

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE INGENIERÍA ENSENADA



**RECONOCIMIENTO DE PATRONES MEDIANTE LA
UTILIZACIÓN DE FIRMAS VECTORIALES**

TESIS

QUE PRESENTA PARA OBTENER EL GRADO DE DOCTOR EN CIENCIAS

JESÚS RAMÓN LERMA ARAGÓN

**DIRECTOR DE TESIS:
DR. JOSUÉ ÁLVAREZ BORREGO**

ENSENADA, BAJA CALIFORNIA, MÉXICO, NOVIEMBRE DE 2010.

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
FACULTAD DE INGENIERÍA ENSENADA

RECONOCIMIENTO DE PATRONES MEDIANTE LA
UTILIZACIÓN DE FIRMAS VECTORIALES

TESIS

QUE PRESENTA PARA OBTENER EL GRADO DE DOCTOR EN CIENCIAS

JESÚS RAMÓN LERMA ARAGÓN

APROBADA POR:

Dr. Josué Álvarez Borrego
Director de tesis

Dra. Selene Solorza Calderón
Miembro del comité

Dr. José Ángel González Fraga
Miembro del comité

Dr. Roberto Machorro Mejía
Miembro del comité

Dr. Juan Crisóstomo Tapia Mercado
Miembro del comité

ENSENADA, BAJA CALIFORNIA, MÉXICO, NOVIEMBRE DE 2010.

RESUMEN de la tesis de Jesús Ramón Lerma Aragón, presentada como requisito parcial para la obtención del grado de DOCTOR EN CIENCIAS. Ensenada, Baja California, México. Noviembre de 2010.

RECONOCIMIENTO DE PATRONES MEDIANTE LA UTILIZACIÓN DE FIRMAS VECTORIALES

Resumen aprobado por:

Dr. Josué Álvarez Borrego
Director de tesis

Ante la creciente demanda de automatización de procesos, la importancia del reconocimiento de patrones en las imágenes aumenta.

Como una nueva contribución frente a esta problemática, la presente investigación está dedicada al desarrollo de un sistema digital basado en la utilización de firmas vectoriales que son invariantes a cambios de posición, rotación y escala empleadas para el reconocimiento de patrones en imágenes en escala de grises y color. El sistema fue construido aprovechando las propiedades que presentan las transformadas de Fourier, Mellin y Escala.

Para imágenes en escala de grises se analizaron letras del alfabeto, números y formas de hojas; a ellas se les aplicaron rotaciones y cambios de escala. Se analizaron imágenes reales de varias especies de copépodos. Para cada imagen se obtuvieron dos firmas vectoriales; y el grado de semejanza entre ellas se obtuvo comparando sus firmas vectoriales mediante la distancia euclidiana. Los resultados obtenidos nos indican que podemos distinguir la imagen de referencia con un alto nivel de confianza y, para el caso concreto de copépodos, además su especie y sexo.

En el caso de imágenes en color se estudiaron dos conjuntos consistentes en letras de colores y diferentes especies de mariposas, aplicándoseles también cambios en rotación y escala como en el caso anterior. Después de realizar simulaciones computacionales encontramos que este método permite distinguir de manera efectiva objetos que tienen forma similar pero diferente color y a su vez, en el caso de imágenes de mariposas, diferenciar especies que son del mismo género.

Los resultados obtenidos nos permiten asegurar que en la mayoría de los casos podemos distinguir el patrón de referencia con al menos 95.4 % de nivel de confianza.

ABSTRACT of the thesis presented by Jesús Ramón Lerma Aragón in order to obtain the DOCTOR IN SCIENCES degree. Ensenada, Baja California, México. November of 2010.

VECTORIAL SIGNATURES FOR PATTERN RECOGNITION

Approved by:

Dr. Josué Álvarez Borrego
Thesis Advisor

With the increasing demand for process automation, the importance of pattern recognition in images increases.

As a new contribution to address this problem, the present investigation is dedicated to the development of a digital system based on the utilization of vectorial signatures that are invariants to changes of position, rotation and scale used for the pattern recognition of images in grayscale and color. The system was developed by taking advantage of the properties of the Fourier, Mellin and Scale transforms.

Grayscale images were analyzed, including letters of the alphabet, numbers and shapes of leaves, applying rotations and scale changes. We analyzed real images of several species of copepods. For each image, we obtained two one-dimensional vectors in the target as well as the input scene. To recognize a target, signatures were compared, calculating the Euclidean distance. The results show us that we can distinguish the reference image with a high level of confidence and, for the specific case of copepods, specie and sex.

In case of color images, two sets consisting of letters in colors and different species of butterflies were studied, applying changes in rotation and scale as in the previous case. After carrying out computer simulations, we found that this method effectively distinguishes objects that have similar shape but different color and, in the case of images of butterflies, we can distinguish species that are of the same gender.

The obtained results allow us to assure that in most cases we can distinguish the pattern of reference with at least 95.4 % of confidence level.

DEDICATORIA

A mi esposa Paty, por su amor, comprensión y apoyo, sin lo que hubiese sido imposible lograr terminar estos estudios.

A mi hijo Jesús Daniel por ser la fuente de mi inspiración y motivación para superarme cada día más.

A mis papas Jesús y Gloria que me dieron la vida y han estado conmigo en todo momento.

AGRADECIMIENTOS

Aquí está resumido el trabajo de investigación de mis últimos años, en los que he recibido apoyo científico y moral de muchas personas, sin el cual hubiera sido difícil realizarlo. A todas ellas les manifiesto mi más sincero agradecimiento. En especial a aquellas que de una u otra forma me han acompañado en esta empresa más de cerca.

Al Dr. Josué Álvarez Borrego, director de esta tesis, quién con su esfuerzo, conocimientos y dedicación supo despejar dudas y orientarme a la realización exitosa de este proyecto, por el tiempo que me ha dedicado y su constante persuasión para que terminara el trabajo.

A los miembros del comité de tesis: Dra. Selene Solorza Calderón, Dr. José Ángel González Fraga, Dr. Roberto Machorro Mejía y Dr. Juan Crisóstomo Tapia Mercado por todos sus valiosos comentarios, sugerencias, asesorías brindadas y sobre todo por brindarme su amistad en todo momento.

A mi esposa Paty por sus valiosos comentarios y sugerencias con respecto a la realización de este escrito, así como a las horas de desvelo que pasó a mi lado, sus palabras de aliento y apoyo incondicional que me brindó en todo momento, haciendo posible que yo pudiera finiquitar con éxito este proyecto. Es la persona que más directamente ha sufrido las consecuencias del trabajo realizado. Nunca le podré estar suficientemente agradecido.

A mi hijo Jesús Daniel por tener la paciencia y comprensión por el poco tiempo que he pasado con él, debido a la realización de este trabajo.

Quiero agradecer a mis padres, pilares fundamentales en mi vida. Sin ellos, jamás hubiese podido conseguir lo que hasta ahora. Su tenacidad y lucha insaciable han hecho de ellos el gran ejemplo a seguir y destacar.

A mis hermanos Jaime, Maye y Rocio gracias por estar conmigo y apoyarme siempre.

A las Autoridades de la Facultad de Ciencias quienes me brindaron la oportunidad, para realizar mis estudios de doctorado.

Al Director y personal de la Facultad de Ingeniería Ensenada por brindarme todas las facilidades para poder alcanzar esta meta.

Por último, agradezco al proyecto de Ciencia Básica financiado por el CONACYT: “Correlación invariante a posición, rotación, escala, ruido e iluminación para identificación de organismos y estructuras microscópicas y macroscópicas”, número 102007 a cargo del Dr. Josué Álvarez Borrego.

ÍNDICE

	<i>Página</i>
RESUMEN EN ESPAÑOL.....	I
RESUMEN EN INGLÉS.....	III
DEDICATORIAS.....	V
AGRADECIMIENTO.....	VI
ÍNDICE.....	VIII
LISTA DE FIGURAS.....	XI
LISTA DE TABLAS.....	XIV
CAPÍTULO I. INTRODUCCIÓN.....	1
1.1 Motivación.....	2
1.2 Objetivos.....	4
1.3 Antecedentes.....	5
1.4 Orígenes del procesado digital de imágenes.....	9
1.5 Reconocimiento de patrones.....	13
1.6 Sistemas para reconocer objetos biológicos.....	16
1.7 Estructura de la tesis.....	19
CAPÍTULO II. MARCO TEÓRICO.....	21
2.1 Introducción.	22
2.2 Antecedentes de la transformada de Fourier.....	22
2.3 Transformadas Integrales.....	23
2.4 Transformada de Fourier en una dimensión.....	23
2.4.1 Transformada bidimensional y su inversa.....	25
2.5 Transformada de Mellin.....	28
2.5.1 Definición y relación con otras transformaciones.	28
2.5.2 Transformada de Fourier-Mellin.	31
2.6 Transformada de Escala.	33

ÍNDICE (continuación)

CAPÍTULO III. METODOLOGÍA.....	36
3.1 Introducción.....	37
3.2 Implementación.....	38
3.3 Aplicación del algoritmo.....	40
CAPÍTULO IV. RECONOCIMIENTO DE IMÁGENES EN ESCALA DE GRISES.....	46
4.1 Análisis de letras.....	47
4.1.1 Resultados para la rotación.....	47
4.1.2 Resultados para cambio de escala.....	53
4.2 Análisis de números.....	54
4.3 Análisis de formas de hojas.....	57
4.4 Copépodos.....	59
4.4.1 Simulaciones utilizando imágenes de los copépodos.....	60
4.4.2. Firmas promedio.....	64
4.4.3. Experimento de identificación de especies.....	67
4.5 Imágenes de color transformadas a escala de grises.....	68
CAPÍTULO V. RECONOCIMIENTO DE IMÁGENES EN COLOR.....	73
5.1 Introducción.....	74
5.2 Metodología.....	76
5.3 Simulaciones numéricas.....	79
5.3.1. Estudio numérico para el caso de letras de colores.....	79
5.3.2. Estudio numérico para el caso de mariposas.....	81
5.4 Trabajos derivados de la tesis.....	84

ÍNDICE (continuación)

CAPÍTULO VI. CONCLUSIONES.....	86
BIBLIOGRAFÍA.....	89

LISTA DE FIGURAS

Figura	Página	
1.1	Imagen digital producida en 1921 con una cinta codificada e impresa en un teletipo (Gonzalez, 2008).	10
1.2	Una imagen digital realizada en 1922 de una cinta perforada después de que las señales habían cruzado el Océano Atlántico dos veces (Gonzalez, 2008).	10
1.3	La primera fotografía de la luna tomada de una nave espacial estadounidense (Gonzalez, 2008).	12
3.1	Diagrama de bloques del procedimiento utilizado.	39
3.2	(a) Imagen original y (b) su espectro de Fourier.	41
3.3	Espectro de Fourier mejorado visualmente.	41
3.4	Filtro parabólico.	42
3.5	Figura 3.5 Resultado de aplicar a) Filtro parabólico al espectro de Fourier. b) El Factor de escala \sqrt{r} a la Fig. 3.5 (a).	42
3.6	Figura 3.6 a) Imagen 3.5 (b) en coordenadas log-polar. b) Subimagen seleccionada.	43
3.7	(a) Imagen tridimensional de Fig. 3.6 (b). (b) Vector de la sumatoria sobre el eje rotación. (c) Vector de la sumatoria sobre el eje escala.	44
3.8	Firmas vectoriales. (a) Firma de escala. (b) Firma de rotación.	45
4.1	Análisis de comportamiento de la d_E de una imagen con respecto a la rotación.	47
4.2	Análisis de comportamiento de la d_E de todas las letras con respecto a la rotación.	48
4.3	Análisis de comportamiento de la d_E .	49
4.4	Comportamiento estadístico de la d_E .	50

LISTA DE FIGURAS (continuación)

Figura		Página
4.5	Comportamiento de la distancia ante cambios de escala de 80 % a 120 %.	53
4.6	Cambios en la distancia versus ángulo de rotación.	55
4.7	Comportamiento estadístico de la distancia para rotación.	56
4.8	Comportamiento estadístico para cambios de escala de 80 % a 120 %.	56
4.9	Patrones utilizados con diferentes formas de las hojas.	57
4.10	Comportamiento estadístico de la d_E para rotación.	58
4.11	Comportamiento estadístico de la d_E para cambios de escala de 80 % a 120 %.	58
4.12	Especies de copépodos utilizadas: a) <i>Calanus pacificus</i> hembra. b) <i>C. pacificus</i> macho. c) <i>Rhincalanus nasutus</i> hembra. d) <i>R. nasutus</i> macho. e) <i>Centropages furcatus</i> hembra. f) <i>C. furcatus</i> macho. g) <i>Pleuromamma gracilis</i> hembra. h) <i>P. gracilis</i> macho. i) <i>Temora discaudata</i> hembra. j) <i>T. discaudata</i> macho. k) <i>Acartia tonsa</i> hembra. l) <i>A. tonsa</i> macho. m) <i>Centropages hamatus</i> hembra. n) <i>C. hamatus</i> macho.	60
4.13	Comportamiento de la d_E para un ejemplar <i>Calanus pacificus</i> hembra.	62
4.14.	Resultado del preprocesado de una imagen.	62
4.15	Cambios del ángulo de rotación contra d_E .	63
4.16	Comportamiento estadístico para los diferentes copépodos de la Figura 4.12, usando como imagen de referencia al del inciso (a).	64
4.17	Firma vectorial promedio.	65
4.18	Comportamiento estadístico de la d_E , para la firma vectorial compuesta de diez imágenes de <i>Calanus pacificus</i> hembra (Grupo a).	66
4.19	Comportamiento estadístico de la d_E , para la firma vectorial compuesta de diez imágenes de <i>Centropages furcatus</i> hembra (Grupo e).	66
4.20	Resultados para identificación de ejemplares <i>Calanus pacificus</i> hembra.	68

LISTA DE FIGURAS (continuación)

Figura		Página
4.21	Mariposas utilizadas. (a) <i>Caligo eurylochus sulanos</i> . (b) <i>Caligo memnon</i> . (c) <i>Caligo oileus scamander</i> (Mariposa a).	69
4.22	Cambios del ángulo de rotación contra d_E .	70
4.23	Comportamiento estadístico de la d_E tomando como referencia a <i>Caligo eurylochus sulanos</i> .	70
4.24	Ejemplares (a) <i>Ansyluris jurgennseni</i> . (b) <i>Archaeoprepona amphimachu amphiktion</i> . (c) <i>Archaeoprepona demophon centralis</i> . (d) <i>Baeotus baeotus</i> .	71
4.25	Cambios del ángulo de rotación contra d_E .	72
4.26	Comportamiento estadístico de la d_E tomando como referencia a <i>Archaeoprepona amphimachus amphiktion</i> (Mariposa b).	72
5.1	Diagrama de bloques del procedimiento utilizado.	77
5.2.	Diagrama detallado del cálculo de las firmas vectoriales.	78
5.3.	Letras en colores.	79
5.4.	Comportamiento estadístico de la d_E tomando como referencia H_R .	79
5.5	Mariposas utilizadas. (1) <i>Actinote guatemalena guerrerensis</i> . (2) <i>Actinote stratonice oaxaca</i> . (3) <i>Agraulis vanillae incarnata</i> . (4) <i>Agrias amydon oaxacata</i> . (5) <i>Anae aidea</i> . (6) <i>Ancyluris inca mora</i> . (7) <i>Anetia thirza</i> . (8) <i>Ansyluris jurgennseni</i> . (9) <i>Arawacus sito</i> . (10) <i>Archaeoprepona amphimachus amphiktion</i> . (11) <i>Archaeoprepona demophon centralis</i> . (12) <i>Baeotus baeotus</i> . (13) <i>Basilarchia archippus</i> . (14) <i>Biblis hyperia aganissa</i> . (15) <i>Bolboneura Sylphis sylphis</i> . (16) <i>Caligo eurylochus sulanos</i> . (17) <i>Caligo memnon</i> . (18) <i>Caligo oileus scamander</i> .	82
5.6	Comportamiento estadístico de la d_E para rotación de las todas las mariposas, tomando como referencia a la mariposa 10.	83
5.7.	Comportamiento estadístico de la d_E para cambios de escala del 70 % al 130 %, tomando como referencia a la mariposa 3.	84

LISTA DE TABLAS

Tabla		Página
1.1	Algunas aplicaciones de los sistemas de reconocimiento de patrones.	14
2.1	Propiedades de la transformada de Fourier.	27
2.2	Propiedades de la transformada de Mellin.	31
4.1	Nivel de confianza porcentual para las letras A-M.	51
4.2	Nivel de confianza porcentual para las letras N-Z.	52
4.3	Nivel de confianza porcentual para cambios de escala.	54

CAPÍTULO I

INTRODUCCIÓN

- 1.1 Motivación.
- 1.2 Objetivos.
- 1.3 Antecedentes.
- 1.4 Orígenes del procesamiento digital de imágenes.
- 1.5 Reconocimiento de patrones.
- 1.6 Sistemas para reconocer objetos biológicos.
- 1.7 Estructura de la tesis.

1.1 Motivación

En la actualidad los sistemas de reconocimiento de patrones son aplicados en diferentes ámbitos. Sus fundamentos y adopción datan de finales del siglo XIX; hoy en día se emplean de forma rutinaria como medio de identificación.

A partir de los años sesenta del siglo pasado se empiezan a desarrollar sistemas de reconocimiento automático de patrones, cuyo uso ha permitido extender su utilización como medio de identificación para un gran número de aplicaciones, incluyendo aplicaciones civiles y militares.

En un principio, las técnicas de identificación automática de patrones trataron de emular las técnicas que los expertos utilizaban manualmente para analizar imágenes. Estas técnicas se basaban en encontrar los puntos singulares y su posición relativa dentro de la imagen para poder compararlas.

Hoy en día, los métodos de captura han avanzado considerablemente, proporcionando imágenes de muy alta calidad. Sin embargo, se necesitan nuevos métodos que sean capaces de comparar toda la información contenida en ellas y explotarla para realizar un reconocimiento más robusto.

Las diferentes técnicas utilizadas en la actualidad presentan dificultades, siendo la principal de ellas que los cálculos conllevan una alta complejidad computacional. Además, en este caso, la complejidad de la operación realizada es directamente proporcional a la resolución de las imágenes utilizadas. El coste computacional lleva inevitablemente asociado un incremento en el tiempo necesario para la comparación. Esto hace ineludible la utilización de nuevos procesadores para poder alcanzar una eficiencia razonable, lo que implica a su vez un incremento en el coste computacional en cualquier sistema comercial.

El considerar el tamaño de las imágenes y la posibilidad de que éstas contengan traslaciones, rotaciones y cambio de escala, hace que las actuales técnicas estén limitadas para obtener una respuesta en tiempo aceptable. En consecuencia y a pesar de su relevancia teórica, estas técnicas han recibido tradicionalmente poca atención y su estudio ha sido minoritario.

Por las razones expuestas anteriormente, el estudio y mejora de las técnicas de reconocimiento reviste un gran interés, particularmente en atención a las modernas generaciones de dispositivos de adquisición de imágenes, los cuales han experimentado un desarrollo extraordinario, a tal grado que la clásica imagen registrada en película fotográfica ha sido reemplazada por la imagen digital con una mayor resolución y calidad. Para estos casos es posible realizar una comparación de imágenes más precisa si los sistemas de reconocimiento consideran toda la información disponible.

Por otra parte, en los últimos años se tiene un desarrollo notable en los recursos de hardware, útil y eficiente para realizar tareas de procesamiento digital de señales, tales como multiplicadores y bloques de memoria, entre otros. Estas arquitecturas pueden contribuir a un incremento de la precisión de los algoritmos de comparación, ya que con ellas sería posible realizar más operaciones en un menor tiempo.

Cabe destacar que varias de las aplicaciones actuales, requieren una evaluación e interpretación de la información contenida en imágenes digitales. Dichos análisis tradicionalmente se llevan a cabo mediante la visión humana de técnicos entrenados o especialistas, y son susceptibles de mejora incrementando la objetividad y automatización, y es en esta dirección que se ha orientado el trabajo en esta tesis.

1.2 Objetivos

El objetivo fundamental es desarrollar un nuevo algoritmo que sea invariante a rotación, escala y posición, para el reconocimiento de objetos y la identificación de patrones en imágenes en escala de grises o color. De forma más detallada, los objetivos particulares se indican a continuación.

- Estudiar y analizar las propiedades de invariancia que poseen diferentes transformaciones matemáticas factibles de ser utilizadas para el reconocimiento de patrones.
- Proponer, diseñar e implementar un nuevo algoritmo invariante a rotación, escala y posición para el reconocimiento e identificación de patrones. Realizar la evaluación de la capacidad del sistema para reconocimiento entre pares de muestras, analizando su problemática específica y aportando soluciones que mejoren su precisión y eficiencia.
- Emplear el sistema de reconocimiento de imágenes en aplicaciones útiles para la ejecución de tareas de identificación de organismos, generalmente exclusivas de técnicos y especialistas (i.e. taxónomos).
- Desarrollar un sistema digital basado en la utilización de firmas vectoriales invariantes a traslación, rotación y escala, aplicadas al reconocimiento de patrones en imágenes en color y aplicarlo para el reconocimiento de imágenes reales.

En el planteamiento de estos objetivos, es conveniente destacar que se ha considerado reducir al mínimo la utilización de algoritmos de preprocesado con el fin de determinar la eficacia de las técnicas por sí mismas. Es evidente que su utilización correcta

proporciona generalmente una mayor calidad de imagen y por tanto favorece la comparación, pero también puede introducir términos que afecten negativamente dicha comparación.

1.3 Antecedentes

El término “reconocimiento de patrones” abarca una vasta área del procesamiento de información, desde el reconocimiento de voz o de escritura cursiva, hasta la detección de fallas en maquinaria o diagnóstico de enfermedades. Todos estos problemas los resolvemos los humanos de manera tan natural que nos puede ser difícil comprender qué tan compleja es esta tarea para las computadoras.

La mayoría de las personas aprenden a leer y escribir durante sus primeros años de vida. Conforme van avanzando en su vida académica, van adquiriendo nuevas habilidades en la lectura y escritura, incluyendo la capacidad de poder leer cualquier tipo de escritura, sea ésta clara o confusa, manuscrita o impresa, independientemente del tipo de fuentes y estilos que se utilicen.

La mayoría de las personas no tienen algún problema en la lectura de documentos, inclusive cuando se trata de fotocopias con poca calidad, impresiones al revés, caracteres o anuncios con decoraciones, dañados o con partes fragmentadas, palabras mal escritas, palabras escritas con estilo artístico o con diseños figurativos. A veces, los caracteres y las palabras pueden aparecer distorsionados y, sin embargo, la mayoría de la gente todavía puede descubrir de qué se trata. Por el contrario, a pesar de más de cinco décadas de investigación intensiva, la habilidad de lectura de la computadora está aún por detrás de la de los seres humanos.

Para entender los fenómenos descritos anteriormente, tenemos que buscar en la historia del reconocimiento óptico de caracteres (OCR por sus siglas en inglés *Optical Character Recognition*) su desarrollo, métodos de reconocimiento, las tecnologías de cómputo y las diferencias entre humanos y máquinas (D’ateur, 1971, Schantz, 1982,

Bunke y Wang, 1997, Mori *et al.*, 1999, Rice y Nartker, 1999, Suen, 1986, Cheriet, *et al.*, 2007).

Siempre es fascinante tratar de encontrar la forma en que una computadora puede ser capaz de imitar las funciones humanas: leer, escribir e identificar objetos. La investigación y desarrollo del OCR se remonta a la década de 1950, cuando los científicos trataron de captar las imágenes de caracteres y textos, primero utilizando métodos mecánicos y ópticos por medio de la rotación de discos y fotomultiplicadores; después, con el escáner de punto que utilizaba un punto de luz como fuente para la digitalización, seguido por fotoceldas y arreglos de ellas.

Al principio, la operación del escáner era lenta y sólo una línea de caracteres podían ser digitalizados en un momento, ya sea moviendo el escáner o el soporte del papel. Posteriormente, con las invenciones del tambor llegaron los escáneres planos, con lo que se extendió la exploración para una página completa. Los avances en circuitos integrados digitales permitieron tener mayor velocidad en la exploración y en la conversión digital. Estas importantes mejoras aceleraron enormemente la velocidad de reconocimiento de caracteres, y con la reducción de los costos se abrieron las posibilidades de procesar una gran variedad de formas y documentos.

A lo largo de los años 1960 y 1970, las nuevas aplicaciones de OCR surgieron en los negocios minoristas, bancos, hospitales, oficinas de correos, seguros, ferrocarriles, aeronaves, editores de periódicos y muchas otras industrias (D'auteur, 1971).

Cada una de las diferentes técnicas actuales aplicadas en el OCR, tiene sus fortalezas y debilidades. La tendencia en la investigación y en el uso de estas herramientas ha pasado de tratar de reconocer texto con tipografía manual y de máquina a ser un problema de reconocimiento de patrones en señales; no sólo es un campo de la computación sino un proceso fundamental que se encuentra en casi todas las acciones humanas.

El punto esencial del reconocimiento de patrones es la clasificación: se quiere clasificar una señal dependiendo de sus características. Señales, características y clases pueden ser de cualquier forma; por ejemplo, se pueden clasificar imágenes digitales de

letras en las clases “A” a “Z” dependiendo de sus píxeles, o ruidos de cantos de los pájaros en clases de órdenes aviares dependiendo de las frecuencias. El objetivo es clasificar patrones con base en un conocimiento a priori o información estadística extraída del patrón de referencia; los patrones a clasificar suelen ser grupos de medidas u observaciones, que nos definen puntos en un espacio multidimensional apropiado.

Un sistema de reconocimiento de patrones completo consiste en un sensor que recoge las observaciones a clasificar, un sistema de extracción de características que transforma la información observada en valores numéricos o simbólicos, y un sistema de clasificación o descripción que, basado en las características extraídas, clasifica la medición. Para realizarla se utiliza habitualmente uno de los siguientes procedimientos:

- Geométrico: En éste enfoque se emplea el cálculo de distancias, geometría de formas, vectores numéricos, puntos de atracción, etc.
- Estadístico: Se basa en la teoría de la probabilidad y la estadística, utiliza análisis de varianzas, covarianzas, dispersión, distribución, etc.

Para la clasificación se puede usar un conjunto de aprendizaje, del cual ya se conoce a priori la clase a la que pertenecen y que servirán para entrenar al sistema, siendo la estrategia resultante conocida como aprendizaje supervisado. Éste puede ser también no supervisado; el sistema no tiene un conjunto para aprender a catalogar la información de antemano, sino que se basa en cálculos estadísticos para clasificar los patrones.

Entre las aplicaciones del reconocimiento de patrones están la clasificación de documentos, el reconocimiento de voz, escritura, caras humanas y muchas más. Los dos últimos ejemplos son representativos del análisis de imágenes, un subconjunto del reconocimiento de patrones que toma imágenes digitales como entradas del sistema.

El procesamiento digital de señales (en inglés *digital signal processing*, DSP) es un área de la ingeniería que se dedica al análisis y procesamiento de señales (audio, voz, imágenes, video) que son discretas. Aunque comúnmente las señales en la naturaleza nos llegan en forma analógica, también existen casos en que éstas son de naturaleza digital.

El reconocimiento y clasificación de imágenes es una de las áreas más activas en el amplio campo de las ciencias de la imagen y la ingeniería. La razón es evidente: la necesidad de reemplazar la capacidad humana visual con una máquina es muy importante y hay diversas aplicaciones. La idea principal es inspeccionar una escena de la imagen mediante el procesamiento de datos obtenidos de los sensores. Tales máquinas pueden reducir sustancialmente la carga de trabajo y mejorar la precisión de la toma de decisiones de los operadores humanos en diversos campos, sistemas de ingeniería biomédica, vigilancia de la salud, cirugías, sistemas de transporte inteligentes, manufactura, robótica, entretenimiento y los sistemas de seguridad incluyendo a los militares.

El reconocimiento y clasificación de imágenes es un campo multidisciplinario. Ello requiere la contribución de diversas tecnologías y conocimientos en sensores, sistemas de imagen, algoritmos de procesamiento señal-imagen, hardware, software y de sistemas de integración.

Si bien ha habido avances significativos en el reconocimiento de imágenes y en las tecnologías de clasificación, existen todavía problemas técnicos y desafíos que afrontar en este campo. Estos incluyen grandes variaciones en la firma del objeto inspeccionado debido a las variaciones geométricas, condiciones ambientales, el envejecimiento, y el comportamiento del sensor (por ejemplo, las fluctuaciones de la firma de infrarrojos térmicos, ángulos de reflexión, etc.). Por lo tanto, el progreso es necesario en el avance de los sensores, los algoritmos y sistemas compactos que integran los sensores, el hardware y los algoritmos de software para proporcionar nuevas y mejoradas funciones para la alta velocidad del reconocimiento de imágenes y su clasificación.

En la actualidad, no existe una técnica única que pueda resolver todas las clases de problemas de clasificación y reconocimiento. En la mayoría de los casos, las soluciones que aparecen dependen de la aplicación y puede combinar varios de estos enfoques para adquirir los resultados deseados. De igual manera, es un área en la que participan investigadores de varias disciplinas como lo son ingenieros en sistemas eléctricos o con interés en procesamiento de imágenes, ingenieros ópticos, informáticos, ingenieros biomédicos, físicos aplicados, matemáticos aplicados, etc.

Uno de los beneficios principales del DSP es que las transformaciones de señales son más sencillas de realizar. Una de las más importantes transformadas que se utiliza es la transformada de Fourier discreta (TFD). Esta transformada convierte la señal del dominio del tiempo al dominio de la frecuencia, lo cual permite un análisis más sencillo y eficaz sobre la frecuencia, especialmente en aplicaciones de eliminación de ruido y en otros tipos de filtrado (pasa bajos, pasa altos, pasa banda, de rechazo de banda, etc.).

1.4 Orígenes del procesado digital de imágenes

A principios de 1920, se desarrolló una de las primeras aplicaciones de un sistema digital de transmisión de imágenes para su utilización en la industria de la prensa; éste fue conocido como el sistema de Bartlane, el cual contaba con equipos de impresión especializados que codificaban la imagen para transmitirla y luego reconstruirla al otro extremo, permitiendo enviar una fotografía mediante líneas telegráficas. Las primeras imágenes utilizaron el cable de transmisión submarino del Atlántico: el sistema podía transmitir una fotografía entre Nueva York y Londres en aproximadamente tres horas con la resolución suficiente para poder utilizarla en la impresión de un periódico gracias a lo cual se redujo el tiempo necesario para transportarla (usualmente era más de una semana).

Algunos de los problemas iniciales para mejorar la calidad visual de las imágenes se relacionan con la selección de los procedimientos de impresión y la distribución de los niveles de intensidad. En la figura 1.1 se muestra una imagen transmitida por el sistema de Bartlane e impresa utilizando un teletipo (sistema de transmisión de textos por línea telegráfica mediante un teclado que permite emitir y recibir mensajes e imprimirlos). Dicho método de impresión fue abandonado a finales de 1921 en favor de una técnica basada en la reproducción fotográfica a partir de las cintas perforadas que se reciben en la terminal del telégrafo. La figura 1.2 muestra un ejemplo utilizando dicho método.



Figura 1.1. Imagen digital producida en 1921 con una cinta codificada e impresa en un teletipo (Gonzalez, 2008).

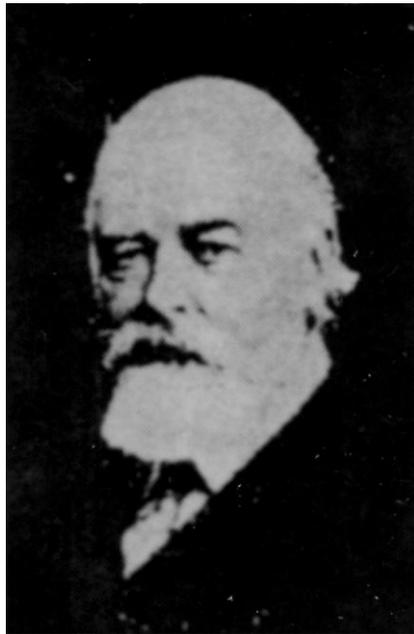


Figura 1.2. Una imagen digital realizada en 1922 de una cinta perforada después de que las señales habían cruzado el Océano Atlántico dos veces (Gonzalez, 2008).

Los primeros sistemas Bartlane eran capaces de codificar imágenes en cinco diferentes niveles de gris. Con posterioridad, en 1929, se ampliaron estos niveles a 15 tonos, gracias a los cuales una especie de máquina de escribir imprimía la fotografía al otro lado del cable al cabo de tres horas (Gonzalez, 2008).

Aunque los ejemplos citados implican imágenes digitales, no se pueden considerar resultados del tratamiento digital de imágenes en el contexto de nuestra definición, porque las computadoras no estuvieron involucradas en su creación. Las mejoras de los métodos para la transmisión de imágenes continuaron durante los 35 años siguientes y no fue sino hasta la aparición de las primeras computadoras digitales que el procesamiento de imágenes se consolidó como disciplina científica.

La historia del procesamiento digital de imágenes está íntimamente ligada al desarrollo de la computadora digital. De hecho, las imágenes digitales requieren tanta capacidad de almacenamiento y potencia computacional que los avances en el campo han estado supeditados al desarrollo de las computadoras y las tecnologías asociadas que incluyen el almacenamiento de datos, visualización y transmisión.

Sin embargo, la base de lo que llamamos una computadora digital moderna se remonta sólo a la década de 1940 con la introducción por parte de John von Neumann del concepto de programa almacenado, que permitió la lectura de un programa dentro de la memoria de la computadora, y después la ejecución de las instrucciones del mismo sin tener que volverlas a escribir. La primera computadora en usarlo fue la llamada EDVAC (*Electronic Discrete Variable Automatic Computer*, es decir “computadora automática electrónica de variable discreta”), desarrollada por von Neumann, Eckert y Mauchly. Los programas almacenados dieron a las computadoras flexibilidad y confiabilidad, haciéndolas más rápidas y menos sujetas a errores que los programas mecánicos.

Las primeras computadoras lo suficientemente potentes como para llevar a cabo el procesamiento de imágenes significativo aparecieron a comienzos de 1960. El nacimiento de lo que llamamos el procesamiento digital de imágenes se remonta a la disponibilidad de

las máquinas y al inicio del programa espacial durante ese período. Fue la combinación de estos dos acontecimientos lo que puso en relieve el potencial de los conceptos del procesamiento digital de imágenes.

Los trabajos sobre el uso de técnicas informáticas para mejorar las imágenes de una sonda espacial se iniciaron en el *Jet Propulsion Laboratory* (Pasadena, California) en 1964 cuando las primeras imágenes de la luna, transmitida por *Ranger 7* fueron procesadas por computadoras para corregir los diversos tipos de distorsión de la imagen inherente en la cámara de televisión a bordo. La figura 1.3 muestra la primera imagen de la luna tomada por el *Ranger 7*, el 31 de julio de 1964 (Gonzalez, 2008). Estas técnicas sirvieron como base a los métodos utilizados para la mejora y restauración de imágenes de los programas *Surveyor*, *Mariner* y *Apollo*. A partir de 1967 estas experiencias se aplicaron al campo de la salud pública con programas de investigación realizados por la escuela de medicina de UCLA, de las cuales se hizo eco el *National Institute of Health* que subvencionó al JPL en tres líneas de investigación: las imágenes en microscopía óptica, electrónica y de rayos X, para el estudio de la arteriosclerosis, el cáncer cervical, y otras patologías.

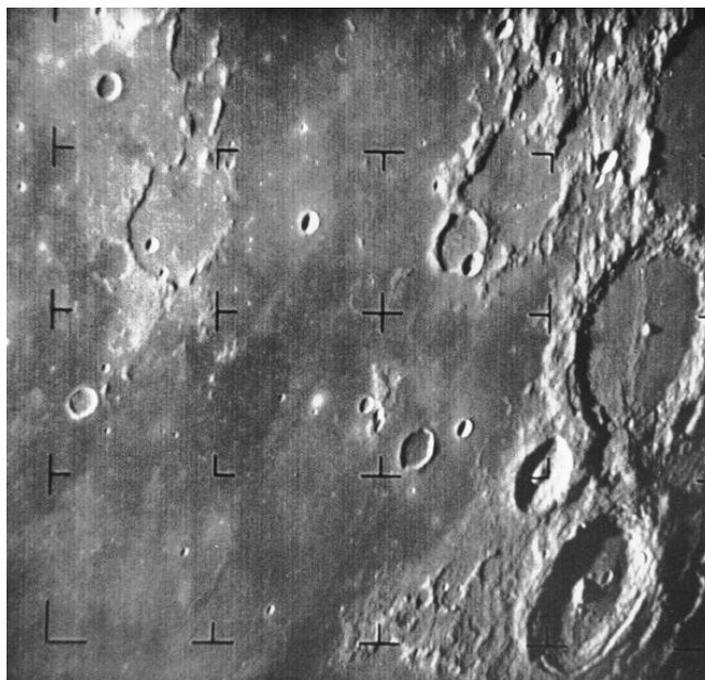


Figura1.3. La primera fotografía de la luna tomada de una nave espacial Estadounidense (Gonzalez, 2008).

1.5 Reconocimiento de patrones

Marques de Sá, (2001) define el término *reconocimiento de patrones* como “una rama del conocimiento, de carácter multidisciplinario, cuyo objeto de estudio son los procesos de identificación, caracterización, clasificación y reconstrucción sobre conjuntos de objetos o fenómenos, así como el desarrollo de teorías, tecnologías y metodologías relacionadas con dichos procesos”.

Dependiendo de la aplicación, estos objetos pueden ser imágenes, formas de onda de diferentes señales o cualquier tipo de mediciones que necesiten ser clasificadas. Nos referiremos a ellos con el término genérico de *patrones* (Theodoridis, 2009).

El reconocimiento de patrones es un área fértil de investigación, con múltiples enlaces a muchas otras disciplinas, que requiere la participación de profesionales de diversas áreas. La gran diversidad de tareas del reconocimiento de patrones que pueden ser automatizadas apunta a un repertorio de posibles acercamientos igualmente amplio. En las situaciones más simples, los patrones pueden estar representados por vectores de números obtenidos, por ejemplo, de muestrear una señal o de escanear una imagen en puntos determinados; las clases pueden estar representadas por uno o varios patrones prototipos.

Los patrones del mundo real muestran una gran variabilidad. Sin embargo, parece plausible suponer que los puntos que representen características similares, tenderán a agruparse en la misma zona del espacio. A la inversa, las representaciones de patrones distintos o de diferentes clases permanecerán en distintas regiones del espacio. Esta interpretación de patrones y clases fue el marco bajo el que se desarrollaron los primeros trabajos en este campo.

Las aplicaciones de los sistemas de reconocimiento de patrones y sus técnicas son numerosas y abarcan una amplia gama de actividades. En la tabla 1.1 se enumeran algunos ejemplos referentes a diversas actividades profesionales.

Tabla 1.1. Algunas aplicaciones de los sistemas de reconocimiento de patrones

<i>Disciplina</i>	<i>Aplicación</i>
Agricultura:	
	Análisis de cosechas.
	Evaluación de suelos.
Astronomía:	
	Análisis de imágenes telescópicas.
	Espectroscopia automatizada.
Biología:	
	Citología automatizada.
	Las propiedades de cromosomas.
	Estudios genéticos.
Administración municipal:	
	Análisis de tráfico vehicular y control.
	Evaluación de crecimiento urbano.
Economía:	
	Previsión de cambios en las acciones.
	Análisis del rendimiento empresarial.
Ingeniería:	
	Detección de defecto en la fabricación de productos.
	Reconocimiento de caracteres.
	Reconocimiento de voz.
	Sistemas automáticos de navegación.
	Análisis de contaminación.
Geología:	
	Clasificación de rocas.
	Estimación de los recursos mineros.
	Análisis de recursos utilizando imágenes de satélite.
	Análisis sísmico.
Medicina:	
	Análisis de electrocardiogramas.
	Análisis de electroencefalogramas.
	Análisis de imágenes médicas.
Militares:	
	Análisis de fotografías aéreas.
	Detección y clasificación de señales de radar y sonar.
	Reconocimiento automático objetivo.
Seguridad:	
	Identificación de huellas digitales.
	Sistemas de vigilancia y alarma.

A lo largo de todos estos años se han desarrollado diversas metodologías para el reconocimiento de patrones. En términos generales, éstos se pueden clasificar de la siguiente manera:

- Geométrico: Los patrones deben ser graficables; en este enfoque se emplea el cálculo de distancias, geometría de formas, vectores numéricos, puntos de atracción, etc.
- Estadístico: Se basa en la teoría de la probabilidad y la estadística, utiliza análisis de varianzas, covarianzas, dispersión, distribución, etc.
- Sintáctico-Estructural: Estudia la estructura de los objetos, es decir, usa teoría de lenguajes formales, gramáticas, teoría de autómatas, etc.
- Neuro-Reticular: Se utilizan redes neuronales que se “entrenan” para dar una cierta respuesta ante determinados valores.
- Lógico-Combinatorio: Se basa en la idea de que la modelación del problema debe ser lo más cercana posible a la realidad del mismo, sin hacer suposiciones que no estén fundamentadas. Se utiliza para conjuntos difusos y utiliza lógica simbólica, circuitos combinacionales y secuenciales, etc.

Históricamente, todas las medidas mencionadas anteriormente han tenido un gran valor en sus respectivos campos. Como se puede ver, existen numerosas formas de abordar un problema tan vasto como el reconocimiento automático de patrones. El estudio de todas ellas y su posible aplicación al problema que nos atañe, excede con mucho el propósito de esta tesis. Por ello, de ahora en adelante nos centraremos en el estudio del reconocimiento geométrico de patrones, base de este trabajo.

1.6 Sistemas para reconocer de manera invariante objetos biológicos

Desde la primera aparición de los filtros espaciales complejos u holográficos del tipo de Vander Lugt a mediados de la década de los sesenta del siglo XX, se han logrado avances sustanciales en el campo del reconocimiento de patrones ópticos, en particular con los métodos de correlación. Las primeras investigaciones para la identificación y conteo de microorganismos – en particular diatomeas (una especie de microalgas) mediante un microscopio – se realizaron en sistemas ópticos coherentes (Almeida *et al.*, 1972) y, utilizando la técnica de filtraje espacial óptico con filtros holográficos (Cairns *et al.*, 1972), lograron reconocer patrones de varias diatomeas. Continuando con esta misma línea de trabajo, en 1976 Almeida y Kim-Tzong presentaron resultados aplicando la técnica de correlación óptica convencional con filtros espaciales acoplados (hogramas de transformada de Fourier) tanto en la identificación de especímenes biológicos, como en el monitoreo de la contaminación del agua, la cual está directamente relacionada con el número de diatomeas contenidas en una muestra dada, y contadas en función de las especies en un cierto período de tiempo.

Fujii y Almeida (1979a) elaboraron filtros espaciales acoplando patrones simulados y obteniendo correlaciones parciales ópticas entre el patrón simulado (contornos, contornos con estrías y características morfológicas de importancia taxonómica) y los microorganismos, mostrando que este filtro es menos sensible a las variaciones de tamaño del objeto, lo cual reduce el número de filtros requeridos para el análisis; en este caso sólo estuvieron considerando el contorno de las diatomeas. El problema consiste en que un taxónomo no solamente reconoce el contorno del objeto, sino también su estructura interna, como son la densidad de estrías, la dirección y formas de éstos. Generalmente estos organismos nunca son iguales; por ejemplo, las diatomeas que pertenecen a una misma especie son muy similares en forma pero pueden ser diferentes en tamaño y en el patrón de estrías, por lo que es muchas veces más importante encontrar similitudes en vez de diferencias para la identificación o clasificación de dichos organismos.

Para tratar de vencer el problema del tamaño de los microorganismos, Fujii *et al.* (1980) desarrollaron un método para el reconocimiento de formas microbiológicas, rotando el filtro espacial acoplado mediante un prisma de cuña motorizado y controlado por una computadora. Este análisis está basado en el conteo de los píxeles de los picos de correlación con un programa de computación sencillo.

Los esfuerzos por lograr la realización de un dispositivo óptico para la identificación y conteo de microorganismos por los métodos anteriormente descritos no fueron satisfactorios debido, principalmente, a que utilizaron la correlación óptica convencional, la cual tiene el inconveniente de ser muy sensible a pequeños cambios de escala y orientación de los objetos (Casasent y Psaltis, 1976a).

Casasent y Psaltis (1976b), mostraron que si una de las imágenes se rota en 3 % y se varía su tamaño en 2 %, entonces la razón señal a ruido (SNR) del pico de correlación entre la función de entrada u objeto y el filtro espacial holográfico decae por un factor de 10. Esto significa que la habilidad de comparar las imágenes con esta técnica convencional de correlación es muy pobre y éstas aparecen muy borrosas.

Para vencer dichos problemas, Casasent y Psaltis (1976c) propusieron el uso de la transformada de Mellin, que por sus propias características es invariante a la escala pero no a translación. Ya que la operación básica realizada por una lente en un sistema óptico es la transformada de Fourier (Goodman, 1996), tenemos que realizar la transformada de Mellin ópticamente vía transformada de Fourier (la cual es invariante a posición). Sin embargo es posible combinar las dos transformaciones para realizar la operación de correlación óptica invariante a escala y corrimiento; también demostraron la invariancia a rotación. Sus resultados confirmaron que es posible correlacionar ópticamente dos funciones u objetos que difieran grandemente en escala y rotación.

Schwartz (1977) muestra en sus estudios cómo el cerebro tiene la habilidad de reconocer las formas de los objetos. La luz del mundo exterior que incide sobre la retina (tejido sensible a la luz situado en la superficie interior del ojo) pasa a través del nervio óptico en forma de impulsos eléctricos; cada célula retinal está unida con una neurona en la

corteza visual cerebral, de tal manera que la imagen es mapeada o proyectada sobre una superficie del cerebro y descrita por una transformación geométrica del espacio de coordenadas cartesianas a coordenadas polares con escalamiento logarítmico. Schwartz (1977) señala cómo el mapeo de la escena visual a través del logaritmo complejo es importante para el reconocimiento de patrones y hace referencia a la solución propuesta en los trabajos de Casasent y Psaltis (1976c) de utilizar la transformación de coordenadas de la función de entrada u objeto al plano polar con escalamiento logarítmico para poder realizar la operación de correlación. Los estudios realizados por Schwartz (1994) muestran que existe una fuerte evidencia fisiológica del mapeo polar logarítmico entre la retina y la corteza visual en muchos sistemas visuales incluyendo el humano.

Todas las ideas expresadas anteriormente se han aplicado por parte de integrantes de la comunidad científica para obtener un sistema de monitoreo automático de organismos; por ejemplo, Pech-Pacheco y Álvarez-Borrego, J., (1998), hicieron estudios en la identificación de cinco especies de fitoplancton del género *Ceratium*, utilizando un sistema óptico digital y aplicando la técnica de correlación invariante descrita por Casasent y Psaltis (1976a, b, c); la razón de haber elegido este género es por la dificultad que representa para el taxónomo el discriminar sus especies. Siguiendo con estos estudios, Castro-Longoria *et al.*, (2001) identificaron especies de Copépodos *Calanoides* usando el algoritmo del módulo al cuadrado de la transformada de Fourier. Forero-Vargas *et al.*, (2002) usaron un método basado en la extracción de imagen por proyección del mapeo polar logarítmico a vectores en 1D aplicado a imágenes de bacterias de tuberculosis.

Álvarez-Borrego *et al.*, (2002) usaron la técnica de correlación invariante para identificar la bacteria del cólera *Vibrio cholerae* O1. Las imágenes de éstas fueron grabadas con una cámara CCD y descompuestas en tres canales de colores: rojo, verde y azul, llamados RGB. Con esta técnica realizaron una simulación numérica para identificar esta bacteria en diferentes muestras, utilizando la transformada de escala digital mediante la transformada de Mellin para realizar un proceso que garantizara un alto grado de discriminación de los filtros, tomando en cuenta sólo la información del color y la forma.

Álvarez-Borrego y Castro-Longoria (2003) hicieron una discriminación entre las especies *Acartia* (crustáceos marinos microscópicos) utilizando sus patrones de difracción en una correlación digital invariante a rotación y posición.

La identificación y conteo de microorganismos a través de un microscopio ha sido y sigue siendo un método tedioso para el taxónomo, ya que le consume bastante tiempo de trabajo con el consecuente cansancio físico y mayores errores en sus resultados. Además, el error en la identificación es diferente para cada persona, ya que se involucran criterios de errores subjetivos visuales para reconocer una especie, por lo tanto, esta técnica es impráctica y poco eficiente. Lo deseable sería que existieran sistemas automáticos para el conteo e identificación de objetos en tiempo real, reduciendo así considerablemente los márgenes de error y el tiempo de análisis.

1.7 Estructura de la tesis.

El presente trabajo se ha organizado en seis capítulos. En este primer capítulo se explica la motivación, objetivos y antecedentes.

En el segundo capítulo se resume el marco teórico en el cual se sustenta el trabajo desarrollado. Se define el concepto de transformada integral, posteriormente se muestran las definiciones y principales características de las transformadas de Fourier, Mellin y Escala, que luego van a jugar un papel importante cuando las apliquemos al reconocimiento de patrones.

En el tercer capítulo presenta el algoritmo para el reconocimiento de patrones, con la característica de ser invariante a cambios de posición, rotación y escala, proponiendo una técnica novedosa para la comparación de imágenes, basada en la utilización de firmas vectoriales.

El cuarto capítulo expone los resultados obtenidos del trabajo de investigación utilizado en imágenes en escala de grises. Los conjuntos que se analizaron corresponden a imágenes de letras del abecedario, números, formas de hojas e imágenes reales de ejemplares en etapa adulta de diferentes especies de copépodos; los resultados corresponden a las simulaciones realizadas para cambios en la escala y rotación de las imágenes. Seguidamente se muestra el análisis realizado para imágenes en color que fueron transformadas a escala de grises.

En el quinto capítulo se encuentra el análisis de imágenes de color, utilizando una variante del algoritmo presentado en el tercero. Los conjuntos de imágenes utilizados para su estudio corresponden a letras escritas en color y diferentes especies de mariposas. Los resultados obtenidos en el capítulo anterior y en éste han sido ya publicados en revistas arbitradas y/o presentados en congresos internacionales y nacionales los cuales se reseñan en la última sección.

Finalmente, en el sexto capítulo se detallan las principales conclusiones resumiendo los logros alcanzados.

CAPÍTULO II

MARCO TEÓRICO

2.1 Introducción.

2.2 Antecedentes de la transformada de Fourier.

2.3 Transformadas Integrales.

2.4 Transformada de Fourier en una dimensión.

2.4.1 Transformada bidimensional y su inversa.

2.5 Transformada de Mellin.

2.5.1 Definición y relación con otras transformaciones.

2.5.2 Transformada de Fourier-Mellin.

2.6 Transformada de Escala.

2.1 Introducción

En este capítulo se repasan algunos conceptos básicos, como la definición y propiedades de las transformadas de Fourier, Mellin y Escala, herramientas que son de gran utilidad en: tratamiento de imagen y audio, reducción de ruido en señales (como el ruido blanco), análisis en frecuencia de cualquier señal discreta, análisis de materiales y estadística, síntesis mediante la transformada inversa de Fourier, etc.

Desde la década de los sesentas la comunidad científica en el área de óptica ha utilizado la transformada de Fourier para el reconocimiento de patrones, y de manera reciente ésta ha sido empleada como un instrumento para el procesamiento digital de señales e imágenes, aprovechando que el espectro resultante de algunas de ellas es invariante a la rotación, translación o escala.

2.2 Antecedentes de la transformada de Fourier

El físico y matemático Jean-Baptiste Joseph Fourier nació en 1760 en el pueblo de Auxerre, localizado entre París y Dijon. La mayor contribución por la cual es recordado se esbozó en 1807 en unas memorias y fue publicada en 1822 en su célebre libro "*La Théorie Analytique de la Chaleur*" (teoría analítica del calor). Seguidor de la teoría matemática de la conducción del calor, estableció la ecuación diferencial parcial que gobierna la difusión del calor solucionándola con el uso de series infinitas de funciones trigonométricas. Dicho trabajo introduce la representación de una función como una serie de senos y cosenos, ahora conocidas como las series de Fourier. Desde entonces, las series han sido aplicadas a diversas ramas de la ciencia y son una de las principales herramientas para el análisis de las señales; la transformada (o integral) de Fourier es una simple extensión de las series.

La transformada discreta de Fourier se popularizó durante los años 40 y 50 al tiempo que se desarrollaron las computadoras. En 1965 Cooley y Tukey publican un artículo poco prometedor titulado "*Un algoritmo de las series complejas de Fourier para máquinas computadoras*" o de forma más refinada DFT. En él se explotan algunas

propiedades especiales y proporciona un nuevo algoritmo conocido como la transformada rápida de Fourier (Fast Fourier Transform o FFT) el cual reduce de forma considerable el número de multiplicaciones requeridas para el cálculo. Reduciendo en un factor de 100 las operaciones, la FFT representó todo un hito en el desarrollo del procesamiento digital de las señales. La fuente original de esta técnica parece provenir del matemático alemán Gauss en 1806, pero fueron Cooley y Tukey los que hicieron popular el uso de la FFT.

2.3 Transformadas Integrales

Gran número de importantes funciones del análisis matemático pueden expresarse como integrales de la forma

$$g(\alpha) = \int_{x_1}^{x_2} f(x)K(\alpha, x) dx . \quad (2.1)$$

Una función g definida por la ecuación (2.1) (en la que la variable α puede ser real o compleja) se llama transformada integral de f . La función K se denomina kernel o núcleo de la transformada. Como un ejemplo de las transformadas integrales, podemos mencionar la Transformada de Fourier.

2.4. Transformada de Fourier en una dimensión

La transformada de Fourier $\mathcal{F}(u)$ de una función continua $f(x)$ de una sola variable, está definida como

$$\mathcal{F}(u) = \int_{-\infty}^{\infty} f(x)e^{-j2\pi ux} dx, \quad (2.2)$$

donde $j = \sqrt{-1}$. Consecuentemente, la transformada de Fourier inversa (IFT) de $\mathcal{F}(u)$, resulta en la función original, descrita por

$$f(x) = \int_{-\infty}^{\infty} \mathcal{F}(u) e^{j2\pi ux} du . \quad (2.3)$$

Las ecuaciones (2.2) y (2.3) son conocidas como par de transformaciones de Fourier; existen si $f(x)$ es continua e integrable y $\mathcal{F}(u)$ es integrable.

Para la función discreta $f(x)$, $x=0,1,\dots, M-1$, la transformada de Fourier discreta (DFT) está dada por

$$\mathcal{F}(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) e^{-j2\pi \frac{ux}{M}} , \quad (2.4)$$

para $u = 0, 1, 2, 3, \dots, M-1$.

De manera similar al caso de funciones continuas, dada $\mathcal{F}(u)$, podemos recuperar la función original $f(x)$, haciendo uso de la transformada de Fourier discreta inversa (IDFT), esto es

$$f(x) = \sum_{u=0}^{M-1} \mathcal{F}(u) e^{j2\pi \frac{ux}{M}} , \quad (2.5)$$

con $x = 0, 1, 2, 3, \dots, M-1$.

Las ecuaciones (2.4) y (2.5) son el par de la transformada discreta de Fourier. Como se puede ver, son muy similares a las ecuaciones (2.2) y (2.3), con la diferencia de que las integrales han sido remplazadas por sumatorias y escaladas por el factor $\frac{1}{M}$; este factor en algunas ocasiones es colocado en la ecuación (2.5) en vez de la ecuación (2.4).

La transformada de Fourier tiene la propiedad de ser invariante a translación. Si consideramos dos funciones $f(x)$ y $f_T = f(x - x_0)$, donde f_T es la función $f(x)$ con un término x_0 de desplazamiento, la relación entre las transformadas de Fourier de las dos funciones está dada por

$$\mathcal{F}[f(x - x_0)] = e^{-j2\pi ux_0} \mathcal{F}(u), \quad (2.6)$$

por lo tanto

$$|\mathcal{F}[f(x)]|^2 = |\mathcal{F}[f(x - x_0)]|^2. \quad (2.7)$$

Al escalar la función $f(x)$ por una constante (a) real distinta de cero, la transformada de Fourier de $f(ax)$ es

$$\mathcal{F}[f(ax)] = \frac{1}{|a|} \mathcal{F}\left(\frac{u}{a}\right), \quad (2.8)$$

por lo cual la transformada de Fourier (2.2) no es invariante a cambios de escala.

2.4.1 Transformada bidimensional y su inversa

La transformada de Fourier puede extenderse fácilmente para una función $f(x,y)$ de dos variables (Goodman, 1996). Si $f(x,y)$ es continua e integrable y $F(u,v)$ es integrable, se tiene los siguientes pares de transformaciones de Fourier:

$$\mathcal{F}(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy, \quad (2.9)$$

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{F}(u, v) e^{j2\pi(ux+vy)} du dv. \quad (2.10)$$

La extensión a dos dimensiones de la transformada discreta de Fourier unidimensional, se hace de manera similar que en el caso de señales continuas. Sea $f(x, y)$ una función bidimensional (imagen) de tamaño $M \times N$, su transformada de Fourier está dada por

$$\mathcal{F}(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)}, \quad (2.11)$$

para valores de $u = 1, 2, 3, \dots, M - 1$ y $v = 1, 2, 3, \dots, N - 1$. De manera similar, si tenemos $\mathcal{F}(u, v)$, podemos obtener $f(x, y)$ por medio de la transformada discreta de Fourier inversa, dada por

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \mathcal{F}(u, v) e^{j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)}, \quad (2.12)$$

para valores de $x = 1, 2, 3, \dots, M - 1$ y $y = 1, 2, 3, \dots, N - 1$.

De la misma manera que en el caso unidimensional, la ubicación del factor $\frac{1}{MN}$ no es importante, y en ocasiones también se encuentra dividido entre la transformada y su inversa como $\frac{1}{\sqrt{MN}}$.

Si consideramos dos funciones $f_1(x, y)$ y $f_2(x, y)$ cuyas transformadas de Fourier son $\mathcal{F}_1(u, v)$ y $\mathcal{F}_2(u, v)$ respectivamente, en la tabla 2.1 se muestran algunas de sus propiedades principales.

Tabla 2.1. Propiedades de la transformada de Fourier.

Propiedad	
Linealidad	$\mathcal{F}[\alpha f_1(x, y) + \beta f_2(x, y)] = \alpha \mathcal{F}_1(u, v) + \beta \mathcal{F}_2(u, v)$
Escala	$\mathcal{F}[f(ax, by)] = \frac{1}{ ab } \mathcal{F}\left(\frac{u}{a}, \frac{v}{b}\right)$
Desplazamiento	$\mathcal{F}[f(x - x_0, y - y_0)] = e^{-j2\pi(ux_0 + vy_0)} \mathcal{F}(u, v)$
Conjugado simétrico	$\mathcal{F}(u, v) = \mathcal{F}^*(-u, -v)$
Simetría en magnitud	$ \mathcal{F}(u, v) = \mathcal{F}(-u, -v) $
Diferenciación en el dominio espacial	$\mathcal{F}\left[\frac{\partial^n f(u, v)}{\partial x^n}\right] = (ju)^n \mathcal{F}(u, v)$
Diferenciación en el dominio de la frecuencia	$\mathcal{F}[(-jx)^n f(x, y)] = \frac{\partial^n \mathcal{F}(u, v)}{\partial u^n}$
Transformada de la integral	$\mathcal{F}\left[\int_{-\infty}^x f(\tau) d\tau\right] = \frac{\mathcal{F}(u)}{ju} + \pi \mathcal{F}(0)\delta(u)$
Transformada de la convolución	$\mathcal{F}[f_1(x, y) * f_2(x, y)] = \mathcal{F}_1(u, v)\mathcal{F}_2(u, v)$
Teorema de Parseval	$\int_{-\infty}^{\infty} f(x) ^2 dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(u) ^2 du$

2.5 Transformada de Mellin

En contraste con las transformadas de Fourier y Laplace, que se introdujeron para resolver problemas físicos, la transformada de Mellin surgió en un contexto matemático. De hecho, la primera aparición de la transformación se encuentra en una memoria de Riemann, donde la utilizó para estudiar la famosa función zeta. Las referencias relativas a este trabajo y sus detalles están dados por Titchmarsh (1975). Sin embargo, es el matemático finlandés R.H. Mellin (1854-1933) quién fue el primero en dar una formulación sistemática de la transformación y su inversa. Trabajando en la teoría de funciones especiales, desarrolló aplicaciones para la solución de las ecuaciones diferenciales hipergeométricas y la derivación de las expansiones asintóticas. Una biografía de R.H. Mellin, incluido un esbozo de sus obras, puede encontrarse en el libro escrito por Elfving (1981).

En realidad, la transformación de Mellin también puede ser colocada en otro marco que, en algunos aspectos, se ajusta más a las ideas originales de Riemann. Los trabajos desarrollados por Vilenkin (1968) y Reid (1992), nos muestran este enfoque, la transformación es vista como una transformación de Fourier en el grupo multiplicativo de los números reales positivos y su evolución es paralela a la presentación teórica de grupo de la transformada de Fourier habitual.

2.5.1 Definición y relación con otras transformaciones

Sea $f(t)$ una función definida en el intervalo $0 < t < I$. La transformación de Mellin \mathcal{M} es la operación del mapa de f en la función F definida en el plano complejo por la relación

$$\mathcal{M}(s) \equiv F(s) = \int_0^{\infty} f(t)t^{s-1}dt. \quad (2.13)$$

La función $\mathcal{M}(s)$ se le conoce como la transformada de Mellin de f . En general, la integral existe sólo para valores complejos $s = a + jb$ tal que $a_1 < a < a_2$, donde a_1 y a_2 dependen de la función $f(t)$ a transformar. Esto introduce lo que se denomina la franja de definición de la transformación de Mellin que se denota por $S(a_1, a_2)$. En algunos casos, esta banda se puede extender a un semiplano ($a_1 = -\infty$ o $a_2 = \infty$) o para todo el plano complejo s ($a_1 = -\infty$ y $a_2 = \infty$).

La transformada de Mellin está estrechamente relacionada con una forma extendida de la transformada de Laplace, con el cambio de variable definido por

$$t = e^{-x}, \quad dt = -e^{-x} dx, \quad (2.14)$$

que convierte la integral (2.13) en

$$F(s) = \int_{-\infty}^{\infty} f(e^{-x}) e^{-sx} dx. \quad (2.15)$$

Al definir la función

$$g(x) \equiv f(e^{-x}), \quad (2.16)$$

se reconoce en (2.15) la transformada de Laplace \mathcal{L} de g

$$\mathcal{L}(s) = \int_0^{\infty} g(x) e^{-sx} dx. \quad (2.17)$$

De manera simbólica se escribe como

$$\mathcal{M}[f(t); s] = \mathcal{L}[f(e^{-x}); s]. \quad (2.18)$$

Para obtener la relación entre la transformada de Mellin y la de Fourier, ahora escribimos $s = a + 2\pi j\beta$ en (2.15), esto es

$$F(s) = \int_{-\infty}^{\infty} f(e^{-x})e^{-ax}e^{-j2\pi\beta x} dx. \quad (2.19)$$

En notación compacta la ec. (2.19) se escribe como

$$\mathcal{M}[f(t); a + j2\pi\beta] = \mathcal{F}[f(e^{-x})e^{-ax}; \beta], \quad (2.20)$$

donde \mathcal{F} representa la transformada de Fourier definida por

$$\mathcal{F}[f; \beta] = \int_{-\infty}^{\infty} f(x)e^{-j2\pi\beta x} dx. \quad (2.21)$$

La transformada de Mellin es especialmente útil por ser invariante a la escala (Bracewell, 1978). 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. En la tabla 2.2 se presentan las principales propiedades de la transformada de Mellin.

Tabla 2.2. Propiedades de la transformada de Mellin.

<i>Propiedad</i>	
Escala	$\mathcal{M}[f(ax)] = \frac{1}{a^{-u}} F(u)$
Multiplicación por x^a	$\mathcal{M}[x^a f(x)] = F(u + a)$
El aumento de la variable independiente a una potencia real	$\mathcal{M}[f(x^a)] = a^{-1} F\left(\frac{u}{a}\right), a > 0$
Inversa de la variable independiente	$\mathcal{M}[x^{-1} f(x^{-1})] = F(1 - u)$
Multiplicación por $\ln(x)$	$\mathcal{M}[\ln(x)f(x)] = \frac{d}{du} F(u)$
Derivada	$\mathcal{M}\left[\frac{d^k}{du^k} f(x)\right] = (-1)^k (u - k)_k F(u - k),$ <p>donde $(u - k)_k = \frac{(u-1)!}{(u-k-1)!} = \frac{\Gamma(u)}{\Gamma(u-k)}$</p>

2.5.2 Transformada de Fourier-Mellin

La transformada de Fourier-Mellin (TFM) se refiere principalmente al estudio de la similitud ante transformaciones (producto directo de la rotación y los cambios de la escala).

El análisis armónico en el grupo proporciona la base para la definición de otras transformadas vía la transformada de Fourier adaptadas a grupos más generales de transformaciones geométricas. En las últimas décadas esta transformada ha sido estudiada de manera extensa en el campo de procesamiento de imágenes.

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

$$\mathcal{M}(ju, jv) = \int_0^{\infty} \int_0^{\infty} f(x, y) x^{-ju-1} y^{-jv-1} dx dy. \quad (2.22)$$

Siendo sencilla la generalización a dos variables sólo consideraremos el caso unidimensional

$$\mathcal{M}(jw) = \int_0^{\infty} f(x) x^{-jw-1} dx. \quad (2.23)$$

Por comodidad, de aquí en adelante $\mathcal{M}(jw)$ será escrita como $\mathcal{M}(w)$, involucrando la operación de transformada de Mellin vía la transformada de Fourier, para tener la invariancia a escala y a desplazamiento. Esto es posible debido a que cualquier transformada integral que cumpla con la condición de linealidad, puede expresarse en términos generales como la ec. (2.1) (Arfken, 1981). Para la transformada de Mellin tendremos

$$g(\alpha) = \int_{-\infty}^{\infty} f(x) x^{\alpha-1} dx, \quad (2.24)$$

donde $x^{\alpha-1}$ es el núcleo de la transformación. Podemos entonces pasar de la transformada integral de Mellin a la transformada de Fourier, haciendo un cambio de variable de la forma $x = e^q$ en la ecuación (2.23). Entonces tenemos

$$\mathcal{M}(jw) = \int_0^{\infty} f(e^q) e^{-jwq} dq. \quad (2.25)$$

La transformada de Mellin de la función $f(x)$ queda escrita ahora como la transformada de Fourier de la función $f(e^q)$. De esta forma tenemos la capacidad de realizar la operación de transformada de Mellin vía la transformada de Fourier.

Una característica de esta transformada es que los módulos de las transformadas de Mellin de dos funciones que difieren en tamaño, son iguales (Casasent y Psaltis, 1976a). Esta invariancia en la escala, como vimos anteriormente, no se presenta en la transformada de Fourier.

$$|\mathcal{M}[f(x)]|^2 = |\mathcal{M}[f(ax)]|^2. \quad (2.26)$$

$$|\mathcal{M}[f(x)]| \neq |\mathcal{M}[f(x - x_0)]|. \quad (2.27)$$

donde a es el factor de escala. Así que, para realizar la correlación invariante a cambios de escala y corrimiento de la función de entrada, es necesario realizar primeramente la transformada de Mellin de la función de entrada y después obtener la transformada de Fourier de ésta. Lo anterior equivale a decir que la invariancia simultánea a escala y corrimiento puede expresarse como $\mathcal{F}\{\mathcal{M}[f(x)]\}$

2.6 Transformada de Escala

En este trabajo vamos a utilizar la transformada de Escala desarrollada por Cohen (1993), quien encontró que ésta es una restricción de la transformada de Mellin (Casasent, 1976c), cuya propiedad principal es ser invariante a los cambios de escala.

Si se define c como la variable de escala, entonces la transformada de Escala y su inversa están dadas de acuerdo a Pech-Pacheco *et al.* (2001), por las ecuaciones:

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(x) x^{-\frac{1}{2}} e^{-jc \ln x} dx, \quad (2.28)$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} D(c) x^{-\frac{1}{2}} e^{jc \ln x} dc; \quad x \geq 0. \quad (2.29)$$

Ahora consideremos dos funciones $f(x)$ y $f_{ESC} = \sqrt{a}f(ax)$, donde f_{ESC} es la versión de $f(x)$ aumentada o reducida de tamaño. Tenemos una ampliación o reducción dependiendo del rango de valores de a (si $0 \leq a \leq 1$ entonces se amplifica y si $1 \leq a \leq \infty$ se reduce). La relación entre las transformadas de Escala de las dos funciones está relacionada por

$$D_{ESC}(x) = e^{j\ln(a)}D(c). \quad (2.30)$$

Por lo tanto

$$|D_{ESC}(c)|^2 = |D(c)|^2. \quad (2.31)$$

Esto demuestra que la transformada de Escala es insensible a cambios en la escala, el espectro de variancia de las dos funciones es igual, independientemente del valor del factor a .

La transformada de Escala puede reescribirse como

$$D_s(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f_s(x) x^{(-jc - \frac{1}{2})} dx. \quad (2.32)$$

Como se puede ver, ésta es la transformada de Mellin con el argumento complejo $-jc + \frac{1}{2}$. Para realizar esta transformada, hacemos el mapeo logarítmico en la función de entrada, seguida de una transformada de Fourier; como sabemos que existe una relación directa entre la transformada de Mellin y la transformada de Fourier, entonces es fácil notar que la transformada de escala – expresada como una transformada de Mellin – la podemos reescribir en términos de la transformada de Fourier al tomar la siguiente función

$$f_s(x) = \frac{1}{\sqrt{x}} f(\ln x), \quad (2.33)$$

y sustituirla en (2.32), obtenemos

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \frac{1}{\sqrt{x}} f(\ln x) x^{(-jc-\frac{1}{2})} dx, \quad (2.33)$$

haciendo uso de las propiedades de los exponentes, la ecuación (2.33) se puede escribir como

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \frac{1}{\sqrt{x}} f(\ln x) x^{(-jc)} x^{(-\frac{1}{2})} dx, \quad (2.34)$$

simplificando

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(\ln x) x^{(-jc)} \frac{dx}{x}. \quad (2.35)$$

Haciendo un cambio de variable $u = \ln(x)$, $du = \frac{dx}{x}$, por lo tanto los límites son ahora de $-\infty$ a ∞ , entonces (2.35) se reescribe como

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{u(-jc)} du. \quad (2.36)$$

Esto es $D(c) = \mathcal{F}(c)$.

En dos dimensiones la transformada de Escala y su inversa se pueden escribir de la forma (Cristobal y Cohen, 1996).

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \int_0^{\infty} f(x, y) \frac{e^{(-jc_x \ln x - jc_y \ln y)}}{\sqrt{xy}} dx dy, \quad (2.37)$$

$$f(x, y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(c_x, c_y) \frac{e^{(jc_x \ln x + jc_y \ln y)}}{\sqrt{xy}} dc_x dc_y. \quad (2.38)$$

CAPÍTULO III

METODOLOGÍA

- 3.1 Introducción.
- 3.2 Implementación.
- 3.3 Aplicación del algoritmo.

3.1 Introducción

Este capítulo se enfoca en la descripción del algoritmo desarrollado, basado en la utilización de firmas vectoriales como método para el reconocimiento de imágenes, considerando que éstas poseen características para realizar el reconocimiento invariante a cambios de posición, escala y rotación.

Dada la diversidad de formas y tamaños que presentan tanto organismos vivientes como objetos inertes, ha surgido la necesidad de buscar sistemas automatizados de identificación tanto en la industria como en el quehacer científico.

En los últimos años las transformadas han sido utilizadas como una herramienta para el procesamiento digital de señales e imágenes. Recientemente se han publicado varios libros (Pratt, 2007, González *et al.*, 2008 y 2009, Cheriet *et al.*, 2007, y Obinata, 2007), que nos muestran los avances y el desarrollo del reconocimiento de patrones, así como las diferentes herramientas para el preprocesado de imágenes, la extracción, selección y creación de características, métodos de clasificación, utilización de diferentes tipos de transformadas, etc. Sin embargo, algunos trabajos se han centrado en resolver el problema de la invariancia a la rotación y escala (Cohen, 1993, Casasent, 1976a y Pech-Pacheco *et al.*, 2001).

Para la implementación práctica de la teoría descrita en el capítulo anterior se debe considerar que de acuerdo a Casasent y Psaltis (1976 a,b,c) es más fácil hacer un sistema que sea invariante a escala por medio de la manipulación de la transformada de Fourier que cambiando directamente la función de entrada. También debemos considerar que la realización práctica de la transformación de Mellin está dada por un mapeo logarítmico de la escena de entrada seguida de una transformada de Fourier. Schwartz (1994) encontró que existe una fuerte evidencia fisiológica y psicofísica en muchos sistemas visuales (incluyendo el humano) del uso de este tipo de mapeo logarítmico entre la retina y la corteza visual.

3.2 Implementación

En este trabajo utilizaremos una transformada de Escala bidimensional que definimos en el capítulo anterior por la ecuación (2.37) escrita ahora en coordenadas polares $D(c_r, c_\theta)$, debido a que esta transformada es invariante a los cambios de tamaño y rotación

$$D(c_r, c_\theta) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \int_0^{2\pi} f(r, \theta) r^{(-jc_r-1/2)} e^{(-jc_\theta\theta)} dr d\theta. \quad (3.1)$$

Tomando el logaritmo de la coordenada radial, es decir $\lambda = \ln r$, tenemos

$$D(c_\lambda, c_\theta) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \int_0^{2\pi} e^{\lambda/2} f(\lambda, \theta) r^{(-jc_r-1/2)} e^{-j(\lambda c_\theta + \theta c_\theta)} d\lambda d\theta. \quad (3.2)$$

Aplicando la teoría descrita en el capítulo II y en las ecuaciones (3.1) y (3.2), se construyó el algoritmo. En la figura 3.1 se muestra el diagrama de bloques con la metodología utilizada en la presente investigación; en el primer paso, se selecciona la imagen de tamaño $M \times N$, que se desea utilizar como referencia (I^i) a la que se denota con $f(x, y)$; después se le aplica la transformada de Fourier y se obtiene su módulo $|F(w_x, w_y)|$ (paso 2) donde w_x y w_y son las frecuencias angulares. A éste se le aplica un filtro parabólico (paso 3); de esta manera, se atenúan las bajas frecuencias, y las altas se realzan en proporción a $(w_x)^2$, $(w_y)^2$ (Pech-Pacheco *et al.*, 2003); esto resalta los detalles finos del objeto. Como paso 4, el resultado obtenido anteriormente se multiplica por el factor \sqrt{r} , donde r es la frecuencia espacial radial, indispensable para diferenciar la transformada de Escala de la transformada de Mellin.

Después de lo anterior, como paso 5, se convierte el resultado a coordenadas logarítmicas polares, para obtener invariancia a rotación. De la imagen obtenida como paso 6 se selecciona una subimagen que se denota como $M(\lambda, \theta)$ (Lerma *et al.*, 2008), que es el fragmento que contiene la mayor parte de la información de referencia; finalmente se realiza una interpolación bilineal para regresar la imagen a su tamaño original; con esto se evita fuga de información debida al cambio de coordenadas.

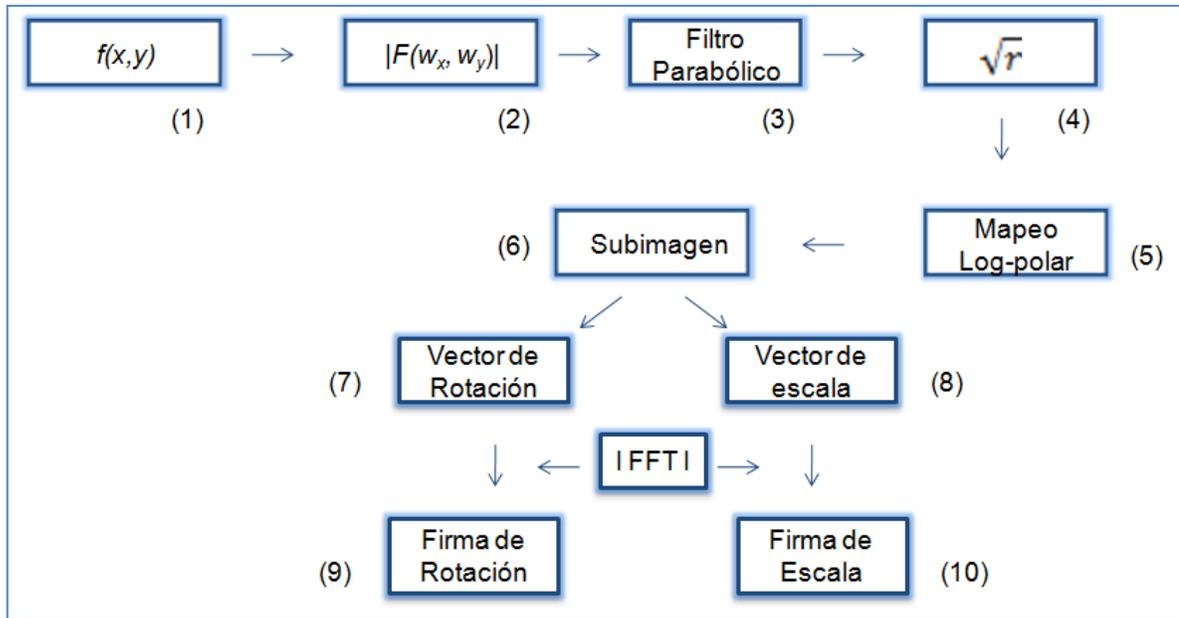


Figura 3.1. Diagrama de bloques del procedimiento utilizado.

De la subimagen de tamaño $M \times N$, se obtienen dos vectores unidimensionales al hacer la proyección de $M(\lambda, \theta)$ en los ejes de coordenadas.

$$M(\lambda_m) = \sum_{k=1}^m M(\lambda_m, \theta_k), \quad \text{vector de rotación.} \quad (3.3)$$

(paso 7)

$$M(\theta_n) = \sum_{k=1}^m M(\lambda_k, \theta_n), \quad \text{vector de escala.} \quad (3.4)$$

(paso 8)

Se calcula el módulo de la transformada de Fourier de los vectores obtenidos, lo que permite definir las firmas vectoriales de rotación y escala mediante las siguientes ecuaciones

$$V_1(w_\lambda) = |F[M(\lambda_n)]|, \quad \text{firma de rotación.} \quad (3.5)$$

(paso 9)

$$V_2(w_\theta) = |F[M(\theta_n)]|, \quad \text{firma de escala.} \quad (3.6)$$

(paso 10)

Para evaluar este algoritmo, se selecciona una imagen de referencia (I^i) y una imagen problema (I^j) ya sea rotada o escalada, a la que se le aplica el procedimiento descrito anteriormente. Para determinar la semejanza entre las imágenes I^i e I^j de la base de datos, se calcula la distancia Euclidiana (d_E) entre sus firmas promedio por medio de la siguiente ecuación

$$d_E = \sqrt{\sum [V_{1f}(w_\lambda) - V_{1g}(w_\lambda)]^2 + [V_{2f}(w_\theta) - V_{2g}(w_\theta)]^2}. \quad (3.7)$$

De esta manera, todas las imágenes j de la base de datos pueden ser comparadas con cualquier imagen i que se desee reconocer.

3.3 Aplicación del algoritmo

Para poder aplicar el algoritmo basado en la teoría descrita anteriormente, se considera en primer término una imagen de referencia; en este caso particular, consideramos una imagen sencilla (un rectángulo) de tamaño de 256 X 256 píxeles y fondo negro, como se muestra en el inciso (a) de la Figura 3.2, a la que se le calcula el módulo de la transformada de Fourier (Espectro de Fourier), el cual se puede observar en (b). Se utilizó el comando `colormap` de Matlab para cambiar a un mapa de colores y con ello visualizar de manera clara cada uno de los procedimientos aplicados.

Como el rango de valores en este espectro es tan grande en comparación con los de la pantalla, dando como resultado que los valores brillantes del centro dominen el resultado,

para mejorar los detalles visuales del espectro de Fourier aplicamos una transformación logarítmica (Gonzalez, 2009); el resultado se muestra en la figura 3.3.

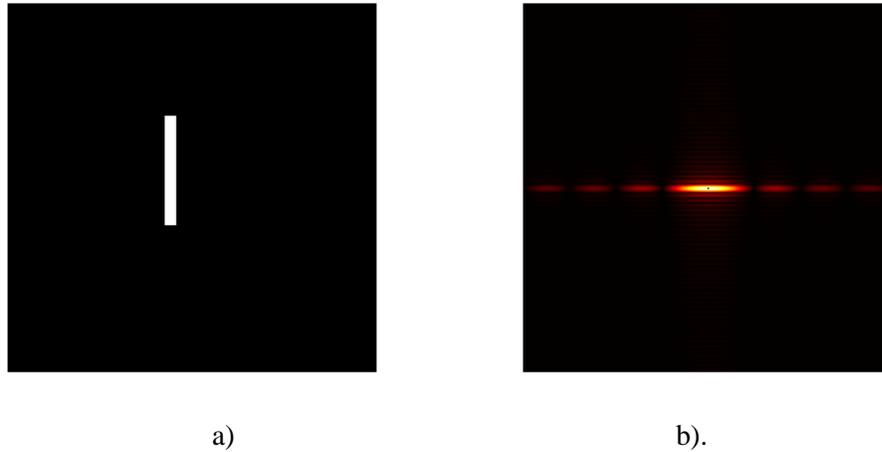


Figura 3.2. (a) Imagen original y (b) su espectro de Fourier.

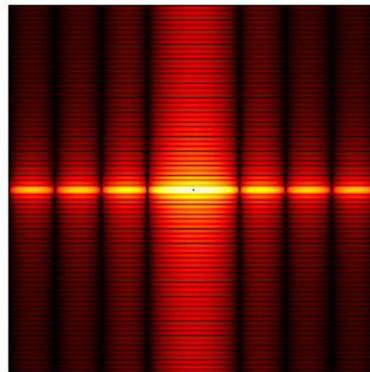


Figura 3.3. Espectro de Fourier mejorado visualmente.

El uso de la transformación logarítmica en el espectro de Fourier de la Fig. 3.2 (b) sólo tiene la finalidad de mejorar la presentación de las imágenes, y en ningún momento está incluido dentro del algoritmo. Al espectro de Fourier se le aplica un filtro parabólico (Figura 3.4); de esta manera se atenúan las bajas frecuencias y las altas frecuencias se

realzan como se muestra en la figura 3.5 (a); en la figura 3.5 (b) se muestra el resultado de multiplicar la figura 3.5 (a) por el factor de escala \sqrt{r} .

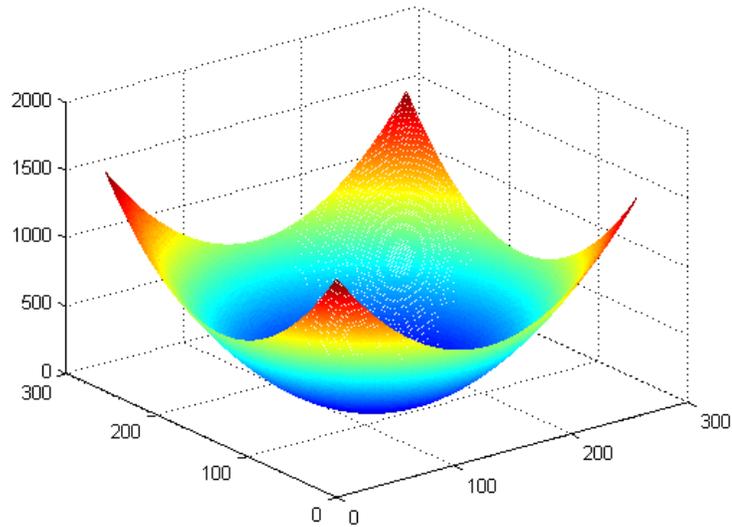
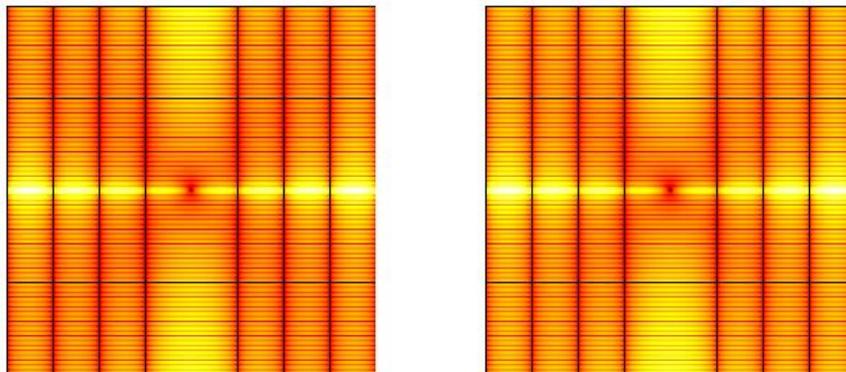


Figura 3.4. Filtro parabólico.



(a)

(b)

Figura 3.5. Resultado de aplicar a) Filtro parabólico al espectro de Fourier.
b) El Factor de escala \sqrt{r} a la Fig. 3.5 (a)

Después de realizar los pasos anteriores, convertimos la imagen de la figura 3.5 (b) a coordenadas logarítmicas polares y elegimos una subimagen localizada en el intervalo [64, 210], correspondiente a la zona que contiene la mayor cantidad de información. En las figuras 3.6 (a) y (b) se observan los resultados.

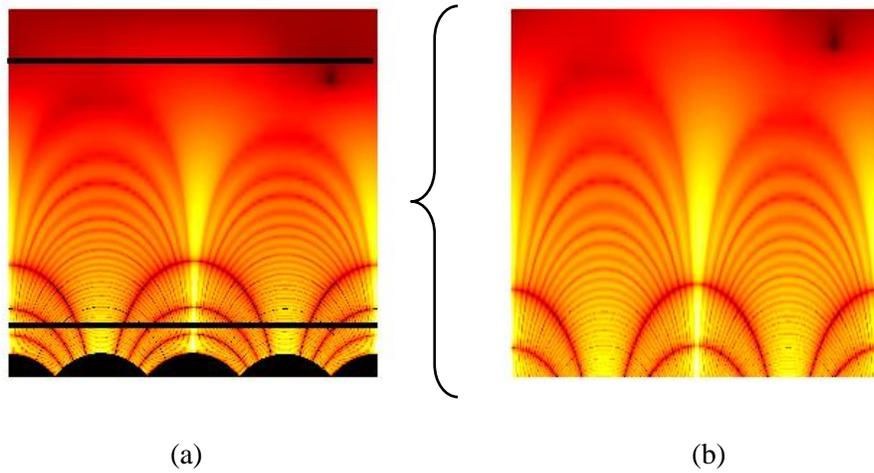


Figura 3.6. a) Imagen 3.5 (b) en coordenadas log-polar. b) Subimagen seleccionada.

A la subimagen $M(\lambda, \theta)$ que se observa en la Figura 3.6 (b) le calculamos sus funciones marginales utilizando las siguientes ecuaciones

$$g(x) = \sum_{\theta} M(\lambda, \theta). \quad (3.8)$$

$$h(x) = \sum_{\lambda} M(\lambda, \theta). \quad (3.9)$$

En la figura 3.7(a) se muestra la subimagen, vista en tres dimensiones, y en 3.6 (b) y 3.6 (c) se observan las funciones marginales, con lo cual obtenemos los vectores de escala y rotación que caracterizan a esta imagen.

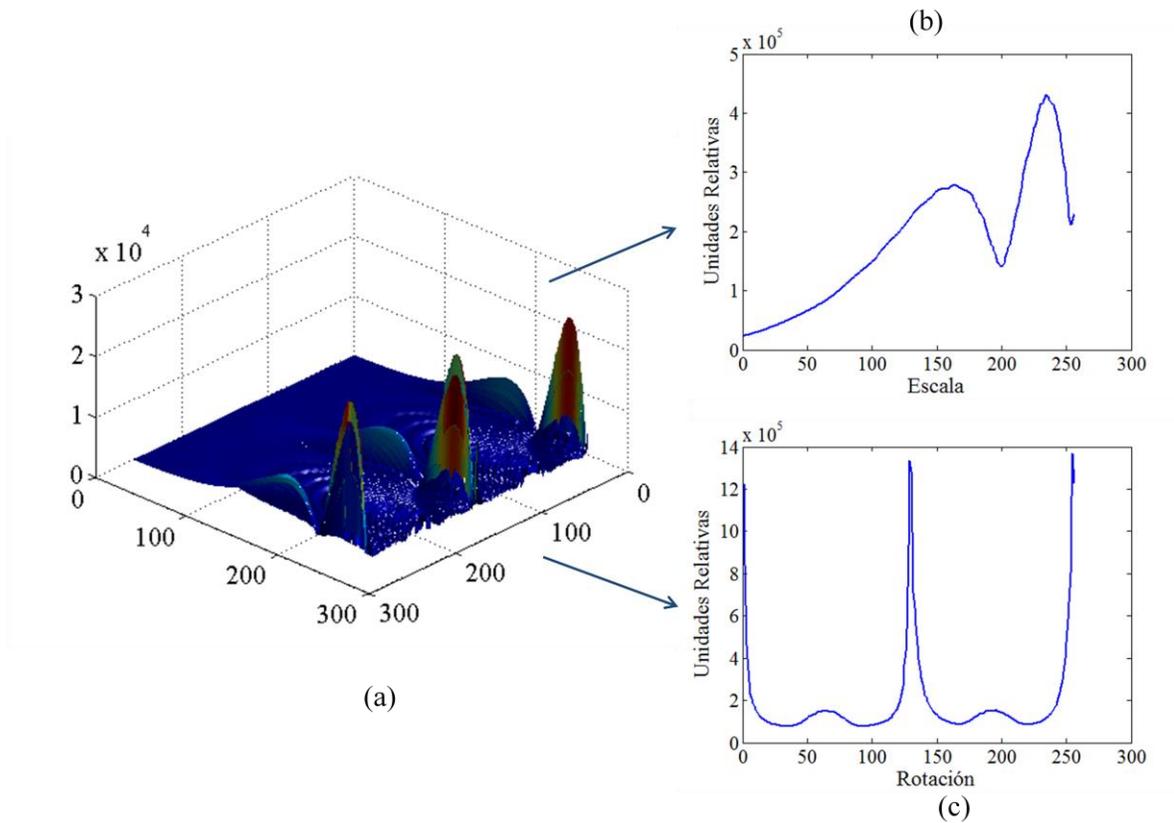


Figura 3.7. (a) Imagen tridimensional de Figura 3.6 (b). (b) Vector de la sumatoria sobre el eje rotación. (c) Vector de la sumatoria sobre el eje escala.

Finalmente, en la Figura 3.8, se presentan los módulos de la transformada de Fourier de los vectores mostrados en las Figuras 3.7 (b) y 3.7 (c) respectivamente, lo que define las firmas vectoriales de la imagen, y son las que se utilizan para hacer la comparación con respecto a otra imagen.

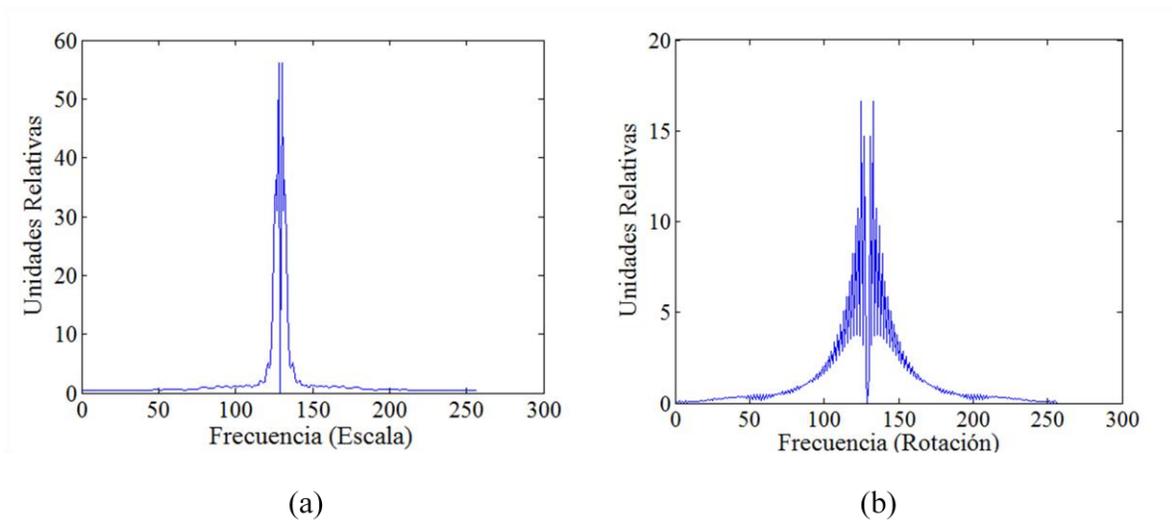


Figura 3.8. Firmas vectoriales. (a) Firma de escala. (b) Firma de rotación.

CAPÍTULO IV

RECONOCIMIENTO DE IMÁGENES EN ESCALA DE GRISES

4.1 Análisis de letras.

4.1.1. Resultados para la rotación.

4.1.2. Resultados para cambios de escala.

4.2 Análisis de números.

4.3 Análisis de formas de hojas.

4.4 Copépodos.

4.4.1 Simulaciones utilizando imágenes de los copépodos.

4.4.2. Firmas promedio.

4.4.3. Experimento de identificación de especies.

4.5 Imágenes de color transformadas a escala de grises.

4.1 Análisis de letras

En este capítulo se tomaron como primer objeto de análisis las imágenes de las diferentes letras que forman el alfabeto en idioma inglés. En primer término se eligió la imagen de la letra E en tipo Arial tamaño 72, escrita en color blanco con fondo negro para utilizarla como referencia; a ésta se le aplicó el procedimiento descrito anteriormente en el capítulo III.

4.1.1. Resultados para la rotación

Lo primero que se analizó fue el comportamiento que tiene la distancia Euclidiana de la imagen de referencia (E) con respecto a sí misma para rotaciones de 0 a 360° con variaciones de un grado.

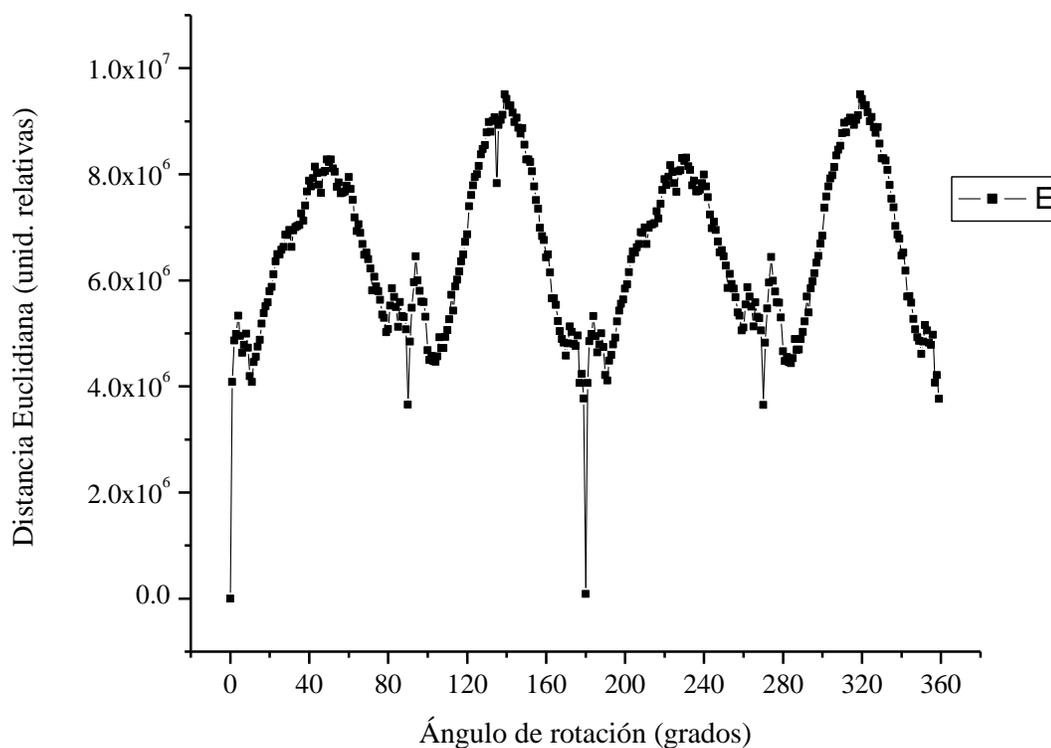


Figura 4.1. Análisis de comportamiento de la d_E de una imagen con respecto a la rotación.

Dada la simetría de la imagen de referencia, podemos observar que el comportamiento de la curva es periódico, es decir, se repite cada 180 grados.

Para evaluar la capacidad de poder distinguirla, la comparamos con las 26 letras del alfabeto, rotadas cada una 360° , con variación de un grado. Se realizó la simulación para determinar la diferencia de la d_E entre ellas, utilizando las 9360 imágenes (26×360), como se puede observar en la Figura 4.2; el resultado muestra una separación muy clara entre los valores de la distancia euclidiana de la letra E comparada con el resto de las letras, lo que permite poder identificarla.

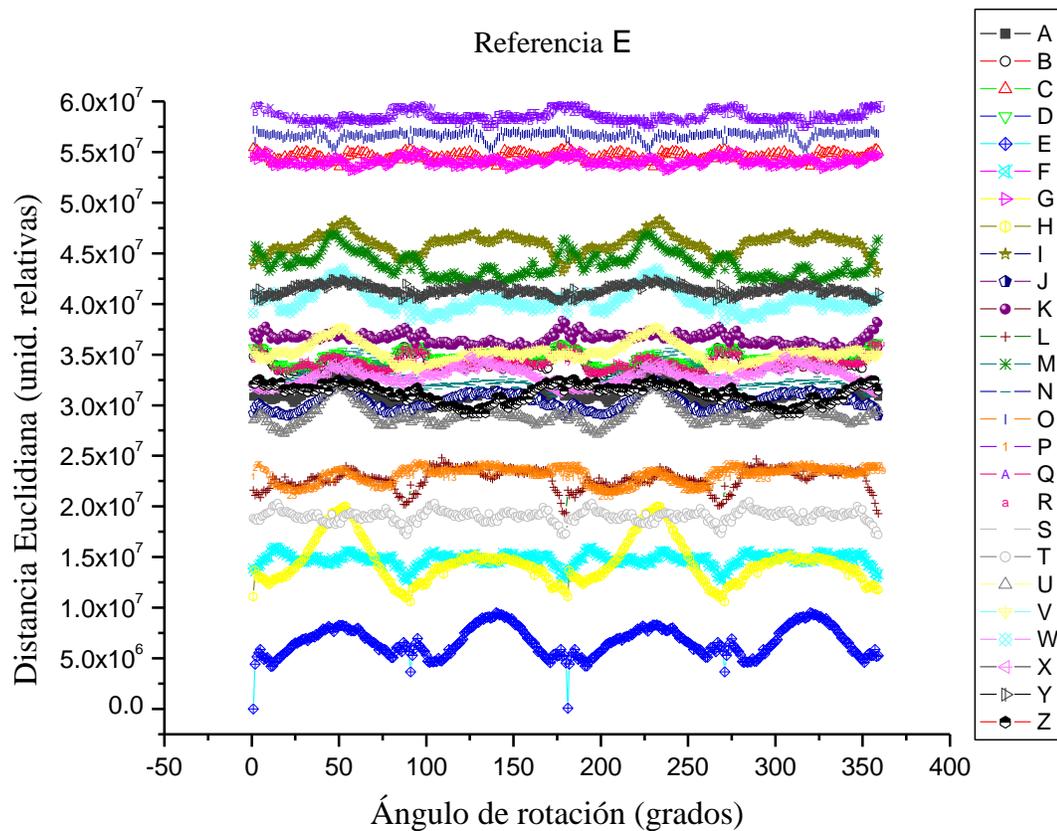


Figura 4.2. Análisis de comportamiento de la d_E de todas las letras con respecto a la rotación.

Si ahora consideramos sólo las letras que tienen las distancias euclidianas más cercanas a la letra de referencia E, obtenemos la Figura 4.3, en donde se puede observar que son las letras F, H, L, P, y T, lo cual era de esperar dado el parecido geométrico que se tiene con respecto a la letra E.

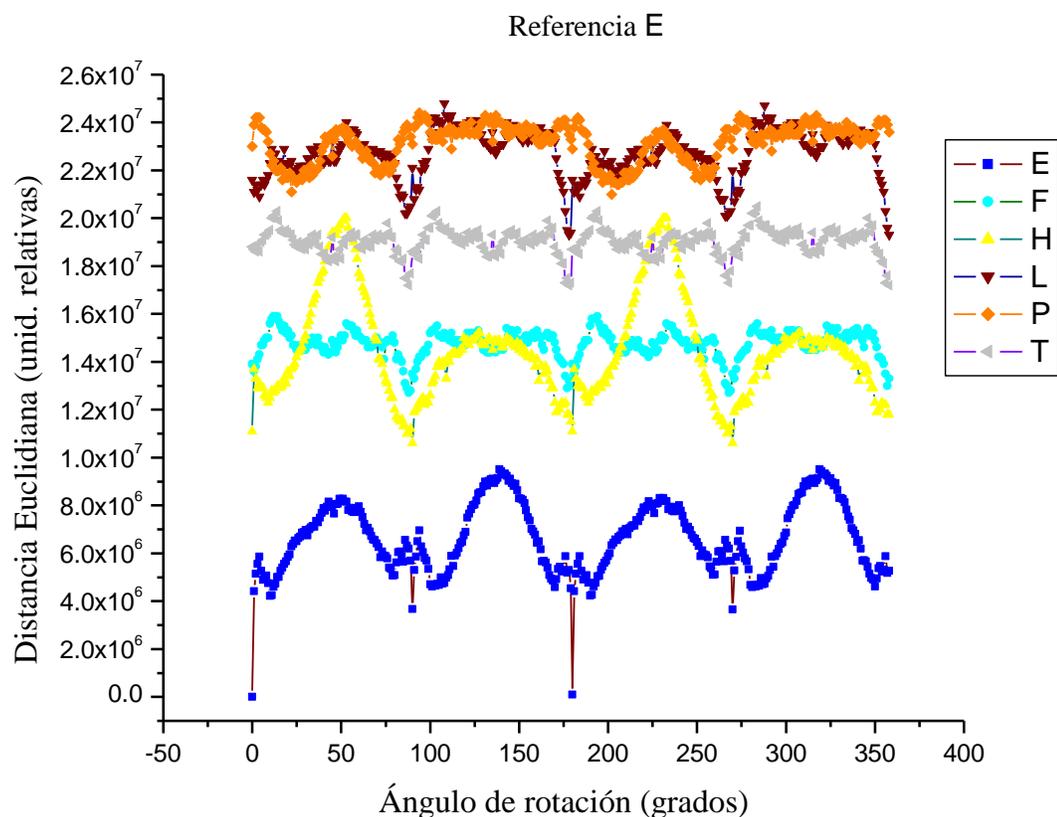


Figura 4.3. Análisis de comportamiento de la d_E .

Es importante hacer al análisis estadístico de todos los datos para determinar su nivel de confiabilidad. En la Figura 4.4 se observan los valores promedios de la distancia, obtenidos para cada una de las letras, así como las barras de error a una y dos veces el error estándar (EE). Esto nos permite visualizar los resultados a un 95.4 % de confiabilidad. Como se puede ver en la figura 4.4, las letras, F, H, L, P, y T, estadísticamente hablando, son las más parecidas a la letra E.



Figura 4.4. Comportamiento estadístico de la d_E .

De las Figuras 4.2 a 4.4 se analizó el caso particular de la letra E, utilizada como referencia. Para probar la metodología, se efectuó el análisis completo para cada una de las letras del alfabeto, obteniéndose el nivel de confianza de cada una de ellas, definido como el cociente del número de puntos traslapados de la letra de referencia y la letra de prueba, entre el número total de imágenes analizadas. En el caso de rotación, para cada una de las 26 letras se realizaron rotaciones con variaciones de un grado entre cada una, hasta completar los 360 grados, analizando en total 243,360 imágenes (26 x 26 x 360). Los resultados obtenidos se muestran de manera porcentual en las tablas (4.1) y (4.2).

Tabla 4.1. Nivel de confianza porcentual para las letras A-M.

		Letras evaluadas													
Letras de Referencia		A	B	C	D	E	F	G	H	I	J	K	L	M	
	A	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	B	100	100	100	77.78	100	100	100	100	100	100	100	100	100	100
	C	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	D	100	81.11	100	100	100	100	100	100	100	100	100	100	100	100
	E	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	F	100	100	100	100	100	100	100	100	100	100	100	97.78	100	
	G	100	100	100	100	100	100	100	100	100	100	100	100	100	
	H	100	100	100	100	57.78	78.89	100	100	100	100	100	86.67	100	
	I	100	100	100	100	100	100	100	100	100	100	100	100	100	
	J	100	100	100	100	100	100	100	100	100	100	100	100	100	
	K	100	100	100	100	100	100	100	100	100	100	100	100	100	
	L	100	100	100	100	100	81.67	100	100	100	100	100	100	100	
	M	100	100	100	100	100	100	100	100	100	100	100	100	100	
	N	100	100	100	100	100	100	100	100	100	100	100	100	100	
O	100	100	100	100	100	100	100	100	100	100	100	100	100		
P	100	100	100	100	100	100	100	100	100	100	100	100	100		
Q	100	100	100	100	100	100	100	100	100	100	100	100	100		
R	100	100	100	89.17	100	100	100	100	100	100	100	100	100		
S	100	100	100	100	100	100	100	100	100	100	100	100	100		
T	100	100	100	100	100	49.72	100	100	100	100	100	33.33	100		
U	100	100	100	100	100	100	100	100	100	100	100	100	100		
V	86.94	100	100	100	100	100	100	100	100	100	100	100	100		
W	100	100	100	100	100	100	100	100	100	100	100	100	100		
X	100	100	100	100	100	100	100	100	100	100	100	100	100		
Y	100	100	100	100	100	100	100	100	100	100	100	100	100		
Z	100	100	100	100	100	100	100	100	100	100	100	100	100		

En la tabla 4.1 se muestran los resultados obtenidos para las letras de referencia (A-Z) al compararlas con las letras evaluadas (A-M). Por ejemplo, si seleccionamos el renglón que corresponde a la letra E cuando se utiliza como referencia, podemos observar que, al compararla con todas las letras evaluadas, podemos distinguirla al 100 %. Si tomamos como referencia a la letra B tenemos problemas al compararla sólo con la letra D, obteniendo un nivel de confianza del 77.78 %, debido a que algunos puntos se encuentran traslapados; si ahora elegimos a la letra D como referencia, podemos ver que ésta tiene problemas de identificación únicamente con la letra B con un nivel de confianza al 81.11 %, esto nos indica que cuando se tiene a la letra D como referencia, el nivel de confianza es mayor que cuando se tiene como referencia a la B.

Tabla 4.2. Nivel de confianza porcentual para las letras N-Z.

		Letras evaluadas													
Letras de Referencia		N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
	A	100	100	100	100	100	100	100	100	100	85.56	100	100	100	100
	B	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	C	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	D	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	E	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	F	100	100	100	100	100	100	50	100	100	100	100	100	100	100
	G	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	H	100	100	100	100	100	100	85.56	100	100	100	100	100	100	100
	I	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	J	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	K	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	L	100	100	100	100	100	100	100	65.83	100	100	100	100	100	100
	M	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	N	100	100	100	100	100	100	100	100	100	80.56	100	100	100	100
	O	100	100	100	100	100	100	100	100	100	100	100	100	100	100
P	100	100	100	100	100	100	100	100	100	100	100	100	100	100	
Q	100	100	100	100	100	100	100	100	100	100	100	100	100	100	
R	100	100	100	100	100	100	100	100	100	100	100	100	100	100	
S	100	100	100	100	100	100	100	100	100	100	100	100	100	100	
T	100	100	100	100	100	100	100	100	100	100	100	100	100	100	
U	100	100	95.56	100	100	100	100	100	100	100	100	100	100	100	
V	100	100	100	100	100	100	100	100	100	100	100	100	100	100	
W	93.33	100	100	100	100	100	100	100	100	100	100	100	100	100	
X	100	100	100	100	100	100	100	100	100	100	100	100	100	100	
Y	100	100	100	100	100	100	100	100	100	100	100	100	100	100	
Z	100	100	100	100	100	100	100	100	100	100	100	100	100	100	

En la tabla 4.2 se muestran los resultados obtenidos para el resto de las letras (comprendidas entre N y Z). De nuevo en la mayoría de ellas tenemos un nivel de confianza del 100 %, con algunos problemas de traslapes de puntos en algunas letras, siendo el caso más crítico el de F comparada con T donde, dado su parecido geométrico, el nivel de confianza obtenido es del 50 %.

Se observa que no existe simetría en las tablas 4.1 y 4.2; por ejemplo, si se toma la letra E como referencia, se puede identificar al cien por ciento, sin ningún problema, al compararla con todas las letras; en cambio, si se toma la letra H, ésta tiene traslapes con las letras E, F, L y T debido a su parecido geométrico.

4.1.2. Resultados para cambio de escala

Para analizar el comportamiento ante los cambios de escala, se realizaron pruebas haciendo modificaciones del 80 % al 120 % del tamaño de la imagen de referencia, con cambios del 0.5 %. En la Figura 4.5 se muestran los resultados obtenidos tomando como referencia la letra E y, comparada de nuevo con todas las letras del alfabeto, puede observarse que para cambios de $\pm 10\%$ sólo hay problemas con la letra H, pero al incrementar este margen aparecen traslapes con otras letras.

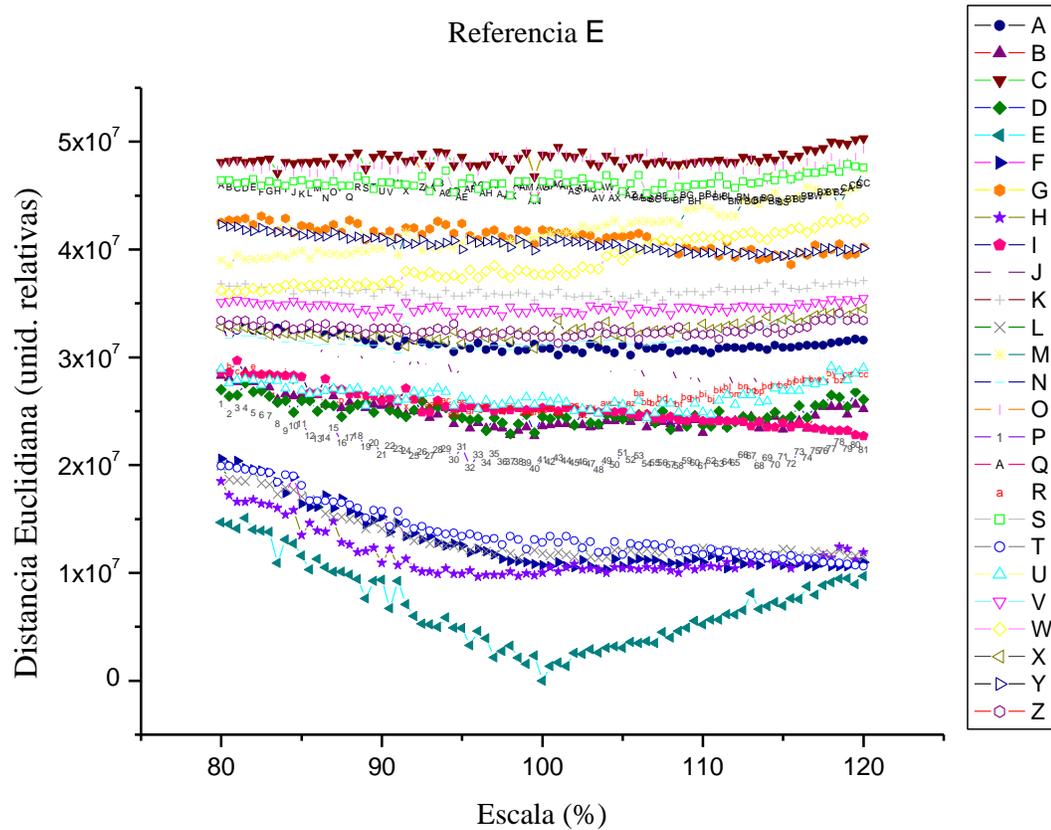


Figura 4.5. Comportamiento de la distancia ante cambios de escala de 80 % a 120 %.

Tabla 4.3. Nivel de confianza porcentual para cambios de escala.

	Escala 90-110 %	Escala 80-120 %		Escala 90-110 %	Escala 80-120 %
A	100	100	N	100	100
B	100	100	O	100	100
C	100	100	P	100	100
D	100	100	Q	100	100
E	100	100	R	100	100
F	100	70.73	S	100	100
G	100	100	T	100	82.93
H	90.48	63.41	U	100	100
I	100	100	V	100	100
J	100	100	W	100	100
K	100	100	X	100	100
L	100	90.24	Y	100	100
M	100	100	Z	100	100

En la Tabla 4.3 se muestran los resultados de los niveles de confianza porcentuales obtenidos de las simulaciones al tomar como referencia la letra E y compararla con todas las letras del abecedario, haciendo en primer término modificaciones en la escala de $\pm 10\%$ observando que solo se tiene problemas con la letra H y posteriormente cambios de $\pm 20\%$, donde ahora tenemos traslapes con las letras F, H, L y T que de nuevo son las que geoméricamente tiene mayor parecido.

4.2 Análisis de números

Se consideró un segundo grupo de imágenes consistente en los dígitos del sistema decimal (0 – 9) escritos en color blanco con fondo negro, utilizando tipo de letra Arial, tamaño 144; de nuevo todas las imágenes son de 256 x 256 píxeles.

Se seleccionó la imagen del número 1 para utilizarla como referencia, y se le aplicó el procedimiento descrito en la sección 3.2. Realizando la simulación numérica para determinar la d_E y evaluar la capacidad del sistema para distinguirlo, al compararlo con todos los dígitos del sistema decimal (0 – 9), rotados cada uno 360 grados – como se puede observar en la Figura 4.6 – el resultado muestra una separación muy clara entre los valores

de la distancia de la imagen del 1 con el resto de los números, lo que nos permite poder identificarla. Es importante hacer el análisis estadístico de todos los datos para determinar su nivel de confiabilidad.

En las Figuras 4.7 se observan respectivamente los valores promedio de la distancia de cada número, así como las barras de error a una y dos veces el error estándar. En el caso de cambios en el ángulo de rotación se puede observar cómo no existe ningún traslape y la distancia euclidiana promedio del número de referencia es diferente a la del resto de los números.

Para el caso de cambios en la escala de 80 % a 120 %, se puede observar en la figura 4.8 que la distancia euclidiana promedio sigue manteniéndose inferior, permitiéndonos asegurar que es posible distinguirlo con un nivel de confiabilidad del 95.4 %.

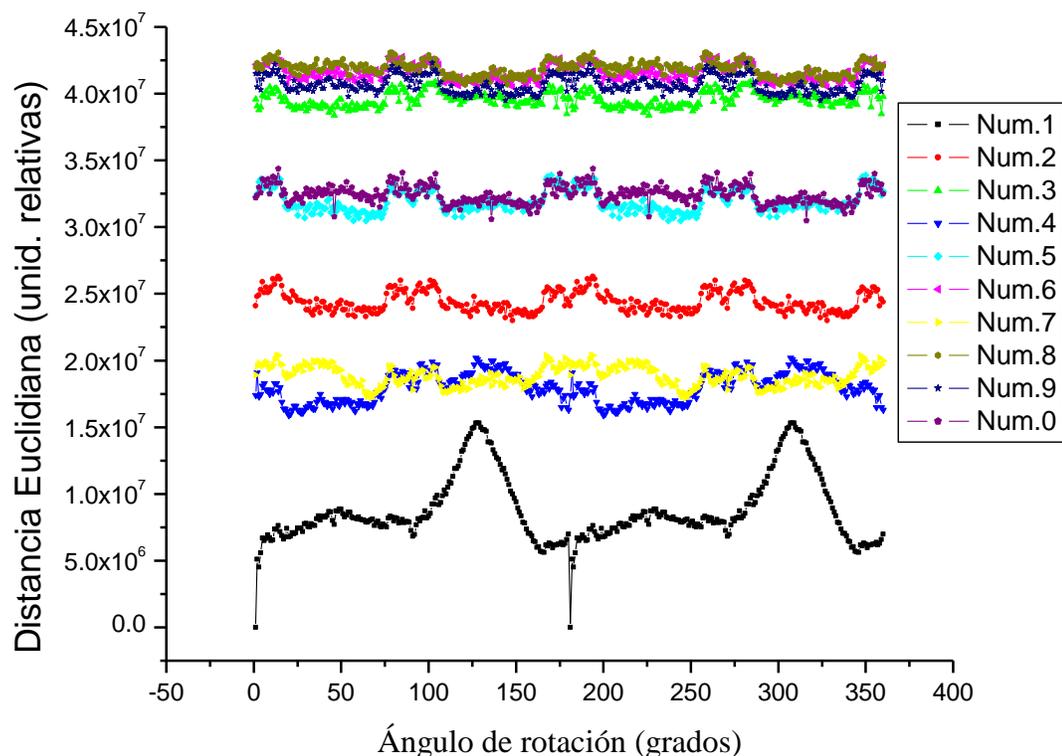


Figura 4.6. Cambios en la distancia versus ángulo de rotación.

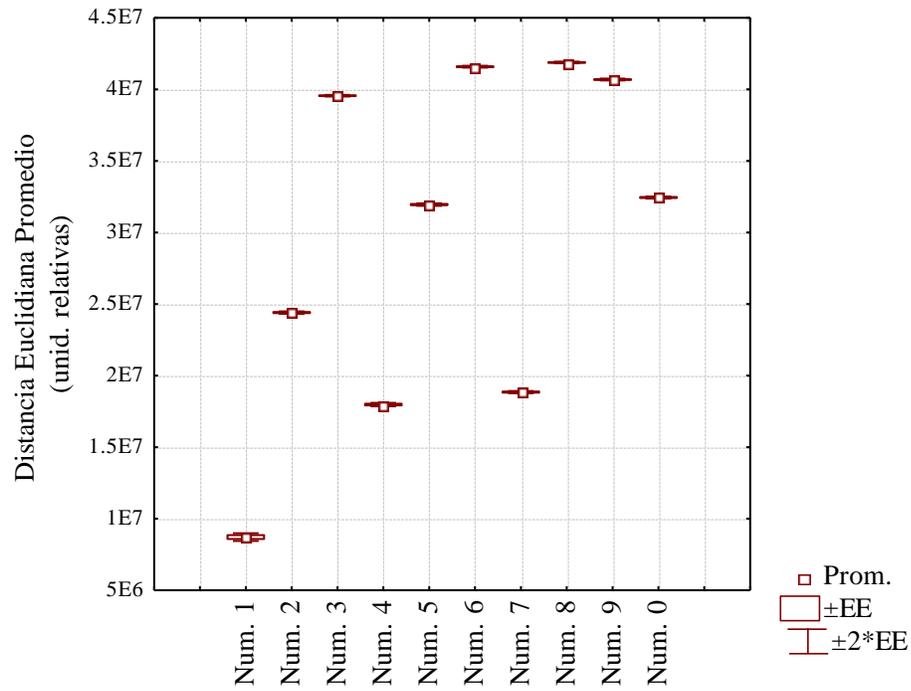


Figura 4.7. Comportamiento estadístico de la distancia para rotación.

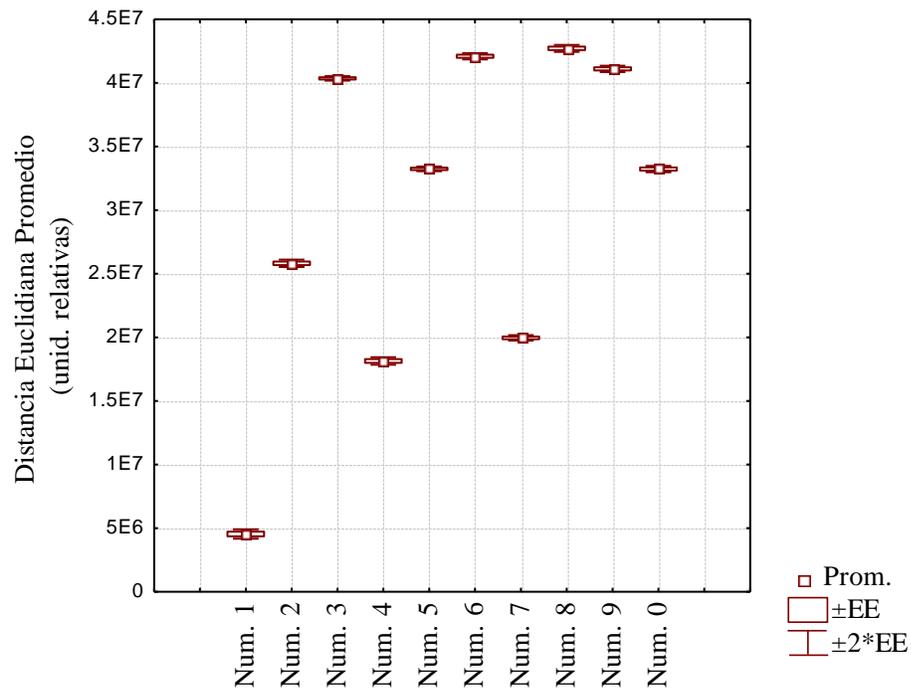


Figura 4.8. Comportamiento estadístico para cambios

de escala de 80 % a 120 %.

4.3 Análisis de formas de hojas

En tercer término se analizaron las imágenes de nueve diferentes formas de hojas de planta (Figura 4.9), tomando la forma ovada (forma 1) como referencia. En la Figura 4.10 se muestra el comportamiento de la d_E para las imágenes rotadas 360 grados con variación de un grado; aquí se aprecia que existe una separación de la forma de referencia elegida con respecto al resto de las hojas, lo que permite identificarla sin dificultad, independientemente de cualquier cambio en la rotación de la imagen. Para el caso de cambios de escala, de igual forma que con los números, se analizaron variaciones del 80 % hasta el 120 % con cambios de 0.5 % (los resultados se muestran en Figura 4.11). Considerando de nuevo la forma 1 como referencia, puede observarse la separación de la d_E de la imagen de referencia con respecto al resto, lo que nos permite poder identificarla.

		
a) ovada	b) oblonga	c) orbicular
		
d) obovada	e) elíptica	f) lanceolada
		
g) triangular	h) linear	i) espatulada

Figura 4.9. Patrones utilizados con diferentes formas de las hojas.

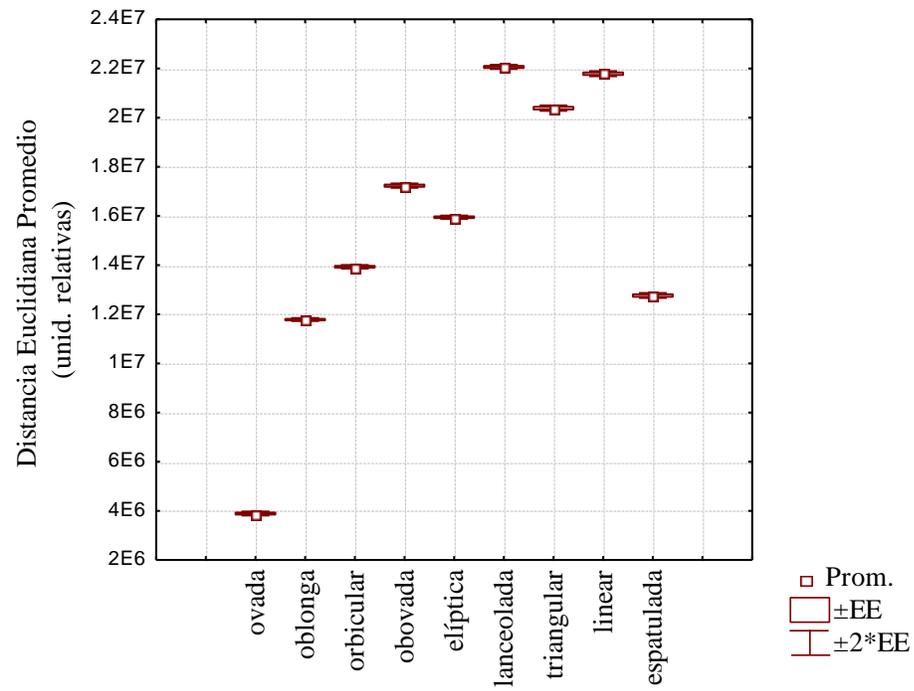


Figura 4.10. Comportamiento estadístico de la d_E para rotación

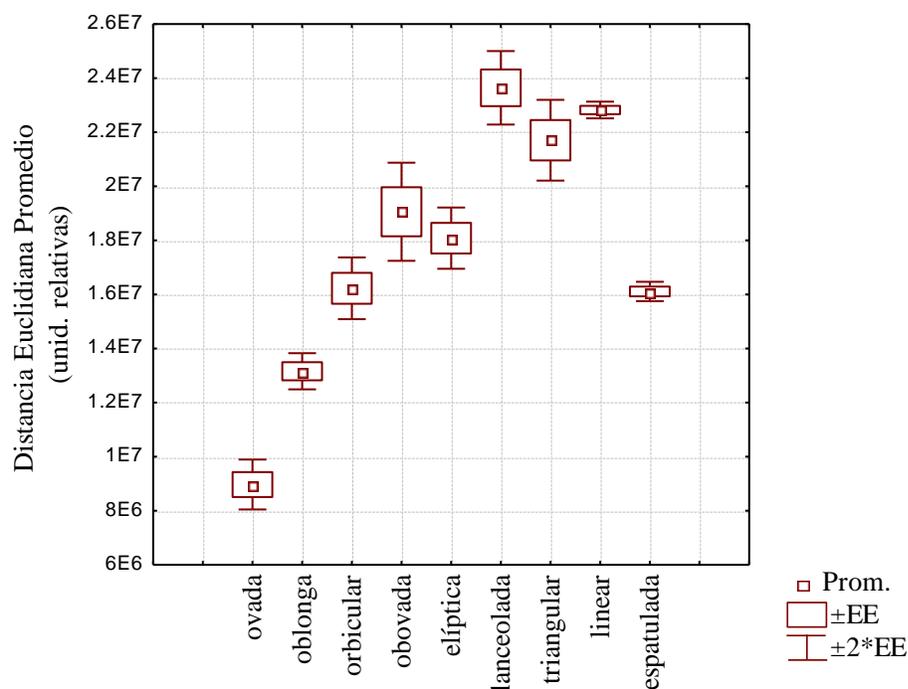


Figura 4.11. Comportamiento estadístico de la d_E para cambios de escala de 80 % a 120 %.

4.4 Copépodos

Los copépodos son el grupo dominante en el zooplancton marino, constituyen por lo menos el 70 % de la fauna planctónica. Este grupo de organismos tiene una gran diversidad. Hay alrededor de 11,500 especies de copépodos descritas por Humes (1994) y el número sigue en constante aumento con la descripción de nuevas especies, como los ejemplares de las cuevas *anchialine* (Fosshagen, 1991), las fuentes hidrotermales (Humes, 1991), y los publicados bajo el nombre de otras especies (Bradford, 1976, Soh, 2000).

En los ecosistemas marinos los copépodos son de primordial importancia. La mayoría se alimentan de fitoplancton, formando un vínculo directo entre la producción primaria y los peces de importancia comercial, como la sardina y el arenque. Son también la principal fuente de alimento para una gran variedad de invertebrados. Estudios sobre la abundancia y composición de especies de copépodos son especialmente importantes, porque la mayoría de las larvas de los peces comerciales se alimentan de ellos, por lo tanto, los cambios en la abundancia de estos copépodos de año en año, puede determinar las

fluctuaciones interanuales de las poblaciones de peces explotadas comercialmente en una región en particular.

El hecho de que algunas especies se puedan confundir con otras, nos muestra que la correcta identificación de un organismo es de suma importancia, pero ésta puede tardar mucho tiempo, por lo que el desarrollo de nuevas técnicas de identificación de especies puede ser muy útil para el taxónomo (Zavala-Hamz y Álvarez-Borrego, 1997, Castro-Longoria *et al.*, 2001, Álvarez-Borrego y Castro-Longoria, 2003, Guerrero-Moreno y Álvarez-Borrego, 2009, Solorza y Álvarez-Borrego, 2010, Álvarez-Borrego y Solorza, 2010).

4.4.1 Simulaciones utilizando imágenes de los copépodos

Para este estudio se eligieron ejemplares de diferentes especies de copépodos en etapa adulta, los cuales fueron separados de varias muestras de plancton. Las muestras se observaron con un microscopio óptico y las imágenes fueron capturadas digitalmente utilizando una cámara CCD. Para probar el algoritmo propuesto se utilizó un conjunto de imágenes de siete diferentes especies de copépodos, éstas se muestran en la Figura 4.1.2.

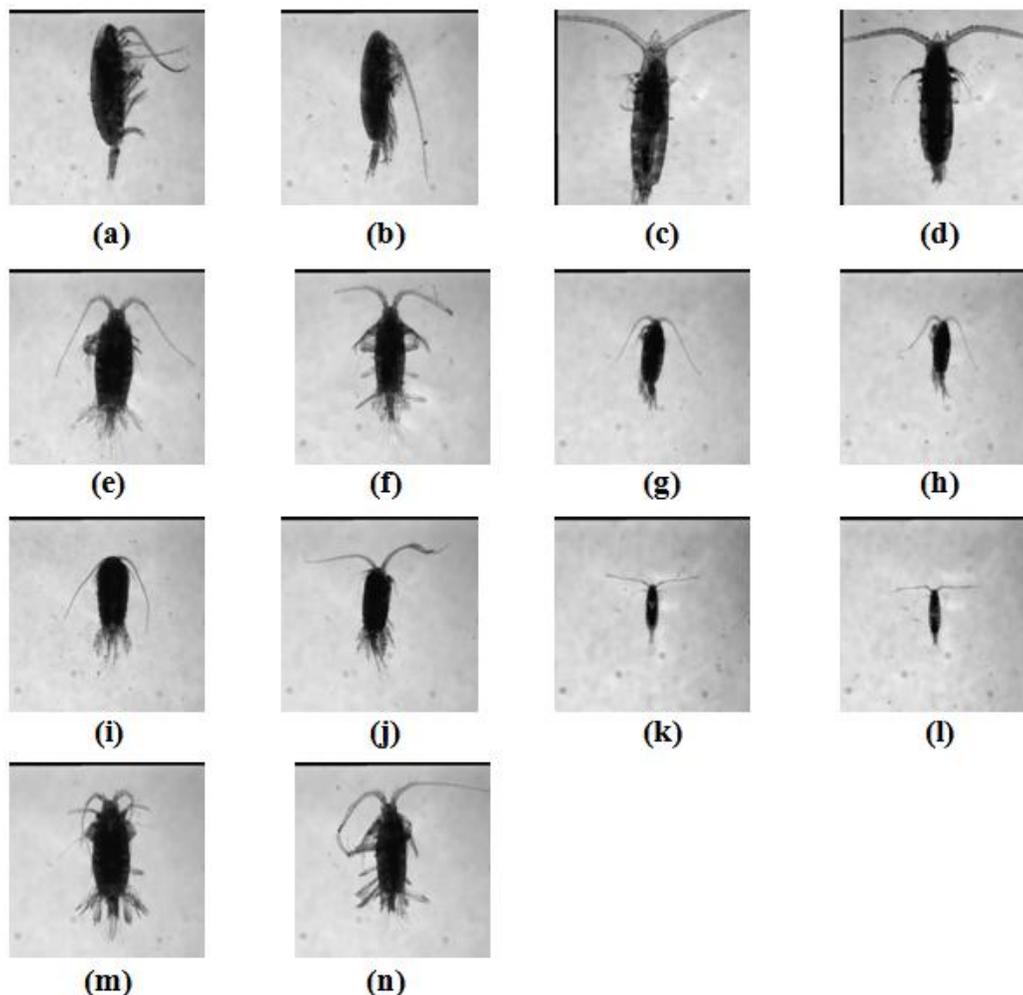


Figura 4.12. Especies de copépodos utilizadas. a) *Calanus pacificus* hembra. b) *C. pacificus* macho. c) *Rhincalanus nasutus* hembra. d) *R. nasutus* macho. e) *Centropages furcatus* hembra. f) *C. furcatus* macho. g) *Pleuromamma gracilis* hembra. h) *P. gracilis* macho. i) *Temora discaudata* hembra. j) *T. discaudata* macho. k) *Acartia tonsa* hembra. l) *A. tonsa* macho. m) *Centropages hamatus* hembra. n) *C. hamatus* macho.

El ejemplar *Calanus pacificus* hembra mostrado en la Figura 4.12(a) fue utilizado como referencia para evaluar el método y valorar si podemos discriminar su especie y sexo al compararla con la base de datos de todos estos organismos. El tamaño de las imágenes es de 256 x 256 píxeles con fondo negro. La imagen de entrada fue rotada 360 grados en incrementos de 1°, las firmas fueron comparadas usando la distancia euclidiana (d_E) entre las imágenes de referencia y problema; como la variación en el tamaño de los copépodos adultos es mínima, no se analizaron los cambios de escala.

Se analizó el comportamiento de la d_E para la imagen de referencia, al compararla consigo misma, para lo cual hice modificaciones en la imagen al rotarla de 0 a 359 grados; la curva que representa la d_E es cíclica y con simetría de 180 ° (Figura 4.13).

Es necesario considerar que cuando se gira una imagen esta presenta distorsiones geométricas, debido a que están formadas como una matriz rectangular de píxeles (éstos aparecen como pequeños cuadrados o rectángulos en matices de gris).

Con el fin de ver si esta metodología tiene un buen rendimiento para reconocer la imagen objetivo cuando se compara con otras imágenes que no corresponden al copépodo elegido como referencia, se realizaron cálculos numéricos para todo el conjunto de las 14 imágenes mostradas en la Figura 4.12. Para las diferentes especies de copépodos utilizadas hay un ejemplar hembra y macho respectivamente; la imagen de referencia se comparó con todas las imágenes. Cada una fue preprocesada para limpiarla, de tal forma que se elimina la mayor parte posible del fondo, mientras se conserva el detalle del objeto que deseamos analizar, con el fin de mejorar la información de los copépodos (Guerrero-Moreno and Álvarez-Borrego, 2009).

En la Figura 4.14 se muestra un ejemplo de este procedimiento para un espécimen *Rhincalanus Nasutus* hembra (Figura 4.12(c)), donde podemos observar cómo la imagen original se limpia, e invertimos los colores para resaltar la información.

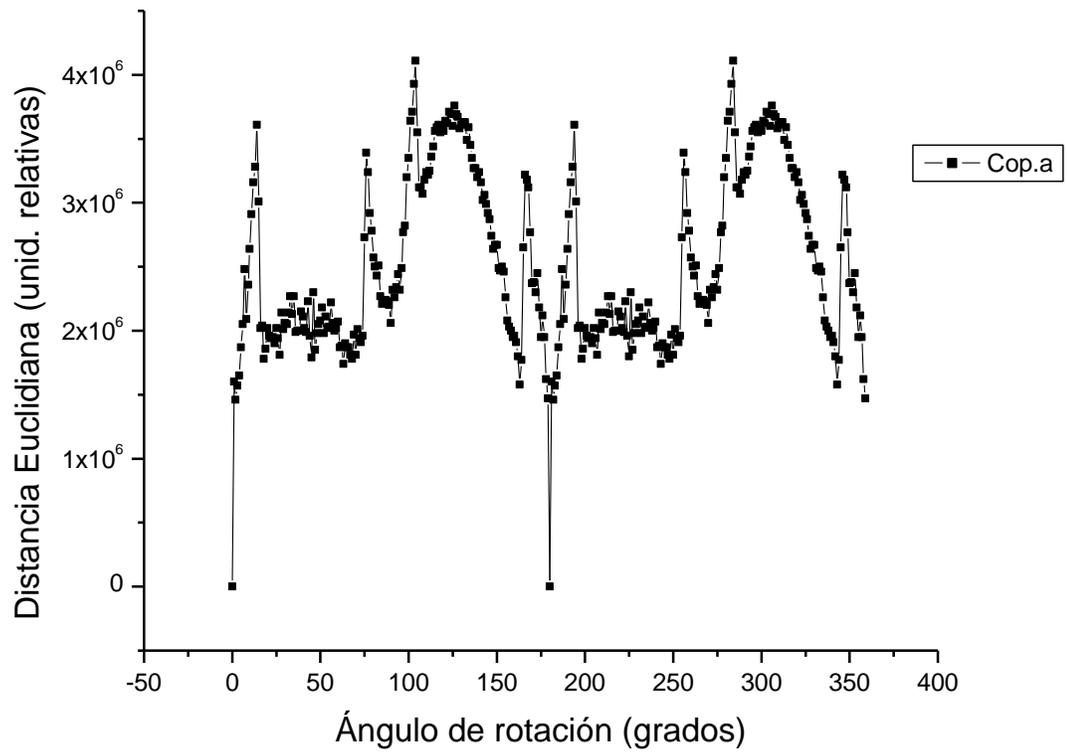
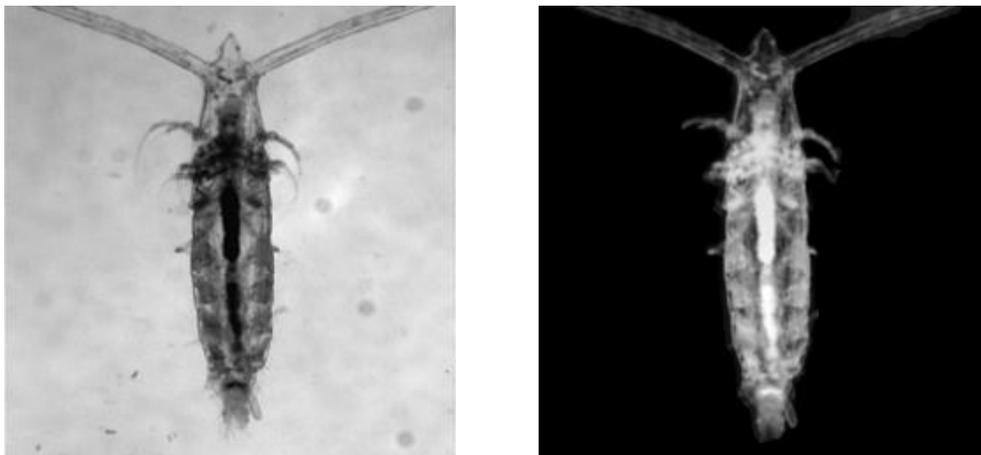


Figura 4.13. Comportamiento de la d_E para un ejemplar *Calanus pacificus* hembra.



(a) Imagen original

(b) imagen procesada

Figura 4.14. Resultado del preprocesado de una imagen.

La Figura 4.15 muestra el comportamiento de la d_E para las catorce imágenes giradas los 360 grados, con variaciones de un grado, considerando como referencia un ejemplar *Calanus pacificus* hembra (Fig. 4.12(a)).

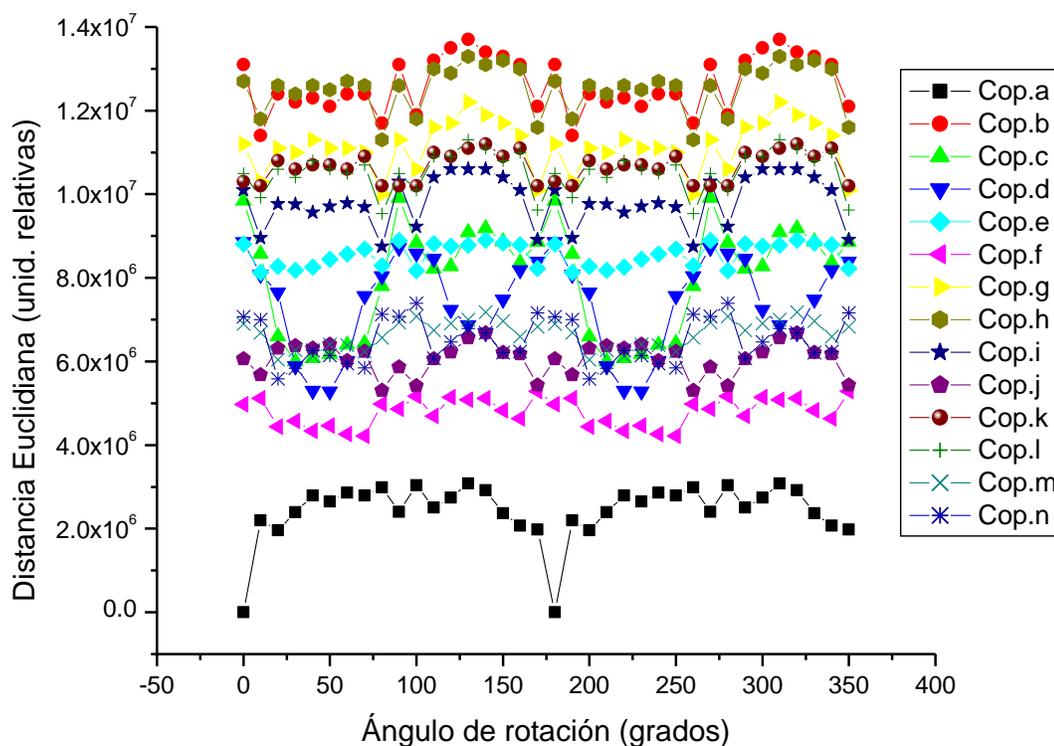


Figura 4.15. Cambios del ángulo de rotación contra d_E .

Se puede distinguir una buena separación de la d_E para la imagen de referencia (copépodo a) con respecto a las distancias de las imágenes que no corresponden a ella. Como puede verse en la Figura 4.15, la curva del copépodo (Figura 4.13(a)) representa el valor mínimo de la d_E , lo que permite distinguirla de las curvas que les corresponden a los demás copépodos. Haciendo un análisis estadístico de todos los datos obtenidos y considerando de nuevo el valor promedio $\pm 2EE$ (dos errores estándar) para cada variable, se puede observar en la Figura 4.16 que este algoritmo tiene al menos el 95.4 % de nivel de confianza para este caso.

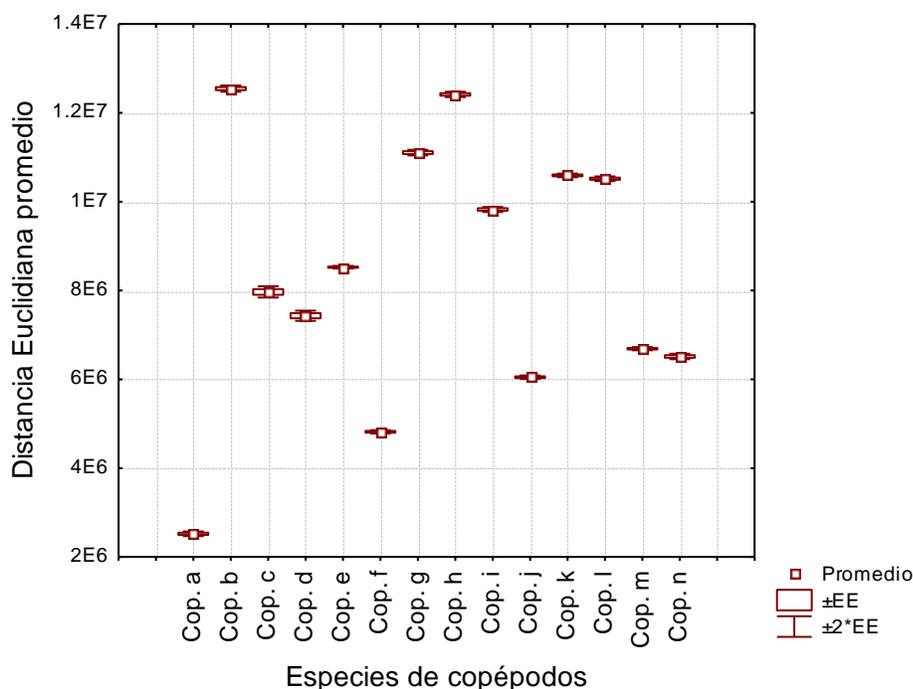


Figura 4.16. Comportamiento estadístico para los diferentes copéodos de la Figura 4.12, usando como imagen de referencia al del inciso (a).

4.4.2. Firmas promedio.

Como segundo conjunto de prueba para los copéodos se eligieron como imágenes de prueba 30 ejemplares hembras y 30 machos de cada una de las 7 especies de copéodos mostrados en la Figura 4.12, formando así una base de datos de 420 imágenes.

Cuando se trabaja con imágenes reales es lógico pensar que, al momento de tomar las fotografías, cada ejemplar presenta una visualización diferente al otro aunque pertenezcan a la misma especie y género; además cada uno puede tener una posición diferente, lo cual nos complica el análisis. Por tal motivo se decidió tomar varias imágenes como referencia, proponiendo el uso de un sistema de identificación que utiliza las características de varios copéodos. Con las firmas vectoriales de cada uno de ellos se obtienen las firmas vectoriales promedio, las cuales son el valor de la sumatoria de las firmas de n copéodos diferentes divididas entre n , obteniéndose con ello su promedio, que es ahora la firma vectorial de referencia. El procedimiento se muestra en la Figura 4.17.

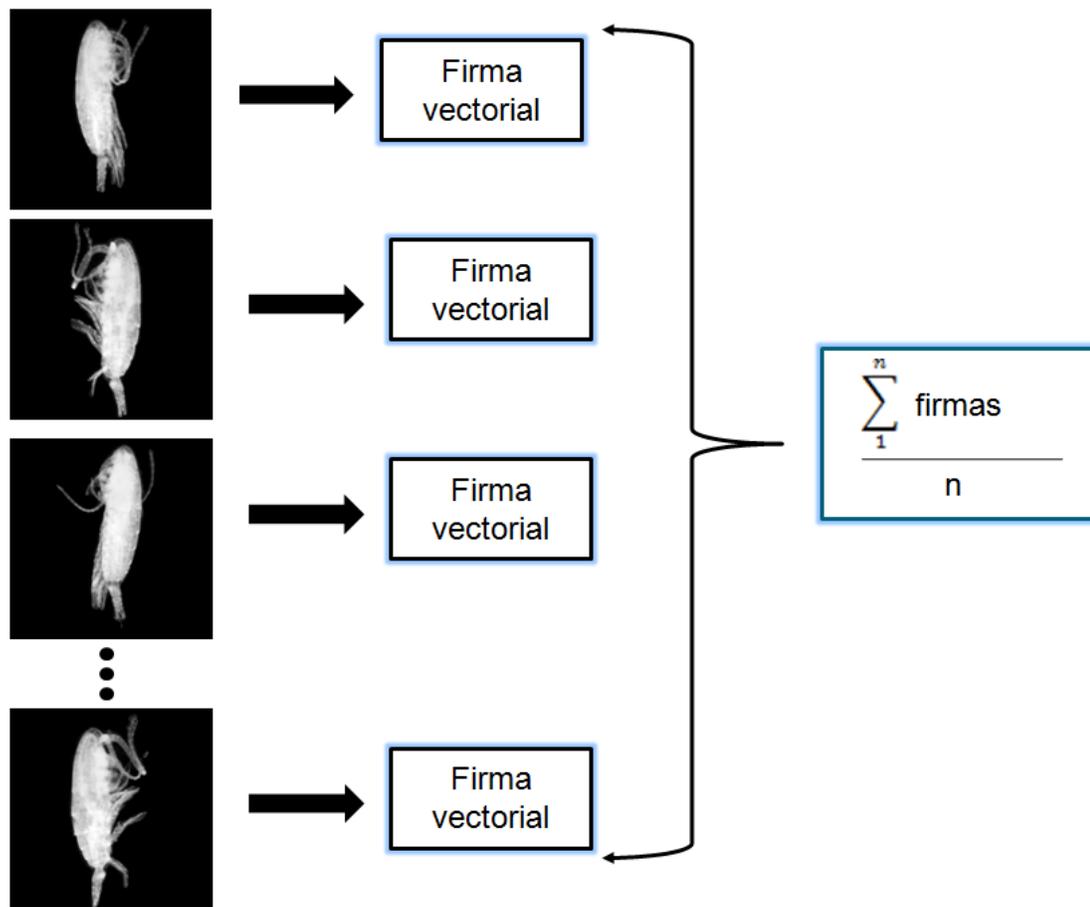


Figura 4.17. Firma vectorial promedio.

Para realizar nuestro análisis, a cada conjunto de copéodos se les denotó como grupo siguiendo la notación empleada en la Figura 4.12. Por ejemplo el grupo **a** corresponde a 30 ejemplares *Calanus pacificus* hembra. Las Figuras 4.18 y 4.19 muestran dos ejemplos de la aplicación de este método. Aquí podemos ver una excelente separación de la distancia euclidiana promedio, correspondiente en el primer caso a una firma vectorial promedio obtenida de 10 copéodos del grupo **a**, mientras que en el segundo caso se utilizaron 10 copéodos del grupo **e** (*Centropages furcatus* hembra).

En ambos casos es posible identificar la especie y el sexo del grupo seleccionado, con un nivel de confianza del 95.4 % al considerar el valor de la distancia euclidiana promedio $\pm 2EE$ (dos errores estándar) para cada grupo.

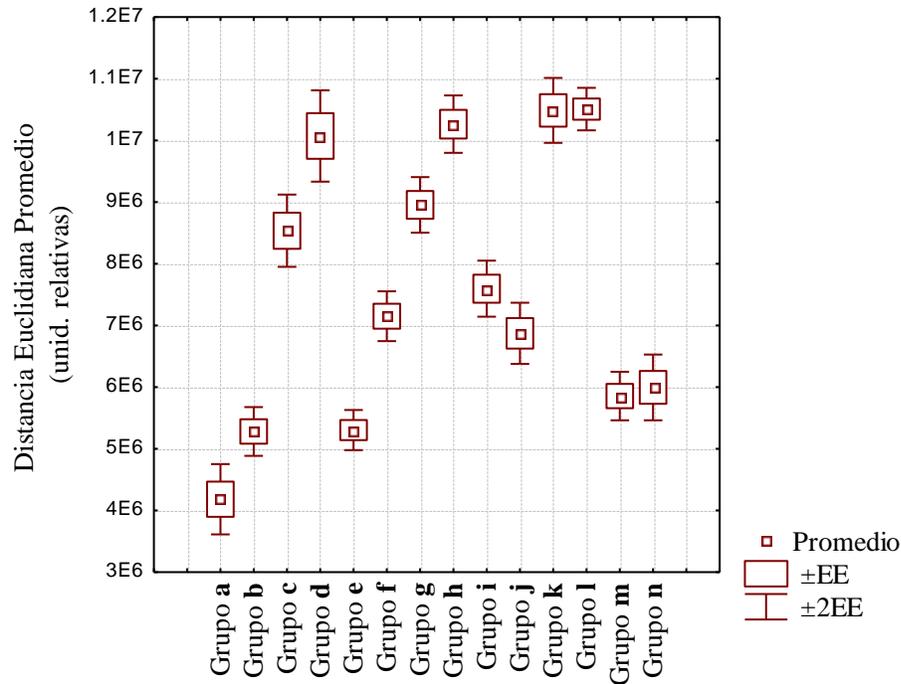


Figura 4.18. Comportamiento estadístico de la d_E , para la firma vectorial compuesta de 10 imágenes de *Calanus pacificus* hembra (Grupo a).

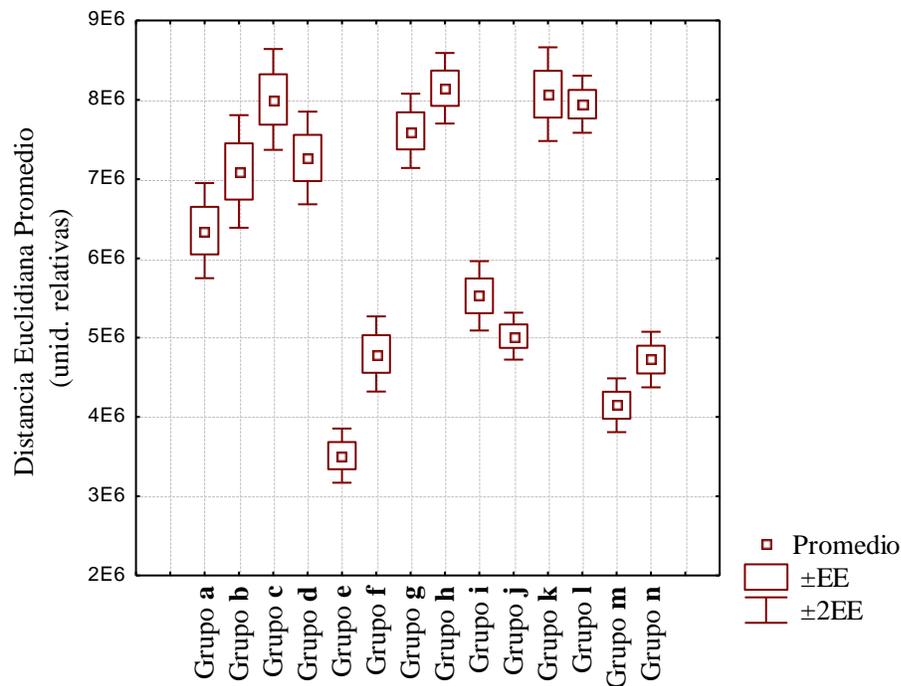


Figura 4.19. Comportamiento estadístico de la d_E , para la firma vectorial compuesta de 10 imágenes de *Centropages furcatus* hembra (Grupo e).

4.4.3. Experimento de identificación de especie

Se realizó una simulación numérica para estudiar la confianza de las firmas vectoriales como un método efectivo de identificación. Como problema se estableció encontrar ejemplares hembra de la especie *Calanus pacificus*.

La metodología utilizada fue la siguiente: en primer lugar se generó un patrón de referencia para la d_E formado por las firmas vectoriales promedio de un conjunto de 30 ejemplares de copépodos de la especie y género seleccionado. Se eligieron para su análisis 20 imágenes de diferentes copépodos, conociendo de antemano que cinco ejemplares pertenecen al grupo que quiere identificar. Cada una de las imágenes se giraron 180 grados para encontrar su d_E promedio. La Figura 4.20 muestra los resultados obtenidos; en ella la primera de las cajas contiene la información del patrón de referencia y posteriormente están todas las imágenes problema.

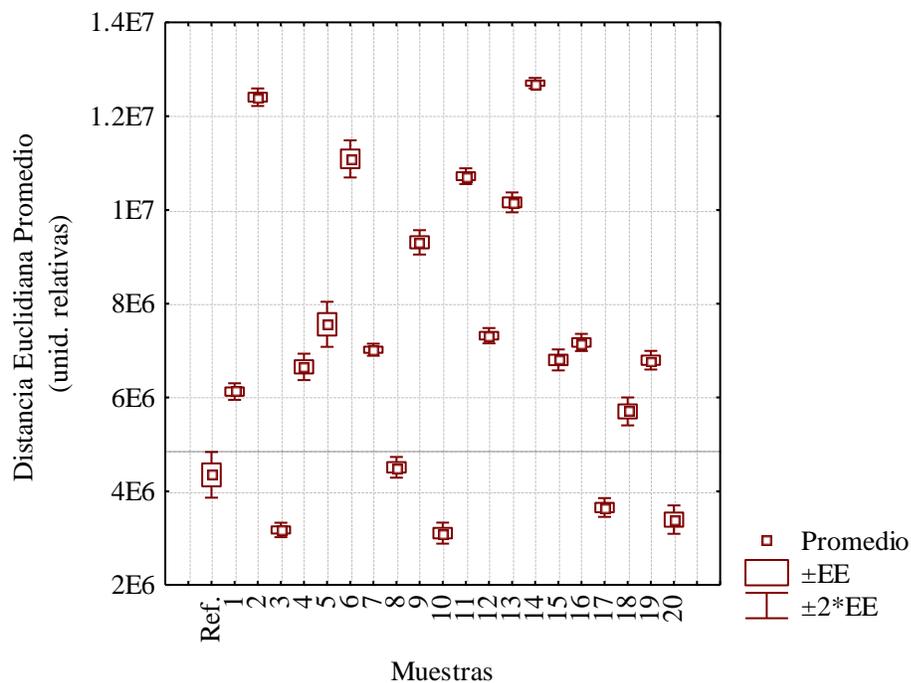


Figura 4.20. Resultados para identificación de ejemplares *Calanus pacificus* hembra.

Es posible observar en la Figura 4.20 que se tienen cinco muestras en las que la distancia es similar o menor a la distancia de referencia, lo que indica que pertenecen al grupo que se desea identificar (muestras 3, 8, 10, 17 y 20). La precisión del 95.4 % para la clasificación es comparable a lo que un biólogo entrenado puede lograr mediante el uso de las técnicas manuales convencionales, con la ventaja de realizarlo de manera rápida y sin necesidad de tener un entrenamiento previo.

4.5 Imágenes de color transformadas a escala de grises

Los resultados obtenidos en los apartados anteriores muestran que el algoritmo funciona muy bien para imágenes en escala de grises, pero surgió la duda respecto a como funcionaria para imágenes en color, para lo cual se aprovechó que éstas se pueden transformar a escala de grises.

La conversión de imágenes en color a escala de grises, es una técnica que se ha sido utilizada frecuentemente, donde se convierte el vector de tres dimensiones (3-D) del espacio RGB a un vector de una sola dimensión (1-D) en escala de grises, por ejemplo al utilizar una impresora en blanco y negro para imprimir imágenes en color.

Lee *et al.*, (2010) explican varios de los métodos utilizados para convertir imágenes en color a imágenes en escala de grises. Sin embargo, su funcionamiento aún no ha sido evaluado teóricamente porque la percepción humana del color no es sólo una relación con las características físicas de un individuo, sino también con sus características psicológicas, por lo que es difícil evaluar los métodos de conversión teóricamente. Las características psicológicas implican el significado implícito, la experiencia pasada, y la memoria. Cadik (2008) propone una evaluación de la apreciación subjetiva de la percepción de las conversiones de color a escala de grises, pero no ofrece un análisis teórico.

Para probar el algoritmo, fue seleccionado un conjunto de imágenes que consiste en mariposas del género *Caligo*, conocidas vulgarmente como mariposas búho, compuesto por 20 especies propias de las selvas localizadas desde el sur de México hasta el sur de Brasil, y

en Trinidad y Tobago. En la Figura 4.21 se muestran los tres ejemplares seleccionados que consisten en *Caligo eurylochus sulanos* (a), *Caligo memnon* (b), *Caligo oileus scamander* (c), los cuales tiene gran similitud en su forma y color.

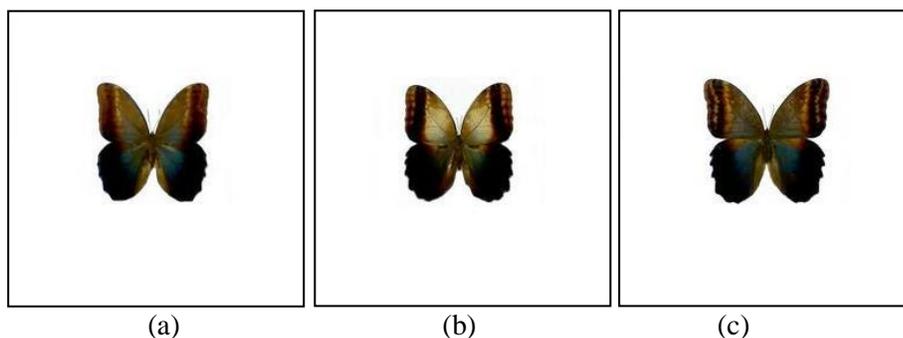


Figura 4.21. Mariposas utilizadas. (a) *Caligo eurylochus sulanos*. (b) *Caligo memnon*. (c) *Caligo oileus scamander*.

Los resultados obtenidos al realizar la simulación tomando como referencia el ejemplar *Caligo eurylochus sulanos*, para el caso de rotación de 360 grados con variaciones cada grado, se muestran en la Figura 4.22. Como se puede observar, tenemos un traslape de las curvas de las tres mariposas, indicándonos que existe confusión para la detección de la imagen de referencia.

Para tener una comparación mas detallada, con los datos obtenidos realizamos el análisis estadístico calculando su distancia euclidiana promedio y, al igual que en los casos anteriores utilizamos un diagrama en forma de cajas y bigotes, en el que se muestra el valor promedio $\pm 2EE$ (dos errores estándar). Como era de esperarse tenemos un traslape de las distancias promedio de las tres mariposas (Figura 4.23).

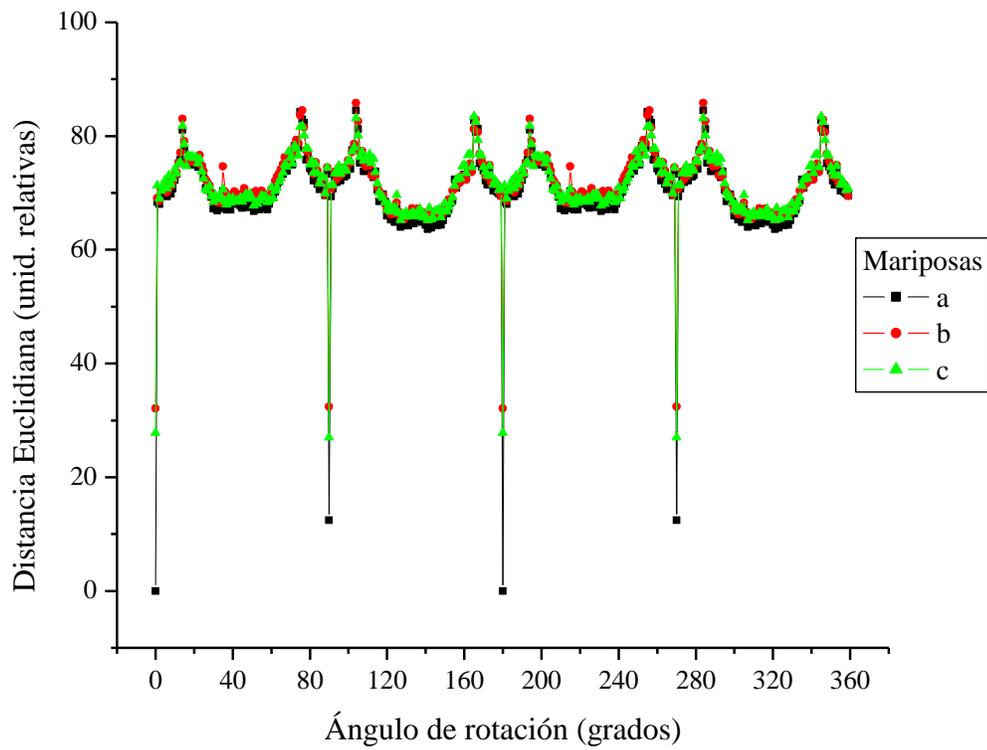


Figura 4.22. Cambios del ángulo de rotación contra d_E .

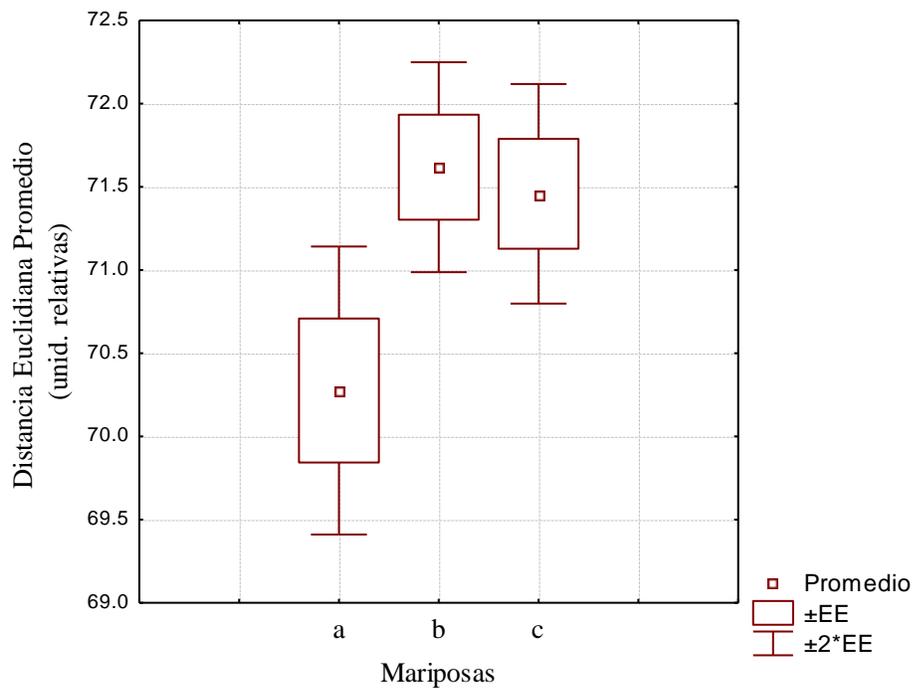


Figura 4.23. Comportamiento estadístico de la d_E tomando como referencia a *Caligo eurylochus sulanos* (Mariposa a).

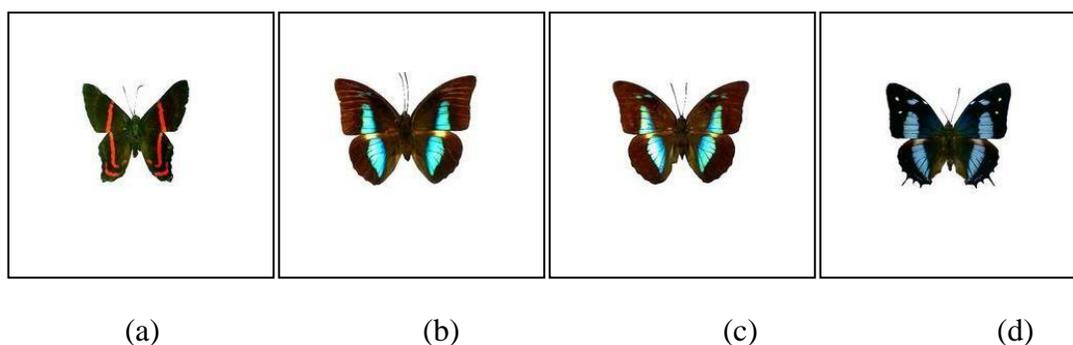


Figura 4.24. Ejemplares (a) *Ansyluris jurgennseni*. (b) *Archaeoprepona amphimachus amphiktion*. (c) *Archaeoprepona demophon centralis*. (d) *Baeotus baeotus*.

El segundo conjunto de imágenes seleccionado consiste en mariposas de diferente género, tomando como referencia un ejemplar de *Archaeoprepona amphimachus amphiktion* (una mariposa perteneciente a la familia Nymphalidae) la cual se encuentra desde México hasta Bolivia en las selvas tropicales y bosques húmedos de hojas caducas. En la Figura 4.24 se muestran los ejemplares seleccionados: (a) *Ansyluris jurgennseni*. (b) *Archaeoprepona amphimachus amphiktion*. (c) *Archaeoprepona demophon centralis*. (d) *Baeotus baeotus*.

En la Figura 4.25 tenemos los resultados de la simulación tomando como referencia *Archaeoprepona amphimachus amphiktion* (mariposa b) para el caso de rotación considerando los mismos parámetros del primer conjunto. Se puede apreciar que existe un traslape en la mayor parte de la gráfica. En la Figura 4.26 podemos ver el comportamiento estadístico, dando como resultado que no es posible identificar la imagen seleccionada, inclusive la menor de las distancias euclidianas no corresponde al ejemplar de referencia.

Es conocido que el ojo humano es bastante sensible al color, así que el nivel de escala de grises que puede discriminarse como tal es bastante más pequeño que el número de colores. Tomando esto en consideración, el siguiente capítulo corresponde al estudio de imágenes en color.

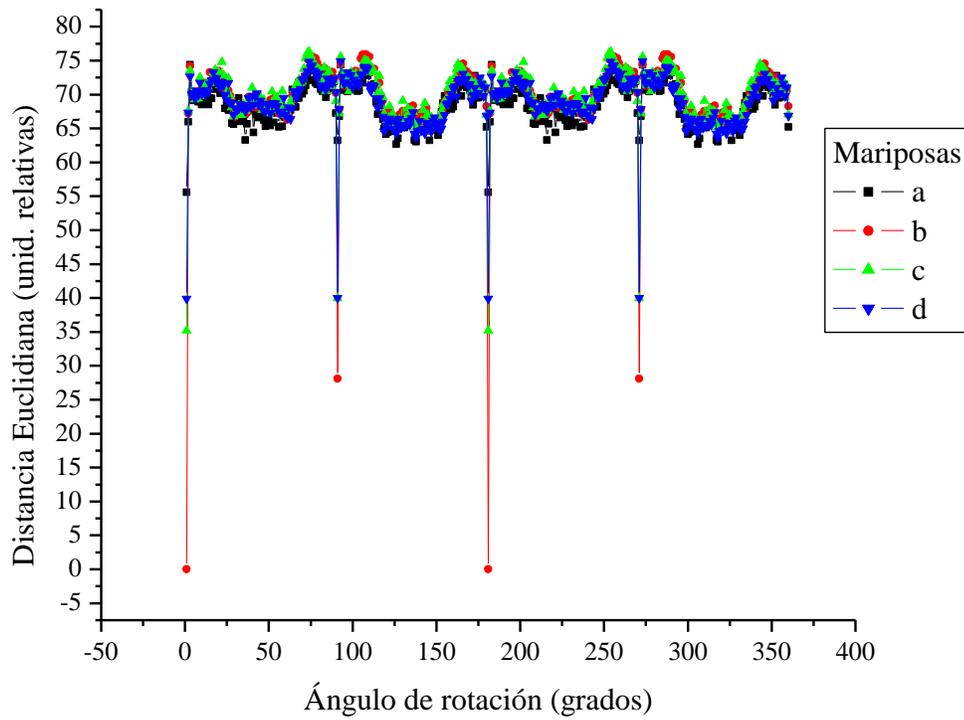


Figura 4.25. Cambios del ángulo de rotación contra d_E .

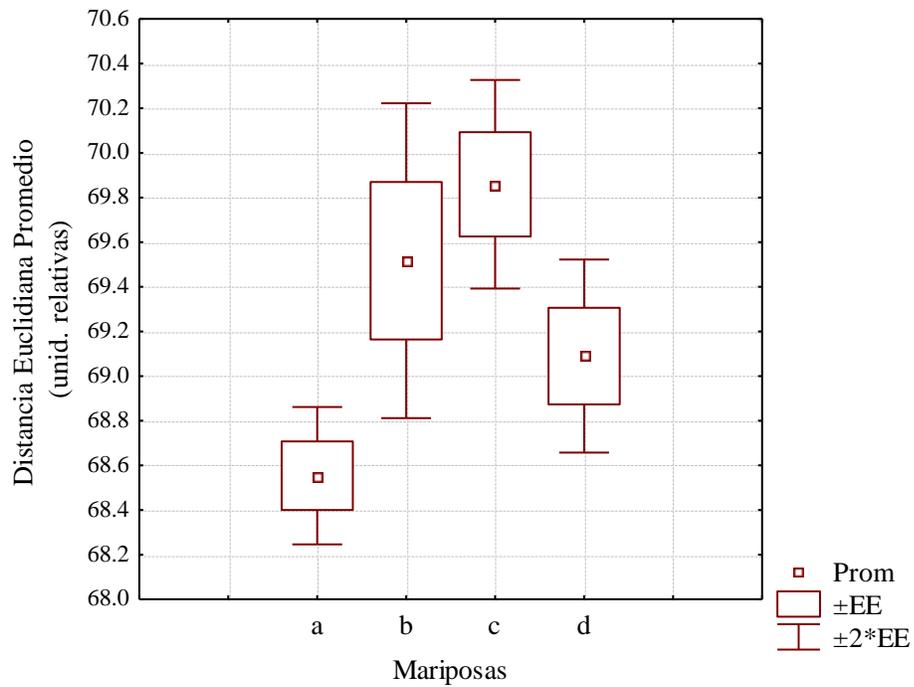


Figura 4.26. Comportamiento estadístico de la d_E tomando como referencia a *Archaeoprepona amphimachus amphiktion* (Mariposa b).

CAPÍTULO V

RECONOCIMIENTO DE PATRONES EN COLOR

5.1 Introducción.

5.2 Metodología.

5.3 Simulaciones numéricas.

5.3.1. Letras de colores.

5.3.2. Mariposas.

5.1 Introducción

Una de las características más importantes con las que se cuenta visualmente para el reconocimiento de patrones es el color, y ha sido utilizado en una gran cantidad de aplicaciones en diferentes áreas del conocimiento, como son la evaluación e identificación de textiles, flores, imágenes microscópicas, superficies de corrosión, reconocimiento de especies, etc. (Corbalán *et al.*, 2008, Bueno *et al.*, 2008, Garzón *et al.*, 2006, Couldridge y Alexander, 2002, Zhu, 1994, Yamaba, 1993).

En nuestra vida cotidiana, nuestra visión y acciones están influenciadas por una gran variedad de formas geométricas y sus colores. La introducción del color aumenta el contenido de la información en el reconocimiento de patrones, por lo que la discriminación puede ser mejorada considerablemente.

Al ser el color un fenómeno que se percibe y no una dimensión física como puede ser la longitud o la temperatura; en el procesamiento de imágenes es necesario contar con una forma adecuada de representarlo. Para esto se utilizan diferentes modelos que se conocen como espacios de color (Gonzalez y Woods, 2008, Koschan y Abidi, 2008, Westland y Ripamonti, 2004), los cuales deben estar adaptados a las exigencias matemáticas de los algoritmos de procesamiento de imágenes en color, a las condiciones técnicas de los diferentes dispositivos y, por supuesto, también con la percepción humana del color. Uno de los espacios de color más utilizado es el RGB (del inglés Red, Green, Blue); este modelo está basado en la mezcla aditiva de los tres colores primarios: rojo (R), verde (G) y azul (B).

El color incrementa la capacidad de discriminación porque hay objetos reales que difieren sólo en su distribución. Yu (1984) describió una técnica que se aplica a un sistema de correlación óptico coherente iluminado por luz policromática: utilizó tres fuentes coherentes colimadas (rojo, verde y azul); colocó una rejilla de difracción sobre la transparencia y en el plano de Fourier obtuvo separadamente la distribución de frecuencias para cada uno de los tres colores. Este tipo de técnicas pueden ser implementadas en tiempo

real en un sistema óptico coherente utilizando una pantalla de cristal líquido (Javidi *et al.*, 1988).

Badiqué *et al.*, (1987) y Badiqué *et al.*, (1988) convirtieron un vector de tres dimensiones (R, G, B) en uno de dos dimensiones mediante la proyección del vector sobre un plano de color, con la obtención de un plano de color óptimo, logrando una mejor discriminación. Pei y Hsu (1993) propusieron el uso de un mapeo logarítmico de los objetos de color, tomando como base la estructura de la retina humana para obtener un sistema de reconocimiento capaz de medir cambios en la magnitud de la escala y rotación.

Gu *et al.*, (1987) presentaron el diseño de un filtro que tiene información acerca del color y forma del objeto para el reconocimiento estadístico de imágenes a color. En la generación del filtro ellos consideran una transformación que extrae los rasgos de una clase y pasa del espacio de observación a un espacio de decisión con un mínimo de error, El correlador utiliza iluminación incoherente y la información de color es descompuesta en tres canales: rojo, verde y azul. Los resultados muestran que el poder de clasificación del algoritmo se incrementa cuando el filtro contiene información tanto del color como de la forma, en lugar de sólo tener información de la forma.

Millán *et al.*, (1989) y Millán *et al.*, (1992) identificaron objetos de color utilizando filtros lineales clásicos y sólo de fase para reconocer objetos con variación de forma y color y objetos únicamente con variación de color. También trabajaron en el reconocimiento de objetos con una característica y una combinación de color dada y reconocimiento de objetos sin importar la combinación de color.

Ninguno de estos trabajos estudió el problema del reconocimiento del objeto invariante a posición, rotación y a escala. Álvarez-Borrego *et al.*, (2002) propusieron un sistema de correlación digital invariante a posición, rotación y a escala para imágenes a color y Mouriño *et al.*, (2006) aplicaron la transformada de Escala y mapeo polar en un sistema de correlación digital para reconocer imágenes de color de la bacteria del cólera. Las imágenes de color fueron tomadas de muestras de laboratorio y de muestras

ambientales, logrando el reconocimiento de la bacteria con un nivel de confianza de al menos 95 %.

La diferencia fundamental entre las imágenes en color y escala de grises, es que en un espacio de color se tiene que a cada píxel de la imagen se le asigna un vector de color (que generalmente consta de tres componentes), mientras que en una imagen en escala de grises, a cada píxel se le asigna un solo escalar.

En muchas aplicaciones tecnológicas (astronomía, imágenes médicas, imágenes de satélite, exploración, minería) se utiliza la técnica de falso color. Ésta permite la transferencia de los valores de gris a pseudocolores, los cuales tienen como propósito hacer más visibles las áreas correspondientes.

El que un algoritmo tenga un buen desempeño en el reconocimiento de objetos en blanco y negro no significa que lo tenga también para el reconocimiento de imágenes en color. Por lo tanto el objetivo de este trabajo es demostrar que el trabajo desarrollado por Lerma-Aragón y Álvarez-Borrego (2009) puede también tener un buen desempeño en el reconocimiento de imágenes en color. El uso de firmas vectoriales en el reconocimiento de objetos da lugar a un bajo costo computacional, ya que el reconocimiento se lleva a cabo con vectores y no con matrices.

5.2 Metodología

En la Figura 5.1 se muestra el diagrama de bloques con la metodología aplicada a espacios de color: en el primer paso, se selecciona la imagen en color que se desee utilizar como referencia (I^i) a la que se denota con $f_i(x, y)$, donde $i=R, G, B$; después, cada imagen se descompone en sus respectivos canales RGB [$f_R(x, y), f_G(x, y), f_B(x, y)$]. A cada una de las tres componentes $f_i(x, y)$ se les calculan sus firmas vectoriales de rotación y escala utilizando las ecuaciones 3.5 y 3.6 obteniendo $V_{1f}(w_\lambda)$ y $V_{2f}(w_\theta)$ para cada uno de los canales; posteriormente se obtienen sus firmas vectoriales promedio. Se repite el proceso

anterior para las imágenes de prueba (I^j) representadas por $g_j(x, y)$. Para determinar la similitud entre las dos imágenes I^i y I^j , se calcula la distancia Euclidiana (d_E) entre ellas (ecuación 3.7).

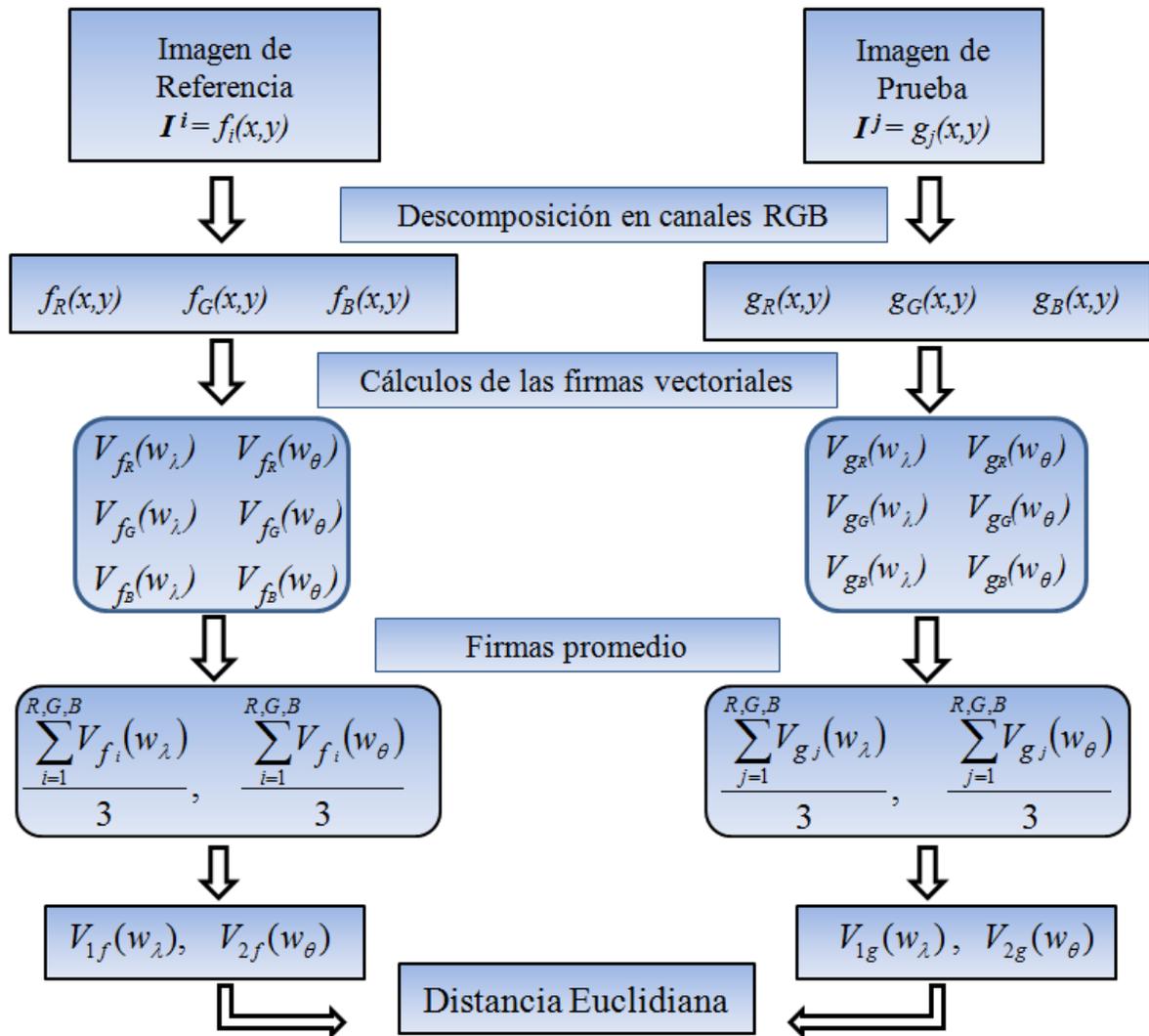


Figura 5.1. Diagrama de bloques del procedimiento utilizado.

El proceso para calcular las firmas vectoriales se puede visualizar como un conjunto de pasos representados en la Figura 5.2. A cada una de las imágenes (de referencia y problema) se le aplica la transformada de Fourier y se obtienen sus módulos $|F(w_x, w_y)|$,

donde w es la frecuencia angular. A éstos se les aplica un filtro parabólico (Pech-Pacheco *et al.*, 2003); de esta manera, se atenúan las bajas frecuencias, y las altas se realzan en proporción a $(w_x)^2$, $(w_y)^2$; esto da lugar a resaltar los detalles finos del objeto. Al resultado obtenido se le multiplica por el factor \sqrt{r} , donde r es la frecuencia espacial radial, indispensable para diferenciar la transformada de Escala de la Transformada de Mellin. A la matriz resultante se le transforma a coordenadas logarítmicas polares para obtener invariancia a rotación; en este paso se realiza una interpolación bilineal, evitando así la fuga de información debido al cambio de coordenadas; de esta imagen se selecciona una subimagen denotada por $M(\lambda, \theta)$ (Lerma *et al.*, 2008), zona que contiene la mayor parte de la información de referencia.

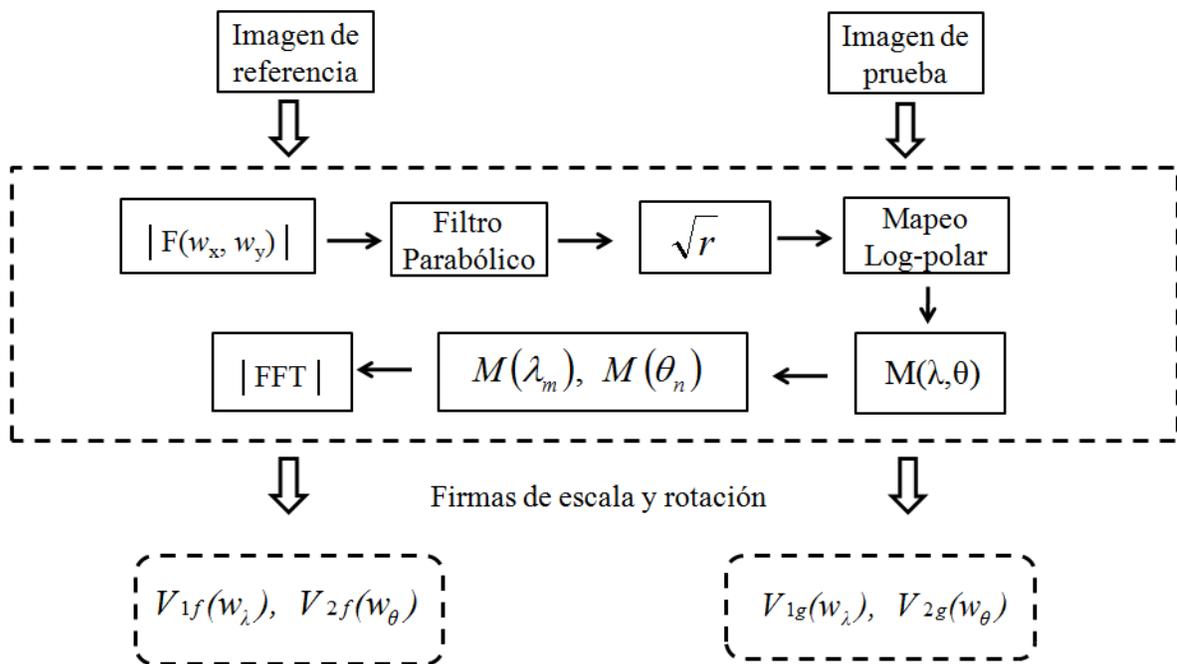


Figura 5.2. Diagrama detallado del cálculo de las firmas vectoriales.

5.3 Simulaciones numéricas

5.3.1. Estudio numérico para el caso de letras de colores

Para probar el algoritmo en primer término se analizaron imágenes de las letras “E, F, H, y T” escritas en letra tipo Arial de tamaño 72, las cuales tienen diferentes colores; éstas se eligieron dado su parecido geométrico. Lerma *et al.*, (2008) encontraron que en escala de grises es posible distinguirlas entre sí, pero no se puede generalizar para objetos de color.

En la Figura 5.3 se muestran las letras utilizadas como primer caso de estudio en colores rojo, verde y azul (R,G y B), con tamaño de 256 x 256 píxeles cada imagen; se procesaron usando fondo negro, de manera que solamente se toma la contribución de la letra y no del fondo; para fines estéticos, en la Figura 5.3 están mostradas con fondo blanco. Para analizar sus variaciones con respecto a rotación, se les aplicaron rotaciones de cero a 360 grados con variaciones de un grado. En la Figura 5.4 se muestran los resultados obtenidos al utilizar como imagen de referencia la letra H en color rojo (denotada por H_R) al compararla con todas las letras de la Figura 5.3 (expresadas con el nombre de la letra y en subíndices el color), donde podemos observar su comportamiento estadístico mediante un diagrama en forma de cajas y bigotes, en el que se muestra el valor promedio $\pm 2EE$ (dos errores estándar) lo que nos permite trabajar con el algoritmo a un nivel de confianza del 95.4 %.

Se puede notar que la menor de las d_E promedio corresponde a la imagen de la letra H roja (H_R), que fue la que utilizamos como referencia y los valores más cercanos a ella corresponden a la H en los otros colores. Es claro que para letras diferentes su d_E es mayor; esto nos permite tener la capacidad de poder distinguirla.

E	E	E	F	F	F
H	H	H	T	T	T

Figura 5.3. Letras en colores.

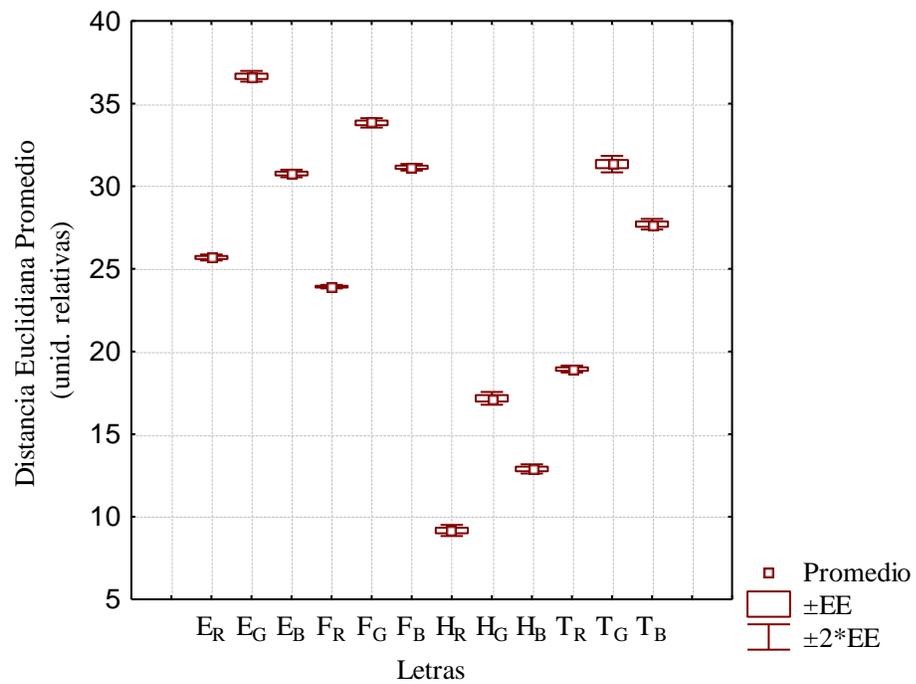


Figura 5.4. Comportamiento estadístico de la d_E tomando como referencia H_R.

5.3.2. Estudio numérico para el caso de mariposas

Después de probar el algoritmo, se tomó un segundo conjunto de imágenes (Figura 5.5), que corresponde a 18 tipos de mariposas diferentes; cada imagen tiene un tamaño de 256 x 256 píxeles. La razón por la cuál se eligió utilizar mariposas es que sus colores son un fenómeno muy interesante, porque provienen de la superficie alar, recubierta de escamas y posee multitud de aristas longitudinales (separadas a veces a menos de $1 \mu m$) que alteran la reflexión de la luz produciendo colores muy llamativos. Además se estima que el número de especies de mariposas en el mundo varía entre 15,000 y 20,000, de las cuales de acuerdo a la *North American Butterfly Association* (NABA), alrededor de 2000 se encuentran en México.

Como se puede observar en el conjunto de imágenes seleccionadas (Figura 5.5), se tienen ejemplares con gran parecido en su forma y color y, en algunos casos, en ambos. Para este conjunto de imágenes de prueba se realizaron simulaciones haciendo de nuevo rotar cada imagen con variación de un grado, partiendo de cero y hasta llegar a los 360 grados.

En la Figura 5.6 se puede observar el comportamiento estadístico de las d_E de todas las mariposas para el caso de rotación, utilizando como referencia la mariposa diez, que corresponde a un ejemplar *Archaeopreona amphimachus amphiktion*, la cual a simple vista tiene gran parecido con otras de las mariposas, por ejemplo con la mariposa once que pertenece a un ejemplar del mismo género pero de diferente especie. Como se esperaba, el valor mínimo de la d_E corresponde a nuestra mariposa de referencia y, estadísticamente hablando, las cinco más parecidas son las mariposas 8, 11, 12, 17 y 18, quedando demostrado que el algoritmo es capaz de distinguir la imagen deseada. Se realizaron simulaciones tomando como referencia cada una de las 18 mariposas y en ninguno de los casos hubo dificultad para reconocerlas.

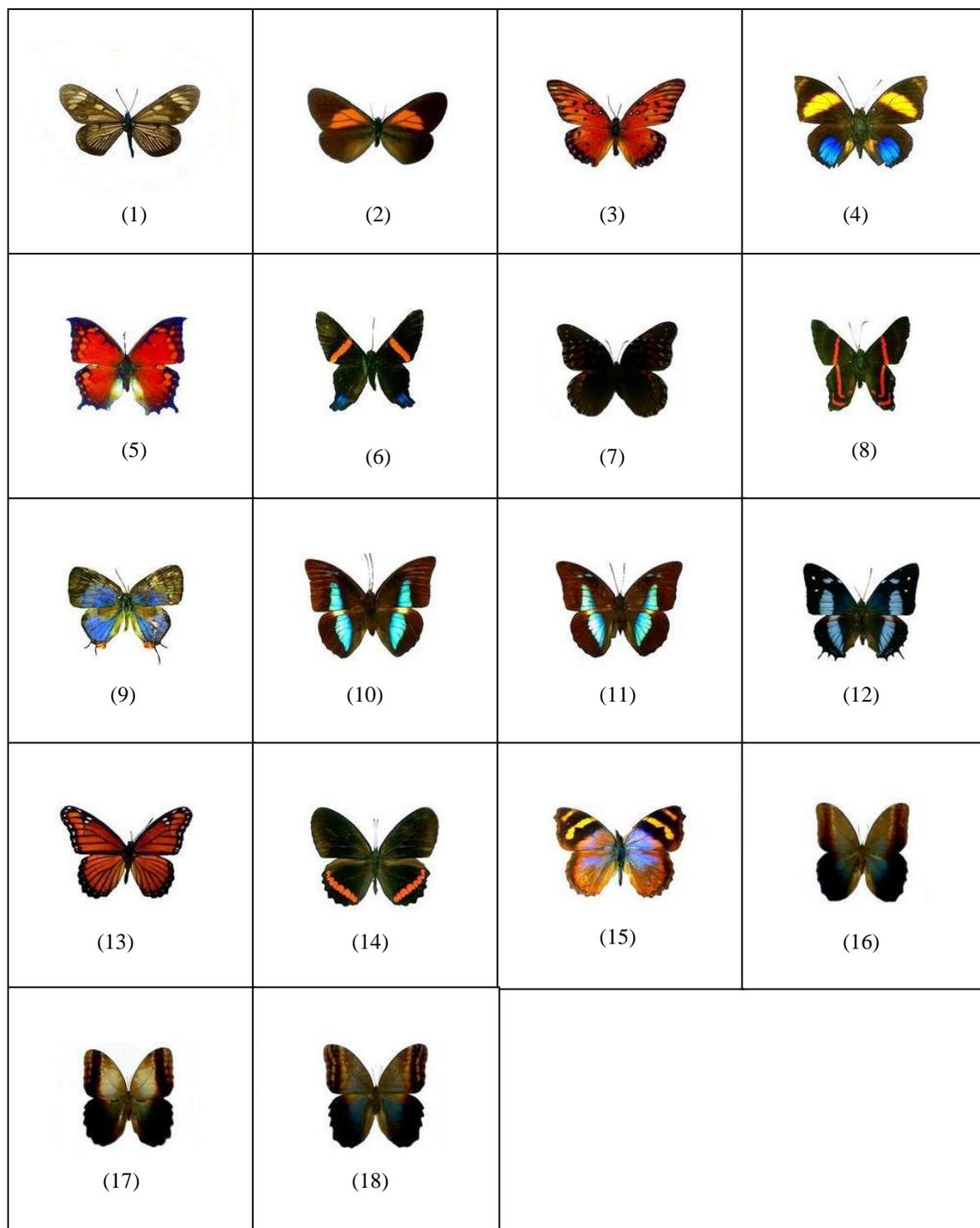


Figura 5.5. Mariposas utilizadas. (1) *Actinote guatemalena guerrensis*. (2) *Actinote stratonice oaxaca*. (3) *Agraulis vanillae incarnata*. (4) *Agrias amydon oaxacata*. (5) *Anae aidea*. (6) *Ancyluris inca mora*. (7) *Anetia thirza*. (8) *Ansyluris jurgennseni*. (9) *Arawacus sito*. (10) *Archaeoprepona amphimachus amphiktion*. (11) *Archaeoprepona demophon centralis*. (12) *Baeotus baeotus*. (13) *Basilarchia archippus*. (14) *Biblis hyperia aganissa*. (15) *Bolboneura Sylphis sylphis*. (16) *Caligo eurylochus sulanos*. (17) *Caligo memnon*. (18) *Caligo oileus scamander*.

Para poder analizar la invariancia a escala, se realizaron cambios del 70 % al 130 %, con variaciones del uno por ciento a cada una de las imágenes de prueba. El análisis del comportamiento estadístico de la distancia se muestra en la Figura 5.7 para el caso en que la mariposa tres es utilizada como referencia. Los resultados obtenidos nos indican que existe separación de la d_E de la imagen de referencia con respecto al resto de las mariposas, lo que permite identificarla sin dificultad, independientemente de los cambios de escala realizados. Se utilizaron como referencia cada uno de los 17 casos restantes y en todos los casos se identificaron las mariposas satisfactoriamente.

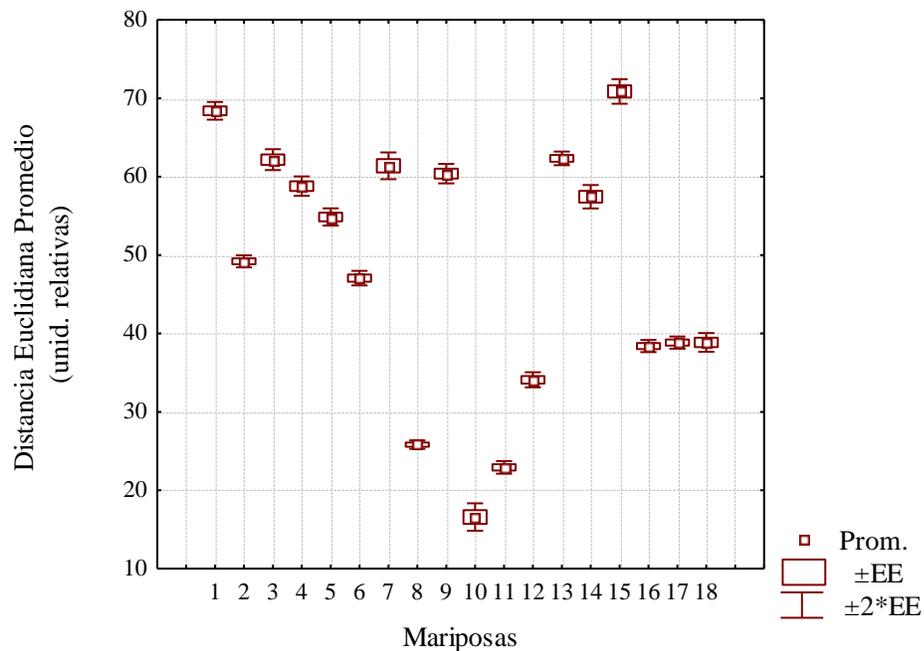


Figura 5.6. Comportamiento estadístico de la d_E para rotación de las todas las mariposas, tomando como referencia a la mariposa 10.

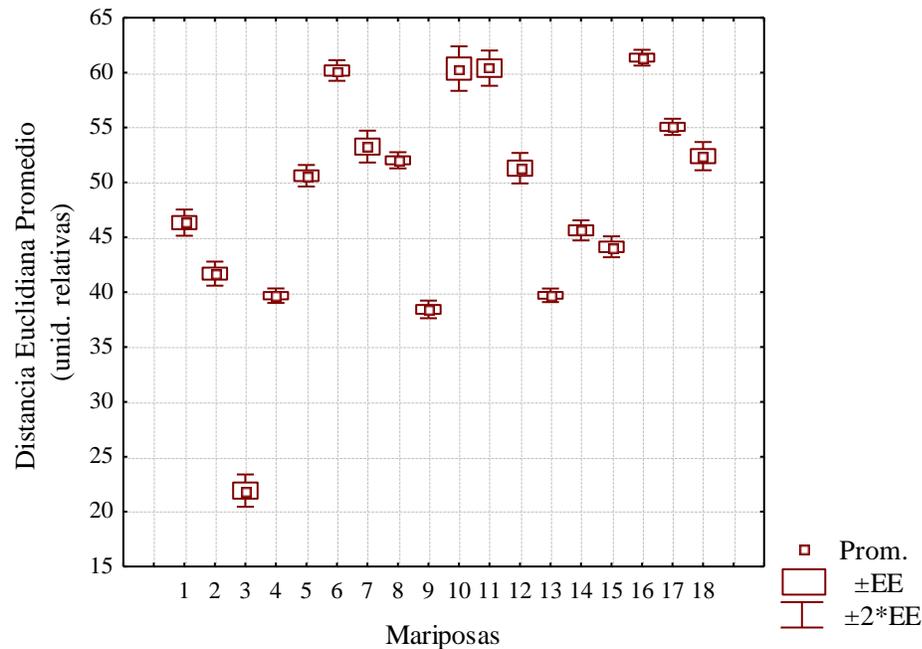


Figura 5.7. Comportamiento estadístico de la d_E para cambios de escala del 70 % al 130 %, tomando como referencia a la mariposa 3.

5.4. Trabajos derivados de la tesis

Parte de las investigaciones desarrolladas ya han sido publicadas en revistas arbitradas o presentadas en congresos internacionales y nacionales:

- Jesús R. Lerma-Aragón and Josué Álvarez-Borrego. “*Vectorial signatures for invariant recognition of position, rotation and scale pattern recognition*”. Journal of Modern Optics. 56. 14. 1598 – 1606. 2009.
- Jesús R. Lerma A., Josué Álvarez Borrego. “*Reconocimiento de caracteres basado en firmas vectoriales*”. Digital Scientific and Technological Journal, e-Gnosis. Vol. Esp. 2009. 1-7. 2009.

- Jesús R. Lerma A., Josué Álvarez-Borrego, José Ángel González-Fraga. “*Invariant recognition to position, rotation and scale considering vectorial signatures*”. Applications of Digital Image Processing XXXI, Proceedings of SPIE. 7073. 707329. 2008.
- Jesús Ramón Lerma Aragón, Josué Álvarez Borrego. José Ángel González-Fraga. “*Reconocimiento digital invariante mediante firmas vectoriales*”. Memorias en extenso arbitradas de la Reunión Anual de Óptica. 2008.
- Jesús Ramón Lerma Aragón, Josué Álvarez Borrego. José Ángel González-Fraga. “*Invariant correlation to position, rotation and scale considering vectorial signatures*”. Optical Engineering + Applications, Part of SPIE Optics +Photonics, Poster Presentation. San Diego, CA, USA. 2008.
- Jesús Ramón Lerma Aragón, Josué Álvarez Borrego. “*Identificación de especies de copépodos utilizando firmas vectoriales*”. Planctología 2009, XV Reunión de la Sociedad Mexicana de Planctología y VIII International Meeting of the Mexican Society of Planktology. 2009.
- Jesús Ramón Lerma Aragón, Josué Álvarez Borrego y José Ángel González-Fraga. “*Correlación digital invariante para el reconocimiento de patrones mediante firmas vectoriales*”. LI Congreso Nacional de Física y II Congreso Latinoamericano de Física. 2008.

CAPÍTULO VI

CONCLUSIONES

La presente investigación cubre un conjunto de objetivos relacionados con el análisis de imágenes. Habiendo estudiado y analizado las propiedades que presentan las transformadas de Fourier, Mellin y Escala se desarrollaron nuevas herramientas que permiten realizar el reconocimiento de patrones, independientemente de cambios en la traslación, rotación y escala que contenga la imagen.

Se desarrolló un nuevo algoritmo para el reconocimiento e identificación de patrones basado en la utilización de firmas vectoriales, desarrollado para ser práctico y preciso; con base en éste se elaboró e implementó un programa para analizar imágenes en escala de grises donde se demostró su eficacia para identificar patrones de diversos tipos; posteriormente fue adaptado para realizar el reconocimiento de imágenes en color. El algoritmo y su adaptación son aportación original de esta tesis.

Los principales resultados obtenidos de las simulaciones numéricas realizadas para analizar imágenes en escala de grises de forma automática son los siguientes:

En el conjunto formado por las 26 letras del abecedario escritas con fuente Arial, se examinaron dos casos: en primer término la invariancia a rotación del objeto, ejecutando para ello variaciones de un grado, iniciando en cero grados y aumentando hasta llegar a 360 grados (se analizaron en total 243,360 imágenes). Los resultados nos muestran que en la mayoría de los casos existe una separación muy clara entre los valores de la distancia euclidiana de la letra que se eligía como referencia al ser comparada con todo el alfabeto, lo que permite poder identificarla. El segundo caso fue invariancia en la escala, donde modificamos el tamaño de las imágenes desde un 80 % hasta llegar al 120 % con aumentos

cada 0.5 %; los resultados obtenidos de la comparación de las distancias euclidianas muestran que es posible identificarlas. Se realizaron además pruebas numéricas con imágenes de los dígitos del sistema decimal y con diferentes formas de hojas, en las que se hicieron variaciones en la escala y rotación similares a las que se hicieron con las letras; los resultados obtenidos de la comparación demuestran que es posible identificar y discriminar la imagen de referencia con un nivel de confianza del 95.4 %.

Tomando en cuenta los buenos resultados, el sistema fue aplicado en una base de datos que contenía imágenes de siete especies de copépodos para determinar su especie y sexo. El primer caso fue el análisis de un ejemplar hembra y un macho de cada una de las especies demostrando que, independientemente de los cambios en rotación, se puede identificar la especie y el sexo. Como todas las muestras corresponden a ejemplares adultos (son los que tienen valor comercial) no se consideraron cambios en el tamaño.

Al trabajar con imágenes de muestras reales, se desarrolló un sistema de firmas vectoriales promedio, en donde se consideró la información de varios ejemplares encontrando que es posible determinar con un nivel de confianza del 95.4 % la especie y el sexo de los copépodos.

Se realizó una modificación del algoritmo anterior para utilizarlo en el reconocimiento de imágenes en color, mediante la utilización de firmas vectoriales promedio obtenidas de la información contenida en los canales RGB. Se realizaron diversas pruebas con letras del alfabeto escritas en diferentes colores similares entre sí, a las cuales se les aplicaron rotaciones de 360 grados igual que en el caso de escala de grises. Los resultados obtenidos de la comparación de las distancias euclidianas demuestran que es posible identificarlas y discriminarlas con un nivel de confianza equivalente a los casos anteriores.

Se aplicó el algoritmo para el reconocimiento de imágenes de mariposas reales, encontrando que el sistema puede identificar la especie que se elija como referencia, al

compararla con las imágenes de diferentes mariposas, para cualquier ángulo de rotación que contenga, así como cambios en la escala del 70 % al 130 %.

La presente investigación contribuye a aumentar el potencial de los sistemas de reconocimiento para ser utilizados en aplicaciones que requieran una evaluación e interpretación de la información, aplicaciones que tradicionalmente se llevan a cabo mediante la visión humana entrenada de técnicos o especialistas y que están todavía lejos de la automatización.

A la luz de los resultados obtenidos, se puede aseverar que las firmas vectoriales, gracias a su precisión, son una técnica prometedora en el campo de la identificación de patrones. Además, se han abierto numerosas vías que podrán ser perfeccionadas con trabajos futuros.

BIBLIOGRAFÍA.

- [1] Almeida, S. P., and Kim-Tzong, J. “*Water pollution monitoring using matched spatial filters*”. *Appl. Opt.* 15. 510-515. 1976.
- [2] Almeida, S. P., Del Balzo, D. R., Cairns, J. Jr., Dickson, K. L., and Lanza, G. R., “*Holographic microscopy of diatoms*”, *Trans. Kans. Acad. Scie.* 74. 257-260. 1972.
- [3] Álvarez-Borrego, J. y Castro-Longoria, E. “*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. 2003.
- [4] Álvarez-Borrego, J. and Solorza, S. “*Comparative analysis of several digital methods to recognize diatoms*”. *Hidrobiológica.* 20. 2. 158-170. 2010.
- [5] Álvarez-Borrego, J. y Chávez-Sánchez, M. “*Detection of IHVN virus in shrimp by digital color correlation*”. *Aquaculture* 194. 1-9. 2001.
- [6] Álvarez-Borrego, J., Mouriño-Pérez, R. R., Cristóbal-Pérez, G., and Pech-Pacheco, J. L. “*Invariant recognition of polychromatic images of Vibrio cholerae 01*”. *Opt. Eng.* 41.4. 827-833. 2002.
- [7] Arfken, G. “*Métodos matemáticos para físicos*”. Ed. Diana, México. 1981
- [8] Badiqué, E., Komiya, Y., Ohyama, N., Tsujiuchi, J., Honda, T. “*Color image correlation*”. *Opt. Commun.* 61(3): 181-186. 1987.
- [9] Badiqué, E., Ohyama, N., Honda, T., Tsujiuchi, J. “*Color image correlation for spatial/spectral recognition and increased selectivity*”. *Opt. Commun.* 68. 2. 91-96. 1988.
- [10] Bracewell, R. N. “*The Fourier transform and its applications*”. McGraw-Hill, 2nd. Ed. New York. Chap.12. 1978.
- [11] Bradford, J. M. “*Partial revision of the Acartia subgenus Acartiura (Copepoda: Calanoida: Acartidae)*”. *New Zealand Journal of Marine and Freshwater Research.* 10.159-202. 1976.
- [12] Bueno, G., González, R., Deniz O. “*Colour model analysis for microscopic image processing*”. *Diagnostic Pathology.* 3. S18. 2008.
- [13] Bunke H. and Wang, P. S. P. “*Handbook of Character Recognition and Document Image Analysis*”. World Scientific Publishing. Singapore. 1997.
- [14] Cadik, M. “*Perceptual evaluation of color-to-grayscale image conversions*” *Comput. Graph. Forum.* 27. 7. 1745–1754. 2008.

- [15] Cairns, J. Jr., Dickson, K. L., and Lanza, G. R. “*Coherent optical spatial filtering of diatoms in water pollution monitoring*”. Arch. Mikrobiol. 83. 141-146. 1972.
- [16] Casasent, D. and Psaltis, D. “*Scale invariant optical correlation using Mellin transforms*”. Opt. Commun. 17. 59-63. 1976a.
- [17] Casasent, D. and Psaltis, D. “*Scale invariant optical transforms*”. Opt. Eng. 15. 258-261. 1976b.
- [18] Casasent, D. and Psaltis, D. “*Position, rotation, and scale invariant optical correlation*”. Appl. Opt. 15. 1795-1799. 1976c.
- [19] Castro-Longoria E., Álvarez-Borrego, J., Pech-Pacheco, J. L. “*Identification of species of calanoid copepods using a new invariant correlation algorithm*”. Crustaceana. 74. 10. 1029-1039. 2001.
- [20] Castro-Longoria, E., Álvarez-Borrego, J., Rocha-Olivares, A., Gómez, S., Kober, V. “*Power of a multidisciplinary approach: use of morphological, molecular and digital methods in the study of harpacticoid cryptic species*”. Mar. Ecol. Prog. Ser. 249. 297-303. 2003.
- [21] Cheriet, M., Kharna N., Cheng-Lin Liu, Suen, C. “*Character recognition systems, A guide for students and practioners*”. John Wiley & Sons, Inc., Hoboken, New Jersey. 2007.
- [22] Cohen, L. “*Time-Frequency Analysis*”. Prentice Hall Signal Processing Series, Upper Saddle River, NJ. 1995.
- [23] Cohen, L. “*The scale representation*”. IEEE Transactions on Signal Processing 41. 3275-3291. 1993.
- [24] Cooley, James W. and Tukey John W. “*An Algorithm for the Machine Calculation of Complex Fourier Series*”. Mathematics of Computation. 19. 90. 297-301. 1965.
- [25] Corbalán, M., Valencia, E. Vega, A. “*Assessment of the color space for the measure of the dominant color*”. En RIAO/OPTILAS: 6th Ibero-American Conference on Optics; 9th Latin- American Meeting on Optics, Lasers and Applications AIP Conference Proceedings. 992. 45-50. 2008.
- [26] Couldridge, V., Graham, A. “*Color patterns and species recognition in four closely related species of Lake Malawi cichlid*”. Behavioral Ecology. 13. 1. 59-64. 2002.
- [27] Cristobal, G. and Cohen, L. “*Scale in image*”. Proc. SPIE. 2846. 251-261. 1996.
- [28] D’ateur, Pas. “*Auerbach on Optical Character Recognition*”. Auerbach Publishers, Inc., Princeton. 1971.

- [29] Elfving, G. *"The History of Mathematics in Finland 1828-1918"*. Frenckell, Helsinki. 1981.
- [30] Forero-Vargas, M., Sroubek, F., Álvarez-Borrego, J., Malpica, N., Cristóbal, G., Santos, A., Alcalá, L., Desco, M. y Cohen, L. *"Segmentation, autofocusing and signature extraction of tuberculosis sputum images"*. The International Symposium on Optical Science and Technology. Seattle, Washington. 4788. 171-182. 2002.
- [31] Fosshagen, A. and Iliffe T. *"A new genus of calanoid copepod from an anchialine caves in Belize"*. Bull. Plankton Soc. Japan, 339-346. 1991.
- [32] Fujii, H. and Almeida, S. P. *"Partially matched spatial filtering with simulated input"*. SPIE. 177. Optical Information Storage. 1979a.
- [33] Fujii, H. and Almeida, S. P. *"Coherent spatial filtering with simulated input"*. Appl. Opt. 18. 1659-1662. 1979b.
- [34] Fujii, H., Almeida, S. P., Dowling, J. E. *"Rotational matched spatial filter for biological pattern recognition"*. Appl. Opt. 19. 1190-1193. 1980.
- [35] Garzón, J., Barrero, C., García, K.E., Pérez, F., Galeano, J., Salazar, A., Lorduy, H. *"Morphological analysis and classification of types of surface corrosion damage by digital image processing"*. Rev. Col. Fís. 38. 2. 557-560. 2006.
- [36] Gerardi, F. R. *"Applications of Mellin and Hankel transforms to networks with time-varying parameters"*. IRE Transactions on Circuit Theory. 6. 197-208. 1959.
- [37] Goodman, J.W., (1996). *"Introduction to Fourier Optics"*. 2a Edition.. McGraw-Hill, New York.
- [38] Gonzalez, Rafael C., Woods, Richard E., Eddins, Steven L. *"Digital Image Processing Using MATLAB[®]"*. Pearson Prentice Hall, chapters 9-12, Upper Saddle, NJ. 2009.
- [39] Gonzalez, Rafael, Woods, Richard E. *"Digital Image Processing"*. 2a Edition. Upper Saddle. NJ. Pearson Prentice Hall. 2008.
- [40] Guerrero-Moreno, R.E. and Álvarez-Borrego, J. *"Nonlinear composite filter performance"*. Opt. Eng. 48. 6. 067201. 2009.
- [41] Gu-Zu, Han., Lee, Sing., Fainman, Y. *"Statistical recognition of color images"*. Appl. Opt. 26. 15. 3145-3152. 1987.
- [42] Humes, A. G. *"How many copepods?"* F. D. Ferrari & B. P. Bradley (eds.). Ecology and morphology of copepods. Hydrobiologia, 292-293. 1994.

- [43] Humes, A. G. “*Zoogeography of copepods at hydrothermal vents in the eastern Pacific Ocean*”. Bull. Plankton Soc. Japan. 383-389. 1991.
- [44] Javidi, B., Kuo, Chung, Jung., Feng, Ying., Ludman, Jacques. “*Color object identification by monochromatic binary correlation*”. Appl. Opt. 27. 5. 949-953. 1988.
- [45] Javidi., B. “*Image Recognition and Classification, Algorithms, Systems, and Applications*”. Marcel Dekker, Inc. New York, NY. 2002.
- [46] Kober, V., Mozerov, M., Álvarez-Borrego, J., and Ovseevich, I. “*An Efficient Algorithm for Suppression of Impulsive Noise in color image*”. Pattern Recognition and Image Analysis. 15. 1. 219–222. 2005.
- [47] Koschan, A., Abidi, M. “*Digital Color Image Processing*”. First Edition. Hoboken, New Jersey. John Wiley & Sons, Inc. 2008.
- [48] Lerma, J.R., Álvarez-Borrego, J., González, Á. “*Invariant recognition to position, rotation and scale considering vectorial signatures*”. Applications of Digital Image Processing XXXI. 7073 707329-1. 2008.
- [49] Lerma, J.R. and Álvarez-Borrego, J. “*Vectorial signatures for invariant recognition of position, rotation and scale pattern recognition*”. Journal of Modern Optics. 56. 14. 1598-1606. 2009.
- [50] Marques de Sá, J. P. “*Pattern Recognition: Concepts, Methods and Applications*”. Springer-Verlag, Berlin Heidelberg New York. 2001.
- [51] Millán, M., Campos, J., Ferreira, C., Yzurl, M. “*Matched filter and phase only filter performance in colour image recognition*”. Opt. Commun. 73. 4. 277-284. 1989.
- [52] Millán, M., Yzuel, M., Campos J., Ferreira, C. “*Different strategies in optical recognition of polychromatic images*”. Appl. Opt. 31. 14. 2560-2567. 1992.
- [53] Mori, S., Nishida H., and Yamada. H. “*Optical Character Recognition*”. Wiley Interscience, New Jersey. 1999.
- [54] Mouriño, R., Álvarez-Borrego, J., Gallardo, C. “*Correlación digital del color para el reconocimiento de Vibrio cholerae 01 en muestras ambientales y de laboratorio*”. Revista de Biología Marina y Oceanografía. 41. 1. 77-86. 2006.
- [55] Obinata, G. and Morris, G. “*Vision systems, segmentation and pattern recognition*”, I-Tech Education and Publishing. 2007.
- [56] Pech-Pacheco, J. L., and Álvarez-Borrego, J. “*Optical-digital system applied to the identification of five phytoplankton species*”. Marine Biology. 132. 357-365. 1998.

- [57] Pech-Pacheco, J., Cristobal, G., Álvarez-Borrego, J., and Cohen, L. “*Automatic system for phytoplanktonic algae identification*”. *Limnetica*. 20. 1. 143-158. 2001.
- [58] Pech-Pacheco, J., Álvarez-Borrego, J., Cristóbal, G., Matthias Keil. “*Automatic object identification irrespective of geometric changes*”. *Opt. Eng.* 42. 551-559. 2003.
- [59] Solorza, S. and Álvarez-Borrego, J. “*Digital system of invariant correlation to position and rotation*”. *Optics Communications*. 283. 19. 3613-3630. 2010.
- [60] Soo-Chang, P. Min-Jor, H. “*Two-dimensional invariant color pattern recognition using a complex log mapping transform*”. *Opt. Eng.* 32. 7. 1616-1622. 1993.
- [61] Pratt, W. K. “*Digital image processing*”. John Wiley & Sons, Inc., Hoboken, NJ. 2007.
- [62] Psaltis, D., Paek, E., Venkatesh, S. Yu-S. “*Optical image correlation with a binary spatial light modulator*”. *Opt. Eng.* 23. 6. 698-704. 1984.
- [63] Reid, C.E. and Passin, T.B. “*Signal Processing in C*”. Wiley. 1992.
- [64] Rice, S. V., Nagy, G. and Nartker, T. A. “*Optical Character Recognition: An Illustrated Guide to the Frontier*”. Kluwer Academic Publishers, Boston. 1999.
- [65] Sawchuk, J. A. “*Space-variant image restoration by coordinate transformations*”. *J. Opt. Soc. Amer.* 138-144. 1974.
- [66] Schantz, H. F. “*The History of OCR*”. Recognition Technologies Users Association, Boston. 1982.
- [67] Schwartz, E. “*Topographic mapping in primate visual cortex: history, anatomy and computation*”. *Visual Science and Engineering. Models and Applications*. Ed. Marcel Dekker, New York. 1994.
- [68] Schwartz, E. L. “*Afferent geometry in the primate visual cortex and the generation of neuronal trigger features*”. *Biol. Cybernetics*. 28. 1-14. 1977.
- [69] Soh, H. Y. & Suh, H.L. “*A new species of Acartia (Copepoda, Calanoida) from the Yellow Sea*”. *J. Plankton Res.* 22. 321-337. 2000.
- [70] Suen, C. Y. Character recognition by computer and applications “*Handbook of Pattern Recognition and Image Processing*”. Academic Press, Inc., Orlando, FL. 569–586. 1986.
- [71] Lee, T. Kim, B.K., Song, W.J. “*Converting color images to grayscale images by reducing dimensions*”. *Opt. Eng.* 49. 5. 057006. 2010.

- [72] Theodoridis, Sergios and Koutroumbas, Konstantinos. "*Pattern Recognition*". Academic Press is an imprint of Elsevier. 2009.
- [73] Titchmarsh, E.C. "*Introduction to the Theory of Fourier Integrals*". Clarendon Press, Oxford. 1975.
- [74] Tsagaris, V., Anastassopoulos, V. "*Multispectral image fusion for improved RGB representation based on perceptual attributes*". International Journal of Remote Sensing. 26. 15. 3241- 3254. 2005.
- [75] Vilenkin, N.Ya. "*Special Functions and the Theory of Group Representation*". Am. Math. Soc. Providence, RI. 1968.
- [76] Westland, Stephen., Ripamonti, Caterina. "*Computational Colour Science using MATLAB*". West Sussex PO19 8SQ, England. John Wiley & Sons Ltd, The Atrium, Southern Gate. 2004.
- [77] Yamaba, K., Miyake, Y. "*Color character recognition method based on human Perception*". Opt. Eng. 32. 1. 33-40. 1993.
- [78] Zavala-Hamz, V. A. and Álvarez-Borrego, J. "*Circular harmonic filters for the recognition of marine microorganisms*". Appl. Opt. 36. 484-489. 1997.
- [79] Zhu, Zimin. "*Color pattern recognition in an image system with chromatic distortion*". Opt. Eng. 33. 9. 3047-3051. 1994.