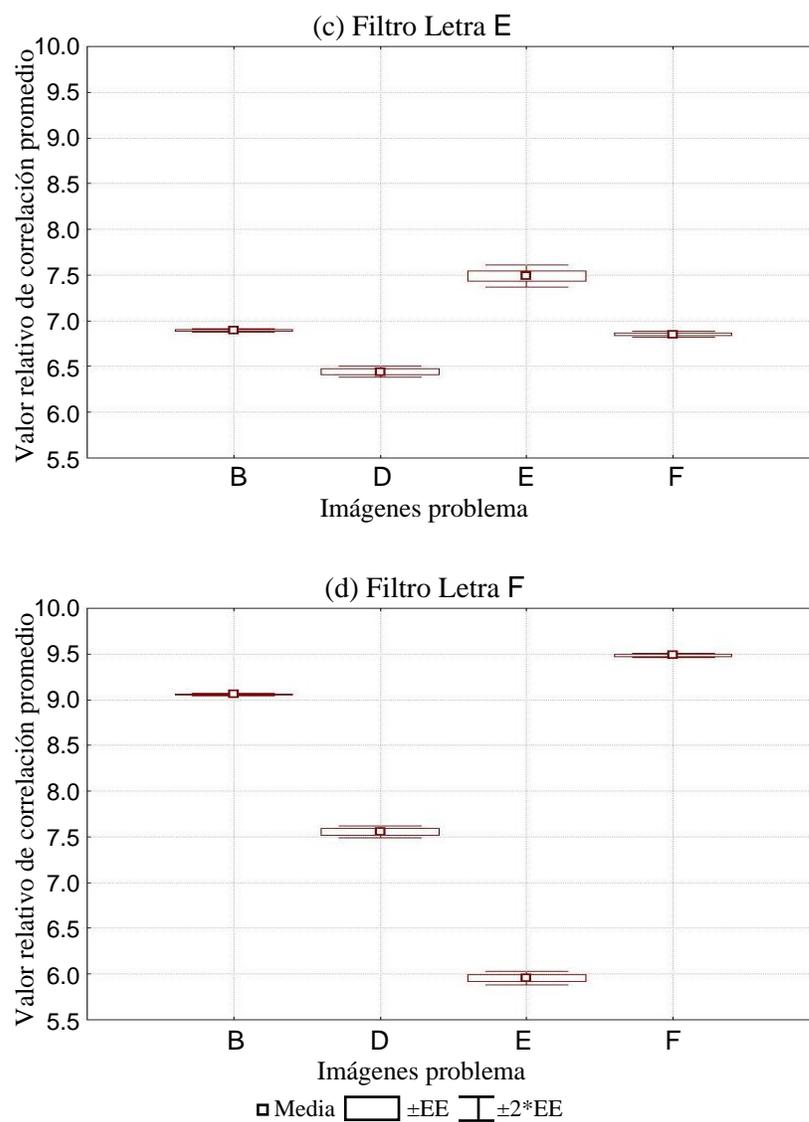


Figura 20. Correlaciones obtenidas utilizando el método iterativo para: (a) filtro letra B, (b) filtro letra D, (c) filtro letra E y (d) filtro letra F

III.1.2 Aplicación de la metodología iterativa a imágenes reales

Como imágenes reales se han seleccionado 5 especies diferentes de diatomeas fósiles, en donde cada una de ellas se gira grado por grado desde 0 hasta 359° formando un total de 1,800 imágenes diferentes con un formato en escala de grises y con un tamaño de 320×320 píxeles. Al igual que en las imágenes binarias, se logró identificar con un nivel de confianza del 100% a las 5 especies de diatomeas fósiles, las especies se etiquetaron con las letras A, B, C, D y E.

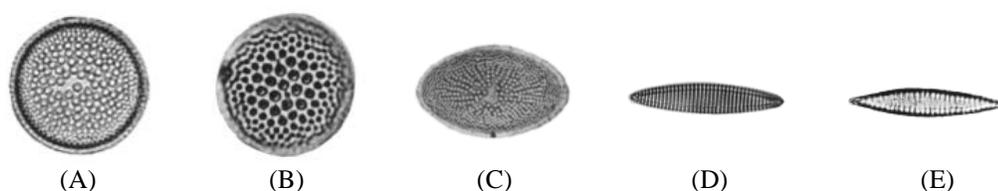
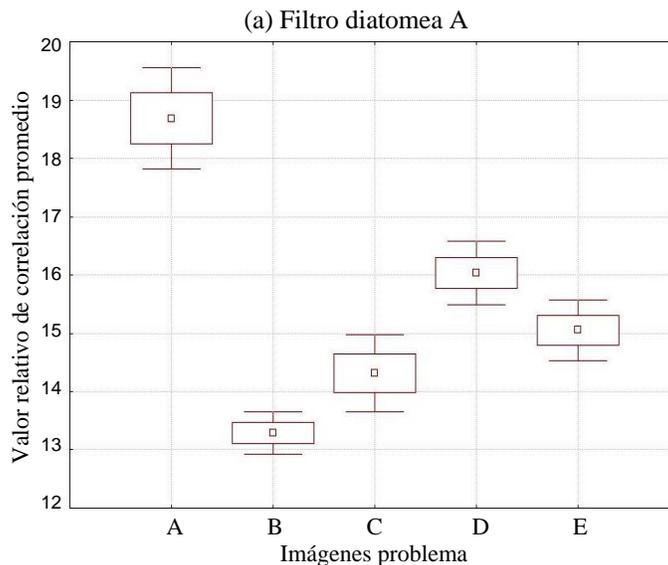


Figura 21. Imágenes de diatomeas utilizadas para probar la metodología iterativa: (A) *Actinocyclus ingens-Ratray*, (B) *Thalassiosira oestruppii var 1*, (C) *Actinocyclus ellipticus-Grunow in van Heurck*, (D) *Nitzschia praereinholdii-Schrader*, (E) *Bogorovia praepaleacea (Schrader) Jouse*

Se seleccionaron las diatomeas fósiles de la figura 21, debido a la similitud visual que existe entre ellas. Por ejemplo, la diatomea C tiene cierta similitud con A y B, pero tiene la característica de estar alargada como D y E. Esperamos además, que el método identifique las diferencias entre las diatomeas A y B ya que ambas tienen una forma circular. Con el mismo propósito se espera que D y E sean correctamente identificadas.



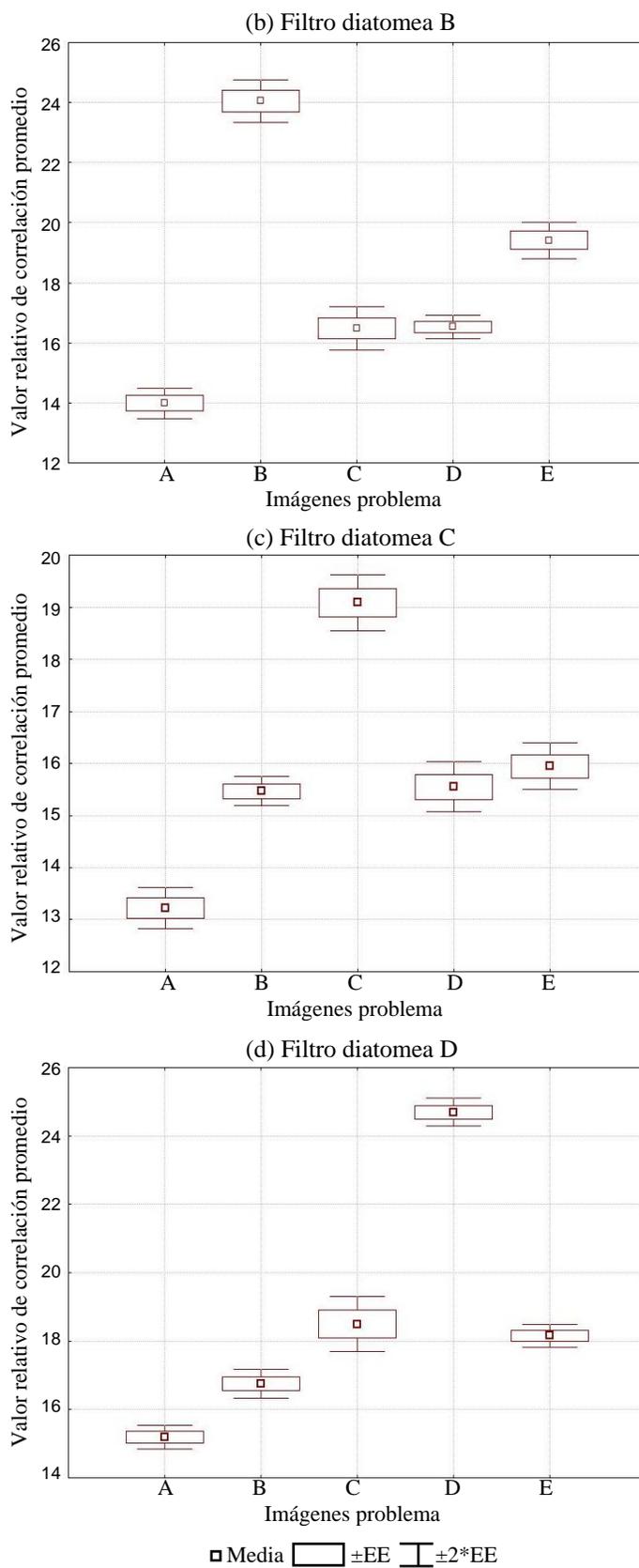


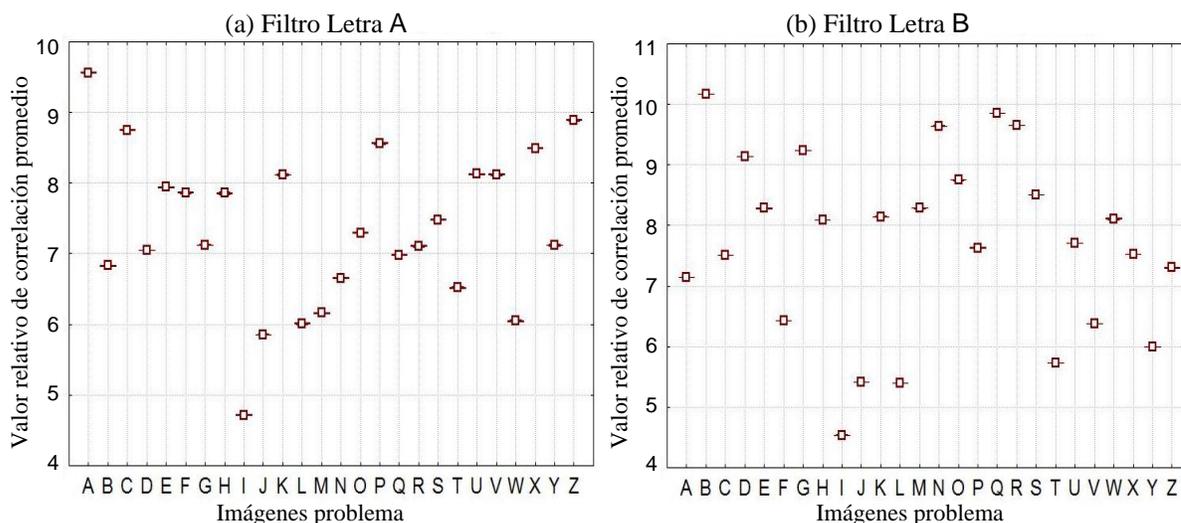
Figura 22. Correlaciones obtenidas utilizando el método iterativo para: (a) filtro diatomea A, (b) filtro diatomea B, (c) filtro diatomea C, (d) filtro diatomea D y (e) filtro diatomea (E)

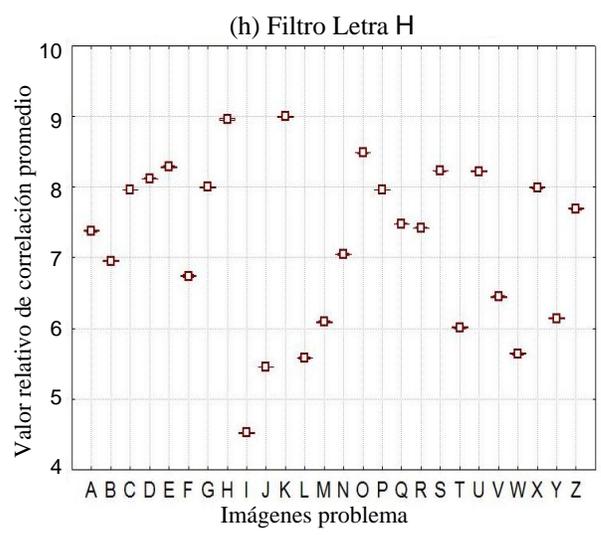
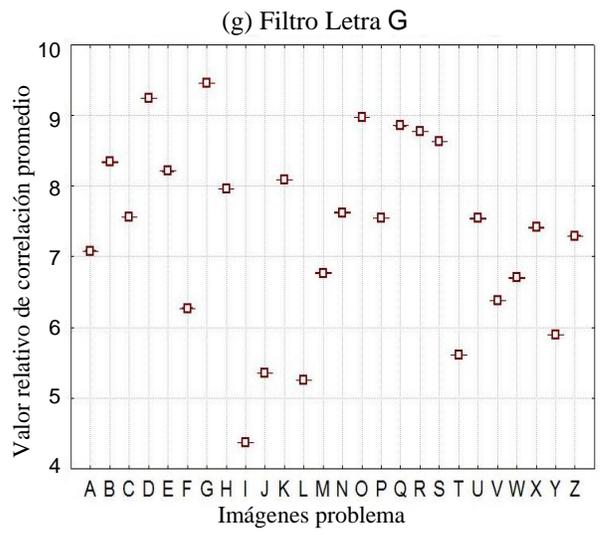
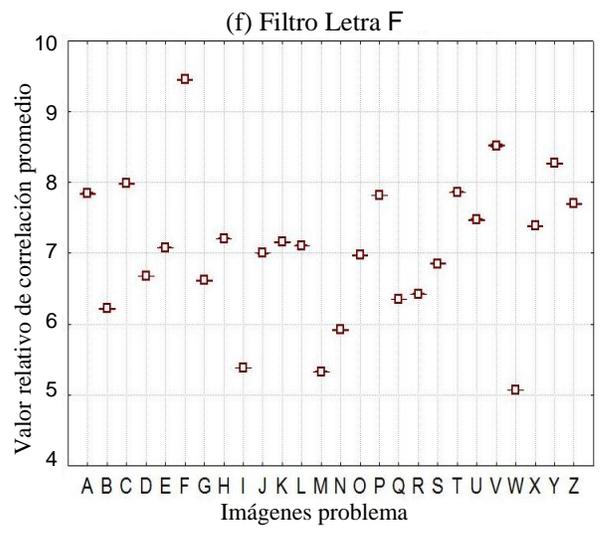
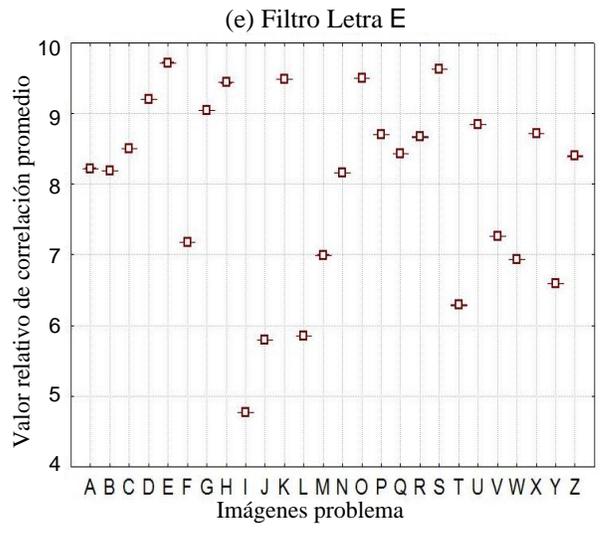
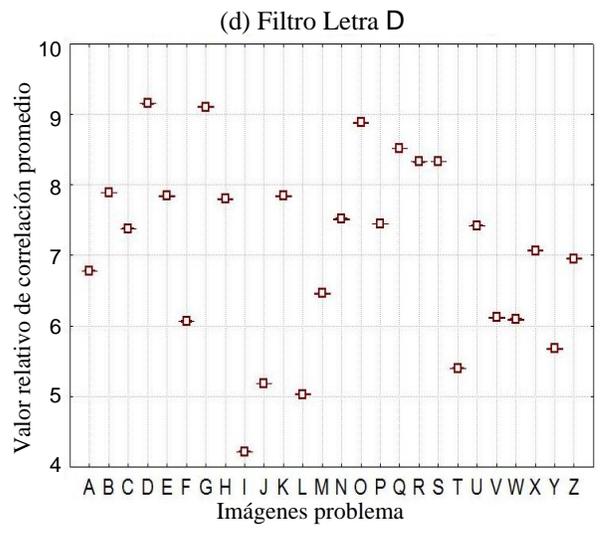
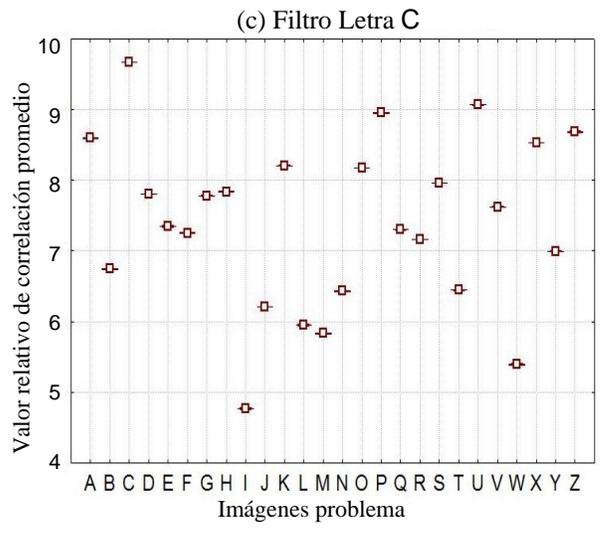
III.2 Reconocimiento de imágenes utilizando la metodología adaptativa

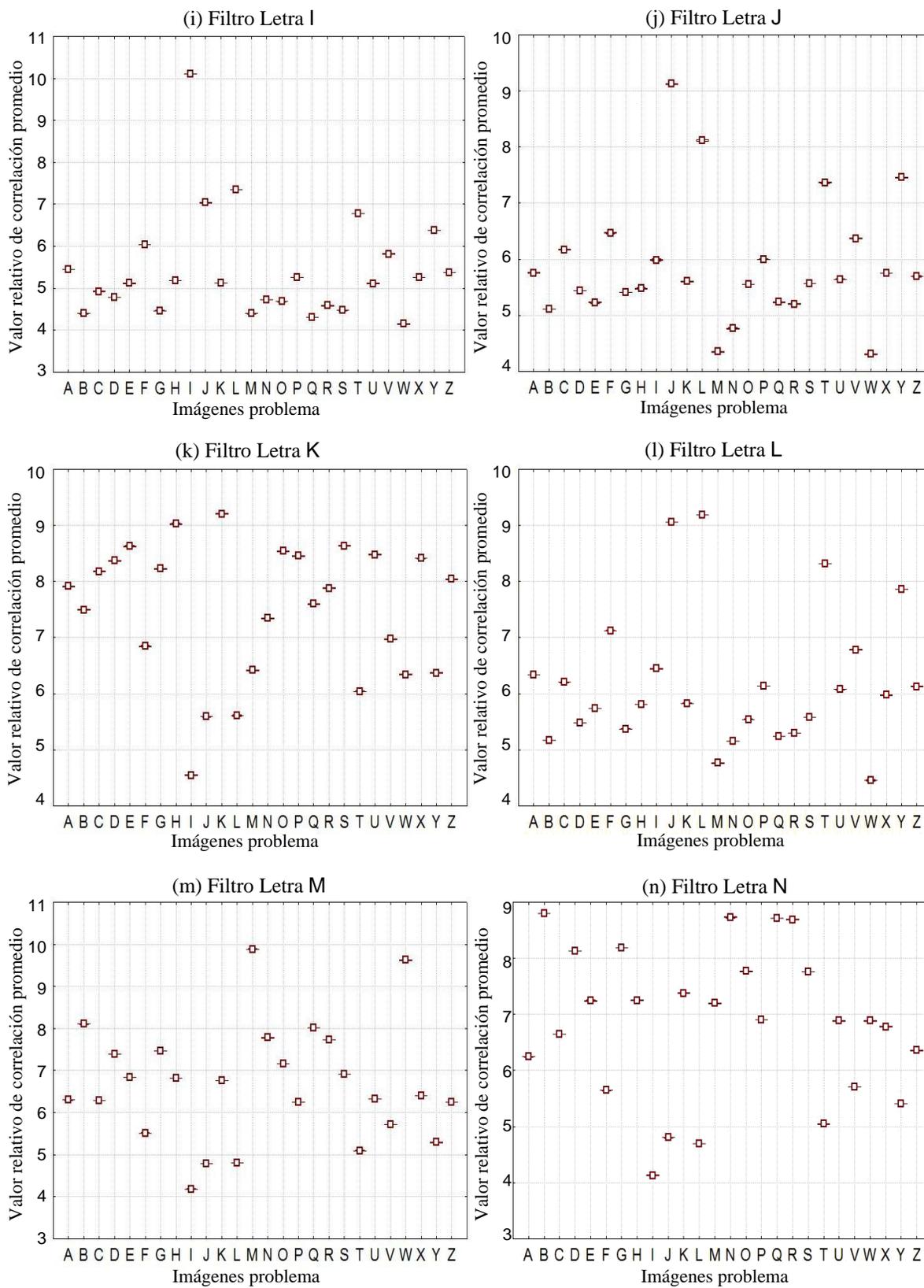
Siguiendo el diagrama de flujo de la figura 17 se obtienen los resultados de la metodología adaptativa, estos resultados se muestran a continuación y son los que finalmente cumplen satisfactoriamente con todos los objetivos, incluyendo la excelente y muy buscada característica de tener un bajo costo computacional. Consiguiendo agilizar el tiempo de cómputo aproximadamente de 39 segundos a 0.19 segundos por cada correlación entre la imagen objetivo y una imagen problema, con respecto a la metodología iterativa mostrada en el diagrama de la figura 9.

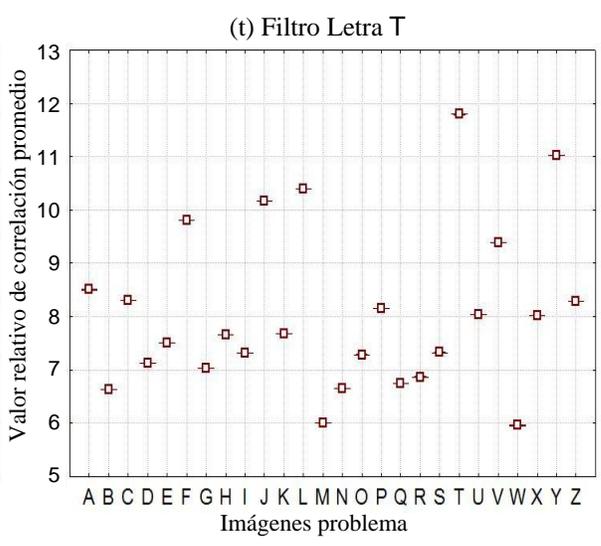
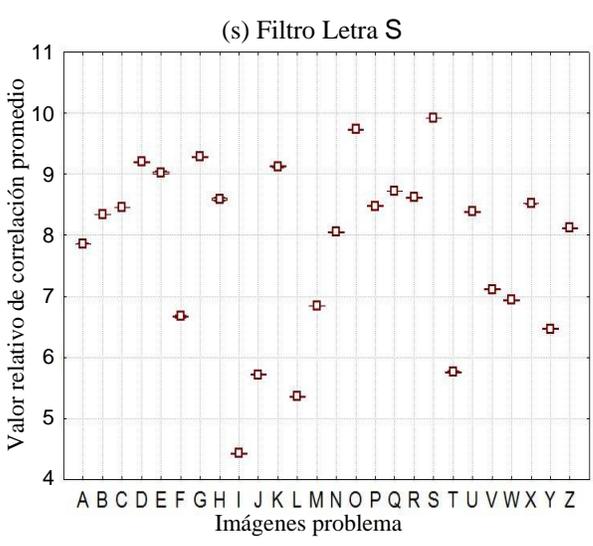
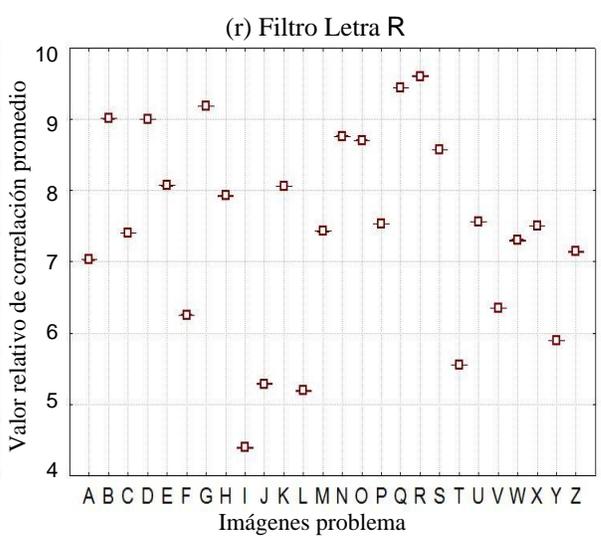
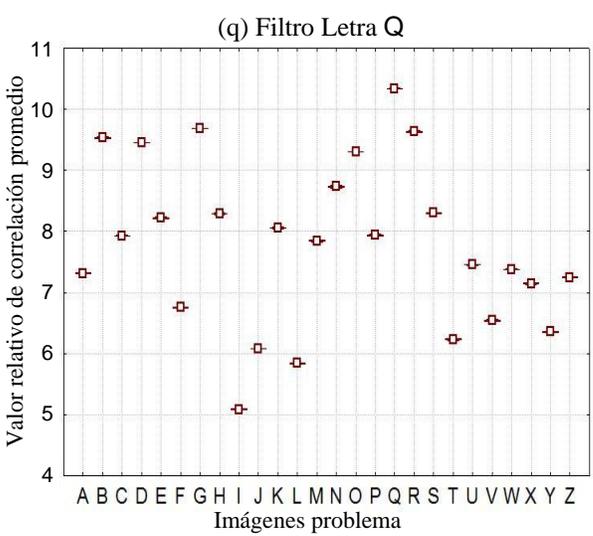
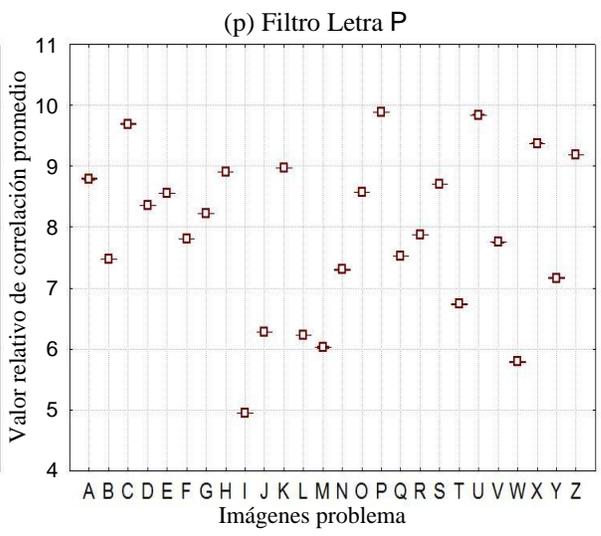
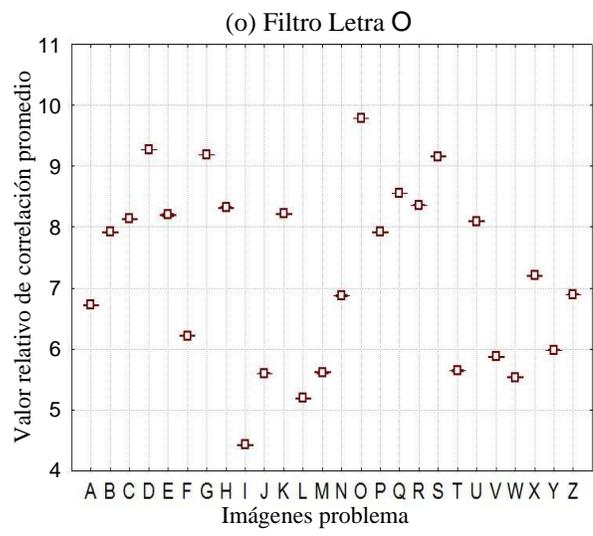
III.2.1 Aplicación de la metodología adaptativa a imágenes binarias

Se utilizaron 26 imágenes que representan todas las letras del abecedario, en un formato blanco y negro con letra tipo Arial. Estas imágenes tienen un tamaño de 256×256 píxeles. Cada una de estas 26 imágenes fue rotada grado por grado desde 0 hasta 359° formando un total de 9,360 imágenes diferentes. Cualquiera de estas imágenes puede ser elegida como imagen objetivo, los experimentos que se presentan en este trabajo utilizaron como imagen objetivo las letras sin rotación. Para completar el experimento fue necesario hacer 243,360 correlaciones diferentes. Estos resultados se muestran en las 26 gráficas de caja, presentadas a continuación.









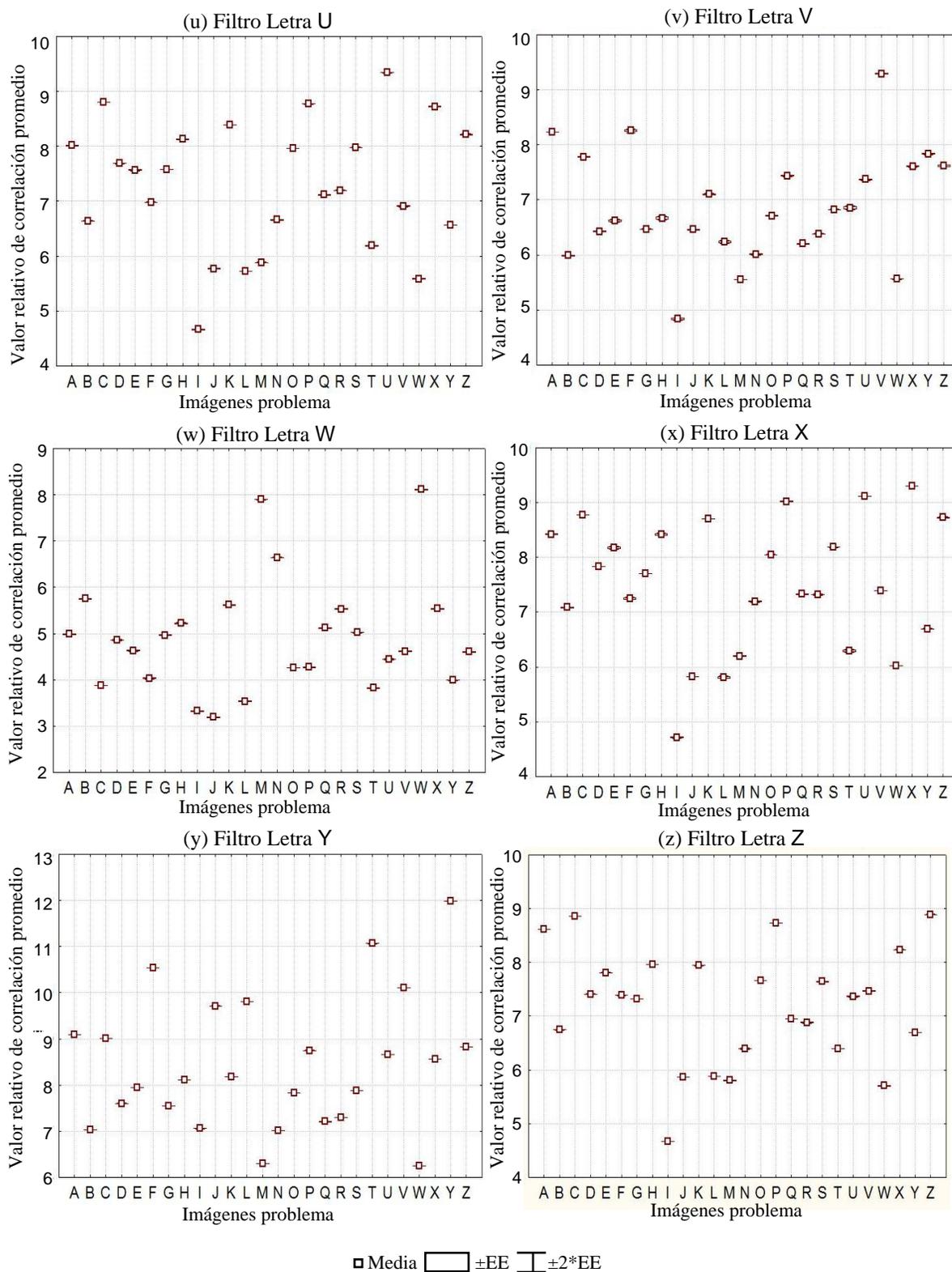
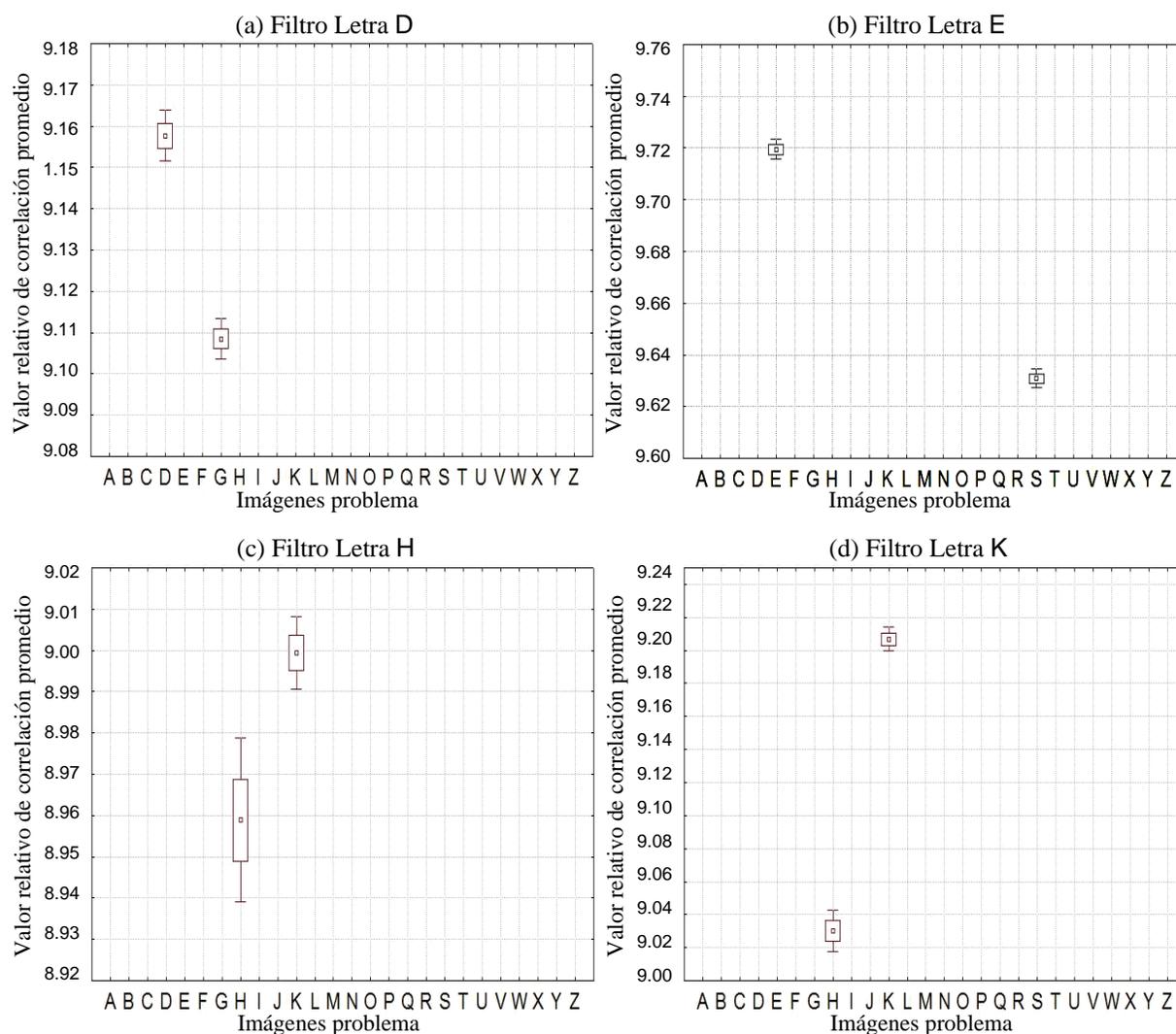
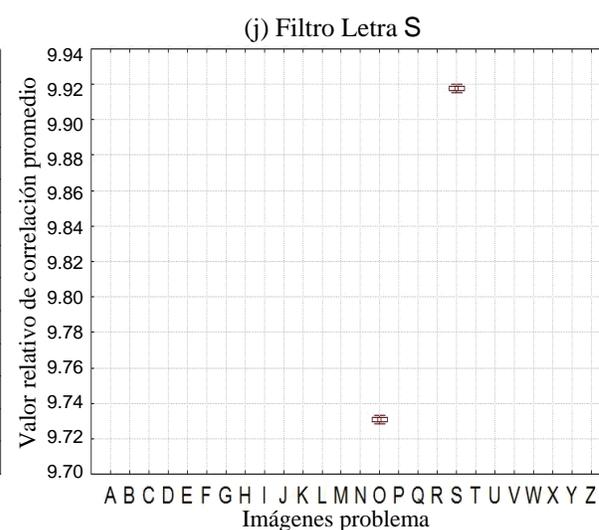
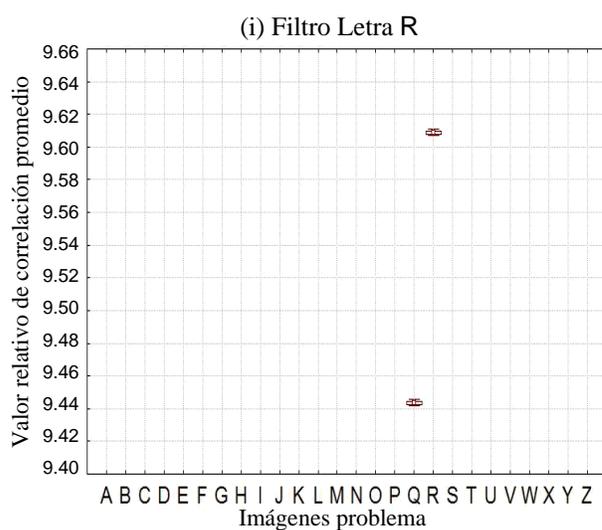
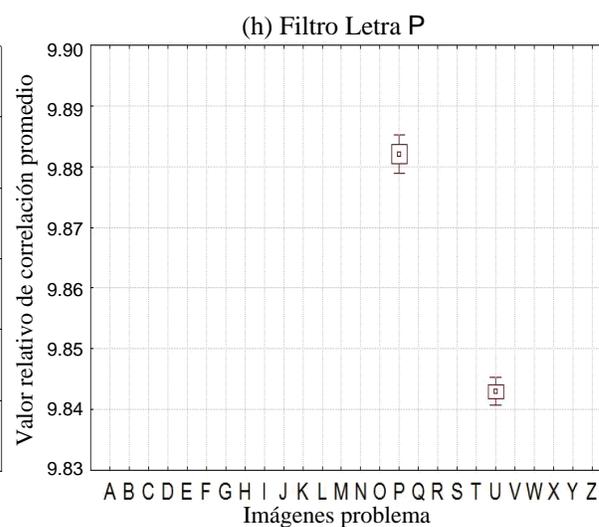
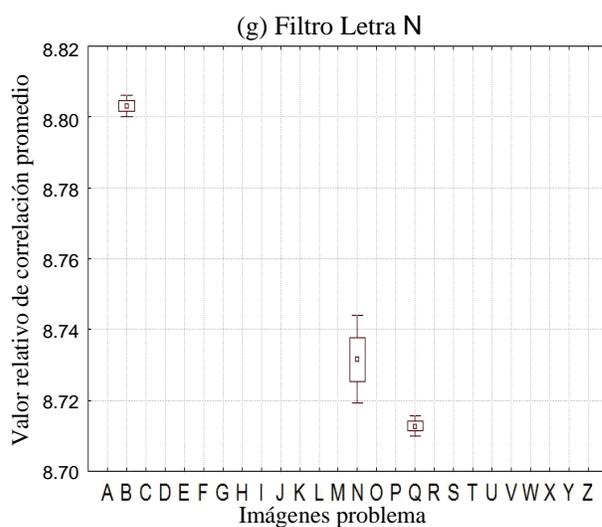
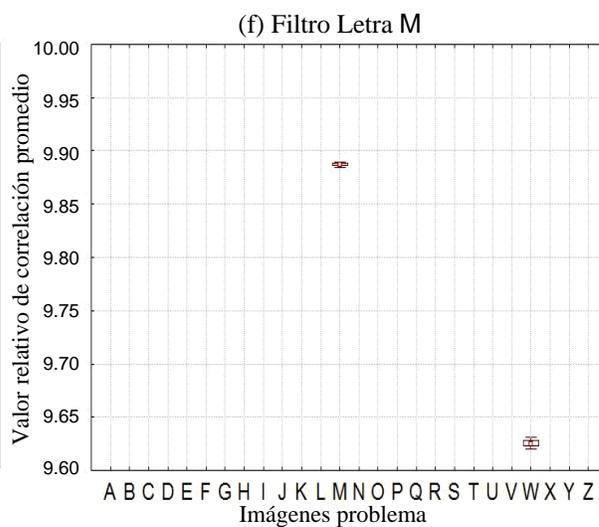
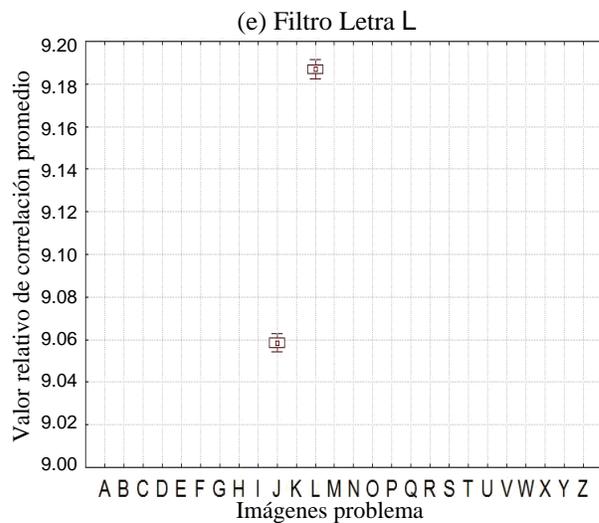


Figura 23. Gráficas de cajas del valor relativo de correlación de las letras del abecedario desde la (A) a la (Z), con respecto al resto, en rotación desde 0 hasta 359° con incrementos de 1°, obtenidas utilizando la metodología adaptativa, donde las cajas representan un error estándar ($\pm EE$) y las barras de las cajas representan dos errores estándar ($\pm 2*EE$) para el pico de correlación.

Los resultados mostrados en la figura 23 han logrado niveles de confianza del 100%. Para verificar con más detalle estos resultados, se realiza un acercamiento a la parte superior de la escala de valores relativos de correlación promedio para poder visualizar con mayor claridad si existe un traslape entre correlaciones correspondientes a imágenes de diferentes letras, asegurando así un nivel de confianza del 100%. Este acercamiento se hace sólo en las gráficas en donde existen valores de correlación promedio cercanos a los valores de correlación promedio resultantes para la imagen objetivo como es el caso de las letras D, E, H, K, L, M, N, P, R, S, X y Z. En la figura 24 se muestran las gráficas con su respectivo acercamiento y se puede apreciar que no existe traslape entre los diferentes valores de correlación.





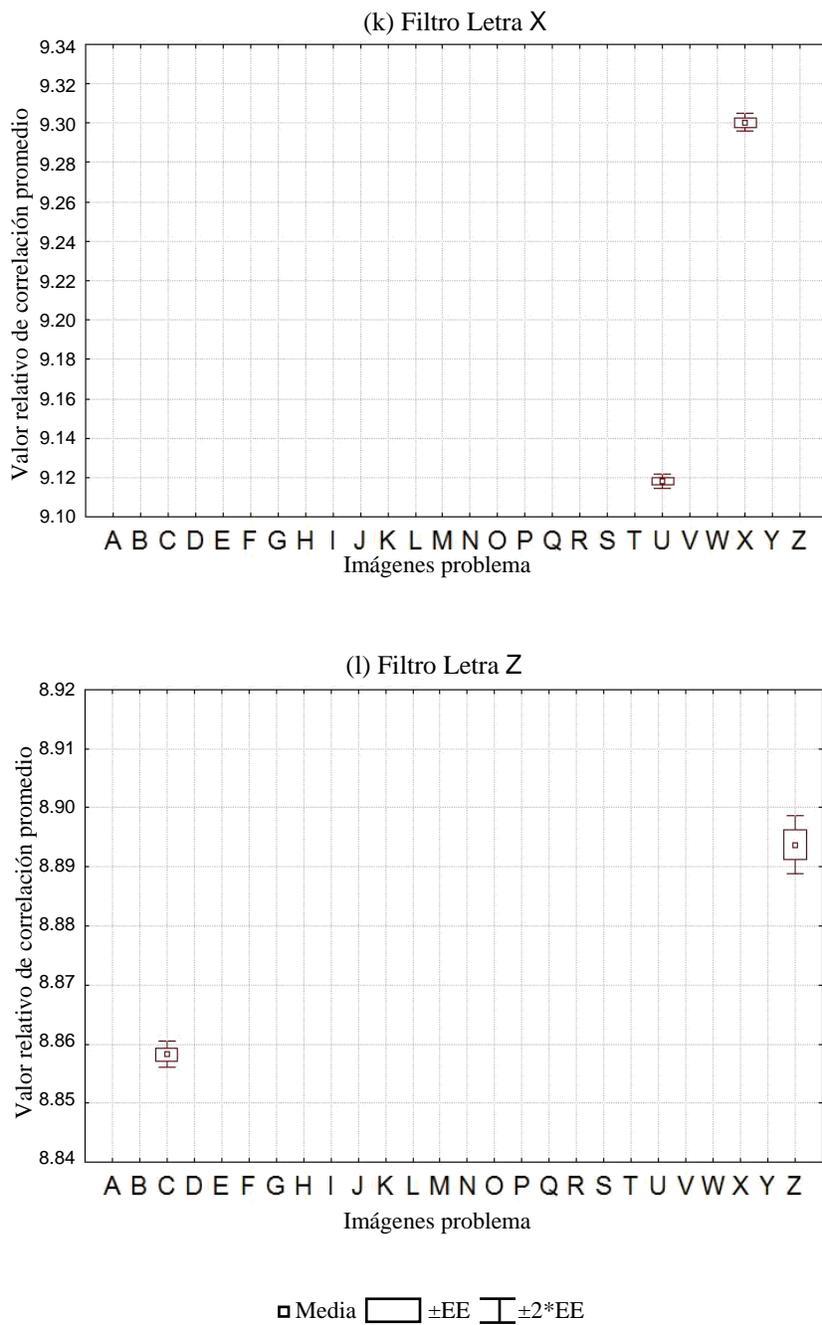


Figura 24. Acercamiento a la escala correspondiente al valor relativo de correlación (Metodología adaptativa), para: (a) Filtro Letra D, (b) Filtro Letra E, (c) Filtro Letra H, (d) Filtro Letra K, (e) Filtro Letra L, (f) Filtro Letra M, (g) Filtro Letra N, (h) Filtro Letra P, (i) Filtro Letra R, (j) Filtro Letra S, (k) Filtro Letra X y (l) Filtro Letra Z.

III.2.2 Aplicación de la metodología adaptativa a imágenes reales

Se analizaron 19 especies diferentes de diatomeas fósiles, en donde cada una de ellas se gira grado por grado desde 0 hasta 359° formando un total de 6,840 imágenes diferentes con un formato en escala de grises y con un tamaño de 320×320 píxeles. Al igual que en las imágenes binarias, se logró identificar con un nivel de confianza del 100% a cada una de las 19 especies de diatomeas fósiles, las especies están ilustradas en la figura 25 y se etiquetaron con las letras A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R y S.

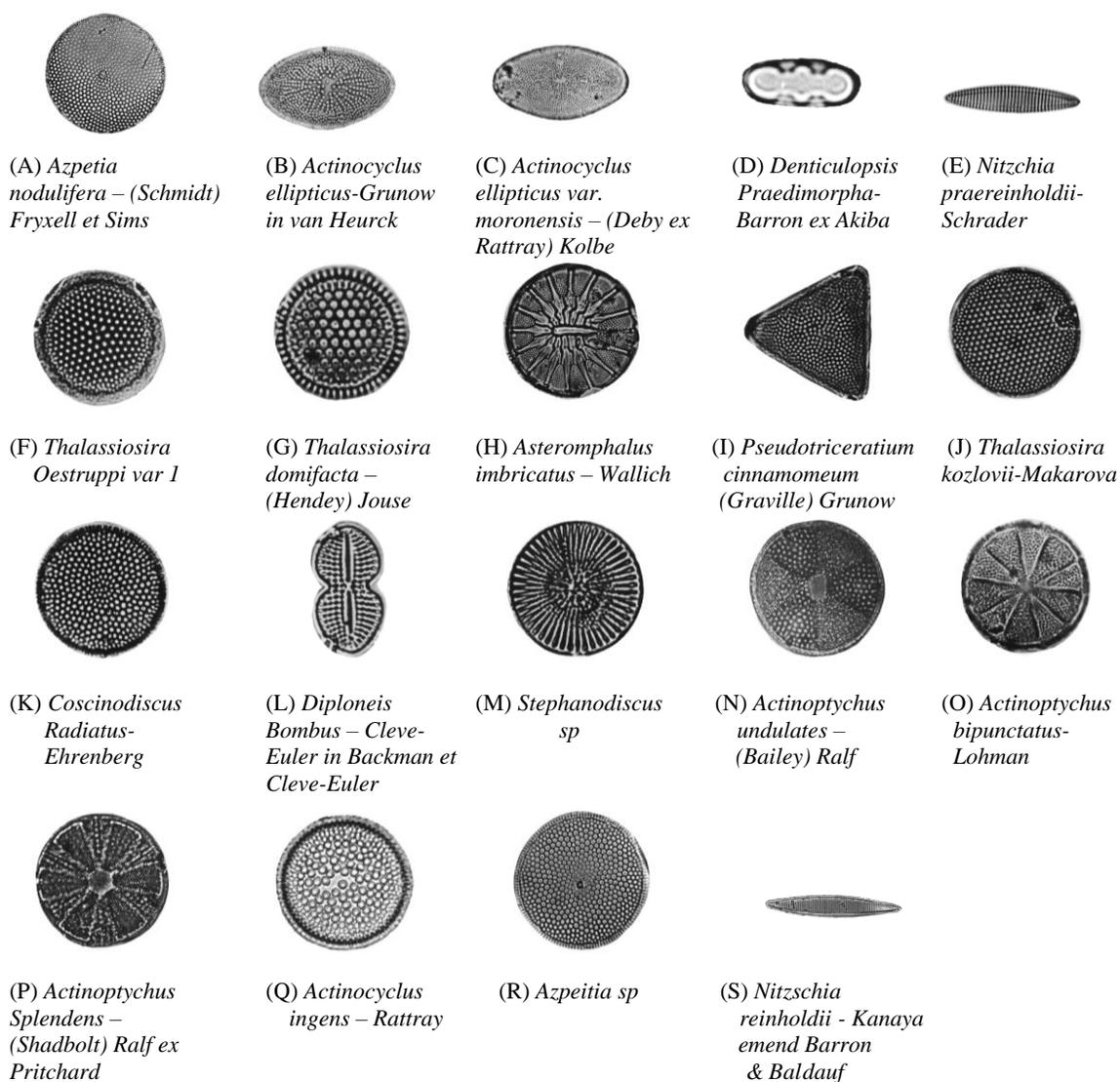
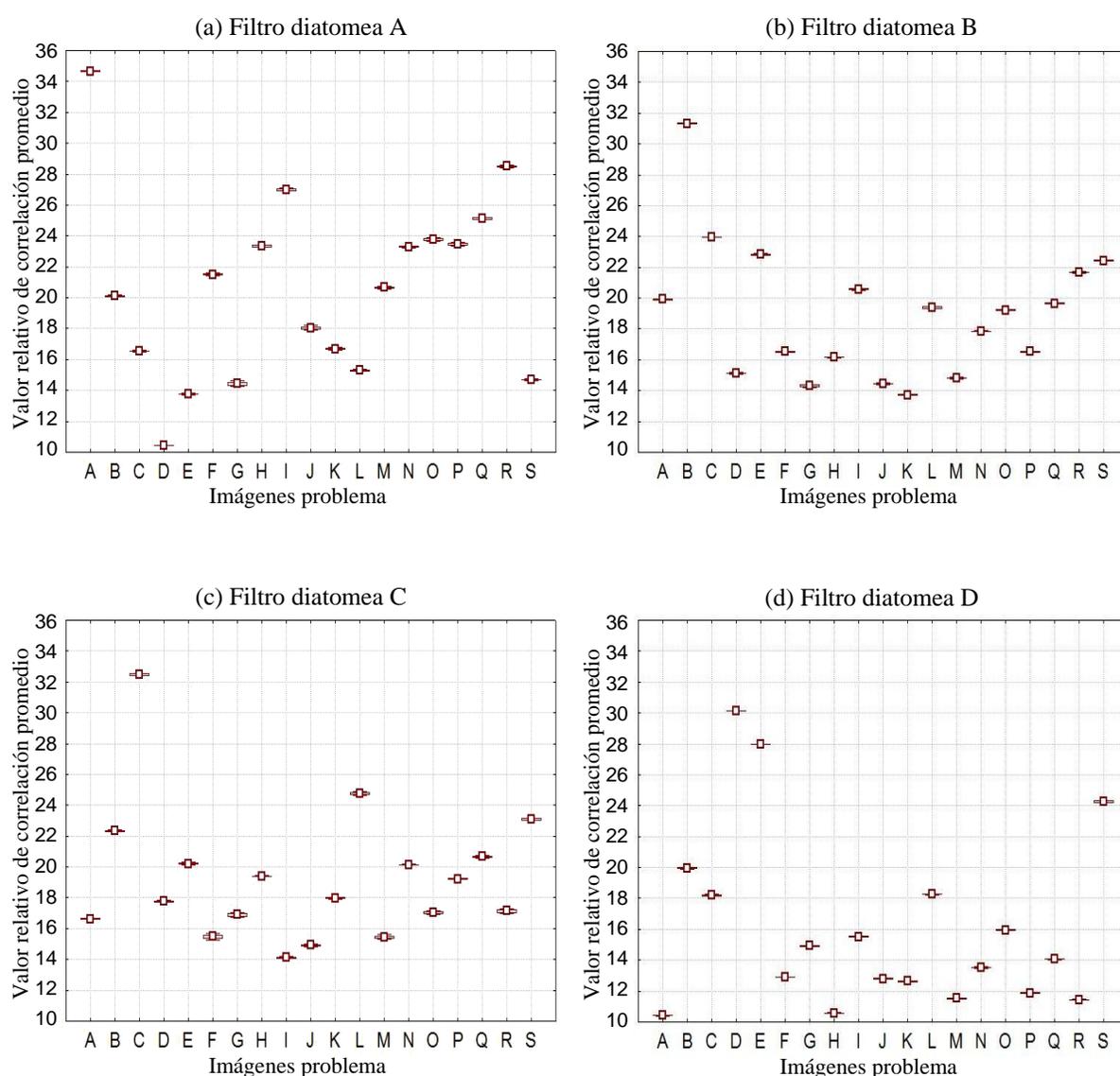
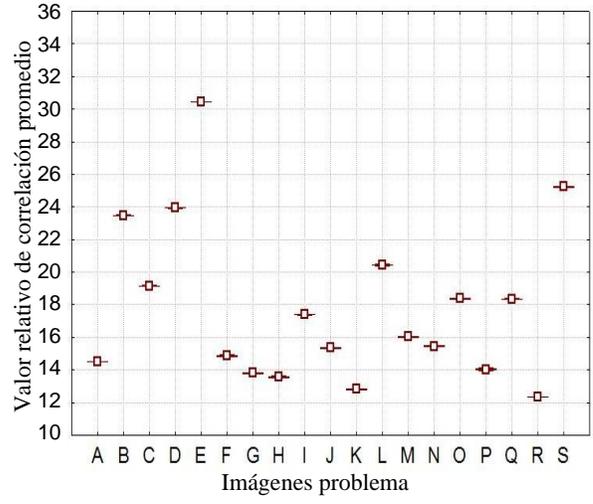


Figura 25. Imágenes de diatomeas utilizadas para probar la metodología adaptativa.

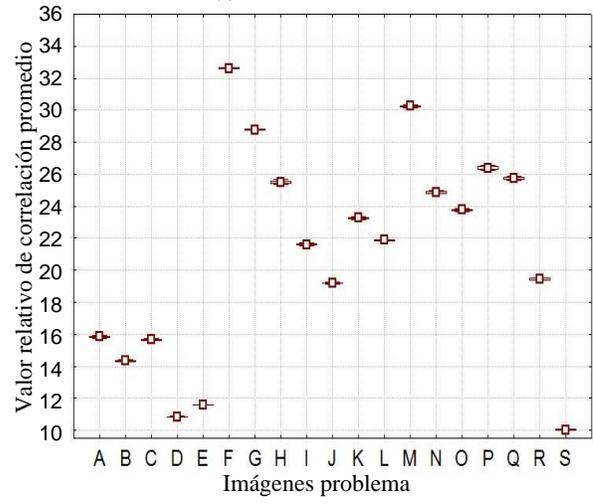
En la figura 26 se muestran los valores de correlación para los filtros de cada una de las diferentes diatomeas en escala de grises. Se nota como cada filtro es capaz de reconocer a la imagen objetivo con un nivel de confianza del 100%, ya que en las gráficas se observa claramente, que los valores de correlación no sufren ningún tipo de traslape. Además, los valores de correlación correspondientes a la imagen objetivo son los de mayor magnitud, para cada filtro. Para completar el experimento fue necesario calcular 6,840 correlaciones para cada filtro, haciendo un total de 243,360 correlaciones diferentes, para la identificación de las 19 especies de diatomeas fósiles.



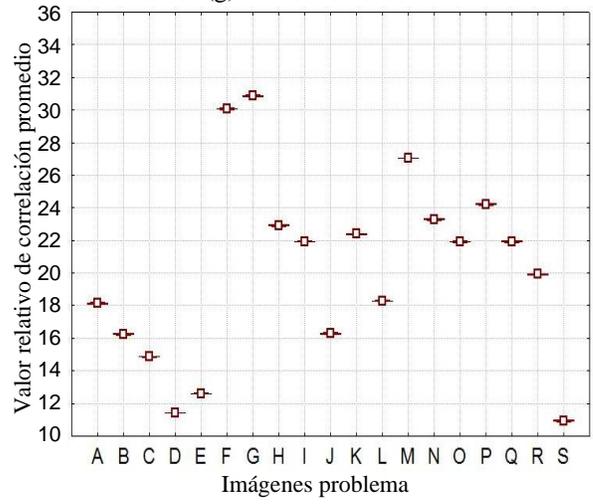
(e) Filtro diatomea E



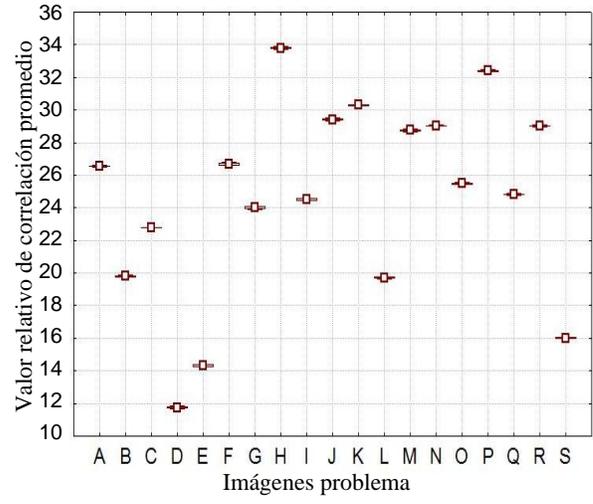
(f) Filtro diatomea F



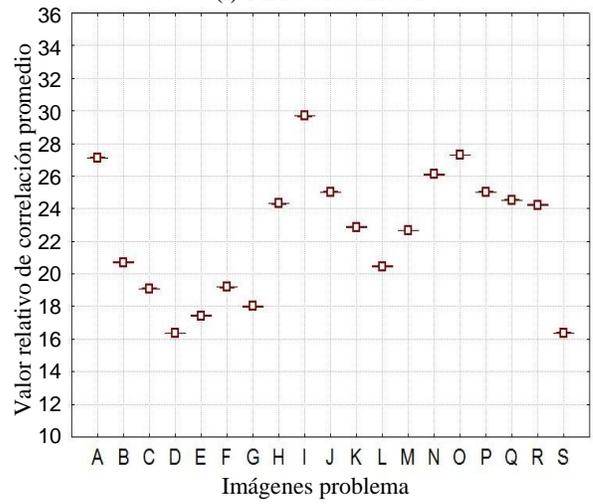
(g) Filtro diatomea G



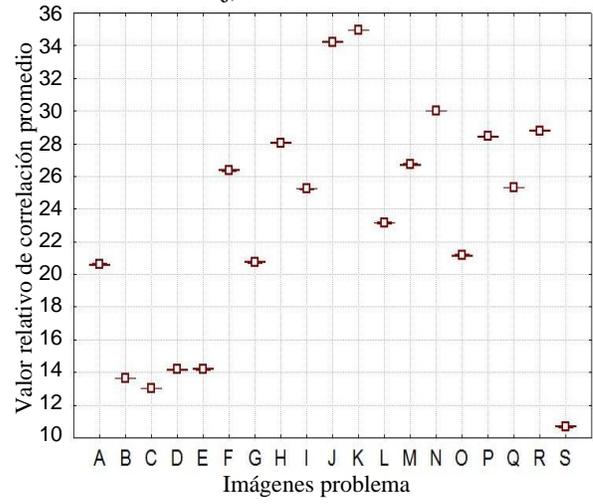
(h) Filtro diatomea H

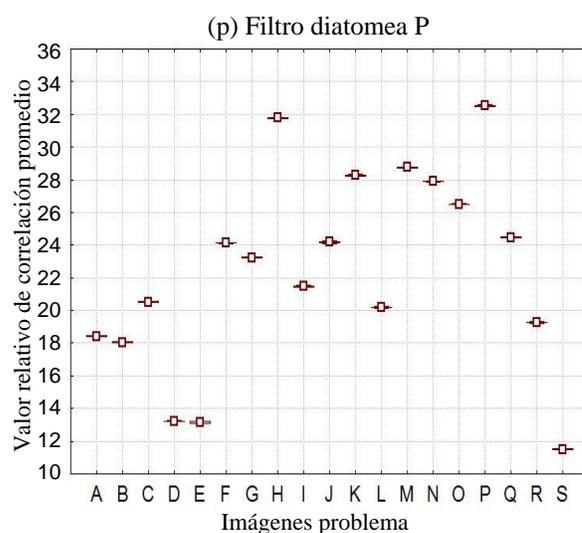
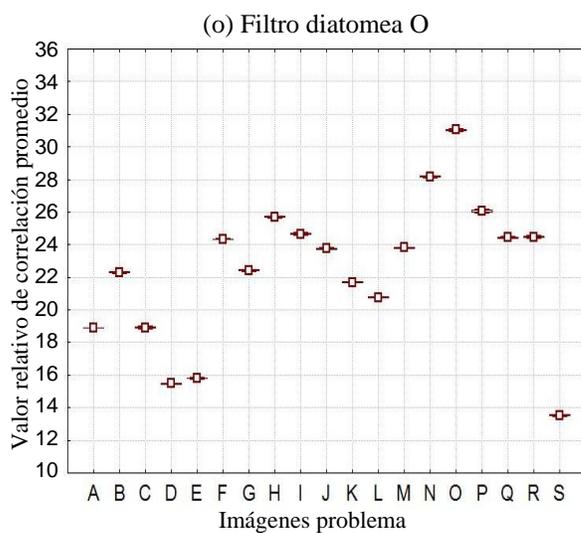
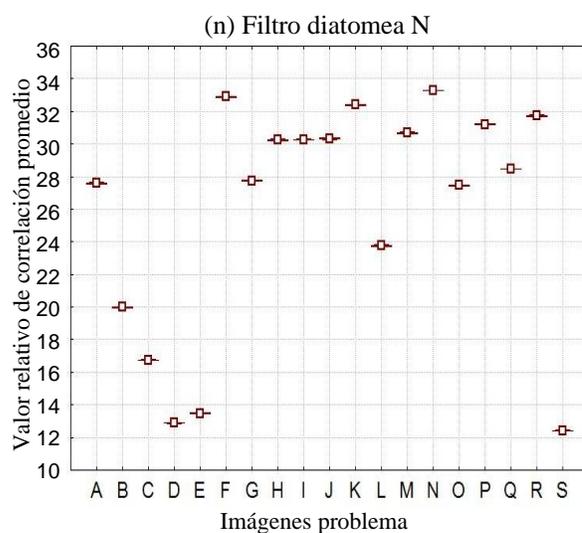
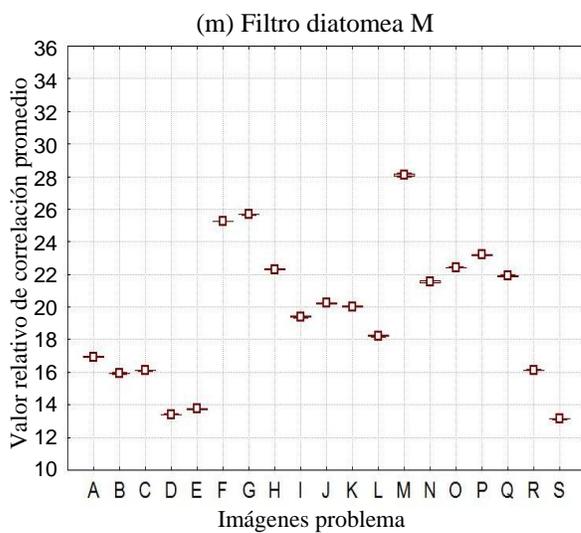
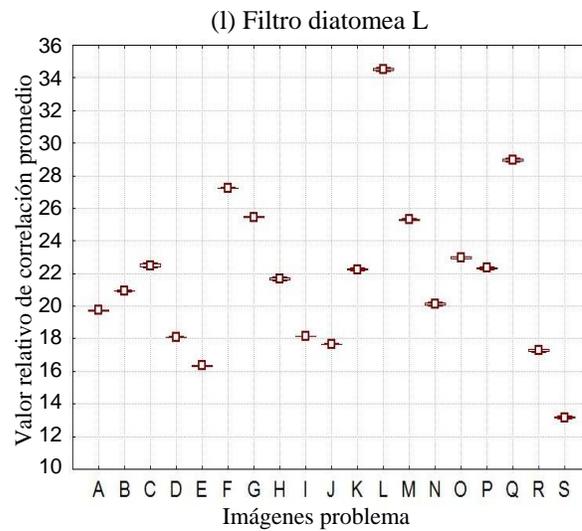
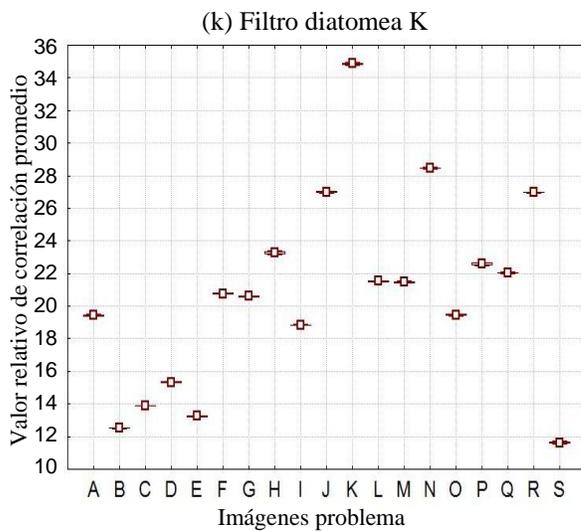


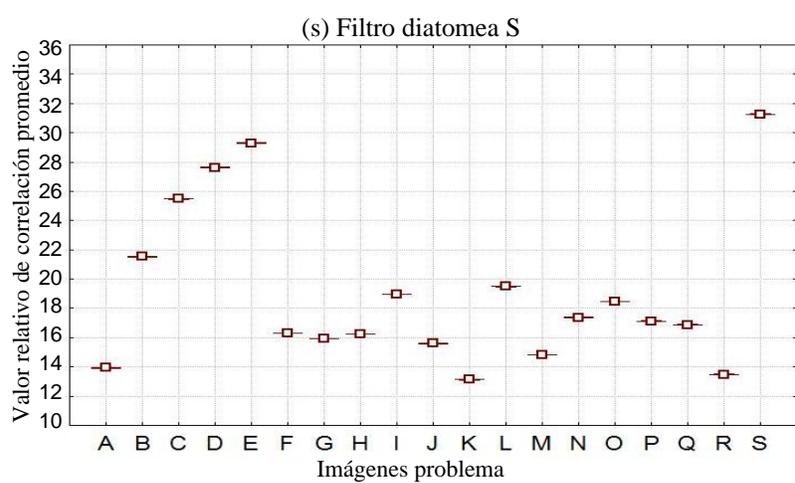
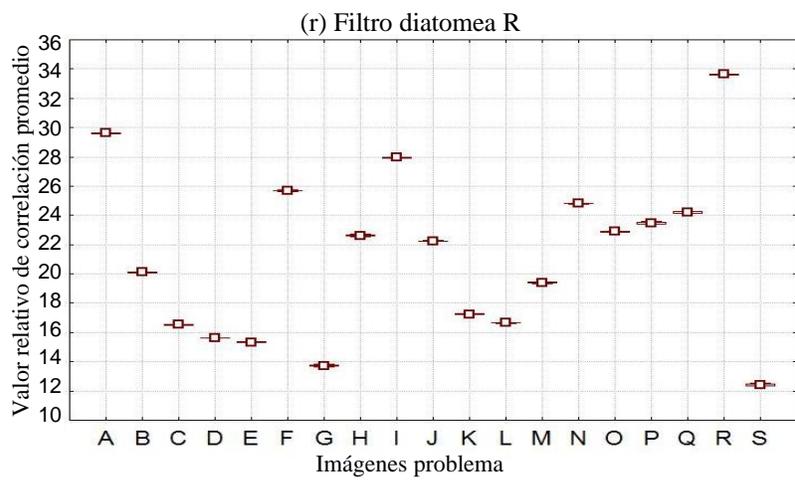
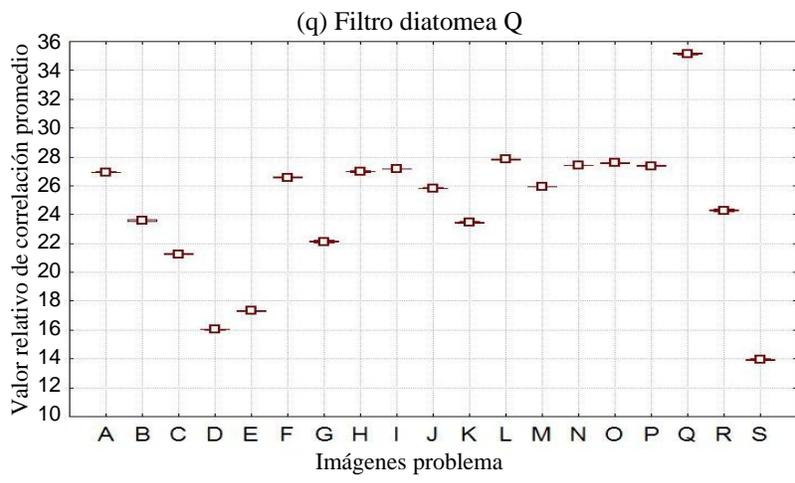
(i) Filtro diatomea I



(j) Filtro diatomea J







□ Media □ $\pm EE$ ▤ $\pm 2*EE$

Figura 26. Gráficas de cajas del valor relativo de correlación de las diatomeas desde la (A) a la (S), con respecto al resto, en rotación desde 0 hasta 359° obtenidas utilizando la metodología adaptativa, con incrementos de 1°, donde las cajas representan un error estándar ($\pm EE$) y las barras de las cajas representan dos errores estándar ($\pm 2*EE$) para el pico de correlación.

Para verificar con más detalle los resultados de la figura 26, se realiza un acercamiento a la parte superior de la escala de valores relativos de correlación promedio para poder visualizar con mayor claridad que existe un traslape entre correlaciones correspondientes a imágenes de diferentes diatomeas, asegurando así un nivel de confianza del 100%. Este acercamiento se hace sólo en las gráficas en donde existen valores de correlación promedio cercanos a los valores de correlación promedio resultantes para la imagen objetivo como es el caso de las diatomeas G, J, N y P. En la figura 27 se muestran las gráficas con su respectivo acercamiento y se puede apreciar que no existe traslape entre los diferentes valores de correlación.

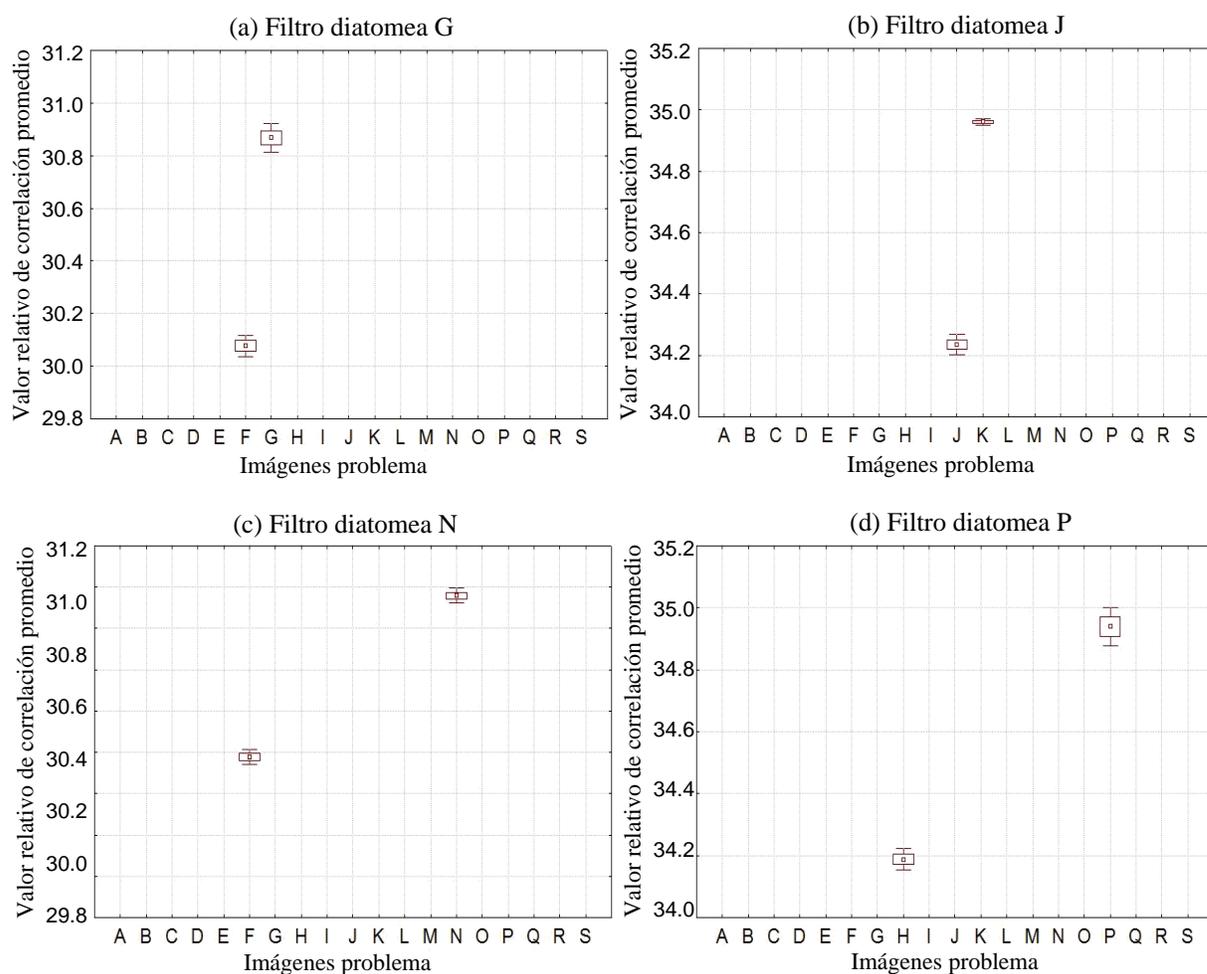


Figura 27. Acercamiento a la escala correspondiente al valor relativo de correlación (Metodología adaptativa) para: (a) Filtro diatomea G, (b) Filtro diatomea J, (c) Filtro diatomea N y (d) Filtro diatomea P.

III.3 Valor óptimo del valor de no linealidad k

Se realizó un experimento estadístico numérico para localizar el valor óptimo de k para el filtro no lineal. Los resultados se muestran usando una gráfica de cajas, del pico de correlación energía (PCE) contra el valor de no linealidad k. Se eligió de manera aleatoria la imagen binaria correspondiente a la letra B Arial como filtro. Este experimento se basa en realizar 360 correlaciones entre la letra B como imagen objetivo y las 360 letras B rotadas desde 0 hasta 359° como imágenes problema. Entonces, utilizando la correlación no lineal C_{NL2} de la Ec. (45) se asignan los valores de $k= 0.10, 0.20, \dots, 0.90$ y se realizaron 360 correlaciones por cada valor de k. De cada una de estas correlaciones se obtiene un valor de PCE utilizando la Ec. (32). El valor de la media de PCE de estos 360 valores por cada valor de k con un error estándar ($\pm EE$) y con dos errores estándar ($\pm 2*EE$), fueron graficados.

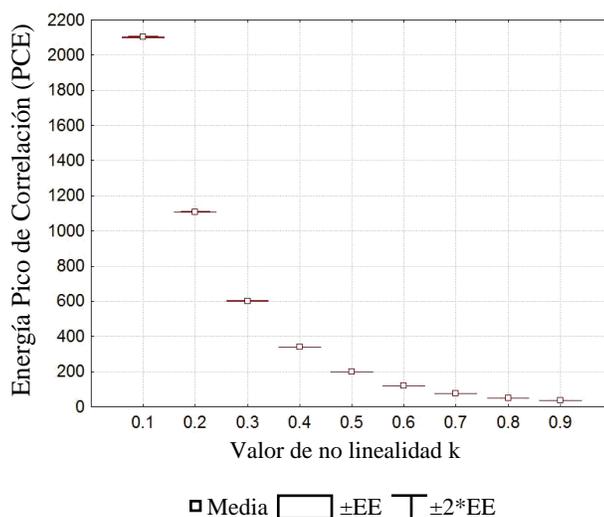


Figura 28. Gráfica de cajas para el pico de correlación energía PCE, contra los valores de no linealidad k, utilizando la metodología adaptativa y la letra B Arial como filtro.

En la figura 28 se observaron valores de PCE más altos, para valores de $k=0.1$, dado que nuestro índice invariante a rotación R_z , debe multiplicarse por el valor de k y dado que $0 < R_z < 1$, se ha elegido como valor óptimo $k=0.1$. Es de esta manera que el producto del índice invariante a rotación con el valor de no linealidad k, tuvo valores mayores a cero y menores a 0.10 y de esta manera se lograron resultados de correlación más limpios y con mejor definición.

III.4 Reconocimiento de imágenes inmersas en ruido

El comportamiento de la metodología adaptativa, para imágenes inmersas en ruido fue medido con el coeficiente de discriminación (CD). Se usaron 36 experimentos numéricos de una imagen afectada con el mismo nivel de ruido a un 95% de nivel de confianza, para un filtro no lineal con $k=0.10$, cuando se considera el ruido gaussiano aditivo con media cero y el ruido impulsivo o sal y pimienta.

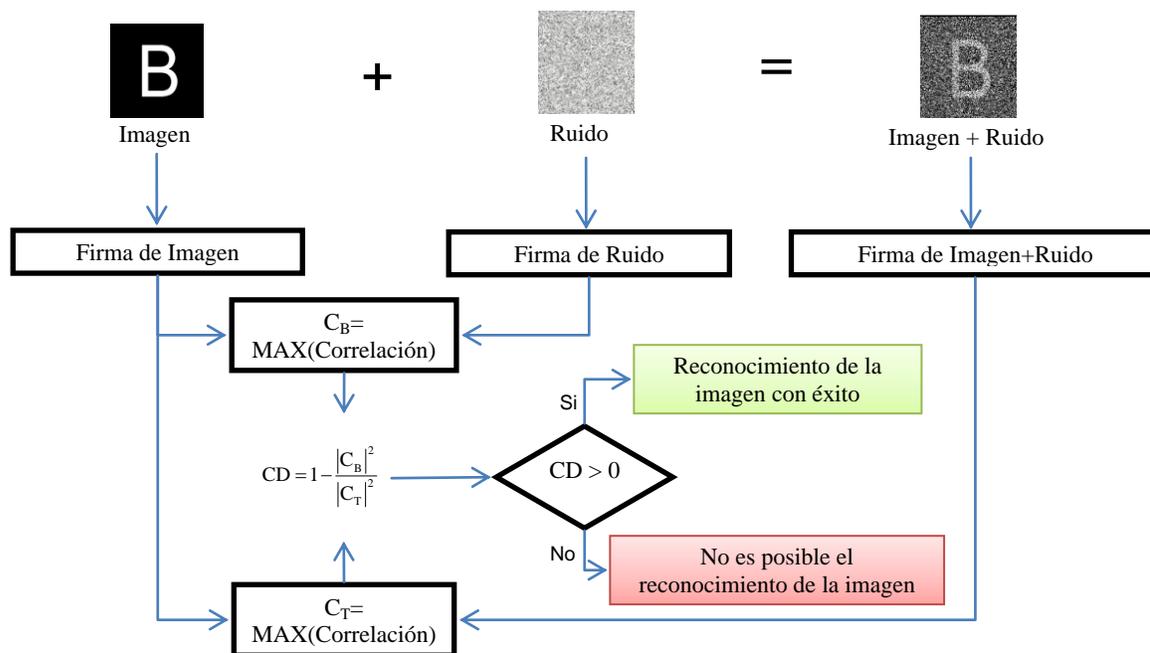


Figura 29. Proceso para obtener un valor de coeficiente de discriminación (CD).

En este experimento se utilizó la imagen binaria de la letra B Arial como filtro. En la figura 30 se muestran imágenes de la letra B Arial afectada por un ruido sal y pimienta con diferentes valores de densidad.

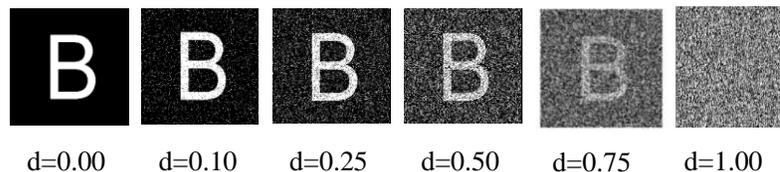


Figura 30. Letra B Arial afectada por ruido sal y pimienta.

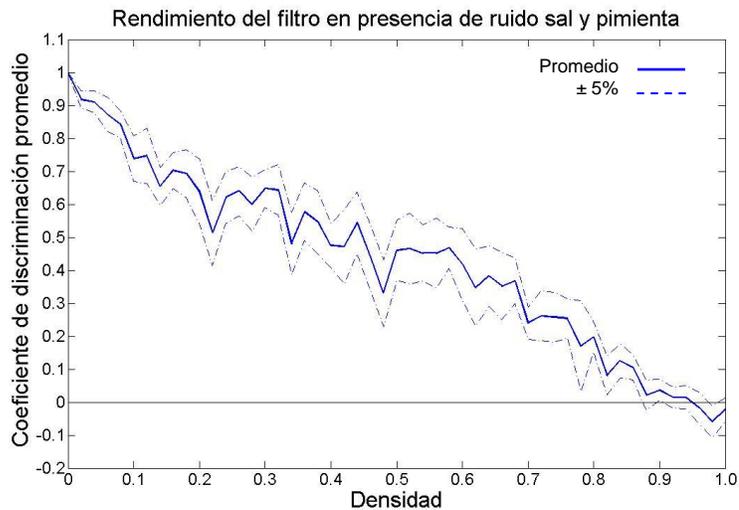


Figura 31. Rendimiento del filtro en presencia de ruido sal y pimienta.

De la figura 31, se observa que el filtro puede reconocer al objeto desde una densidad cero hasta una densidad de ruido sal y pimienta de aproximadamente 0.87.

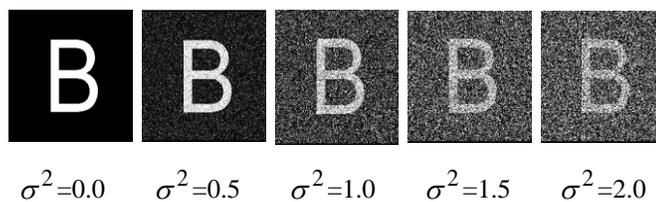


Figura 32. Letra B Arial afectada por gaussiano aditivo de media cero.

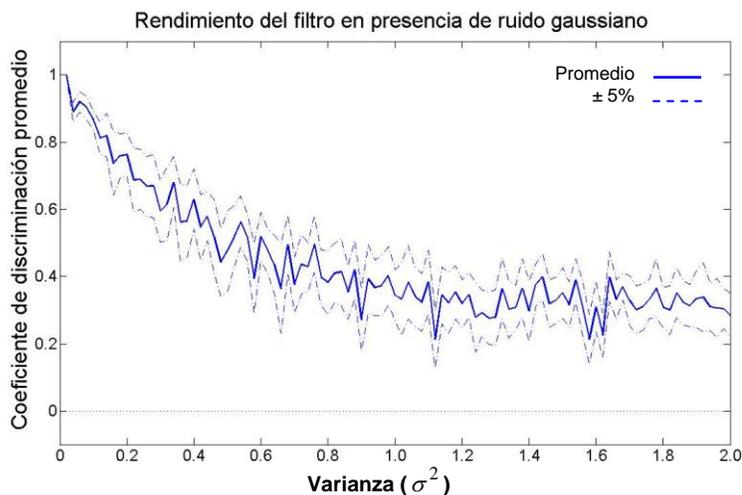


Figura 33. Rendimiento del filtro en presencia de ruido gaussiano de media cero.

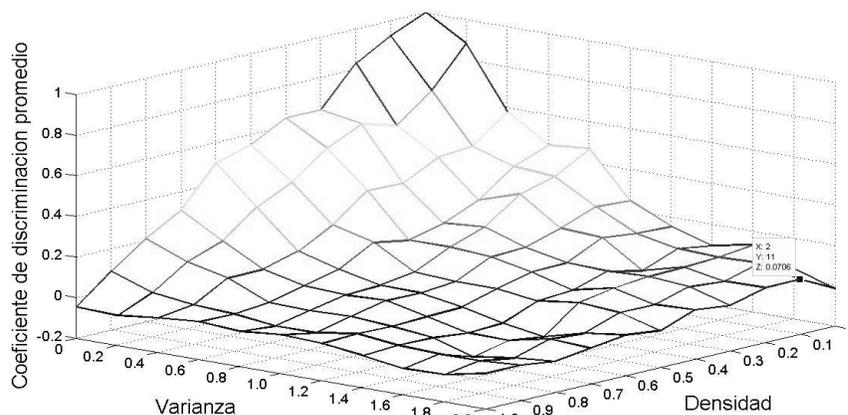


Figura 34. Coeficiente de discriminación del filtro en presencia de ruido gaussiano y ruido impulsivo.

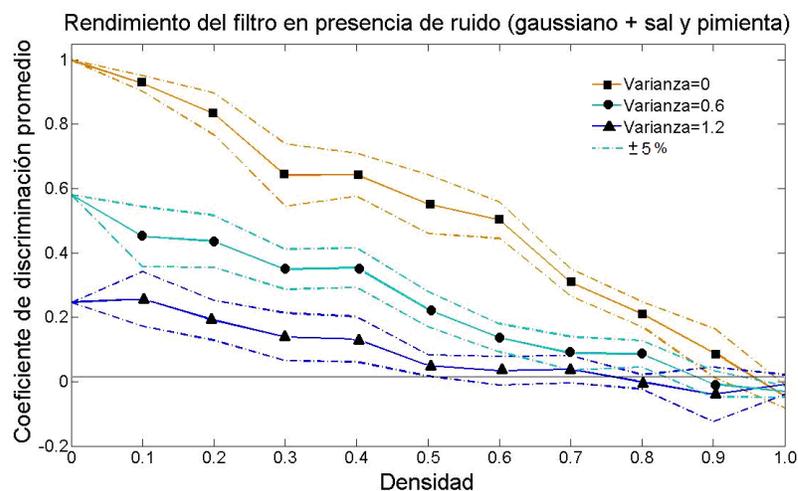


Figura 35. Coeficiente de discriminación del filtro en presencia de ruido gaussiano ($\sigma^2=0$, $\sigma^2=0.6$ y $\sigma^2=1.2$) y ruido impulsivo.

En la figura 33 se observa que, aun cuando se agrega ruido gaussiano a la imagen con varianza de 2, el filtro tiene la capacidad de identificar la imagen sin problemas ya que el valor CD es mayor a cero. La figura 34 muestra un esquema general del comportamiento del filtro en presencia de la combinación de ruido impulsivo y gaussiano. La figura 35 muestra secciones específicas de la figura 34. En estas secciones la combinación de ruido gaussiano con varianza (σ^2)= 0.6 y ruido impulsivo, muestran que la efectividad del filtro deja de ser 95% confiable cuando la densidad se incrementa hasta 0.7. En cambio para una varianza (σ^2)= 1.2, el filtro deja de ser confiable cuando la densidad de ruido impulsivo llega a una densidad de 0.5.

Capítulo IV. Discusión

Las técnicas actuales, a pesar de presentar invariancia a rotación, no ofrecen niveles de confianza al 100% para todos los casos. Por lo que se propone el desarrollo de una nueva metodología para mejorar la invariancia a rotación utilizando técnicas adaptativas.

Se han presentado dos metodologías que como método de comparación utilizan una correlación no lineal en combinación con máscaras adaptativas binarias y logran la identificación de imágenes con invariancia a rotación aun cuando estas se encuentran inmersas en ruido, cumpliendo así, con los objetivos planteados.

La primera metodología desarrollada (metodología iterativa) pretende autogenerar todas las distorsiones adquiridas en la parte real del patrón de difracción cuando la imagen presenta rotación. Entonces el diseño de la máscara binaria se inicia rotando las imágenes de entrada intencionalmente desde 0 hasta 359°. Las operaciones de rotación, transformada rápida de Fourier en dos dimensiones y búsqueda del mejor transecto se repiten 180 veces para cada imagen de entrada, lo cual consume mucho tiempo de cómputo. Realizar sólo una correlación llega a demorar hasta 39 segundos. Debido a esto fue necesario pensar en una modificación que agilizará el algoritmo y fue entonces que utilizando las mismas herramientas, pero eliminando las iteraciones, se diseña una segunda metodología (metodología adaptativa).

La metodología adaptativa, además de utilizar las máscaras binarias de anillos concéntricos adaptativas, utiliza también una correlación no lineal adaptativa, misma que se obtuvo al agregar un índice invariante a rotación en la correlación no lineal original (Ec. 43). Esta nueva metodología, elimina las iteraciones y la búsqueda del mejor transecto, logrando disminuir su tiempo de cómputo a aproximadamente 0.19 segundos por correlación. Además, al agregar el índice invariante a rotación, se cuenta con una referencia adicional que incrementa la capacidad de discriminación tanto para imágenes binarias como para imágenes reales en escala de grises.

Este proyecto de tesis se limitó a lograr la invariancia a rotación y posición, pero es posible agregar la invariancia a escalamiento con el uso de filtros compuestos o implementando la transformada de escala. Como trabajo futuro es posible modificar la nueva correlación no lineal (Ec. 45), que tiene un valor de no linealidad k dependiente de un índice invariante a rotación (Ec. 44), para implementarla con un índice invariante a escalamiento, o con cualquier otro índice con invariancia a alguna otra alteración que presente la imagen problema. Otro futuro trabajo utilizando las mismas bases es desarrollar metodologías para aplicarlas a la identificación de objetos parcialmente ocultos.

Conclusiones

Se desarrolló un sistema de correlación digital que cumple con los objetivos de la tesis y que es capaz de identificar imágenes contaminadas con altos niveles de ruido gaussiano e impulsivo. Se realizaron experimentos en donde se utiliza la métrica de capacidad de discriminación para demostrar que imágenes contaminadas con ruido gaussiano de media cero y niveles de varianza hasta de $\sigma^2=2$, pueden ser identificadas sin problemas con un nivel de confianza del 95%. Para imágenes contaminadas con ruido impulsivo se logra la identificación de la imagen, aun cuando el 87% de los pixeles que la componen se encuentran afectados. Estos experimentos también demostraron que, aun cuando las imágenes son afectadas con una combinación de ambos ruidos, el sistema es capaz de lograr la identificación de las imágenes.

Al implementar la nueva correlación no lineal con un valor de no linealidad k dependiente de la razón entre los promedios de la imagen objetivo y la imagen problema, se logró hacer más sencillo el algoritmo debido a que se elimina la dependencia de la imagen problema para generar la máscara binaria de anillos concéntricos.

La metodología adaptativa, logró mejorar el tiempo de cómputo de la metodología iterativa de aproximadamente 39 segundos a sólo 0.19 segundos por correlación; gracias al corto tiempo de ejecución de la metodología adaptativa fue posible comprobar y presentar su efectividad, verificando la discriminación que se genera al utilizar como filtro cada una de las 26 letras del abecedario, en el caso de las imágenes binarias y al utilizar como filtro cada una de las 19 diatomeas fósiles, en el caso de imágenes en escala de grises. Además, la identificación de imágenes a alta velocidad nos da la opción de poder aplicar este sistema a aplicaciones en tiempo real.

Se alcanzó un nivel de confianza del 100% en todos los experimentos realizados, mejorando los resultados en cuanto invariancia a rotación publicados en la literatura reciente.

Referencias bibliográficas

Asselin, D., Arsenault, H. H. and Prévost D. (1994). Optical circular sampling system for translation and rotation invariant pattern recognition. *Optical Communications*, 110, 507-513.

Cañada, A. (2006). Fourier y sus coeficientes. *Boletín de la Sociedad Española de Matemática Aplicada*, 36, 125-148.

Chaparro Magallanez, G., Álvarez Borrego, J. and Solorza Calderón S. (2011). Invariant correlation to position, rotation and scale using one-dimensional composite filters. *Proceedings of SPIE*, 8011, 801171

Cooley, J. W. and Tukey, J. W. (1965). An algorithm for machine calculation of complex Fourier series. *Mathematics of Computation*, 19(90), 297-301.

Coronel Beltrán, A. and Álvarez Borrego, J. (2010). Comparative analysis between different font types and styles letters using a nonlinear invariant digital correlation. *Journal of Modern Optics*, 57(1), 58-64.

Fares, A., Bouzid, A. and Hamdi M. (2000). Rotation invariance using diffraction pattern sampling in optical pattern recognition. *Journal of Microwaves and Optoelectronics*, 2(2), 33-39.

Fimbres Castro, C., Álvarez Borrego, J. and Bueno Ibarra, M. (2012). Invariant nonlinear correlation and spectral index for diatoms recognition. *Optical Engineering*, 51(4), 047201, 1-4.

George, N., Thomasson, J. T. and Spendel A. (1972). Photodetector light pattern detector. *US patent*, 3,689,772.

González, R. C. and Woods, R. E. (2002). *Digital image processing* (2ND Ed.). Upper Saddle River, NJ, Prentice Hall.

Horner, J. L. and Gianino, P. D. (1984). Phase-only matched filtering. *Applied Optics*, 23, 812-816.

Hsu, H.P. (1984). *Applied Fourier analysis*. Orlando, FL, Harcourt Brace Jovanovich.

Hsu, Y. N., Arsenault, H. H. and April G. (1982). Rotation-invariant digital pattern recognition using circular harmonic expansion. *Applied Optics*, 21, 4012-4015.

Javidi, B. (2002). *Image recognition and classification algorithms, systems, and applications*. New York, NY, Marcel Dekker.

- Kumar, B. V. and Hassebrook, L. (1990). Performance measures of correlation filters. *Applied Optics*, 29, 2997-3006.
- Lendaris, G. G. and Stanley, G. L. (1965). A optical logical self-organizing recognition system, in *Optical and Electro-Optical information Processing*. Cambridge, M.I.T. Press.
- Lendaris, G. G. and Stanley, G. L. (1970). Diffraction-pattern sampling for automatic pattern recognition. *Proceedings of the IEEE*, 58(2), 198-216.
- Lerma Aragón, J. and Álvarez Borrego, J. (2009). Vectorial signatures for invariant recognition of position, rotation and scale pattern recognition. *Journal of Modern Optics*, 56, 1598-1606.
- Mak, C., Elsner, M., DeFrancesco, L. and Hare P. (2010). Research highlights. *Nature Biotechnology*, 28, 45-46.
- O'Toole, R. K., and Stark, H. (1980). Comparative study of optical-digital vs. all-digital techniques in textural pattern recognition. *Applied Optics*, 19(15), 2496-2506.
- Pech Pacheco, J., Álvarez Borrego, J., Cristóbal, G. and Keil, M. (2003). Automatic object identification irrespective to geometric changes. *Optical Engineering*, 42(2), 551-559.
- Rosen, C. A., (1960). *Graphical Data Processing Research Study and Experimental Investigations*. Menlo Park, CA, Standford Research Institute.
- Solorza Calderón, S. and Álvarez Borrego, J. (2010). Digital system of invariant correlation to position and rotation. *Optical Communications*, 283(19), 3613-3630.
- Solorza Calderón, S. and Álvarez Borrego, J. (2011). Digital system of invariant correlation to position and scale using adaptive ring masks and unidimensional signatures. *Proceedings of SPIE*, 8011 (801171), 1-7.
- Solorza Calderón, S., Álvarez Borrego, J. and Chaparro Magallanez G. (2012). Pattern Recognition of Digital Images by One-Dimensional Signatures. In: Salih Mohammed, *Fourier Transform - Signal Processing*. Rijeka, Croatia, InTech.
- Vander Lugt, A. (1964). Signal detection by complex spatial filtering. *IEEE Transactions on Information Theory*, 10, 139-145.
- Yaroslavsky L. P. (1993). The theory of optical method for localization of objects in pictures. *Progress in Optics*, 32, 145-201.
- Zavala Hamz, V. A. and Álvarez Borrego, J. (1997). CH filters for the recognition of marine microorganisms. *Applied Optics*, 36(2), 484-489.