

**CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE EDUCACIÓN
SUPERIOR DE ENSENADA, BAJA CALIFORNIA**



**PROGRAMA DE POSGRADO EN CIENCIAS
EN ÓPTICA CON ORIENTACIÓN EN ÓPTICA FÍSICA**

**Estudio y análisis comparativo de diferentes sistemas
invariantes de reconocimiento de objetos**

Tesis

para cubrir parcialmente los requisitos necesarios para obtener el grado de

Maestro en Ciencias

Presenta:

Esbanyely Garza Flores

Ensenada, Baja California, México
2014

Tesis defendida por
Esbanyely Garza Flores
y aprobada por el siguiente Comité

Dr. Josué Álvarez Borrego
Director del Comité

Dra. Ernestina Castro Longoria
Miembro del Comité

Dr. Héctor Manuel Escamilla Taylor
Miembro del Comité

Dra. Karina Garay Palmett
Miembro del Comité

Dra. Selene Solorza Calderón
Miembro del Comité

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

Dr. Jesús Favela Vara
Director de Estudios de Posgrado

Noviembre, 2014

Resumen de la tesis de Esbanyely Garza Flores, presentada como requisito parcial para la obtención del grado de Maestro en Ciencias en Óptica con orientación en Óptica física.

Estudio y análisis comparativo de diferentes sistemas invariantes de reconocimiento de objetos

Resumen elaborado por:

Esbanyely Garza Flores

En este trabajo se implementaron tres diferentes metodologías de reconocimiento de patrones invariantes a escala posición y rotación basadas en correlación: correlación no lineal con transformada de escala, firmas vectoriales y máscaras binarias. También se implementó la metodología de índices espectrales, la cual no utiliza correlación. Se realizó un análisis comparativo de las metodologías con imágenes afectadas con ruido gaussiano e impulsivo para determinar su capacidad de reconocimiento ante la presencia de ruido. Se determinó cuál de las metodologías produce un plano de salida menos ruidoso analizando las imágenes de 21 especies diferentes de diatomeas fósiles que fueron rotadas de 0 a 179 grados en pasos de 1 grado y de 80 a 128% en variaciones de escala en gradientes de 1%. Se realizó un análisis con fragmentos de las imágenes de diatomeas fósiles para observar cuánto porcentaje de información requiere una metodología para realizar una identificación. Se calculó el coste computacional para cada metodología.

Se propuso una nueva metodología basada en firmas generadas con máscaras binarias en la que no se requiere ninguna correlación y los resultados de todas las imágenes procesadas se muestran en un único plano de identificación. Esta nueva propuesta toma menos de 2 minutos para procesar 3780 imágenes, mientras que para la metodología del índice espectral que tampoco utiliza la correlación, toma 2.25 minutos.

Palabras clave: índices espectrales, correlación no lineal, transformada de escala, firmas vectoriales, máscaras binarias, ruido, diatomeas.

Abstract of the thesis presented by Esbanyely Garza Flores as a partial requirement to obtain the Master in Science degree in Optics with orientation in Optical physics.

Study and comparative analysis of different invariant object recognition systems

Abstract by:

Esbanyely Garza Flores

In this work different image pattern recognition methodologies invariant to position scale and rotation based in correlation were implemented: nonlinear correlation with scale transform, vectorial signatures and binary masks. The spectral index methodology is also implemented, which doesn't use correlation. A comparative analysis of these methodologies was made when images affected with Gaussian and impulsive noise are compared to determine their recognition capability under these conditions. The correlation output plane of the methodologies was determined analyzing images of 21 different species of fossil diatoms which were rotated from 0 to 179 degrees in steps of 1 degree and 80 to 128% in scale variations in gradients of 1%. Also an analysis with fragments of the fossil diatoms images was made to see how much percentage of information a methodology requires to make a positive identification. The computational cost in terms of the time of execution for each methodology was calculated.

A new methodology based on the vector signatures generated by binary masks was proposed in which no correlation is required, and the results of the processed images are shown in a identification plane. This new proposal took less than 2 minutes to process 3780 images while for the spectral index methodology which also doesn't use correlation, took 2.25 minutes.

Keywords: spectral index, nonlinear correlation, scale transform, vectorial signatures, noise, binary masks, diatoms.

Dedicatoria

A mi madre y hermanos que siempre me han apoyado.

Agradecimientos

Al Dr. Pedro Negrete y miembros del CPP que me dieron la oportunidad de poder cursar mi Maestría en el Departamento de Óptica de CICESE.

A la Dra. Nahara Ayala por su ayuda.

Al Dr. Josué Álvarez Borrego, director de tesis, por su apoyo, paciencia y disposición para orientarme en el desarrollo de mi trabajo.

A los miembros del comité de tesis, la Dra. Ernestina Casto Longoria, Dr. Héctor Manuel Escamilla Taylor, Dra. Karina Garay Palmett y Dra. Selene Solorza Calderón, por su ayuda, tiempo y consejos para lograr el término de este trabajo de maestría.

Al CONACYT por su apoyo económico al otorgarme la beca para realizar mis estudios de maestría, así como también su financiamiento del proyecto 169174: *“Reconocimiento de patrones en imágenes digitales mediante máscaras binarias de anillos concéntricos y firmas unidimensionales invariantes a posición, rotación y escalamiento”*. Ciencia Básica SEP-CONACYT 2011.

Tabla de contenido

	Página
Resumen español	ii
Resumen inglés	iii
Dedicatorias	iv
Agradecimientos	v
Lista de Figuras	viii
Lista de Tablas	xii
Capítulo 1. Introducción	1
1.1 Antecedentes.....	2
1.2 Motivación.....	4
1.3 Objetivo General.....	4
1.4 Objetivos particulares.....	4
1.5 Organización de la tesis.....	5
Capítulo 2. Fundamentos	6
2.1 Imágenes digitales.....	6
2.2 Transformada de Fourier.....	6
2.2.1 Transformada de Fourier para una función unidimensional.....	7
2.2.2 Transformada de Fourier bidimensional.....	7
2.3 Propiedades de la Transformada de Fourier.....	8
2.3.1 Linealidad.....	9
2.3.2 Translación en el dominio del espacio.....	9
2.3.3 Translación en el dominio de la frecuencia.....	9
2.3.4 Teorema de escalamiento.....	9
2.3.5 Simetría.....	10
2.4 Teorema de Parseval.....	10
2.5 Teorema de muestreo.....	10
2.6 Transformada discreta de Fourier.....	11
2.6.1 Transformada discreta de Fourier para una función unidimensional.....	12
2.6.2 Transformada discreta de Fourier para una función bidimensional.....	12
2.7 Convolución.....	13
2.8.1 En una dimensión.....	13
2.8.2 En dos dimensiones.....	13
2.8 Teorema de Convolución.....	14
2.9 Correlación.....	14
2.9.1 En una dimensión.....	15
2.9.2 En dos dimensiones.....	15
2.10 Teorema de Correlación.....	16
2.11 Transformada de Mellin.....	16
2.11.1 Invariancia a escala de la transformada de Mellin.....	17
2.11.2 La transformada de Mellin vía la transformada de Fourier.....	18
2.11.3 Transformada de Mellin bidimensional.....	18

Tabla de contenido (continuación)

2.12	Transformada de escala.....	19
2.12.1	La transformada de escala vía la transformada de Fourier.....	20
2.12.2	La transformada de escala bidimensional.....	21
2.13	Filtros de correlación.....	22
2.13.1	Filtro clásico acoplado CMF	22
2.13.2	Filtro sólo de fase POF	23
2.13.3	Filtro no lineal NLF	23
Capítulo 3.	Metodología	24
3.1	Correlación no lineal con transformada de escala invariante a posición, rotación y escala.....	24
3.1.1	Obtención del filtro	24
3.1.2	Correlación no lineal.....	28
3.2	Índices espectrales	28
3.2.1	Índice 1.....	28
3.2.2	Índice 2.....	29
3.3	Correlación invariante a posición, rotación y escala, con firmas vectoriales.....	32
3.4	Correlación invariante a posición, y rotación, con máscaras binarias	34
3.4.1	Máscaras adaptativas.....	42
3.4.2	Utilización del perfil óptimo.....	42
3.5	Comparación de los diferentes métodos.....	47
3.5.1	Evaluación de PCE.....	47
3.5.2	Capacidad para reconocer imágenes afectadas por ruido.	49
3.5.2.1	Imágenes afectadas por ruido gaussiano.....	50
3.5.2.2	Imágenes afectadas por ruido impulsivo.....	51
3.5.3	Capacidad de reconocimiento con fragmentos de información.....	53
3.5.4	Tiempo de ejecución.....	56
Capítulo 4.	Resultados	58
4.1	Resultados de PCE.....	58
4.2	Índices espectrales.....	63
4.3	Tiempo de ejecución.....	65
4.4	Resultados de la capacidad de reconocimiento con fragmentos de información.....	67
4.5	Resultados de la capacidad para reconocer imágenes afectadas por ruido.....	73
Capítulo 5.	Nueva propuesta metodológica	83
Capítulo 6.	Conclusiones	90
	Lista de referencias	92

Lista de figuras

Figura		Página
1	Diagrama de bloques para obtener el filtro no lineal.....	25
2	Imagen blanco y negro de la letra E tipo arial.....	25
3	Módulo de la transformada de Fourier de la letra E tipo arial.....	26
4	Módulo de la transformada de Fourier después de haber sido aplicado el filtro parabólico.....	26
5	Imagen del espectro de frecuencias en coordenadas polares.....	27
6	Imagen del espectro de frecuencias en coordenadas polares después de aplicar $\rho = \ln r$	27
7	Correlación digital mediante filtros no lineales.....	28
8	Procedimiento para obtener el índice 1.	29
9	Procedimiento para la obtención del índice 2	30
10	Obtención de subimagen del espectro de frecuencias.....	30
11	Escalamiento al tamaño de la imagen original.....	31
12	Imágenes de prueba de la metodología de índices espectrales.....	32
13	Resultado de la metodología de índices espectrales donde cada una de las imágenes mostradas en la figura 12 tiene su respectivo espacio dentro del plano de intificación.....	32
14	Diagrama a bloques del método de obtención de firmas vectoriales.....	33
15	Parte real (a) e imaginaria (b) de la imagen de la letra E de la figura 2.....	36
16	a) Perfil de la parte real de $F(u,v)$. b) Perfil de la parte imaginaria de $F(u,v)$	37
17	a) Perfil binario de la parte real de la figura 10a usando un umbral positivo. b) Perfil binario de la parte imaginaria 10b usando un umbral positivo.....	39
18	Cilindros concéntricos producidos al rotar los perfiles binarios, parte real (a), parte imaginaria (b).....	40
19	a) MBRP . b) MBIP.....	40
20	a) MBRNP. b) MBINP	41
21	Procedimiento para obtener la firma de la imagen, donde \otimes significa multiplicación punto a punto.....	42
22	Máscara de disco binaria de 256X256.....	43
23	a) Filtrado de la parte real de la transformada de Fourier por un disco D. b) Filtrado de la parte imaginaria de Fourier por un disco D.....	44
24	Metodología de máscaras binarias.....	45

Lista de figuras (continuación)

Figura		Página
25	Metodología de máscaras binarias adaptativas.....	46
26	Imágenes de 21 especies diferentes de diatomeas fosilizadas.....	48
27	Imagen de la especie de diatomea <i>Actinocyclus Ingens-Rattray</i> inmersa en ruido gaussiano con $\mu = 0$: (a) $\sigma^2 = 0$, (b) $\sigma^2 = 0.1$ y (c) $\sigma^2 = 0.5$	51
28	Imagen de la especie de diatomea <i>Actinocyclus Ingens-Rattray</i> inmersa en ruido sal y pimienta: (a) $d = 0$, (b) $d = 0.1$ y (c) $d = 0.5$	52
29	Imagen de especie de diatomea <i>Actinocyclus Ingens-Rattray</i> y sus 49 fragmentos.....	54
30	Imagen de especie de diatomea especie <i>Azpeitia sp</i> y sus 49 fragmentos.....	54
31	Resultados de correlación de la imagen entera de la especie A (color azul) con las 1050 imágenes problema y resultado de correlación del fragmento 27 de la especie A (color rojo) con las 1050 imágenes problema.....	56
32	Gráfica de cajas de los resultados de PCE para la diatomea A, donde se muestra la media así como ± 1 error estandar (SE) y ± 2 errores estandar ($2*SE$)	59
33	Gráfica de cajas de los resultados de PCE para la diatomea F, donde se muestra la media así como ± 1 error estandar (SE) y ± 2 errores estandar ($2*SE$).....	60
34	Gráfica de cajas de los resultados de PCE para la diatomea P, donde se muestra la media así como ± 1 error estandar (SE) y ± 2 errores estandar ($2*SE$).....	62
35	Resultado estadístico de índices espectrales para las 21 especies de diatomeas, donde se presenta la media de cada índice, así como $\pm 2SE$	64
36	Acercamiento a los resultados estadísticos de índices espectrales de diatomea R y diatomea K.....	65
37	Resultados de correlación de la imagen entera de la diatomea A (color azul) con las 1050 imágenes problema y resultado de correlación del fragmento 35 de la especie A (color rojo) con las 1050 imágenes problema.....	68
38	Gráfica de resultados de correlación de la imagen entera de la diatomea B (color azul) con las 1050 imágenes problema y resultado de correlación del fragmento 30 de la especie B (color rojo) con las 1050 imágenes problema.....	69
39	Gráfica de resultados de correlación de firmas vectoriales de escala de la imagen entera de la diatomea A con las 1050 imágenes problema.....	70

Lista de figuras (continuación)

Figura		Página
40	Gráfica de resultados de correlación de firmas vectoriales de rotación de la imagen entera de la diatomea A con las 1050 imágenes problema.....	71
41	Gráfica de la multiplicación de resultados de correlación de firmas vectoriales de escala y firmas de rotación de la imagen entera de la diatomea A problema.....	71
42	Gráfica del promedio de los resultados del coeficiente de discriminación en color negro y $\pm 1.96SE$ en rojo, con 30 iteraciones para CNLTE con ruido gaussiano.....	74
43	Gráfica del promedio de los resultados del coeficiente de discriminación en color negro y $\pm 1.96SE$ en rojo, con 30 iteraciones para CNLTE con ruido impulsivo.....	75
44	Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ para máscaras binarias con ruido gaussiano.....	76
45	Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ para la metodología de máscaras binarias utilizando el transecto óptimo con ruido impulsivo...	76
46	Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ para la metodología de máscaras binarias utilizando el transecto central con ruido gaussiano..	77
47	Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ para la metodología de máscaras binarias utilizando el transecto central con ruido impulsivo...	78
48	Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ para la metodología de firmas vectoriales con ruido gaussiano para $0 \leq \sigma^2 \leq 0.55$	79
49	Gráfica del promedio de los resultados del coeficiente de discriminación de las firmas de escala y firmas de rotación con ruido gaussiano.....	80
50	Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ para la metodología de firmas vectoriales con ruido impulsivo para $0 \leq \sigma^2 \leq 0.55$	80
51	Gráfica del promedio de los resultados del coeficiente de discriminación de las firmas de escala y firmas de rotación con ruido impulsivo.....	79
52	Gráfica comparativa entre firmas vectoriales, MBRNP_TO y MBIP_TO para ruido gaussiano.....	81
53	Gráfica comparativa entre firmas vectoriales, MBRNP_TO y MBIP_TO para ruido sal y pimienta.....	82
54	Diagrama a bloques de la metodología propuesta.....	84

Lista de figuras (continuación)

Figura		Página
55	Gráfica de la media $\mu \pm 2SE$ del área y varianza de las firmas utilizando MBRP.....	86
56	Acercamiento a los resultados de la media $\mu \pm 2SE$ del área y varianza de las firmas utilizando MBRP.....	86
57	Gráfica de la media $\mu \pm 2SE$ del área y varianza de las firmas utilizando MBRNP.....	87
58	Gráfica de la media $\mu \pm 2SE$ del área y varianza de las firmas utilizando MBIP.....	87
59	Gráfica de la media $\mu \pm 2SE$ del área y varianza de las firmas utilizando MBINP.....	88

Lista de tablas

Tabla		Página
1	Media de PCE para <i>Actinocyclus Ingens-Rattray</i>	60
2	Media de PCE para <i>Denticulopsis Praedimorpha - Barron ex Akiba</i>	61
3	Media de PCE para <i>Diplonei bombus - Cleve</i>	62
4	Tiempo de ejecución.....	66
5	Fragmentos <i>Actinocyclus Ingens-Rattray</i>	72
6	Fragmentos <i>Azpeitia sp</i>	73
7	Tiempos de ejecución de la metodología propuesta comparados con los de índices espectrales.....	89

Capítulo 1. Introducción

El reconocimiento de patrones es un área del procesamiento de imágenes que consiste en la descripción y clasificación de objetos presentes en una imagen, basándose en un conjunto previamente establecido de todos los posibles objetos a reconocer.

Los sistemas de reconocimiento de patrones son utilizados como apoyo de la captación, interpretación y análisis de datos experimentales en diversas áreas como la medicina, biología, taxonomía, física e industria, por mencionar sólo algunas.

La automatización de los procesos de reconocimiento de patrones es cada vez más importante, por ejemplo, en la industria los requerimientos de calidad hacen necesario encontrar un producto defectuoso dentro de millones, lo cual ha orillado a los fabricantes a buscar alternativas para la identificación de estos productos, pues no existe ser humano que pueda repetir la misma actividad millones de veces. Además, la identificación automática elimina el posible error humano al hacer un análisis visualmente, lo cual se hace más importante al trabajar en áreas como la medicina donde se busca la identificación de alguna enfermedad.

En el reconocimiento de patrones se emplean técnicas para medir la similitud entre dos funciones, por lo que la función de correlación es una herramienta idónea para tal acción. Con la invención del láser y el desarrollo del filtro espacial complejo u holográfico propuesto por Vander Lugt a principios de los años 60's, también llamado filtro clásico acoplado, los sistemas de correlación adquirieron un importante impacto en el desarrollo del reconocimiento de patrones en diferentes campos. Dado que la implementación del filtro Vander Lugt es realizada ópticamente, éste permite tener la gran ventaja de realizar el proceso de identificación a alta velocidad (velocidad luz), sin importar la cantidad de datos de entrada, sin embargo, al utilizarse elementos ópticos se tiene la desventaja de la complejidad del arreglo, costo elevado y sensibilidad a variaciones de posición, rotación y escala de las imágenes.

La correlación implementada de manera digital es una alternativa a la implementación óptica, ya que las operaciones realizadas sobre las imágenes son simplemente simulaciones digitales de su equivalente óptico.

El procesamiento digital tiene la ventaja de que resulta más barato y más simple que la implementación óptica, sin embargo, a pesar de los grandes avances en cuanto a velocidad de cómputo, el costo computacional siempre es algo que se debe tener en mente, especialmente si se maneja una gran cantidad de datos, ya que el tiempo de procesamiento aumenta significativamente cuando aumenta la cantidad de datos de entrada, es decir, el método digital es más lento que el método óptico.

Al igual que en el caso de la implementación óptica, la correlación implementada digitalmente también es sensible a variaciones de posición, rotación y escala. Sin embargo, se han desarrollado diferentes técnicas de filtraje encaminadas a superar estas desventajas, lo que ha permitido que la implementación de algoritmos basados en correlación se conviertan en una solución atractiva en el reconocimiento de patrones para establecer si un objeto se encuentra presente en una imagen.

1.1 Antecedentes

En las últimas décadas se han desarrollado diferentes técnicas de correlación que varían dependiendo en gran medida del problema presente.

Almeida et al. (1972) utilizaron la técnica de procesamiento óptico coherente en el reconocimiento y caracterización de especies de microalgas, específicamente diatomeas. Cairns et al. (1972) utilizaron la técnica de filtrado espacial óptico coherente en el reconocimiento de patrones de las diatomeas *Navicula sp.* y *Cyclotella sp.* con filtros espaciales complejos del tipo Vander Lugt construidos con un interferómetro de Rayleigh modificado. Almeida y Eu (1976) presentaron resultados en la identificación de diatomeas, aplicando la técnica de correlación óptica convencional con filtros espaciales acoplados u hologramas con la transformada de Fourier.

Almeida et al. (1978) realizaron análisis de algas en transparencias fotográficas de 35mm utilizando un procesador óptico híbrido, en el cual el procesamiento de la información se realiza ópticamente y digitalmente.

Javidi et al. (1997) utilizaron, filtros compuestos no lineales en el plano de Fourier para lograr el reconocimiento de patrones invariante a distorsiones. Coronel-Beltrán (1988) utilizó un sistema óptico coherente invariante a rotación y ubicación de organismos en una diapositiva, aplicado a las diatomeas. Pech-Pacheco y Álvarez-Borrego (1998), trabajaron en la identificación de fitoplancton mediante la implementación de un sistema óptico-digital. Villalobos-Flores et al. (2001) emplearon un método de correlación invariante a posición, rotación y escala para la identificación de 12 especies de diatomeas fósiles. Castro-Longoria et al. (2001) propusieron un algoritmo para identificar especies de copépodos (crustáceos maxilópodos de tamaño muy pequeño). Javidi, B. (2002), realizó un reporte sobre la factibilidad de implementar un sistema automatizado de reconocimiento de letreros de límite de velocidad en Estados Unidos de Norte America, como asistencia para los conductores. Pech-Pacheco et al. (2003), presenta un método de identificación de objetos basado en el uso de la correlación de fase en el dominio de la transformada de escala para la identificación automatizada de caracteres. Mouriño-Pérez et. al. (2006), utilizaron correlación digital para la detección del virus *Vibrio cholerae* O1. Coronel-Beltrán y Álvarez-Borrego (2009), realizaron un análisis de las letras de tipo fuente Times New Roman y Arial en estilo plano usando un sistema digital de correlación invariante a posición, rotación y escala con un filtro no lineal y un filtro solo de fase. Bueno-Ibarra et. al. (2010), trabajaron en el reconocimiento del virus de la mancha blanca mediante correlación no lineal. Coronel-Beltrán y Álvarez-Borrego (2010), realizaron un análisis comparativo entre letras de tipo fuente Times New Roman, Courier New y Arial utilizando correlación digital no lineal invariante a posición, rotación y escala. Fimbres-Castro et. al. (2012), trabajaron en el reconocimiento de diatomeas mediante correlación no lineal invariante a posición, rotación y escala e índices espectrales. Álvarez-Borrego et. al. (2013), desarrollaron un sistema de correlación invariante a posición y rotación usando máscaras binarias aplicadas a imágenes en niveles de gris.

1.2 Motivación

Debido a la implementación del reconocimiento de patrones en diversas áreas como la biología y medicina, así como las diferentes ventajas y desventajas de los diversos algoritmos de correlación, es de importancia realizar un análisis comparativo de diferentes metodologías. Entre éstas existen cuatro muy recientes que se mencionan en los objetivos particulares.

1.3. Objetivo general

Estudio y análisis de cuatro algoritmos de reconocimiento de patrones con invariancia a posición, rotación y escala, así como una propuesta de una metodología alterna de reconocimiento.

1.4 Objetivos particulares

- a) Implementación de correlación invariante a posición, escala, rotación utilizando índices espectrales (Fimbres-Castro *et. al*, 2012).
- b) Implementación de correlación invariante a posición, rotación y escala, utilizando filtros no lineales con transformada de escala (Coronel-Beltrán y Álvarez-Borrego, 2010).
- c) Implementación de correlación invariante a posición, rotación y escala con firmas vectoriales (Lerma-Aragón y Álvarez-Borrego, 2009).
- d) Implementación de correlación invariante a posición y rotación con máscaras binarias. (Álvarez-Borrego et al., 2013).
- e) Comparación de las metodologías empleadas.
- f) Implementación de una nueva propuesta de reconocimiento.

1.5 Organización de la tesis

Esta tesis está organizada, a partir de este capítulo introductorio, de la siguiente manera: en el capítulo 2, se establece una base teórica en la cual se sustenta el trabajo desarrollado en esta tesis. Primero se describe lo que es una imagen digital. Después, se menciona la transformada de Fourier para funciones continuas y discretas en una y dos dimensiones. Además, también se establecen sus transformadas inversas. Se presentan las principales propiedades de la transformada de Fourier. Posteriormente, se define la convolución y correlación para funciones continuas y discretas en una y dos dimensiones. Se define la transformada de Mellin y también la transformada de escala y la manera en que ésta puede ser implementada vía la transformada de Fourier. Por último, se mencionan los tipos de filtros de correlación.

El capítulo 3 trata sobre las metodologías implementadas en este trabajo. Primeramente se menciona la metodología de correlación no lineal con transformada de escala invariante a posición, rotación y escala. Posteriormente se describe la metodología de índices espectrales. Después se describen las metodologías de firmas vectoriales y máscaras binarias. Por último se describe la forma en que las metodologías serán evaluadas y comparadas.

En el capítulo 4 se presenta los resultados del análisis comparativo realizado de las metodologías implementadas, el cual consiste en determinar qué tan ruidoso es el plano de correlación mediante la métrica de PCE, la capacidad para reconocer imágenes afectadas por ruido mediante la métrica de DC, la capacidad de reconocimiento con fragmentos de información y del tiempo de ejecución.

En el capítulo 5 se presenta la nueva metodología propuesta y resultados de confiabilidad del sistema, así como el tiempo de ejecución.

En el capítulo 6 se presentan las conclusiones de este trabajo.

Capítulo 2. Fundamentos

En este capítulo se presentan las herramientas matemáticas utilizadas en este trabajo, las cuales sirven como fundamento esencial en el procesamiento digital de señales, así como soporte teórico en el trabajo realizado.

2.1 Imágenes Digitales

Una imagen generada por una cámara fotográfica, un radar, etc., puede ser representada por una función continua. Esta representación no es la forma en la que se representa una imagen en formato digital, por lo que se hace necesario convertir la imagen continua a su forma digital. El proceso para la digitalización de la imagen, consiste en muestrear la imagen para producir un arreglo de $N \times M$ píxeles y cuantizar (dividir) sus valores de intensidad en G niveles.

Los G niveles de cuantización, nos dará el valor para cada píxel, el cual es usualmente un número entre 0 y 255. Por lo tanto, para que la imagen sea visible, el valor de cada píxel es convertido a una escala de gris, donde 0 es negro y 255 blanco (256 niveles de gris), y los valores intermedios son matices de gris. Estos 256 niveles de gris son llamados niveles de cuantización.

Comúnmente se utilizan 256 niveles de gris en el procesado de imágenes, lo cual corresponde a un solo byte por píxel. La razón para ello es por conveniencia en el manejo de los datos, debido a que ésta es la forma en que usualmente las computadoras almacenan los datos.

2.2 Transformada de Fourier

La transformada de Fourier es una herramienta importante dentro del procesamiento de señales, la cual es utilizada para extraer la información de la señal que se encuentra en la frecuencia, fase y amplitud de las componentes senoidales que la

integran, permitiéndonos calcular la contribución de cada valor de frecuencia a la formación de dicha señal.

2.2.1 Transformada de Fourier para una función unidimensional

La definición de la transformada de Fourier para una señal continua unidimensional está dada por (Gonzalez y Woods, 1992, p.82)

$$F(u) = \int_{-\infty}^{\infty} f(x) \exp(-j2\pi ux) dx, \quad (1)$$

donde $j = \sqrt{-1}$ y u es la frecuencia.

También, dada la función $F(u)$, podemos a partir de ésta calcular la función $f(x)$, es decir que esta transformación tiene inversa, definida por (Gonzalez y Woods, 1992, p.82)

$$f(x) = \int_{-\infty}^{\infty} F(u) \exp(j2\pi ux) du. \quad (2)$$

2.2.2 Transformada de Fourier bidimensional

Para una función en dos dimensiones expresada como $f(x, y)$, donde x, y son coordenadas espaciales, la transformada de Fourier bidimensional $F(u, v)$ se define por

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp(-j2\pi(ux + vy)) dx dy, \quad (3)$$

donde u, v son frecuencias.

La transformada inversa de Fourier bidimensional se define por

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) \exp(j2\pi(ux + vy)) \, dudv. \quad (4)$$

Se observa que las componentes de la transformada de Fourier son cantidades complejas; esto significa que está conformada por una parte real $Re\{F(u, v)\}$ y una parte imaginaria $Im\{F(u, v)\}$, por lo tanto, al igual que cualquier otro número complejo la transformada de Fourier puede ser expresada en forma polar:

$$F(u, v) = |F(u, v)| \exp(j\phi(u, v)), \quad (5)$$

donde

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

es la *magnitud, módulo o espectro de amplitud* de la transformada de Fourier y

$$\phi(u, v) = \tan^{-1} \left(\frac{Im\{F(u, v)\}}{Re\{F(u, v)\}} \right), \quad (7)$$

es el *ángulo de fase o espectro de fase* de la transformada de Fourier.

La *densidad espectral de potencia o espectro de energía* de una señal, se obtiene mediante el módulo al cuadrado de la transformada de Fourier:

$$E(u, v) = |F(u, v)|^2. \quad (8)$$

2.3 Propiedades de la transformada de Fourier

Desde el punto de vista del procesamiento de imágenes, las siguientes propiedades son importantes (González y Woods, 2002, p.194).

2.3.1 Linealidad

Si $\mathfrak{F}\{af_1(x, y)\} = aF_1(u, v)$ es la transformada de Fourier de la función $f_1(x, y)$ multiplicada por una constante a y $\mathfrak{F}\{bf_2(x, y)\} = bF_2(u, v)$ es la transformada de Fourier de la función $f_2(x, y)$ multiplicada por una constante b , se tiene que

$$\mathfrak{F}\{af_1(x, y) + bf_2(x, y)\} = aF_1(u, v) + bF_2(u, v), \quad (9)$$

La ecuación (9) significa que la transformada de Fourier cumple con las propiedades de homogeneidad y superposición.

2.3.2 Translación en el dominio del espacio

Si $\mathfrak{F}\{f(x, y)\} = F(u, v)$ y x_0, y_0 son constantes reales, se tiene que

$$\mathfrak{F}\{f(x - x_0, y - y_0)\} = F(u, v) \exp(-j2\pi(ux_0 + vy_0)). \quad (10)$$

Esto nos indica que la transformada de Fourier de una función desplazada en el espacio es la misma transformada de Fourier de la función original multiplicada por un factor de fase.

2.3.3 Translación en el dominio de la frecuencia

Si $F(u, v)$ es desplazada por las constantes u_0, v_0 , la transformada inversa de Fourier se expresa:

$$\mathfrak{F}^{-1}\{F(u - u_0, v - v_0)\} = f(x, y) \exp(j2\pi(xu_0 + yv_0)). \quad (11)$$

2.3.4 Teorema de escalamiento

Si $\mathfrak{F}\{f(x, y)\} = F(u, v)$, entonces

$$\mathfrak{T}\{f(ax, by)\} = \frac{1}{|ab|} F\left(\frac{u}{a}, \frac{v}{b}\right), \quad (12)$$

donde a y b son constantes reales; esto es, una expansión de coordenadas en el dominio espacial, resulta en una contracción en el dominio de frecuencias y viceversa, más un cambio en la amplitud total del espectro.

2.3.5 Simetría

Si $f(x, y)$ es real, el módulo de su transformada de Fourier exhibe simetría, esto es,

$$|F(u, v)| = |F(-u, -v)|. \quad (13)$$

2.4 Teorema de Parseval

El teorema de Parseval establece que

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)|^2 dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F(u, v)|^2 du dv, \quad (14)$$

lo cual significa que la energía de una señal en el dominio del espacio, es igual a la energía de la señal en el dominio de la frecuencia, esto es, la energía de la señal se conserva al aplicar la transformada de Fourier, además el área bajo la curva de esta integral es la *varianza* de la función.

2.5 Teorema de muestreo

Este teorema nos indica bajo qué condiciones de muestreo una señal continua puede ser recobrada a partir de un conjunto de muestras.

Sea $f(x)$ una función de banda limitada, esto es, los valores de su transformada de Fourier $F(u)$ son cero para cada valor de frecuencia u fuera del intervalo $[-u_0, u_0]$. La condición necesaria para que $f(x)$ se pueda recuperar a partir de un conjunto de muestras es:

$$\Delta x \leq \frac{1}{2u_0}, \quad (15)$$

donde:

Δx es el intervalo de muestreo, esto es la separación entre cada muestra.

$2u_0 = f_n$, es llamada "frecuencia de Nyquist", y es el valor de frecuencia de muestreo mínimo para poder recuperar una señal continua muestreada.

Lo anterior significa que para poder recuperar una función continua, ésta debe ser muestreada con una frecuencia $f_n \geq 2u_0$.

Para el caso de una función bidimensional $f(x, y)$ como una imagen, el teorema de muestreo es:

$$\Delta x \leq \frac{1}{2u_0} \quad (16)$$

y

$$\Delta y \leq \frac{1}{2v_0}, \quad (17)$$

donde u_0 y v_0 representan las frecuencias límite en las direcciones x y y , respectivamente.

2.6 Transformada discreta de Fourier

La transformada discreta de Fourier (TDF), al igual que la transformada de Fourier para una señal continua, nos permite realizar un análisis en el *dominio de la*

frecuencia, con la diferencia que es utilizada para el caso de señales digitalizadas, como lo son las imágenes con las que realizamos nuestro trabajo.

2.6.1 Transformada discreta de Fourier para una función unidimensional

En el caso de una señal digitalizada, una función $f(x)$ es muestreada tomando N muestras de la misma separadas Δx unidades; esto significa que se obtendrá una sucesión finita de valores discretos.

El par de transformadas discretas de Fourier asociadas a una sucesión finita de valores discretos de una señal digital, se expresa de la siguiente forma:

$$F(u) = \frac{1}{N} \sum_{x=1}^N f(x) \exp(-j2\pi ux / N), \quad (18)$$

para $u = 1, 2, \dots, N$;

$$f(x) = \sum_{u=1}^N F(u) \exp(j2\pi ux / N), \quad (19)$$

para $x = 1, 2, \dots, N$.

2.6.2 Transformada discreta de Fourier de una función bidimensional

En el caso de una función de dos variables, el par de transformadas discretas de Fourier están dadas por las expresiones:

$$F(u, v) = \frac{1}{NM} \sum_{x=1}^M \sum_{y=1}^N f(x, y) \exp\left(-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right), \quad (20)$$

para $u = 1, 2, \dots, M$ y $v = 1, 2, \dots, N$;

$$f(x, y) = \sum_{u=1}^M \sum_{v=1}^N F(u, v) \exp\left(j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right), \quad (21)$$

para $x = 1, 2, \dots, M$ y $y = 1, 2, \dots, N$.

2.7 Convolución

Dentro del procesamiento digital de señales, la convolución es una operación matemática de importancia, debido a que nos relaciona tres señales de interés en un sistema lineal: la señal de entrada, la señal de salida y la respuesta al impulso (Smith, W. S., 1999, pp.107-109).

2.7.1 En una dimensión

Para funciones continuas:

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(x') g(x - x') dx'. \quad (22)$$

Para funciones Discretas:

$$f(x) * g(x) = \sum_{m=1}^M f(m) g(x - m), \quad (23)$$

donde $m = 1, 2, \dots, M$.

2.7.2 En dos dimensiones

La convolución en el dominio espacial entre dos funciones bidimensionales, es expresada matemáticamente por la doble integral

$$f(x, y) * g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') g(x - x', y - y') dx' dy'. \quad (24)$$

En el caso de señales discretas se tiene

$$f(x, y) * g(x, y) = \sum_{m=1}^M \sum_{n=1}^N f(m, n) g(x - m, y - n), \quad (25)$$

donde $m = 1, 2, \dots, M$ y $n = 1, 2, \dots, N$.

2.8 Teorema de Convulación

El teorema de convolución indica que la transformada de Fourier de la convolución de dos funciones es igual al producto de la transformada de Fourier de dichas funciones, esto es,

$$\mathfrak{F}\{f(x, y) * g(x, y)\} = F(u, v)G(u, v), \quad (26)$$

donde $F(u, v)$ y $G(u, v)$ son las transformadas de Fourier de $f(x, y)$ y $g(x, y)$ respectivamente.

Lo anterior implica que podemos calcular la convolución de dos funciones aplicando la transformada inversa de Fourier al resultado de la multiplicación de las correspondientes transformadas de Fourier de cada función, esto es,

$$f(x, y) * g(x, y) = \mathfrak{F}^{-1}\{F(u, v)G(u, v)\}. \quad (27)$$

2.9 Correlación

La correlación es una operación matemática que nos indica el grado de relación entre dos funciones, ya sean éstas funciones continuas o funciones discretas.

2.9.1 En una dimensión

Si tenemos una serie de datos representados por la función $f(m)$, la cual es la función objetivo, y tenemos además una serie de datos, los cuales hemos recopilado de algún experimento, $g(m)$, la correlación nos permite encontrar la correspondencia entre $f(m)$ y $g(m)$, y está definida como

$$c(x) = f(x) \circ g(x) = \sum_{m=0}^{M-1} f^*(m)g(m+x), \quad (28)$$

donde $m=0,1,2,3,\dots,M-1$, \circ significa correlación y $*$ indica el complejo conjugado.

Para una función continua la correlación se expresa como

$$f(x) \circ g(x) = \int_{-\infty}^{\infty} f^*(x')g(x'+x)dx'. \quad (29)$$

2.9.2 En dos dimensiones

Caso continuo:

$$f(x) \circ g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(x', y')g(x'+x, y'+y)dx' dy'. \quad (30)$$

Caso discreto:

$$f(x) \circ g(x, y) = \sum_{m=1}^M \sum_{n=1}^N f^*(m, n)g(x+m, y+n), \quad (31)$$

donde $m = 1, 2, \dots, M$; $n = 1, 2, \dots, N$.

2.10 Teorema de Correlación

El teorema de correlación indica que la transformada de Fourier de la correlación de dos funciones $f(x)$ y $g(x)$ es igual a la multiplicación de los espectros de frecuencia de la transformada de Fourier de $f(x)$ y el conjugado de la transformada de Fourier de $g(x)$:

$$\mathfrak{F}\{f(x) \circ g(x)\} = F(u)G^*(u). \quad (32)$$

En el caso de tener dos funciones bidimensionales $f(x, y)$ y $g(x, y)$, el teorema de correlación se expresa:

$$\mathfrak{F}\{f(x, y) \circ g(x, y)\} = \mathfrak{F}\{F(u, v)G^*(u, v)\}. \quad (33)$$

2.11 Transformada de Mellin

La familia de transformadas, llamada la transformada β -Mellin (De Sena y Rocchesso, 2007), está dada por

$$M\{f(x)\} = F_M(p) = \int_0^{\infty} f(x)x^{p-1}dx, \quad (34)$$

donde se tiene la variable compleja $p = \beta - j\omega$, donde el parámetro β es un número real y $j = \sqrt{-1}$.

Usando un valor de $\beta = 0$ tenemos la transformada de Mellin invariante a compresión/expansión (De Sena y Rocchesso, 2007), en la cual tendremos $p = -j\omega$, donde $\omega = 2\pi u$. En este caso,

$$F_M(j\omega) = \int_0^{\infty} f(x)x^{-j\omega-1} dx = \int_0^{\infty} \frac{1}{x} f(x) \exp(-j\omega \ln x) dx . \quad (35)$$

Por comodidad, de aquí en adelante $F_M(j\omega)$ será escrito como $F_M(\omega)$.

2.11.1 Invariancia a escala de la transformada de Mellin

La transformada de Mellin cuando $\beta = 0$ es especialmente útil por ser invariante a compresiones y expansiones; esto significa que la transformada de Mellin es invariante a la escala. Ésta se ha aplicado en la restauración de imágenes espacialmente variantes (Sawchuk, 1974) y en el análisis de redes que varían con el tiempo (Gerardi, 1959), entre otras aplicaciones.

Partiendo de la transformada de Mellin, dada por la ecuación (35), la transformada de Mellin de $f(\alpha x)$, donde α es un factor de escalamiento es:

$$M\{f(\alpha x)\} = F_M^\alpha(\omega) = \int_0^{\infty} f(\alpha x)x^{-j\omega-1} dx . \quad (36)$$

Si $\eta = \alpha x$ entonces $x = \frac{\eta}{\alpha}$ y $dx = \frac{d\eta}{\alpha}$,

entonces la transformada de Mellin de la función escalada $F_M^\alpha(\omega)$ puede describirse como

$$F_M^\alpha(\omega) = \int_0^{\infty} \left(\frac{\eta}{\alpha}\right)^{-j\omega-1} f(\eta) \frac{d\eta}{\alpha} ; \quad (37)$$

o también se puede escribir de la forma

$$F_M^\alpha(\omega) = \alpha^{j\omega} \int_0^{\infty} \eta^{-j\omega-1} f(\eta) d\eta , \quad (38)$$

donde se puede ver que la integral es la transformada de Mellin sin escalar, esto es,

$$F_M^\alpha(\omega) = \alpha^{j\omega} F_M(\omega), \quad (39)$$

es decir,

$$F_M^\alpha(\omega) = e^{j\omega \ln \alpha} F_M(\omega). \quad (40)$$

Esto significa que un cambio en la escala produce un factor de fase en la transformada de Mellin. Por lo tanto, el módulo de la transformada de Mellin nos producirá una invarianza a la escala debido a que

$$|F_M^\alpha(\omega)| = |F_M(\omega) \exp(j\omega \ln \alpha)| = |F_M(\omega)|. \quad (41)$$

2.11.2 La transformada de Mellin vía la transformada de Fourier

Debido a que en este trabajo se hace uso del programa matlab y éste no dispone de un algoritmo para el cálculo de la transformada de Mellin, es necesario hacerlo vía la transformada de Fourier. Esto se realiza mediante un cambio de variable de la forma $x = \exp(q)$ en la ecuación (35), donde los límites de integración en la nueva variable

son $\lim_{x \rightarrow 0}(\ln x) = -\infty$ y $\lim_{x \rightarrow \infty}(\ln x) = \infty$, entonces se tiene:

$$F_M(\omega) = \int_0^{\infty} f(x) x^{-j\omega-1} dx = \int_{-\infty}^{\infty} f(\exp(q)) \exp(-j\omega q) dq, \quad (42)$$

donde $\omega = 2\pi u$. Por lo tanto, la transformada de Mellin de la función $f(x)$ queda escrita ahora como la transformada de Fourier de la función $f(\exp(q))$.

2.11.3 Transformada de Mellin bidimensional

La transformada de Mellin en dos dimensiones de una función a lo largo del eje imaginario, se define por (Casasent y Psaltis, 1976)

$$M(\omega_x, \omega_y) = \int_0^{\infty} \int_0^{\infty} f(x, y) x^{-j\omega_x - 1} y^{-j\omega_y - 1} dx dy. \quad (43)$$

2.12 Transformada de escala

La transformada de escala fue desarrollada por L. Cohen (1995), quien encontró que ésta es un caso particular de la transformada de Mellin, la cual tiene la propiedad de ser invariante a escala.

Definiendo c como la variable de escala, entonces la transformada de escala $D(c)$ de una función $f(x)$ y su inversa están dadas por (De Sena y Rocchesso, 2004)

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(x) \exp((-jc - 1/2)\ln x) dx, \quad (44)$$

y

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} D(c) \exp((jc - 1/2)\ln x) dc; \quad x \geq 0. \quad (45)$$

Se puede ver de la ecuación (44), que la transformada de escala es la transformada de Mellin con el argumento complejo $p = \beta - j\omega$, donde $\beta = 1/2$.

Desarrollando la ecuación (44) se puede escribir de la siguiente forma:

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \frac{1}{\sqrt{x}} f(x) \exp(-jc \ln x) dx. \quad (46)$$

Si se compara la ecuación anterior con la ecuación (35), se puede observar que sin tomar en cuenta el factor constante $1/\sqrt{2\pi}$, la única diferencia entre la transformada de escala y la transformada de Mellin es un factor $1/\sqrt{x}$. Por lo tanto, si tenemos la

transformada de Mellin F_M de una función $f(x)$, podemos obtener la transformada de escala multiplicando la función $f(x)$ por el factor \sqrt{x} .

2.12.1 La transformada de escala vía la transformada de Fourier

Para poder implementar esta transformada de escala, reemplazamos a x por $\ln(x)$ de la forma

$$f_L(x) = \frac{1}{\sqrt{x}} f(\ln x), \quad (47)$$

y substituyendo en la ecuación (46) tenemos

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \frac{1}{\sqrt{x}} f_L(x) \exp(-jc \ln x) dx, \quad (48)$$

aplicando las propiedades de los exponentes y simplificando se obtiene:

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(\ln x) x^{(-jc)} \frac{dx}{x}. \quad (49)$$

Haciendo $\tau = \ln x$; $d\tau = \frac{dx}{x}$; $dx = x d\tau$, además los límites serán ahora de $-\infty$ a ∞ , entonces se tiene que

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) \exp(-jc\tau) d\tau, \quad (50)$$

$$D(c) = \mathfrak{F}\{f(\tau)\}. \quad (51)$$

Por lo tanto, la transformada de escala de la función $f_L(x)$ es igual a la transformada de Fourier de la función $f(\tau)$; esto es, se puede implementar la transformada de

escala vía la transformada de Fourier realizando un mapeo logarítmico en la función de entrada.

2.12.2 La transformada de escala bidimensional

En dos dimensiones la transformada de escala y su inversa se pueden escribir de la forma

$$D(c_x, c_y) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \int_0^{\infty} \frac{1}{\sqrt{xy}} f(x, y) \exp(-jc_x \ln x - jc_y \ln y) dx dy, \quad (52)$$

$$f(x, y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{xy}} D(c_x, c_y) \exp(jc_x \ln x + jc_y \ln y) dc_x dc_y. \quad (53)$$

En este trabajo se usó la transformada de escala en 2-D en coordenadas polares, la cual está expresada como

$$D(c_r, c_\theta) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \int_0^{2\pi} f(r, \theta) r^{(-jc_r - 1/2)} \exp(-jc_\theta \theta) dr d\theta, \quad (54)$$

donde $r = \sqrt{x^2 + y^2}$, y $\theta = \tan^{-1}\left(\frac{y}{x}\right)$.

Utilizando la transformada de escala de esta manera se obtiene invarianza ante cambios de rotación además de la invarianza a cambios de escala.

Aplicando el logaritmo a la coordenada radial, es decir $\rho = \ln r$, entonces la transformada de escala en coordenadas polares se escribe (Pech-Pacheco, J. L *et. al.* 2003):

$$D(c_\rho, c_\theta) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \int_0^{2\pi} \exp(\rho/2) f(\rho, \theta) \exp(-j(\rho c_\rho + \theta c_\theta)) d\rho d\theta. \quad (55)$$

De esta manera se puede implementar la transformada de escala vía la transformada de Fourier.

2.13 Filtros de correlación

En esta sección se presentan los principales filtros de correlación, los cuales sirven como herramienta en el reconocimiento y localización de objetos dentro de una imagen o escena de prueba. Los filtros que a continuación se describen son derivados del filtro clásico acoplado, introducido por Vander Lugt (1964).

2.13.1 Filtro clásico acoplado CMF (*Classical Matched Filter*)

Este filtro, también conocido como filtro complejo u holográfico, es un filtro lineal propuesto y probado por Vander Lugt (1964), al cual lo describió como una función compleja de la forma

$$H(u, v)_{CMF} = |H(u, v)| \exp(j\phi(u, v)), \quad (56)$$

donde u, v son las variables de frecuencia, $|H(u, v)|$ es la respuesta de amplitud o módulo de la transformada de Fourier del objeto y $\exp(j\phi(u, v))$ es la respuesta de fase y $\phi(u, v)$ es la fase de la transformada de Fourier del objeto.

Algunas de las limitaciones de este filtro clásico son, un pico de correlación amplio y sensibilidad a distorsiones y debido a las consideraciones tomadas en su diseño, presenta un funcionamiento inadecuado para las imágenes que contienen al objetivo espacialmente separado del fondo (Javidi y Horner, 1994).

2.13.2 Filtro sólo de fase POF (*Phase Only Filter*)

Este filtro se expresa como (Horner y Gianino, 1984):

$$H(u, v)_{POF} = \exp(j\phi(u, v)). \quad (57)$$

Con el filtro sólo de fase se obtienen picos de correlación más estrechos y es más selectivo que el CMF.

2.13.3 Filtro no lineal NLF (*Nonlinear Filter*), (ley k)

La no linealidad del filtro es determinada por el exponente k que afecta al módulo de la transformada de Fourier de la imagen de referencia (ley k). La expresión para el filtro no lineal está dada por (Vijaya Kumar y Hassebrook, 1990)

$$H[u, v]_{NLF} = |S(u, v)|^k \exp(j\phi(u, v)), \quad (58)$$

donde $\phi(u, v)$ es la fase de la transformada de Fourier de la imagen de referencia y $|S(u, v)|$ es el módulo de la transformada de Fourier de la imagen de referencia, k puede tomar valores entre $0 < k < 1$. Variando el valor de k podemos variar la capacidad de discriminación del filtro no lineal y de esta manera encontrar el valor óptimo de k para el filtro dependiendo del sistema de reconocimiento que se desee utilizar.

Capítulo 3. Metodología

Con el desarrollo del filtro clásico CMF Vander Lugt en 1964, se logra un avance en el reconocimiento de patrones mediante el método de correlación, sin embargo, este tipo de filtros no funcionan correctamente cuando existen variaciones de escala, posición o rotación. Por otro lado, existen en la actualidad diversas metodologías que son capaces de vencer estos obstáculos. En este trabajo se implementaron y se analizaron cuatro metodologías invariantes a posición, rotación y escala.

3.1 Correlación no lineal con transformada de escala invariante a posición, rotación y escala (Coronel-Beltrán y Álvarez-Borrego, 2010)

En esta metodología se utiliza la transformada de escala en coordenadas polares con el logaritmo de la coordenada radial $\rho = \ln r$, expresada como en la ecuación (55), la cual permite tener una respuesta invariante a escala y rotación. Dado que el módulo de la transformada de Fourier tiene la propiedad de invariancia a posición, esta forma de la transformada de escala se aplica al módulo de la transformada de Fourier de una imagen. Con esta implementación se logra tener una respuesta invariante a posición, rotación y escala. Además, con la intención de tener una mayor claridad en la respuesta de la correlación, se utiliza un filtro no-lineal.

3.1.1 Obtención del filtro

El algoritmo para obtener el filtro no lineal se muestra en la figura 1. Primero tenemos la imagen original u objetivo de dimensiones $N \times N$ descrita por la función bidimensional de la forma $f(x, y)$, donde (x, y) representa un pixel de la imagen, por ejemplo, como la imagen 256×256 de la letra E tipo arial en blanco y negro mostrada en la figura 2.

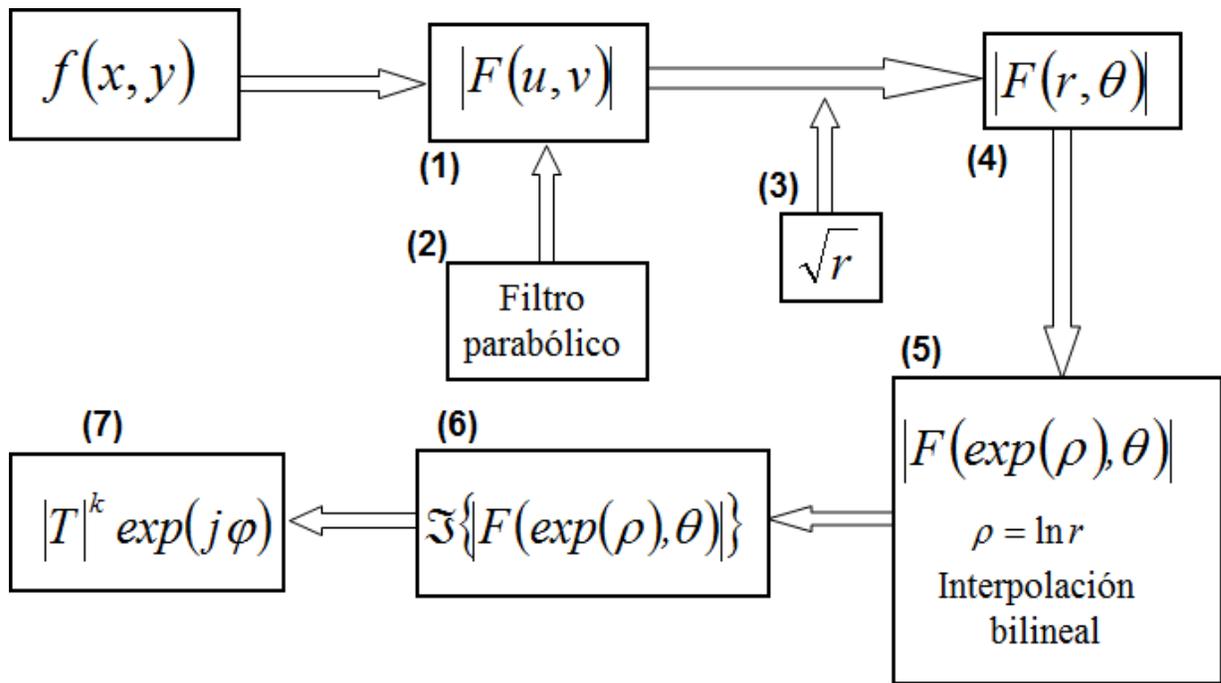


Figura 1. Diagrama de bloques para obtener el filtro no lineal.

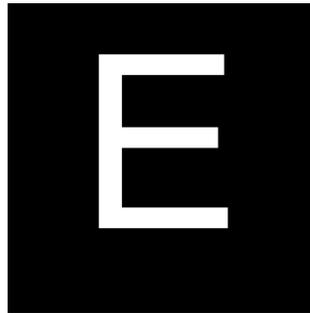


Figura 2. Imagen blanco y negro de la letra E tipo arial.

Los pasos a seguir mostrados en la figura 1 son:

1. Se calcula la transformada de Fourier de la imagen $f(x, y)$ y se obtiene su módulo, denotado como $|F(u, v)|$ y de esta manera obtendremos invarianza a posición. El módulo de la transformada de Fourier de la letra se muestra en la figura 3, donde se grafica el $\log(|F(u, v)| + 1)$ por cuestiones visuales.

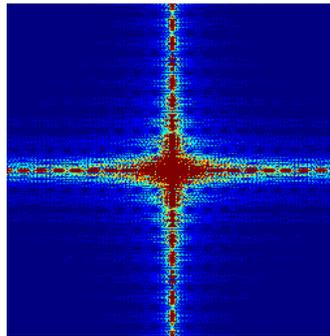


Figura 3. Módulo de la transformada de Fourier de la letra E tipo arial (se grafica el $\log(|F(u,v)| + 1)$ para mejor visualización).

2. En el paso 2 se aplica al módulo de la transformada de Fourier un filtro parabólico con el objetivo de resaltar las altas frecuencias, las cuales determinan los detalles finos de la imagen. El espectro filtrado de la figura 3 se muestra en la figura 4.

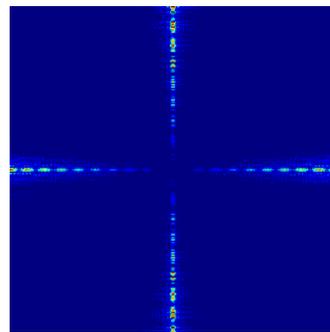


Figura 4. Módulo de la transformada de Fourier después de haber sido aplicado el filtro parabólico.

3. El siguiente paso es multiplicar por el factor \sqrt{r} , donde $r = \sqrt{u^2 + v^2}$ es la frecuencia espacial radial. Este factor es el que hace la diferencia entre la transformada de Mellin y la transformada de escala, por lo que nos permitirá poder aplicar la transformada de escala al módulo de la transformada de Fourier filtrado en el paso 2.

4. Después, se le aplica al módulo de la transformada de Fourier filtrado una conversión de coordenadas cartesianas de las frecuencias espaciales (u,v) a coordenadas polares (r,θ) para expresar el módulo de la transformada de Fourier $|F(u,v)|$ como $|F(r,\theta)|$, donde

$$\theta = \tan^{-1}\left(\frac{v}{u}\right). \quad (59)$$

De esta manera obtendremos invarianza a rotación. En la figura 5 se muestra el módulo de la transformada de Fourier en coordenadas polares.

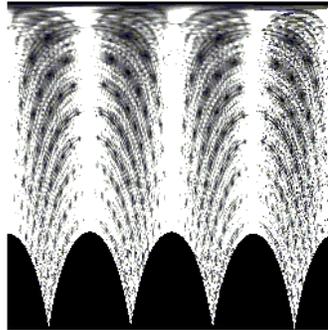


Figura 5. Imagen del espectro de frecuencias en coordenadas polares.

5. Se hace un escalamiento logarítmico en la parte radial de las coordenadas polares, para poder implementar la transformada de escala mediante la transformada de Fourier $F(\exp(\rho), \theta)$. Además se realiza una interpolación bilineal para evitar fugas de información, el resultado de este paso se muestra en la figura 6.

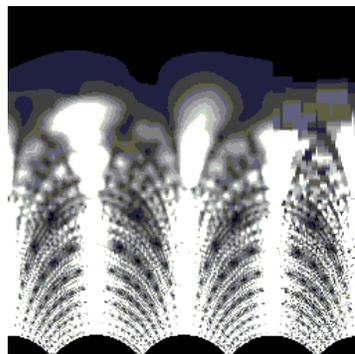


Figura 6. Imagen del espectro de frecuencias en coordenadas polares después de aplicar $\rho = \ln r$.

6. Aplicamos la transformada de Fourier para obtener la transformada de escala.

7. En este paso, a partir de la transformada de escala obtenemos el filtro no lineal invariante a posición, rotación y escala de la imagen de referencia descrito por

$|T|^k \exp(j\varphi)$, donde $|T|^k$ es el módulo de la transformada de escala elevado a la potencia k y φ es la fase.

3.1.2 Correlación no lineal

El filtro no lineal $|T|^k \exp(j\varphi)$ es posteriormente utilizado para ser correlacionado con la información no lineal de la imagen problema, a la cual se le aplica el mismo proceso, obteniéndose el filtro no lineal $|G|^k \exp(j\varphi)$, donde k es el factor de no linealidad y φ es la fase (figura 7). Esta correlación se realiza utilizando un valor de $k = 0.3$, el cual es el valor óptimo para esta metodología (Coronel-Beltrán y Álvarez-Borrego, 2010).

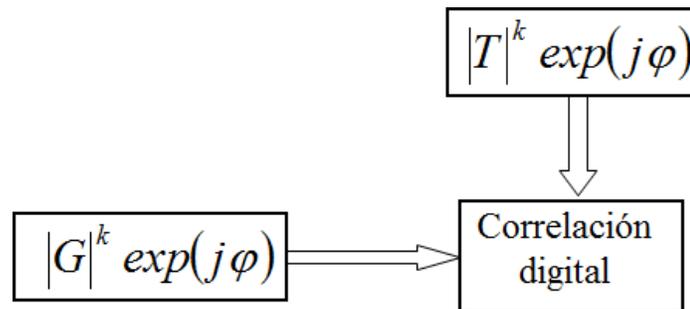


Figura 7. Correlación digital mediante filtros no lineales.

3.2 Índices espectrales (Fimbres-Castro, *et. al.*, Optical Engineering, Vol. 51, No. 4. 2012)

El índice espectral es un valor obtenido a través de procedimientos matemáticos, que representa la imagen denotada por una función bidimensional (Fimbres-Castro, *et. al.*, 2012). En esta metodología se calculan dos índices, los cuales son utilizados para presentar la imagen en un plano, llamado plano de identificación.

3.2.1 Índice 1

El procedimiento para obtener el índice 1 se muestra en la figura (8).

Dada una imagen $N \times N$ del objeto a ser detectado, ésta puede ser expresada como la función bidimensional $f(x, y)$. En el paso 1, se suman cada uno de los valores de la función sin tomar en cuenta el fondo, obteniendo un valor representativo de esta imagen, el cual es llamado *Volumen*:

$$Volumen = \sum_{x=1}^N \sum_{y=1}^N f(x, y). \quad (60)$$

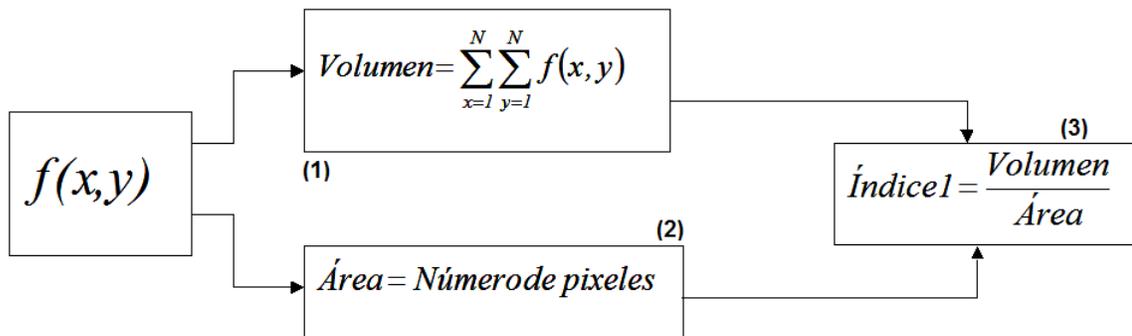


Figura 8. Procedimiento para obtener el índice 1.

Con el cálculo del *Volumen* se obtiene invariancia a rotación, dado que al ser rotada la imagen este valor se conservará. Sin embargo, el valor *Volumen* no se conservará cuando exista un cambio de escala en la imagen de referencia $f(x, y)$, ya que este valor aumentará o disminuirá en caso de que la escala aumente o disminuya, por ello, en el paso 2 se toma en consideración el número de píxeles de la imagen de referencia $f(x, y)$, el cual es denotado como *Área*. Ya que el valor *Área* también varía con algún cambio de escala, en el paso 3, la invariancia a escala se obtendrá dividiendo el *Volumen* entre el *Área*; de esta manera se obtiene el índice 1:

$$Índice 1 = \frac{Volumen}{Área}. \quad (61)$$

3.2.2 Índice 2

El procedimiento de obtención del índice 2 se observa en la figura 9.

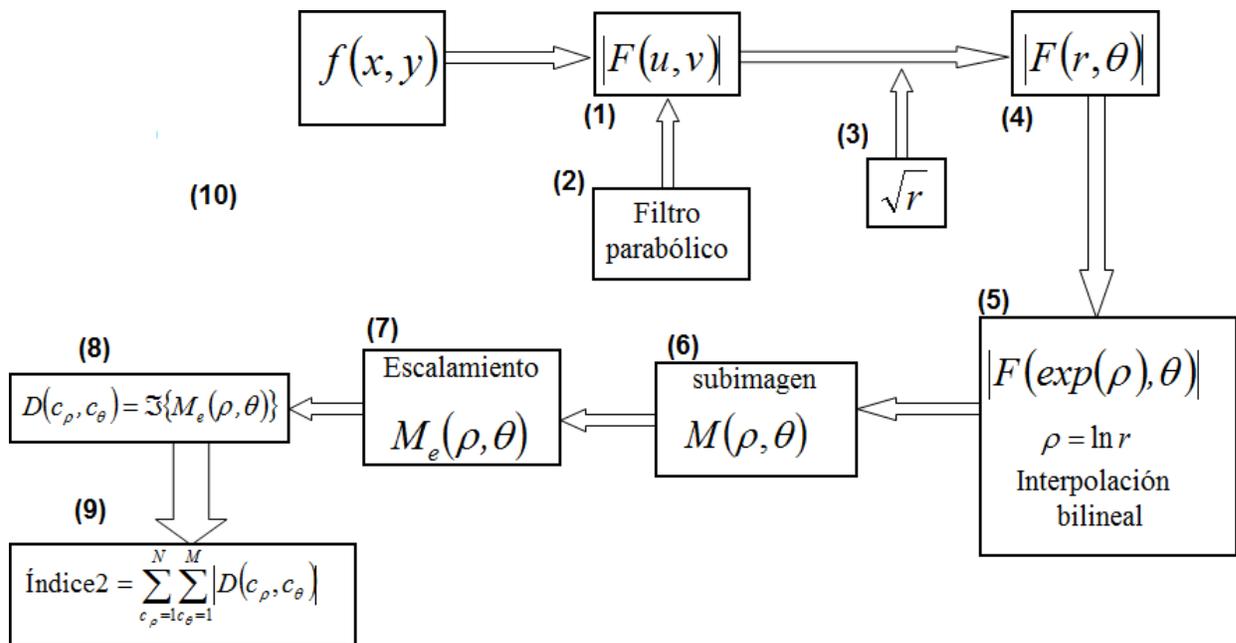


Figura 9. Procedimiento para la obtención del índice 2.

Los pasos 1 al 5 ya fueron explicados en la sección 3.1.1. En el paso 6, a partir del resultado obtenido en el paso 5 se extrae una subimagen $M(\rho, \theta)$ del espectro de frecuencias, realizando un recorte que corresponde a la sección que contiene mayor información. En la figura 10, se muestra el recorte realizado al espectro de frecuencias de la imagen 256×256 de la letra E tipo arial de la figura 6.

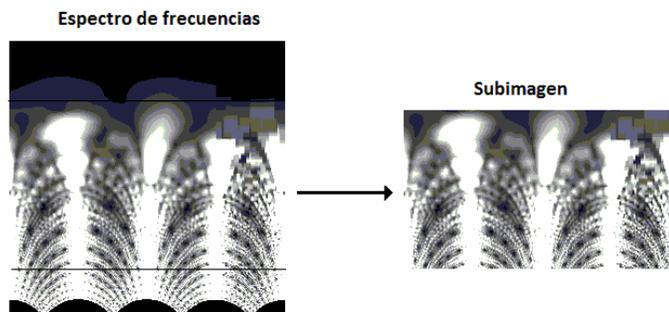


Figura 10. Obtención de subimagen del espectro de frecuencias.

En el paso 7, se realiza un escalamiento de la subimagen $M(\rho, \theta)$ para tener de nueva cuenta el tamaño de la imagen original, figura 11. Este escalamiento es realizado para no tener fuga de información.

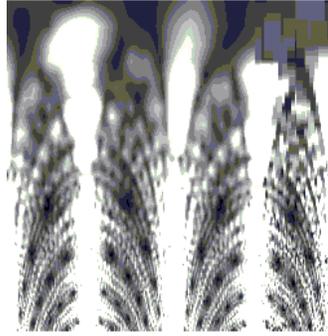


Figura 11. Escalamiento al tamaño de la imagen original.

En el paso 8, se aplica la transformada de Fourier a la subimagen escalada $M_e(\rho, \theta)$, obteniendo así la transformada de escala de dicha subimagen.

Finalmente en el paso 9 se obtiene el índice 2, el cual está definido por la siguiente ecuación (Fimbres-Castro *et. al*, 2012):

$$\text{Índice2} = \sum_{c_\rho=1}^N \sum_{c_\theta=1}^N |D(c_\rho, c_\theta)|, \quad (62)$$

donde $|D(c_\rho, c_\theta)|$ es el módulo de la transformada de escala de la subimagen, donde (c_ρ, c_θ) son las variables de escala en coordenadas polares para $\rho = \ln r$. El índice 2 obtenido de esta manera es invariante a escala, rotación y posición.

El resultado del procedimiento de identificación se muestra en un plano de identificación donde cada una de las imágenes procesadas tienen su respectivo espacio, esto es, no es necesario hacer una comparación o correlación entre imágenes ya que todas quedan desplegadas en su correspondiente espacio determinado por sus respectivos índice 1 e índice 2 que la representan, por ejemplo, para las imágenes mostradas en la figura 12, el resultado de aplicar esta metodología se muestra en el plano de identificación de la figura 13.



Figura 12. Imágenes de prueba de la metodología de índices espectrales.

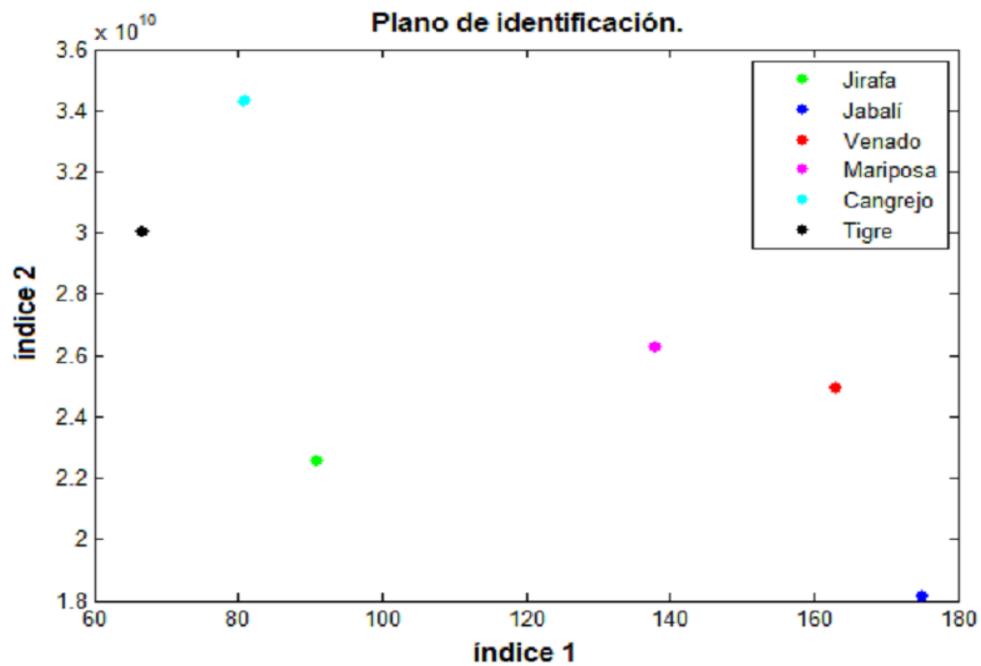


Figura 13. Resultado de la metodología de índices espectrales donde cada una de las imágenes mostradas en la figura 12 tiene su respectivo espacio dentro del plano de identificación.

3.3 Correlación invariante a posición, rotación y escala, con firmas vectoriales (Lerma-Aragón y Álvarez-Borrego, Journal of Modern Optics, Vol. 56, No. 14. 2009)

En la figura 14 se muestra el proceso de firmas vectoriales. Del punto 1 al punto 7, el proceso es el mismo al utilizado en el método por índices espectrales. En los puntos 8

y 9 se obtienen dos vectores unidimensionales a partir de la subimagen $M_e(\rho, \theta)$. Se calculan sus marginales utilizando las siguientes ecuaciones:

$$M_\theta(t) = \sum_{\rho=1}^N M_e(\rho, \theta), \text{ vector de rotación,} \quad (63)$$

y

$$M_\rho(s) = \sum_{\theta=1}^N M_e(\rho, \theta), \text{ vector de escala,} \quad (64)$$

donde $t = 1, \dots, N$ y $s = 1, \dots, N$.

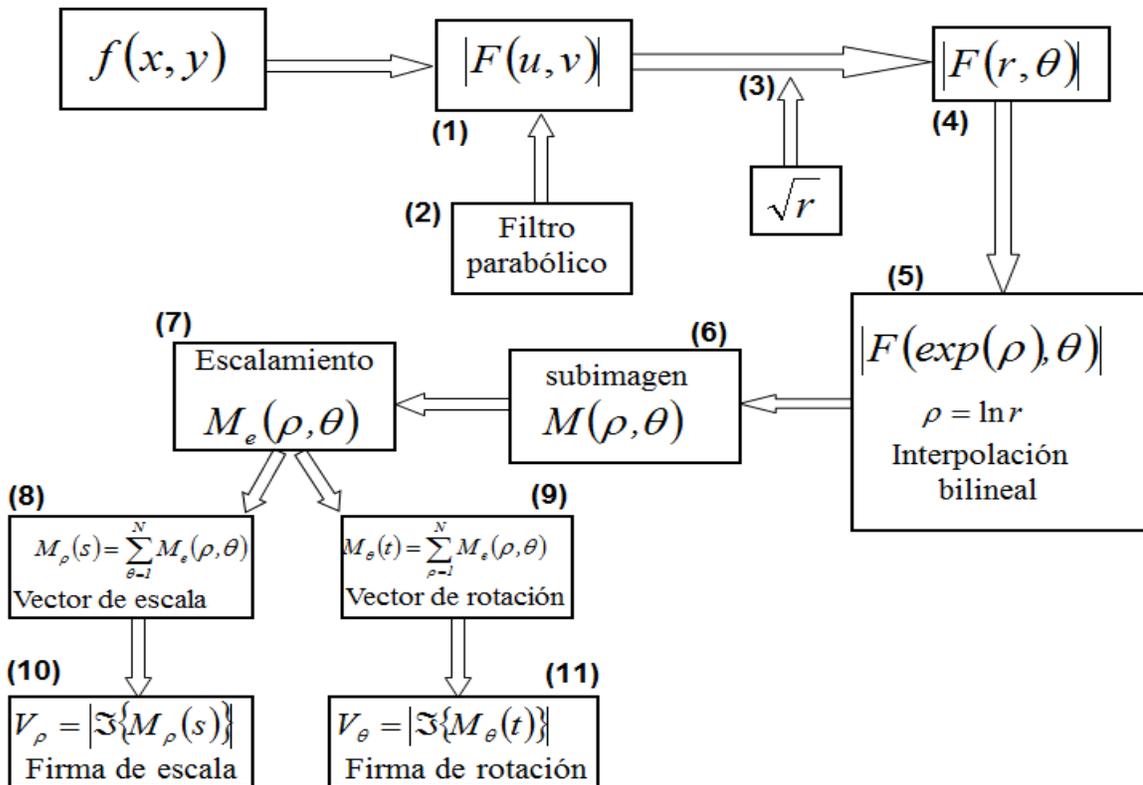


Figura 14. Diagrama a bloques del método de obtención de firmas vectoriales.

En los puntos 10 y 11 para obtener las firmas vectoriales de rotación y de escala se calcula el módulo de la transformada de Fourier, de los marginales unidimensionales $M_\theta(t)$ y $M_\rho(s)$. Estas firmas vectoriales son:

$$V_{\theta} = |\mathfrak{F}\{M_{\theta}(t)\}|, \text{ firma de rotación,} \quad (65)$$

y

$$V_{\rho} = |\mathfrak{F}\{M_{\rho}(s)\}|, \text{ firma de escala.} \quad (66)$$

Este proceso se realiza para una imagen de referencia y una imagen problema, obteniendo para cada una de éstas sus respectivas firmas de rotación y de escala.

Una vez obtenidas las respectivas firmas para la imagen de referencia e imagen problema, se realiza el cálculo de la distancia Euclidiana entre las firmas de la imagen de referencia e imagen problema. Una alternativa es calcular la correlación no lineal de las firmas (Fimbres-Castro, 2012).

3.4 Correlación invariante a posición, y rotación, con máscaras binarias (Álvarez-Borrego *et. al.* Optics Communications. Vol. 294. 2013)

Si se tiene una imagen $N \times N$ dada por la función bidimensional $f(x, y)$ y el pixel central de la imagen es (c_x, c_x) , donde c_x está dado por

$$c_x = \begin{cases} \frac{N}{2} + 1, & \text{si N es par,} \\ \left\lfloor \frac{N}{2} \right\rfloor + 1, & \text{si N es impar,} \end{cases} \quad (67)$$

donde $\left\lfloor \frac{N}{2} \right\rfloor$ se redondea al entero más cercano hacia $-\infty$.

Para una imagen de dimensiones 256×256 el pixel central será $(c_x, c_x) = (129, 129)$ o para una imagen de dimensiones 307×307 el pixel central es $(c_x, c_x) = (154, 154)$.

Para obtener una máscara binaria de una imagen, primeramente, se puede trabajar con la parte real de la transformada de Fourier de una imagen:

$$F_{\text{Re}}(u, v) = \text{Re}\{\mathfrak{F}(x, y)\}, \quad (68)$$

o la parte imaginaria,

$$F_{\text{Im}}(u, v) = \text{Im}\{\mathfrak{F}(x, y)\}. \quad (69)$$

Por ejemplo, para una la imagen de 256×256 como la letra E mostrada en la figura 2, las gráficas de la parte real e imaginaria de la transformada de Fourier se observan en la figura 15a y 15b respectivamente.

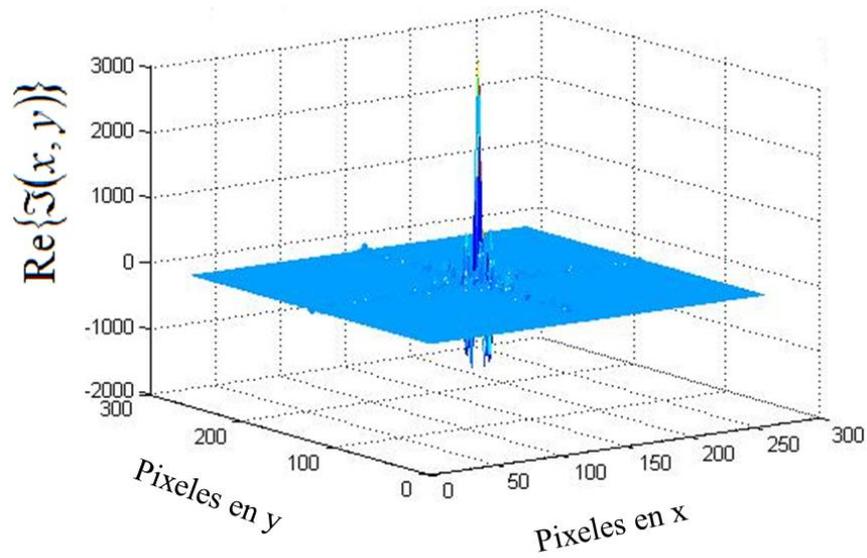
Posteriormente, se obtiene un perfil de la sección transversal central de la parte real o de la parte imaginaria de la transformada de Fourier; este perfil es un vector unidimensional. Para la parte real el perfil central está dado por

$$T_R(c_x, y) = \text{Re}\{\mathfrak{F}(c_x, y)\}, \quad (70)$$

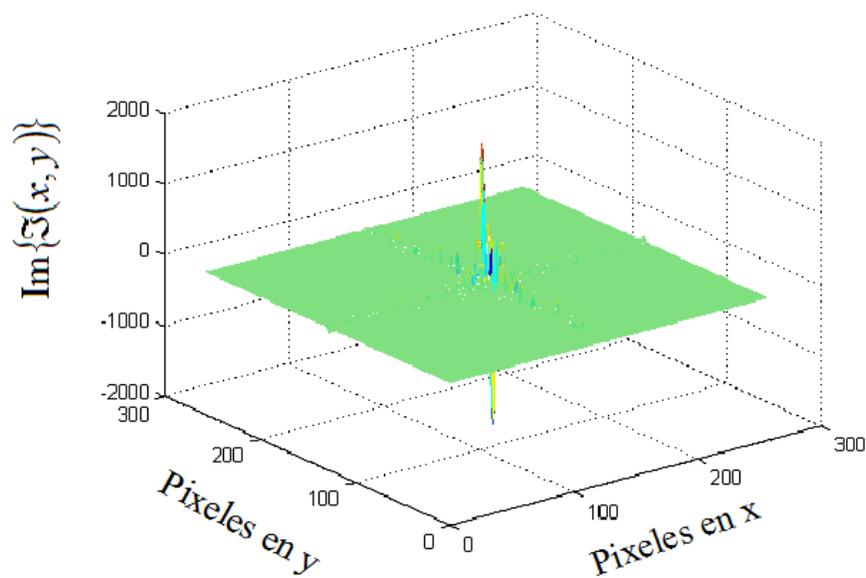
donde c_x es la coordenada en x del pixel central (c_x, c_x) y $1 \leq y \leq N$. El perfil central de la parte imaginaria de la transformada de Fourier está dado por

$$T_I(c_x, y) = \text{Im}\{\mathfrak{F}(c_x, y)\}, \quad (71)$$

donde $1 \leq y \leq N$ y c_x es la coordenada en x del pixel central.



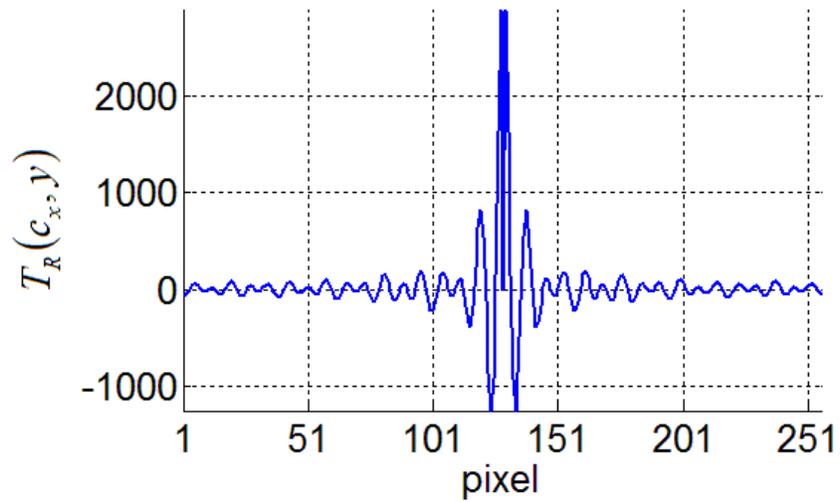
a)



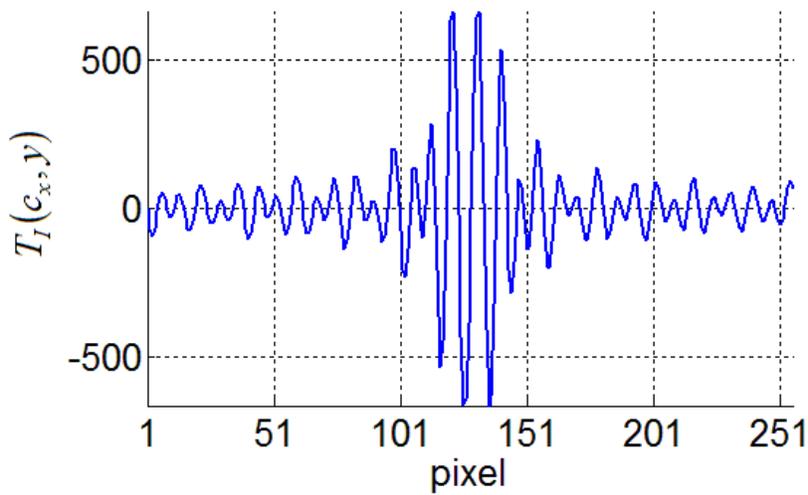
b)

Figura 15. Parte real (a) e imaginaria (b) de la transformada de Fourier de la imagen de la letra E de la figura 2.

En las figuras 16a y 16b se muestran las gráficas para $T_R(c_x, y)$ y $T_I(c_x, y)$ respectivamente, los cuales el perfil de la sección transversal de las gráficas mostradas en la figura 15a y 15b.



a)



b)

Figura 16. a) Perfil de la parte real de $F(u, v)$. b) Perfil de la parte imaginaria de $F(u, v)$.

El siguiente paso es aplicar un umbral al perfil central obtenido, para generar un perfil binario de dicho perfil; en este caso se utiliza el valor de cero como umbral de la siguiente manera:

$$Z_{RP}(y) = \begin{cases} 1, & \text{si } T_R(c_x, y) > 0, \\ 0, & \text{si } T_R(c_x, y) \leq 0, \end{cases} \quad (72)$$

$$Z_{IP}(y) = \begin{cases} 1, & \text{si } T_I(c_x, y) > 0, \\ 0, & \text{si } T_I(c_x, y) \leq 0, \end{cases} \quad (73)$$

donde $Z_{RP}(y)$ y $Z_{IP}(y)$ son el perfil binario positivo para la parte real y parte imaginaria, respectivamente. También se puede tomar el umbral de manera inversa para la parte real y para la parte imaginaria, esto es:

$$Z_{RN}(y) = \begin{cases} 1, & \text{si } T_R(c_x, y) \leq 0, \\ 0, & \text{si } T_R(c_x, y) > 0, \end{cases} \quad (74)$$

$$Z_{IN}(y) = \begin{cases} 1, & \text{si } T_I(c_x, y) \leq 0, \\ 0, & \text{si } T_I(c_x, y) > 0, \end{cases} \quad (75)$$

y $Z_{RN}(y)$ y $Z_{IN}(y)$ corresponden al perfil binario no positivo de la parte real y parte imaginaria, respectivamente, y la variable y toma valores a todo lo ancho de la imagen.

Si se tiene una imagen de 256×256 como la mostrada en la figura 7, entonces Z_{RP} es el perfil binario positivo de la sección transversal de la parte real $T_R(129, y)$, donde $c_x = 129$ y $1 \leq y \leq 256$. Z_{IP} es el perfil binario de la sección transversal de la parte imaginaria $T_I(129, y)$, donde $c_x = 129$ y $1 \leq y \leq 256$. En la figura 17a y 17b se muestran los perfiles real e imaginario tomando un umbral positivo de los perfiles centrales.

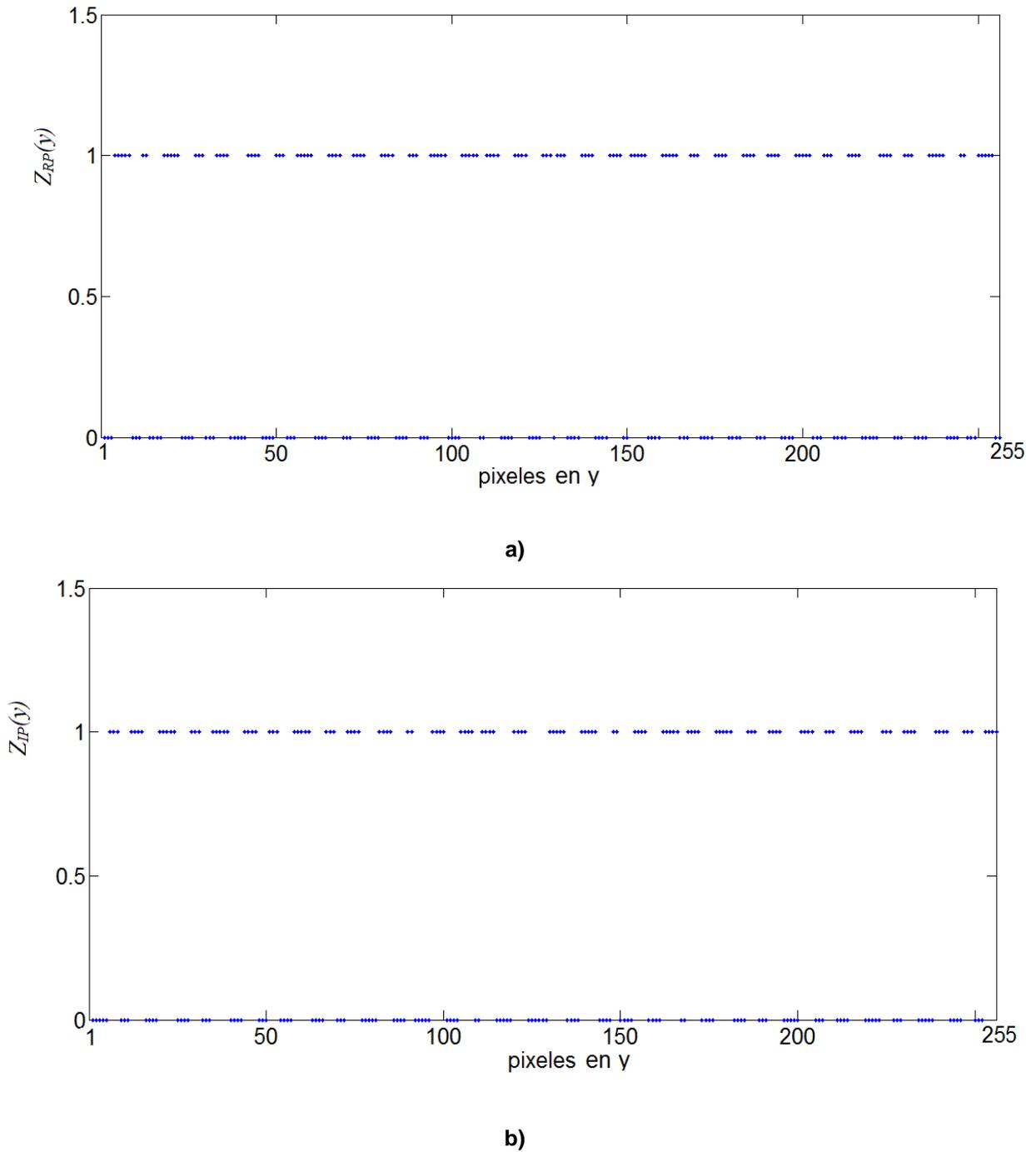


Figura 17. a) Perfil binario de la parte real de la figura 10a usando un umbral positivo. b) Perfil binario de la parte imaginaria 10b usando un umbral positivo.

Posteriormente se toma $c_x = 129$ como eje para hacer una rotación de 180° del perfil binario para obtener cilindros concéntricos; los cuales tendrán una altura de 1.

Estos cilindros concéntricos tendrán diferente grosor y estarán centrados en (c_x, c_x) ; éstos se muestran en la figura 18.

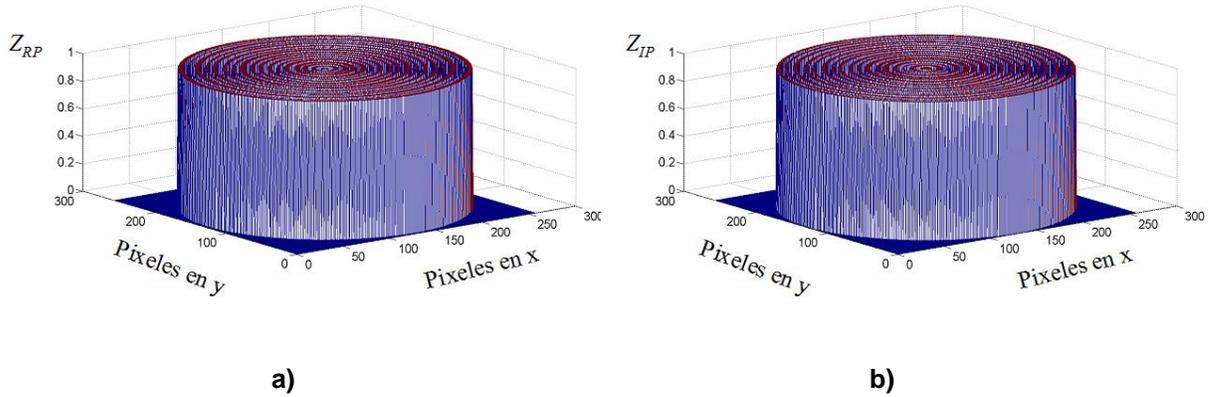


Figura 18. Cilindros concéntricos producidos al rotar los perfiles binarios, parte real (a), parte imaginaria (b).

Mapeando la información tridimensional a un plano bidimensional, se puede observar una máscara binaria bidimensional. En la figura 19a se observan la máscara binaria producida por el perfil binario de la parte real aplicando un umbral positivo, llamada MBRP y en la figura 19b se observa la máscara producida por el perfil binario de la parte imaginaria aplicando un umbral positivo, llamada máscara MBIP.

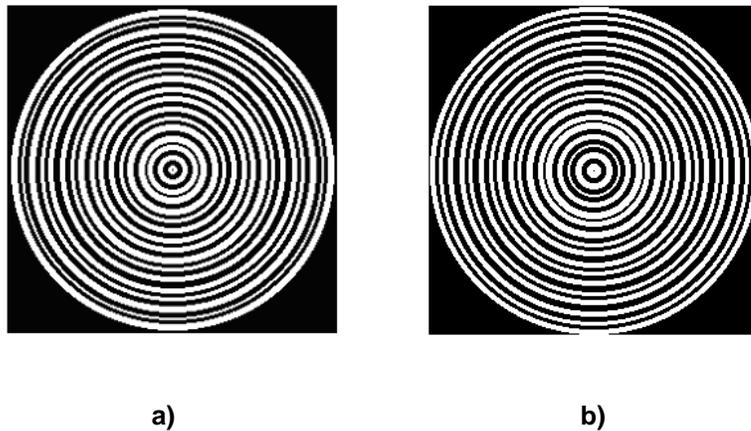


Figura 19. a) MBRP. b) MBIP.

También se puede utilizar un umbral no positivo, como se muestra en las ecuaciones (74) y (75), para la formación de máscaras. En la figura 20a se muestran las máscaras obtenida aplicando un umbral no positivo al perfil central de la parte real (MBRNP), y

en la figura 20b se muestra la máscara obtenida aplicando un umbral no positivo al perfil central de la parte imaginaria (MBINP). Por lo tanto, para una imagen se pueden generar cuatro diferentes máscaras, MBRP, MBRNP, MBIP y MBINP, dependiendo si se trabaja con la parte real o imaginaria y dependiendo de cómo se aplique el umbral.

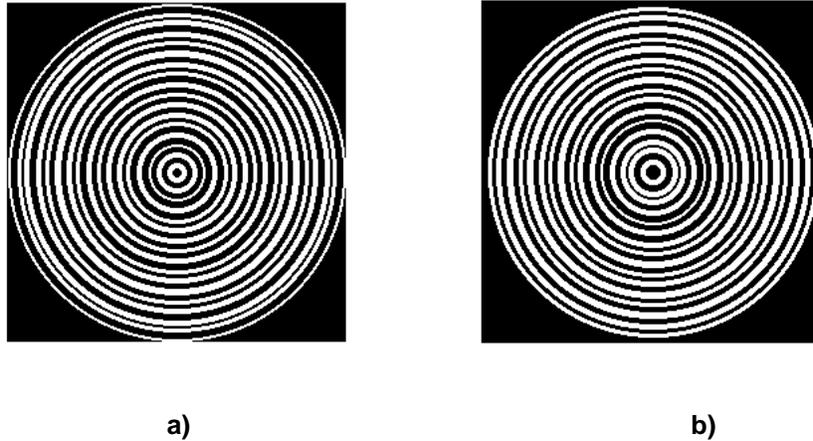


Figura 20. a) MBRNP. b) MBINP.

Una vez obtenida la máscara binaria, ésta es utilizada para multiplicarse con el módulo de la transformada de Fourier de la imagen, como se muestra en la figura 21, por ejemplo. En el paso 1 se tiene una máscara binaria real positiva, en el paso 2 se tiene el módulo de la transformada de Fourier de la letra E graficado en escala logarítmica, el cual se multiplica punto a punto con la MBRP para obtener un muestreo del módulo de la transformada de Fourier, como se muestra en el paso 3. Posteriormente se calculan valores del índice para cada anillo, cada uno de los cuales son el resultado de la suma de los valores del módulo de la transformada de Fourier muestreados por el anillo correspondiente, obteniéndose de esta manera un índice para cada anillo para formar la firma correspondiente a la imagen mostrada en el paso 4.

Para el proceso de identificación se debe obtener la firma de una imagen de referencia, así como la firma de una imagen problema que se desea analizar. La firma de la imagen problema se puede obtener utilizando la misma máscara binaria de la imagen de referencia. Obtenidas las dos firmas se procede a realizar una correlación no lineal de las firmas con un factor $k = 0.1$.

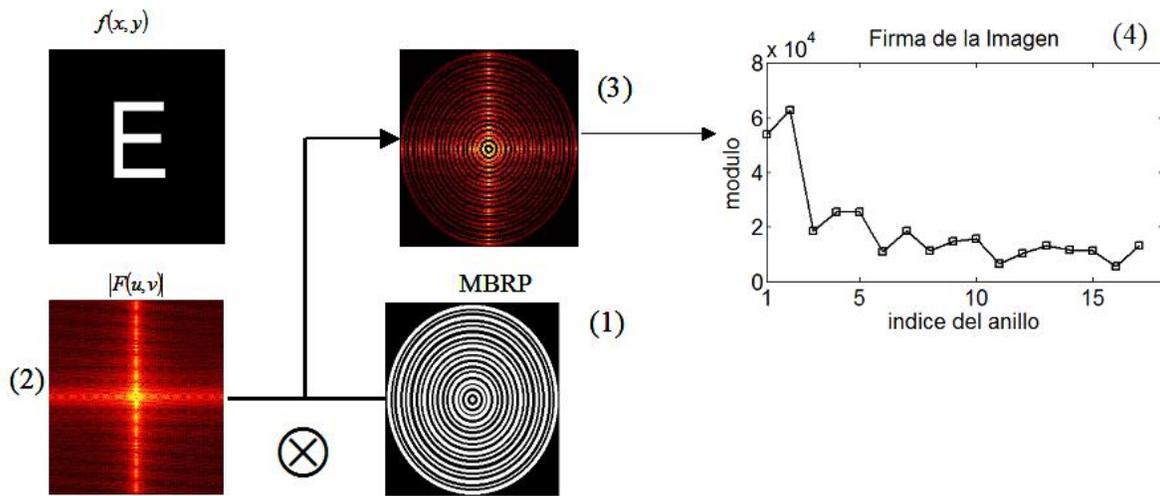


Figura 21. Procedimiento para obtener la firma de una imagen, donde \otimes significa multiplicación punto a punto.

3.4.1 Máscaras adaptativas

También se puede obtener la firma de la imagen problema utilizando una máscara obtenida de sí misma, a este procedimiento se le llama *máscaras adaptativas*, debido a que la imagen de referencia y la imagen problema tienen su propia máscara binaria.

Una vez obtenidas las firmas para la imagen de referencia e imagen problema se procede a realizar una correlación no lineal con un factor $k = 0.1$.

3.4.2 Utilización del perfil óptimo

En el mundo real no sabemos la rotación que tienen las imágenes, por lo que una variación a la metodología de máscaras binarias es el de buscar el perfil óptimo, el cual será representado por el perfil con mayor energía.

Para una imagen $N \times N$ dada por la función bidimensional $f(x,y)$, el pixel central de la imagen es el punto (c_x, c_x) . Se escoge la parte real o la parte imaginaria de la transformada de Fourier y se trazan 180 perfiles, los cuales se intersectan en el punto (c_x, c_x) y están separados por $\Delta\theta = 1^\circ$. Sin embargo, para una imagen no todos los perfiles trazados tendrán la misma longitud, por lo tanto, para tener perfiles de la

misma longitud y analizar la cantidad de energía que tiene cada perfil, la parte real (figura 15a) o imaginaria (figura 15b) de la transformada de Fourier de la imagen se filtran multiplicando por un disco binario D dado por (Solorza-Calderón y Álvarez-Borrego, 2014)

$$D = \begin{cases} 1, & \text{si } d((c_x, c_x), (x, y)) \leq c_x, \\ 0, & \text{cualquier otro caso,} \end{cases} \quad (76)$$

donde $d((c_x, c_x), (x, y))$ es la distancia Euclidiana entre el pixel central (c_x, c_x) y el punto (x, y) , por lo tanto el disco D está centrado en (c_x, c_x) con un diámetro de N pixeles. En la figura 22 se muestra el disco de diámetro 256.



Figura 22. Máscara de disco binaria de 256X256.

En la figura 23a y 23b se muestra la parte real y parte imaginaria filtradas por el disco D respectivamente, este filtraje se expresa como

$$f_R = D \otimes \text{Re } \mathfrak{F}\{f(x, y)\}, \quad (77)$$

$$f_I = D \otimes \text{Im } \mathfrak{F}\{f(x, y)\}, \quad (78)$$

donde \otimes significa multiplicación punto a punto y f_R es el producto del filtraje de la parte real y f_I el producto del filtraje de la parte imaginaria.

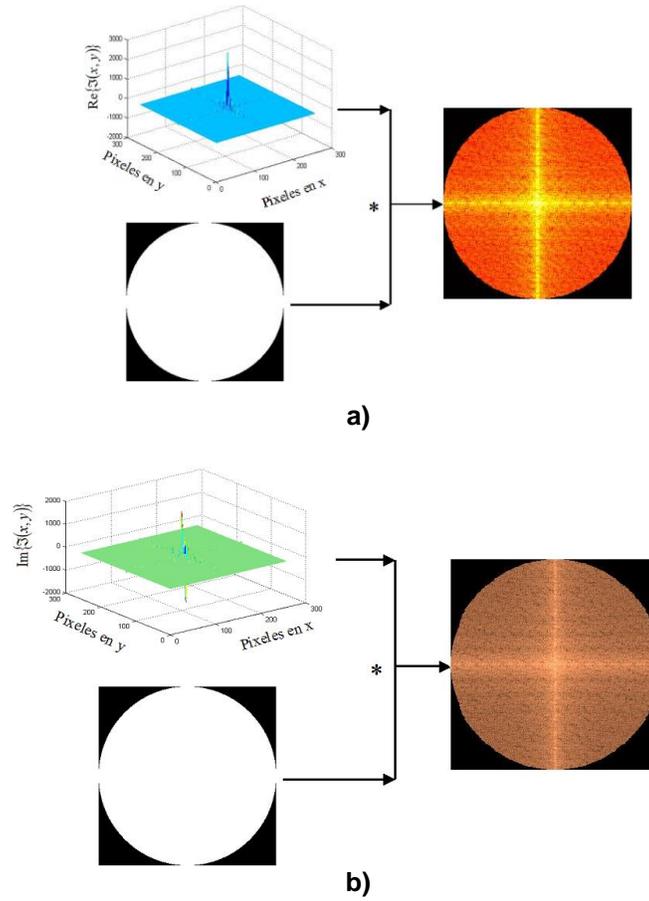


Figura 23. a) Filtrado de la parte real de la transformada de Fourier por un disco D . b) Filtrado de la parte imaginaria de Fourier por un disco D .

Una vez realizado el filtraje con el disco D , se podrán trazar 180 perfiles de la misma longitud, los cuales al estar separados $\Delta\theta=1^\circ$, estarán muestreando el disco completamente. Las ecuaciones de los perfiles para la parte real y la parte imaginaria se expresan como

$$P_R^\theta = \{f_R(d \cos\theta, d \sin\theta)\}, \quad (79)$$

$$P_I^\theta = \{f_I(d \cos\theta, d \sin\theta)\}, \quad (80)$$

donde, $\theta \in \{0, \dots, 179\}$ y $-c_x \leq d \leq c_x$ es el diámetro en pixeles. Después, se realiza la suma del cuadrado de la intensidad de cada perfil:

$$S_R^\theta = \sum_{g=1}^N (P_R^\theta)^2, \quad (81)$$

$$S_I^\theta = \sum_{g=1}^N (P_I^\theta)^2, \quad (82)$$

entonces, el perfil que tenga el valor máximo al calcular la suma, será seleccionado como el perfil óptimo.

Al perfil seleccionado se le aplica un umbral como se indica en las ecuaciones 72-75, para obtener un perfil binario y continuar el proceso para la obtención de la máscara binaria y firma de la imagen. Por lo tanto, para una imagen se pueden obtener cuatro diferentes máscaras: máscara obtenida aplicando un umbral positivo al perfil óptimo de la parte real (MBRP_PO), máscara obtenida aplicando un umbral positivo al perfil óptimo de la parte imaginaria (MBIP_PO), máscara obtenida aplicando un umbral no positivo al perfil óptimo de la parte real (MBRNP_PO) y máscara obtenida aplicando un umbral no positivo al perfil óptimo de la parte imaginaria (MBINP_PO).

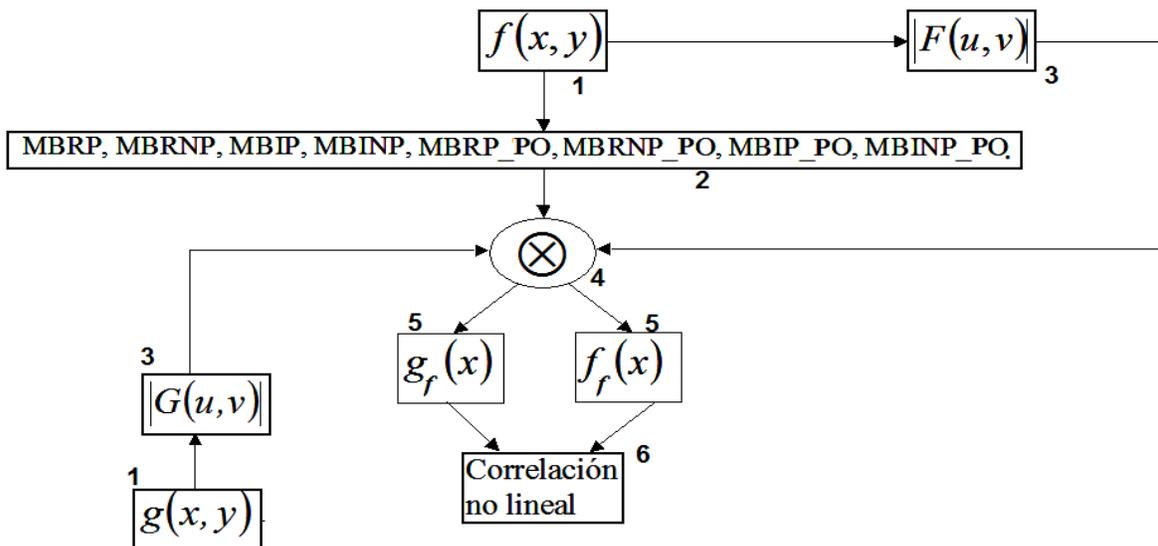


Figura 24. Metodología de máscaras binarias. Donde \otimes significa multiplicación punto a punto.

En la figura 24 se resume la metodología de máscaras binarias. En el paso 1 se tiene una imagen de referencia $f(x,y)$ y una imagen problema $g(x,y)$. En el paso 2 se

obtiene la máscara de la imagen referencia con la que se va a trabajar, donde se define si se utilizara el perfil óptimo o el perfil central, así como el umbral. En el paso 3 se calcula el módulo de la transformada de Fourier de la imagen de referencia $|F(u,v)|$ y de la imagen problema $|G(u,v)|$. En el paso 4 se realiza la multiplicación punto a punto de la máscara binaria y el módulo de la transformada de Fourier. En el paso 5 se obtiene la firma de la imagen de referencia $f_f(x)$ y la firma de la imagen problema $g_f(x)$. Finalmente, en el paso 6 se realiza la correlación de las firmas obtenidas en el paso 5.

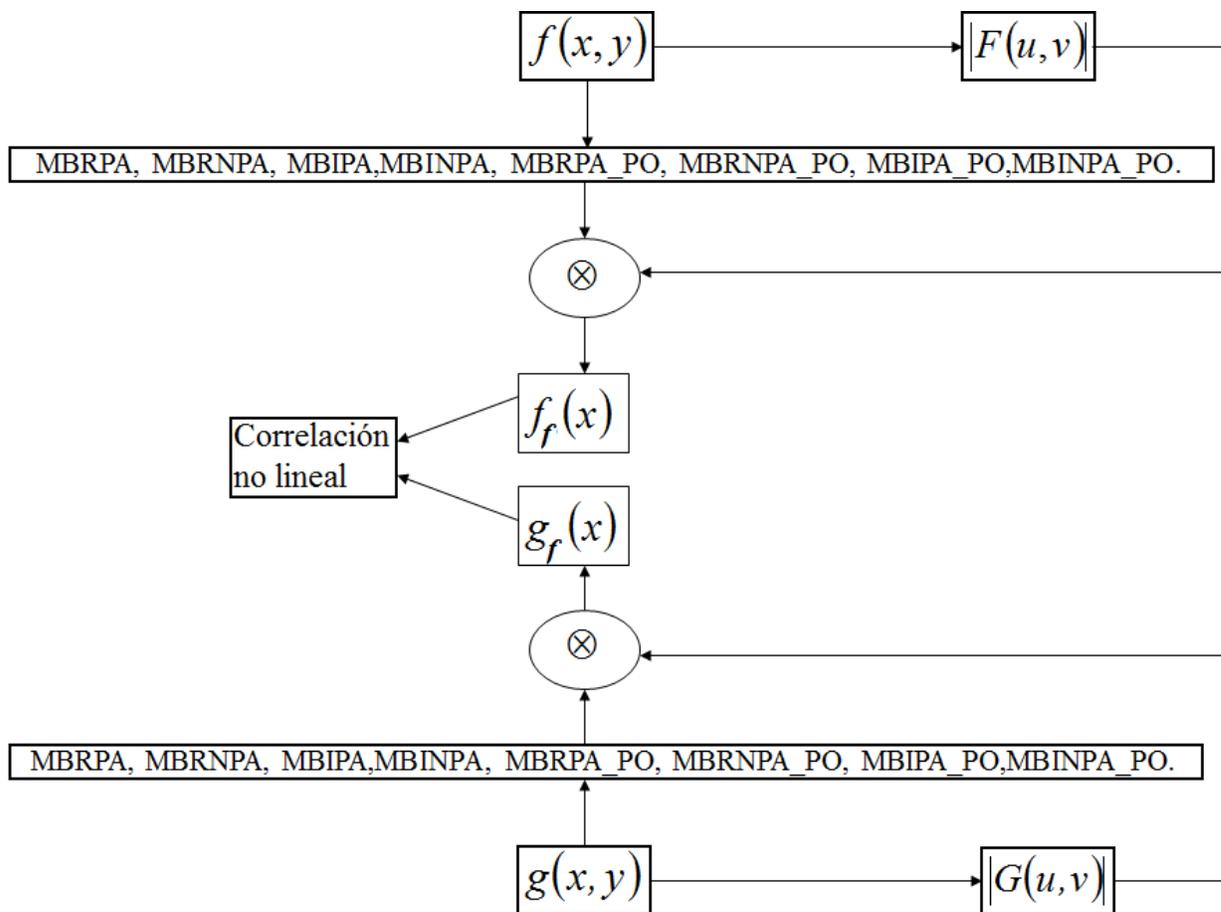


Figura 25. Metodología de máscaras binarias adaptativas. Donde \otimes significa multiplicación punto a punto.

En la figura 25 se muestra la metodología de máscaras binarias adaptativas. En el paso 1 se tiene una imagen de referencia $f(x,y)$ y una imagen problema $g(x,y)$. En el paso 2 se obtiene la máscara de la imagen referencia con la que se va a trabajar; así

mismo, se obtiene la máscara binaria correspondiente a la imagen problema, donde se define si se utilizará el perfil óptimo o el perfil central, así como el umbral para las dos máscaras. En el paso 3 se calcula el módulo de la transformada de Fourier de la imagen de referencia $|F(u,v)|$ y de la imagen problema $|G(u,v)|$. En el paso 4 se realiza la multiplicación punto a punto de las máscaras binarias y el módulo de las transformadas de Fourier de las imágenes. En el paso 5 se obtiene la firma de la imagen de referencia $f_f(x)$ y la firma de la imagen problema $g_f(x)$. Finalmente, en el paso 6 se realiza la correlación de las firmas obtenidas en el paso 5.

3.5 Comparación de los diferentes métodos

Se realizó un análisis comparativo de las metodologías anteriormente descritas, para ello, se trabajó con 21 imágenes de diferentes especies de diatomeas, las cuales se muestran en la figura 26.

3.5.1 Evaluación y comparación de PCE

Aquí se analizaron las metodologías de correlación no lineal con transformada de escala, firmas vectoriales y máscaras binarias.

El objetivo de este análisis es determinar cuál de las metodologías nos brinda una respuesta menos ruidosa. Para llevar a cabo este análisis se utiliza la métrica llamada PCE, la cual nos permite saber que tan ruidoso es el plano de correlación.

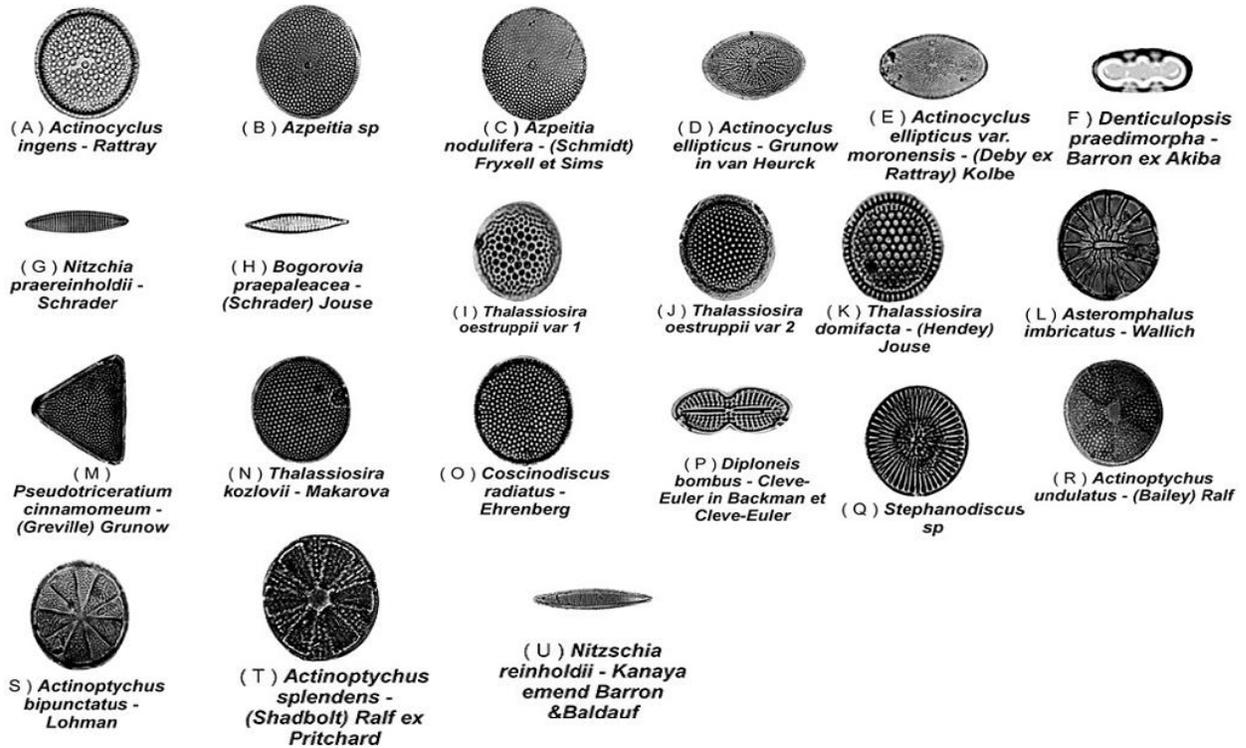


Figura 26. Imágenes de 21 especies diferentes de diatomeas fosilizadas.

Energía Pico de correlación (PCE)

Esta métrica definida por Javidi y Horner (1994), es una relación de la energía del pico de correlación y la energía promedio de todo el plano de correlación:

$$PCE = \frac{|C(c_x, c_x)|^2}{\sum_{x=1}^n \sum_{y=1}^n |C(x, y)|^2}, \quad (83)$$

donde el numerador es la energía del pico de correlación y el denominador es el valor de la energía promedio en el plano de correlación.

Con la métrica PCE se puede determinar qué tan ruidosa es la respuesta de la correlación, ya que al realizar, por ejemplo, una autocorrelación, se tendrá una respuesta de intensidad que consiste solo del pico de correlación, mientras que en el resto del plano de correlación se tendrá una intensidad cero; entonces, al sumar la

energía de todo el plano solo se tendrá el valor del pico de correlación, y en tal caso, el PCE tendrá un valor de 1, lo cual indicaría una respuesta libre de ruido y conforme el ruido va aumentando en el plano de correlación, el valor del PCE comenzará a disminuir.

Metodología para la evaluación de PCE

Para el análisis, cada una de las 21 imágenes de las diferentes especies de diatomeas se escalaron de 80 a 120% en intervalos de 1% y se rotaron de 0 a 179° en intervalos de 1°, esto es, se trabajó con 41 variaciones de escala y 180 variaciones de rotación para cada una de las 21 imágenes, por lo tanto, para cada especie se obtuvieron 221 imágenes problema; así se realizó la correlación no lineal entre la imagen de cada especie con sus respectivas 221 variaciones de sí misma y se calculó el PCE de cada correlación. Por lo tanto, se realizaron 4641 correlaciones para las metodologías de correlación no lineal con transformada de escala y firmas vectoriales, con lo cual se obtuvo una media del valor de PCE para cada metodología. Para la metodología de máscaras binarias solo se consideraron las variaciones de rotación, por lo tanto, con esta metodología se realizaron 3780 correlaciones para la obtención de la media del valor de PCE.

3.5.2 Capacidad para reconocer imágenes afectadas por ruido

Para determinar la capacidad de reconocimiento, se mide la habilidad de las metodologías para distinguir cada una de las especies de diatomeas incrustadas en un fondo ruidoso. Para llevar a cabo este análisis comparativo se implementó la métrica conocida como *DC*.

Coefficiente de discriminación (DC)

Aquí se mide la habilidad de un filtro para distinguir un objeto de entre otros objetos diferentes, o cuando un objeto se encuentra incrustado en un fondo ruidoso. Está definida como

$$DC = 1 - \frac{|C_{I-R}|^2}{|C_{I-IR}|^2}, \quad (84)$$

donde C_{I-R} es el valor máximo del plano de correlación entre la imagen del objeto y la imagen de ruido y C_{I-IR} es el valor máximo del plano de correlación entre la imagen del objeto y la imagen del objeto afectada por ruido. Cuando el valor correspondiente a la correlación entre la imagen del objeto e imagen de ruido sea cero, esto es, no exista ruido, el valor de DC tendrá su valor máximo de 1; sin embargo, al ir aumentando el ruido, el valor de DC decrecerá hasta el punto en que adquirirá un valor desde cero a negativo, lo que nos indicará el valor de C_{I-R} es mucho mayor de C_{I-IR} ; esto implica que bajo esa cantidad de ruido el sistema ya no puede distinguir el objeto.

El cálculo del DC fue realizado mediante la aplicación de ruido aditivo a las imágenes de las especies de diatomeas para determinar el límite de cada metodología en su capacidad de identificación con respecto al ruido aplicado.

3.5.2.1 Imágenes afectadas por ruido gaussiano

El ruido gaussiano es uno de los principales modelos de ruido que se maneja en múltiples sistemas. El ruido Gaussiano se representa matemáticamente como

$$n_g(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad (85)$$

donde x es una variable aleatoria, σ^2 es la varianza y μ es la media.

Para cada una de las imágenes de diatomeas, se aplicó ruido aditivo gaussiano con media cero y se fue aumentando la varianza y se calculó el DC. El valor de $|C_{I-R}|^2$ se obtuvo mediante la correlación no lineal entre la imagen de la diatomea y la imagen del ruido generado al cambiar la varianza, mientras que el valor de $|C_{I-IR}|^2$ se obtuvo

mediante la correlación no lineal entre la imagen de la diatomea y la imagen de la diatomea con ruido. Mediante el cálculo del DC se determinó hasta qué valor de varianza soporta cada una de las metodologías. La figura 27 muestra el efecto del ruido gaussiano en una imagen de diatomea 256×256 píxeles en escala de grises.

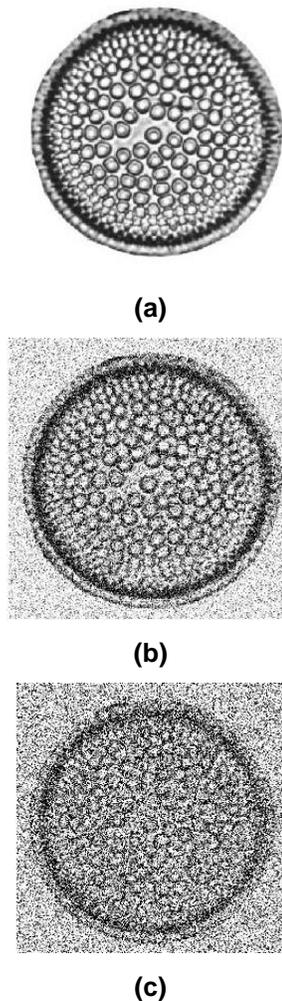


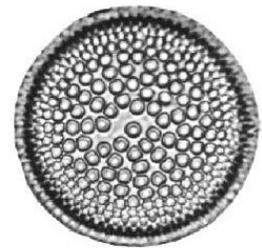
Figura 27. Imagen de la especie de diatomea *Actinocyclus Ingens-Rattray* inmersa en ruido gaussiano con media $\mu = 0$: (a) $\sigma^2 = 0$, (b) $\sigma^2 = 0.1$ y (c) $\sigma^2 = 0.5$.

3.5.2.2 Imágenes afectadas por ruido impulsivo

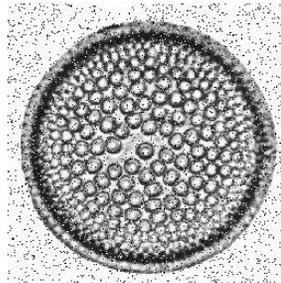
En el ruido impulsivo, o ruido sal y pimienta, el valor que toma el píxel no tiene relación con el valor de la imagen original, sino que toma valores muy altos o muy bajos. “Toma el valor máximo (sal) o el mínimo (pimienta), para una imagen de 256 niveles de gris el valor típico de la pimienta es 0 y de la sal 255. Defectos que contribuyen a este tipo de

ruido son las manchas de polvo dentro de los componentes ópticos de la cámara, o bien un sensor CCD defectuoso, que realizará una captura errónea” (González y Woods, 2002, p. 226).

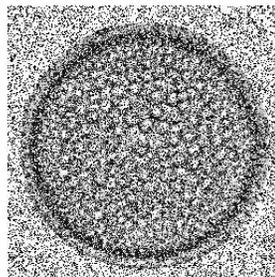
La medida del ruido está determinada por el índice de densidad d , que nos proporciona la información de píxeles afectados de la imagen, que va de 0 a 1. Cuando la densidad $d=1$, todos los píxeles están dañados, una densidad de ruido impulsivo $d=0.5$, equivale al 50% de los píxeles afectados.



(a)



(b)



(c)

Figura 28. Imagen de la especie de diatomea *Actinocyclus Ingens-Rattray* inmersa en ruido sal y pimienta: (a) $d=0$, (b) $d=0.1$ y (c) $d=0.5$.

A cada una de las imágenes de diatomeas se le aplicó ruido sal y pimienta variando el valor de la densidad y se calculó el DC, para lo cual, el valor de $|C_{I-R}|^2$ se obtuvo

mediante la correlación no lineal entre la imagen de referencia y la imagen del ruido generado al cambiar la densidad, y el valor de $|C_{I-IR}|^2$ se obtuvo mediante la correlación no lineal entre la imagen de referencia y la imagen de referencia con ruido. Mediante el cálculo del DC se determinó hasta qué valor de densidad soporta cada una de las metodologías. La figura 28 muestra el efecto del ruido sal y pimienta en una imagen de diatomea 256×256 píxeles en escala de grises.

3.5.3 Capacidad de reconocimiento con fragmentos de información

En algunas aplicaciones existe la limitante de no tener la información completa sobre el objeto a reconocer. Por ejemplo, en el caso de la identificación taxonómica de diatomeas fosilizadas que constituyen las comunidades de fitoplancton y que se utilizan como indicadores ambientales de tiempos pasados en áreas como ecología, evolución y paleoecología, se tiene la limitante de que se encuentran en muestras de sedimento en fragmentos o pulverizadas, lo cual puede producir confusión en la identificación de la especie. Por lo tanto, es de importancia el determinar la capacidad de los métodos de identificación cuando éstos trabajan con fragmentos de información para poder conocer con qué porcentaje de información son capaces de realizar una identificación.

Para este análisis se trabajó con las mismas 21 especies de diatomeas fosilizadas empleadas en el análisis de PCE, pero en este caso, de cada una de las 21 especies se cuenta con 49 fragmentos con diferente porcentaje de información, por lo tanto, se trabajó con 50 imágenes de cada una de las 21 especies, esto es, un total de 1050 imágenes. En la figura 29 se muestran las imágenes de la especie de diatomea A (especie *Actinocyclus Ingens-Rattray*) con sus 49 fragmentos y en la figura 30 la diatomea B (especie *Azpeitia sp.*).

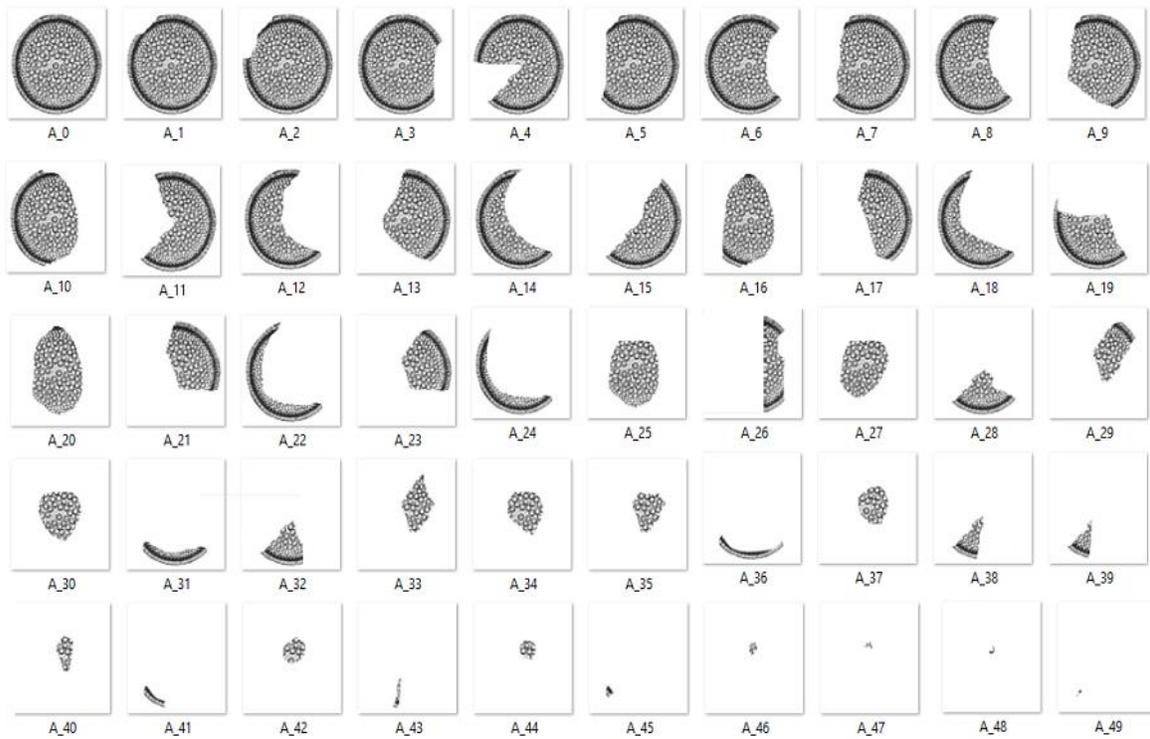


Figura 29. Imagen de la especie de diatomea *Actinocyclus Ingens-Rattray* y sus 49 fragmentos.

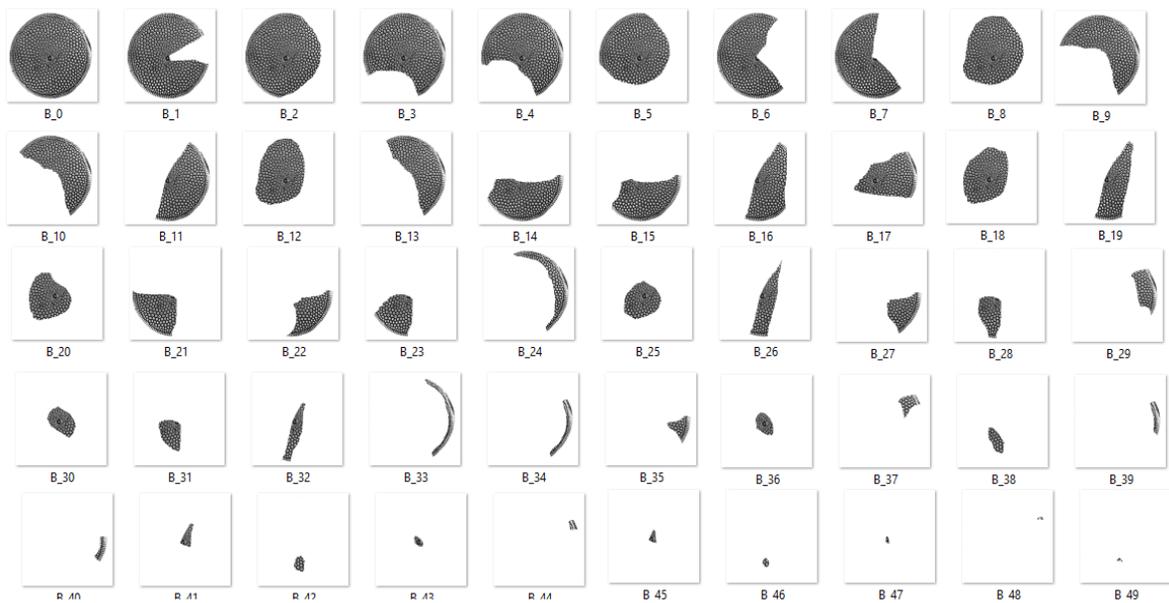


Figura 30. Imagen de la especie de diatomea especie *Azpeitia sp* y sus 49 fragmentos.

Para cada una de las metodologías de correlación se realizó el análisis. Este análisis consistió en que para una determinada especie se realiza la correlación entre la imagen de la diatomea entera seleccionada con cada uno de sus 49 fragmentos, así como con el resto de las imágenes enteras de las demás especies y sus respectivos fragmentos.

Los resultados de estas correlaciones se graficaron como se muestra en la figura 31, donde en color azul se presenta el resultado de la metodología de correlación no lineal y transformada de escala de la imagen entera de la especie *Actinocyclus Ingens-Rattray* con las 1050 imágenes, las cuales se muestran en el eje x (eje horizontal) de la gráfica, donde el fragmento 1 a 50 corresponde a la diatomea A y sus 49 fragmentos, y del fragmento 51 al 60 corresponde a la especie B y sus respectivos fragmentos, y así sucesivamente cada 50 valores en el eje de las x corresponde a cada una de las 21 especies de diatomeas. Además, cada uno de los 49 fragmentos de la especie de diatomea seleccionada se correlacionó con su imagen entera y sus demás fragmentos, así como con el resto de imágenes enteras y fragmentadas de las otras especies. Se muestra en color rojo el resultado de la correlación del fragmento No. 27 de la diatomea A con las 1050 imágenes problema, donde se puede ver que el valor de correlación entre este fragmento y la imagen entera de la diatomea es de 4.197×10^7 . Este resultado se presenta en $x=1$, el cual es superior al valor en el eje y (eje vertical) presentado para $51 \leq x \leq 1050$, que corresponde a los valores de correlación con fragmentos de otras especies. Esto implica que al estar por arriba del resto de los valores no existe confusión de una especie con otra.

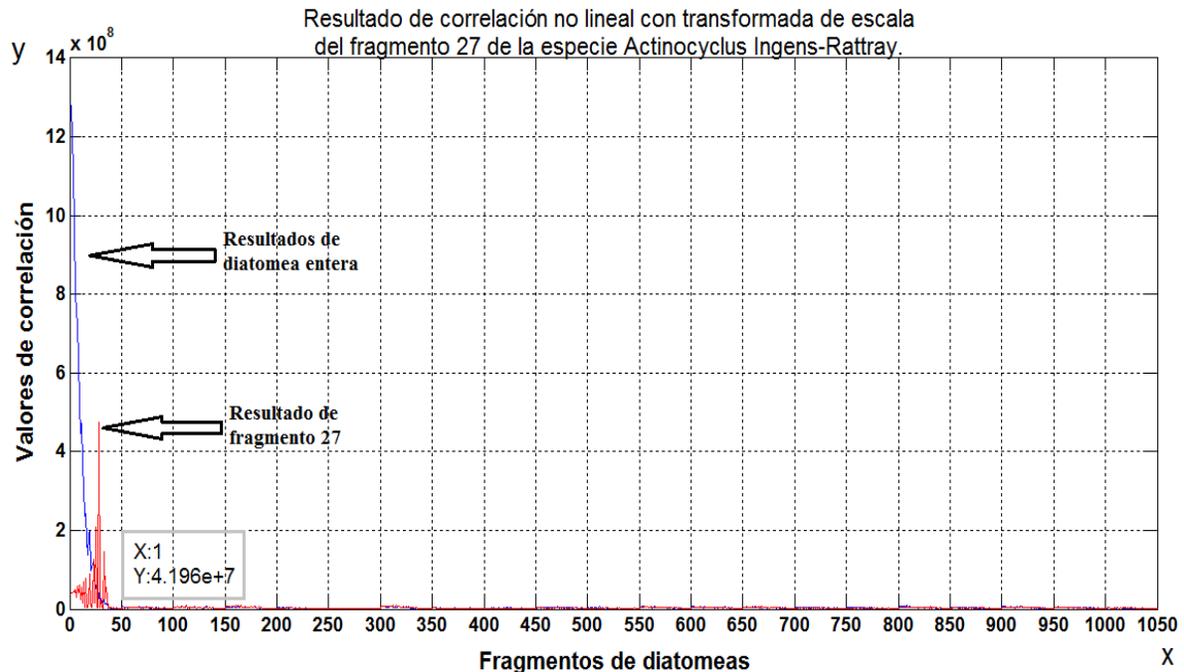


Figura 31. Resultados de correlación de la imagen entera de la especie A (color azul) con las 1050 imágenes problema, y resultado de correlación del fragmento 27 de la especie A (color rojo) con las 1050 imágenes problema.

Mediante este análisis se determina que utilizando el fragmento No. 27 como imagen de referencia, la metodología de correlación no lineal con transformada de escala es capaz de distinguir la especie A del resto de las 21 especies de diatomeas.

Así, para cada especie se realizaron correlaciones de sus 50 imágenes (entera y 49 fragmentos) con el total de 1050 imágenes problema, esto es, para cada especie se realizaron 52500 correlaciones.

3.5.4 Tiempo de ejecución

Se determinó el tiempo de ejecución de cada metodología al calcularse el PCE considerando el procesamiento de 21 especies de diatomeas con 180 variaciones de rotación desde 0 a 179 grados en intervalos de 1 grado y 41 variaciones de escala entre 80% y 120% en intervalos de 1%, para las metodologías de correlación no lineal con transformada de escala y firmas vectoriales, mientras que para la metodología de máscaras binarias solo se consideraron las variaciones de rotación. En el caso de la

metodología de índices espectrales, al no utilizarse correlación, se tomó en cuenta el tiempo de despliegue de los resultados en el plano de identificación, considerando tanto variaciones de rotación y de escala.

Todos los procesos fueron realizados en una computadora personal con procesador Intel Core i7-3770 de 3.40 GHz de velocidad y 8GB (7.84GB utilizable) de memoria RAM.

Capítulo 4. Resultados

En este capítulo se presentan los resultados del análisis comparativo descrito en el capítulo 3.

4.1. Resultados de PCE

El análisis de PCE se realizó para las metodologías de correlación no lineal con transformada de escala (CNLTE), firmas vectoriales de rotación (FVR) y firmas vectoriales de escala (FVE), así como la metodología de máscaras binarias en sus diversas variantes. Todas estas metodologías realizan correlación no lineal, mientras que en el caso de la metodología de índices espectrales no se realiza correlación, por esta razón, esta última metodología mencionada no se considera en este análisis.

Para cada una de las metodologías se calculó el PCE de la correlación de las imágenes de cada especie con sus 221 variaciones. Para ello, al realizar la correlación no lineal se utilizó un valor de k óptimo correspondiente para cada método. Para la correlación no lineal con transformada de escala se utilizó un valor de $k = 0.3$, para el caso de las firmas vectoriales de rotación y de escala se utilizó un valor de $k = 0.3$ y para el caso de las máscaras binarias un valor de $k = 0.1$. Una vez obtenidos, para cada metodología, los 221 resultados de PCE de cada especie, se calculó la media y se realizó una gráfica de cajas tomando en cuenta \pm dos errores estándar, para comparar así las diferentes metodologías

En la figura 32 se muestra el resultado estadístico de PCE para la diatomea A que corresponde a la especie *Actinocyclus Ingens-Rattray*. Se puede observar de estos resultados que la mejor respuesta de PCE la obtuvieron el grupo de metodologías de máscaras binarias. Sin embargo, no se tiene la misma respuesta con el grupo de metodologías de máscaras adaptativas, las cuales tienen una respuesta inferior a la de las firmas vectoriales de escala y firmas vectoriales de rotación. También se observa que la peor respuesta de PCE lo muestra la metodología de correlación no lineal con transformada de escala (CNTE).

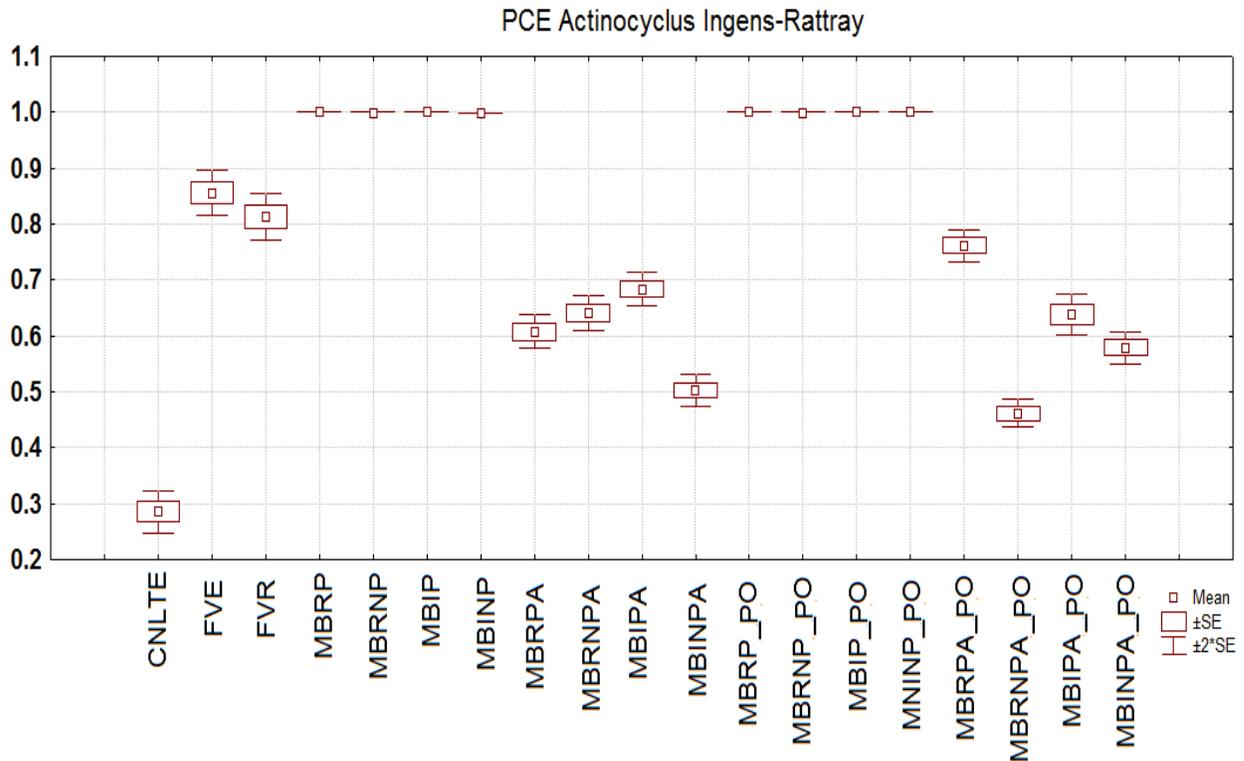


Figura 32. Gráfica de cajas de los resultados de PCE para la diatomea A, donde se muestra la media así como ± 1 error estándar (SE) y ± 2 errores estándar ($2*SE$).

En la Tabla 1 se presentan los valores de la media de los resultados de PCE para cada una de las metodologías, donde se puede ver que la metodología de máscaras binarias real positiva por medio de perfil óptimo (MBRP_PO) es la que tiene un valor de media más cercano a 1.

En la figura 33 se tienen los resultados estadísticos de PCE para la diatomea F (especie *Denticulopsis Praedimorpha - Barron ex Akiba*). Se puede observar que, de nuevo, el grupo de metodologías de máscaras binarias presentan la mejor respuesta de PCE, lo que implica que estas metodologías producen una respuesta menos ruidosa en el plano de correlación. Igualmente que en el resultado para la diatomea A, se observa que al utilizarse el grupo de metodologías de máscaras binarias adaptativas el PCE disminuye, esto es, el plano de correlación es más ruidoso al utilizarse esta variante de máscaras binarias. También, el resultado de CNLTE es la metodología que presenta la peor respuesta de PCE.

Tabla 1. Media de PCE para *Actinocyclus Ingens-Rattray*.

Metodología	μ de PCE
CNLTE	0.285395743
FVE	0.855750992
FVR	0.812423049
MBRP	0.999833581
MBRNP	0.999741755
MBIP	0.999918165
MBINP	0.999426934
MBRPA	0.607386424
MBRNPA	0.640568418
MBIPA	0.684085911
MBINPA	0.502368227
MBRP_PO	0.999984801
MBRNP_PO	0.999652179
MBIP_PO	0.999896108
MBINP_PO	0.999975278
MBRPA_PO	0.761266987
MBRNPA_PO	0.461756146
MBIPA_PO	0.637547063
MBINPA_PO	0.57869921

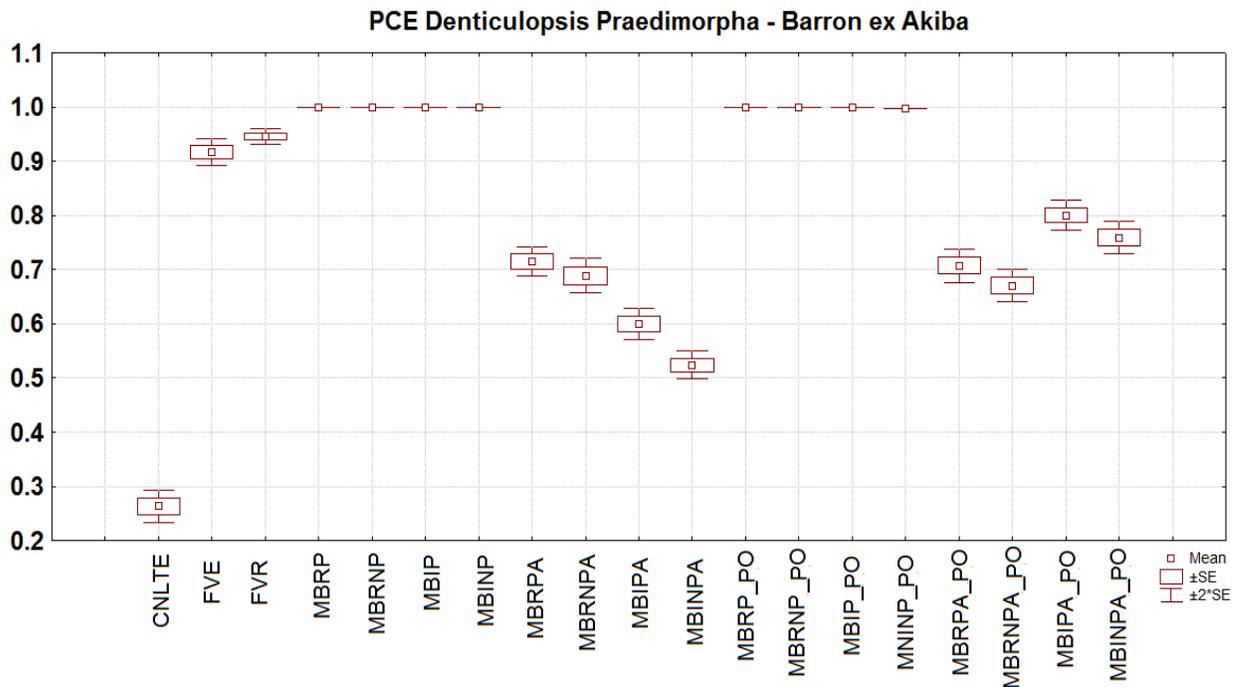


Figura 33. Gráfica de cajas de los resultados de PCE para la diatomea F, donde se muestra la media así como ± 1 error estándar (SE) y ± 2 errores estándar ($2 \cdot SE$).

Tabla 2. Media de PCE para *Denticulopsis Praedimorpha* - *Barron ex Akiba*.

Metodología	μ de PCE
CNLTE	0.263189279
FVE	0.91631037
FVR	0.945855027
MBRP	0.99988421
MBRNP	0.999956107
MBIP	0.999838588
MBINP	0.999674585
MBRPA	0.714987199
MBRNPA	0.688655296
MBIPA	0.599448843
MBINPA	0.523727482
MBRP_PO	0.999982647
MBRNP_PO	0.999199743
MBIP_PO	0.999888839
MBINP_PO	0.997797466
MBRPA_PO	0.707813644
MBRNPA_PO	0.670583244
MBIPA_PO	0.800317514
MBINPA_PO	0.759328033

En la Tabla 2 se muestra los valores de la media del PCE para cada una de las metodologías para la especie *Denticulopsis Praedimorpha* - *Barron ex Akiba*, donde se puede ver que, de nueva cuenta, la metodología de máscaras binarias real positiva obtenida con el perfil óptimo (MBRP_PO) tiene una media de PCE más alta y más cercana a 1, el cual es el valor deseado de PCE, dado que implica una salida en el plano de correlación libre de ruido y más claramente se puede observar el pico de correlación.

En la figura 34 se presentan los resultados estadísticos de PCE para la diatomea P (especie *Diplonei bombus* - *Cleve*) para todas las metodologías. Se puede observar que similarmente a los resultados de las otras dos especies presentadas, las metodologías de máscaras binarias son las que presentan la mejor respuesta de PCE, asimismo, se observa que al utilizarse máscaras binarias adaptativas el PCE disminuye, y de nueva cuenta, se aprecia que la metodología de CNTE es la que presenta valores de PCE más bajos.

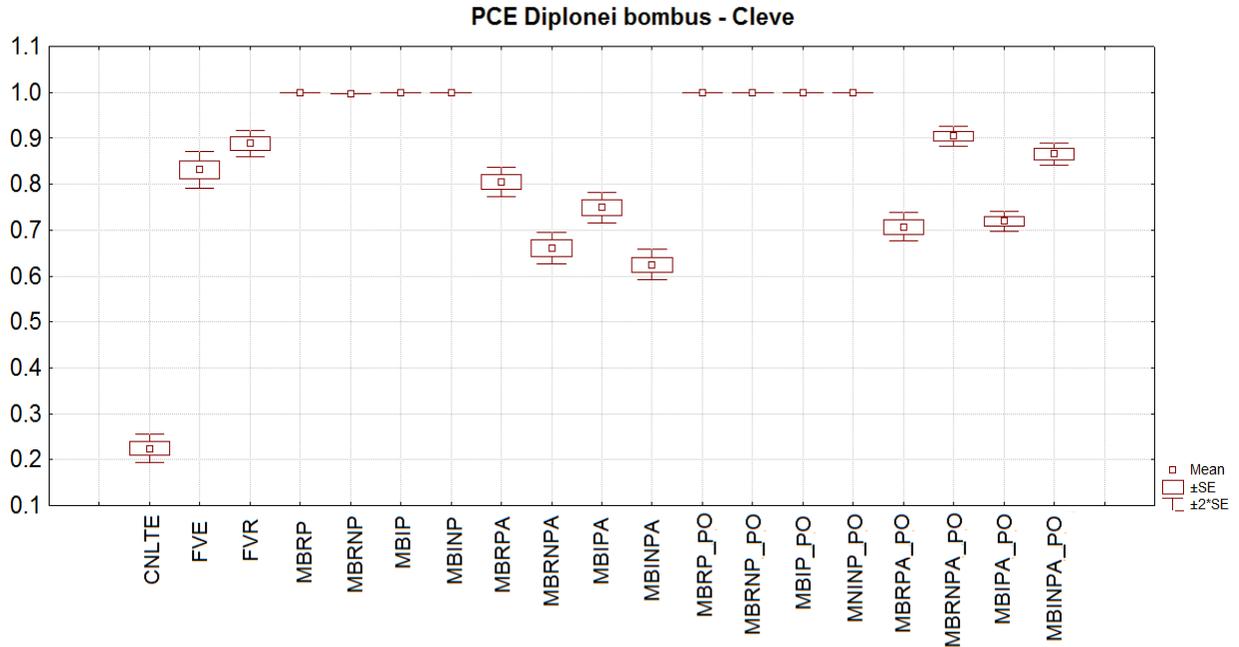


Figura 34. Gráfica de cajas de los resultados de PCE para la diatomea P, donde se muestra la media, así como ± 1 error estándar (SE) y ± 2 errores estándar ($2 \cdot SE$).

Tabla 3. Media de PCE para *Diplonei bombus* - Cleve.

Metodología	μ de PCE
CNLTE	0.224800444
FVE	0.831363255
FVR	0.889004092
MBRP	0.999901686
MBRNP	0.997566396
MBIP	0.999992169
MBINP	0.99963656
MBRPA	0.804479555
MBRNPA	0.660718871
MBIPA	0.749307964
MBINPA	0.624920268
MBRP_PO	0.999128136
MBRNP_PO	0.999976696
MBIP_PO	0.99903996
MBINP_PO	0.999239354
MBRPA_PO	0.706793463
MBRNPA_PO	0.904758402
MBIPA_PO	0.719591967
MBINPA_PO	0.865904343

En la Tabla 3 se muestra los valores de la media del PCE para cada una de las metodologías para la especie *Diplonei bombus* – Cleve, donde se puede ver que en esta ocasión, la metodología de máscaras binarias imaginaria positiva obtenida con el perfil central (MBIP) tiene la media de PCE más alta.

En los resultados anteriormente presentados se puede ver que el grupo de metodologías de máscaras binarias, son las que presentan una respuesta menos ruidosa, esto es, debido a sus altos valores de PCE se puede diferenciar con mayor claridad el pico de correlación con respecto al resto del plano de correlación. Esto debido a que se utiliza la misma máscara binaria para la imagen de referencia e imagen problema, la cual muestrea los mismos componentes de frecuencia en el módulo de la imagen de referencia e imagen problema.

Estos resultados también nos indican que la metodología que tiene una menor claridad en su respuesta es la de correlación no lineal con transformada de escala, ya que es la que en los resultados de las tres especies presentadas es la que tiene valores de PCE más bajos, lo cual indica un plano de correlación muy ruidoso, lo que hace que sea más difícil diferenciar el pico de correlación resultante del resto del plano de correlación.

4.2 Índices espectrales

La metodología de índices espectrales no fue parte del análisis de PCE, debido a que no utiliza correlación, ya que esta metodología nos presenta los resultados de todas las imágenes procesadas en un solo plano de salida y no es necesario realizar una comparación entre una imagen de referencia y una imagen problema. Por lo tanto, para evaluar esta metodología, se procesaron las 21 imágenes de diatomeas, así como sus 42 variaciones de escala y 180 variaciones de rotación, y se realizó una gráfica de cajas dentro del plano de salida.

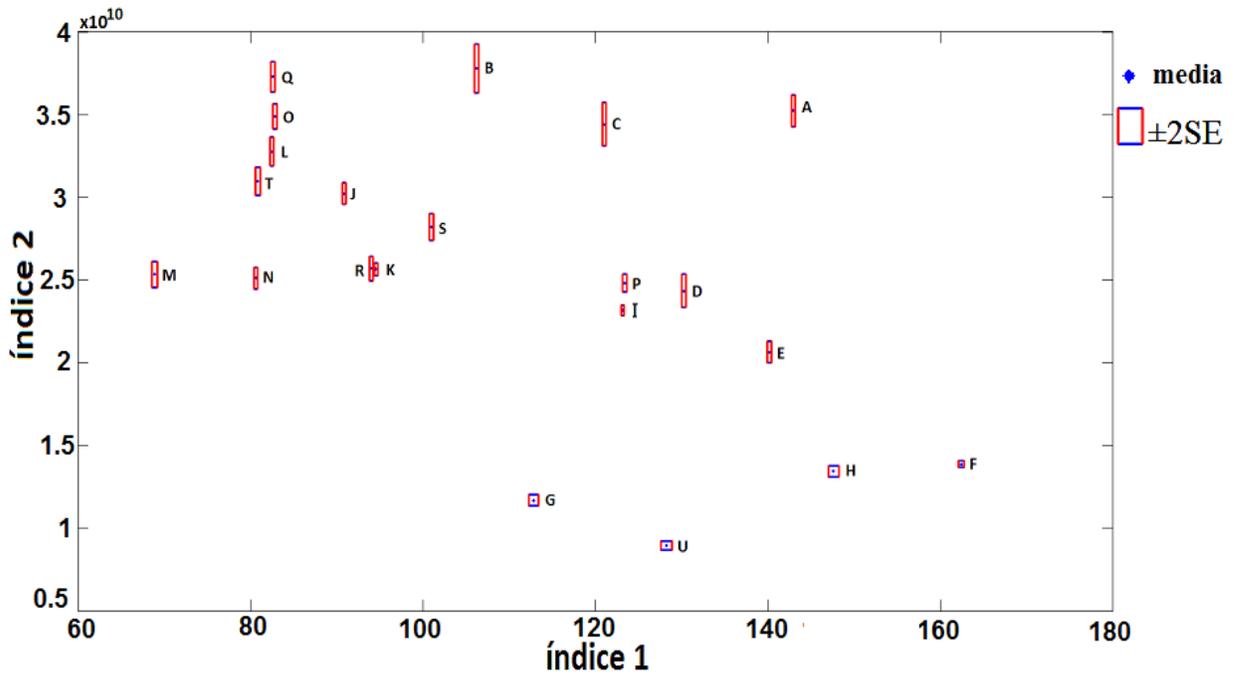


Figura 35. Resultado estadístico de índices espectrales para las 21 especies de diatomeas, donde se presenta la media de cada índice, así como $\pm 2SE$.

El resultado de este proceso se presenta en la figura 35, donde se muestra en un solo plano de identificación el resultado estadístico del cálculo de los índices espectrales de las 21 especies de diatomeas, donde el eje horizontal corresponde al índice 2 y el eje vertical corresponde al índice 1. Para realizar esta gráfica se calculó la media del índice 2 y la media del índice 1, asimismo, se calculó el error estándar y se graficó más-menos dos errores estándar ($\pm 2SE$) de cada uno de los índices espectrales y se graficaron en forma de caja, de manera que se delimita un espacio dentro del plano de identificación que corresponde al 95.4% de confianza para cada índice espectral de cada una de las 21 especies. Se puede ver que el resultado de la diatomea R y la diatomea K son muy cercanos, sin embargo, al realizar un acercamiento a esa zona, como se muestra en la figura 36, se observa que no existe traslape entre estas dos diatomeas. De esta manera, se puede ver que a cada especie le corresponde un espacio exclusivo dentro del plano de identificación definido por $\pm 2SE$.

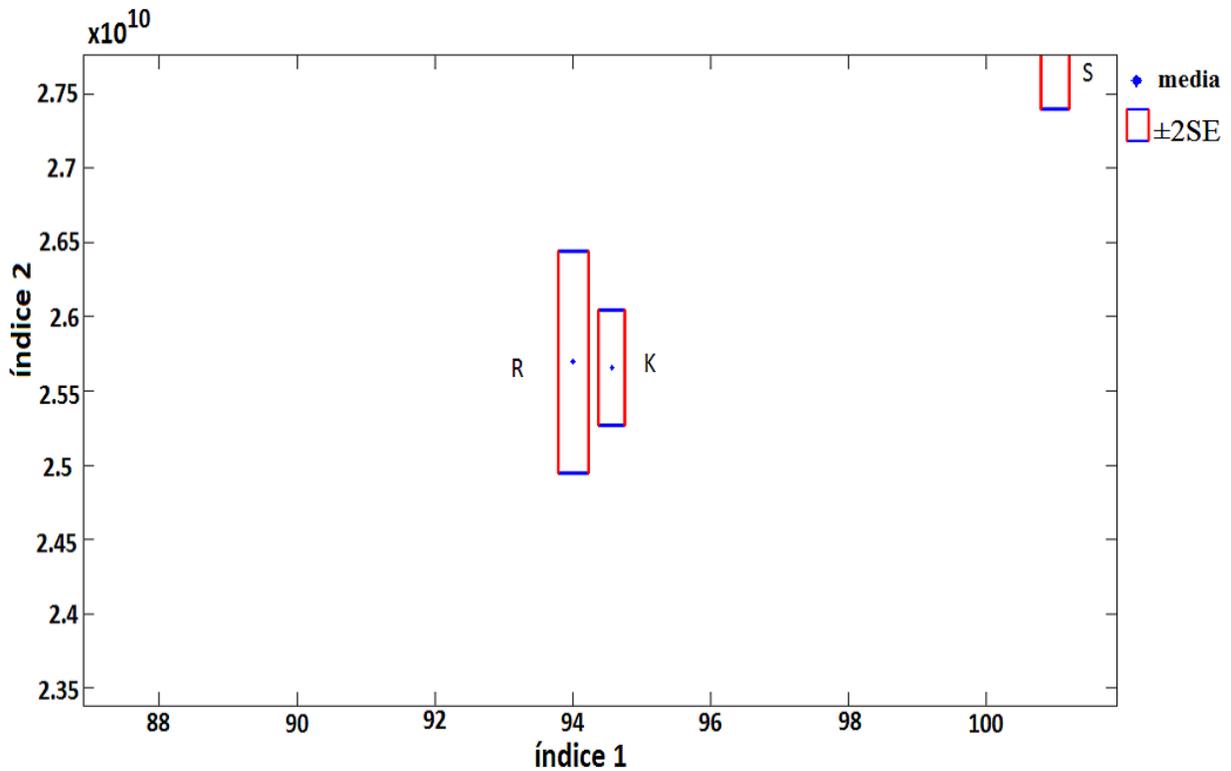


Figura 36. Acercamiento a los resultados estadísticos de índices espectrales de diatomea R y diatomea K.

4.3 Tiempo de ejecución

Para el análisis, cada una de las 21 imágenes de las diferentes especies de diatomeas se escalaron de 80 a 120% en intervalos de 1% y se rotaron de 0 a 179° en intervalos de 1° para cada una de las 21 imágenes, por lo que para cada especie se obtuvieron 221 variaciones, así se realizó el cálculo del PCE para cada especie. Por lo tanto, se realizaron 4641 procesamientos para las metodologías de correlación no lineal con transformada de escala y firmas vectoriales. En el caso de máscaras binarias, debido a que es una metodología que es invariante solo a rotación, no se tomaron en cuenta las variaciones de escala en su análisis, por lo cual se realizaron 3780 procesamientos para esta metodología. Para la metodología de índices espectrales, se calculó el tiempo en que se procesaron las 4661 imágenes y se desplegaron los resultados en el plano de identificación. En la Tabla 4 se muestra los resultados del tiempo de ejecución de cada metodología. En la segunda y en la tercera columna se muestra el

tiempo de ejecución en segundos y en minutos al procesar todas las imágenes y en la última columna se presenta el tiempo correspondiente a una sola imagen.

Tabla 4. Tiempo de ejecución.

Metodología	Tiempo total en segundos	Tiempo total en minutos	Tiempo por imagen en segundos
CNLTE	1134.1562	18.9026	243.3x10 ⁻³
Firmas vectoriales	174.55871	2.9093	37.5x10 ⁻³
MBRP	108.76093	1.8127	28.8x10 ⁻³
MBRNP	101.0695	1.6845	26.7x10 ⁻³
MBIP	110.5006	1.8417	29.2x10 ⁻³
MBINP	103.22249	1.7204	27.3x10 ⁻³
MBRPA	679.87830	11.3313	179.9x10 ⁻³
MBRNPA	753.07069	12.5512	199.2x10 ⁻³
MBIPA	784.09967	13.0683	207.4x10 ⁻³
MBINPA	704.30789	11.7385	186.3x10 ⁻³
MBRP_PO	131.27138	2.1879	34.7x10 ⁻³
MBRNP_PO	128.91804	2.1486	34.1x10 ⁻³
MBIP_PO	114.56566	1.9094	30.3x10 ⁻³
MBINP_PO	112.5696	1.8762	29.8x10 ⁻³
MBRPA_PO	963.36109	16.0560	254.9x10 ⁻³
MBRNPA_PO	950.62948	15.8438	251.5x10 ⁻³
MBIPA_PO	951.22791	15.8538	271.9x10 ⁻³
MBINPA_PO	9.0076520	15.0128	238.3x10 ⁻³
índices espectrales	183.08694	3.0514	39.3x10 ⁻³

Se puede ver por los resultados que el grupo de metodologías de máscaras binarias presentan los mejores resultados en el tiempo de ejecución, ya que la duración de su proceso fue inferior a 35 milésimas de segundo por imagen, requiriéndose un poco más de tiempo en el caso en el que se trabaja con el perfil óptimo, debido al tiempo que lleva el encontrar éste, sin embargo la mayor diferencia es de solo 7.4 milésimas entre la metodologías MBRNP y MBRNP_PO. También los resultados muestran que el grupo de metodologías de máscaras binarias adaptativas tienen un mayor tiempo de ejecución, teniendo el mayor tiempo de ejecución en el caso en el que se calcula la máscara a partir del perfil óptimo, solo la metodología de MBINPA_PO tiene un tiempo menor a la CNTE, mientras que todas las demás tienen un mayor tiempo de ejecución, siendo la metodología MBIPA_PO la que mayor tiempo tiene de todas las metodologías evaluadas con 271.9 milésimas de segundo por imagen.

4.4 Resultados de la capacidad de reconocimiento con fragmentos de información

Aquí se presentan los resultados del análisis con imágenes de fragmentos de diatomeas con la intención de determinar el porcentaje de información requerida para que una determinada metodología pueda realizar una identificación correcta.

En la figura 37 se muestra el resultado de la metodología de correlación con transformada de escala para el fragmento 35 de la diatomea A, correspondiente a la especie *Actinocyclus Ingens-Rattray*, donde se puede ver el resultado de la correlación de este fragmento con la imagen entera de su misma especie, y que se muestra en $x=1$, tiene un valor de $y=1.148 \times 10^7$, mientras que la correlación de este mismo fragmento con el fragmento $x=311$, que corresponde a la especie G, tiene un valor $y=1.496 \times 10^7$, lo cual no es un resultado deseado, ya que se desea que los valores de la correlación con fragmentos de otras especies sea inferior a los valores de correlación de su misma especie. Esto significa que al llegar al fragmento 35 ya no se puede obtener una identificación correcta, ya que se confunde un fragmento de la especie G con los de la especie A. De esta manera se tiene que esta metodología puede obtener resultados correctos hasta el fragmento 34.

Dado que la imagen entera de la diatomea A, sin tomar en cuenta el fondo, tiene un número de píxeles de 42415 y el fragmento 34 de esta especie tiene un número de píxeles de 5087, el fragmento 34 contiene un 11.99% de la información total. Por lo tanto, la metodología de correlación no lineal con transformada de escala puede identificar esta especie teniendo el 11.99% de la información total.

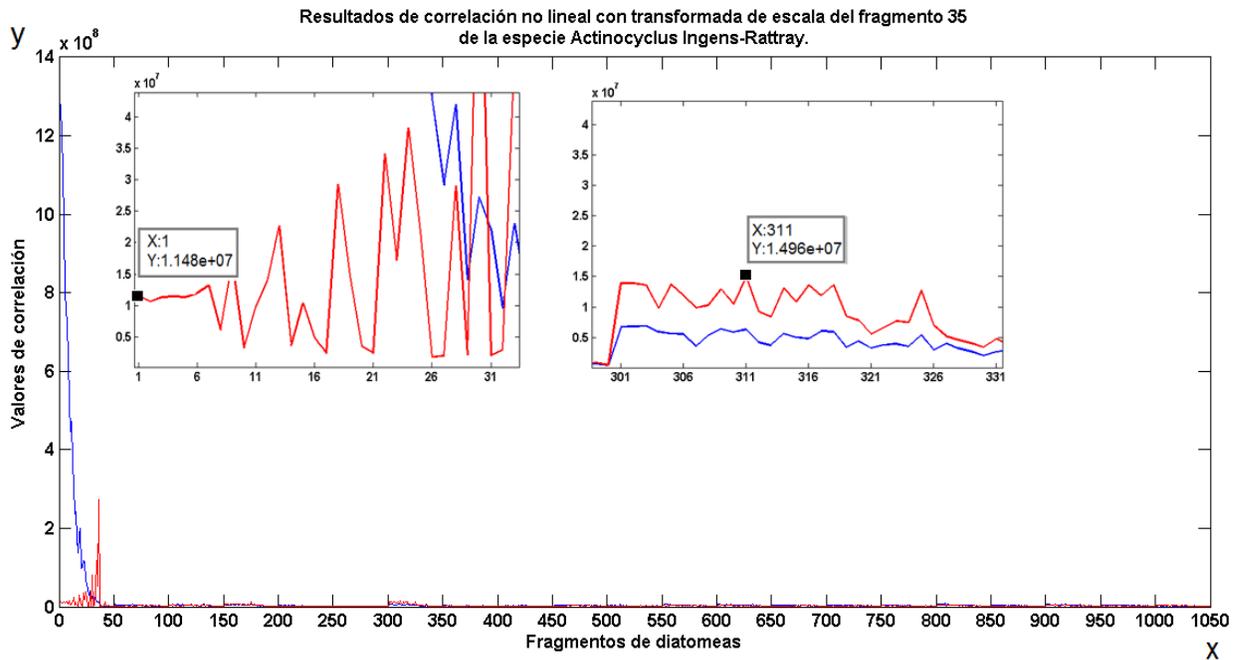


Figura 37. Resultados de correlación de la imagen entera de la diatomea A (color azul) con las 1050 imágenes problema y resultado de correlación del fragmento 35 de la especie A (color rojo) con las 1050 imágenes problema. En el recuadro de la izquierda se muestra un acercamiento al fragmento $x=1$ y en el recuadro de la derecha un acercamiento al fragmento 311.

En la figura 38 se muestra el resultado de la metodología de correlación no lineal con transformada de escala para el fragmento 30 de la diatomea B, correspondiente a la especie *Azpeitia sp*, donde se puede ver el resultado de la correlación de este fragmento con la imagen entera de su misma especie, y que se muestra en $x=51$, tiene un valor de $y=2.853 \times 10^6$, mientras que la correlación de este mismo fragmento con el fragmento $x=111$, que corresponde a la especie C, tiene un valor $y=3.08 \times 10^6$. Esto significa que la utilización de este fragmento como referencia para la identificación mediante esta metodología no produce resultados correctos, al confundirse un fragmento de la especie C con los de la especie B. Por lo tanto, esta metodología puede obtener resultados correctos hasta el fragmento 29.

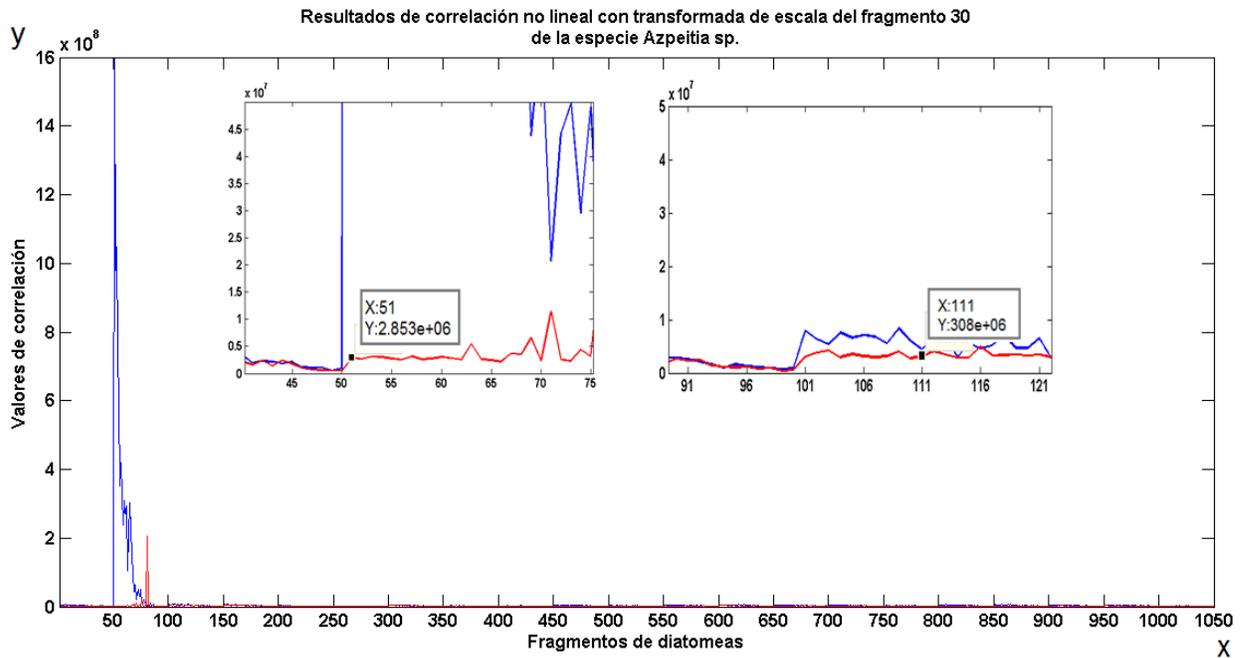


Figura 38. Gráfica de resultados de correlación de la imagen entera de la diatomea B (color azul) con las 1050 imágenes problema y resultado de correlación del fragmento 30 de la especie B (color rojo) con las 1050 imágenes problema. En el recuadro de la izquierda se muestra un acercamiento al fragmento $x=51$ y en el recuadro de la derecha un acercamiento al fragmento 111.

En el caso de la metodología de firmas vectoriales, para poder considerar un resultado como positivo al momento de realizar una identificación, es necesario tener una identificación positiva al realizar la correlación de las firmas vectoriales de escala, así como de la correlación de las firmas vectoriales de rotación simultáneamente. En caso de que solo se tenga una identificación positiva en el resultado de solo una de las firmas, se considera que la identificación es negativa. Por lo tanto, en este análisis, para un determinado fragmento se calcularon las correlaciones de las firmas vectoriales de escala y las correlaciones de las firmas vectoriales de rotación de dicho fragmento y todas las 1050 imágenes. Posteriormente se procedió a realizar la multiplicación punto a punto de los dos resultados, debido a que si se tiene una identificación positiva en ambas firmas, al realizar la multiplicación punto a punto el pico de correlación se mantendrá, ya que se encontrará localizado en el mismo punto en el plano de correlación de cada firma vectorial.

En la figura 39 se muestra el resultado de las correlaciones de las firmas de escala de la imagen entera de la diatomea A con los 1050 fragmentos; en la figura 40 se muestra

el resultado de las correlaciones de las firmas de rotación para la misma imagen entera de la diatomea A y los 1050 fragmentos, y en la figura 41 se muestra el resultado de la multiplicación de los dos resultados. Dado que se están utilizando las firmas de la imagen entera de la diatomea para realizar las correlaciones, se espera que cuando se correlaciona con ella misma, esto es, cuando se realiza la autocorrelación, el valor de correlación sea superior a la correlación con las firmas de cualquier otro fragmento; sin embargo se puede ver que esto no ocurre en los resultados de cada firma por separado. Por ejemplo, en los resultados de la firma de escala mostrados en la figura 39 se observa que el valor de la autocorrelación de la firma de la imagen entera de la diatomea A que se encuentra en $x=1$ tiene un valor de 6.369×10^{12} , mientras que para el fragmento $x=51$, que corresponde a la diatomea B, se tiene un valor de correlación de $y=7.764 \times 10^{12}$. En el caso de las firmas de rotación de la figura 40 se tiene que el valor para $x=1$ es de 3.268×10^{11} , mientras que para $x=261$, que corresponde a la diatomea F, el valor de correlación es 3.752×10^{11} .

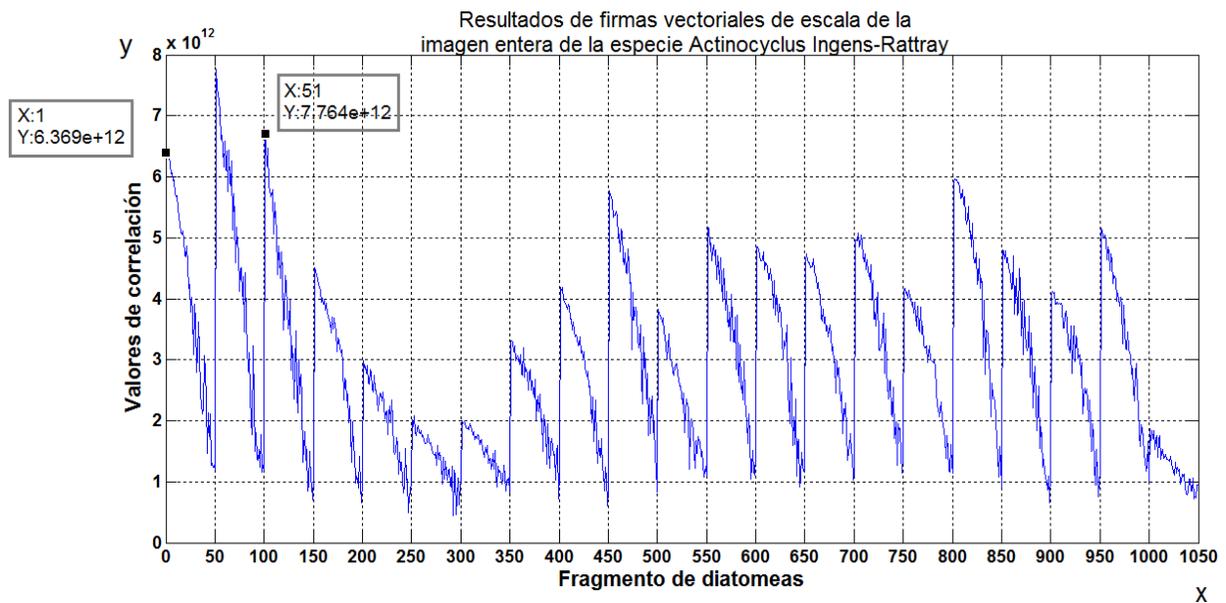


Figura 39. Gráfica de resultados de correlación de firmas vectoriales de escala de la imagen entera de la diatomea A con las 1050 imágenes problema.

Los picos falsos no ocurren en el mismo lugar en ambos resultados, por lo que al realizar la multiplicación punto a punto se eliminan, permaneciendo como pico máximo el correspondiente a $x=1$, el cual es el que está presente en ambos resultados por

separado, como es requerido para determinar una identificación positiva. Esto se muestra en la figura 41.

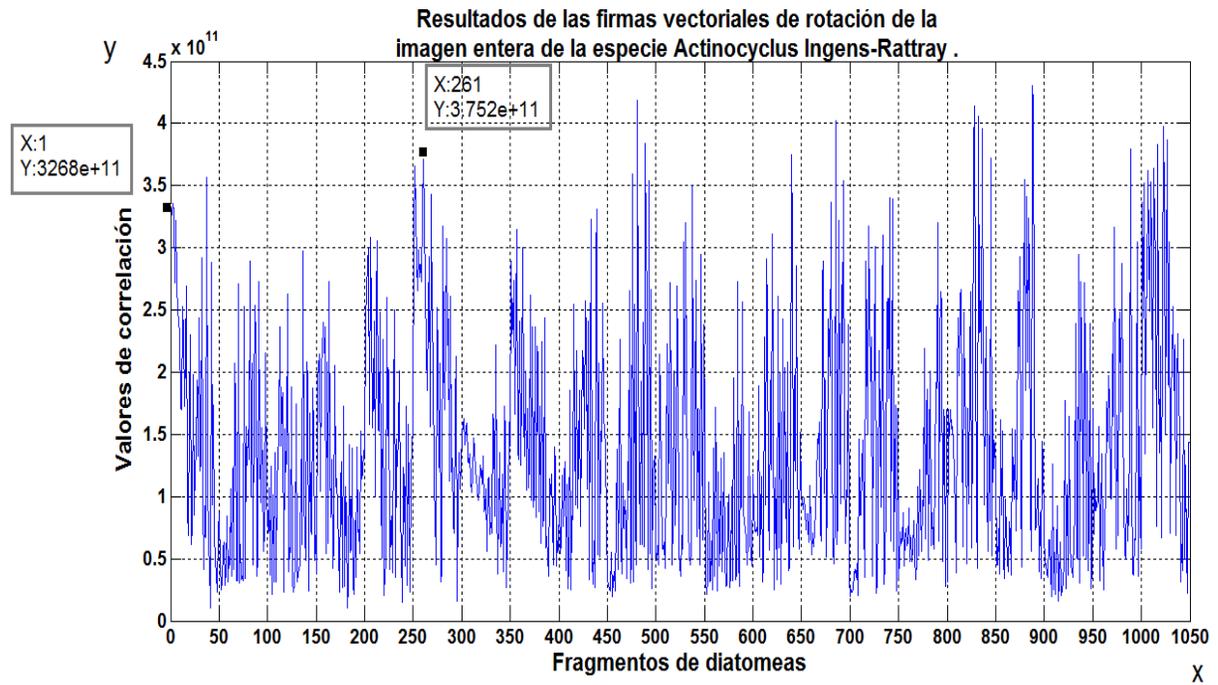


Figura 40. Gráfica de resultados de correlación de firmas vectoriales de rotación de la imagen entera de la diatomea A con las 1050 imágenes problema.

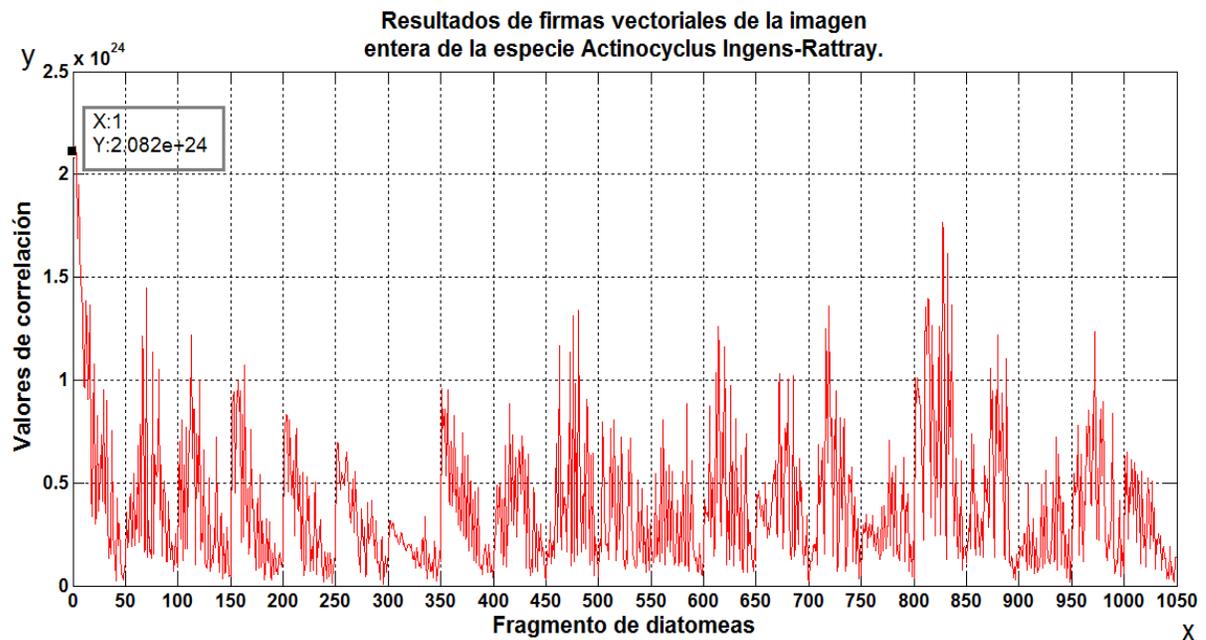


Figura 41. Gráfica de la multiplicación de resultados de correlación de firmas vectoriales de escala y firmas de rotación de la imagen entera de la diatomea A problema.

De esta manera se determinó cuál es el fragmento límite para esta metodología, el cual fue hasta el fragmento 4, lo que indica que se necesita tener el 86.4458% de la información total.

Este análisis se realizó también para la metodología de máscaras binarias y máscaras binarias adaptativas en todas sus variantes.

En la Tabla 5 se muestran los resultados correspondientes a la especie de diatomea A para todos los métodos, así como el porcentaje de información requerido por cada uno de éstos. En la Tabla 6 se muestran los resultados correspondientes a la especie de diatomea B para todos los métodos, así como el porcentaje de información requerido por cada uno de éstos.

Tabla 5. Fragmentos *Actinocyclus Ingens-Rattray*.

METODOLOGÍA	Número de fragmento	Porcentaje de información requerido.
CNLTE	29	13.1439 %
Firmas vectoriales	4	86.4458%
MBRP	4	86.4458%
MBRNP	7	79.7383%
MBIP	0	100%
MBINP	12	59.7784%
MBRPA	0	100%
MBRNPA	2	95.6737%
MBIPA	0	100%
MBINPA	1	98.5807%
MBRP_PO	0	100%
MBRNP_PO	0	100%
MBIP_PO	0	100%
MBINP_PO	0	100%
MBRPA_PO	0	100%
MBRNPA_PO	0	100%
MBIPA_PO	0	100%
MBINPA_PO	0	100%

De los resultados obtenidos se puede ver que la metodología de correlación no lineal con transformada de escala (CNLTE) es la que tiene la mayor capacidad de reconocimiento con información parcial. Esto es debido, quizá, a que con esta metodología, durante el proceso de correlación se trabaja con la información completa

de las imágenes, mientras que en las demás metodologías, firmas vectoriales y máscaras binarias se correlacionan vectores unidimensionales. Esto significa que durante el proceso de estas metodologías no se utiliza toda la información de la imagen procesada, por lo tanto, de antemano ya existe pérdida de información al aplicar estas metodologías al análisis con fragmentos de información, lo cual no permite que tengan la misma capacidad que la metodología CNLTE.

Tabla 6. Fragmentos *Azpeitia sp.*

METODOLOGÍA	Número de fragmento	Porcentaje de información requerido.
CNLTE	34	11.99%
Firmas vectoriales	2	86.4458%
MBRP	6	68.4952%
MBRNP	16	38.0604%
MBIP	10	50.1741%
MBINP	0	100%
MBRPA	0	100%
MBRNPA	0	100%
MBIPA	0	100%
MBINPA	0	100%
MBRP_PO	0	100%
MBRNP_PO	0	100%
MBIP_PO	2	86.4458%
MBINP_PO	0	100%
MBRPA_PO	0	100%
MBRNPA_PO	0	100%
MBIPA_PO	0	100%
MBINPA_PO	0	100%

4.5 Resultados de la capacidad para reconocer imágenes afectadas por ruido

En este análisis se utilizó como referencia la imagen de la diatomea A correspondiente a la especie *Actinocyclus Ingens-Rattray*, la cual se afectó con ruido gaussiano y ruido impulsivo, o sal y pimienta, para determinar la respuesta de cada metodología ante ruido. Para el ruido gaussiano se modificó la varianza y se calculó el coeficiente de discriminación de acuerdo a la ecuación (84). Este cálculo se repitió 30 veces para

poder así obtener la media μ de los resultados de DC, así como su error estándar (SE), y crear una gráfica de la $\mu \pm 1.96SE$ del coeficiente de discriminación que corresponden al 95% de nivel de confianza. Este mismo proceso se realizó para el ruido impulsivo, pero en este caso modificando su densidad.

La figura 42 muestra el resultado de la media del coeficiente de discriminación para la metodología de correlación no lineal con transformada de escala (CNLTE), para ruido gaussiano de media cero y donde se va incrementando la varianza desde 0 a 0.3 en intervalos de 0.01. En color negro se muestra la media del coeficiente de discriminación y en color rojo se indica $\pm 1.96SE$, para 30 iteraciones para correlación no lineal con $k=0.3$. Se observa que la respuesta se vuelve negativa cuando la varianza es de 0.01, esto es, no es capaz de detectar la imagen de la diatomea A bajo estas condiciones de ruido.

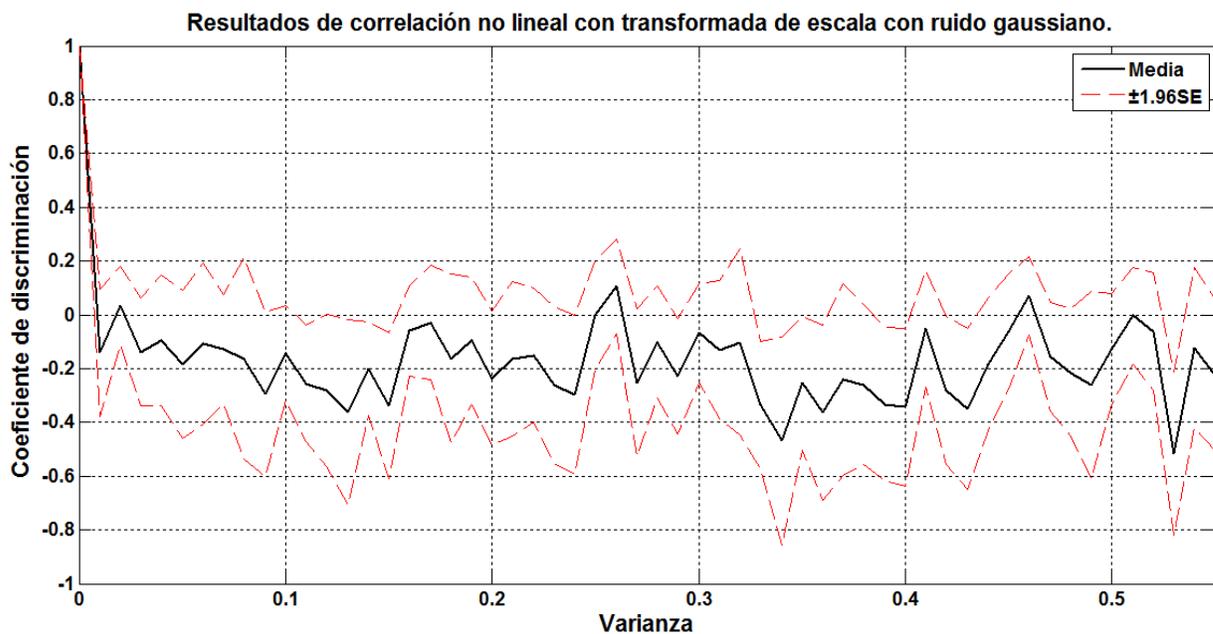


Figura 42. Gráfica del promedio de los resultados del coeficiente de discriminación en color negro y $\pm 1.96SE$ en rojo, con 30 iteraciones para CNLTE con ruido gaussiano.

En la figura 43 se muestra el resultado con ruido impulsivo para la misma metodología. Se muestra en negro la media del coeficiente de discriminación, y en rojo, $\pm 1.96SE$ de 30 iteraciones variando la densidad entre 0 y 0.3 en intervalos de

0.01. Se puede observar que la media de la respuesta del coeficiente de discriminación se vuelve negativa para 0.01, lo cual nos indica que la metodología de correlación no lineal con transformada de escala no es capaz de reconocer la imagen de la diatomea A cuando está afectada con ruido impulsivo.

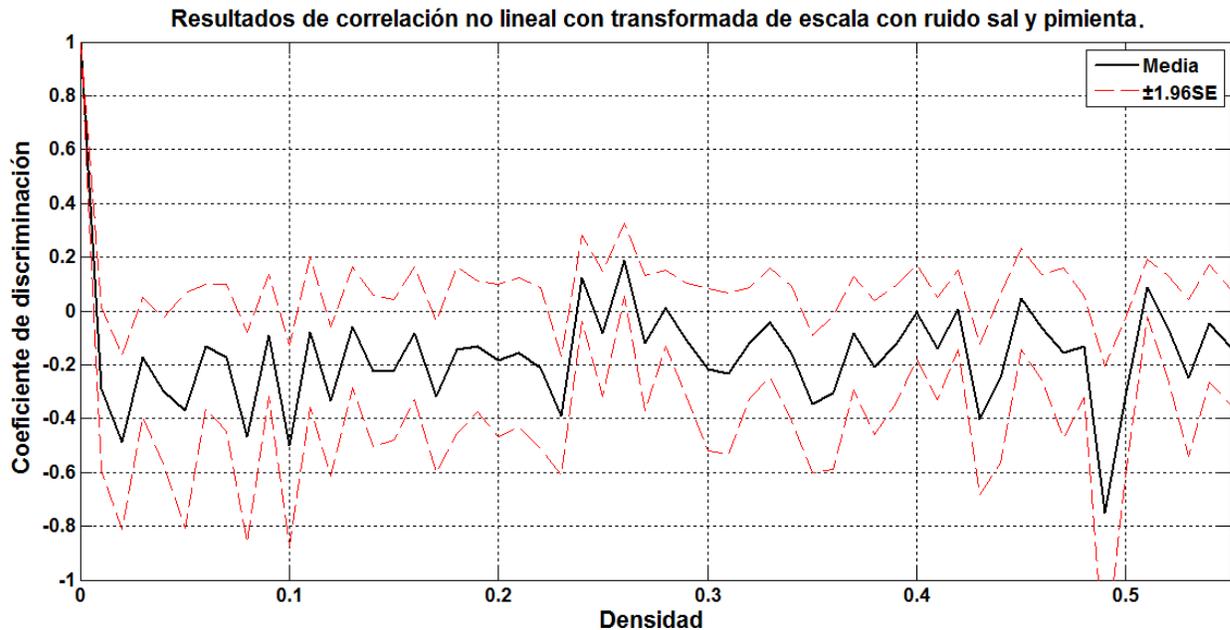


Figura 43. Gráfica del promedio de los resultados del coeficiente de discriminación en color negro y $\pm 1.96SE$ en rojo, con 30 iteraciones para CNLTE con ruido impulsivo.

En la figura 44 se muestran los resultados para el grupo de máscaras binarias utilizando el perfil óptimo, con la imagen afectada con ruido gaussiano, donde se muestra cómo varía la media del coeficiente de discriminación $\pm 1.96SE$ con respecto a la varianza con 30 iteraciones realizadas. Se puede observar que los tipos de máscaras que presentan una mejor respuesta ante la presencia de ruido gaussiano son las MBRNP_PO, en color rojo en la gráfica, y MBIP_PO, en color verde en la gráfica. La peor respuesta la muestran las máscaras binarias con umbral no positivo aplicado a la parte imaginaria (MBINP_PO), mostrado de color negro en la gráfica.

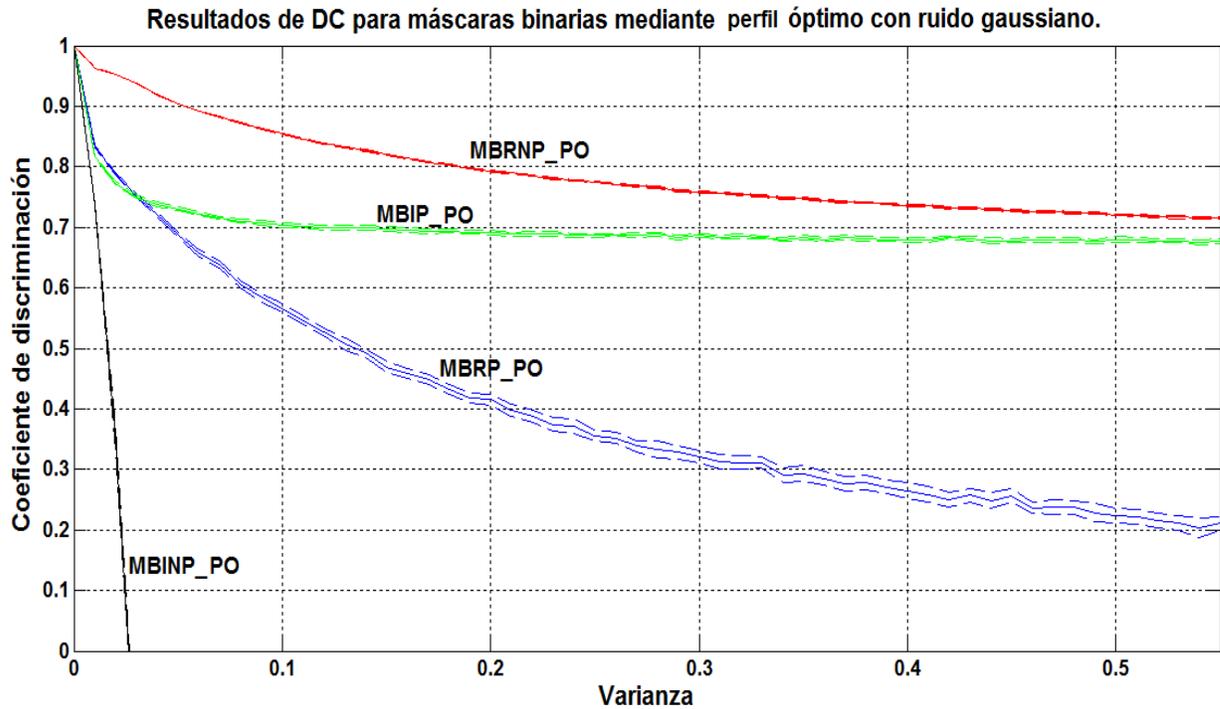


Figura 44. Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ con ruido gaussiano para máscaras binarias utilizando el perfil óptimo.

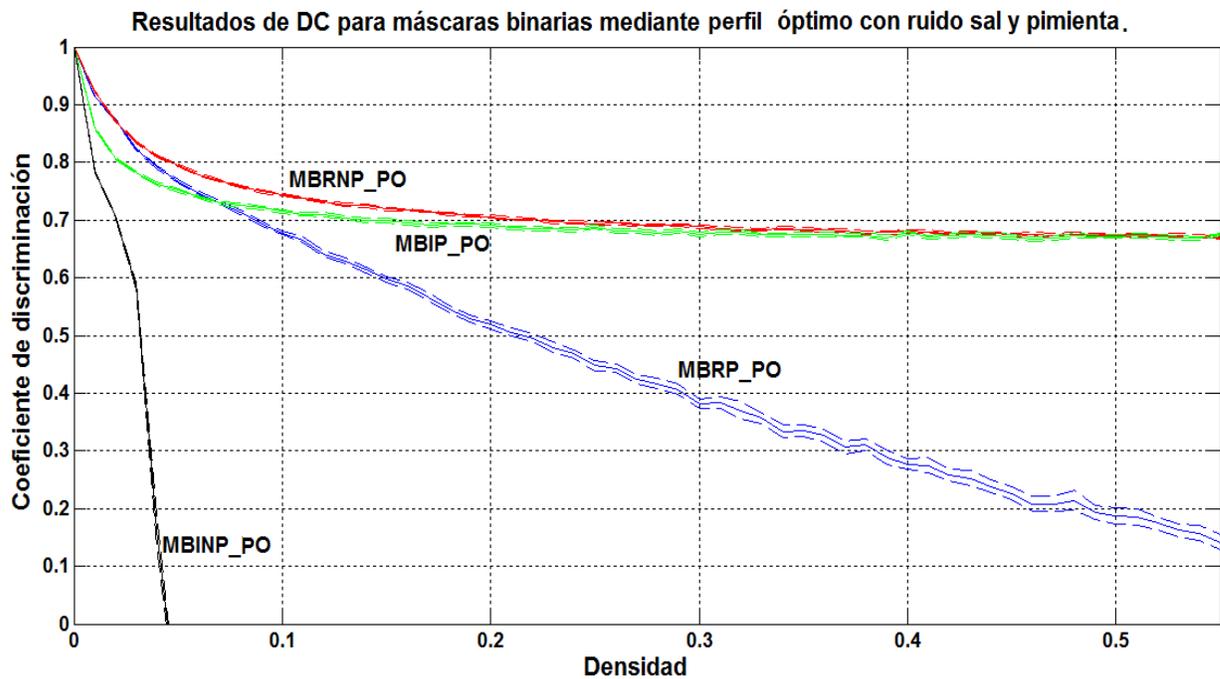


Figura 45. Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ con ruido impulsivo para la metodología de máscaras binarias utilizando el perfil óptimo.

En la imagen 45 se muestran ahora los resultados cuando la imagen es afectada con ruido sal y pimienta, donde se muestra cómo varía la media del coeficiente de discriminación $\pm 1.96SE$ con respecto a la densidad, $0 \leq d \leq 0.55$, realizando 30 iteraciones. De nueva cuenta se observa que al igual que en el caso de ruido gaussiano, también se obtiene una mejor respuesta al utilizarse MBRNP_PO, en color rojo en la gráfica y MBIP_PO en color verde en la gráfica. Asimismo, las que tienen una peor respuesta ante la presencia de ruido son las máscaras binarias con umbral no positivo aplicado a la parte imaginaria (MBINP_PO), comportamiento mostrado de color negro en la gráfica.

La figura 46 muestra la respuesta para el grupo de metodologías de máscaras binarias utilizando el perfil central con ruido gaussiano, donde MBRP está en color azul, MBRNP en color rojo, MBIP en color verde y MBINP en color negro. Se puede ver que al igual a lo que sucede al utilizar el perfil óptimo, las máscaras MBRNP y MBIP tienen mejor respuesta.

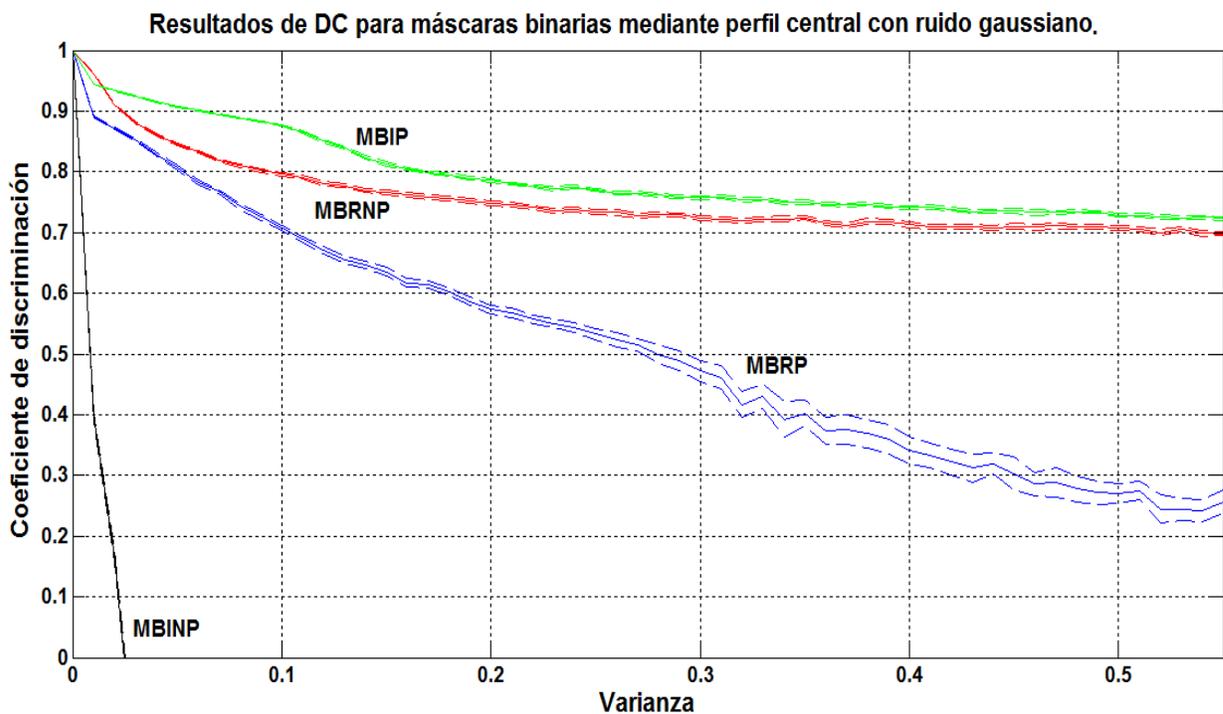


Figura 46. Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ para la metodología de máscaras binarias utilizando el perfil central con ruido gaussiano.

En la figura 47 se muestra la respuesta para el grupo de metodologías de máscaras binarias utilizando el perfil central con ruido sal y pimienta, donde MBRP está en color azul, MBRNP en color rojo, MBIP en color verde y MBINP en color negro. De nueva cuenta, se observa que al igual que en el caso del perfil óptimo, las máscaras MBRNP y MBIP tienen una mejor respuesta ante ruido impulsivo.

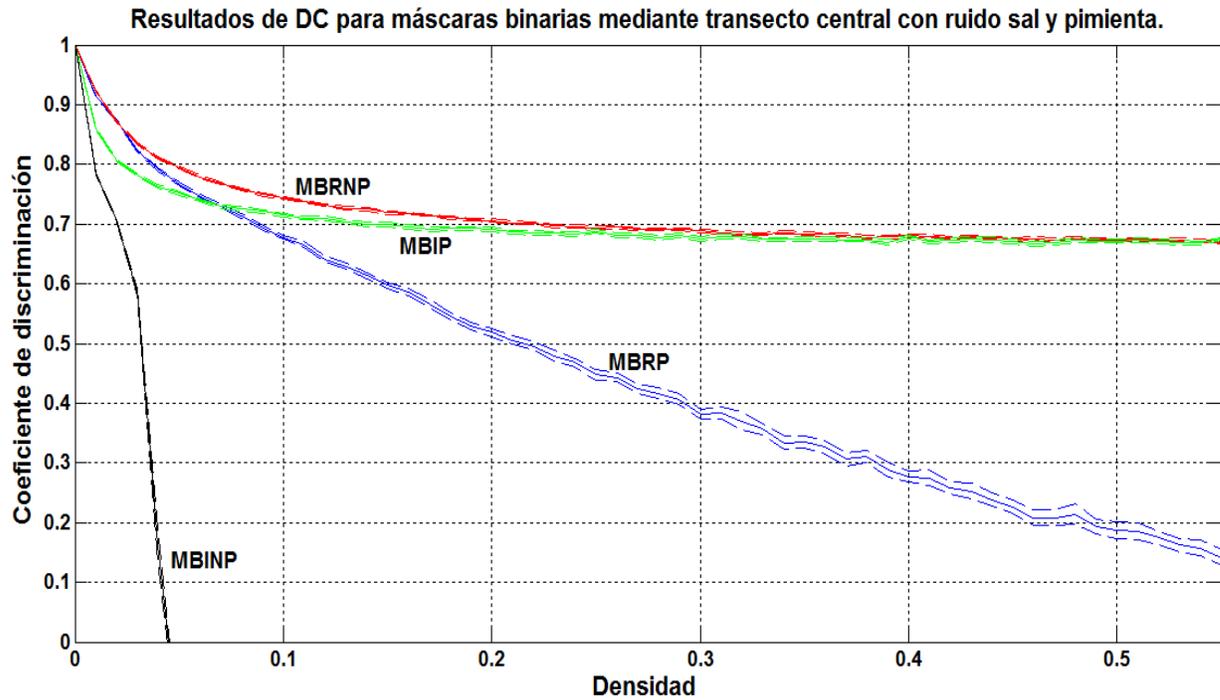


Figura 47. Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ para la metodología de máscaras binarias utilizando el perfil central con ruido impulsivo.

En la figura 48 se presenta el resultado con ruido gaussiano de la metodología de firmas vectoriales, donde se muestra la gráfica de la $\mu \pm 1.96SE$ de los resultados del coeficiente de discriminación al modificar la varianza entre 0 y 0.55, con 30 iteraciones, donde la μ se muestra en color negro, y en rojo se muestra $\pm 1.96SE$. En la figura 47 se muestra el promedio de las firmas de escala y las firmas de rotación.

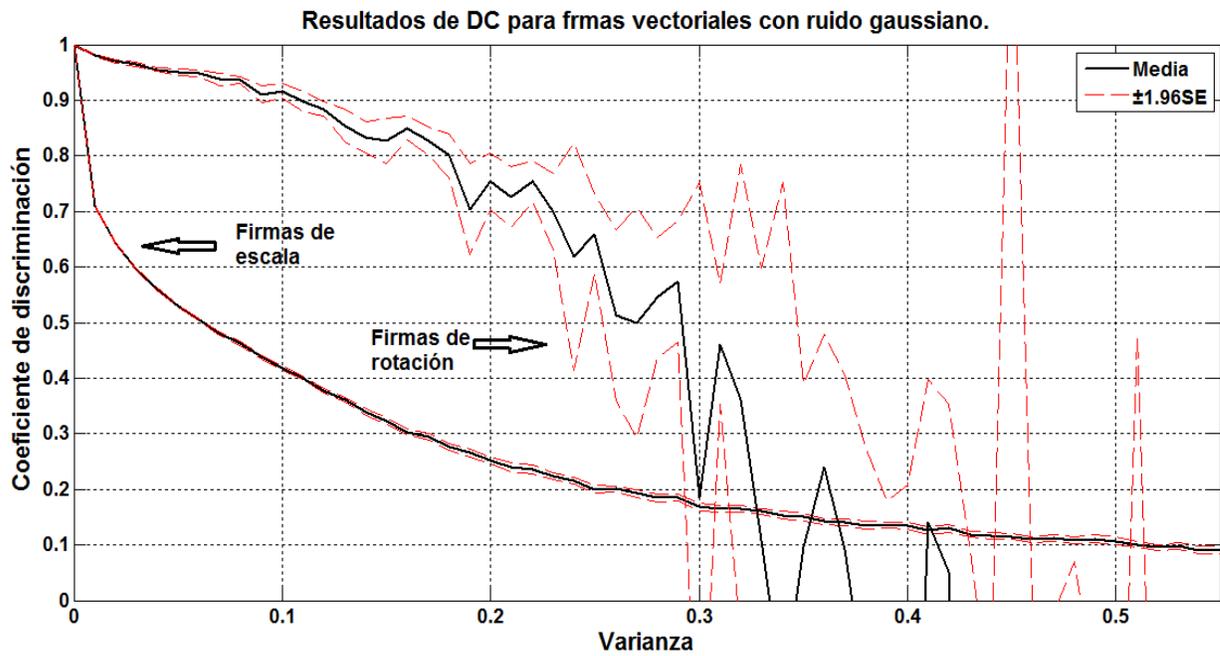


Figura 48. Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ para la metodología de firmas vectoriales con ruido gaussiano para $0 \leq \sigma^2 \leq 0.55$.

Dado que en esta metodología se deben considerar para la identificación los resultados de la firma vectorial de escala y la firma vectorial de rotación simultáneamente, en la figura 49 se presenta el promedio de la respuesta ante ruido gaussiano de las dos firmas.

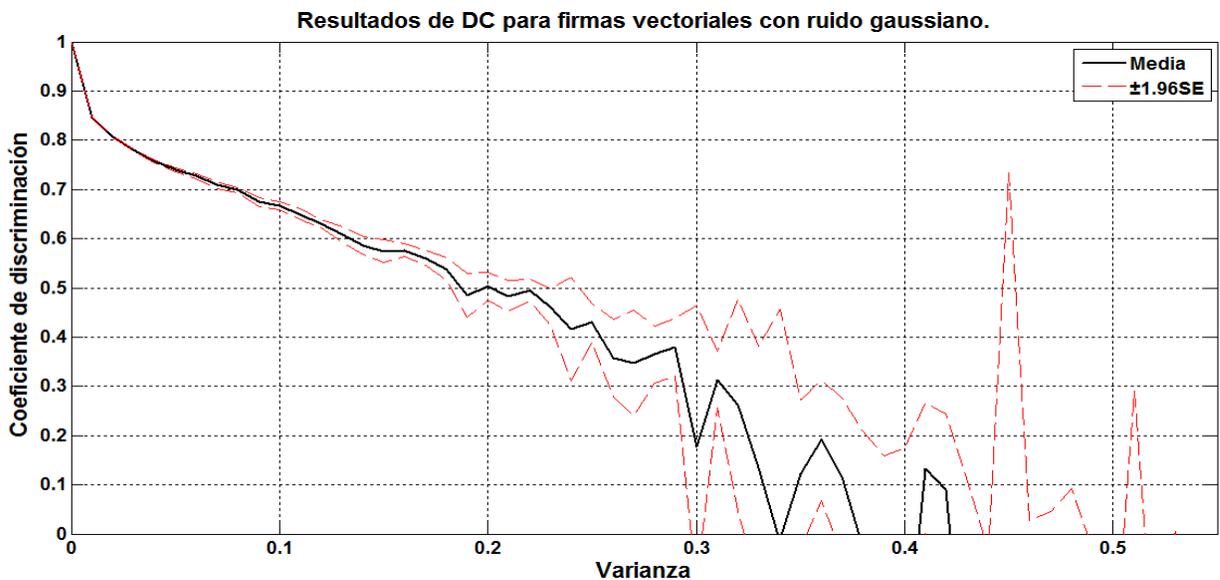


Figura 49. Gráfica del promedio de los resultados del coeficiente de discriminación de las firmas de escala y firmas de rotación con ruido gaussiano.

En la gráfica 50 se muestra el resultado para firmas vectoriales con ruido sal y pimienta con 30 iteraciones y variando la densidad $0 \leq d \leq 0.55$, y en la gráfica 51 se muestra el promedio de las dos firmas.

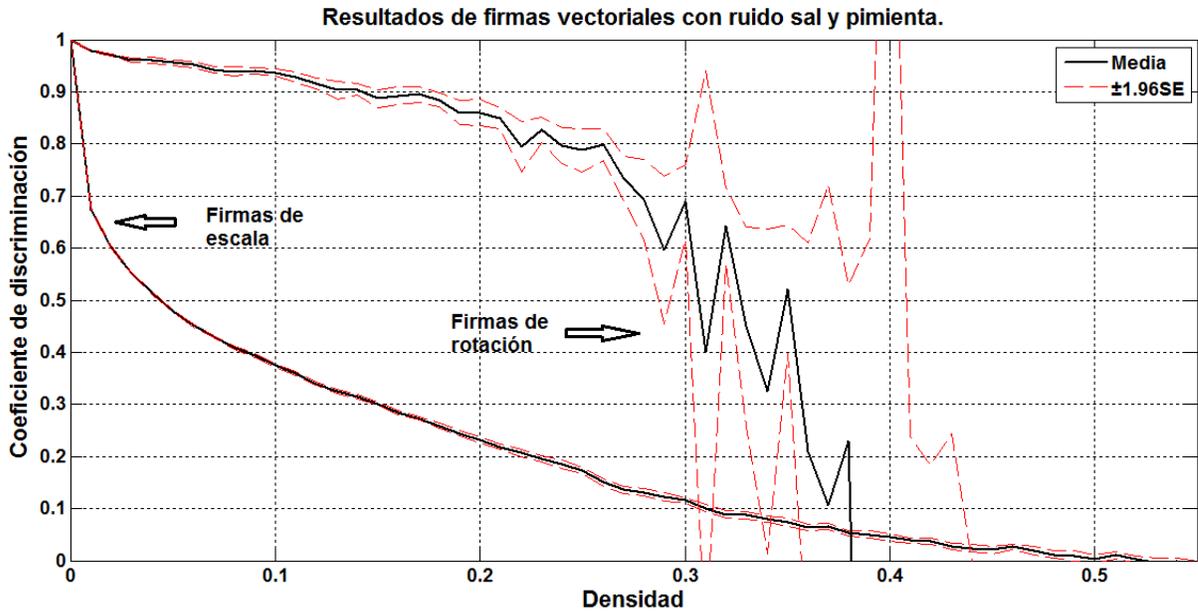


Figura 50. Gráfica del promedio de los resultados del coeficiente de discriminación $\pm 1.96SE$ para la metodología de firmas vectoriales con ruido impulsivo para $0 \leq \sigma^2 \leq 0.55$.

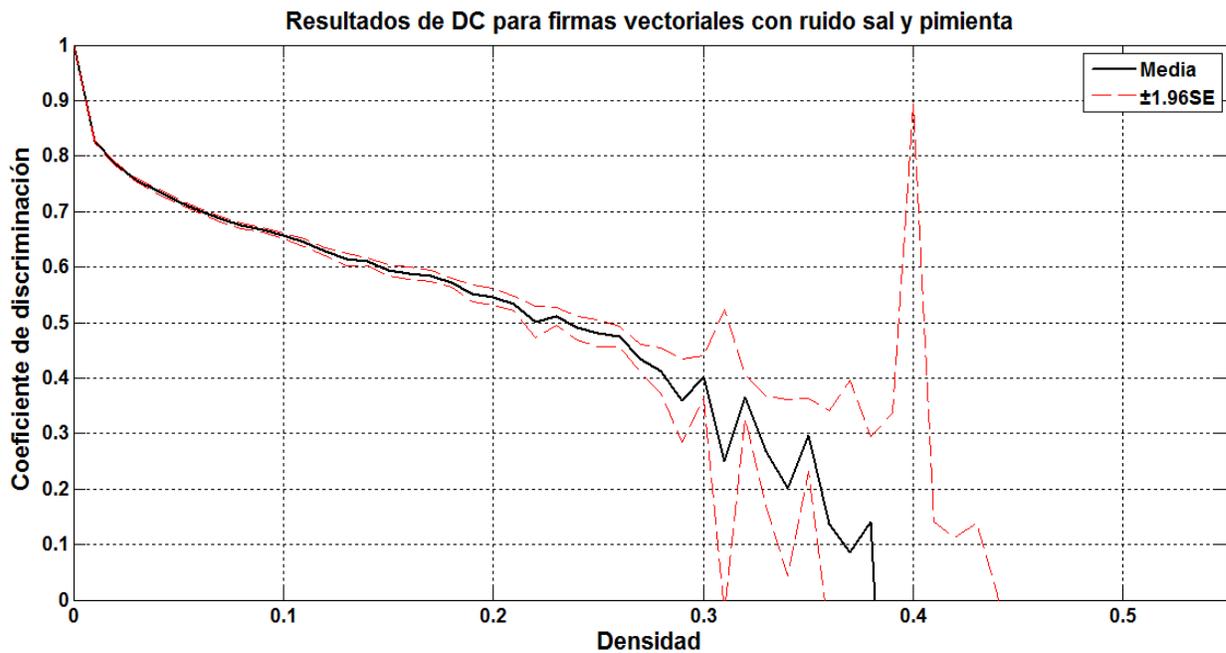


Figura 51. Gráfica del promedio de los resultados del coeficiente de discriminación de las firmas de escala y firmas de rotación con ruido impulsivo.

En la gráfica 52 se muestra una comparativa entre la metodología de firmas vectoriales con la de máscaras binarias MBRNP_PO y MBIP_PO, las cuales fueron las que mostraron la mejor respuesta dentro de todas las variantes de máscaras binarias mediante perfil óptimo.

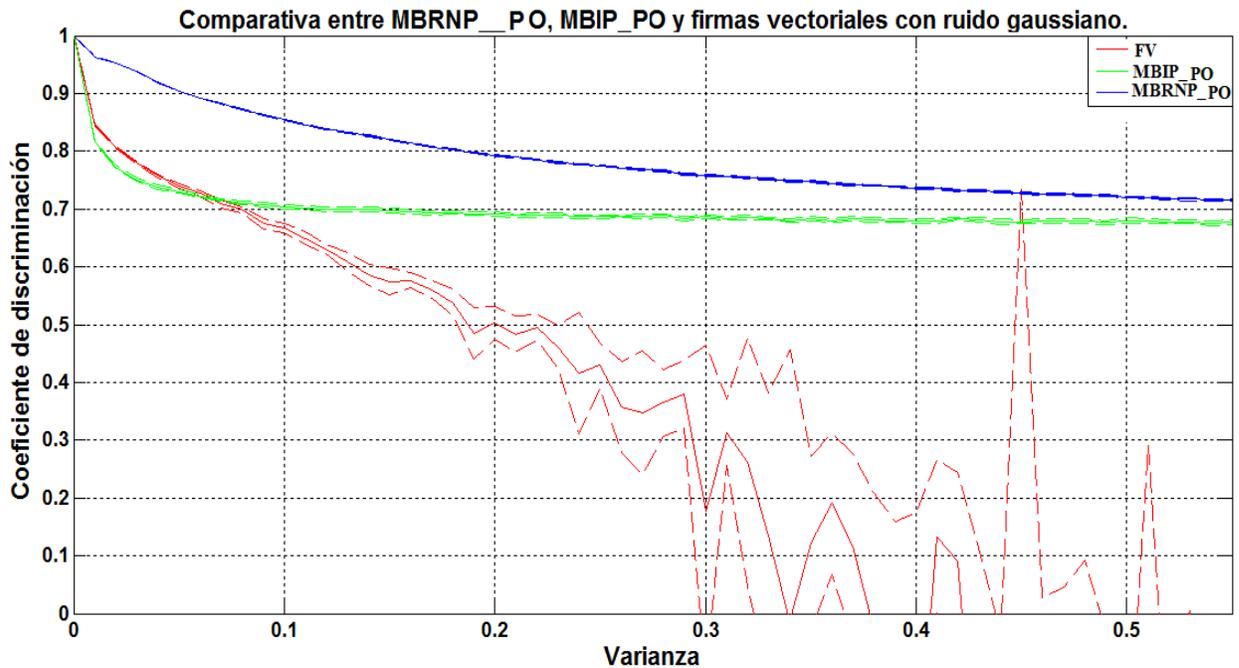


Figura 52. Gráfica comparativa entre firmas vectoriales, MBRNP_PO y MBIP_PO, para ruido gaussiano.

En la gráfica se muestra el promedio de las firmas de las dos firmas vectoriales, ya que deben ser consideradas simultáneamente dentro del proceso de identificación; las gráficas corresponden a la respuesta de las metodologías ante la presencia de ruido gaussiano. Se puede ver que las dos variantes de máscaras binarias presentadas en la gráfica tienen una mejor respuesta, ya que la respuesta de las firmas vectoriales cae a cero aproximadamente cuando la varianza es de 0.37, mientras que la respuesta de las máscaras binarias sigue estando por arriba de cero al llegar a una varianza de 0.55.

En la figura 53 se presenta ahora la misma comparativa, pero ahora con ruido sal y pimienta.

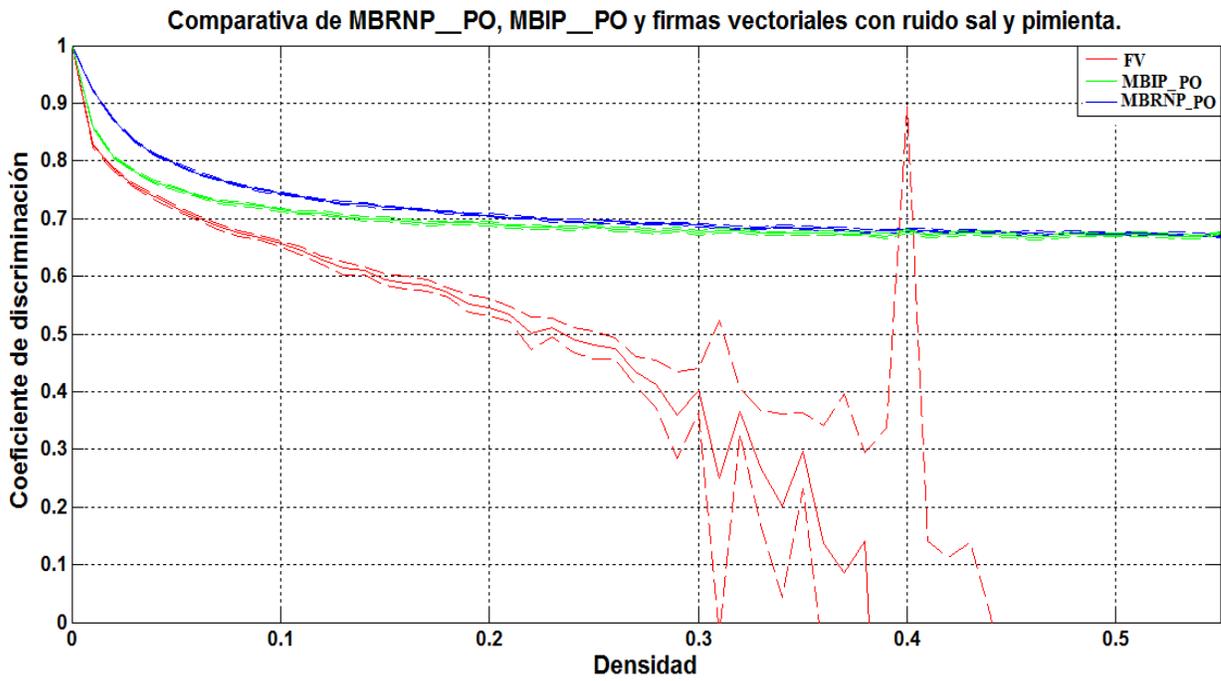


Figura 53. Gráfica comparativa entre firmas vectoriales, MBRNP_PO y MBIP_PO, para ruido sal y pimienta.

Con esta comparativa se observa que de todas las metodologías, las que tienen mejor respuesta ante los dos tipos de ruido son las de máscaras binarias cuando se utiliza la parte real con umbral no positivo y cuando se utiliza la parte imaginaria y un umbral positivo.

Capítulo 5. Nueva propuesta metodológica

En este capítulo se propone una nueva metodología de reconocimiento.

De los resultados obtenidos del análisis de las metodologías se pudo ver que la metodología de índices espectrales tiene la gran ventaja de que no utiliza la correlación, lo que permite obtener el resultado para todas las imágenes procesadas al mismo tiempo en un plano bidimensional, sin embargo, su tiempo de ejecución es el mayor de todas las metodologías analizadas, ya que por imagen fue de más de 19 segundos, mientras que la metodología de máscaras binarias tuvo una duración de milésimas de segundo por imagen. Por lo tanto, la idea es tener una metodología que nos muestre una respuesta a todas las imágenes en un mismo plano, pero con el menor tiempo de ejecución posible. Para ello se propone trabajar con las firmas unidimensionales generadas con las máscaras binarias, y a partir de éstas calcular dos parámetros que sean representativos de cada imagen, para ser graficados uno contra el otro en un plano bidimensional, y así obtener al mismo tiempo una respuesta para todas las imágenes.

Área bajo la firma

Los parámetros propuestos son por un lado el área bajo la firma; esto es, si tenemos una firma generada por una máscara binaria dada por $f_f(x)$, donde N es la longitud de la firma, el área de la firma está dada por

$$\text{área} = \frac{I}{N} \sum_{x=1}^N f_f(x). \quad (86)$$

Este parámetro es representativo de la imagen, ya que la firma generada por la máscara binaria es por sí misma representativa de la imagen procesada.

Varianza

El otro parámetro a utilizarse es la varianza de la firma. La varianza representa la media de la dispersión de los datos, además, el área del espectro de densidad de potencia de una señal es la densidad de varianza de la señal. Por lo tanto, la varianza total de la firma la obtendremos de la siguiente manera:

$$\text{varianza} = \frac{I}{N} \sum_{k=1}^N |F_f(k)|^2 \quad (87)$$

donde $|F_f(k)|^2$ es el módulo cuadrado de la transformada de Fourier de $f_f(x)$, el cual representa el espectro de densidad de varianza de la firma generada por la máscara binaria. Una vez calculado el área y la varianza de la firma, estos parámetros se grafican en un plano.

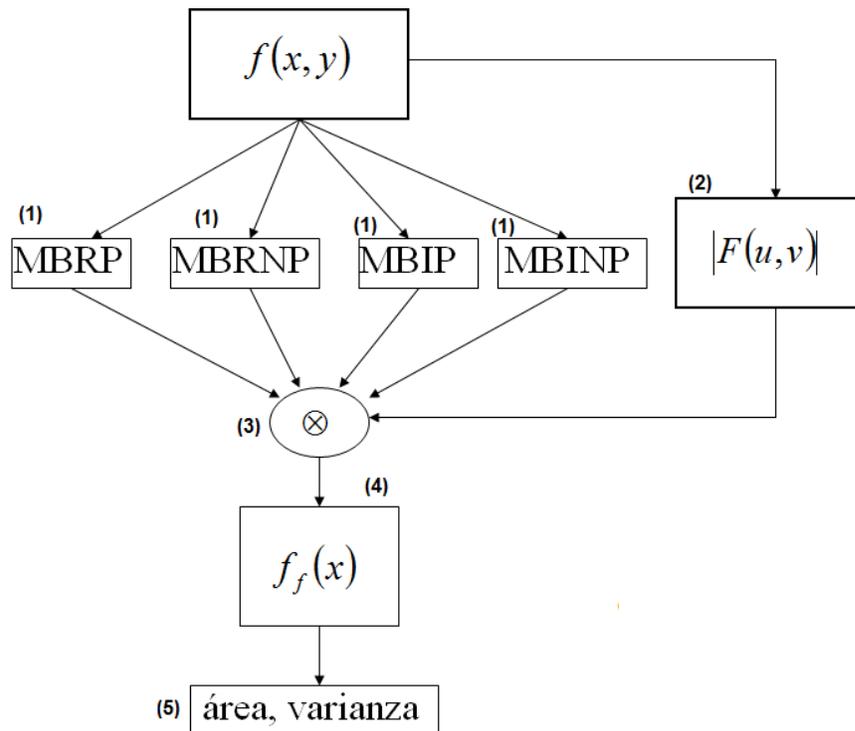


Figura 54. Diagrama a bloques de la metodología propuesta.

En la figura 54 se presenta el diagrama de los pasos a seguir en la metodología. En el paso 1 se obtiene la máscara binaria a utilizar: MBRP, MBRNP, MBIP, MBINP. En el paso 2 se calcula el módulo de la transformada de Fourier de la imagen $f(x, y)$. En el paso 3 se realiza la multiplicación punto a punto de la máscara binaria y el módulo de la transformada de Fourier. En el paso 4 se obtiene la firma de la imagen $f_f(x)$. En el paso 5 se calculan el área y varianza de la firma de la imagen.

Para implementar esta metodología se utilizaron las 21 imágenes de especies diferentes de diatomeas, las cuales se rotaron de la misma manera que antes, procesando un total de 3780 imágenes. Las firmas de las cuales se obtienen su área y su varianza son obtenidas de la misma manera en la que se obtiene en la metodología de máscara binarias, esto es, se utiliza una máscara binaria para muestrear el módulo de la transformada de Fourier de la imagen y obtener así un índice de anillo para formar la firma de cada imagen.

En la figura 55 se presenta el resultado del área y la varianza de las firmas generadas mediante una máscara MBRP, esto es, una máscara obtenida utilizando el perfil central de la parte real aplicando un umbral positivo. Se muestra en un plano de identificación el resultado estadístico del cálculo del área y la varianza para las 21 especies de diatomeas, donde el eje horizontal corresponde al área y el eje vertical corresponde a la varianza. Para realizar esta gráfica se calculó la media del área y la media de la varianza. Asimismo, se calculó el error estándar y se graficó más-menos dos errores estándar ($\pm 2SE$) para el área y la varianza en forma de cajas, de manera que se delimita un espacio dentro del plano de identificación que corresponde a un 95.4% de confianza de los valores de área y varianza para cada una de las 21 especies.

Se puede ver que el espacio correspondiente para cada diatomea no es muy grande y que aparentemente existe un traslape entre la diatomea M y la diatomea S, pero haciendo un acercamiento a esa zona, como se muestra en la figura 56, se puede ver que el espacio para cada diatomea está bien definido y que no existe traslape alguno entre éstas.

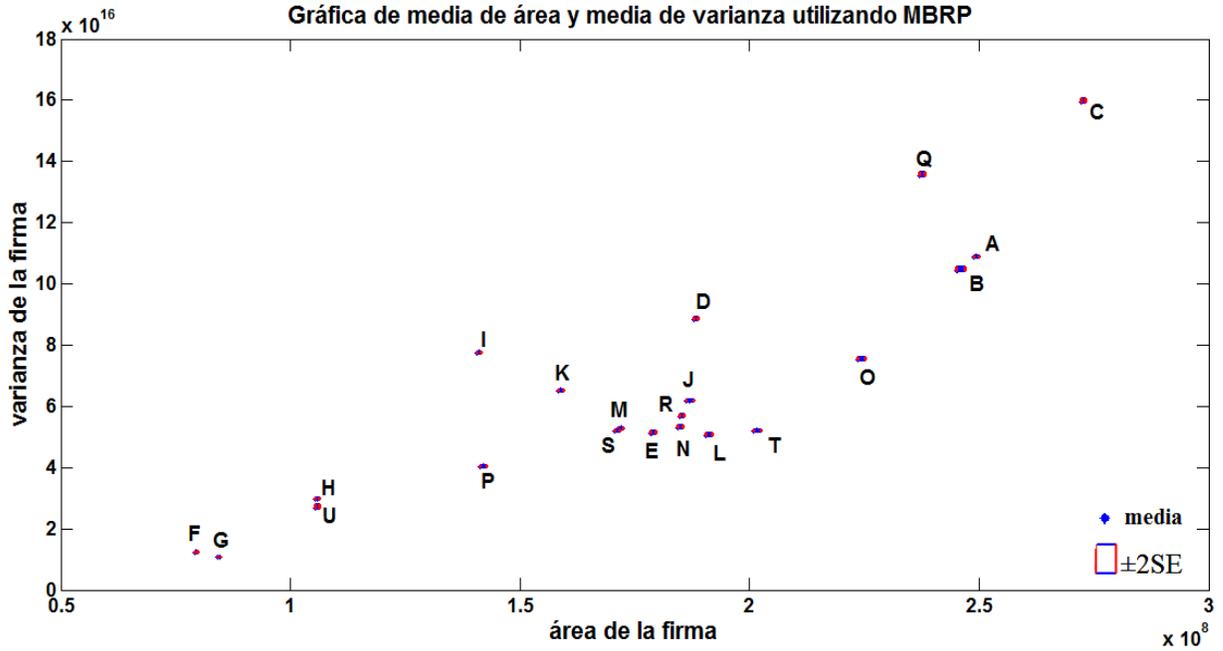


Figura 55. Gráfica de la media $\mu \pm 2SE$ del área y varianza de las firmas utilizando MBRP.

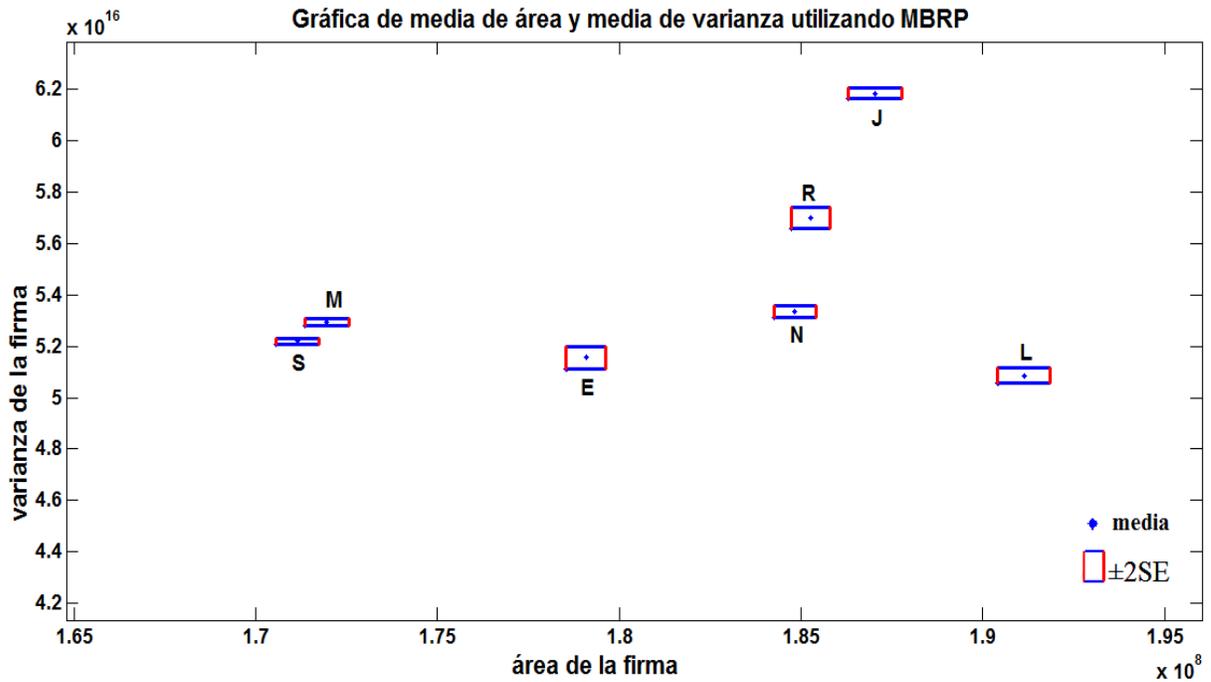


Figura 56. Acercamiento a los resultados de la media $\mu \pm 2SE$ del área y varianza de las firmas utilizando MBRP.

En la figura 56 se muestra el resultado cuando la firma es generada al utilizar una MBRNP. En este caso se muestra que en el plano de identificación existe una mayor separación entre las medias de las imágenes, lo que hace que el espacio para cada imagen sea más claro de observar que en el caso donde se utiliza MBRP.

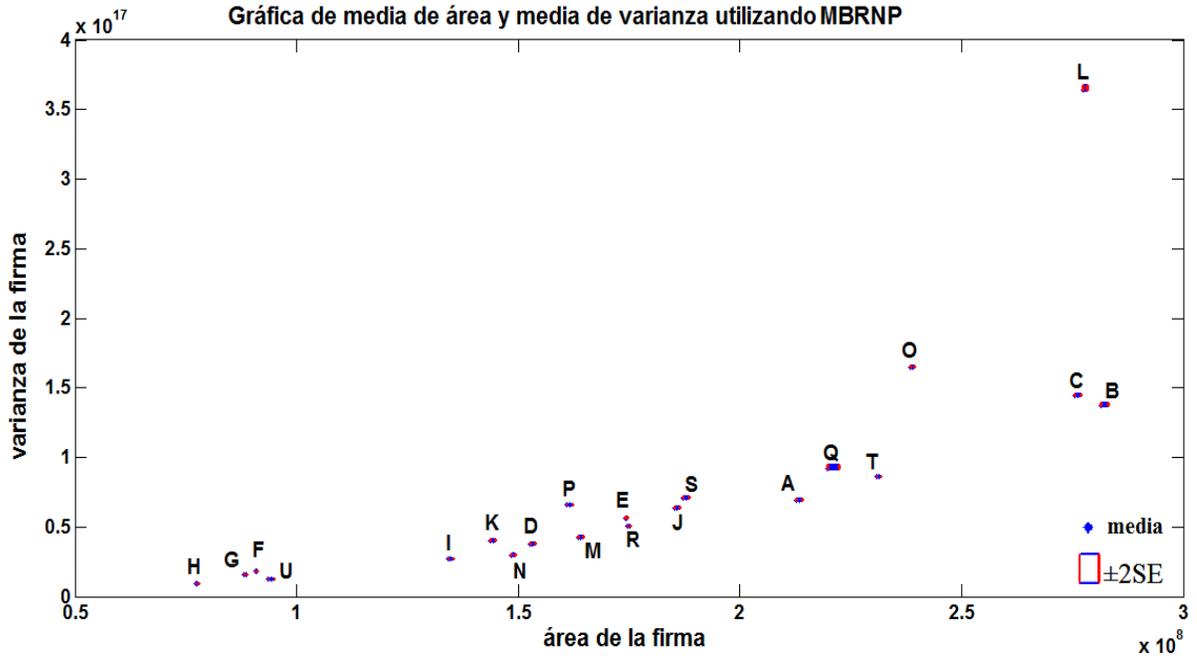


Figura 57. Gráfica de la media $\mu \pm 2SE$ del área y varianza de las firmas utilizando MBRNP.

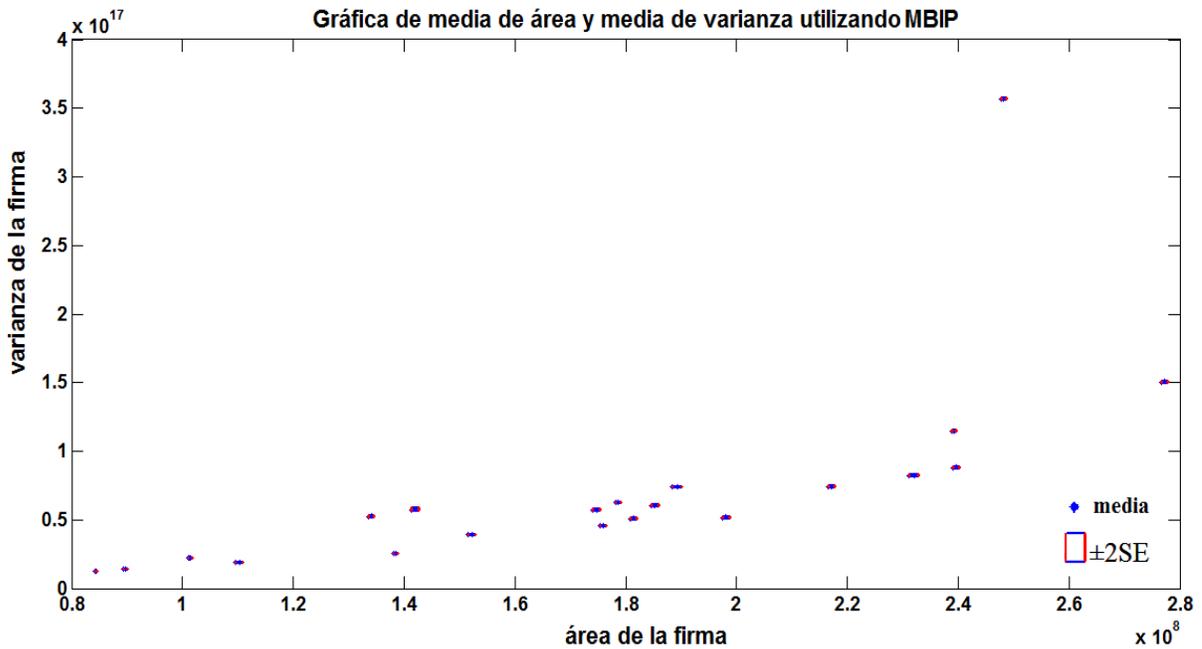


Figura 58. Gráfica de la media $\mu \pm 2SE$ del área y varianza de las firmas utilizando MBIP.

En la figura 58 se muestra el resultado cuando la firma es generada al utilizar una MBIP. Se puede ver que en el plano de identificación quedan desplegadas claramente cada una de las 21 diatomeas en base a la media del área y la media de varianza de su firma.

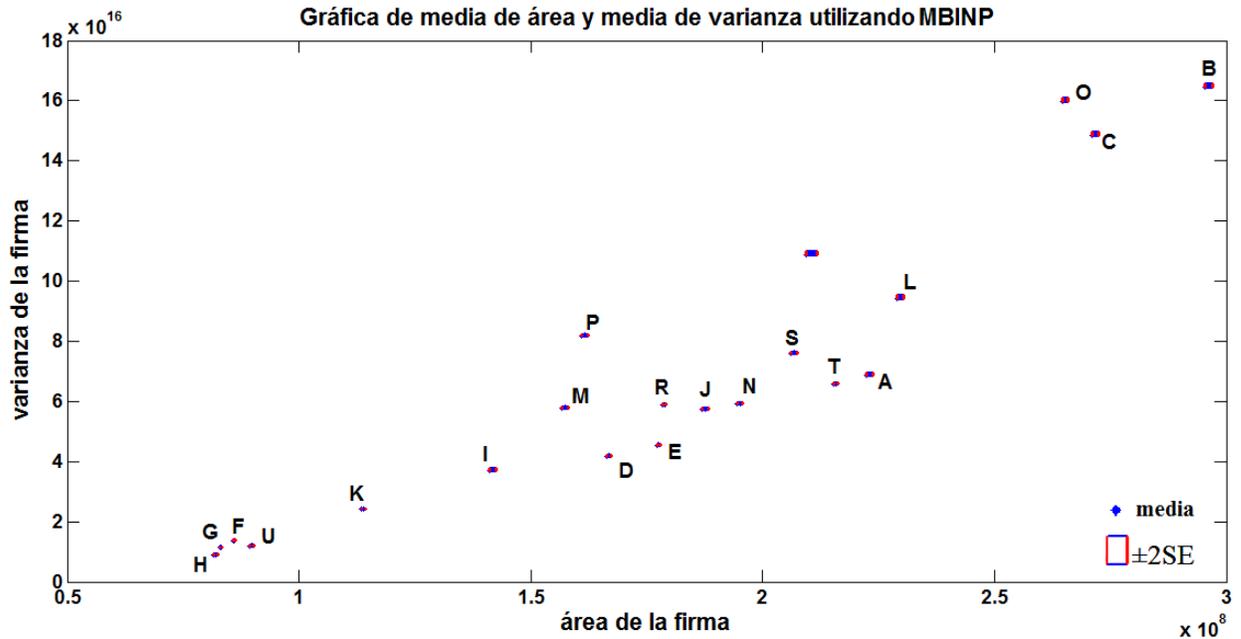


Figura 59. Gráfica de la media $\mu \pm 2SE$ del área y varianza de las firmas utilizando MBINP.

Los resultados cuando se utiliza MBINP para generar la firma de imagen se presentan en la figura 59. De nueva cuenta se observa que en el plano de identificación quedan desplegadas claramente los resultados de todas las imágenes.

Tiempo de ejecución

El tiempo de ejecución se calculó al procesar 3780 imágenes. El tiempo de ejecución para las cuatro variantes presentadas se muestra en la Tabla 7, donde también se muestra el tiempo de duración que tomó la metodología de índices espectrales al procesar la misma cantidad de imágenes que procesó la metodología propuesta.

Tabla 7. Tiempos de ejecución de la metodología propuesta comparados con los de índices espectrales.

Metodología	Tiempo total en segundos	Tiempo total en minutos	Tiempo por imagen en segundos
índices espectrales	135.03400	2.2506	35.7×10^{-3}
MBRP	101.65803	1.6943	26.9×10^{-3}
MBRNP	99.512801	1.6585	26.3×10^{-3}
MBIP	103.17923	1.7197	27.3×10^{-3}
MBINP	114.27260	1.9045	30.2×10^{-3}

Se puede observar que la metodología propuesta tiene un tiempo de ejecución menor que el que tiene la metodología de índices espectrales, la cual tardó 2.25 minutos en procesar la misma cantidad de imágenes que la metodología propuesta, la cual procesó la información en menos de 2 minutos.

Capítulo 6. Conclusiones

Se cumplió con el objetivo de implementación de la metodología de correlación no lineal con transformada de escala, firmas vectoriales, índices espectrales y máscaras binarias en sus diversas variaciones. Del análisis realizado, se llega a la conclusión de que no se puede señalar alguna de las metodologías analizadas como la mejor en un contexto general, ya que una metodología puede presentar un gran desempeño en alguna área pero tener un pobre desempeño en otra, así pues, se observó que en el caso de imágenes afectadas por ruido, el mejor desempeño lo tuvo la metodología de máscaras binarias cuando éstas se implementaban con la parte real y umbral no positivo e implementadas con la parte imaginaria aplicando umbral positivo, tanto si se utilizaba el perfil central o el perfil óptimo. La metodología de firmas vectoriales también mostró un buen desempeño, y la que tuvo el peor desempeño de todas fue la de correlación no lineal con transformada de escala.

También el análisis de PCE mostró que de las metodologías que utilizan correlación no lineal, la de máscaras binarias cuando se aplica la misma máscara a la imagen referencia e imagen problema presenta un mejor plano de salida, esto es, su salida es menos ruidosa, no siendo así para las metodologías de máscaras binarias adaptativas, mientras que la correlación no lineal con transformada de escala fue la que peor desempeño tuvo en esta evaluación.

En cuestión de tiempo de cómputo, también el grupo de metodologías de máscaras binarias fueron las que tuvieron el mejor desempeño ya que se utiliza la misma máscara para la imagen de referencia e imagen problema. Debido que al implementar las metodologías de máscaras binarias adaptativas es necesario calcular una máscara para la imagen referencia y otra para la imagen problema, el tiempo de ejecución, se incrementa en este grupo de metodologías, al grado que al utilizar el perfil óptimo al calcular la máscara adaptativa se tuvieron los tiempos de ejecución mayores.

La metodología de índices espectrales, la cual no utiliza correlación, da la ventaja de tener la respuesta de todas las imágenes involucradas en un solo plano.

A pesar de que la metodología de correlación no lineal con transformada de escala tuvo un mal desempeño en lo que respecta a la presencia de ruido, plano de salida y además de ser una de las que más tiempo consume, en el caso en el que se tiene solo un fragmento de información, tuvo el mejor desempeño de todas las metodologías.

Por lo tanto, la mejor metodología dependerá de los requerimientos del problema a enfrentar y así determinar que metodología es la más adecuada para dicho problema.

Con base en los resultados obtenidos se propuso una metodología que cuenta con la ventaja de proporcionarnos la respuesta de todas las imágenes a ser analizadas al mismo tiempo en un solo plano bidimensional, pero con un menor tiempo de ejecución que el requerido en el cálculo de los índices espectrales. Esto se logró al utilizar el área y la varianza de las firmas generadas con máscaras binarias, aprovechando su corto tiempo de ejecución, reduciendo así de 35.7 milésimas de segundo por imagen en el caso de índices espectrales a menos de 31 milésimas de segundo por imagen en la metodología propuesta.

Lista de referencias

- Almeida, S. P., Case, S. K., Fournier, J. M., Fugii, H., Cairns Jr, J., Dickson, K. L., y Pryfogle, P. (1978). Analysis of algae samples using coherent optical processing. Proceedings of the International Commission for Optics, ICO-11, Madrid, Spain, 351-354.
- Almeida, S. P., Del Balzo, D. R., Cairns Jr, J., Dickson, K. L., y Lanza, G. R. (1971). Holographic microscopy of diatoms. Transactions of the Kansas Academy of Science (1903), 257-260.
- Almeida, S. P., y Eu, J. K. T. (1976). Water pollution monitoring using matched spatial filters. Applied optics, 15(2), 510-515.
- Alvarez-Borrego, J., y Castro-Longoria, E. (2003). Discrimination between *Acartia* (Copepoda: Calanoida) species using their diffraction pattern in a position, rotation invariant digital correlation. Journal of Plankton Research, 25(2), 229-233.
- Álvarez-Borrego, J., y Solorza, S. (2010). Comparative analysis of several digital methods to recognize diatoms. Hidrobiologica, 20(2), 158-170.
- Álvarez-Borrego, J., Solorza, S., y Bueno-Ibarra, M. A. (2013). Invariant correlation to position and rotation using a binary mask applied to binary and gray images. Optics Communications, 294, 105-117.
- Bueno-Ibarra, A., Chávez-Sánchez, M. C., y Álvarez-Borrego, J. (2010). Nonlinear law spectral technique to analyze white spot syndrome virus infection. Biosciences world, 30-33.
- Bueno-Ibarra, M. A., Chávez-Sánchez, M. C., y Álvarez-Borrego, J. (2011). K-law spectral signature correlation algorithm to identify white spot syndrome virus in shrimp tissues. Aquaculture, 318(3), 283-289.
- Cairns Jr, J., Dickson, K. L., Lanza, G. R., Almeida, S. P., y Del Balzo, D. (1972). Coherent optical spatial filtering of diatoms in water pollution monitoring. Archiv fuer Mikrobiologie, 83(2), 141-146.
- Cairns, J., Almeida, S. P., y Fujii, H. (1982). Automated identification of diatoms. Bioscience, 32(2), 98-102.
- Casasent, D., y Psaltis, D. (1976). Scale invariant optical correlation using Mellin transforms. Optics Communications, 17(1), 59-63.
- Casasent, D., y Psaltis, D. (1976). Scale invariant optical transform. Optical Engineering, 15(3), 153258-153258.
- Casasent, D., y Psaltis, D. (1976). Position, rotation, and scale invariant optical correlation. Applied optics, 15(7), 1795-1799.

- Castro-Longoria, E., Alvarez-Borrego, J., y Pech-Pacheco, J. L. (2001). Identification of species of calanoid copepods using a new invariant correlation algorithm. *Crustaceana-International Journal of Crustacean Research*-, 74(10), 1029-1040.
- Cohen, L. (1995). *Time-frequency analysis* (Vol. 778). Englewood Cliffs, NJ: Prentice Hall PTR.
- Coronel Beltrán, Á. (1988) *Correlacionador óptico invariante aplicado a la identificación de microorganismos fitoplanctónicos*. Tesis de maestría en ciencias. Centro de Investigación Científica y Educación Superior de Ensenada. 110 p.
- Coronel-Beltrán, Á., y Álvarez-Borrego, J. (2009). Sistema digital de correlación invariante aplicado a la comparación de dos fuentes de letras. *e-Gnosis*, 1-7.
- Coronel-Beltran, A., y Álvarez-Borrego, J. (2010). Comparative analysis between different font types and letter styles using a nonlinear invariant digital correlation. *Journal of Modern Optics*, 57(1), 58-64.
- Coronel-Beltrán, A., y Álvarez-Borrego, J. (2008, August). Nonlinear filter for pattern recognition using the scale transform. In *Optical Engineering+ Applications* (pp. 70732H-70732H). International Society for Optics and Photonics.
- Chaparro Magallanes, G. (2011) *Correlación invariante mediante el uso de firmas unidimensionales*. Tesis de maestría en ciencias Centro de Investigación Científica y Educación Superior de Ensenada. 63 p.
- De Sena, A., y Rocchesso, D. (2007). A fast Mellin and scale transform. *EURASIP Journal on Advances in Signal Processing*.
- De Sena, A., y Rocchesso, D. (2004, October). A fast mellin transform with applications in dafx. In *Proc. of the 7th Int. Conference on Digital Audio Effects (DAFx'04)* (pp. 65-69).
- Fimbres-Castro, C., Álvarez-Borrego, J., y Bueno-Ibarra, M. A. (2011, August). Invariant correlation by using vectorial signatures and spectral index. In *International Commission for Optics (ICO 22)* (pp. 801171-801171). International Society for Optics and Photonics.
- Fimbres-Castro, C., Álvarez-Borrego, J., y Bueno-Ibarra, M. A. (2012). Invariant nonlinear correlation and spectral index for diatoms recognition. *Optical Engineering*, 51(4), 047201-1.
- Gerardi, F. (1959). Application of Mellin and Hankel transforms to networks with time-varying parameters. *Circuit Theory, IRE Transactions on*, 6(2), 197-208.
- Gonzalez, R. C., y Woods, R. E. (1992). *Digital Image Processing*. Addison-Wesley.

- Gonzalez, R. C., y Woods, R. E. (2002). *Digital Image Processing (2nd ed.)*. New Jersey, Prentice Hall.
- Guerrero Moreno, R. (2008) *Correlación invariante de objetos utilizando filtros compuestos no lineales*. Tesis de maestría en ciencias Centro de Investigación Científica y Educación Superior de Ensenada. 90 p.
- Guerrero-Moreno, R. E., y Álvarez-Borrego, J. (2009). Nonlinear composite filter performance. *Optical Engineering*, 48(6), 067201-067201.
- Horner, J. L., y Gianino, P. D. (1984). Phase-only matched filtering. *Applied optics*, 23(6), 812-816.
- Javidi, B. (1989). Nonlinear joint power spectrum based optical correlation. *Applied Optics*, 28(12), 2358-2367.
- Javidi, B. (1989). Synthetic discriminant function-based binary nonlinear optical correlator. *Applied optics*, 28(13), 2490-2493.
- Javidi, B. (1989). Nonlinear matched filter based optical correlation. *Applied optics*, 28(21), 4518-4520.
- Javidi, B. (1990). Generalization of the linear matched filter concept to nonlinear matched filters. *Applied optics*, 29(8), 1215-1224.
- Javidi, B. (1990). Comparison on nonlinear joint transform correlator and nonlinear matched filter based correlator. *Optics communications*, 75(1), 8-13.
- Javidi, B., Castro, M. A., Kishk, S., y Perez, E. (2002). Automated detection and analysis of speed limit signs. University of Connecticut, Tech. Rep. JHR, 02-285.
- Javidi, B., y Horner, J. L. (Eds.). (1994). *Real-time optical information processing*. Academic Press.
- Javidi, B., Wang, W., y Zhang, G. (1997). Composite Fourier-plane nonlinear filter for distortion-invariant pattern recognition. *Optical Engineering*, 36(10), 2690-2696.
- Kumar, B. V. K., y Hassebrook, L. (1990). Performance measures for correlation filters. *Applied optics*, 29(20), 2997-3006.
- Lerma-Aragón, J. (2010) *Reconocimiento de patrones mediante la utilización de firmas vectoriales*. Tesis de doctorado en ciencias. Universidad Autónoma de Baja California. 92 p.
- Lerma Aragon, J. R., y Álvarez-Borrego, J. (2009). Vectorial signatures for invariant recognition of position, rotation and scale pattern recognition. *Journal of Modern Optics*, 56(14), 1598-1606.

- Mouriño-Pérez, R. R., Alvarez-Borrego, J., y Gallardo-Escárate, C. (2006). Digital color correlation for the recognition of *Vibrio cholerae* 01 in laboratory and environmental samples. *Revista de Biología Marina y Oceanografía*, 41(1), 77.
- Pech-Pacheco, J. L., y Alvarez-Borrego, J. (1998). Optical–digital system applied to the identification of five phytoplankton species. *Marine Biology*, 132(3), 357-365.
- Pech-Pacheco, J. L., Álvarez-Borrego, J., Gabriel Cristóbal y Matthias Keil. (2003). Automatic object identification irrespective to geometric changes. *Optical Engineering*, Vol. 42, No. 2, pp 551-559.
- Sawchuk, A. A. (1974). Space-variant image restoration by coordinate transformations. *JOSA*, 64(2), 138-144.
- Smith, W. S. (1999). *The Scientist and Engineer's Guide to Digital Signal Processing*. (2nd ed.). San Diego Cal, California Technical Publishing.
- Solorza, S., y Álvarez-Borrego, J. (2010). Digital system of invariant correlation to position and rotation. *Optics Communications*, 283(19), 3613-3630.
- Selene, S., y Álvarez-Borrego, J. (2014). Digital system invariant to position and rotation to object recognition in images by intensity profiles. *Revista Cubana de Física*, Vol. 31, No. 1.
- Villalobos-Flores, C. E., Alvarez-Borrego, J., Kober, V., Cristobal-Perez, G., y Castro-Longoria, E. (2002). Study of 21 fragmented fossil diatoms using a digital invariant correlation. In *International Symposium on Optical Science and Technology* (pp. 528-533). International Society for Optics and Photonics.