

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



**PROGRAMA DE POSGRADO EN CIENCIAS
EN ÓPTICA**

**Metodología basada en la transformada de escala y técnicas
de correlación no lineal con invariancia a distorsiones
geométricas e iluminación**

Tesis

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

Presenta:

Alfredo Solís Ventura

Ensenada, Baja California, México
2015

Tesis defendida por
Alfredo Solís Ventura

y aprobada por el siguiente Comité

Dr. Josué Álvarez Borrego
Director del Comité

Dra. Selene Solorza Calderón
Miembro del Comité

Dra. Diana Tentori Santa Cruz
Miembro del Comité

Dr. Héctor Manuel Escamilla Taylor
Miembro del Comité

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

Dr. Jesús Favela Vara
Director de Estudios de Posgrado

Agosto, 2015

Resumen de la tesis que presenta Alfredo Solís Ventura como requisito parcial para la obtención del grado de Doctor en Ciencias en Óptica con orientación en Óptica Física.

Metodología basada en la transformada de escala y técnicas de correlación no lineal con invariancia a distorsiones geométricas e iluminación.

Resumen elaborado por:

Alfredo Solís Ventura

En este trabajo se desarrolla una metodología con invariancia a posición, rotación y escala utilizando las propiedades de la transformada de escala, la cual es un caso particular de la transformada de Mellin. Esta nueva metodología se inicia centrando la imagen objetivo y obteniendo una máscara binaria de anillos concéntricos creada a partir del transecto central del módulo de la transformada de escala de la imagen objetivo. Al realizar una multiplicación punto por punto de la máscara de anillos y el módulo de la transformada de escala de la imagen se obtienen una serie de anillos concéntricos que contienen muestras radiales del módulo de la transformada de escala de la imagen. Cada uno de estos anillos contiene información invariante a la escala de la imagen. La invariancia a rotación se consigue al sumar los valores contenidos en cada anillo y de esta manera se obtiene una firma vectorial capaz de representar a la imagen que se desea identificar, a pesar de estar desplazada, rotada y con cambios de escala; que además es robusta ante la presencia de ruido y cambios de iluminación. Como método de comparación entre firmas vectoriales se utiliza una correlación no lineal, a la cual se le agrega la característica de ser adaptativa mediante un índice invariante a rotación y a escala, que se obtiene con una relación entre el área y el volumen de una imagen problema con la imagen objetivo. Se demuestra la efectividad del método utilizando imágenes de 21 diatomeas fósiles y de 30 especies de fitoplancton en escala de grises; se utilizaron versiones escaladas y giradas de estas imágenes. Además se presentan resultados de identificación de imágenes cuando estas se encuentran incompletas, ante iluminaciones no uniformes o inmersas en ruido.

Palabras clave: **Transformada de escala, correlación no lineal, máscara binaria, firmas vectoriales.**

Abstract of the thesis presented by Alfredo Solís Ventura as a partial requirement to obtain the Doctor of Science degree in Optics with orientation Optical Physics

Methodology based in the scale transform and nonlinear correlation techniques with invariance to geometric distortions and illumination.

Abstract by:

Alfredo Solís Ventura

In this work a methodology with invariance to position, rotation and scale is developed using the properties of the scale transform, this is a particular case of the Mellin transform. The new methodology starts centering the target image and obtaining a binary mask of concentric rings created from the central transect of the scale transform modulus of the target image. Calculating the point to point multiplication of the binary mask with the modulus of the scale transform of the image, a set of concentric rings were obtained with the specific sampled values of the modulus of the scale transform of the image. The information in these rings contain information invariant to the scale of the image and also give us the property of rotation invariance. Making the sum of the values in each ring, a vectorial signature is obtained. This signature can represent a target image with invariance to position, rotation and scale even in presence of noise or with changes in the illumination. As comparison method between vectorial signatures, a nonlinear correlation was used. An adaptive characteristic was added to this nonlinear correlation with the aid of a scale and rotation invariant index. This was obtained from a ratio between the mean and the volume values of the problem image and the target image. The effectiveness of the method was demonstrated using images of 21 fossil diatoms and 30 phytoplankton species in gray scale, scaled and rotated versions of each of these images were used. Also results of the identification of images when these were incomplete, with no uniform illumination or immersed in noise.

Keywords: **Scale Transform, nonlinear correlation, binary mask, vectorial signatures**

Dedicatoria

A mi esposa Yaneth Pino.
Mi cómplice, mi mejor amiga y la mujer de mi vida.

A mis hijos Vanessa y Santiago.
*Mis tesoros más preciados e invaluables fuentes de amor,
son mi principal motivo de superación.*

Agradecimientos

Al CICESE, porque además de tener una excelente infraestructura, durante mis estudios de maestría y doctorado, me abrió sus puertas y me permitió utilizar sus instalaciones, haciéndome sentir como en casa.

Al Dr. Josué Álvarez Borrego, que ha sido sin duda, la pieza más importante en mi formación científica, gracias por su tiempo y consejos para lograr el término de este trabajo.

A los miembros del comité de tesis, Dra. Selene Solorza Calderón, Dra. Diana Tentori Santa Cruz y Dr. Héctor Manuel Escamilla Taylor. Ya que además de darme sus consejos y apoyo durante los avances y la revisión de esta tesis, los considero un ejemplo a seguir.

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

A mis padres Alicia Ventura Baltazar y Alfredo Solís Reyes, quienes son mi ejemplo de vida y superación. Gracias por apoyarme incondicionalmente durante mis estudios de doctorado.

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

Tabla de contenido

	Página
Resumen español.....	ii
Resumen inglés	iii
Dedicatorias.....	iv
Agradecimientos	v
Lista de Figuras.....	viii
Lista de Tablas.....	xii
Capítulo 1. Introducción	1
1.1 Máscaras binarias.....	2
1.2 Filtros de correlación basados en la transformada de escala	5
1.3 Motivación	6
1.4 Objetivos.....	6
1.5 Estructura de la tesis	7
Capítulo 2. Marco teórico.....	9
2.1 Transformada de Fourier	9
2.2 Transformada discreta de Fourier	11
2.3 Propiedades de la transformada discreta de Fourier 2D	12
2.3.1 Propiedad de desplazamiento	12
2.3.2 Propiedad de linealidad	13
2.3.3 Propiedad de escalamiento.....	13
2.3.4 Propiedad de rotación.....	13
2.3.5 Propiedad de periodicidad.....	14
2.3.6 Propiedad de simetría conjugada	14
2.3.7 Propiedad de separabilidad	14
2.3.8 Teorema de convolución	14
2.3.9 Teorema de correlación.....	15
2.4 Transformada rápida de Fourier.....	15
2.5 Transformada de Mellin	16
2.6 Transformada de escala	17

Tabla de contenido (continuación)

	Página
2.7 Filtros de correlación	20
2.7.1. Filtro de acoplamiento clásico	20
2.7.2 Filtro solo de fase.....	21
2.7.3 Filtro no lineal	21
2.8 Métricas de desempeño.....	22
2.8.1 Capacidad de discriminación.....	22
2.8.2 Energía pico de correlación (PCE)	22
2.9 Ruido en la imagen.....	23
2.9.1 Ruido gaussiano	23
2.9.2 Ruido Impulsivo.....	24
Capítulo 3. Metodología.....	25
3.1. Máscara binaria de anillos concéntricos	25
3.2. Firma vectorial de la imagen	29
3.3. Correlación no lineal adaptativa	31
3.4 Metodología invariante a distorsiones geométricas causadas por cambios en posición, rotación y escala	34
Capítulo 4. Resultados.....	36
4.1 Identificación de diatomeas fósiles.....	36
4.2 Identificación de fitoplancton.....	40
4.3 Imágenes con diatomeas fósiles incompletas.....	49
4.3.1 Fragmentos de diatomeas fósiles como Imágenes problema.....	49
4.3.2. Fragmentos de diatomeas fósiles como imagen objetivo	53
4.4 Capacidad para reconocer imágenes afectadas por ruido.....	55
4.5 Reconocimiento de imágenes con iluminación no homogénea	58
4.6 Reconocimiento de imágenes reales con rotación, escala, iluminación y fondo natural	61
Capítulo 5. Conclusiones.....	67
Lista de referencias	70
Apéndice A.....	75

Lista de figuras

	Página
Figura 1. Ubicación de la máscara binaria en un sistema óptico.	2
Figura 2. Detector RWD. (a) RWD con 32 anillos y 32 transectos (George et al, 1972). (b) RWD modificado (Ganotra et al, 2004).	3
Figura 3. Cuatro diferentes máscaras adaptativas asociadas a la imagen de la diatomea fósil <i>Actinocyclus ingens</i> . Estas máscaras se generaron a partir de, (a) la parte real y positiva de la transformada de Fourier, (b) la parte real y negativa de la transformada de Fourier, (c) la parte imaginaria y positiva de la transformada de Fourier, (d) la parte imaginaria y negativa de la transformada de Fourier (Solorza y Álvarez-Borrego, 2015).	5
Figura 4. Implementación de la transformada de escala separable en 2D vía la transformada rápida de Fourier.	26
Figura 5. Procedimiento para obtener la máscara de anillos concéntricos.	27
Figura 6. Digitalización de la transformada de escala. (a) Imagen de la “letra T Arial”. (b) Módulo de la versión separable de la transformada de escala en 2D correspondiente a la “letra T Arial”. (c) Perfil $ D_{IO}(c_x, c_y) $. (d) Digitalización a partir de las pendientes de $z(c_y)$ donde $c_y=1, \dots, 128$	28
Figura 7. Máscara binaria de anillos para la “letra T Arial”.	28
Figura 8. Procedimiento para obtener la firma vectorial de la imagen. (a) Firma vectorial de la “letra O Arial”. (b) Firma vectorial de la “letra T Arial”.	29
Figura 9. Efecto en las firmas vectoriales debido a distorsiones en la imagen causadas por cambios en rotación o escala. (a) Firma vectorial para las “letras T Arial” rotadas. (b) Firma vectorial para las “letras T Arial” escaladas. (c) Firma vectorial para las “letras O Arial” rotadas. (d) Firma vectorial para las “letras O Arial” escaladas.	30
Figura 10. Procedimiento para obtener un índice invariante a rotación y escala, donde $f(x,y)$ representa a una imagen digital.	33
Figura 11. Procedimiento para lograr el reconocimiento de imágenes invariante a distorsiones geométricas causadas por cambios en posición, rotación y escala a partir de la transformada de escala en 2D y una correlación no lineal adaptativa.	35
Figura 12. Imágenes de diatomeas utilizadas para probar la metodología.	38

Lista de figuras (continuación)

Página

Figura 13. Resultados de correlaciones obtenidas utilizando la nueva metodología basada en la transformada de escala con invariancia a posición, rotación y escala (Figura11). (a) Filtro para diatomea B. (b) Filtro diatomea C. (c) Filtro diatomea I. (d) Filtro diatomea S. (e) Filtro diatomea T.	39
Figura 14. Imágenes de las 30 especies de fitoplancton usadas para probar la metodología adaptativa.	41
Figura 15. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo <i>Acanthogonyaulax spinifera</i> (IM01). (a) Usando correlación no lineal C_{NL} . (b) Usando correlación no lineal adaptativa C_{AD}	44
Figura 16. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo <i>Histioneis</i> (IM05). (a) Usando correlación no lineal C_{NL} . (b) Usando correlación no lineal adaptativa C_{AD}	45
Figura 17. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo <i>Podolampas palmipes</i> (IM13). (a) Usando correlación no lineal C_{NL} . (b) Usando correlación no lineal adaptativa C_{AD}	46
Figura 18. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo <i>Podolampas spinifer 2</i> (IM14). (a) Usando correlación no lineal C_{NL} . (b) Usando correlación no lineal adaptativa C_{AD}	47
Figura 19. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo <i>Podolampas spinifer 2</i> (IM03). (a) Usando correlación no lineal C_{NL} . (b) Usando correlación no lineal adaptativa C_{AD}	48
Figura 20. Ejemplo de cinco fragmentos de la diatomea <i>Actinocyclus ingens</i> , utilizados en el experimento.	50
Figura 21. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo la diatomea (A) entera y como imágenes problemas las 21 diatomeas y sus 49 fragmentos rotados y escalados.	51
Figura 22. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo la diatomea (B) entera y como imágenes problemas las 21 diatomeas y sus 49 fragmentos rotados y escalados.	51

Lista de figuras (continuación)

	Página
Figura 23. Resultados de correlación para la imagen con un fragmento del 73% de la diatomea A. (a) Utilizando C_{NL} . (b) Utilizando C_{AD}	53
Figura 24. Resultados de correlación para la imagen con un fragmento del 6% de la diatomea A. (a) Utilizando C_{NL} . (b) Utilizando C_{AD}	54
Figura 25. Proceso para determinar el valor del coeficiente de discriminación (CD), donde el símbolo * especifica una multiplicación punto a punto.	55
Figura 26. Rendimiento del filtro en presencia de ruido sal y pimienta.	57
Figura 27. Rendimiento del filtro en presencia de ruido gaussiano.	57
Figura 28. Ejemplos de la diatomea fósil <i>Actinocyclus ingen</i> , afectada por ruido aditivo. (a) Sal y pimienta. (b) Gaussiano.....	58
Figura 29. Patrones de iluminación no homogénea.	59
Figura 30. Ejemplos de la diatomea fósil <i>Actinocyclus ingens</i> , afectada por los diferentes patrones de iluminación no homogénea.....	59
Figura 31. Reconocimiento de imágenes con iluminación no homogénea. (a) Filtro para diatomea B. (b) Filtro diatomea C. (c) Filtro diatomea I. (d) Filtro diatomea S. (e) Filtro diatomea T.	60
Figura 32. Imágenes reales de fitoplancton con rotación, escala, iluminación y fondo natural.	62
Figura 33. Resultado de la multiplicación de la imagen real por el disco binario, donde * representa a la multiplicación punto a punto.	62
Figura 34. Resultados del valor promedio de los máximos de correlación utilizando imágenes en escenarios reales como imágenes problema y como imagen objetivo una imagen limpia de la especie de fitoplancton, (A) <i>Acanthogonyaulax spinifera</i> . (a) Utilizando la correlación no lineal (C_{NL}). (b) Utilizando la correlación no lineal adaptativa (C_{AD}).....	64
Figura 35. Resultados del valor promedio de los máximos de correlación utilizando imágenes en escenarios reales como imágenes problema y como imagen objetivo una imagen limpia de la especie de fitoplancton, (B) <i>Ornithocercu armata</i> . (a) Utilizando la correlación no lineal (C_{NL}). (b) Utilizando la correlación no lineal adaptativa (C_{AD}).....	64

Lista de figuras (continuación)

	Página
Figura 36. Resultados del valor promedio de los máximos de correlación utilizando imágenes en escenarios reales como imágenes problema y como imagen objetivo una imagen limpia de la especie de fitoplancton, (C) <i>Dinophysis hastate</i> . (a) Utilizando la correlación no lineal (C_{NL}). (b) Utilizando la correlación no lineal adaptativa (C_{AD}).....	65
Figura 37. Resultados del valor promedio de los máximos de correlación utilizando imágenes en escenarios reales como imágenes problema y como imagen objetivo una imagen limpia de la especie de fitoplancton, (D) <i>Ornithocercus magnificus</i> . (a) Utilizando la correlación no lineal (C_{NL}). (b) Utilizando la correlación no lineal adaptativa (C_{AD}).....	65
Figura 38. Resultados del valor promedio de los máximos de correlación utilizando imágenes en escenarios reales como imágenes problema y como imagen objetivo una imagen limpia de la especie de fitoplancton, (E) <i>Podolampas bipes 1</i> . (a) Utilizando la correlación no lineal (C_{NL}). (b) Utilizando la correlación no lineal adaptativa (C_{AD}).....	66

Lista de tablas

Página

Tabla 1.	Resultados del porcentaje de confiabilidad obtenido con la correlación no lineal (C_{NL}) y la versión adaptativa (C_{AD}).	42
Tabla 2.	Resultados del porcentaje de confiabilidad obtenido con la correlación no lineal (C_{NL}) y las diatomeas incompletas como imágenes problema.	52

Capítulo 1. Introducción

Los seres humanos tenemos una gran capacidad para identificar un objeto dentro de una imagen sin importar la posición, la rotación, el tamaño y la iluminación que este objeto presente en la imagen, inclusive en ocasiones podemos llegar a identificar un objeto aun cuando solo tenemos presente una pequeña parte de este, de la misma manera podemos llegar a distinguir un objeto parcialmente oculto y bajo condiciones de ruido. Sin embargo, hay casos en donde se requiere de expertos taxónomos para la identificación y conteo de microorganismos a través de imágenes, lo cual es una tarea tediosa ya que requiere de bastante tiempo de trabajo, con el consecuente desgaste físico y con la probabilidad del error humano. Por esto, es deseable desarrollar una metodología para la identificación automática de objetos en imágenes digitales que trabaje con rapidez.

Las investigaciones sobre reconocimiento de objetos estudian diferentes maneras para que las máquinas logren observar el entorno que las rodea para aprender y distinguir entre diferentes objetos de interés y realizar decisiones acertadas en cuanto a la clasificación del objeto observado. Esto se logra gracias a la extracción de diferentes características y atributos contenidos en los objetos, que permiten desarrollar algoritmos robustos que sean capaces de tomar en cuenta el conjunto de características del objeto. Algunos de estos atributos pueden incluso no ser detectables por el ojo humano.

La metodología realizada en este trabajo se basa en la extracción de características del módulo de la transformada de escala de las imágenes. Utilizando máscaras binarias adaptativas de anillos concéntricos, se toman muestras del módulo de la transformada de escala. A partir del contenido de cada anillo, se genera una firma vectorial que representa a una imagen con las características principales del objeto a reconocer.

La comparación entre las firmas de la imagen objetivo con las firmas de diferentes imágenes problema se lleva a cabo mediante una correlación digital no lineal, la cual

nos da información de la similitud entre diferentes firmas de imágenes con gran capacidad para discriminar objetos.

A continuación se presenta como han evolucionado las máscaras binarias y el desarrollo de filtros de correlación basados en la transformada de escala. Cabe señalar que la metodología desarrollada para este trabajo se basa principalmente en el diseño de un nuevo filtro de correlación no lineal de firmas vectoriales obtenidas a partir de máscaras binarias con invariancia a posición, rotación y escala.

1.1 Máscaras binarias

Una máscara binaria es un filtro espacial que se utiliza para tomar muestras del patrón de difracción y reducir la cantidad de datos que representan a una imagen. Estos datos pueden formar una firma vectorial unidimensional, facilitando y agilizando cualquier tipo de procesamiento posterior. La figura 1 muestra la ubicación de la máscara binaria en un sistema óptico.

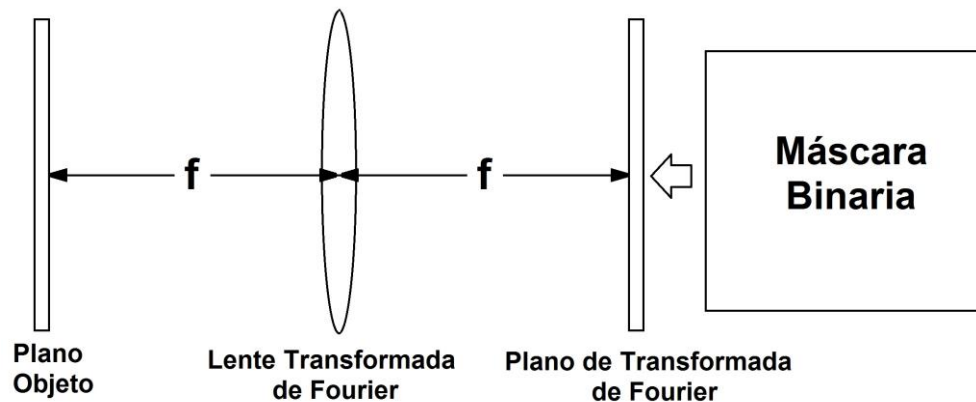


Figura 1. Ubicación de la máscara binaria en un sistema óptico.

La representación de una imagen a partir de muestras de su patrón de difracción fue una aportación de Lendaris y Stanley (1965), quienes tomaron muestras a partir de círculos o líneas rectas. Esta técnica se llegó a utilizar para agilizar el análisis de imágenes aéreas. Años más tarde utilizaron dos máscaras binarias: una de anillos concéntricos y otra de transectos de circunferencia (Lendaris y Stanley, 1970). El muestreo con la máscara de anillos concéntricos logró muestras de un rango de frecuencias espaciales específico del patrón de difracción en dos dimensiones y la firma

obtenida tiene la característica de ser invariante a posición y rotación. Con las muestras obtenidas a partir de los transectos de circunferencia separados 9° entre sí, se obtuvo una invariancia a posición y escala logrando obtener la aproximación de la posición del mejor transecto, lo cual da una referencia en grados de la rotación de la imagen problema. Máscaras como estas fueron utilizadas en el análisis de radiografías para el diagnóstico de enfermedades pulmonares (Krueger et al., 1974) y diferentes aplicaciones de tipo industrial tales como inspección de agujas para jeringas y detección de defectos de fábrica en telas (Kasdan, 1979). El dispositivo de la figura 2a fue diseñado para tomar muestras de un patrón de difracción, y es el detector RWD (*Ring Wedge Detector*) que consiste en dos regiones de medio círculo; un semicírculo está compuesto por una serie de anillos concéntricos y el otro se forma con transectos de circunferencia separados por ángulos equidistantes (George et al., 1972). Este dispositivo tiene como objetivo tomar muestras invariantes a rotación con los semicírculos y toma muestras invariantes a escala con los sectores de círculo. La figura 2b muestra una versión modificada del detector RWD, esta modificación reduce la cantidad de muestras y ha sido aplicada en la detección de huellas digitales (Ganotra et al. 2004). Una máscara tipo RWD tiene la ventaja de ser invariante a rotación y escala, pero cuando distorsiones generadas por rotación y escalamiento están presentes al mismo tiempo, interpretar los resultados de muestreo resulta complicado. Esto se debe a que si el patrón de difracción sufre cambios de rotación, las muestras recopiladas por los anillos permanecen constantes, pero los datos de los transectos circulares separados formando ángulos equidistantes serán muy variables.

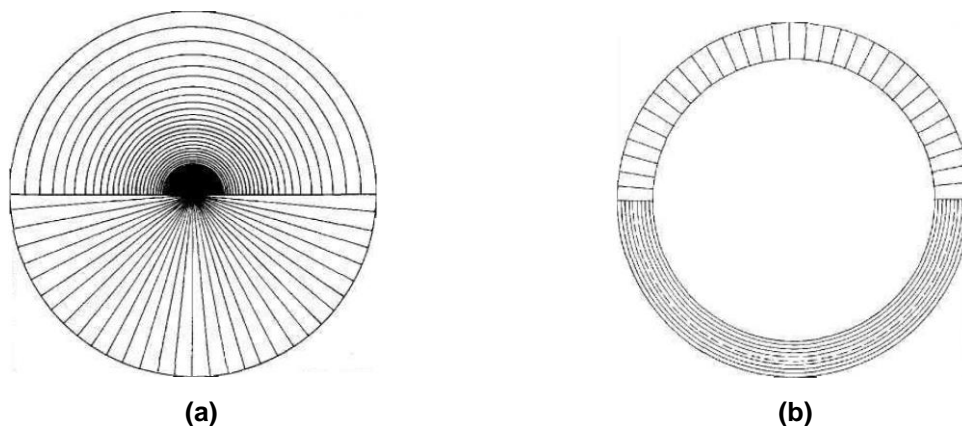


Figura 2. Detector RWD. (a) RWD con 32 anillos y 32 transectos (George et al, 1972). (b) RWD modificado (Ganotra et al, 2004).

Ahora, si suponemos cambios de escala, las muestras obtenidas de los transectos circulares permanecerán constantes pero las muestras recopiladas por los anillos variarán (Asselin et al. 1994). Por este motivo, si solo se requiere una invariancia a rotación, se recomienda utilizar máscaras de anillos concéntricos. Una investigación en donde solo se han utilizado máscaras binarias de anillos concéntricos fue desarrollada por Fares et al. (2000), quienes obtienen firmas de las imágenes tomando muestras radiales del módulo de la transformada de Fourier y mejoran sus resultados tomando muestras radiales de la fase de la transformada de Fourier.

En investigaciones recientes realizadas por el grupo de procesamiento de imágenes del departamento de óptica del CICESE, se han desarrollado diferentes sistemas de correlación digital para el reconocimiento de patrones utilizando filtros simples y compuestos, lineales y no lineales, invariantes a posición y rotación, mediante el uso de máscaras binarias. Iniciando con la primera máscara binaria adaptativa de anillos concéntricos creada a partir de la parte real de la transformada de Fourier, se toman únicamente muestras radiales de las frecuencias espaciales más importantes contenidas en la imagen, logrando la identificación de objetos muy similares entre sí, aun cuando estos están rotados aleatoriamente (Solorza y Álvarez-Borrego, 2010). Estas máscaras adaptativas también pueden ser creadas a partir de la parte imaginaria de la transformada de Fourier de la imagen objetivo, y se puede conseguir invariancia a escala utilizando filtros compuestos (Chaparro-Magallanes, 2011; Solorza et al., 2012). Además, se ha comprobado que hay diferentes alternativas para generar estas máscaras binarias de anillos concéntricos y mejorar la invariancia a rotación (Solís-Ventura, 2012; Solorza y Álvarez-Borrego, 2013; Solorza y Álvarez-Borrego, 2015). La figura 3 se presenta cuatro máscaras binarias adaptativas diferentes, asociadas a una misma imagen.

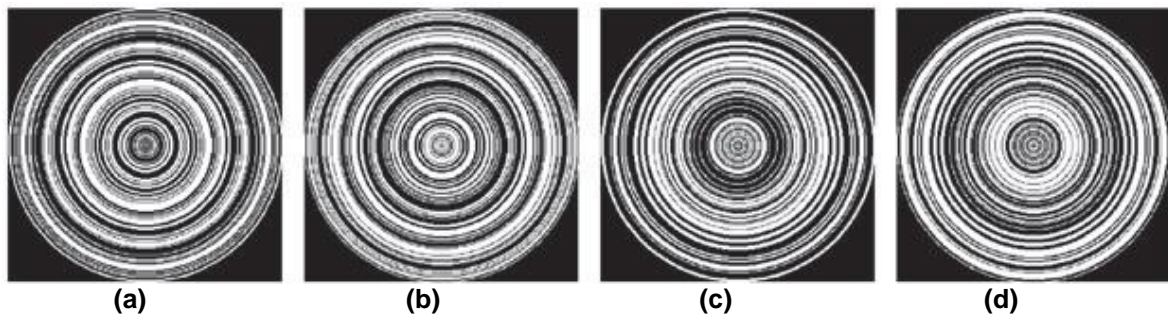


Figura 3. Cuatro diferentes máscaras adaptativas asociadas a la imagen de la diatomea fósil *Actinocyclus ingens*. Estas máscaras se generaron a partir de, (a) la parte real y positiva de la transformada de Fourier, (b) la parte real y negativa de la transformada de Fourier, (c) la parte imaginaria y positiva de la transformada de Fourier, (d) la parte imaginaria y negativa de la transformada de Fourier (Solorza y Álvarez-Borrego, 2015).

1.2 Filtros de correlación basados en la transformada de escala

Diferentes investigaciones relacionadas con la invariancia a escala utilizan como pre-procesamiento un mapeo log-polar seguido de una transformada de Fourier para lograr cierta invarianza a escala (Brousil y Smith, 1967; Robbins y Huang, 1972; Casasent y Psaltis, 1976; Schwartz, 1977; Wieman y Chaiken, 1979). Robbins y Huang (1972) fueron los primeros en utilizar la transformada discreta de Fourier seguida de un mapeo log-polar de una imagen de entrada para aproximar la transformada de Mellin.

Casasent y Psaltis (1976) propusieron el uso de la transformada de Mellin, en combinación con la transformada de Fourier, para realizar la operación de correlación óptica invariante a posición y escala. También demostraron la invariancia a rotación. Sus resultados confirmaron que es posible correlacionar ópticamente dos objetos o funciones que difieran considerablemente en escala y rotación.

La transformada de escala fue desarrollada por Cohen (1993), quien encontró que esta es un caso particular de la transformada de Mellin, la cual tiene la propiedad de ser invariante a cambios de escala. Diferentes implementaciones digitales de la transformada de Mellin y la transformada de escala han sido publicadas (Zwicke y Kiss, 1983; Zalubas y Williams, 1995; Djurdjanovic et al., 1999), pero fueron De Sena y Rocchesso (2007) quienes publicaron la transformada rápida de Mellin y la transformada rápida de escala logrando reducir el costo computacional al implementar la transformada rápida de Fourier en su algoritmo. Esta metodología ha sido utilizada en

los trabajos de Pech-Pacheco et al. (2003), en donde se realizó la identificación de objetos basándose en la correlación de fase en el dominio de la transformada de escala para el reconocimiento automático de caracteres, logrando invariancia a posición, rotación y escala. Lerma-Aragón (2009) desarrolló firmas vectoriales, basadas en las propiedades de las transformadas de Fourier y de escala, para el reconocimiento de patrones invariante a posición, rotación y escala, con las que se logró identificar por especie y sexo diferentes especies de copépodos. Coronel-Beltrán et al. (2010) compararon diferentes tipos de letras utilizando una correlación digital no lineal invariante a rotación y escala. Fimbres-Castro et al. (2012) diseñaron un sistema invariante utilizando índices espectrales y firmas vectoriales obtenidas a partir de las transformadas de Fourier y de escala, logrando la identificación de 21 diatomeas fósiles.

1.3 Motivación

A pesar del avance tecnológico y de las numerosas herramientas, métodos y algoritmos existentes para el reconocimiento de objetos, aún no se ha logrado desarrollar un algoritmo capaz de indicarle a una computadora como identificar todo tipo de objetos, ya que para la máquina, un mismo objeto en diferente posición, rotación o escalamiento representará otra imagen diferente. Esto dificulta una identificación automática, ya que las distorsiones causadas por diferentes transformaciones geométricas, pueden generar un número infinito de imágenes de un mismo objeto. Por esto, en esta tesis se desarrolla una metodología invariante a distorsiones generadas en las imágenes digitales, causadas por transformaciones geométricas, tales como, cambios en posición, rotación y escala que además pueden incluir variaciones de iluminación y ruido.

1.4 Objetivos

El objetivo general de esta tesis es desarrollar un nuevo sistema no lineal de identificación de objetos que conserve la invariancia a posición y rotación, e incluya además la invariancia a escala, para el reconocimiento de imágenes, aun cuando estas estén bajo condiciones de iluminación no uniforme, inmersas en ruido, parcialmente ocultas o incompletas.

Los objetivos particulares son:

1. Estudiar y analizar las propiedades de invariancia que poseen diferentes transformaciones matemáticas.
2. Realizar un nuevo banco de imágenes de prueba, el cual además contendrá imágenes reales de diatomeas fósiles utilizadas en trabajos previos.
3. Desarrollar un nuevo sistema no lineal de identificación de objetos invariante a traslación, rotación, escala e iluminación. Que sea capaz de identificar imágenes inmersas en ruido.
4. Generar firmas vectoriales invariantes a rotación y escala a partir de máscaras binarias de anillos concéntricos.
5. Implementar una correlación no lineal adaptativa para determinar el grado de similitud entre firmas vectoriales.
6. Incorporar al sistema diseñado la identificación de imágenes parcialmente ocultas o incompletas.
7. Evaluar la efectividad del algoritmo al utilizar las imágenes de prueba inmersas en diferentes tipos de ruido.

1.5 Estructura de la tesis

La tesis comprende cinco capítulos y está estructurada de la siguiente manera:

El capítulo 1 inicia con una introducción general a esta tesis y con los principales antecedentes en temas de máscaras binarias y de filtros de correlación basados en la transformada de escala. También se presentan los objetivos y la estructura de las tesis.

El marco teórico se presenta en el capítulo 2, en esta sección se describen las principales propiedades y características de las transformadas integrales bidimensionales que son utilizadas en los algoritmos desarrollados para lograr los objetivos de este trabajo. Además se presentan diferentes tipos de filtros de correlación y las herramientas de evaluación del desempeño de los filtros de correlación utilizados en el reconocimiento de imágenes.

En el capítulo 3 se encuentra la descripción de las aportaciones y de los algoritmos desarrollados para lograr una metodología basada en la transformada de escala y técnicas de correlación no lineal, con invariancia a distorsiones causadas por transformaciones geométricas en las imágenes digitales, debidas a cambios de posición, rotación y escala, que además es robusta ante cambios de iluminación y bajo la presencia de diferentes tipos de ruido.

Los resultados obtenidos están reportados en el capítulo 4. En esta sección se muestra la efectividad de la metodología utilizando 21 imágenes de diatomeas fósiles y 30 especies de fitoplancton en escala de grises. Cada una de estas imágenes es sometida a diferentes tipos de iluminación no uniforme, a diferentes tipos de ruido y a distorsiones geométricas por cambios de rotación y escala. Se realizan comparaciones de resultados obtenidos entre la correlación no lineal y la propuesta de una nueva correlación no lineal adaptativa que utiliza un índice invariante a posición, rotación y escala, que depende de una razón entre el área y el volumen de la imagen objetivo y la imagen problema. Por último, en el capítulo 5 se presentan las conclusiones.

Capítulo 2. Marco teórico

2.1 Transformada de Fourier

Las series de Fourier, permiten representar a una función periódica como la suma de componentes sinusoidales que tienen diferentes frecuencias. Cualquier función periódica $f(x)$ con un periodo T , se puede expresar mediante la serie compleja de Fourier (Hsu, 1998):

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{jnu_0x}, \quad (1)$$

donde $u_0 = 2\pi / T$ y

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{-jnu_0x} dx \quad (2)$$

Las condiciones, bajo las cuales es posible la representación en serie de Fourier de una función $f(x)$ dada son:

- 1) La función $f(x)$ tiene un número finito de discontinuidades en un periodo.
- 2) La función $f(x)$ tiene un número finito de máximos y mínimos en un periodo.
- 3) La integral del valor absoluto de $f(x)$ en un periodo es finita, es decir,

$$\frac{1}{T} \int_{-T/2}^{T/2} |f(x)| dx < \infty. \quad (3)$$

Las series de Fourier constituyen un poderoso instrumento en el tratamiento de diversos problemas que implican funciones periódicas. Sin embargo, debido a que muchos problemas prácticos no involucran funciones periódicas, es necesario utilizar una herramienta que incluya funciones no periódicas. La herramienta que nos proporciona una relación exacta entre los dominios de la frecuencia y del espacio para funciones periódicas y no periódicas, es la transformada de Fourier. Esta transformada

y sus propiedades se encuentran muy bien documentadas en los libros de Brigham (1988), Hsu (1998), Bracewell (2000) y Goodman (2005), y está definida mediante el siguiente par de integrales:

$$F(u) = \int_{-\infty}^{\infty} f(x) e^{-j2\pi ux} dx, \quad (4)$$

$$f(x) = \int_{-\infty}^{\infty} F(u) e^{j2\pi ux} du, \quad (5)$$

donde $j = \sqrt{-1}$, $f(x)$ representa una función continua con dominio en el espacio y $F(u)$ representa la transformada de Fourier de $f(x)$; x y u son variables continuas que representan espacio y frecuencia espacial, respectivamente.

La condición para que exista $F(u)$ generalmente está dada por

$$\int_{-\infty}^{\infty} |f(x)| dx < \infty. \quad (6)$$

En otros términos, la integral del valor absoluto de $f(x)$ debe ser finita. Esta es una condición suficiente pero no necesaria para la existencia de $F(u)$; funciones que no satisfacen la ecuación (6) pueden tener transformadas de Fourier. Ejemplos de este caso se presentan en funciones como, el escalón unitario y en las funciones periódicas, debido a que la integral del valor absoluto de estas es igual a infinito; pero su transformada de Fourier existe en el sentido de una función generalizada, tales como la función impulso unitario y sus derivadas (Hsu, 1998, p.110).

Para representar una imagen se utilizan funciones continuas de dos variables independientes $f(x, y)$ y el par de transformadas de Fourier está dado por,

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy, \quad (7)$$

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} du dv, \quad (8)$$

donde, (x, y) son variables continuas que representan el espacio, (u, v) son variables continuas que representan las frecuencias, $f(x, y)$ representa la imagen y $F(u, v)$ representa la transformada de Fourier de la imagen.

La transformada de Fourier nos permite una mayor manipulación de una señal debido a que nos da acceso a su contenido espectral. En este podemos seleccionar información específica, propia de la imagen que se analiza. Además nos facilita realizar operaciones de filtrado que atenúan o refuerzan algunas de las componentes senoidales asociadas a los atributos de nuestro interés, logrando diferentes efectos en el dominio espacial de la señal original.

Las ecuaciones anteriores requieren de funciones y variables continuas, situación que no es conveniente si se requiere procesar la información de manera digital. Por tal motivo, en la práctica es común utilizar aproximaciones como la transformada discreta de Fourier (*Discrete Fourier Transform*, DFT) o la transformada rápida de Fourier (*Fast Fourier Transform*, FFT), las cuales se revisan a continuación.

2.2 Transformada discreta de Fourier

La transformada discreta de Fourier o DFT representa la versión discreta de la transformada de Fourier. La DFT relaciona N muestras de una señal en el espacio con N muestras de su espectro. El par de transformadas discretas de Fourier asociadas a una sucesión finita de valores discretos de una función se expresa de la siguiente forma (González y Woods, 2002)

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j2\pi ux/N}, \quad (9)$$

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

donde x es una variable discreta espacial y u es una variable discreta de frecuencia espacial, para valores de $x=0,1,2,\dots,N-1$, $u=0,1,2,\dots,N-1$, donde $f(x)$ representa una función en el dominio espacial de tamaño N , $F(u)$ es la transformada de Fourier

discreta de la función,. El par de transformadas discretas para el caso bidimensional están dadas por (González y Woods, 2002)

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

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M+vy/N)}, \quad (12)$$

donde (x, y) son las variables espaciales discretas y (u, v) son las variables de frecuencia espacial discretas, para valores de $x=0, 1, 2, \dots, M-1$, $y=0, 1, 2, \dots, N-1$, $u=0, 1, 2, \dots, M-1$, $v=0, 1, 2, \dots, N-1$, donde $f(x, y)$ representa a una imagen digital de entrada en el dominio espacial de tamaño $M \times N$ y $F(u, v)$ es la transformada de Fourier discreta de la imagen.

2.3 Propiedades de la transformada discreta de Fourier 2D

En esta sección, se presentan las propiedades de la transformada discreta de Fourier en dos dimensiones, que se relacionan con el desarrollo de este trabajo (González y Woods, 2002, p. 194).

2.3.1 Propiedad de desplazamiento

La multiplicación por un exponencial en el dominio espacial corresponde a traslación en frecuencia,

$$f(x, y) e^{j2\pi(u_0x/M+v_0y/N)} \Leftrightarrow F(u-u_0, v-v_0). \quad (13)$$

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

$$f(x-x_0, y-y_0) \Leftrightarrow F(u, v) e^{-j2\pi(x_0u/M+y_0v/N)}. \quad (14)$$

2.3.2 Propiedad de linealidad

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

$$\mathfrak{F}[f_1(x, y) + f_2(x, y)] = \mathfrak{F}[f_1(x, y)] + \mathfrak{F}[f_2(x, y)], \quad (15)$$

donde \mathfrak{F} representa la transformada de Fourier. La transformada de Fourier es distributiva respecto a la suma pero no al producto,

$$\mathfrak{F}[f_1(x, y) \cdot f_2(x, y)] \neq \mathfrak{F}[f_1(x, y)] \cdot \mathfrak{F}[f_2(x, y)]. \quad (16)$$

2.3.3 Propiedad de escalamiento

Para dos escalares a y b se tiene,

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

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

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

2.3.4 Propiedad de rotación

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

$$f(r, \theta + \theta_0) \Leftrightarrow F(\omega, \phi + \theta_0), \quad (19)$$

al girar la función $f(x, y)$ a un ángulo θ_0 , se produce un giro con el mismo ángulo en $F(u, v)$. Esta propiedad también se da a la inversa, es decir, si la transformada se rota en un determinado ángulo, la transformada inversa también se verá rotada ese mismo ángulo.

2.3.5 Propiedad de periodicidad

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

$$F(u, v) = F(u + M, v) = F(u, v + N) = F(u + M, v + N), \quad (20)$$

$$f(x, y) = f(x + M, y) = f(x, y + N) = f(x + M, y + N). \quad (21)$$

2.3.6 Propiedad de simetría conjugada

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

$$F(u, v) = F^*(-u, -v). \quad (22)$$

2.3.7 Propiedad de separabilidad

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

$$F(u, v) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) e^{-j2\pi ux/M} \frac{1}{N} \sum_{y=0}^{N-1} f(y) e^{-j2\pi vy/N}. \quad (23)$$

2.3.8 Teorema de convolución

La convolución discreta de dos funciones $f(x, y)$ y $h(x, y)$ de tamaño $M \times N$, está definida por la ecuación,

$$f(x, y) * h(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) h(x-m, y-n). \quad (24)$$

La convolución de dos funciones en el dominio espacial equivale a multiplicar sus transformadas individuales, si $\mathfrak{F}[f(x, y)] = F(u, v)$ y $\mathfrak{F}[h(x, y)] = H(u, v)$, entonces

$$f(x, y) * h(x, y) \Leftrightarrow F(u, v) H(u, v), \quad (25)$$

$$f(x, y) h(x, y) \Leftrightarrow F(u, v) * H(u, v). \quad (26)$$

2.3.9 Teorema de correlación

La correlación discreta de dos funciones $f(x,y)$ y $h(x,y)$ de tamaño $M \times N$, está definida por la ecuación,

$$f(x,y) \circ h(x,y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f^*(m,n) h(x+m, y+n), \quad (27)$$

donde f^* representa el complejo conjugado de f .

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

$$f(x,y) \circ h(x,y) \Leftrightarrow F^*(u,v) H(u,v), \quad (28)$$

$$f^*(x,y) h(x,y) \Leftrightarrow F(u,v) \circ H(u,v). \quad (29)$$

2.4 Transformada rápida de Fourier

La transformada rápida de Fourier o FFT (*Fast Fourier Transform*), es un algoritmo que calcula con rapidez la transformada discreta de Fourier (DFT). El número de operaciones aritméticas requeridas para realizar el cálculo con la DFT es del orden de N^2 para el caso de una dimensión y de $(MN)^2$ para el caso de dos dimensiones (González y Woods, 2007, p. 321). Sin embargo, debido a las propiedades de la transformada de Fourier, el algoritmo de la transformada rápida de Fourier desarrollado por Cooley y Tukey, 1965, hace posible reducir el número de operaciones al orden de $N \log_2 N$ para el caso de una dimensión y al orden $MN \log_2(MN)$ para el caso de dos dimensiones. En el desarrollo de esta tesis toda transformada de Fourier será calculada digitalmente, ya que gracias al avance de la tecnología, cualquier computadora comercial puede realizar una transformada rápida de Fourier en una fracción de segundo, dando una alternativa digital al procesamiento óptico, aportando portabilidad, flexibilidad de operación y reducción de costos.

2.5 Transformada de Mellin

La transformada de Mellin, es una transformada integral que está fuertemente relacionada con la transformada de Fourier. La transformada de Mellin de una función $f(x)$ está definida como (Brecewell, 2000),

$$F_M(s) = \int_0^{\infty} f(x)x^{s-1}dx, \quad (30)$$

donde s es una variable compleja. Podemos definir la transformada de Mellin sobre el eje imaginario $s = j\omega$, en este caso

$$F_M(\omega) = \int_0^{\infty} f(x)x^{j\omega-1}dx. \quad (31)$$

La propiedad más útil de la transformada de Mellin es que cambios de escala en el dominio espacial se trasladan a cambios de fase en el dominio de Mellin. Esto es análogo a como la traslación de una señal en el dominio espacial se traslada a un cambio de fase en el dominio de Fourier. La invariancia a escala se puede probar de la siguiente manera.

Si expresamos la transformada de Mellin de la función $f(\alpha x)$ como

$$F_M^\alpha(\omega) = \int_0^{\infty} f(\alpha x)x^{j\omega-1}dx, \quad (32)$$

y sustituimos $\eta = \alpha x$, o $x = \eta/\alpha$, la transformada de Mellin de la señal escalada puede entonces ser reescrita como

$$F_M^\alpha(\omega) = \int_0^{\infty} f(\eta)\left(\frac{\eta}{\alpha}\right)^{j\omega-1} \frac{d\eta}{\alpha}, \quad (33)$$

y simplificando

$$F_M^\alpha(\omega) = \alpha^{-j\omega} \int_0^{\infty} f(\eta)\eta^{j\omega-1}d\eta, \quad (34)$$

y sustituyendo la definición dada por la ecuación (31)

$$F_M^\alpha(\omega) = \alpha^{-j\omega} F_M(\omega) = e^{-j\omega \ln \alpha} F_M(\omega). \quad (35)$$

Esto significa que cambios de escala se manifiestan en términos solo de cambios de fase en el dominio de Mellin. Por lo tanto, la magnitud de la transformada de Mellin es invariante a cambios de escala:

$$|F_M(\omega)| = |F_M^\alpha(\omega)|, \quad (36)$$

es decir, los módulos de las transformadas de Mellin de dos funciones que difieren solo en tamaño son iguales. En cambio, el módulo de la transformada de Fourier es invariante a corrimiento.

Para obtener la transformada de Mellin de una imagen debemos aplicar la transformada bidimensional de Mellin (Casasent y Psaltis, 1977)

$$F_M(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) x^{-ju-1} y^{-jv-1} dx dy. \quad (37)$$

2.6 Transformada de escala

El concepto de transformada de escala fue propuesto por Cohen (1993). Podemos considerar la escala como un atributo físico de una señal. Por ejemplo, en el caso de la transformada de Fourier, la magnitud asociada es la frecuencia, de modo análogo, se desarrolla la representación de la escala donde la magnitud principal a representar es, como indica su nombre, la escala. El concepto de escala aparece en numerosos campos de investigación tales como: el procesado de señales multidimensionales y filtrado de imágenes, Cristobal et al., (1998) eliminaron el ruido en imágenes aplicando filtros en el dominio de la escala, también compararon la efectividad de diferentes filtros, aplicándolos en el dominio de Fourier y en el dominio de la transformada de escala. El análisis del habla, Mertis y Rademacher, (2005) haciendo uso de la transformada de escala, lograron una metodología de reconocimiento automático de voz, invariante a la

longitud del tracto vocal, lo que les permitió identificar palabras sin importar el sexo o edad de la persona que habla. La identificación de bacterias, Mouriño Pérez, (1999) identifica la bacteria responsable de la enfermedad del cólera (*Vibrio cholerae* O1), utilizando la transformada de escala para lograr la invariancia al tamaño de la bacteria. Análisis de vibraciones mecánicas, Combet et al., (2009) logran estimar la velocidad relativa de una fluctuación en una señal de vibración mecánica, analizando la señal en el dominio de la transformada de escala. Procesamiento de audio, debido a que una misma melodía puede ser tocada a diferente velocidad, Holzapfel y Stylianou, (2011) usan la transformada de escala para identificar señales de música, ya que representan la velocidad como un factor de escala. Y otras áreas como en el control de temperatura (Harley y Moura, 2012) y en la detección de fallas en motores de inducción (Antonino Daviu, et al., 2013).

Si tomamos c como la variable de escala, entonces la transformada de escala y su inversa están dadas por (Cohen, 1993)

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(x) \frac{e^{-jc \ln x}}{\sqrt{x}} dx, \quad (38)$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} D(c) \frac{e^{jc \ln x}}{\sqrt{x}} dc; x \geq 0. \quad (39)$$

La ecuación (38) también la podemos escribir como

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(x) x^{-jc-1/2} dx, \quad (40)$$

que es equivalente a la transformada de Mellin con el argumento complejo $s = -jc + 1/2$.

El par de transformadas de escala en dos dimensiones está dado por (Cristobal, et al., 1998),

$$D(c_x, c_y) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \int_0^{\infty} f(x, y) \frac{e^{-jc_x \ln x - jc_y \ln y}}{\sqrt{xy}} dx dy, \quad (41)$$

$$f(x, y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(c_x, c_y) \frac{e^{jc_x \ln x + jc_y \ln y}}{\sqrt{xy}} dc_x dc_y. \quad (42)$$

Para obtener la invariancia a cambios de escala de un objeto dentro de una imagen, la metodología desarrollada en este trabajo utiliza la ecuación (41), para calcular la transformada de escala en dos dimensiones. La aplicación práctica de esta transformada se realizará vía la transformada rápida de Fourier.

Si en la ecuación (41), se realiza un cambio de variable $x = \exp(x')$, $y = \exp(y')$, la transformada de escala se relaciona con la transformada de Fourier, ya que después del cambio de variable se obtiene,

$$D(c_x, c_y) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \int_0^{\infty} f(e^{x'}, e^{y'}) e^{(x'+y')/2} e^{-jc_x x' - jc_y y'} dx' dy', \quad (43)$$

así

$$D(c_x, c_y) = \mathfrak{F} \left(f(e^{x'}, e^{y'}) e^{(x'+y')/2} \right). \quad (44)$$

Sea la función

$$g(x, y) = f(x, y) \frac{1}{\sqrt{xy}}, \quad (45)$$

podemos calcular su transformada de escala,

$$D(c_x, c_y) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \int_0^{\infty} g(x, y) \frac{e^{-jc_x \ln x - jc_y \ln y}}{\sqrt{xy}} dx dy, \quad (46)$$

sustituyendo la ecuación (45) en la ecuación (46) y haciendo el cambio de variable $x = \exp(u)$, $y = \exp(v)$, obtenemos:

$$D(c_x, c_y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(e^u, e^v) e^{-jc_x u - jc_y v} du dv, \quad (47)$$

$$F_M(u, v) = \mathfrak{F} \left(f(e^u, e^v) \right), \quad (48)$$

por lo tanto, las ecuaciones que relacionan la transformada de Mellin con la transformada de escala en dos dimensiones son

$$D[g(x, y)] = D \left[f(x, y) \frac{1}{\sqrt{xy}} \right] = F_M[f(x, y)], \quad (49)$$

$$F_M[g(x, y)] = F_M[\sqrt{xy} f(x, y)] = D[f(x, y)]. \quad (50)$$

2.7 Filtros de correlación

El filtrado espacial consiste en insertar máscaras o filtros en el plano de Fourier para seleccionar o evitar la llegada de frecuencias espaciales al plano imagen. El primero en aplicar este tipo de procesamiento óptico fue Ernst Abbe, quien mejoró la resolución de un microscopio aplicando un filtro espacial, debido a que concluyó que la formación de la imagen en un microscopio depende de los rayos difractados, pues si estos se suprimen o se alteran, la proyección aumentada del objeto se modifica o simplemente no se forma (Abbe, 1873). Como resultado de esto, el procesamiento óptico se basa en las propiedades de difracción de la luz, la cual se obtiene utilizando lentes para realizar la transformada de Fourier.

2.7.1. Filtro de acoplamiento clásico

El desarrollo del filtro de acoplamiento clásico o CMF (*Classic Matched Filter*), propuesto por VanderLugt (1964), inició un interés por desarrollar nuevos filtros cada vez más eficientes, logrando así un gran avance en los métodos de correlación óptica. Algunas limitaciones de este filtro son producir: picos de correlación anchos, lóbulos laterales grandes, baja discriminación relativa y poca eficiencia de luz en su implementación óptica.

La respuesta en frecuencia del filtro de acoplamiento clásico está dada por

$$H(u, v) = \beta \frac{T^*(u, v)}{P_n(u, v)} \equiv H_{CMF}(u, v), \quad (51)$$

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

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

$$H_{CMF}(u, v) = T^*(u, v). \quad (52)$$

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

$$H_{CMF}(u, v) = |H(u, v)| e^{j\phi(u, v)}, \quad (53)$$

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

2.7.2 Filtro solo de fase

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

$$H_{POF}(u, v) = e^{j\phi(u, v)}, \quad (54)$$

es decir, solo toma valores en el círculo unitario del plano complejo. Con el POF (*Phase Only Filter*) se obtienen picos de correlación más estrechos y es más selectivo que el CMF (*Classic Matched Filter*).

2.7.3 Filtro no lineal

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

$$H_{NLF}(u, v) = |T(u, v)|^k e^{j\phi(u, v)}, \quad (55)$$

donde $|T(u, v)|$ es el módulo de la transformada de Fourier del objeto que deseamos reconocer, $\phi(u, v)$ es la fase de la transformada de Fourier, k es el factor de no linealidad que toma valores $0 < k < 1$. Al variar los valores de k en este intervalo, podemos manipular las capacidades discriminantes del filtro y determinar el mejor valor k . Cuando el operador no lineal k modifica la transformada de Fourier, tanto de la imagen problema como la de la imagen objetivo, consideramos que tenemos un

procesador no lineal. El filtro no lineal tiene ventajas comparado con los filtros clásicos acoplados, los filtros solo de fase y otros filtros lineales, debido a su gran capacidad para discriminar objetos, a que su máximo pico de correlación está bien localizado y a que el plano de salida es menos ruidoso.

2.8 Métricas de desempeño

Una métrica de desempeño para filtros de correlación cuantifica qué tan bien puede definirse ese pico de correlación (Kumar y Hassebrook, 1990). En este trabajo se utiliza: la capacidad de discriminación y la energía pico de correlación.

2.8.1 Capacidad de discriminación

El coeficiente de discriminación (CD) evalúa la habilidad de un filtro para detectar un objeto deseado y rechazar cualquier otro objeto; por ejemplo, el fondo o el ruido. Este coeficiente está dado por (Vijaya y Hassebrook, 1990),

$$CD = 1 - \frac{|C_{I-R}|^2}{|C_{I-IR}|^2}, \quad (56)$$

Donde C_{I-R} es el valor máximo del plano de correlación entre la imagen del objeto y la imagen de fondo (ruido) y C_{I-IR} es el valor máximo del plano de correlación entre la imagen del objeto y la imagen del objeto afectada por ruido. El valor máximo de esta medida es la unidad y se presenta en imágenes sin ruido, mientras que valores negativos ocurren en presencia de imágenes que se encuentran muy afectadas por ruido, indicando que el objeto no puede ser reconocido por el filtro.

2.8.2 Energía pico de correlación (PCE)

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

$$PCE = \frac{|E\{C(0,0)\}|^2}{E\{|C(x,y)|^2\}}, \quad (57)$$

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

2.9 Ruido en la imagen

Las imágenes se obtienen mediante la captura de una escena del mundo real y es de esperar que estas imágenes estén distorsionadas por la acción de al menos una fuente de ruido. El ruido es una mezcla de señales aleatorias no deseadas, que están presentes en la imagen pero que son ajenas a la imagen original. Es un proceso estadístico con datos sin significado que de alguna manera indeseada se transmiten a la imagen deseada. El ruido se genera en el momento de la captura de la imagen de un objeto, debido a variaciones del medio ambiente y a variaciones en la sensibilidad del sensor detector que convierte una imagen óptica en una señal eléctrica. En este trabajo, se agrega ruido a las imágenes intencionalmente, con el fin de evaluar el desempeño de los filtros.

2.9.1 Ruido gaussiano

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

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{\sigma^2}}, \quad (58)$$

donde x es la variable aleatoria y representa el nivel de gris, μ es el valor promedio de x y σ^2 es la varianza de x . Cuanto mayor es el valor de la varianza, más se dispersan los datos en torno a la media. Un valor pequeño de este parámetro indica una gran probabilidad de obtener datos cercanos al valor medio de la distribución.

2.9.2 Ruido Impulsivo

En el ruido impulsivo (ruido sal y pimienta) el valor que toma el píxel no tiene relación con el valor de la imagen original, sino que toma valores muy altos o muy bajos. “Toma el valor máximo (sal) o el mínimo (pimienta); para una imagen de 256 niveles de gris el valor típico de la pimienta es 0 y de la sal 255. Defectos que contribuyen a este tipo de ruido son las manchas de polvo dentro de las componentes ópticas de la cámara, o bien un sensor CCD defectuoso, que realizará una captura errónea” (González y Woods, 2002, p. 226).

La medida del ruido está determinada por un índice de densidad “d”, que nos proporciona la información de píxeles afectados de la imagen que va de 0 a 1. Cuando la densidad es 1, la totalidad de los píxeles están dañados. Una densidad de ruido impulsivo de 0.5 equivale al caso en donde la mitad de los píxeles están siendo afectados.

Capítulo 3. Metodología

En este capítulo se muestran los algoritmos desarrollados y la metodología basada en la transformada de escala y en técnicas de correlación no lineal, con invariancia a distorsiones causadas por transformaciones geométricas en imágenes digitales, debidas a cambios de posición, rotación y escalamiento.

Esta metodología utiliza máscaras de anillos concéntricos construidas a partir de las pendientes contenidas en el transecto central del módulo de la transformada de escala de la imagen problema. Al realizar una multiplicación punto por punto de la máscara de anillos y el módulo de la transformada de escala de la imagen problema, se estará obteniendo una serie de anillos que ahora contienen muestras radiales del módulo de la transformada de escala. El contenido en estos anillos, tiene la propiedad de ser invariante a posición, rotación y escala. Después de realizar la suma de los valores en cada anillo, se obtiene una firma vectorial unidimensional capaz de representar a la imagen que se desea identificar, a pesar de estar desplazada, rotada y escalada.

Como método de comparación entre firmas vectoriales se utiliza una correlación no lineal (ley k). Hay casos en donde las firmas que representan a objetos que no corresponden a la imagen objetivo, resultan ser muy similares a la firma de la imagen objetivo, esto provoca que el valor del máximo pico de correlación que resulta de la correlación no lineal (ley k), no logre la identificación automática del objeto en la imagen. Una aportación importante se logra modificando esta correlación para dar lugar a una correlación no lineal adaptativa a cambios de rotación y escala, que incrementa la confiabilidad en la identificación de las imágenes y ayuda en la identificación automática de la imagen objetivo.

3.1. Máscara binaria de anillos concéntricos

Debido a que la versión separable de la transformada de escala en 2D representada por las ecuaciones (41) y (42) carece de invariancia a rotación, se ha diseñado una máscara binaria de anillos concéntricos creada a partir del módulo de la transformada

de escala 2D en su forma separable, creando una alternativa de invariancia a rotación y escala, además de la que ofrece la versión no separable de la transformada de escala, representada en coordenadas log-polares (ρ, θ) , donde $\rho = \ln r$, la cual está expresada como (Cristóbal y Cohen, 1998)

$$D(c_\rho, c_\theta) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \int_0^{2\pi} e^{\rho/2} f(\rho, \theta) e^{-j(\rho c_\rho + \theta c_\theta)} d\rho d\theta. \quad (59)$$

La figura 4 muestra la implementación práctica de la versión separable de la transformada de escala en 2D dada por la ecuación (41). Esto se logra realizando un mapeo exponencial de una imagen de entrada y multiplicando por el factor $\exp[(x'+y')/2]$. Al hacer esto tenemos una versión modificada de la imagen que nos sirve para calcular el módulo de la transformada de escala de una manera práctica, a través del cálculo de su transformada rápida de Fourier.

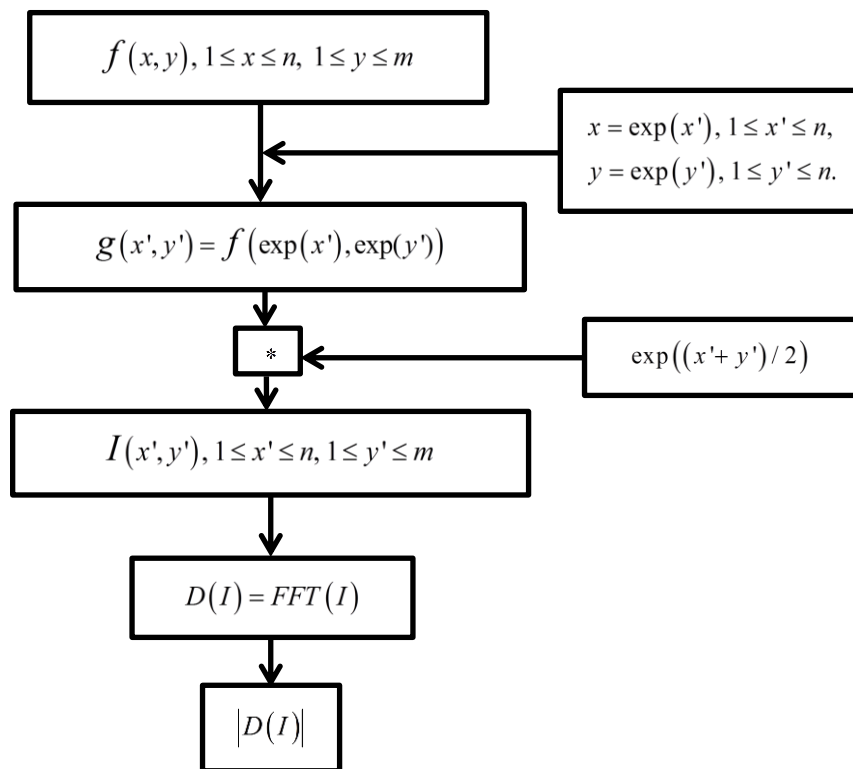


Figura 4. Implementación de la transformada de escala separable en 2D vía la transformada rápida de Fourier.

El diagrama de flujo de la figura 5 muestra los pasos para obtener la máscara, a partir de una imagen objetivo (figura 6a). Este procedimiento se inicia obteniendo de una manera práctica el módulo de la transformada de escala de la imagen objetivo

$|D_{IO}(c_x, c_y)|$ (figura 6b), del cual se selecciona el perfil central $y(c_y)=|D_{IO}(C, c_y)|$ mostrado en la figura 6c. El perfil central se digitaliza a partir del valor de la pendiente que se obtiene con

$$m = \frac{y(c_y) - y(c_y + 1)}{c_y - (c_y + 1)}, \quad c_y = 1, \dots, C-1, \quad (60)$$

en donde m representa el valor de la pendiente en un punto c_y del perfil central $y(c_y)$. En el paso 2 de la figura 5, la función binaria $z(c_y)$ es calculada como

$$z(c_y) = \begin{cases} 1, & m < 0 \\ 0, & m \geq 0 \end{cases}, \quad c_y = 1, \dots, C-1. \quad (61)$$

Después en el paso 3, se toma $c_y = C$ como eje de rotación, la función de $z(c_y)$ se rota 360° para obtener cilindros concéntricos con altura igual a uno, de diferentes grosores y centrados en el píxel (C, C) . Finalmente en paso 4, al tomar una sección transversal con respecto al eje de simetría, se genera la máscara binaria correspondiente a la imagen objetivo. En la figura 6d podemos apreciar un ejemplo de la digitalización del perfil central de la transformada de escala y en la figura 7 se muestra la máscara binaria resultante, que fue creada girando la función digital $z(c_y)$.

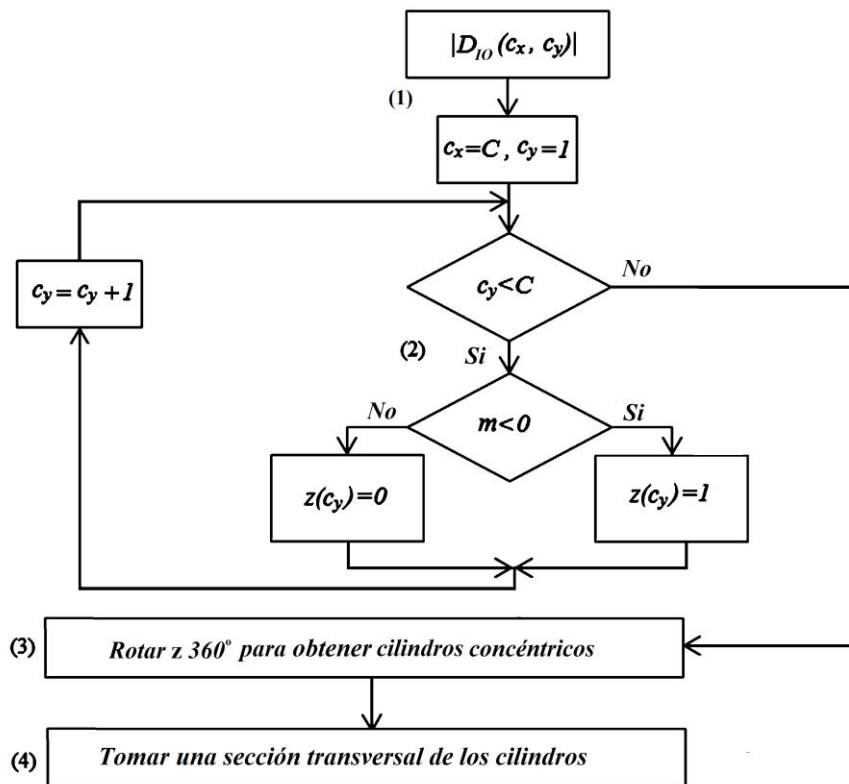


Figura 5. Procedimiento para obtener la máscara de anillos concéntricos.

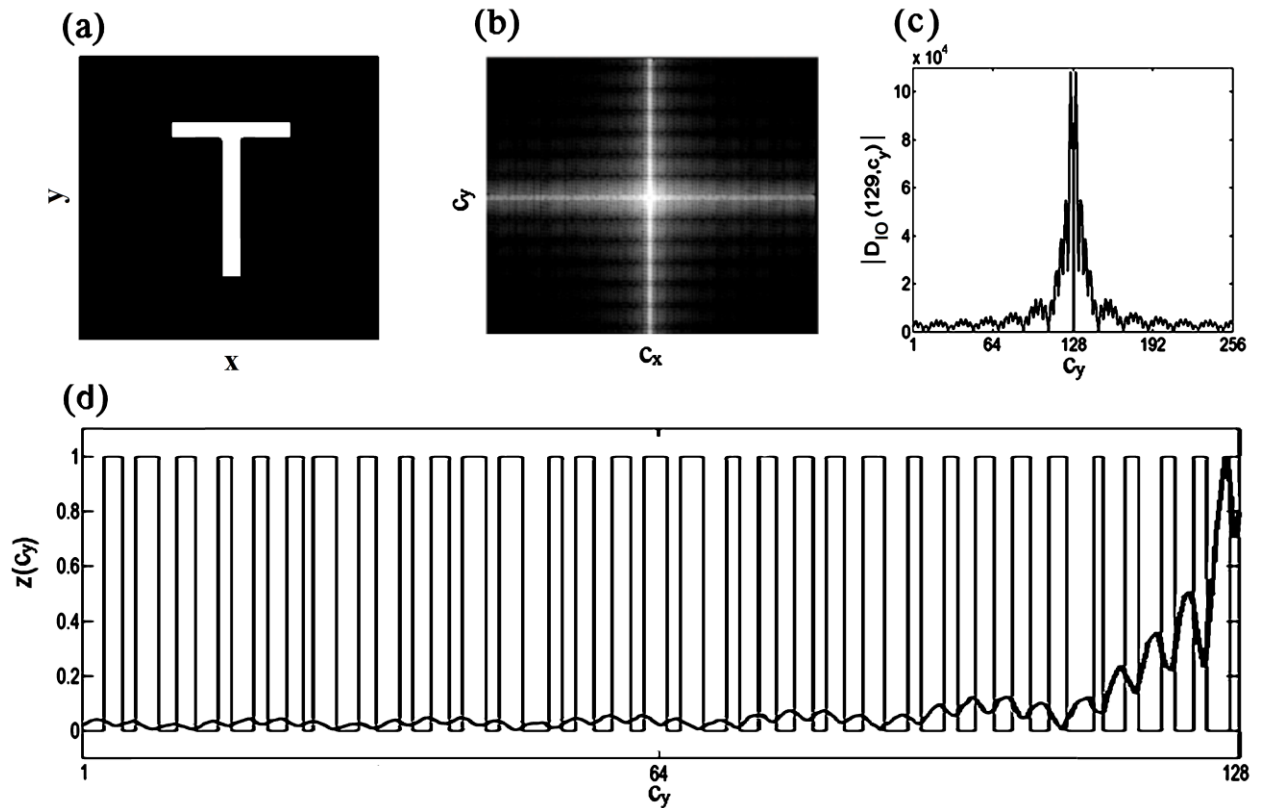


Figura 6. Digitalización de la transformada de escala. (a) Imagen de la “letra T Arial”. (b) Módulo de la versión separable de la transformada de escala en 2D correspondiente a la “letra T Arial”. (c) Perfil $|D_{10}(c_x, c_y)|$. (d) Digitalización a partir de las pendientes de $z(c_y)$ donde $c_y=1, \dots, 128$.

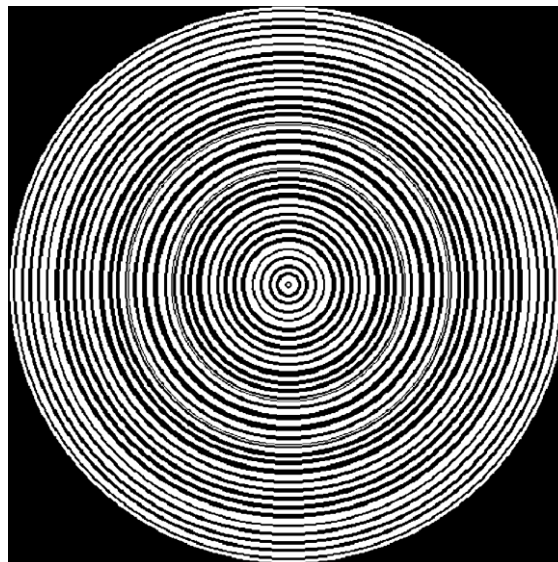


Figura 7. Máscara binaria de anillos para la “letra T Arial”.

3.2. Firma vectorial de la imagen

La firma vectorial invariante es generada a partir de las muestras recolectadas por la máscara binaria en el dominio del plano de escala. Este proceso se presenta en la figura 8, la cual muestra la firma obtenida cuando el objetivo es una letra O Arial (figura 8a) y cuando es una letra T Arial (figura 8b). En el paso 2 se muestra el módulo de la transformada de escala y en el paso 3 tenemos la máscara binaria correspondiente a la imagen objetivo. Al realizar una multiplicación punto por punto del módulo de la transformada de escala con la máscara binaria se obtendrán anillos concéntricos con muestras radiales del plano de escala de la imagen. El último paso para generar la firma vectorial consiste en sumar las cantidades de cada anillo. A cada anillo se le asigna un índice de anillo, donde el primer índice corresponde al anillo más interno (paso 5). En la figura 8 se puede apreciar que para la imagen correspondiente a la letra O Arial se obtienen 38 anillos, mientras que para la letra T Arial se obtiene una firma proveniente de una máscara con 36 anillos.

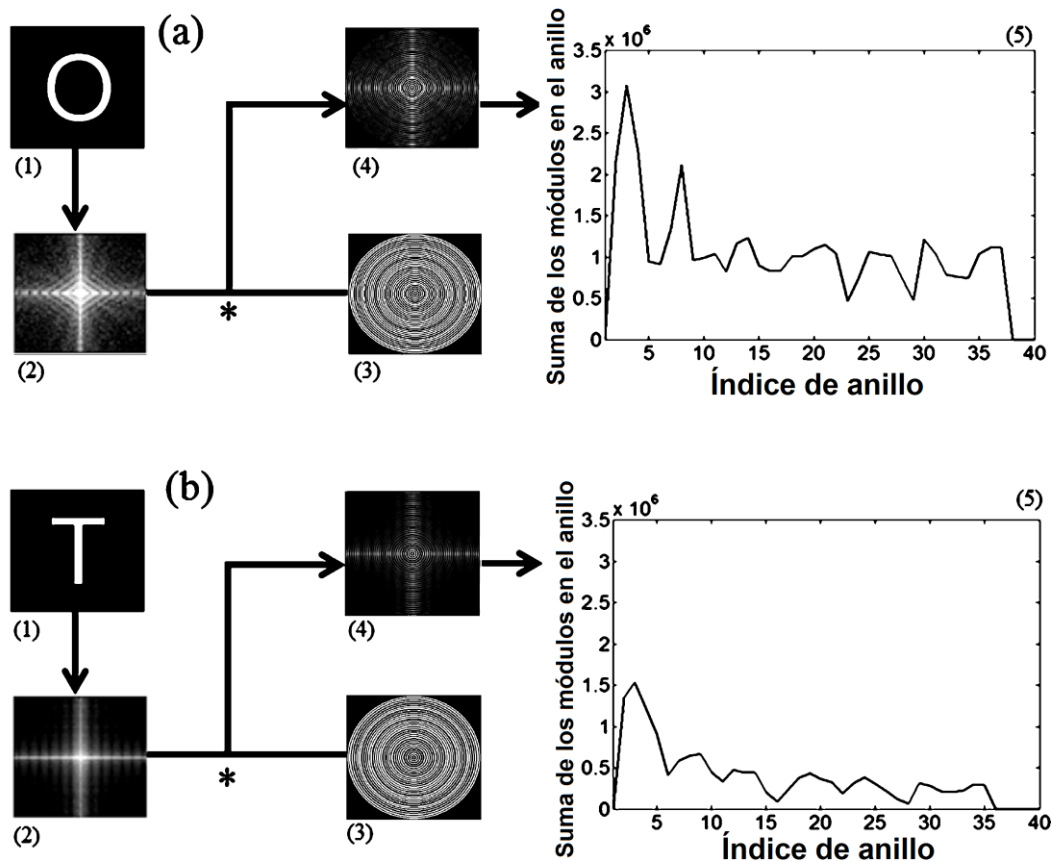


Figura 8. Procedimiento para obtener la firma vectorial de la imagen. (a) Firma vectorial de la “letra O Arial”. (b) Firma vectorial de la “letra T Arial”.

Cada imagen objetivo forma su firma particular con invariancia a posición, rotación y escala. Las imágenes problema al tener una escala y rotación aleatoria no formarán firmas idénticas a la firma de la imagen objetivo, debido a las distorsiones geométricas provocadas por los cambios en rotación y en escala. La figura 9 muestra como las imágenes problema de un mismo tipo de letra, que tienen diferentes rotaciones desde 0 hasta 180° y diferentes escalamientos desde 80% hasta 120% generan diferentes variaciones con respecto a la firma de la imagen objetivo, que es una imagen sin rotación y con escala original del 100% . Se puede apreciar que la rotación de la letra T provoca mayores distorsiones a la firma debido a la distorsión con efecto diente de sierra que se induce al tener las líneas rectas que forman la letra T.

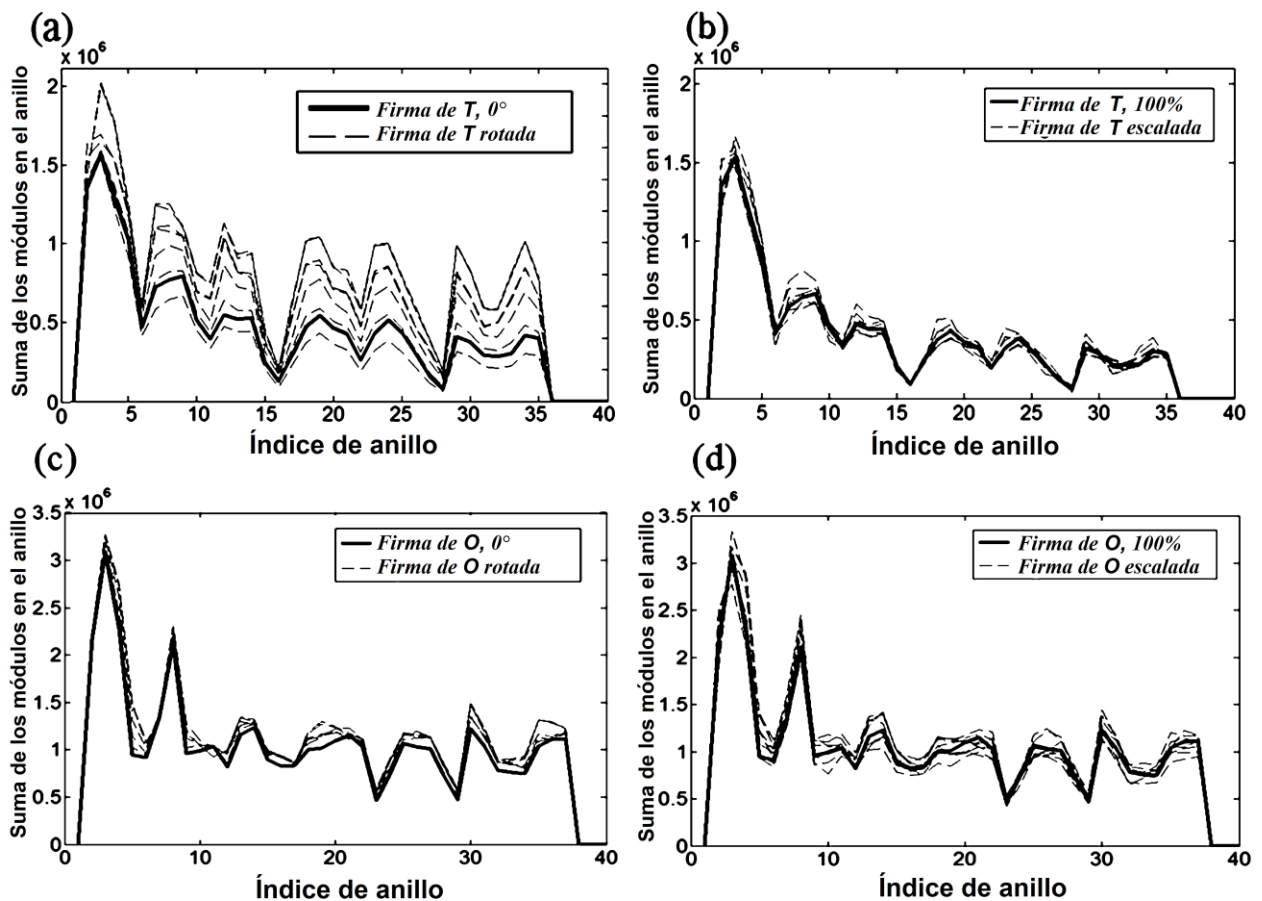


Figura 9. Efecto en las firmas vectoriales debido a distorsiones en la imagen causadas por cambios en rotación o escala. (a) Firma vectorial para las “letras T Arial” rotadas. (b) Firma vectorial para las “letras T Arial” escaladas. (c) Firma vectorial para las “letras O Arial” rotadas. (d) Firma vectorial para las “letras O Arial” escaladas.

3.3. Correlación no lineal adaptativa

Los filtros no lineales se caracterizan por tener un mejor desempeño en comparación de los filtros lineales en cuanto a capacidad de discriminación, picos de correlación más definidos y robustez en presencia de diferentes tipos de ruido (Javidi, 1990). Como método de comparación entre la firma objetivo (F_{IO}) y las diferentes firmas problema (F_{IP}), se ha seleccionado la correlación no lineal ley k

$$C_{NL}(F_{IO}, F_{IP}) = FFT^{-1} \left(|FFT(F_{IP})|^k \exp(i\phi_{IP}) |FFT(F_{IO})|^k \exp(-i\phi_{IO}) \right), \quad (62)$$

donde ϕ_{IO} y ϕ_{IP} son la fase de la transformada de Fourier de la firma de la imagen objetivo y de la firma de imagen problema respectivamente, k representa el factor de no linealidad que toma valores mayores a cero y menores a la unidad. Para obtener el valor óptimo de k , se realizó un experimento que se basa en calcular 779 correlaciones entre la firma vectorial generada con imagen de la diatomea fósil *Actinocyclus ingens* (diatomea A de la figura 12) como imagen objetivo y las 779 firmas vectoriales generadas, con versiones escaladas y rotadas de la diatomea fósil *Actinocyclus ingens*, como imágenes problema. Las imágenes problema son versiones escaladas desde un 80% hasta un 120% de su tamaño original, con incrementos del 1%. Además, cada versión escalada fue rotada desde 0 hasta 180° con incrementos de 10°. Utilizando la correlación no lineal ley k se asignan los valores de $k= 0.10, 0.20, \dots, 0.90$ y se realizaron 779 correlaciones por cada valor de k . De cada una de estas correlaciones se obtiene un valor de PCE utilizando la ecuación (57). El valor de la media de PCE más alto se obtuvo con $k= 0.10$, por lo tanto, el valor de no linealidad para el que se consigue una mejor capacidad de discriminación, en esta metodología es $k= 0.1$.

Al utilizar la misma máscara binaria, para generar las firmas correspondientes a todas las imágenes problema, las correlaciones entre la firma de la imagen objetivo y las diferentes firmas correspondientes a las imágenes problema, pueden resultar con máximos de correlación similares. Esto quiere decir que la correlación entre dos firmas será alta aun cuando la imagen problema no corresponde a la imagen objetivo. Para

solucionar este problema se ha propuesto una versión adaptativa de la correlación no lineal, ley k , que se logra agregando un factor R_z , cuyo valor depende de la razón entre dos índices invariantes a rotación. En la ecuación (63) se eligen como índices invariantes a rotación el valor promedio de la imagen objetivo y de la imagen problema. Esta razón entre promedios se calcula como lo muestra la siguiente ecuación (Solís-Ventura, 2012),

$$R_z = \begin{cases} \frac{\text{Área}(IP)}{\text{Área}(IO)}, & \text{si } \text{Área}(IP) \leq \text{Área}(IO), \\ \frac{\text{Área}(IO)}{\text{Área}(IP)}, & \text{si } \text{Área}(IO) < \text{Área}(IP). \end{cases} \quad (63)$$

Por lo tanto, el valor de R_z siempre será menor a la unidad y se acercará su valor a uno solo cuando la imagen problema sea casi idéntica a la imagen objetivo. Será igual a uno cuando la imagen objetivo sea exactamente igual a la imagen problema.

El factor R_z , toma valores $0 \leq R_z \leq 1$ y multiplica al factor de no linealidad k , correspondiente al módulo de la transformada de Fourier de la firma de la imagen problema de la ecuación (62), formando una versión adaptativa de esta, dando como resultado la siguiente correlación adaptativa,

$$C_{AD}(F_{IO}, F_{IP}) = FFT^{-1} \left(|FFT(F_{IP})|^{R_z \cdot k} \exp(i\phi_{IP}) |FFT(F_{IO})|^k \exp(-i\phi_{IO}) \right), \quad (64)$$

el valor R_z de la ecuación (63) solo considera cambios de rotación y no funciona ante posibles cambios de escala en las imágenes problema. Por este motivo se propuso una nueva variable que considera la razón entre dos índices invariantes a rotación y escala, los cuales se construyen relacionando el área y el volumen de las imágenes. El área corresponde al número de píxeles ocupados por el objeto a identificar en la imagen, sin incluir el fondo de la imagen y el volumen se calcula sumando el valor de todos los píxeles en la imagen. Los índices invariantes a rotación y escala, se calculan siguiendo el diagrama de la figura 10 (Fimbres-Castro et al., 2012).

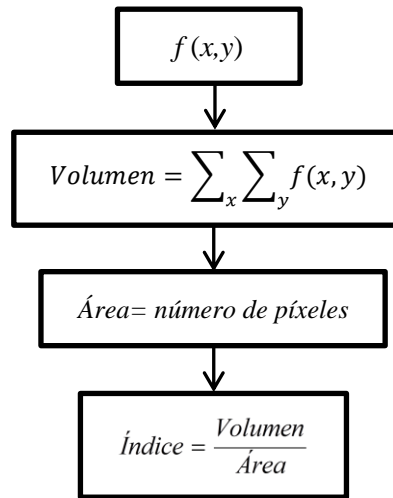


Figura 10. Procedimiento para obtener un índice invariante a rotación y escala, donde $f(x,y)$ representa a una imagen digital.

Siguiendo el diagrama de la figura 10 se generan dos índices invariantes a rotación y a escala a partir de la imagen objetivo y de la imagen problema, dados por,

$$\text{Índice}_{IO} = \frac{\text{Volumen}(IO)}{\text{Área}(IO)}, \quad (65)$$

$$\text{Índice}_{IP} = \frac{\text{Volumen}(IP)}{\text{Área}(IP)}. \quad (66)$$

De esta manera, la razón entre los dos índices invariantes a rotación y a escala de las ecuaciones (65) y (66), que dará valores a la variable R_z , seguirá cumpliendo con su objetivo inicial de agregar una mayor capacidad de discriminación a la correlación no lineal (ley k) de la ecuación (62). Sin embargo, ahora tendrá la capacidad de aceptar imágenes problema que además de estar rotadas, también tengan cambios en escala con respecto a la imagen objetivo. La razón entre estos índices está dada por

$$R_z = \begin{cases} \frac{\text{Índice}_{IP}}{\text{Índice}_{IO}}, & \text{si } \text{Índice}_{IP} \leq \text{Índice}_{IO}, \\ \frac{\text{Índice}_{IO}}{\text{Índice}_{IP}}, & \text{si } \text{Índice}_{IO} < \text{Índice}_{IP}. \end{cases} \quad (67)$$

Siguiendo la ecuación (67), la variable R_z tendrá valores menores a uno, y al igual que en la ecuación (63) solo será igual a uno cuando la imagen objetivo sea exactamente igual a la imagen problema. Esta variable agregará una propiedad adaptativa que hace más robusta ante cambios de rotación y escala, a la correlación no lineal adaptativa de la ecuación (64).

3.4 Metodología invariante a distorsiones geométricas causadas por cambios en posición, rotación y escala

La figura 11 muestra el diagrama de flujo que contiene los pasos a seguir para lograr una metodología invariante a distorsiones geométricas causadas por cambios en posición, rotación y escala. La metodología se basa en la forma separable de la transformada de escala en dos dimensiones, y se calcula aplicando la transformada rápida de Fourier, después de realizar un mapeo exponencial de la imagen. Con esta forma de calcular la transformada de escala, se pierde la propiedad de invariancia a posición. Es por esto, que para agregar invariancia a posición en los pasos 1 y 2 de la figura 11, se centran las imágenes objetivo y problema con respecto a sus centros de masa. La máscara binaria de anillos se genera a partir de la imagen objetivo, paso 3. Después, en los pasos 4 y 5 se obtienen los módulos de la transformada de escala para la imagen objetivo y la para la imagen problema (llamados $|D_{IO}|$ y $|D_{IP}|$, respectivamente), logrando así, la invariancia a cambios de posición y escala. Con el propósito de obtener una firma con invariancia a rotación, en los pasos 6 y 7, los módulos se filtran con la máscara binaria de anillos concéntricos, que se obtiene con el procedimiento mostrado en la figura 5. Esto se logra realizando una multiplicación punto por punto de los módulos de la transformada de escala con la máscara binaria, dando como resultado un conjunto de anillos concéntricos, que contienen muestras radiales del módulo de la transformada de escala. En los pasos 8 y 9, se realizan las sumas del contenido de cada uno de estos anillos, para la imagen objetivo y para la imagen problema, respectivamente. La suma del contenido en cada uno de los anillos, forma un número real, que se asigna a la función firma evaluada en el correspondiente índice de anillo. En el paso 10, se obtiene la firma vectorial para representar a la imagen objetivo. Se sigue el mismo procedimiento para obtener la firma vectorial que representa a la imagen problema con invariancia a posición, rotación y escala (paso 11). Así la imagen problema puede ser una versión escalada y rotada de la imagen objetivo o una imagen completamente diferente. Por lo tanto, la firma de la imagen problema, puede ser similar o diferente a la firma de la imagen objetivo.

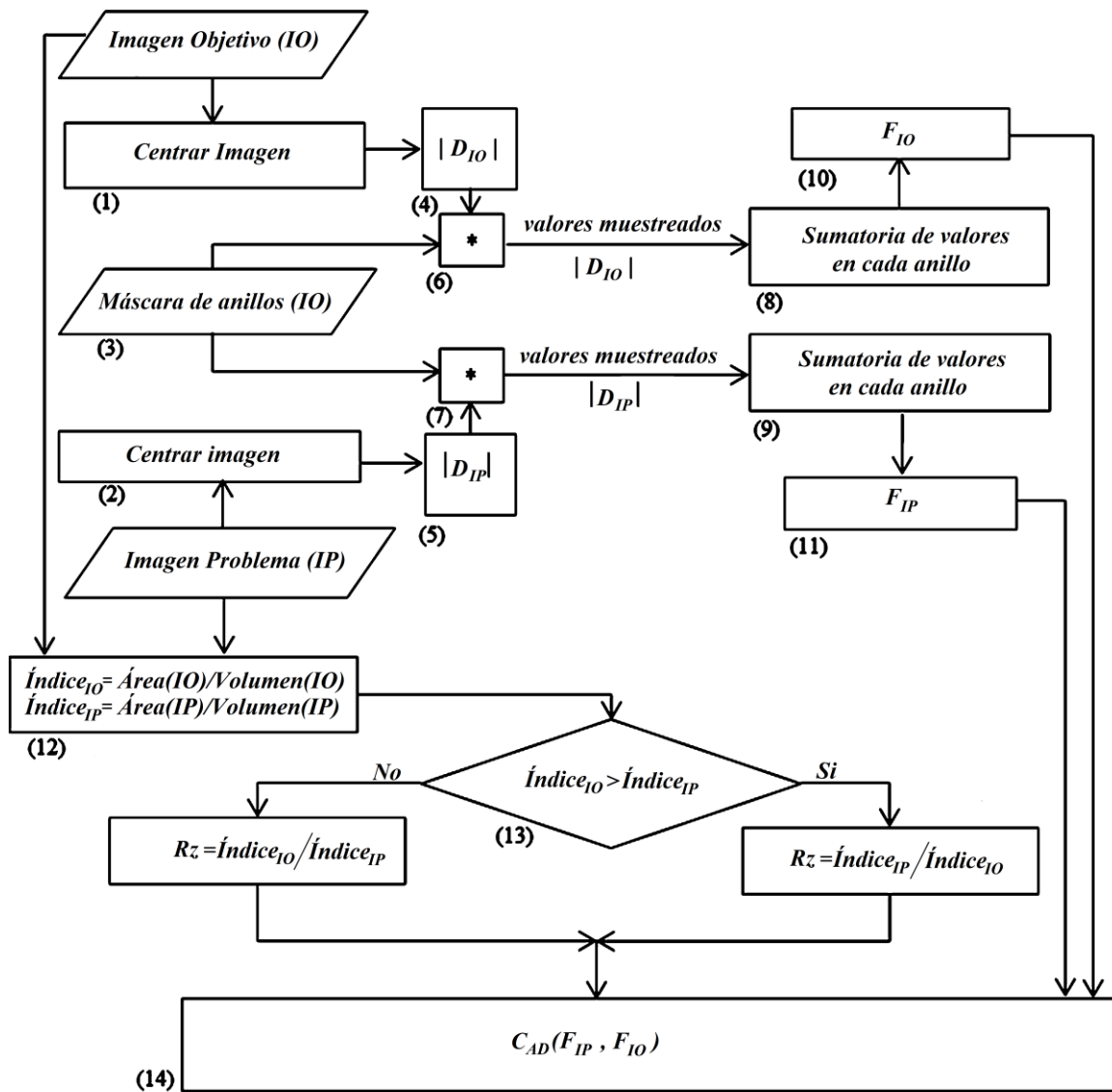


Figura 11. Procedimiento para lograr el reconocimiento de imágenes invariante a distorsiones geométricas causadas por cambios en posición, rotación y escala a partir de la transformada de escala en 2D y una correlación no lineal adaptativa.

En el paso 12 se calculan los dos índices invariantes a posición, rotación y escala. Estos son una razón entre área y volumen de la imagen objetivo y de la imagen problema, de tal manera que en el paso 13, aplicando la ecuación (67) se logran valores menores a la unidad para la variable R_z . Esta es utilizada para formar la versión adaptativa de la correlación no lineal (ley k) de la ecuación (64). La correlación no lineal adaptativa (C_{AD}) entre las firmas de la imagen objetivo y de la imagen problema se realiza en el paso 14.

Capítulo 4. Resultados

En este capítulo, además de mostrar la capacidad de la metodología propuesta, para la identificación de objetos con invariancia a diferentes distorsiones geométricas causadas por cambios de posición, rotación y escala, se añade la posibilidad de que las imágenes se vean afectadas por estar incompletas, por estar bajo diferentes condiciones de ruido y por tener diferentes tipos de iluminación no uniforme.

Debido a que las correlaciones no lineales superan a las correlaciones lineales en cuanto a la discriminación entre objetos (Solorza y Álvarez-Borrego, 2010), en este trabajo, solo se hacen comparaciones entre las firmas de las imágenes, a partir de correlaciones no lineales.

Como imágenes de prueba se utilizaron las imágenes de diatomeas fósiles mostradas en la figura 12 y las imágenes de 30 especies de fitoplancton mostradas en la figura 14. Con el fin de tener un banco de imágenes de prueba más completo, todas estas imágenes fueron rotadas y escaladas digitalmente con las funciones de Matlab *imrotate* e *imresize*, respectivamente.

4.1 Identificación de diatomeas fósiles

La primera prueba que se realizó siguiendo la metodología propuesta (figura 11), fue con las diatomeas fósiles. Las diatomeas pueden vivir en ambientes muy diferentes, tanto en aguas dulces, como salinas, en el hielo o en aguas termales. También se les puede encontrar fuera del agua, sobre suelo húmedo o sobre cortezas de árboles. Su distribución está estrechamente relacionada con las características físicas y químicas del ambiente y por esta razón son consideradas como valiosos indicadores biológicos. La utilización de las diatomeas como indicadores necesita de un buen conocimiento de su biología, su ecología y su clasificación, lo que hace necesaria la participación de un especialista, ya que son muchas las especies que pueden ser halladas (Maidana, 2013).

Regularmente la identificación de diatomeas requiere de muestras que contienen un gran número de ejemplares de diferentes especies, y muchos de estos ejemplares se encuentran deteriorados con diferentes grados de fragmentación, lo que dificulta aún más la identificación y el requerimiento en tiempo de trabajo de expertos taxónomos. Es por esto que la identificación y clasificación automática de diatomeas es de gran importancia. En la figura 12 se presentan las 21 diatomeas fósiles, las cuales fueron obtenidas de la Cuenca de San Lázaro en Baja California Sur, México (Esparza, 1999).

La primer prueba se realizó con imágenes de 320×320 píxeles de las diatomeas *Coscinodiscus azpeitia*, *Azpeitia nodulifera*, *Thalassiosira oestrupii var 1*, *Actinoptychus bipunctatus* y *Actinoptychus splendens* de la figura 12, las cuales se etiquetan con las letras B, C, I, S y T, respectivamente. Se eligieron estas imágenes debido a la similitud que existe entre ellas. Con cada una de estas diatomeas se generaron versiones escaladas desde un 80% hasta un 120% de su tamaño original, con incrementos del 1%. Además, cada versión escalada fue rotada desde 0 hasta 180° con incrementos de 10° , lo que nos generó un banco de 3,895 imágenes problema con 779 versiones rotadas y escaladas para cada imagen.

La figura 13 muestra una identificación del 99% para las cinco diatomeas, los resultados son representados en gráficas con diagramas de caja, cada una de estas gráficas representa la estadística correspondiente a los resultados de 3,895 imágenes problema. El promedio de estos valores está representado por un cuadrado (\square) dentro de la caja, a ± 1 error estándar (EE); representado por la caja tenemos un valor de confianza de 68.3%, y a ± 2 errores estándar (2EE), representado por los bigotes de la caja, el nivel de confianza es del 95.4%. Finalmente, los valores atípicos “outliers” (datos numéricamente distantes del resto de los datos que forman la estadística) son representados por un círculo. Si no se presenta un cruce entre estos valores el nivel de confianza es de al menos 99%. Las gráficas de caja fueron realizadas con el programa STATISTICA 6. La metodología para lograr el reconocimiento de imágenes invariante a distorsiones geométricas causadas por cambios en posición, rotación y escala fue programada en MATLAB.

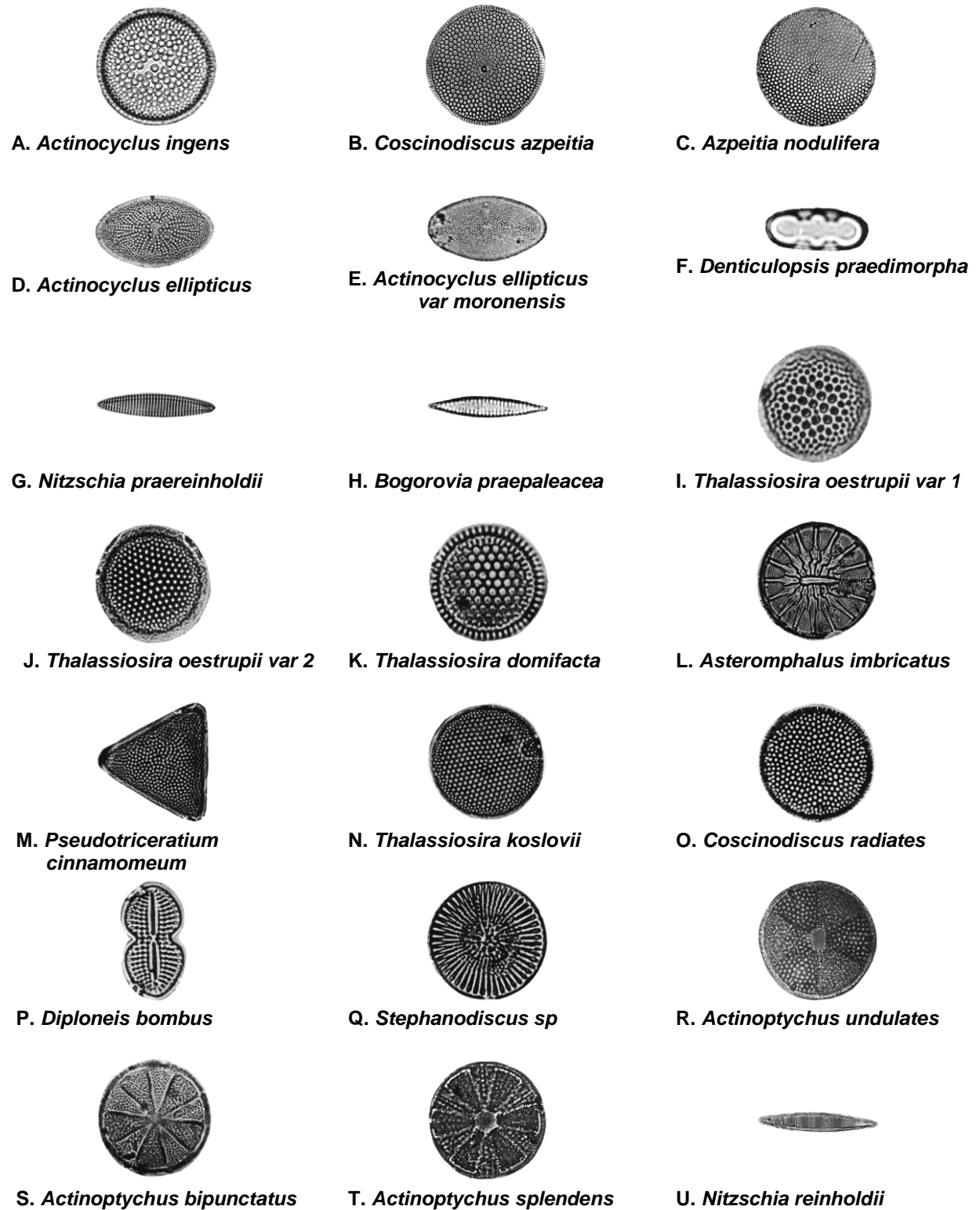


Figura 12. Imágenes de diatomeas utilizadas para probar la metodología.

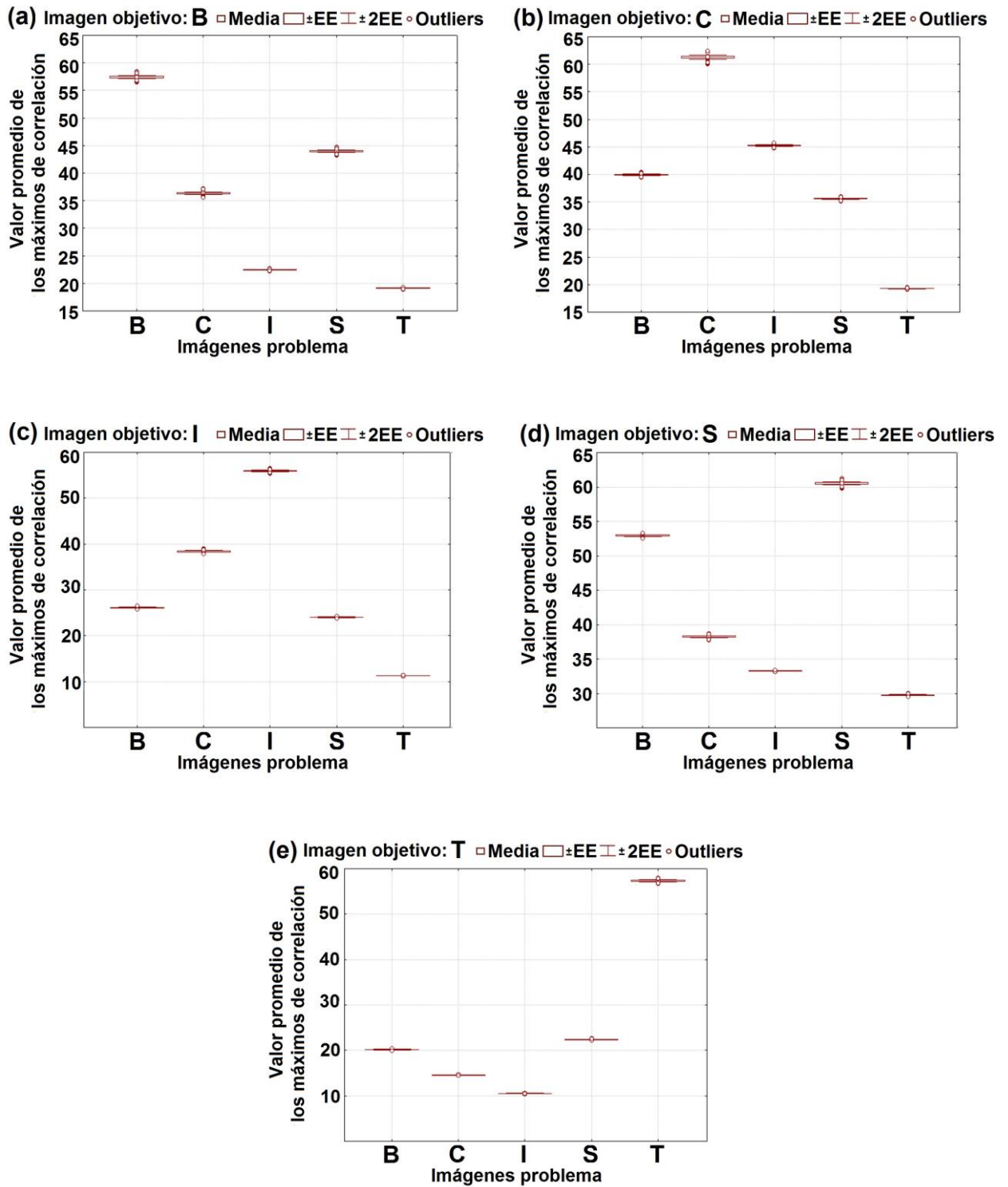


Figura 13. Resultados de correlaciones obtenidas utilizando la nueva metodología basada en la transformada de escala con invariancia a posición, rotación y escala (Figura11). (a) Filtro para diatomea B. (b) Filtro diatomea C. (c) Filtro diatomea I. (d) Filtro diatomea S. (e) Filtro diatomea T.

4.2 Identificación de fitoplancton

Como la metodología respondió con éxito al utilizar las cinco diatomeas fósiles que forman un banco de imágenes de 3,895 imágenes, se procedió a trabajar con un nuevo grupo de imágenes reales de 30 especies de fitoplancton con un tamaño de 320×320 píxeles (figura 14). Se define como fitoplancton la comunidad de microorganismos, en su mayoría fotosintéticos, que viven suspendidos en las masas de agua dulce y agua salada. Los estudios del fitoplancton son de gran importancia para la biología, ecología y conservación del ecosistema marino. Sin embargo, estas investigaciones implican un proceso de alto costo en términos de tiempo y esfuerzo, requiriendo personal entrenado para la identificación, conteo y medición de numerosos grupos taxonómicos, así como para la estimación de biomasa. Muchos estudios quedan incompletos debido al alto número de muestras colectadas, las cuales terminan almacenadas, sin ser analizadas en su totalidad, como resultado del requerimiento de personal especializado y dedicación necesaria para la realización de análisis confiables de la composición taxonómica. Es por esto que actualmente existe un creciente interés por la implementación de nuevas técnicas que permitan automatizar los análisis, incorporando el reconocimiento automático de imágenes digitalizadas, para acelerar el procesamiento de las muestras de fitoplancton y obtener resultados más rápidos en la medición, conteo e identificación de organismos (Rodríguez, 2005).

Para el experimento con este grupo de imágenes se tomaron las 30 imágenes de fitoplancton y se sometieron a cambios de escala desde 90% hasta 110% de su tamaño original, con incrementos del 1%, y cada versión escalada fue rotada desde 0 hasta 180° con incrementos de 10° . Durante este experimento se generaron 60 gráficas. En 30 de estas se utilizó como método de comparación entre firmas vectoriales la correlación no lineal (C_{NL}) dada por la ecuación (62), como ejemplo se presentan las figuras 15a, 16a, 17a, 18a y 19a; y en las otras 30 se utilizó la correlación adaptativa (C_{AD}) dada por la ecuación (64), como ejemplo se presentan las figuras 15b, 16b, 17b, 18b y 19b. En cada una de las 60 gráficas se compara una imagen objetivo contra las 11,970 imágenes problema.

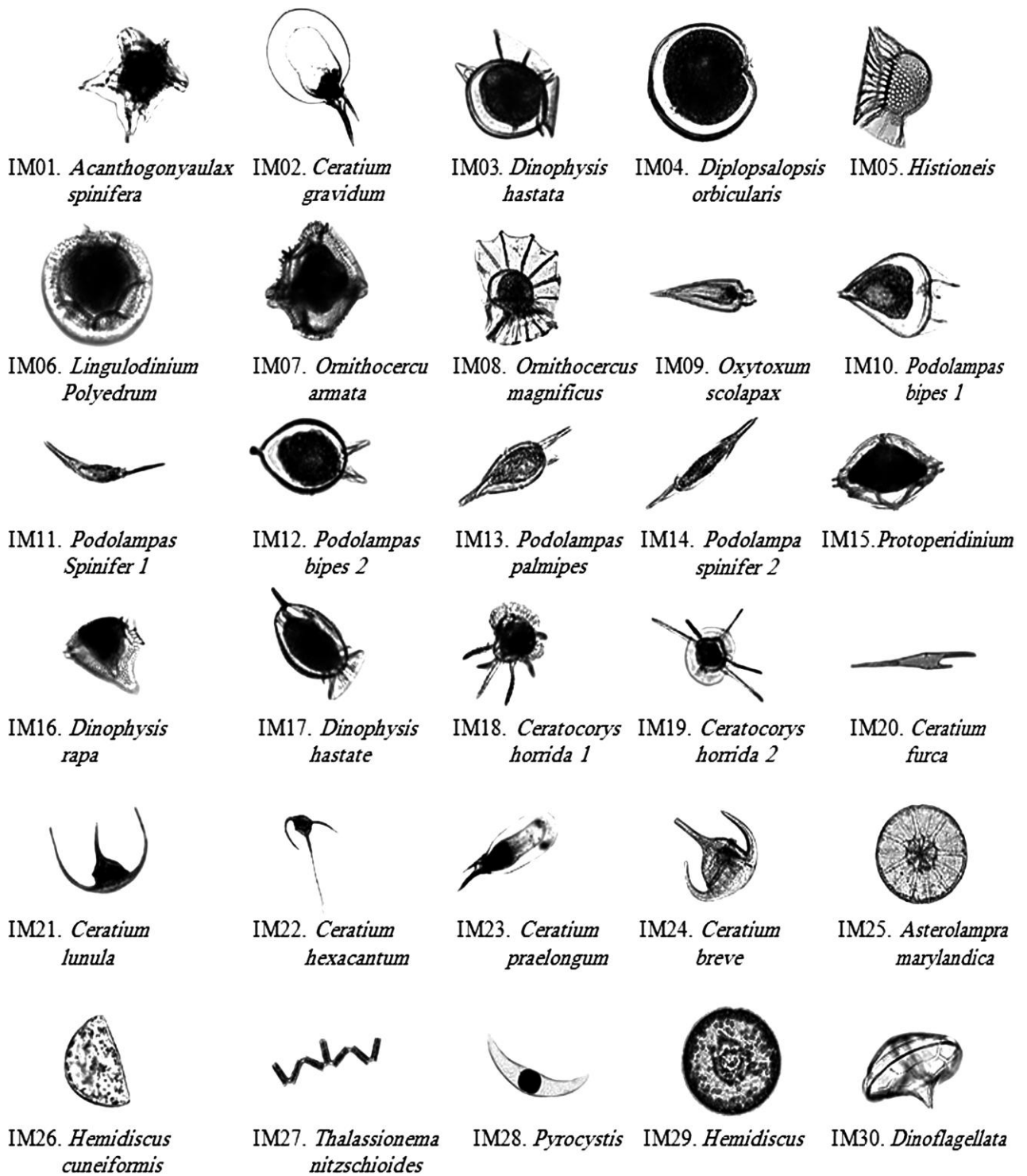


Figura 14. Imágenes de las 30 especies de fitoplancton usadas para probar la metodología adaptativa.

En la tabla 1 se hace un resumen de los resultados de las 60 gráficas, en donde se presenta el porcentaje de confiabilidad para la detección de la imagen objetivo. Este porcentaje fue obtenido de los diagramas de cajas, que contienen información de la

media de los valores del máximo de correlación. Se observa claramente que la mayoría de los resultados dados por la correlación C_{NL} son mejorados por la versión adaptativa C_{AD} . Además, se comprueba que en el peor de los casos la versión adaptativa es igual a la clásica correlación no lineal (ley k).

Tabla 1. Resultados del porcentaje de confiabilidad obtenido con la correlación no lineal (C_{NL}) y la versión adaptativa (C_{AD})

Objetivo	C_{NL}	C_{AD}
IM01	99% sin IM05	99%
IM02	95.4%	95.4%
IM03	95.4%	99%
IM04	99%	99%
IM05	99% sin IM10	99%
IM06	95.4%	99%
IM07	95.4%	99%
IM08	99%	99%
IM09	95.4%	99%
IM10	95.4%	99%
IM11	95.4%	99%
IM12	99%	99%
IM13	99% sin IM26	99%
IM14	95.4% sin IM05	95.4%
IM15	99%	99%
IM16	95.4%	99%
IM17	99%	99%
IM18	95.4%	99%
IM19	99%	99%
IM20	95.4%	95.4%
IM21	99%	99%
IM22	95.4%	95.4%
IM23	99%	99%
IM24	95.4%	99%
IM25	99%	99%
IM26	95.4%	99%
IM27	95.4%	99%
IM28	95.4%	99%
IM29	99%	99%
IM30	95.4%	99%

En la tabla 1, hay 30 casos de comparación entre la correlación no lineal (C_{NL}) y su versión adaptativa (C_{AD}). Se presentan 17 casos en donde se logra mejor porcentaje de confiabilidad al utilizar C_{AD} que al utilizar C_{NL} . En cuatro de estos casos (los casos de las

imágenes objetivo IM01, IM05, IM13 e IM14), los resultados de la correlación C_{NL} mostraron traslapes entre resultados con imágenes problema que no pertenecen a la especie de la imagen objetivo. En la figura 15a se encuentra el caso en que los resultados de la imagen objetivo *Acanthogonyaulax spinifera* (IM01) se traslapan con los resultados de *Histioneis* (IM05). En la figura 16a se aprecia el traslape de resultados para la imagen objetivo *Histioneis* (IM05) con los resultados de *Podolampas bipes 1* (IM10). En la figura 17a los resultados de *Podolampas palmipes* (IM13) se traslapan con los resultados de *Ceratium praelongum* (IM26), y en la figura 18a los resultados de *Podolampas spinifer 2* (IM14) se traslapan con los resultados de *Histioneis* (IM05). Los resultados con traslape son un caso no deseado, ya que nos indican que existen imágenes problema que serán clasificadas erróneamente en el conjunto de imágenes objetivo. En las figuras 15b, 16b, 17b y 18b se aprecia que los problemas por traslape no existen en las gráficas generadas con los resultados que utilizan C_{AD} como método de correlación.

Un ejemplo en donde al utilizar C_{NL} no se generaron traslapes, pero que con el uso de C_{AD} se mejoró el porcentaje de confiabilidad del 95.4% al 99%, lo muestran las gráficas de las figuras 19a y 19b. En los resultados de la gráfica de la figura 19a se aprecia que los “outliers” del diagrama de caja que representa a los resultados de *Dinophysis hastata* (IM03) se traslapan con los “outliers” correspondientes a *Lingudolium Polyedrum* (IM06), lo que da como resultado una confiabilidad de 94.5%. En la figura 19b se aprecia como todos los diagramas de caja se alejan del diagrama correspondiente a los resultados de *Dinophysis hastata* (IM03), logrando así una confiabilidad de 99%.

Solo para los casos de las imágenes objetivo *Ceratium gravidum* (IM02), *Ceratium furca* (IM20) y *Ceratium hexacatum* (IM22), C_{AD} no logra mejorar el porcentaje de confiabilidad del 95.4% al 99%, pero conserva el mismo porcentaje de confiabilidad obtenido con C_{NL} .

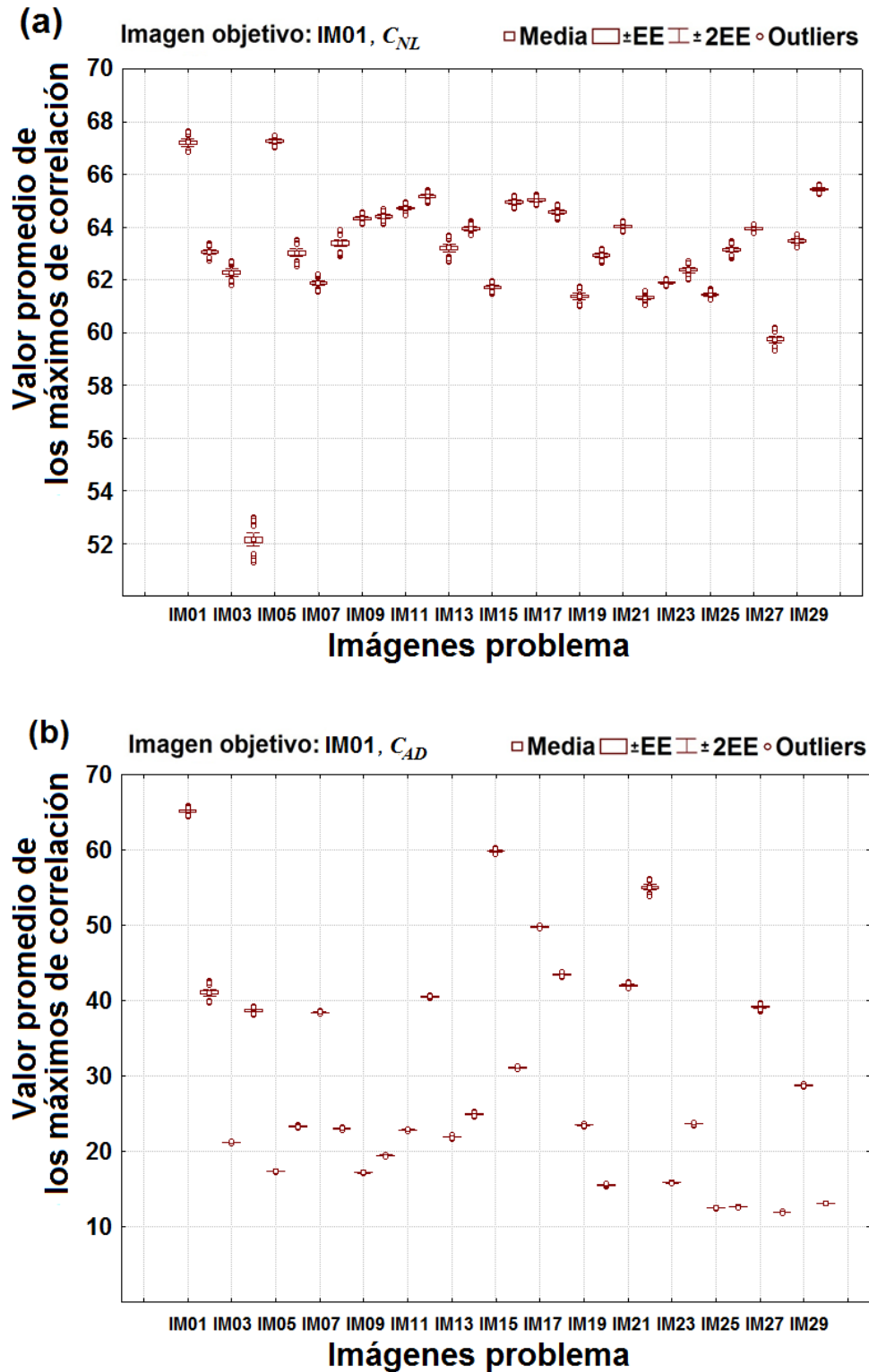


Figura 15. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo *Acanthogonyaulax spinifera* (IM01). (a) Usando correlación no lineal C_{NL} . (b) Usando correlación no lineal adaptativa C_{AD} .

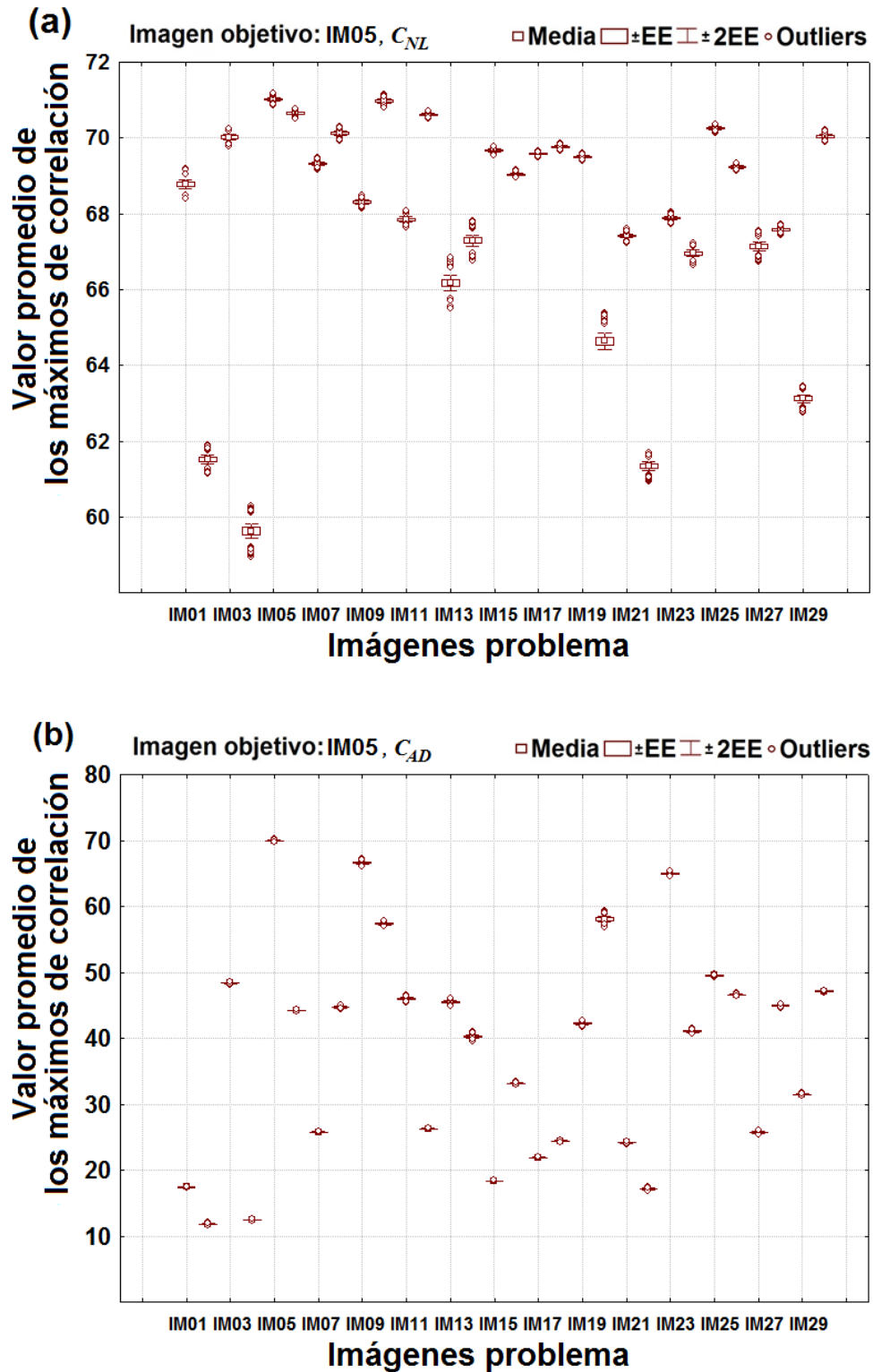


Figura 16. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo *Histioneis* (IM05). (a) Usando correlación no lineal C_{NL} . (b) Usando correlación no lineal adaptativa C_{AD} .

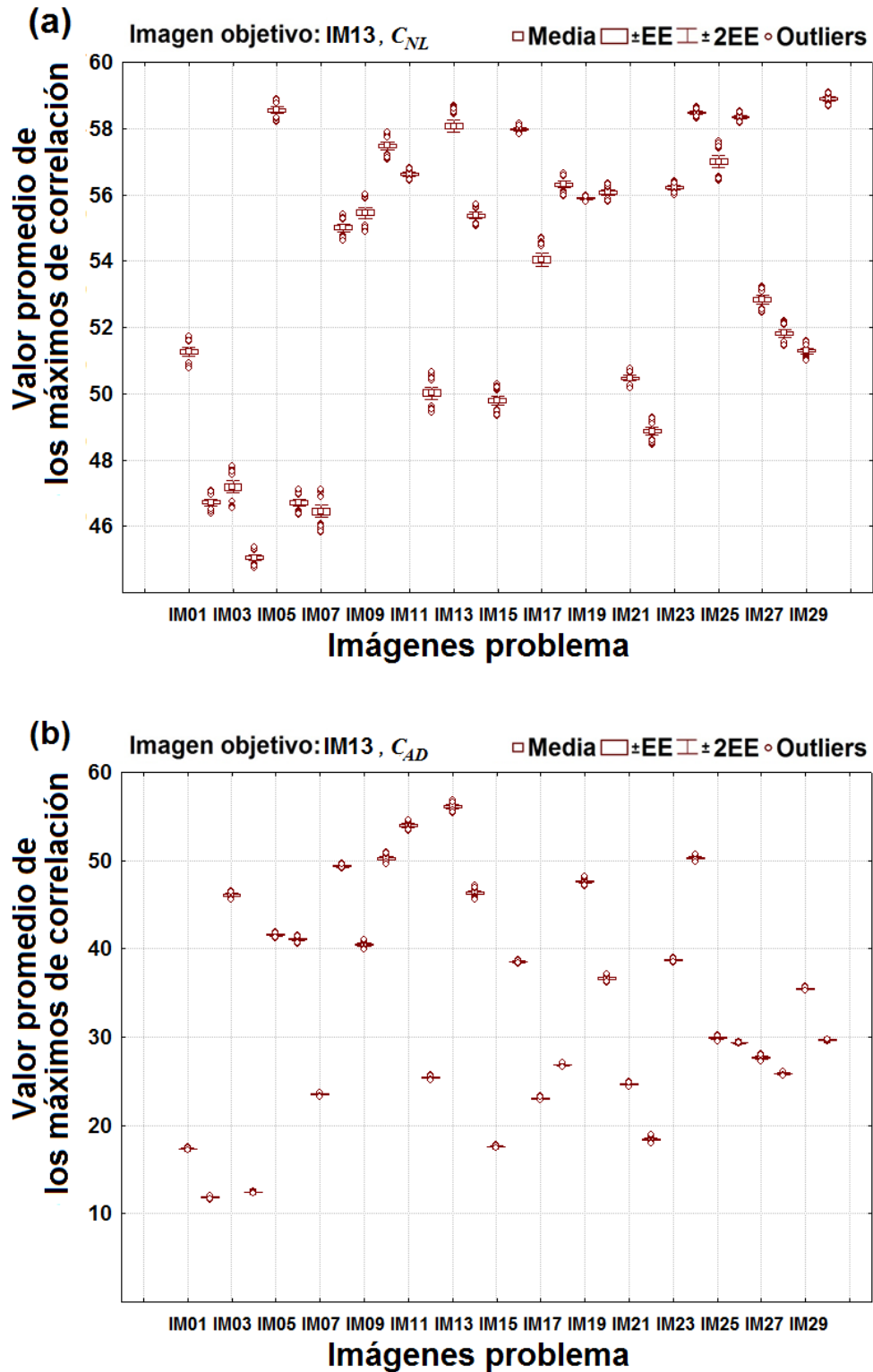


Figura 17. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo *Podolampas palmipes* (IM13). (a) Usando correlación no lineal C_{NL} . (b) Usando correlación no lineal adaptativa C_{AD} .

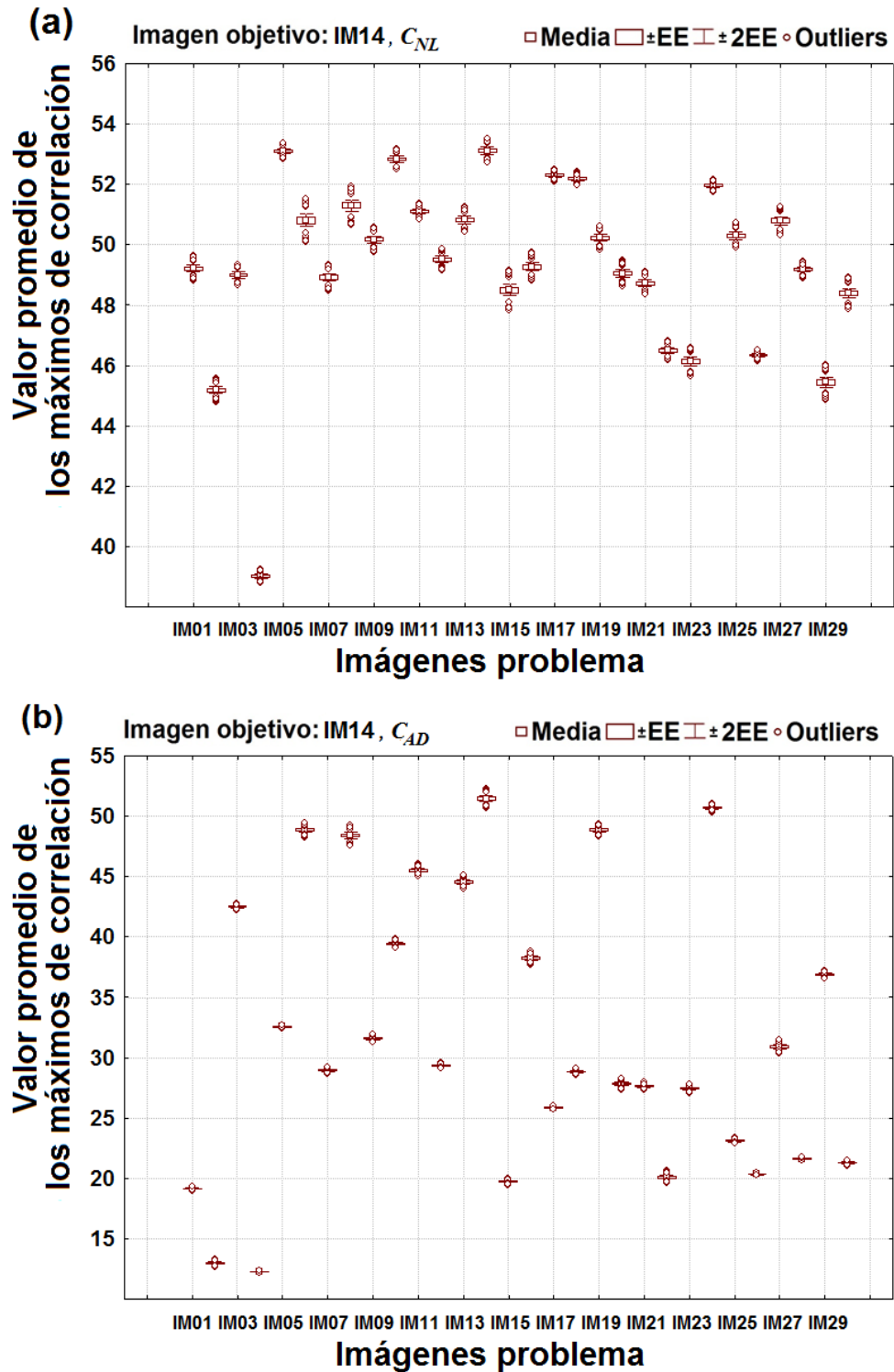


Figura 18. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo *Podolampas spinifer 2* (IM14). (a) Usando correlación no lineal C_{NL} . (b) Usando correlación no lineal adaptativa C_{AD} .

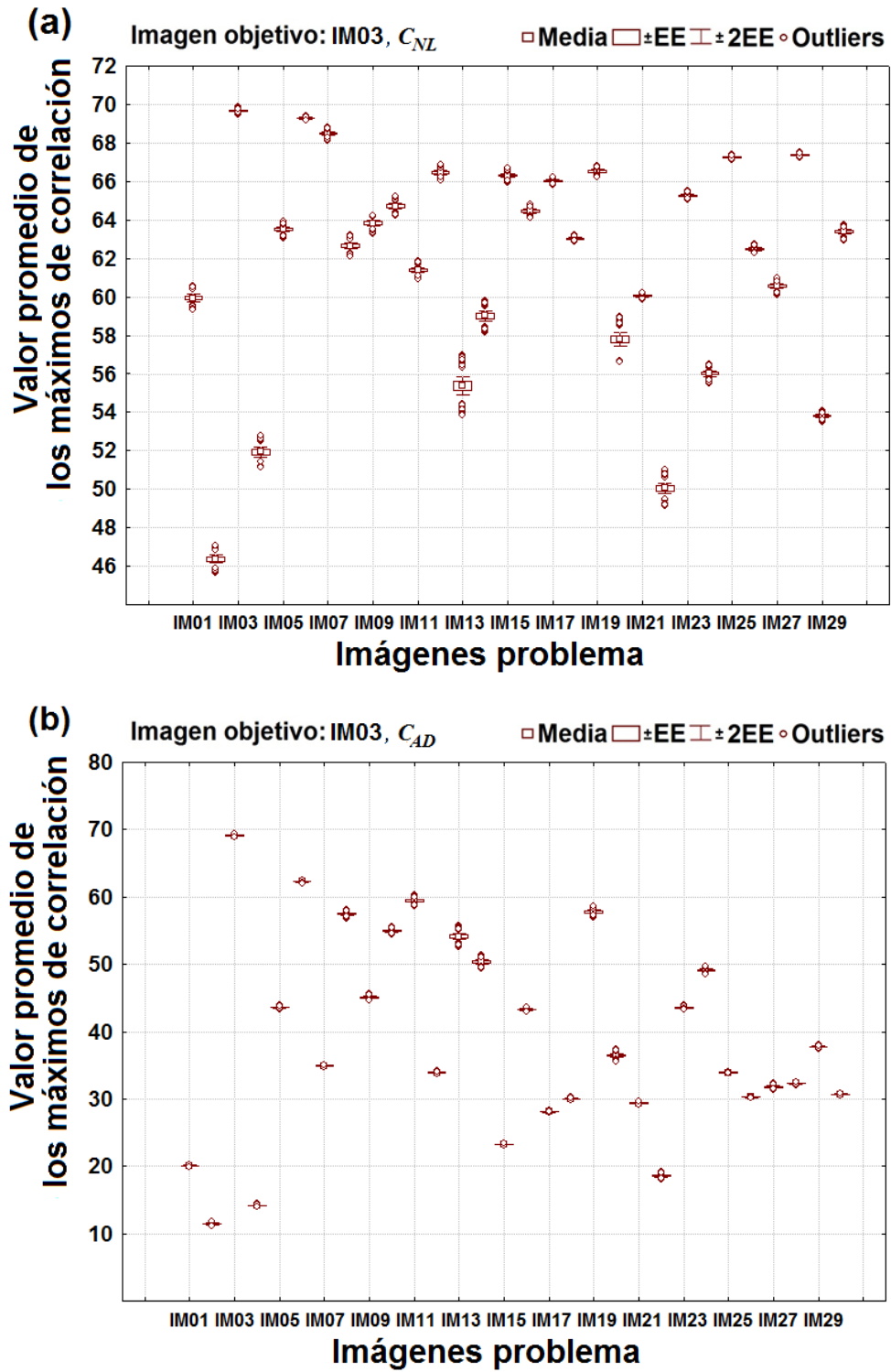


Figura 19. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo *Podolampas spinifer* 2 (IM03). (a) Usando correlación no lineal C_{NL} . (b) Usando correlación no lineal adaptativa C_{AD} .

4.3 Imágenes con diatomeas fósiles incompletas

Es muy común que al identificar imágenes reales nos encontremos con imágenes incompletas. Por ejemplo, en la identificación de huellas dactilares de investigaciones forenses encontramos este caso muy común, ya que las imágenes de la base de datos contienen únicamente imágenes completas y las imágenes de huellas dactilares recolectadas en la escena del crimen, que son llamadas huellas latentes, se encuentran en la mayoría de los casos degradadas e incompletas. Esto imposibilita en la mayoría de los casos una identificación automática por computadora (Jain y Feng, 2011).

En la actualidad las diatomeas fósiles, son utilizadas como indicadores ambientales de tiempos pasados. Se encuentran en muestras de sedimento, por lo que es muy común encontrarlas en fragmentos o pulverizadas, lo cual puede producir confusión en la identificación de la especie (Hajós, 1967).

Realizar un examen basado en la búsqueda de diatomeas en los órganos vitales de una persona fallecida, es en la actualidad una práctica de rutina para el diagnóstico de muerte por sumersión (o ahogamiento húmedo), en donde es común encontrar fragmentos de diatomeas (Maidana, 2013).

Hay problemas en donde las imágenes que deseamos reconocer se encuentran parcialmente ocultas, como puede ser en un caso de identificación de rostros en donde algunas imágenes problema contienen personas con lentes oscuros. En este caso podríamos considerar que la imagen problema está incompleta, ya que la información del rostro correspondiente a los ojos se encuentra ausente, causando problemas de identificación y la necesidad de una identificación más robusta (Latha et. al., 2015).

4.3.1 Fragmentos de diatomeas fósiles como Imágenes problema

Se pueden presentar problemas en donde contamos con imágenes que contienen solamente una parte del objeto a identificar, es decir, el objeto se encuentra incompleto en la imagen, la metodología desarrollada se sometió a una prueba con fragmentos de nuestra base de datos de diatomeas fósiles como imágenes problema. Contamos con 49 fragmentos de diferentes tamaños para cada una de las 21 diatomeas con un

tamaño de 256×256 píxeles (apéndice A). La figura 20 muestra cinco de los 49 fragmentos de la diatomea A utilizados en el experimento. Los 49 fragmentos de las 21 diatomeas fueron rotados desde 0 hasta 180° y escalados desde 90% hasta 110% de su tamaño original, obteniendo resultados que nos dan información sobre la capacidad de identificación cuando la metodología trabaja con fragmentos de información.

Para que la aplicación de C_{AD} sea exitosa, se requiere que las imágenes problema correspondientes a la imagen objetivo, mantengan una relación con poca variabilidad en los índices que dependen de su área y volumen. Esto no sucede con los fragmentos de diatomeas, ya que la razón dada entre los índices invariantes a posición, rotación y escala es muy variable para cada uno de los 49 fragmentos de una misma diatomea, por lo que hacer una comparación con C_{AD} no es efectivo.

En la figura 21 se utiliza la C_{NL} como método de comparación entre firmas y se muestran los resultados obtenidos a partir de la comparación de la imagen objetivo diatomea entera A contra las 231,241 imágenes problema, que son las 50 imágenes que contienen diatomeas en fragmentos para cada una de las 21 diatomeas, a su vez cada una de estas imágenes tiene 19 rotaciones y cada una de estas versiones rotadas se escala 21 veces. En la figura 22 se muestra este mismo caso pero con la diatomea B como imagen objetivo.

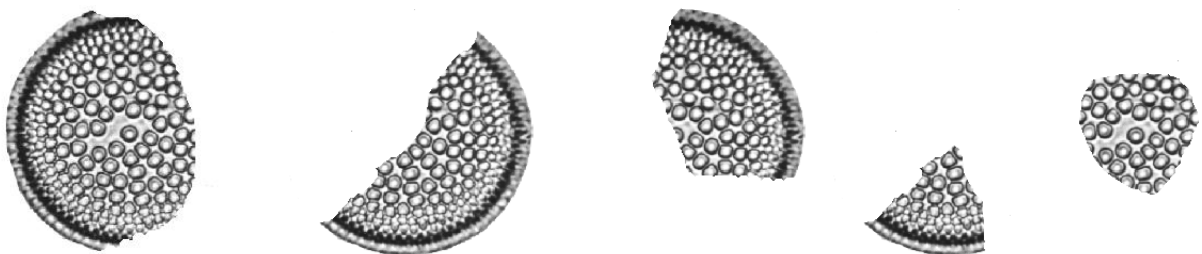


Figura 20. Ejemplo de cinco fragmentos de la diatomea *Actinocyclus ingens*, utilizados en el experimento.

La tabla 2 muestra los resultados obtenidos después de comparar cada una de las imágenes de diatomeas fósiles y sus fragmentos contra las versiones rotadas y escaladas de todas las imágenes contenidas en la base de datos con fragmentos de diatomeas (Apéndice A).

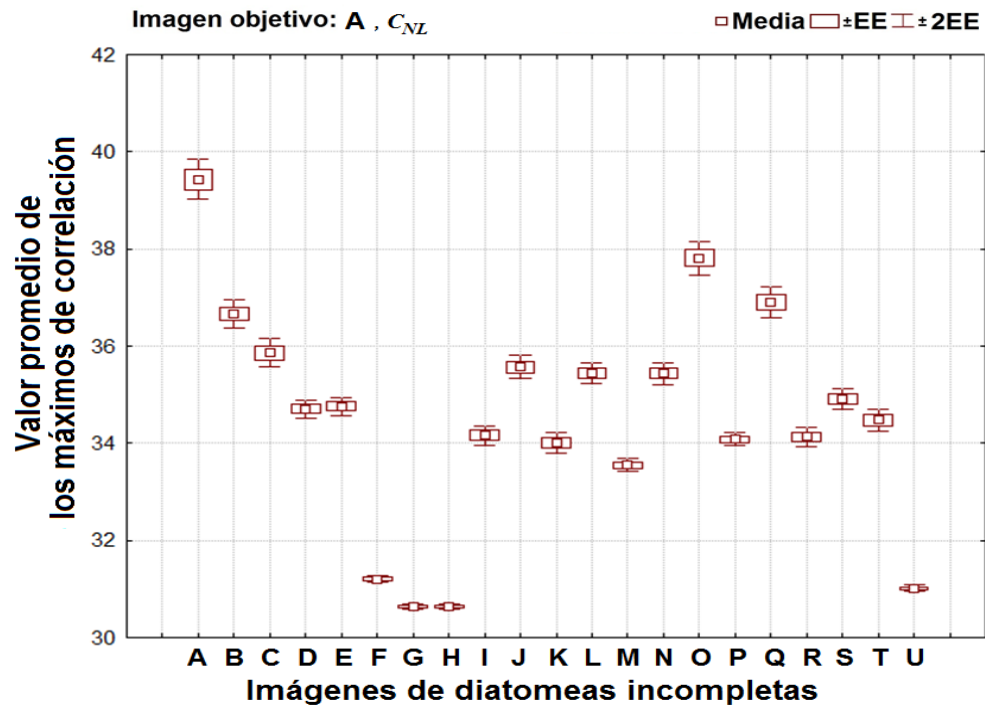


Figura 21. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo la diatomea A entera y como imágenes problemas las 21 diatomeas y sus 49 fragmentos rotados y escalados.

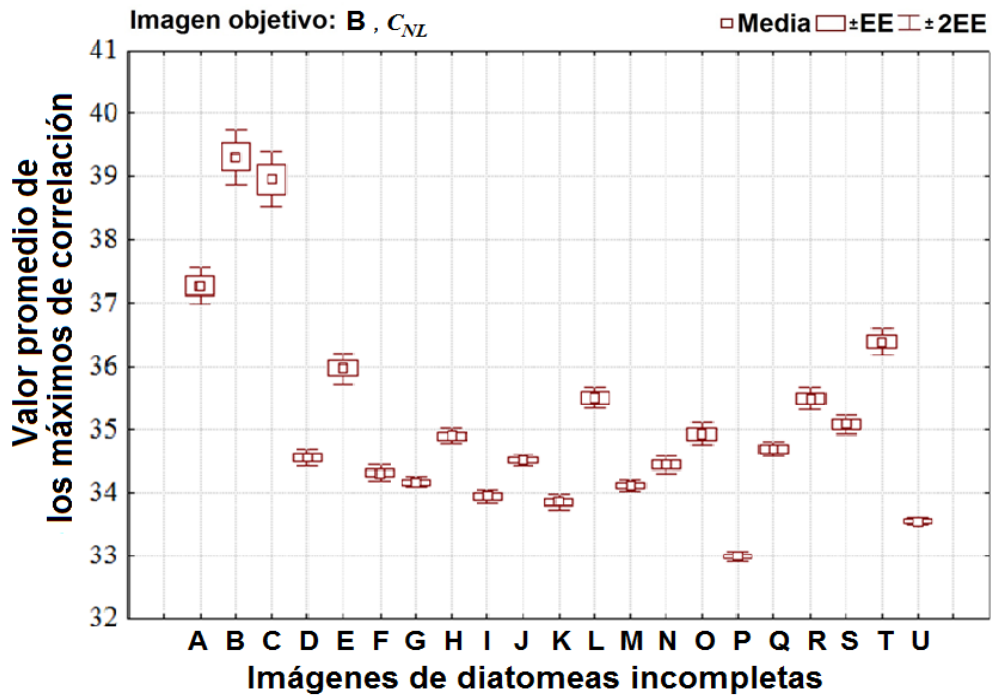


Figura 22. Resultados del valor promedio de los máximos de correlación utilizando como imagen objetivo la diatomea B entera y como imágenes problemas las 21 diatomeas y sus 49 fragmentos rotados y escalados.

En la primera columna de la tabla 2 están los nombres de las diatomeas utilizadas como imagen objetivo. La segunda columna nos muestra el porcentaje de confiabilidad obtenido al utilizar como objetivo una diatomea entera y como en la mayoría de los casos existe traslape de resultados. En la tercera columna se indica la condición especial que permite tener la confiabilidad mostrada en la tabla, por ejemplo, en el primer renglón no hay condición especial porque la caja de la imagen objetivo no se traslapa con ninguna caja de las imágenes problema. En este primer renglón, se encuentran los datos obtenidos de la figura 21, donde se utiliza la diatomea A entera, como imagen objetivo. En el segundo renglón se encuentran los datos obtenidos de la figura 22, en donde se tiene la condición especial “sin C”, que nos indica que debemos excluir los resultados de la diatomea C para obtener el resultado mostrado en la tabla, ya que los resultados de la diatomea A se traslapan con los resultados de la diatomea C. Por último, la quinta columna indica el mínimo tamaño del fragmento expresado en porcentaje, que seguirá obteniendo el mismo porcentaje de confiabilidad que logra la diatomea entera, al ser utilizado como imagen objetivo.

Tabla 2. Resultados del porcentaje de confiabilidad obtenido con la correlación no lineal (C_{NL}) y las diatomeas incompletas como imágenes problema

Nombre de la diatomea (imagen objetivo)	Confiabilidad	Condición especial	Fragmento Mínimo
A. <i>Actinocyclus ingens</i>	95.4%		84%
B. <i>Coscinodiscus azpeitia</i>	95.4%	sin C	50%
C. <i>Azpeitia nodulifera</i>	95.4%	sin B	76%
D. <i>Actinocyclus ellipticus</i>	95.4%	sin R, S	76%
E. <i>Actinocyclus ellipticus var moronensis</i>	68.3%	sin C, N, O, Q	85%
F. <i>Denticulopsis praedimorpha</i>	68.3%	sin Q, R	85%
G. <i>Nitzschia praereinholdii</i>	95.4%		45%
H. <i>Bogorovia praepaleacea</i>	95.4%		69%
I. <i>Thalassiosira oestruppii var 1</i>	95.4%	sin O	90%
J. <i>Thalassiosira oestruppii var 2</i>	95.4%	sin A	80%
K. <i>Thalassiosira domifacta</i>	68.3%	sin I, J	62%
L. <i>Asteromphalus imbricatus</i>	68.3%	sin B, J, N	45%
M. <i>Pseudotriceratium cinnamomeum</i>	68.3%	sin L, Q, T	58%
N. <i>Thalassiosira kozlovii</i>	95.4%	sin J, L	52%
O. <i>Coscinodiscus radiatus</i>	68.3%		73%
P. <i>Diploneis bombus</i>	68.3%		73%
Q. <i>Stephanodiscus sp</i>	95.4%		87%
R. <i>Actinoptychus undulatus</i>	68.3%	sin B, C, D, S	90%
S. <i>Actinoptychus bipunctatus n</i>	68.3%	sin B, D, P	89%
T. <i>Actinoptychus splendens</i>	68.3%	sin A, N, O	89%
U. <i>Nitzschia reinholdii</i>	95.4%	sin G, H	53%

4.3.2. Fragmentos de diatomeas fósiles como imagen objetivo

Se realizaron pruebas en las que se utiliza una diatomea incompleta como imagen objetivo y esta se compara contra el banco de imágenes formado por las versiones rotadas y escaladas de diatomeas en la figura 12 (diatomeas completas). En este caso, la relación entre área y volumen del fragmento utilizado como objetivo es constante, y todas las imágenes problema al ser diatomeas enteras también mantendrán poca variabilidad en su relación entre área y volumen, así que la correlación no lineal adaptativa puede ser utilizada.

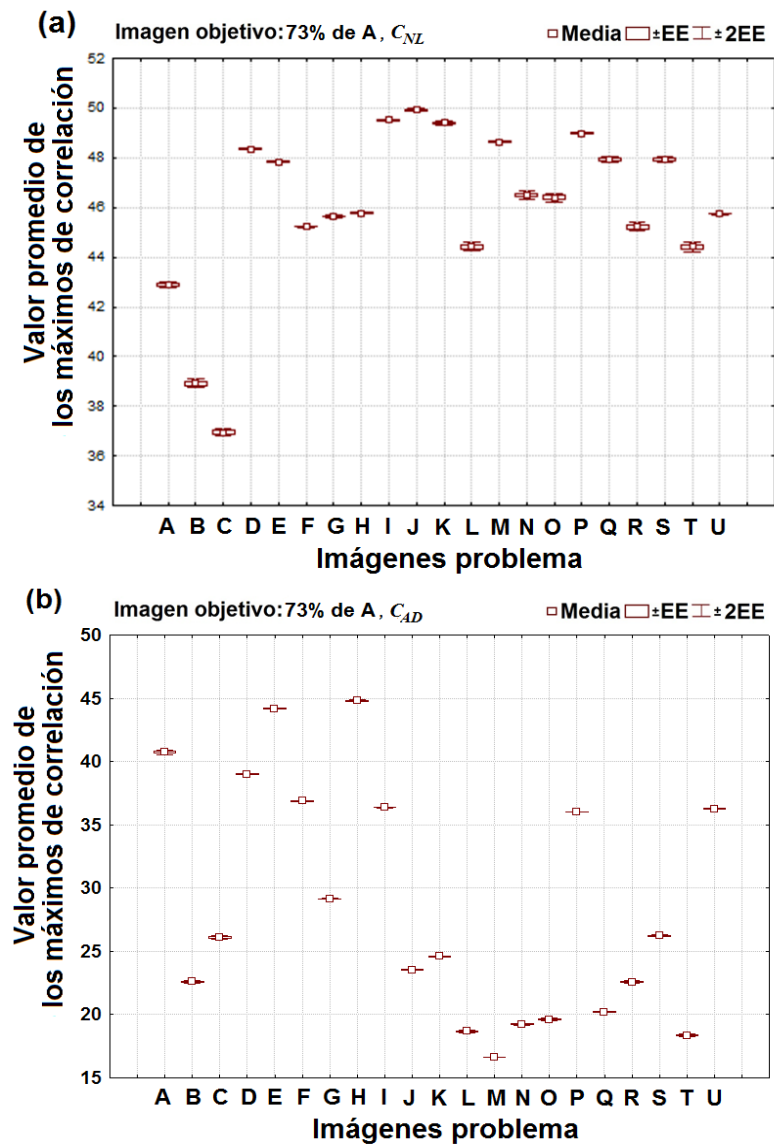


Figura 23. Resultados de correlación para la imagen con un fragmento del 73% de la diatomea A. (a) Utilizando C_{NL} . (b) Utilizando C_{AD} .

Las figuras 23a y 23b, presentan los resultados obtenidos con C_{NL} y con C_{AD} , respectivamente, para el fragmento 8 de la diatomea A como imagen objetivo, que corresponde a un fragmento del 73% de la diatomea A entera. En las figuras 24a y 24b se presentan los resultados obtenidos con C_{NL} y con C_{AD} , respectivamente, utilizando el fragmento 39 de la diatomea A como imagen objetivo, que corresponde a un fragmento del 6% de la diatomea A entera. Para los casos presentados en las figuras 23 y 24, las dos correlaciones no lineales logran resultados satisfactorios.

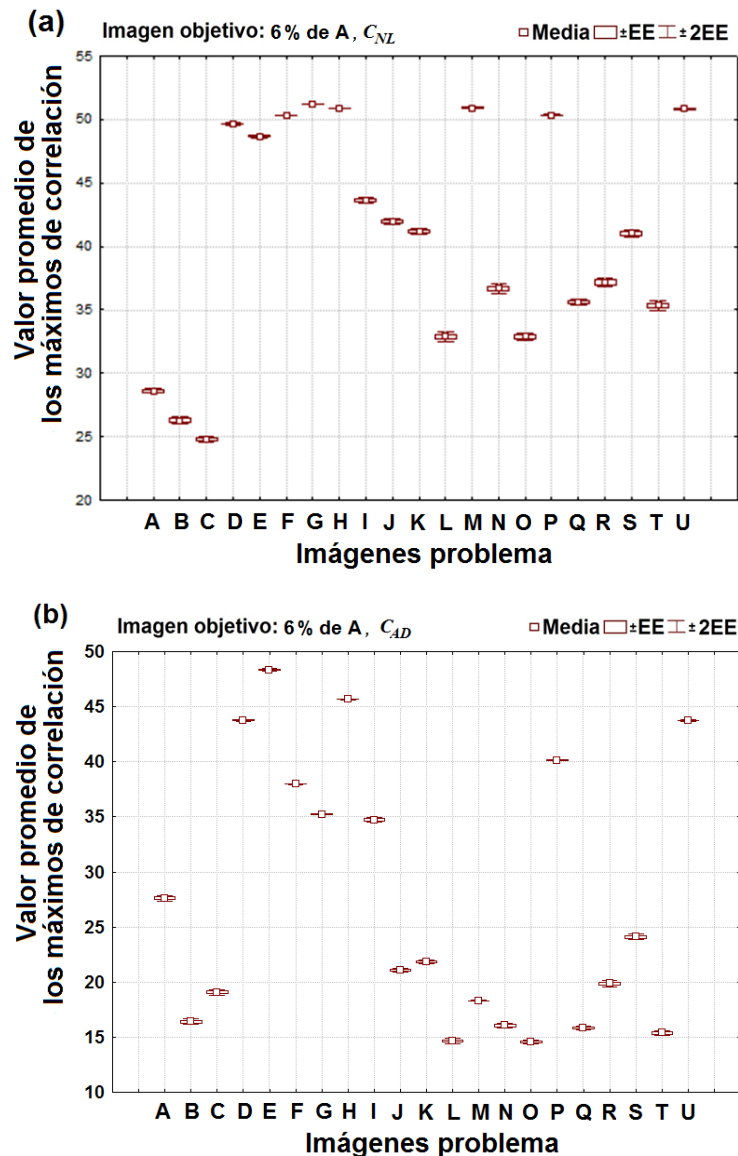


Figura 24. Resultados de correlación para la imagen con un fragmento del 6% de la diatomea A. (a) Utilizando C_{NL} . (b) Utilizando C_{AD} .

4.4 Capacidad para reconocer imágenes afectadas por ruido

La capacidad para reconocer imágenes inmersas en ruido fue medida con el coeficiente de discriminación (CD). Para estas pruebas se utilizó como filtro la imagen de la diatomea fósil *Actinocyclus ingens* (Diatomea A de la figura 12) y la correlación no lineal adaptativa con $k=0.1$, considerando ruido gaussiano aditivo con media cero y ruido impulsivo, o sal y pimienta.

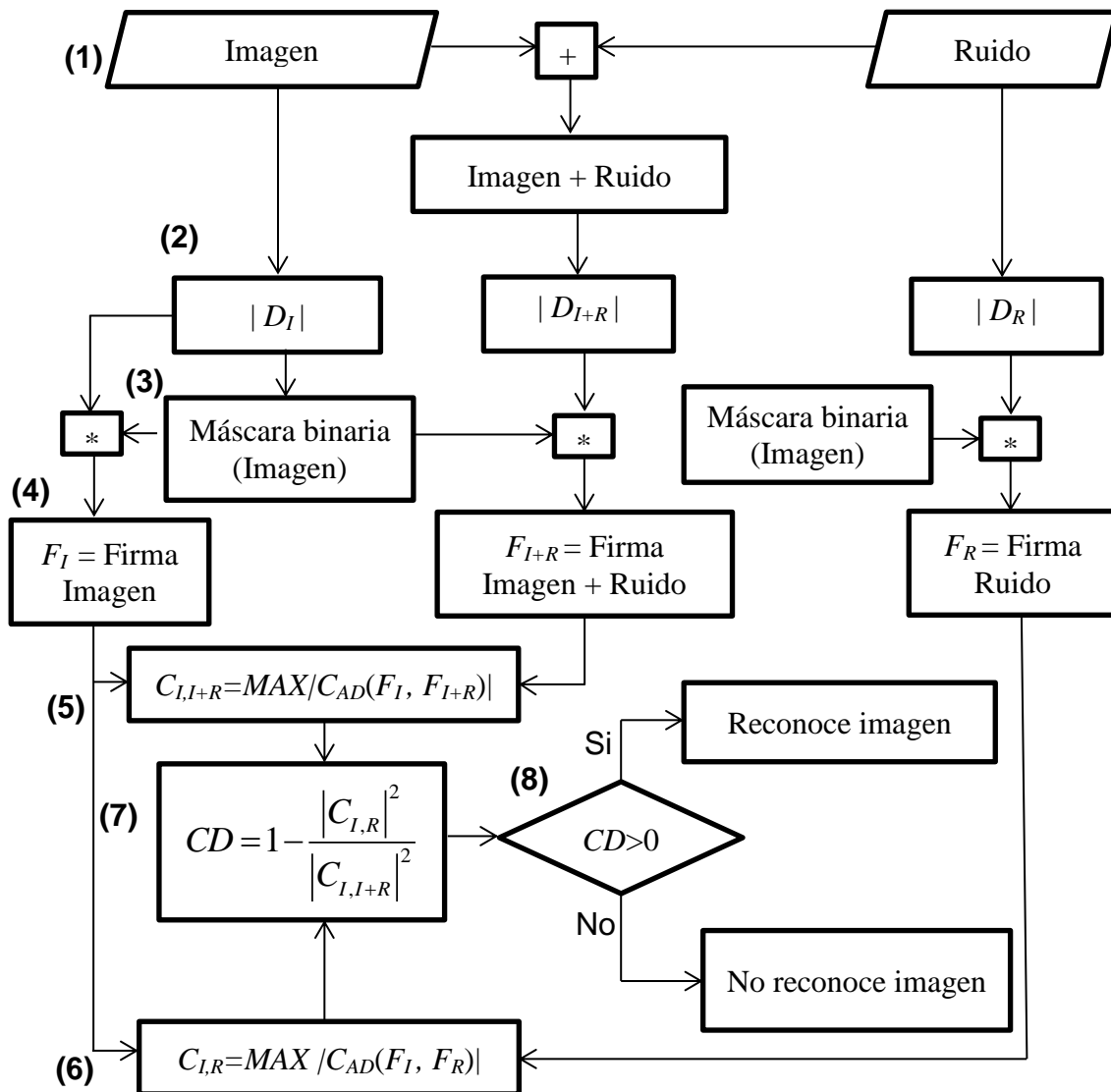


Figura 25. Proceso para determinar el valor del coeficiente de discriminación (CD), donde el símbolo * especifica una multiplicación punto a punto.

En la figura 25 se muestra el proceso para obtener el valor del coeficiente de discriminación a partir de una imagen sin ruido y un cierto nivel de ruido, que se suman y resulta una imagen con ruido aditivo. En el paso 2 se obtiene el módulo de la transformada de escala de la imagen sin ruido ($|D_I|$), de una imagen vacía con fondo ruidoso ($|D_R|$) y de la imagen con ruido aditivo ($|D_{I+R}|$). En el paso 3, y siguiendo el diagrama de la figura 5, se utiliza el módulo de la transformada de escala de la imagen sin ruido para construir una máscara de anillos binarios. En el paso 4, a partir de las multiplicaciones punto a punto de la máscara de anillos con los módulos de las transformadas de escala de las imágenes, se generan las firmas que representan a la imagen sin ruido (F_I), al fondo con ruido (F_R) y a la imagen con ruido aditivo (F_{I+R}). En los pasos 5 y 6 se realiza una correlación no lineal adaptativa entre la imagen sin ruido y la imagen con ruido aditivo ($C_{I,I+R}$) y una correlación no lineal adaptativa entre la imagen sin ruido y el ruido ($C_{I,R}$), respectivamente. Con estos valores de correlación, en el paso 7 se calcula el coeficiente de discriminación utilizando la ecuación (56). Finalmente, en el paso 8, un valor positivo de CD nos indica que la metodología evaluada con cierto nivel de ruido es capaz de reconocer el objeto en la imagen con éxito, y cualquier valor negativo nos indica que es imposible reconocer el objeto presente en la imagen.

La figura 26 muestra una gráfica con el resultado del coeficiente de discriminación promedio con un nivel de confianza del 95.4% para la diatomea A, afectada por ruido aditivo de sal y pimienta. En dicha gráfica se varió la densidad desde 0 hasta 1 con incrementos de 0.01. En la figura 27 se encuentran los resultados del coeficiente de discriminación promedio con un nivel de confianza del 95.4% para la misma imagen, pero ahora afectada por ruido aditivo gaussiano, en donde se modificó la varianza desde 0 hasta 2. En estos experimentos se realizaron 36 iteraciones para cada valor de densidad en el caso del ruido sal y pimienta y 36 iteraciones para cada valor de varianza en el caso del ruido gaussiano. De ellas se obtuvieron el valor promedio y el error estándar en los resultados de los coeficientes de discriminación calculados.

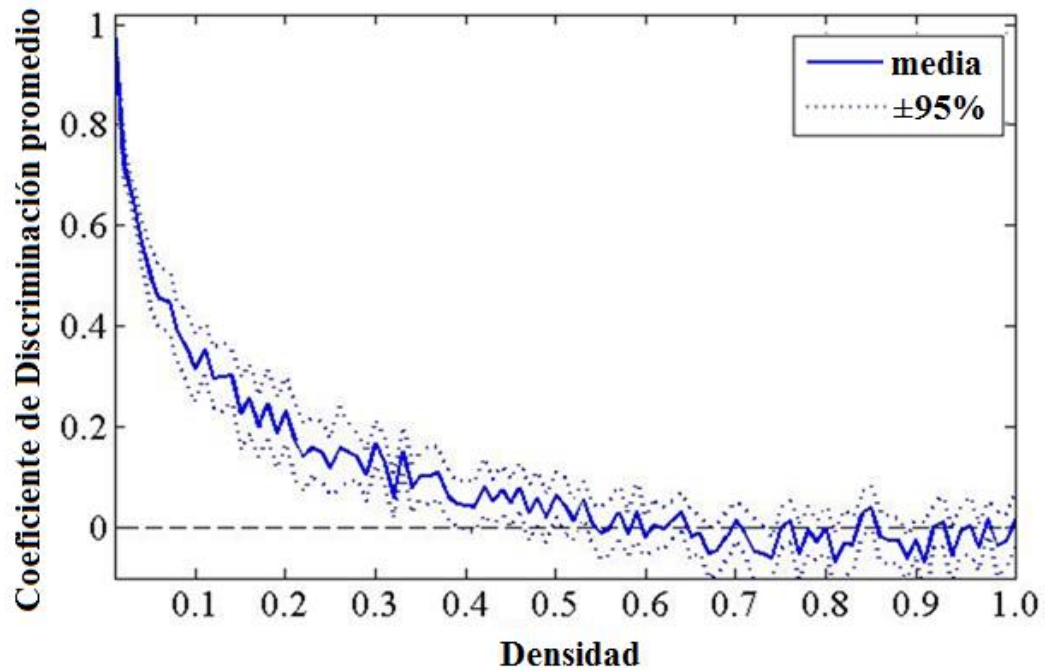


Figura 26. Rendimiento del filtro en presencia de ruido sal y pimienta.

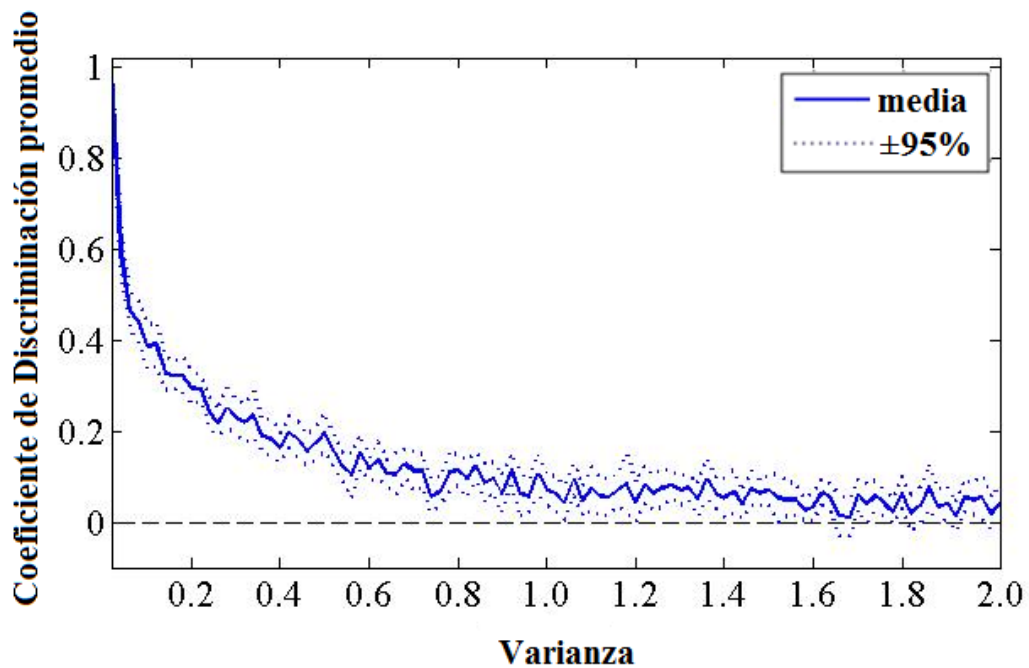


Figura 27. Rendimiento del filtro en presencia de ruido gaussiano.

Las gráficas de las figuras 26 y 27 indican que la metodología desarrollada en esta tesis tiene un límite de identificación a una densidad de 0.4 en el caso del ruido sal y pimienta y un límite de identificación para una varianza de 1.6 en el caso del ruido gaussiano, ya que en estos valores de densidad y varianza, para una confiabilidad de al menos 95%, el coeficiente de discriminación promedio llega a cero.

En la figuras 28a y 28b se aprecian ejemplos de la diatomea A afectada por diferentes niveles de ruido sal y pimienta y ruido gaussiano, respectivamente.

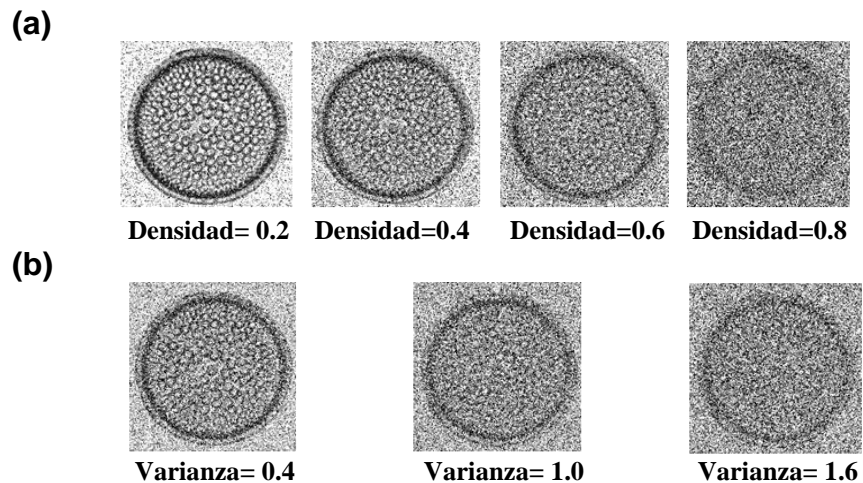


Figura 28. Ejemplos de la diatomea fósil *Actinocyclus ingen*, afectada por ruido aditivo. (a) Sal y pimienta. (b) Gaussiano.

4.5 Reconocimiento de imágenes con iluminación no homogénea

Los índices que se calculan siguiendo el procedimiento indicado en la figura 10, se sustituyen en las ecuaciones (46), (47) y (48), se usan para conseguir una correlación no lineal adaptativa a cambios de rotación y escala, pero los cambios en iluminación provocan variabilidad en la relación de área y volumen de las imágenes. Es por esto, que en este caso se ha utilizado la correlación no lineal (C_{NL}) como método de comparación entre firmas vectoriales.

Esta metodología fue evaluada con los ocho diferentes patrones de iluminación no homogénea mostrados en la figura 29, y se utilizaron imágenes de 256×256 píxeles de las diatomeas *Coscinodiscus azpeitia*, *Azpeitia nodulifera*, *Thalassiosira oestrupii var 1*, *Actinoptychus bipunctatus* y *Actinoptychus splendens* que se etiquetan con las letras B, C, I, S y T, respectivamente (figura 12). Estas imágenes se escalaron desde un 80% hasta 120% de su tamaño original con incrementos del 1%. Además, cada versión escalada fue rotada desde 0 hasta 180° con incrementos de 10° , lo que nos generó 3,895 imágenes problema rotadas y escaladas para cada patrón de iluminación no homogénea, haciendo un banco de 31,160 imágenes de prueba.

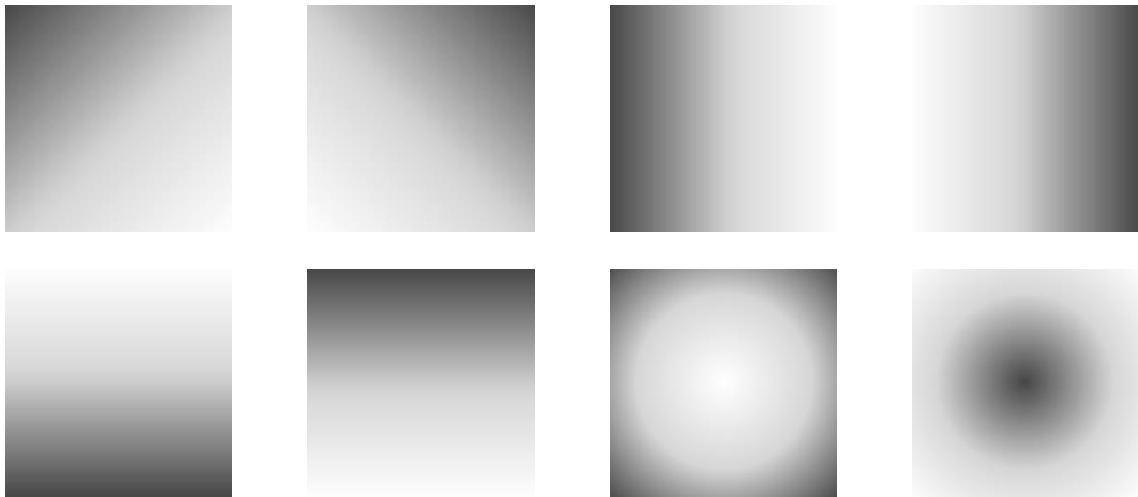


Figura 29. Patrones de iluminación no homogénea.

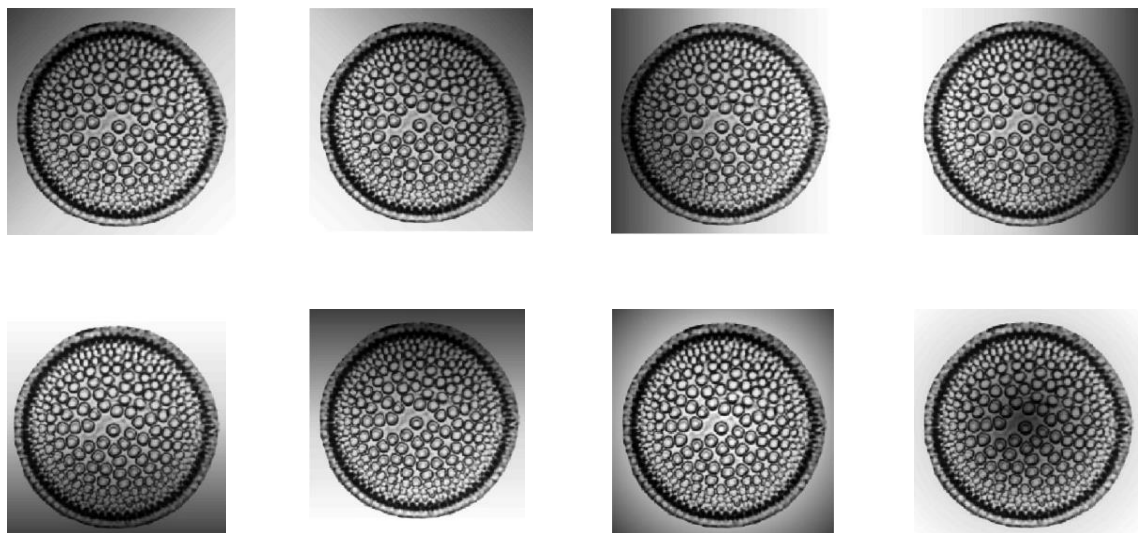


Figura 30. Ejemplos de la diatomea fósil *Actinocyclus ingens*, afectada por los diferentes patrones de iluminación no homogénea.

Las gráficas con diagramas de caja de la figura 31 muestran los resultados con un nivel de confianza de al menos 95.4%. Para todos los casos, las imágenes objetivo son imágenes de diatomeas limpias, tal como se muestra en la figura 12. Estas imágenes no están rotadas, no contienen cambios de escala y tampoco están afectadas por iluminación no homogénea.

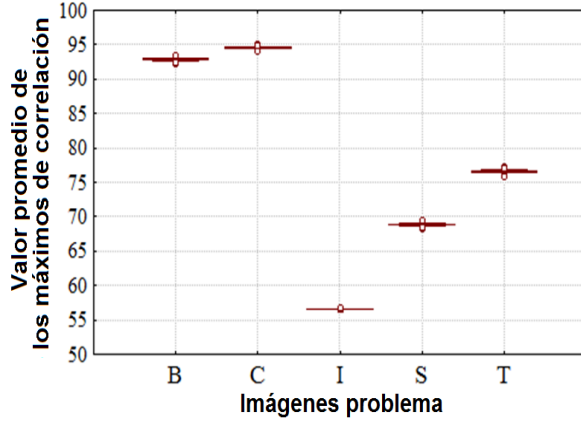
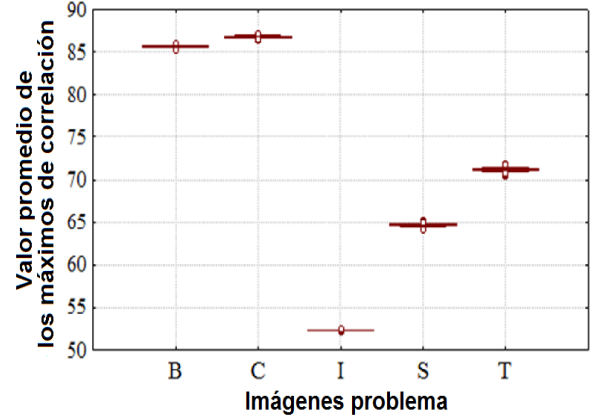
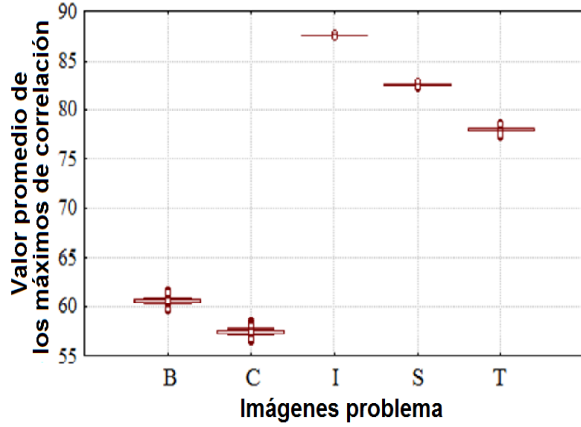
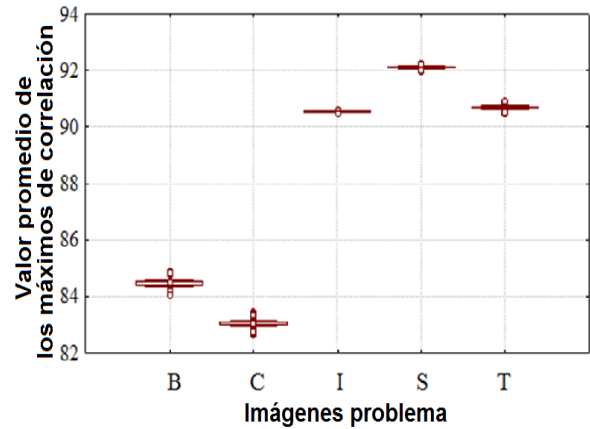
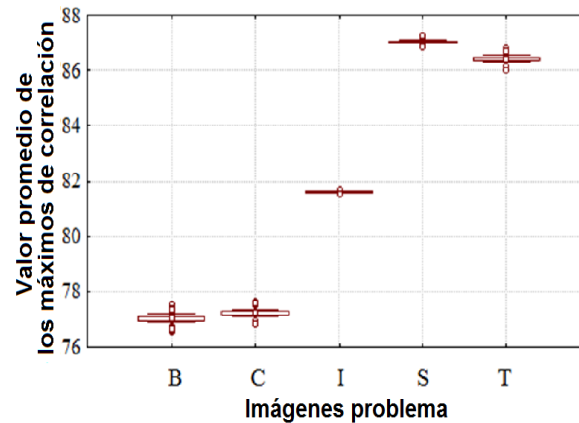
(a) Imagen objetivo: B, C_{NL} □ Media □ ±EE ±2EE ◦ Outliers(b) Imagen objetivo: C, C_{NL} □ Media □ ±EE ±2EE ◦ Outliers(c) Imagen objetivo: I, C_{NL} □ Media □ ±EE ±2EE ◦ Outliers(d) Imagen objetivo: S, C_{NL} □ Media □ ±EE ±2EE ◦ Outliers(e) Imagen objetivo: T, C_{NL} □ Media □ ±EE ±2EE ◦ Outliers

Figura 31. Reconocimiento de imágenes con iluminación no homogénea. (a) Filtro para diatomea B. (b) Filtro diatomea C. (c) Filtro diatomea I. (d) Filtro diatomea S. (e) Filtro diatomea T.

4.6 Reconocimiento de imágenes reales con rotación, escala, iluminación y fondo natural

Para demostrar que la metodología basada en la transformada de escala y técnicas de correlación no lineal con invariancia a distorsiones geométricas e iluminación, es efectiva para identificar imágenes en escenarios reales. Se desarrolló una prueba con imágenes reales con cinco variantes en su escenario real, con rotaciones naturales y con diferentes tamaños naturales para cinco diferentes especies de fitoplancton: *Acanthogonyaulax spinifera*, *Ornithocercu armata*, *Dinophysis hastate*, *Ornithocercus magnificus* y *Podolampas bipes* 1. Estas imágenes se muestran en la figura 32 etiquetadas como A, B, C, D y E, respectivamente, todas con un tamaño de 256×256 píxeles. Estas imágenes además fueron alteradas digitalmente en rotación desde 0 hasta 179° (180 rotaciones) y escaladas desde 81% hasta 100% (20 cambios de escala), generando 18,000 imágenes diferentes de cada especie, por lo cual, en esta base de datos tenemos 90,000 imágenes.

Las imágenes de la figura 32 incluyen su fondo natural, el cual en algunas de estas imágenes es más sucio que en otras, y cuando se hace la manipulación digital para cambio de escala y rotación, parte del fondo se pierde al amplificar la imagen (ya que se conserva el tamaño de 256×256 píxeles), y al reducir la imagen, se ganan espacios en blanco. Por lo anterior, en la manipulación digital de imágenes que conservan su fondo original, solo se hacen pruebas con cambios de escala desde 81% hasta 100%. Además se ha utilizado una preparación previa de la imagen, ya que ninguna de las esquinas de las imágenes de prueba contienen información relevante del objeto a identificar, se limita el fondo de la imagen a un círculo, entonces, a todas estas imágenes se les realiza una multiplicación punto a punto con un disco binario DB , dado por (Solorza y Álvarez-Borrego, 2014)

$$DB = \begin{cases} 1, & \text{si } d((C,C),(x,y)) \leq C, \\ 0, & \text{cualquier otro caso,} \end{cases} \quad (68)$$

donde $d((C,C),(x,y))$ es la distancia euclidiana entre el píxel central (C,C) y el punto (x,y) . En este caso particular este disco tiene un diámetro de 256 píxeles. En la figura 33 se ilustra el resultado de esta multiplicación.

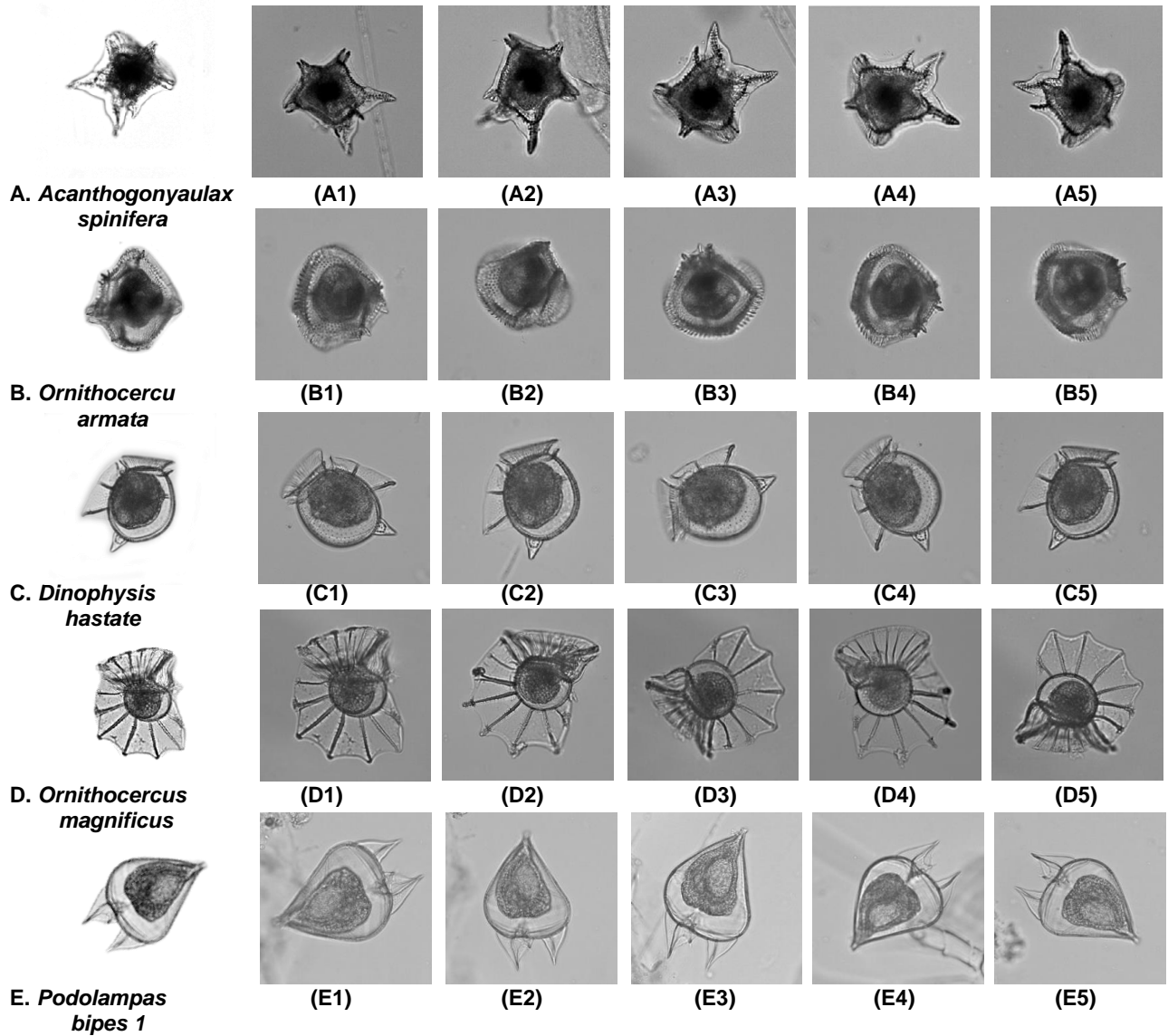


Figura 32. Imágenes reales de fitoplancton con rotación, escala, iluminación y fondo natural.

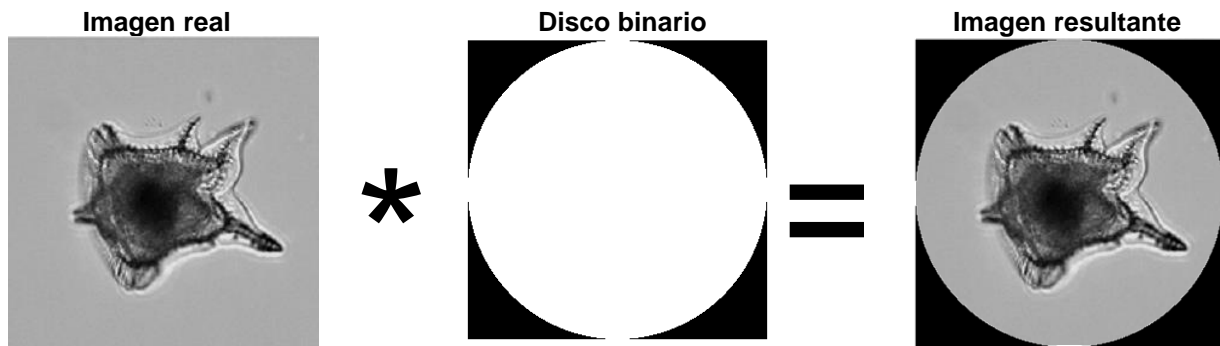


Figura 33. Resultado de la multiplicación de la imagen real por el disco binario, donde * representa a la multiplicación punto a punto.

A este tipo de imágenes no se les pueden aplicar los índices invariantes a rotación y escala debido a que la relación de área y volumen no permanece constante. Además se complica calcular el área y volumen, ya que la imagen contiene información en el fondo. Utilizar una razón entre las desviaciones estándar de las imágenes objetivo y problema es una alternativa que nos permite utilizar la correlación no lineal adaptativa (C_{AD}), ya que es un parámetro que tiene poca variabilidad entre imágenes de fitoplancton de la misma especie y además también es invariante a los cambios de rotación y escala. Esta razón está dada por,

$$Rz = \begin{cases} \frac{Ds_{IP}}{Ds_{IO}}, & \text{si } Ds_{IP} \leq Ds_{IO}, \\ \frac{Ds_{IO}}{Ds_{IP}}, & \text{si } Ds_{IO} < Ds_{IP}. \end{cases} \quad (69)$$

donde Ds_{IP} y Ds_{IO} son las desviaciones estándar de la imagen problema y de la imagen objetivo, respectivamente.

En este experimento se diseñaron filtros a partir de las imágenes sin ruido A, B, C, D y E de la figura 32, que serán tomadas como imágenes objetivo y a partir de estas se construirá la máscara binaria de anillos concéntricos siguiendo el mismo procedimiento de la figura 5. Las firmas objetivo generadas serán comparadas contra las firmas de las imágenes problema que forman el banco de 90,000 imágenes en sus escenarios reales. Como método de comparación entre firmas vectoriales se utilizaron las dos correlaciones no lineales C_{NL} y C_{AD} .

En las figuras 34, 35, 36, 37 y 38 se muestran los resultados de los datos estadísticos para los casos en que se utilizan como imagen objetivo, las imágenes de fitoplancton A, B, C, D y E de la figura 31 respectivamente. Estos datos son representados en diagramas de caja.

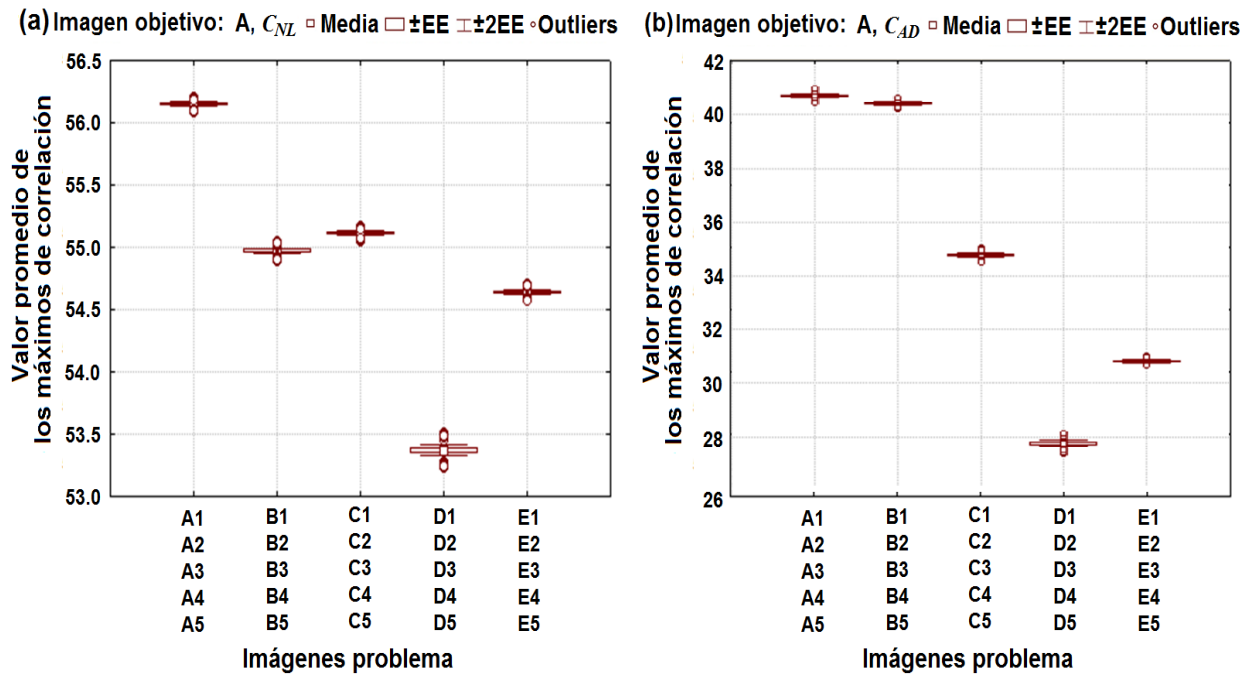


Figura 34. Resultados del valor promedio de los máximos de correlación utilizando imágenes en escenarios reales como imágenes problema y como imagen objetivo una imagen limpia de la especie de fitoplancton, A (*Acanthogonyaulax spinifera*). (a) Utilizando la correlación no lineal (C_{NL}). (b) Utilizando la correlación no lineal adaptativa (C_{AD}).

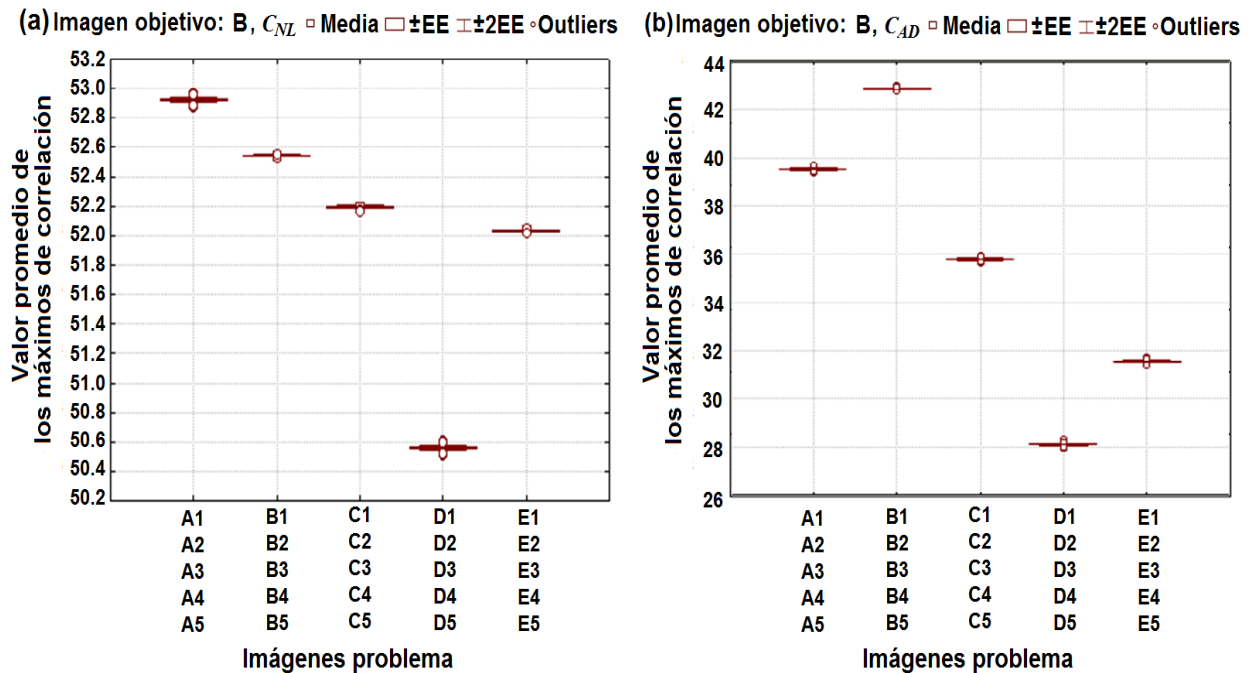


Figura 35. Resultados del valor promedio de los máximos de correlación utilizando imágenes en escenarios reales como imágenes problema y como imagen objetivo una imagen limpia de la especie de fitoplancton, B (*Ornithocercu armata*). (a) Utilizando la correlación no lineal (C_{NL}). (b) Utilizando la correlación no lineal adaptativa (C_{AD}).

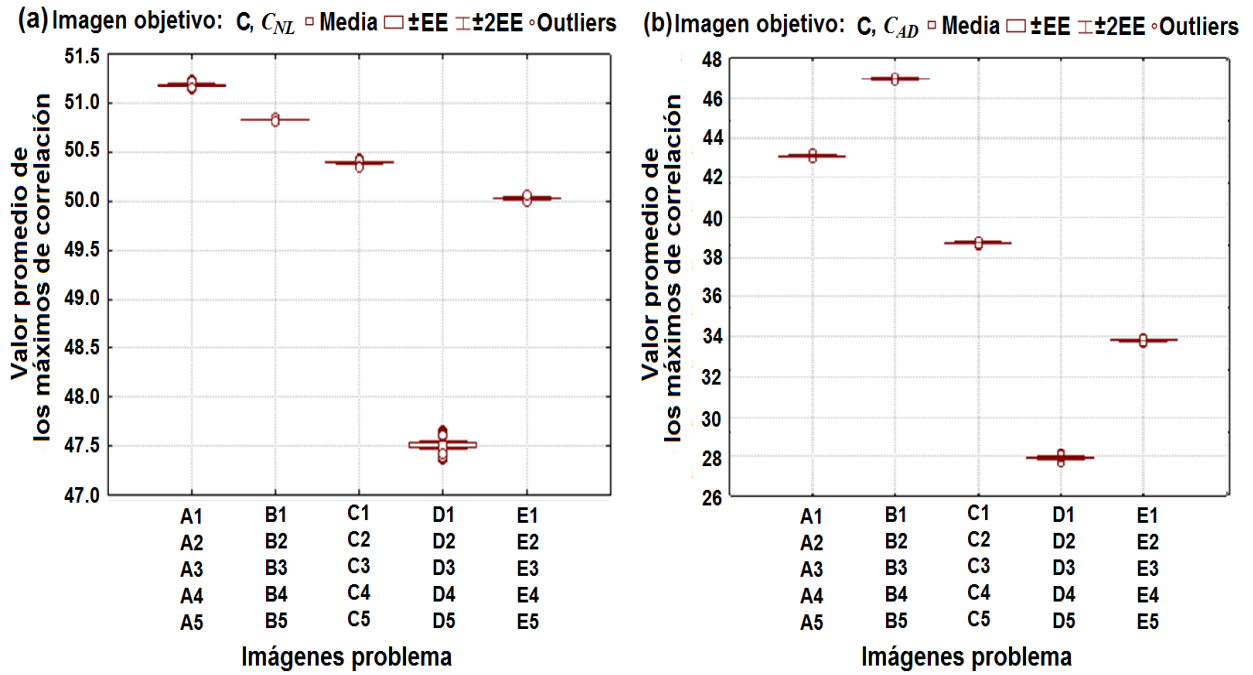


Figura 36. Resultados del valor promedio de los máximos de correlación utilizando imágenes en escenarios reales como imágenes problema y como imagen objetivo una imagen limpia de la especie de fitoplancton, C (*Dinophysis hastate*). (a) Utilizando la correlación no lineal (C_{NL}). (b) Utilizando la correlación no lineal adaptativa (C_{AD}).

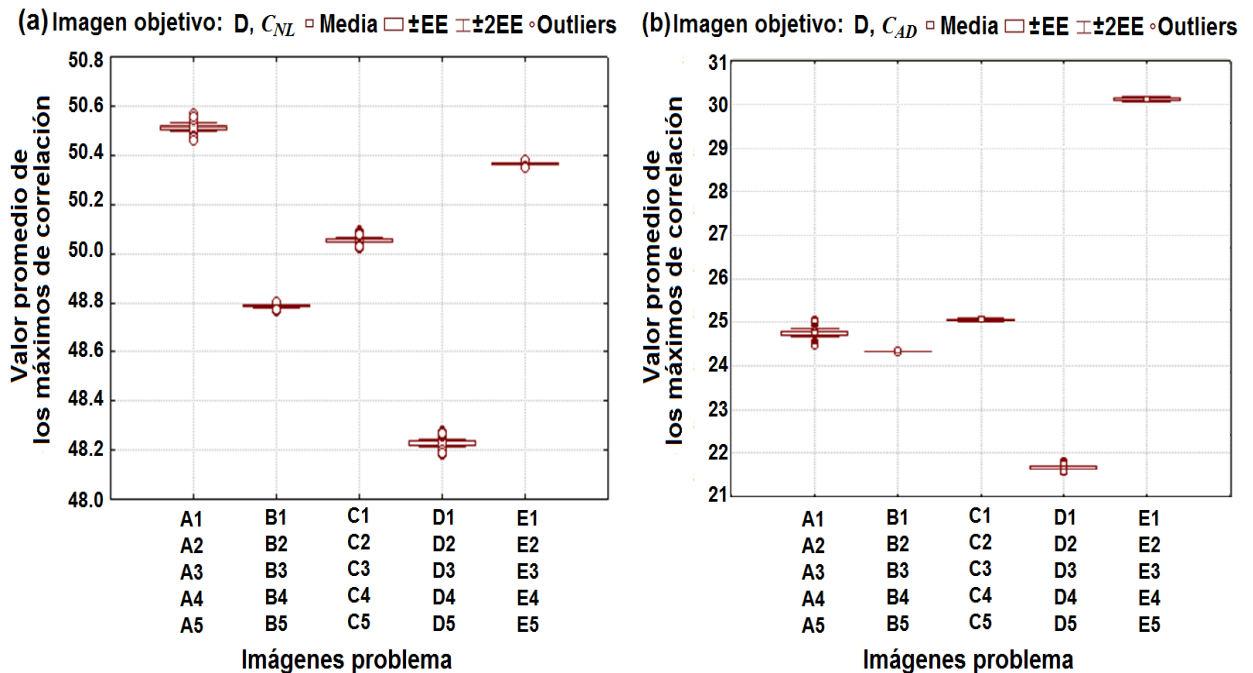


Figura 37. Resultados del valor promedio de los máximos de correlación utilizando imágenes en escenarios reales como imágenes problema y como imagen objetivo una imagen limpia de la especie de fitoplancton, D (*Ornithocercus magnificus*). (a) Utilizando la correlación no lineal (C_{NL}). (b) Utilizando la correlación no lineal adaptativa (C_{AD}).

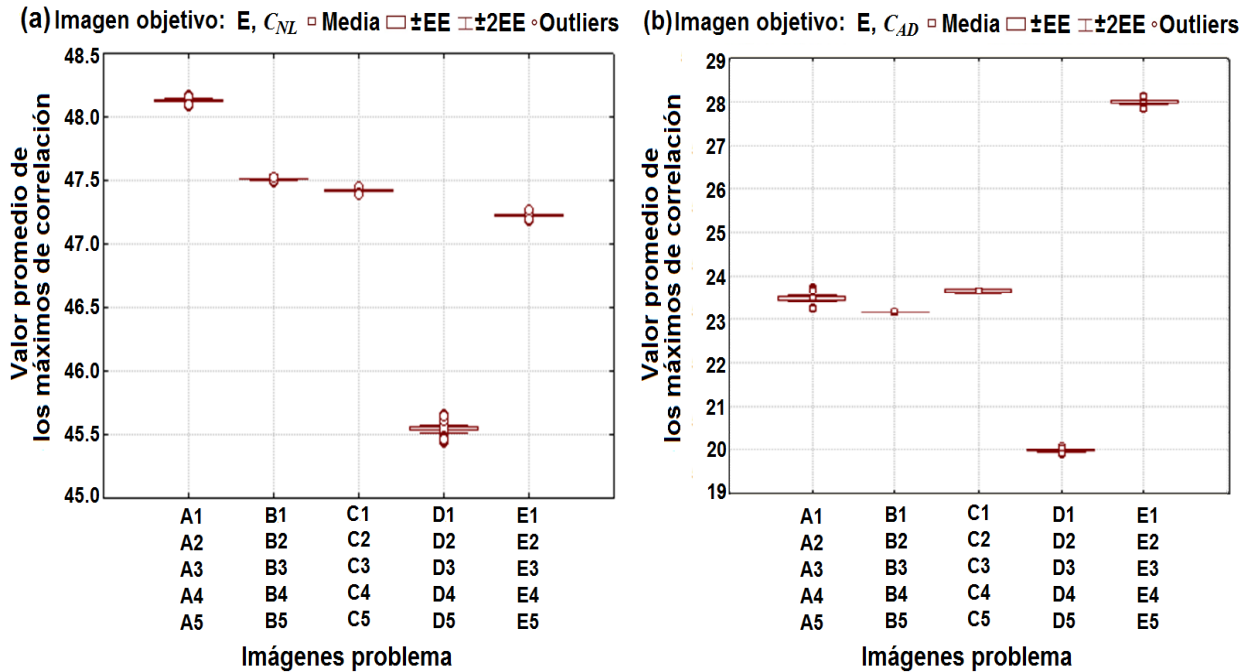


Figura 38. Resultados del valor promedio de los máximos de correlación utilizando imágenes en escenarios reales como imágenes problema y como imagen objetivo una imagen limpia de la especie de fitoplancton, E (*Podolampas bipes* 1). (a) Utilizando la correlación no lineal (C_{NL}). (b) Utilizando la correlación no lineal adaptativa (C_{AD}).

En las figuras 34, 35, 36, 37 y 38 se puede apreciar claramente una identificación exitosa, para cada una de las imágenes con rotación, escala, iluminación y fondo natural, ya que no hay ningún traslape entre los conjuntos de las imágenes objetivo con el resto de los diagramas de caja correspondientes a las imágenes problema. Con la incorporación de la razón entre las desviaciones estándar de las imágenes dada por la ecuación (69) a la correlación no lineal adaptativa (C_{AD}), se lograron buenos resultados. Sin embargo estos no son mejores que los obtenidos usando una correlación no lineal (C_{NL}). Por ejemplo, en la gráfica de la figura 34a, los resultados para la imagen A están libres de traslapes utilizando C_{NL} , mientras que, en la figura 34b se utiliza C_{AD} y no se logra mejorar el resultado, ya que hay un traslape entre los diagramas de caja de las imágenes A y B. En cambio, en las figuras 35b y 36b se puede apreciar un mejor rendimiento con la correlación C_{AD} , ya que sus resultados de máximos de correlación más altos, son los que corresponden a la imagen objetivo.

Capítulo 5. Conclusiones

La metodología propuesta cumple con los objetivos de esta tesis ya que se logró la identificación de objetos con invariancia a posición, rotación y escala, aun cuando las imágenes se encuentren incompletas, ante iluminaciones no uniformes o inmersas en ruido. En resumen, esto se logró aplicando la transformada de escala a cada una de las imágenes, la cual se obtuvo vía la transformada rápida de Fourier, obteniendo así la invariancia a escala. Las máscaras binarias de anillos concéntricos, se construyen a partir del módulo de la transformada de escala, consiguiendo así, representar a las imágenes utilizando firmas vectoriales invariantes a posición rotación y escala. Finalmente, como método de comparación entre firmas vectoriales se utilizan correlaciones no lineales.

El módulo de la transformada de Fourier es invariante a posición, y cuando se realiza la suma del contenido de cada uno de los anillos resultantes de la multiplicación punto a punto del módulo de la transformada de Fourier de una imagen con una máscara binaria de anillos concéntricos, se consigue una firma vectorial invariante a la rotación y a la posición del objeto en la imagen, sin embargo, esta firma carece de invariancia a cambios de escala. De manera análoga, al trabajar con la transformada de escala 2D en su forma separable, cuando se realiza la multiplicación punto a punto de una máscara binaria con el módulo de la transformada de escala, se logra obtener como resultado un conjunto de anillos concéntricos, con los que se realizan muestras radiales del módulo de la transformada de escala, y al sumar el contenido de cada uno de los anillos se crea una firma vectorial invariante a la rotación y a los cambios de escala del objeto en la imagen, pero que carece de invariancia a posición.

Para obtener la transformada de escala, vía la transformada de Fourier, es necesario hacer un mapeo exponencial de la imagen antes de aplicar la transformada de Fourier. Por lo tanto, aunque el módulo de la transformada de Fourier es invariante a posición, el mapeo exponencial no lo es. Por esta causa, la invariancia a posición del objeto en la

imagen se solucionó con un pre-procesamiento que centra el objeto en la imagen problema mediante su centro de masa, de esta manera se obtuvo una transformada de escala que es invariante a posición y escala.

Previo a este trabajo las máscaras binarias adaptativas de círculos concéntricos se obtenían a partir de la transformada de Fourier de una imagen, en este trabajo las máscaras se obtuvieron a partir de las pendientes de un transecto del módulo de la transformada de escala. Estas máscaras, lograron tomar muestras del módulo de la transformada de escala, lo que ayudó a formar firmas vectoriales con invariancia a posición, rotación y escala, capaces de representar al objeto dentro de la imagen. El uso de firmas vectoriales logra un éxito en cuanto al tiempo de cómputo, ya que las imágenes que se utilizaron en las pruebas tienen un tamaño de al menos 256×256 píxeles, lo cual quiere decir que cada una de estas imágenes contiene 65,536 datos, que logran ser representados con una cantidad mucho menor de datos colocados en una firma vectorial unidimensional, agilizando así la comparación y la identificación automática de imágenes.

Como método de comparación entre firmas vectoriales se utilizaron correlaciones no lineales en su versión original y en su versión adaptativa; el valor óptimo de $k=0.1$ para estas correlaciones se obtuvo experimentalmente, seleccionando el valor de no linealidad k que logro un valor PCE de mayor magnitud. Se logró modificar la versión adaptativa para ser adaptable a cambios de rotación y a cambios de escala mediante el uso de índices que relacionan el área y el volumen de la imagen objetivo y de la imagen problema.

En la identificación de todas las imágenes de prueba, la correlación no lineal adaptativa (CAD), obtuvo resultados con mayor nivel de confianza que la CNL . Sin embargo, se encontraron casos en donde la nueva razón entre índices no se pudo implementar debido a que en esas situaciones se altera la relación entre el área y el volumen de las imágenes. Estos casos fueron dos: el primer caso fue cuando se incluyeron fragmentos en las imágenes problema en donde la correlación no lineal original (CNL) fue la mejor opción a utilizar como método de comparación, y el segundo fue cuando las imágenes

problema fueron expuestas a iluminación no homogénea, en donde la correlación no lineal original (C_{NL}) logró al menos un 95.4% de confiabilidad. En todos los demás casos la correlación no lineal adaptativa (C_{AD}) logró incrementar el nivel de confiabilidad con respecto al obtenido con la correlación C_{NL} .

Se realizaron experimentos en donde se utiliza la métrica de capacidad de discriminación para demostrar que imágenes contaminadas con ruido gaussiano de media cero y niveles de varianza hasta de 1.6 pueden ser identificadas sin problemas con un nivel de confianza del 95.4%. Para imágenes contaminadas con ruido impulsivo se logra la identificación de la imagen aun cuando el 40% de los píxeles que la componen se encuentran afectados.

Se demostró la robustez de la metodología ante iluminaciones no homogéneas, exponiendo cinco diatomeas fósiles a ocho diferentes tipos de iluminación no homogénea, en donde esta nueva metodología logró identificar con éxito cada una de estas cinco diatomeas con una confiabilidad de al menos el 95.4%.

Como prueba final, se utilizó un banco de imágenes en donde se seleccionaron cinco especies diferentes de fitoplancton, cada una en situaciones reales y con cinco variantes en su escenario real, con rotaciones naturales y con diferentes tamaños naturales. Además, estas imágenes se alteraron digitalmente para generar intencionalmente una variedad más extensa de distorsiones geométricas causadas por cambios debidos a rotación y escala. Esto hace de esta prueba el escenario más representativo de una situación real. La correlación C_{NL} fue capaz de clasificar con una confiabilidad de al menos 99% cada una de las cinco especies. Al repetir el experimento con la correlación C_{AD} , el uso de una razón entre índices calculados a partir del cociente entre el área y el volumen de las imágenes de entrada, para dar adaptabilidad a cambios de rotación y de escala a la correlación no lineal, no funcionó en estos casos. La solución que dio la oportunidad de utilizar la C_{AD} se basa en utilizar una razón entre las desviaciones estándar de las imágenes de entrada. De esta manera fue posible aplicar con éxito la C_{AD} y también se logró clasificar a las cinco especies con una confiabilidad de al menos 99%.

Lista de referencias

- Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für mikroskopische Anatomie*, 9(1), 413-420.
- Álvarez-Borrego, J., Solorza Calderón, S., and Bueno-Ibarra, M. A. (2013). Invariant correlation to position and rotation using a binary mask applied to binary and gray images. *Optics Communications*, 294, 105-117.
- Antonino-Daviu, J., Aviyente, S., Strangas, E. G., and Riera-Guasp, M. (2013). Scale invariant feature extraction algorithm for the automatic diagnosis of rotor asymmetries in induction motors. *Industrial Informatics, IEEE Transactions on*, 9(1), 100-108.
- Asselin, D., Arsenault, H. H. and Prévost, D. (1994). Optical circular sampling system for translation and rotation invariant pattern recognition. *Optical Communications*, 110, 507-513.
- Bracewell, R.N., (2000). *The Fourier Transform and its Applications*. New York, McGraw Hill, p. 254.
- Brigham, E. O., (1988). *The Fast Fourier Transform and its Applications*. Englewood Cliffs, New Jersey, Prentice Hall, p. 32, p. 50-51, p. 240.
- Brousil, J. K., and Smith, D. R. (1967). A threshold logic network for shape invariance. *Electronic Computers, IEEE Transactions on*, (6), 818-828.
- Casasent, D., and Psaltis, D. (1976). Position, rotation, and scale invariant optical correlation. *Applied optics*, 15(7), 1795-1799.
- Casasent, D., and Psaltis, D. (1977). New optical transforms for pattern recognition. *Proceedings of the IEEE*, 65(1), 77-84.
- Chaparro-Magallanes, G. (2011). Correlación invariante mediante el uso de firmas unidimensionales. Tesis de maestría en ciencias. Centro de Investigación Científica y de Educación Superior de Ensenada, Ensenada, B.C. 63 p.
- Cohen, L. (1993). The scale representation. *Signal Processing, IEEE Transactions on*, 41(12), 3275-3292.
- Combet, F., and Zimroz, R. (2009). A new method for the estimation of the instantaneous speed relative fluctuation in a vibration signal based on the short time scale transform. *Mechanical Systems and Signal Processing*, 23(4), 1382-1397.
- Cooley, J. W., and Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Mathematics of computation*, 19(90), 297-301.

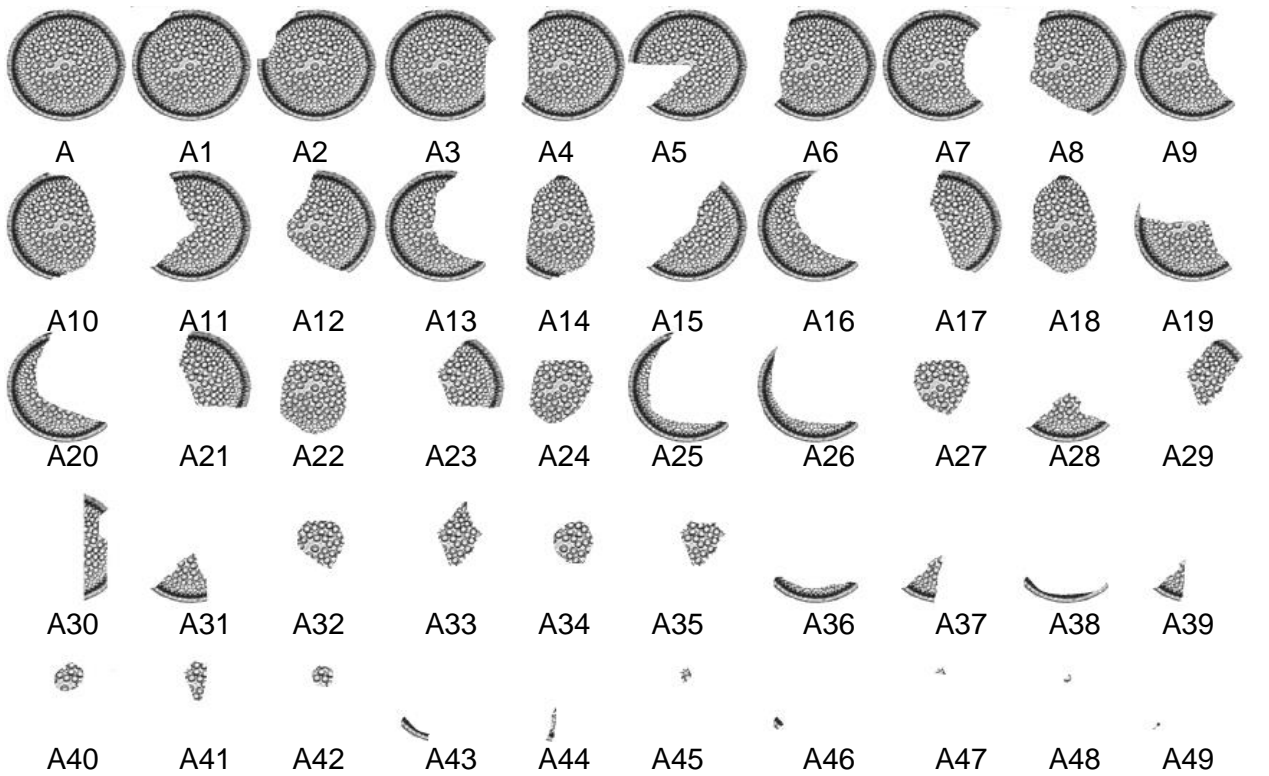
- Coronel-Beltrán, Á., and Álvarez-Borrego, J. (2010). Comparative analysis between different font types and letter styles using a nonlinear invariant digital correlation. *Journal of Modern Optics*, 57(1), 58-64.
- Cristobal, G., and Cohen, L. (1997). Scale analysis in images through invariant operators. En *Proc. IEEE Int. Conf. on Image Processing* (Vol. 2).
- Cristbal, G., Cuesta, J., and Cohen, L. (1998). Image filtering and denoising through the scale transform. In *Time-Frequency and Time-Scale Analysis, 1998. Proceedings of the IEEE-SP International Symposium on* (pp. 617-620).
- De Sena, A., and Rocchesso, D. (2007). A fast Mellin and scale transform. *EURASIP Journal on Advances in Signal Processing*, 2007.
- Djurdjanovic, D., Williams, W. J., and Koh, C. K. (1999). Discrete implementations of scale transform. En *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*. International Society for Optics and Photonics, 522-533.
- Esparza, M.A., (1999) "Variabilidad de la comunidad de diatomeas en los sedimentos de la Cuenca de San Lázaro, Baja California Sur, México," Tesis de maestría en ciencias. Centro de Investigación Científica y de Educación Superior de Ensenada, Ensenada, B.C., 113 p
- Fares, A., Bouzid, A. and Hamdi M. (2000). Rotation invariance using diffraction pattern sampling in optical pattern recognition. *Journal of Microwaves and Optoelectronics*, 2(2), 33-39.
- Fimbres-Castro, C., Álvarez-Borrego, J., and Bueno-Ibarra, M. A. (2012). Invariant nonlinear correlation and spectral index for diatoms recognition. *Optical Engineering*, 51(4), 047201-1-047201-11.
- Ganotra, D., Joseph, J., and Singh, K. (2004). Modified geometry of ring-wedge detector for sampling Fourier transform of fingerprints for classification using neural networks. *Optics and lasers in engineering*, 42(2), 167-177.
- George, N., Thomasson, J. T. and Spendel A. (1972). Photodetector light pattern detector. *US patent*, 3,689,772.
- Gonzalez, R. C., y Woods, R. E. (2002). *Digital Image Processing 2ed*. New Jersey, Prentice Hall. p. 793
- Gonzalez, R. C., y Woods, R. E. (2007). *Digital Image Processing 3ed*. New Jersey, Prentice Hall. p. 976.
- Goodman J. W. (2005). *Introduction to Fourier Optics*. (3rd ed.) Greenwood Village, CO. Roberts and Company Publishers. p. 393

- Hajós, M. (1976). Upper eocene and lower oligocene diatomaceae, archaemonadaceae and silicoflagellates in southwestern pacific sediments. *DSSP, leg, 29*, 817-883.
- Harley, J. B., and Moura, J. M. (2012). Scale transform signal processing for optimal ultrasonic temperature compensation. *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on*, 59(10).
- Holzapfel, A., and Stylianou, Y. (2011). Scale transform in rhythmic similarity of music. *Audio, Speech, and Language Processing, IEEE Transactions on*, 19(1), 176-185.
- Horner, J. L., and Gianino, P. D. (1984). Phase-only matched filtering. *Applied optics*, 23(6), 812-816.
- Hsu, H. P. (1998). *Análisis de Fourier*, México, Prentice Hall. p. 274
- Kumar, B. V. K., and Hassebrook, L. (1990). Performance measures for correlation filters. *Applied optics*, 29(20), 2997-3006.
- Lendaris, G. G. and Stanley, G. L. (1965). A optical logical self-organizing recognition system, in *Optical and Electro-Optical information Processing*. Cambridge, M.I.T. Press.
- Lendaris, G. G. and Stanley, G. L. (1970). Diffraction-pattern sampling for automatic pattern recognition. *Proceedings of the IEEE*, 58(2), 198-216.
- Lerma Aragón, J. R., and Álvarez-Borrego, J. (2009). Vectorial signatures for invariant recognition of position, rotation and scale pattern recognition. *Journal of Modern Optics*, 56(14), 1598-1606.
- Maidana, N. I. (2013). El test de diatomeas en el diagnóstico de muerte por sumersión. *Acta Nova*, 6(1-2), 70-81.
- Mertins, A., and Rademacher, J. (2005). Vocal tract length invariant features for automatic speech recognition. En *Automatic Speech Recognition and Understanding, 2005 IEEE Workshop on* (pp. 308-312). IEEE.
- Mouriño-Pérez, R. R. 1999. Correlación óptica a color para la identificación de diferentes estadios de *Vibrio cholerae* 01. Tesis de doctorado en ciencias. Centro de Investigación Científica y de Educación Superior de Ensenada, Ensenada, B.C., 120 p.
- Jain, A. K., and Feng, J. (2011). Latent fingerprint matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(1), 88-100.
- Javidi, B. (1990). Comparison on nonlinear joint transform correlator and nonlinear matched filter based correlator. *Optical Communications*, 75(1), 8-13.
- Javidi, B. (2002). *Image recognition and classification algorithms, systems, and applications*. New York, NY, Marcel Dekker.

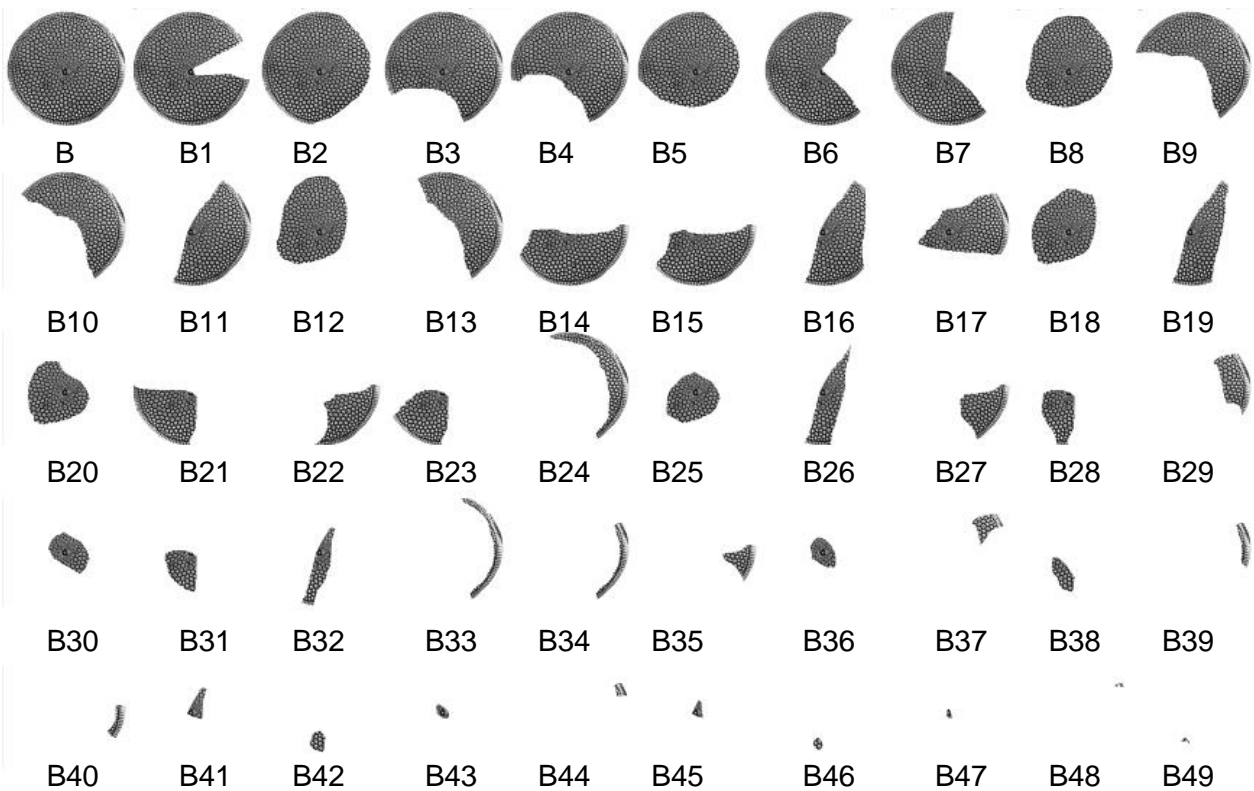
- Kasdan, H. L. (1979). Industrial applications of diffraction pattern sampling. *Optical Engineering*, 18(5), 185496-185496.
- Kruger, R. P., Thompson, W. B., and Turner, A. F. (1974). Computer diagnosis of pneumoconiosis. *Systems, Man and Cybernetics, IEEE Transactions on*, (1), 40-49.
- Kumar, B. V. K., and Hassebrook, L. (1990). Performance measures for correlation filters. *Applied optics*, 29(20), 2997-3006.
- Latha, P., Meffiya, N., and Dayana, B. D. (2015). Robust Face Recognition for Uncontrolled Environment. *International Journal for Innovative Research in Science and Technology*, 1(9), 133-141.
- Pech-Pacheco, J. L., Alvarez-Borrego, J., Cristóbal, G., and Keil, M. 2003. Automatic object identification irrespective to geometric changes, *Optical Engineering*, Vol. 42, No. 2, pp 551-559
- Rodríguez, J. (2005). La estructura de tamaños del plancton: un tópico interdisciplinar y Margalefiano. *Revista Ecosistemas*, 14(1), 1-11.
- Robbins, G. M., and Huang, T. S. (1972). Inverse filtering for linear shift-variant imaging systems. *Proceedings of the IEEE*, 60(7), 862-872.
- Schwartz, E. L. (1977). Afferent geometry in the primate visual cortex and the generation of neuronal trigger features. *Biological Cybernetics*, 28(1), 1-14.
- Solís-Ventura, A. (2012). *Correlación no lineal invariante a posición y rotación utilizando máscaras adaptativas binarias*. Tesis de maestría en ciencias. Centro de investigación científica y de educación superior de Ensenada, Ensenada, B.C. 64 p.
- Solorza Calderón, S. and Álvarez Borrego, J. (2010). Digital system of invariant correlation to position and rotation. *Optical Communications*, 283(19), 3613-3630.
- Solorza Calderón, S., Álvarez Borrego, J. and Chaparro Magallanes, G. (2012). Pattern Recognition of Digital Images by One-Dimensional Signatures. En: Salih Mohammed, *Fourier Transform - Signal Processing*. Rijeka, Croatia, InTech.
- Solorza-Calderón, S. and Álvarez-Borrego, J. (2013). Pattern Recognition by Bessel Mask and One-Dimensional Signatures. En: Darrell B. Vincent, *Pattern Recognition: Practices, Perspectives and Challenges*. United States, Nova Science Publishers, Inc.
- Vander Lugt, A. (1964). Signal detection by complex spatial filtering. *Information Theory, IEEE Transactions on*, 10(2), 139-145.
- Weiman, C. F. y Chaikin, G. (1979). Logarithmic spiral grids for image processing and display. *Computer Graphics and Image Processing*, 11(3), 197-226.

- Yaroslavsky L. P. (1993), The theory of optimal methods for localization of objects in pictures. En: XXXII E. Wolf, *Progress in Optics*. Amsterdam, Elsevier Science Publishers.
- Zalubas, E. J., and Williams, W. J. (1995, May). Discrete scale transform for signal analysis. En *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on* (Vol. 3, pp. 1557-1560). IEEE.
- Zwicke, P. E., and Kiss, I. (1983). A new implementation of the Mellin transform and its application to radar classification of ships. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2), 191-199.

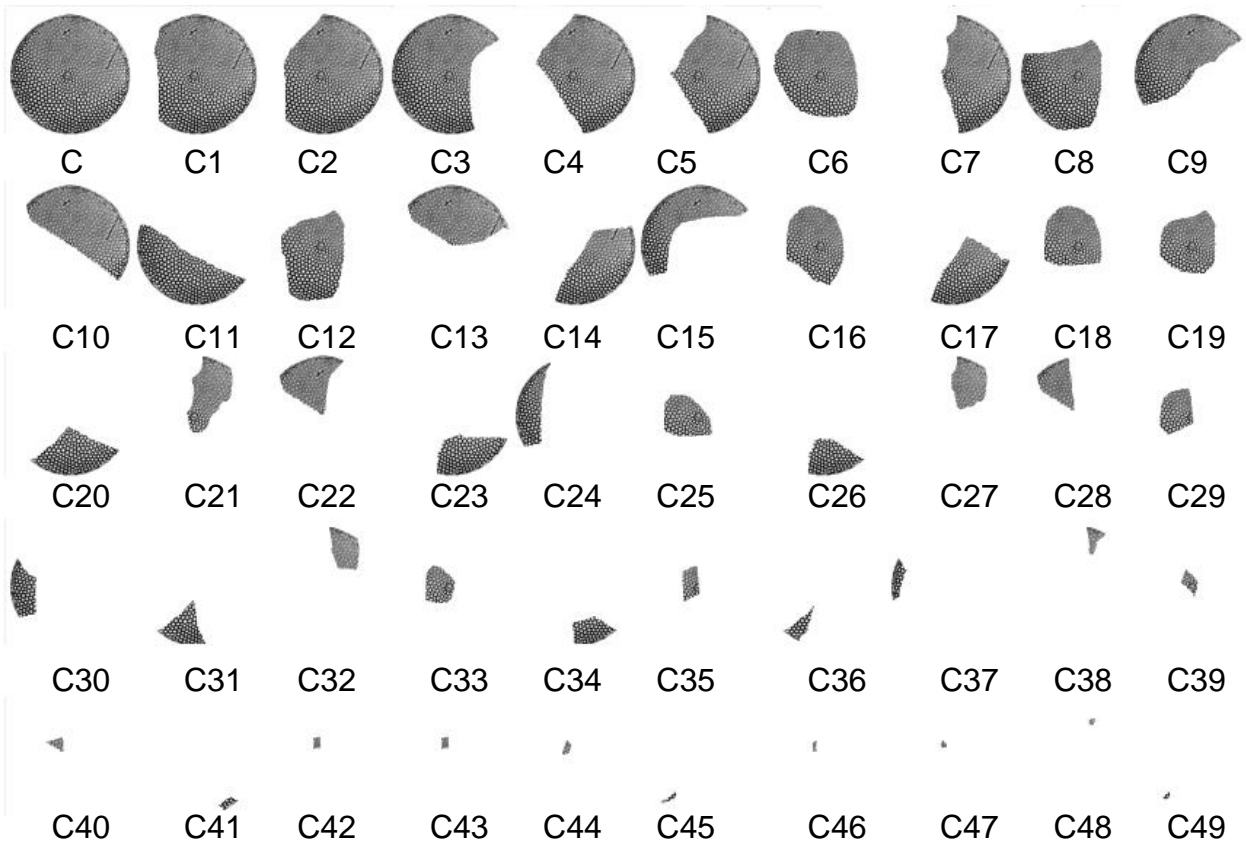
Apéndice A



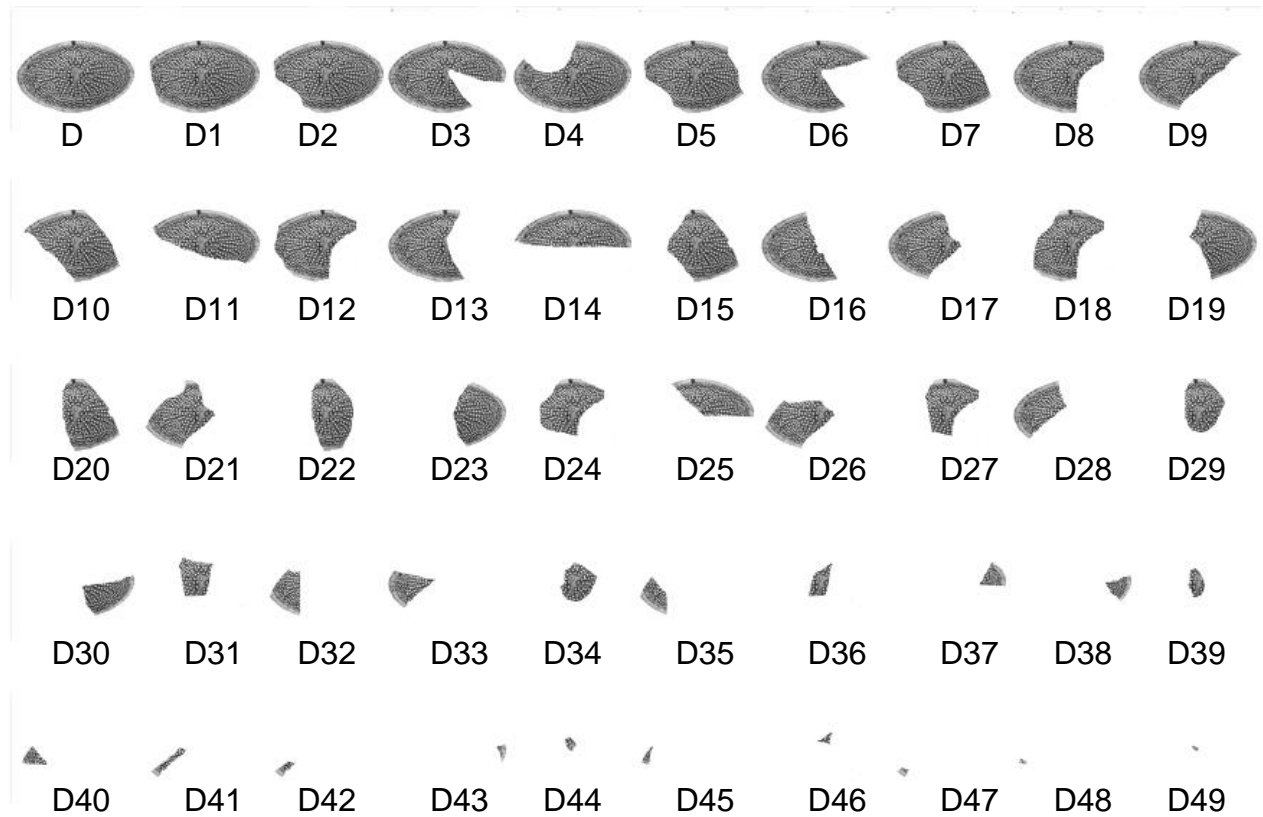
Diatomea fósil A y sus 49 fragmentos.



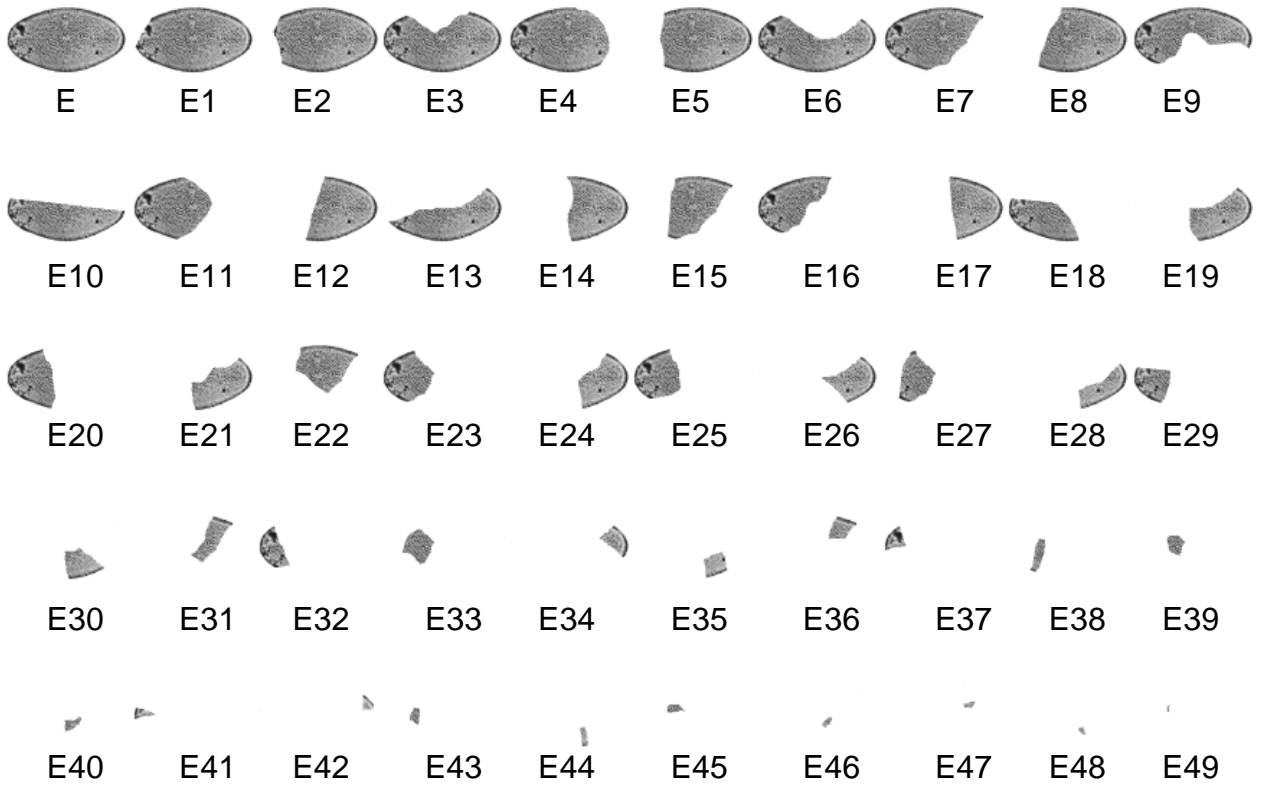
Diatomea fósil B y sus 49 fragmentos.



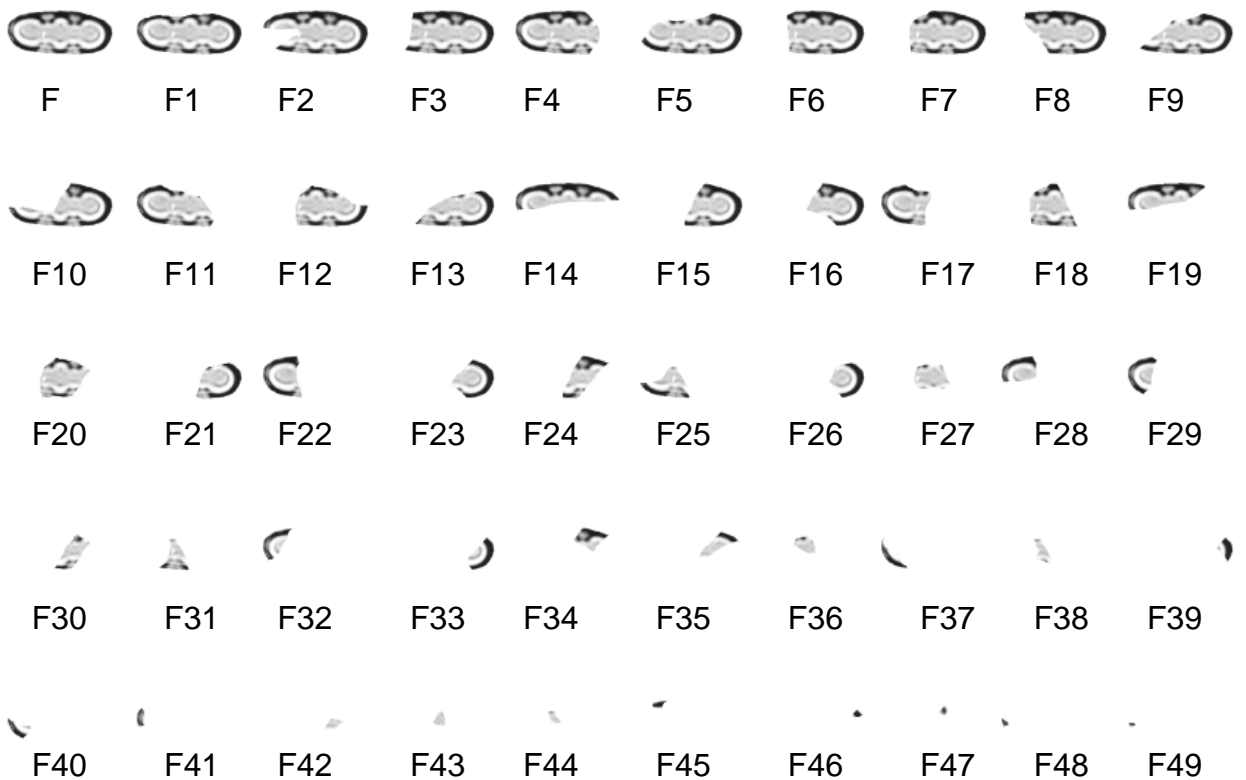
Diatomea fósil C y sus 49 fragmentos.



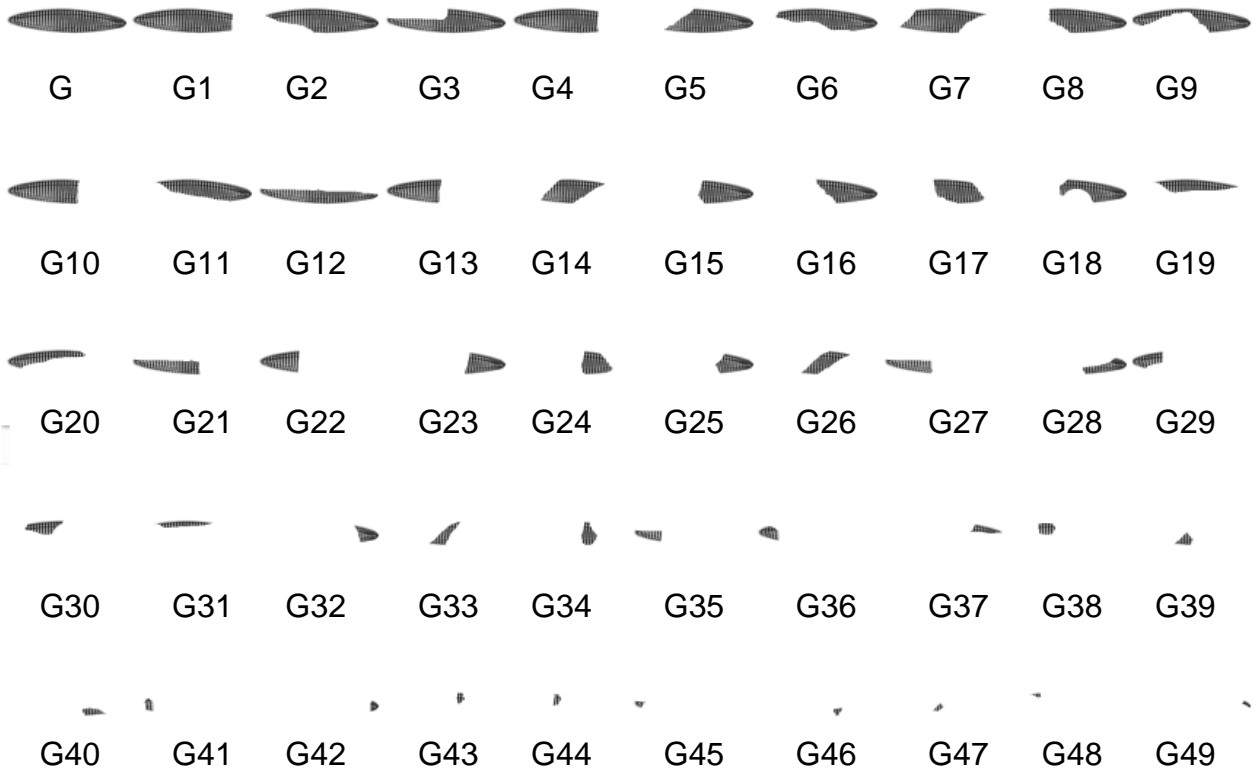
Diatomea fósil D y sus 49 fragmentos.



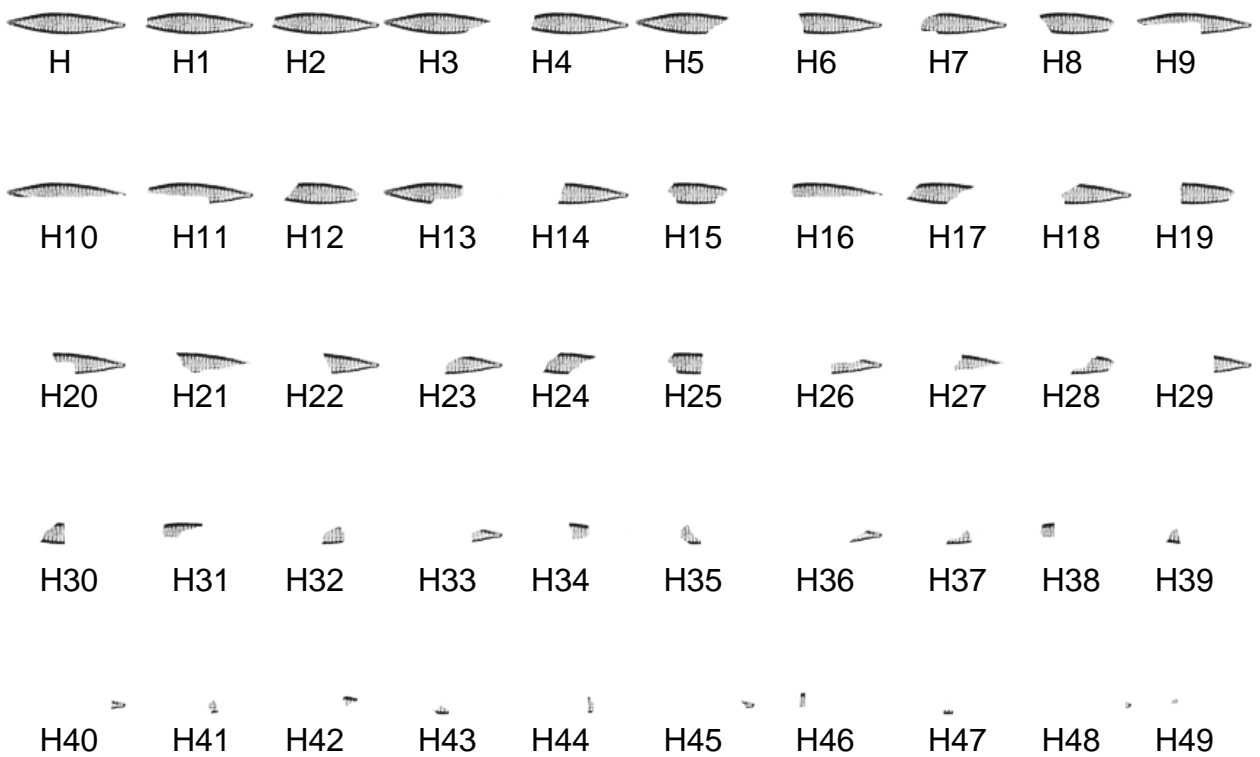
Diatomea fósil E y sus 49 fragmentos.



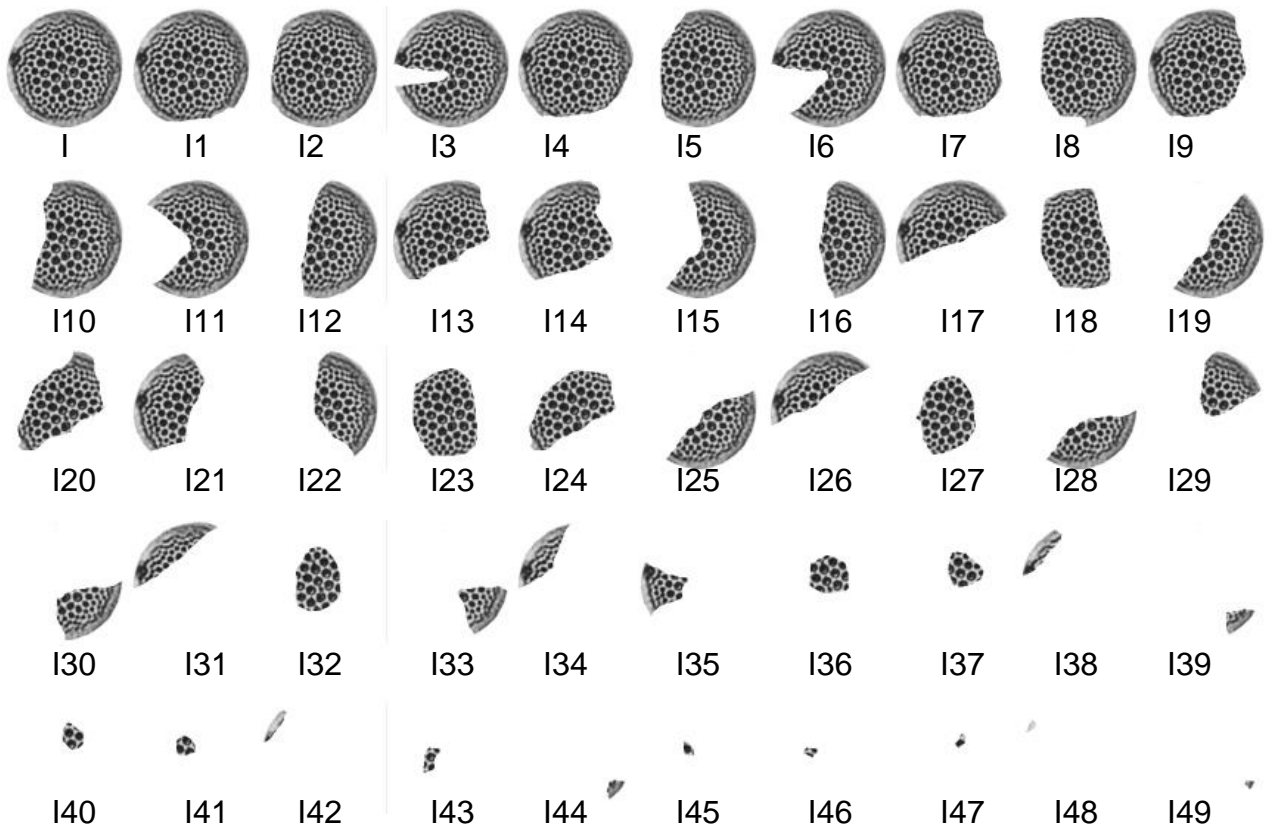
Diatomea fósil F y sus 49 fragmentos.



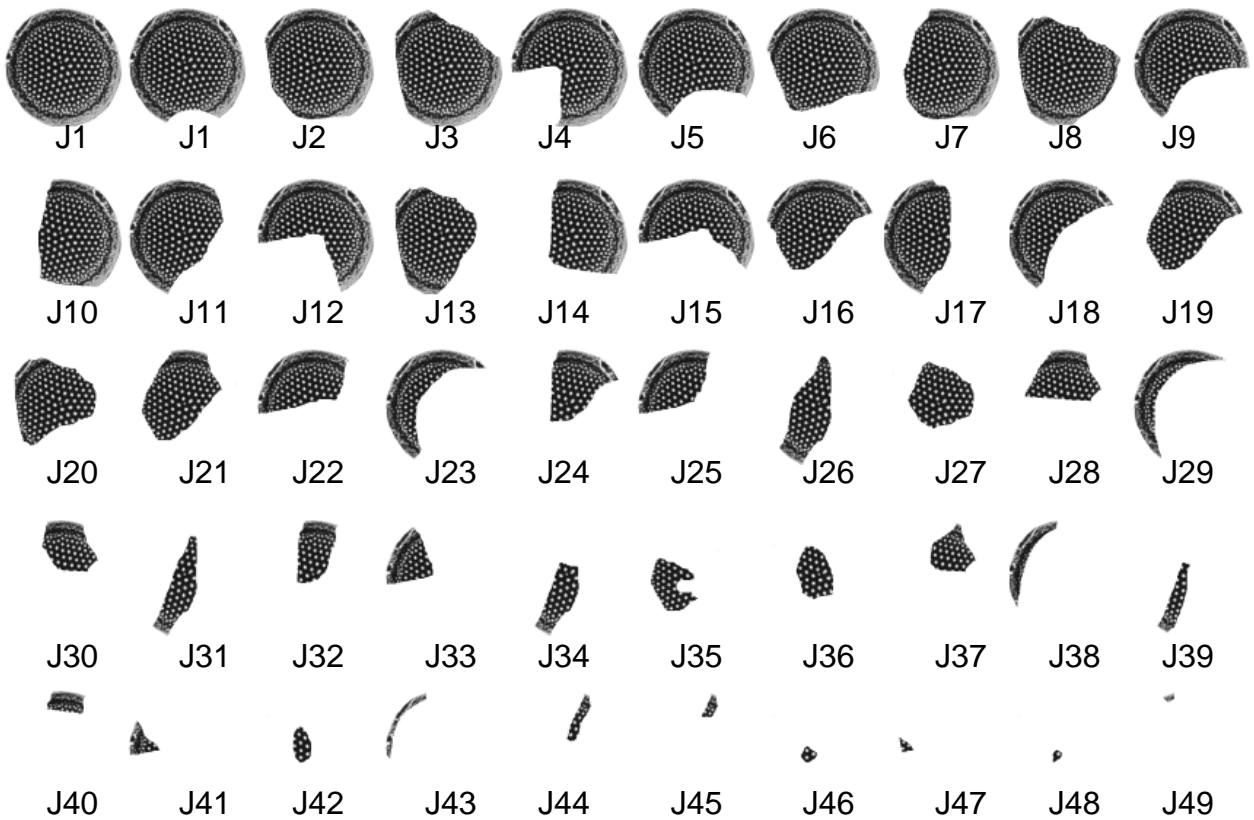
Diatomea fósil G y sus 49 fragmentos.



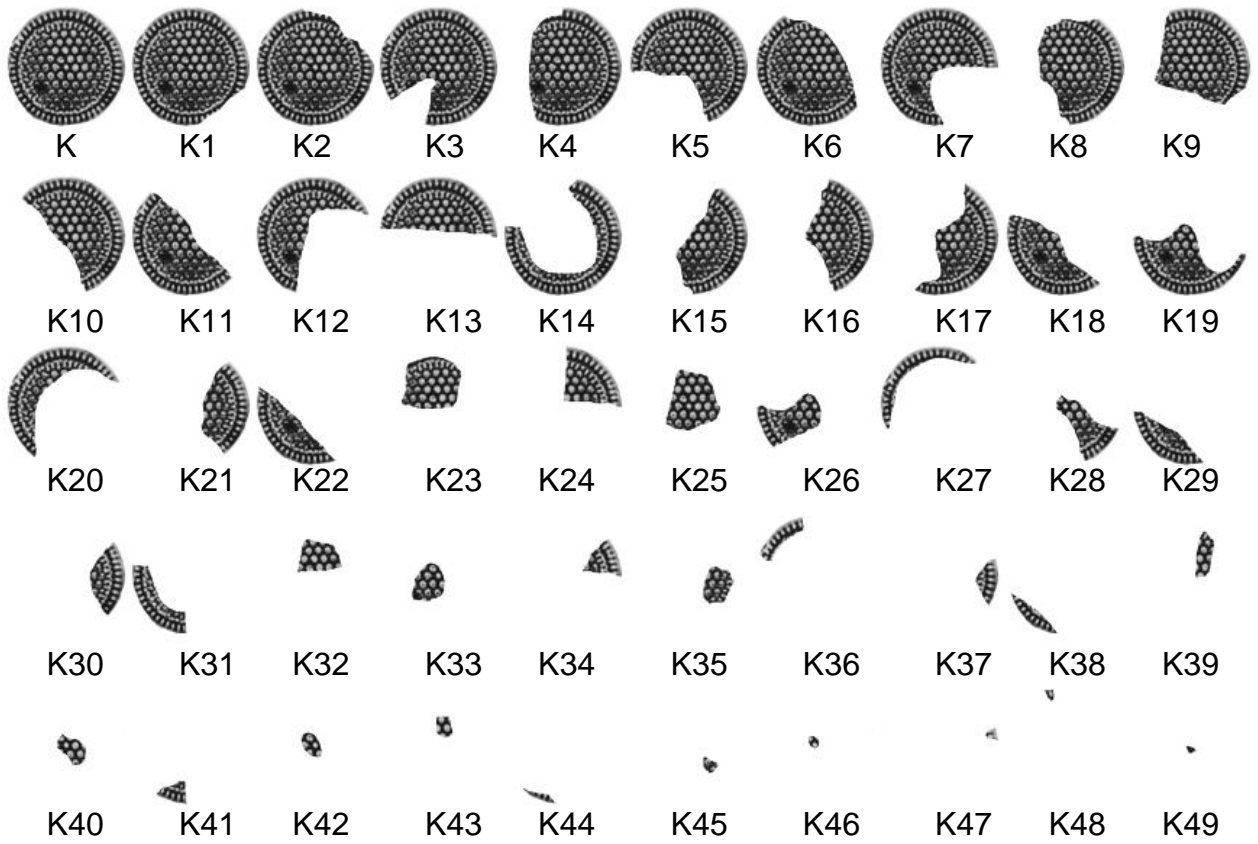
Diatomea fósil H y sus 49 fragmentos.



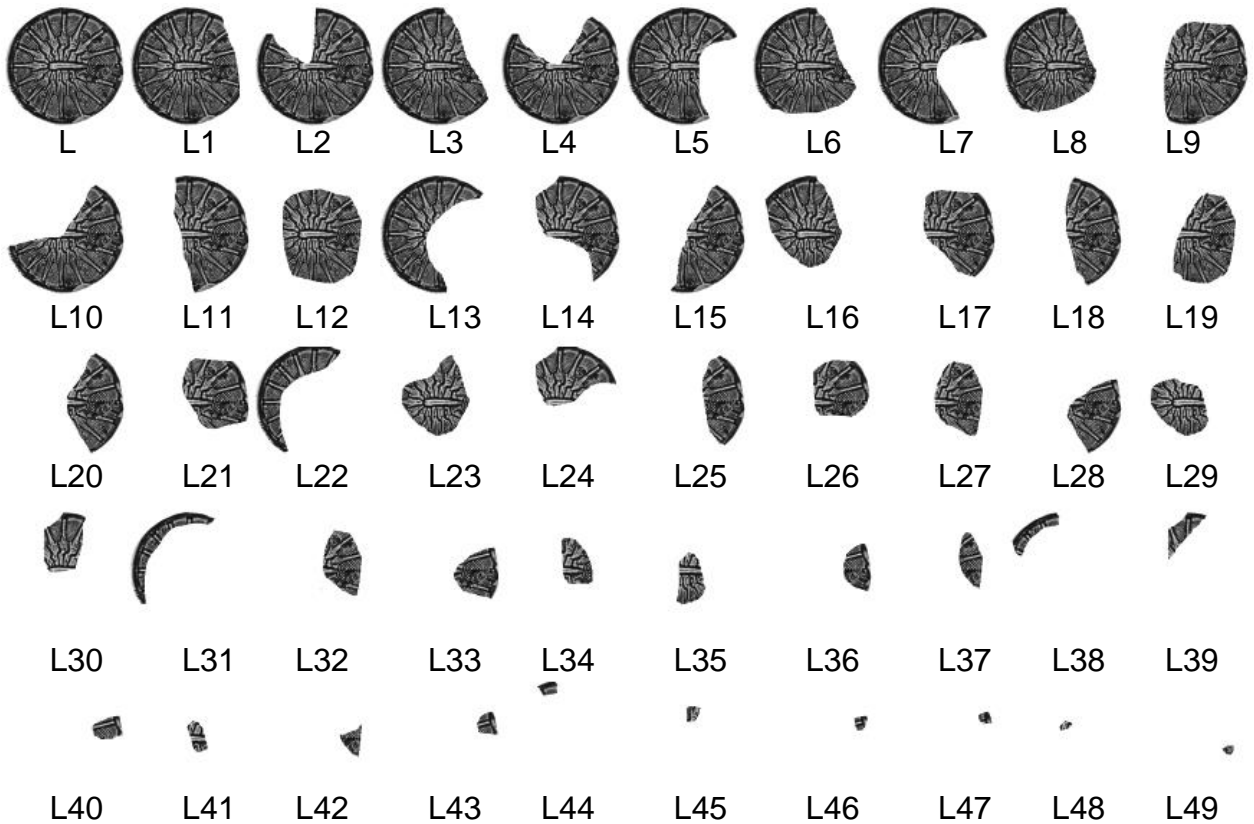
Diatomea fósil I y sus 49 fragmentos.



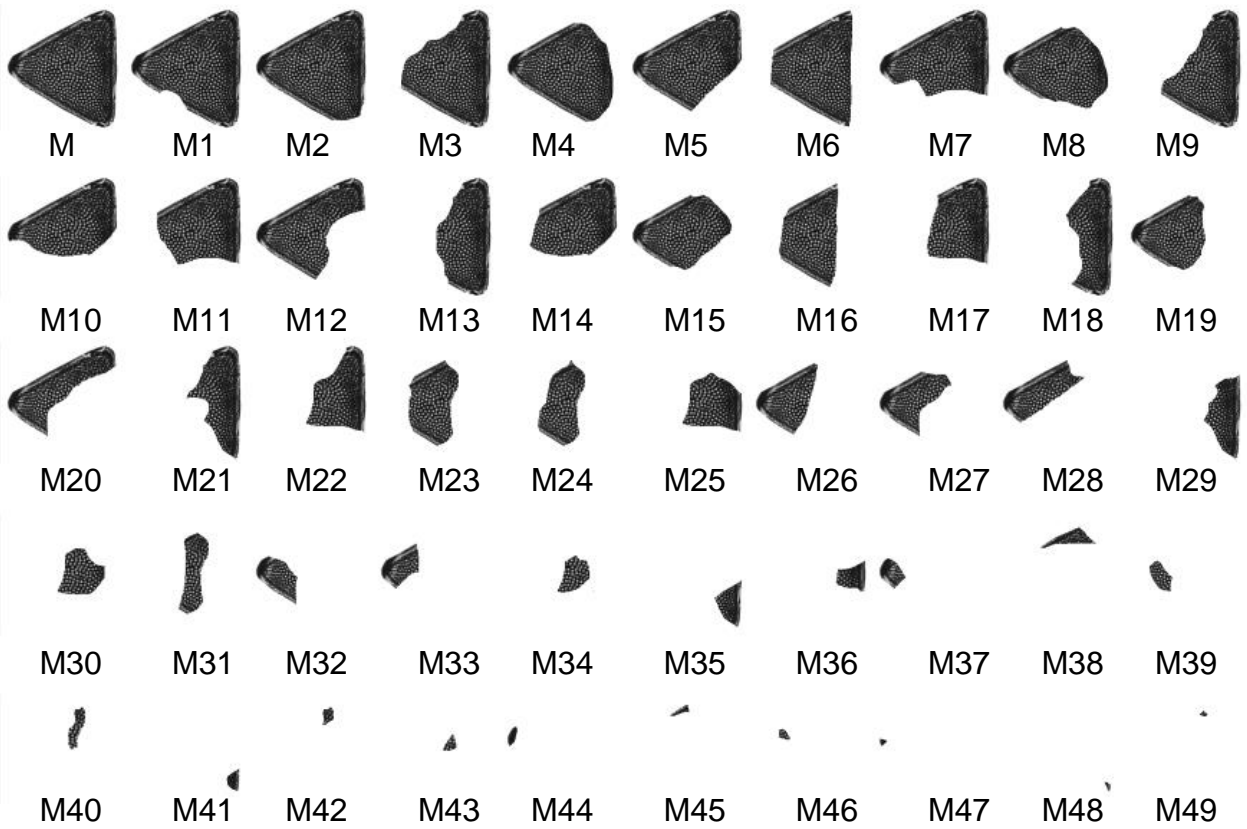
Diatomea fósil J y sus 49 fragmentos.



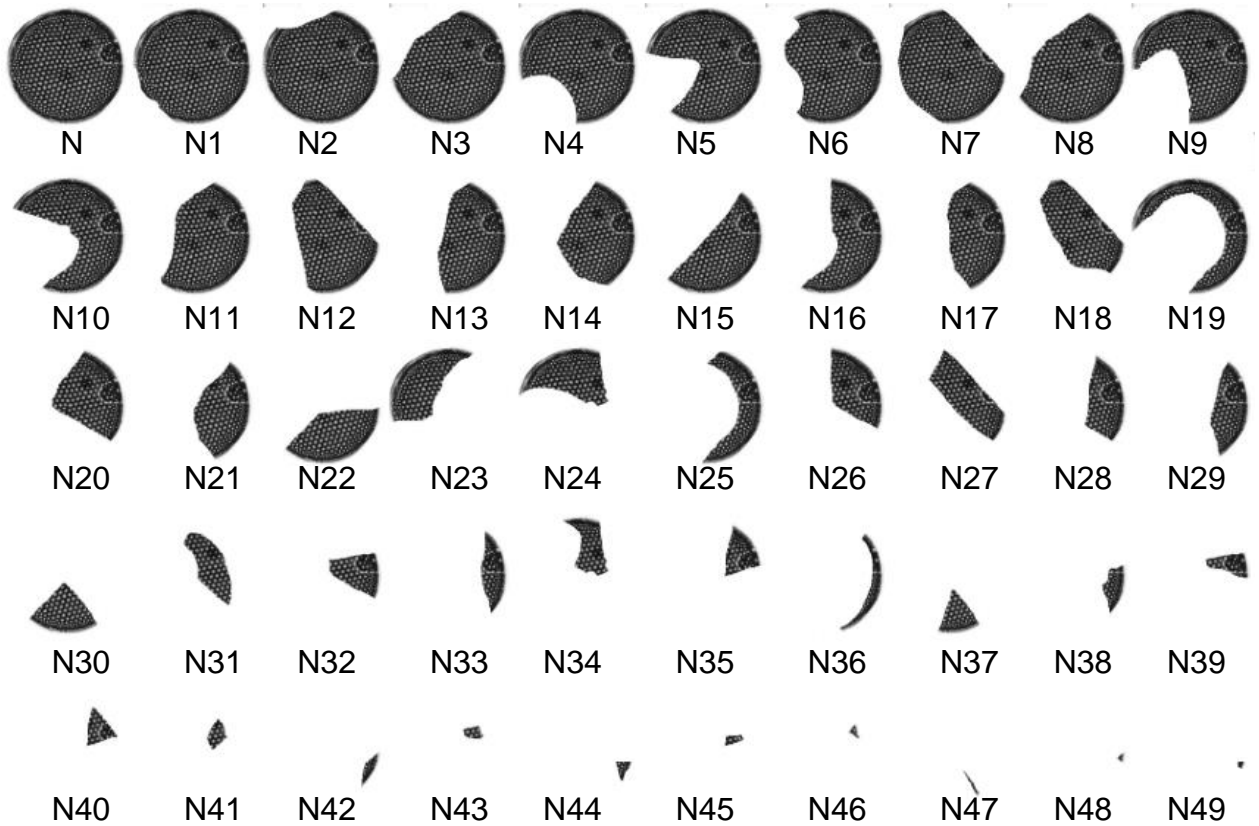
Diatomea fósil K y sus 49 fragmentos.



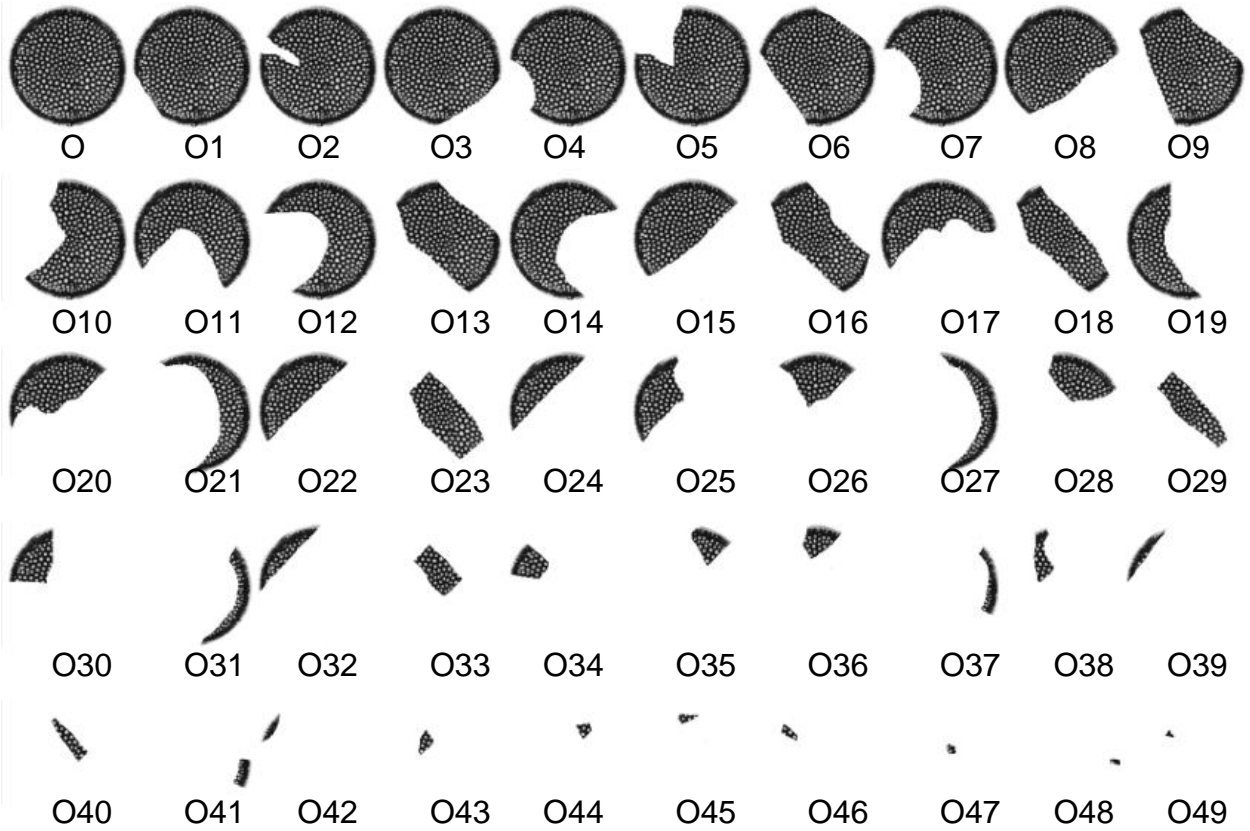
Diatomea fósil L y sus 49 fragmentos.



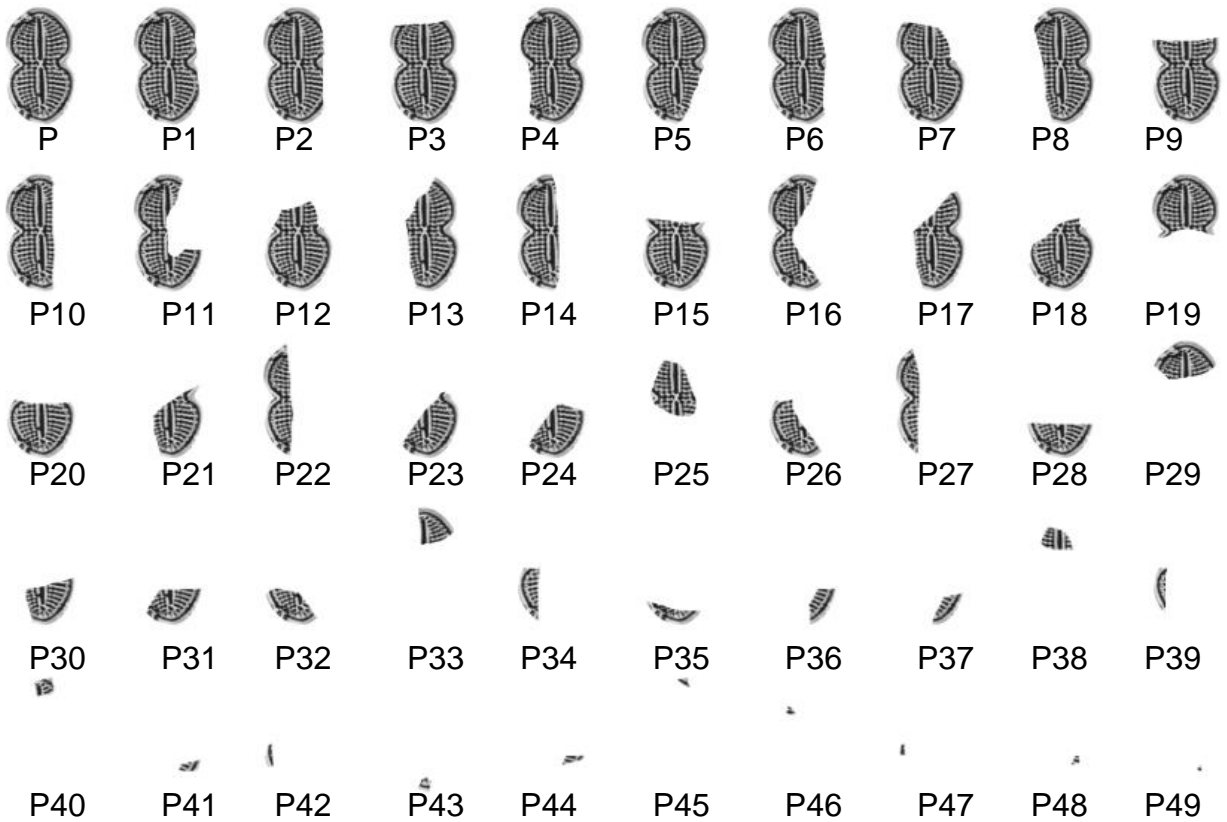
Diatomea fósil M y sus 49 fragmentos.



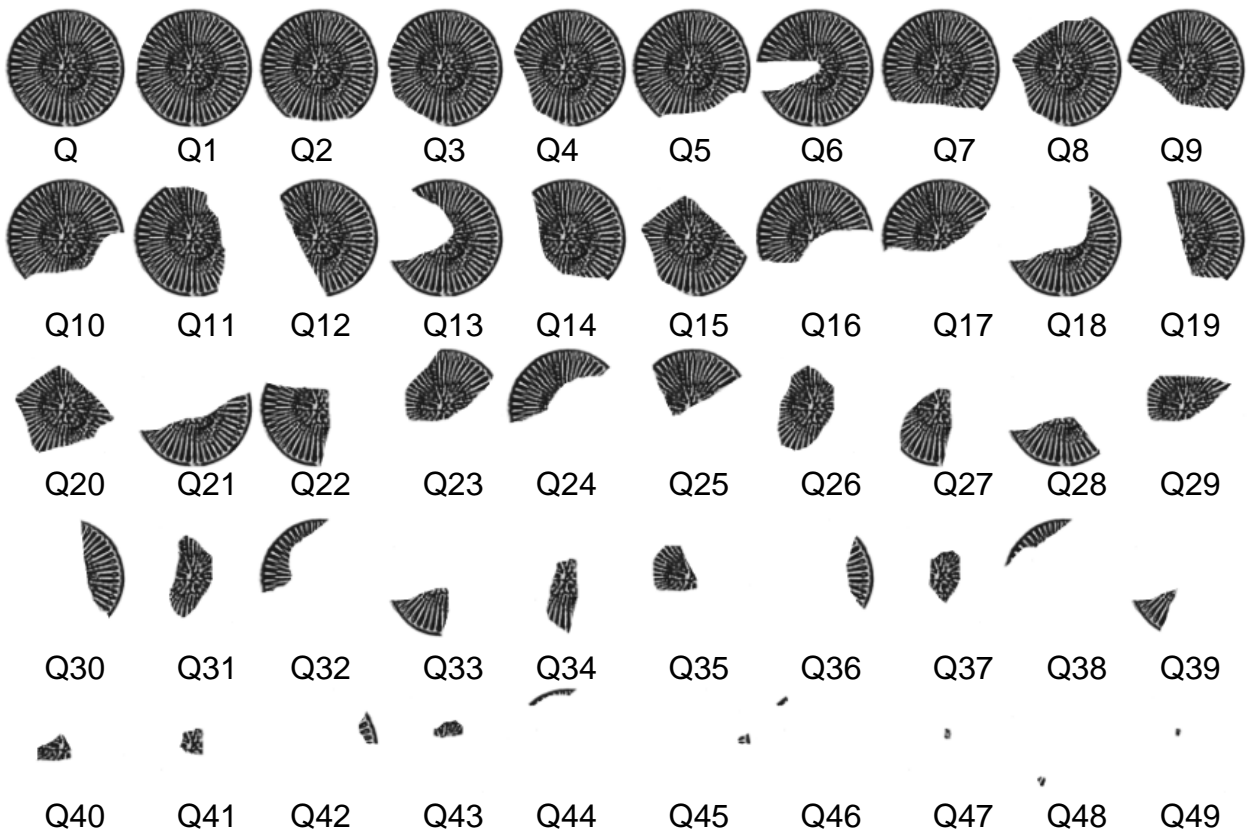
Diatomea fósil N y sus 49 fragmentos.



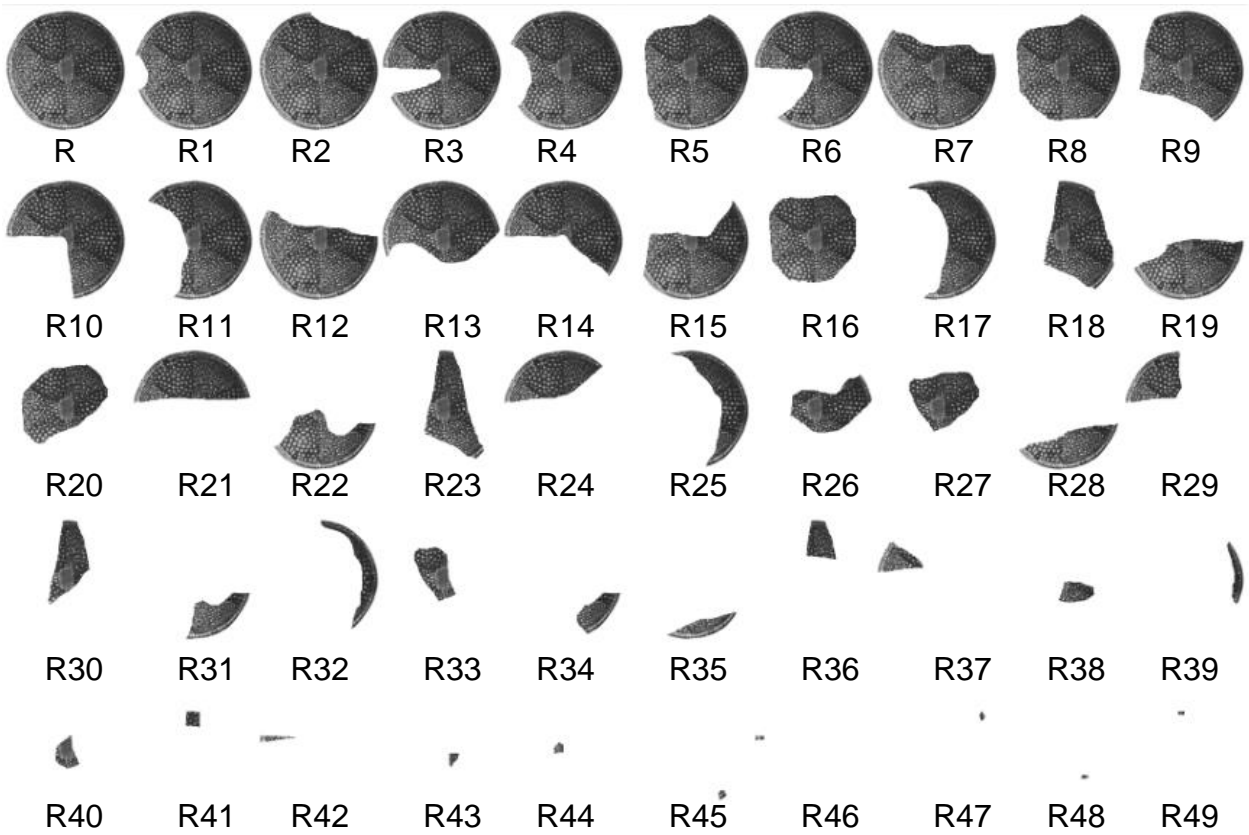
Diatomea fósil O y sus 49 fragmentos.



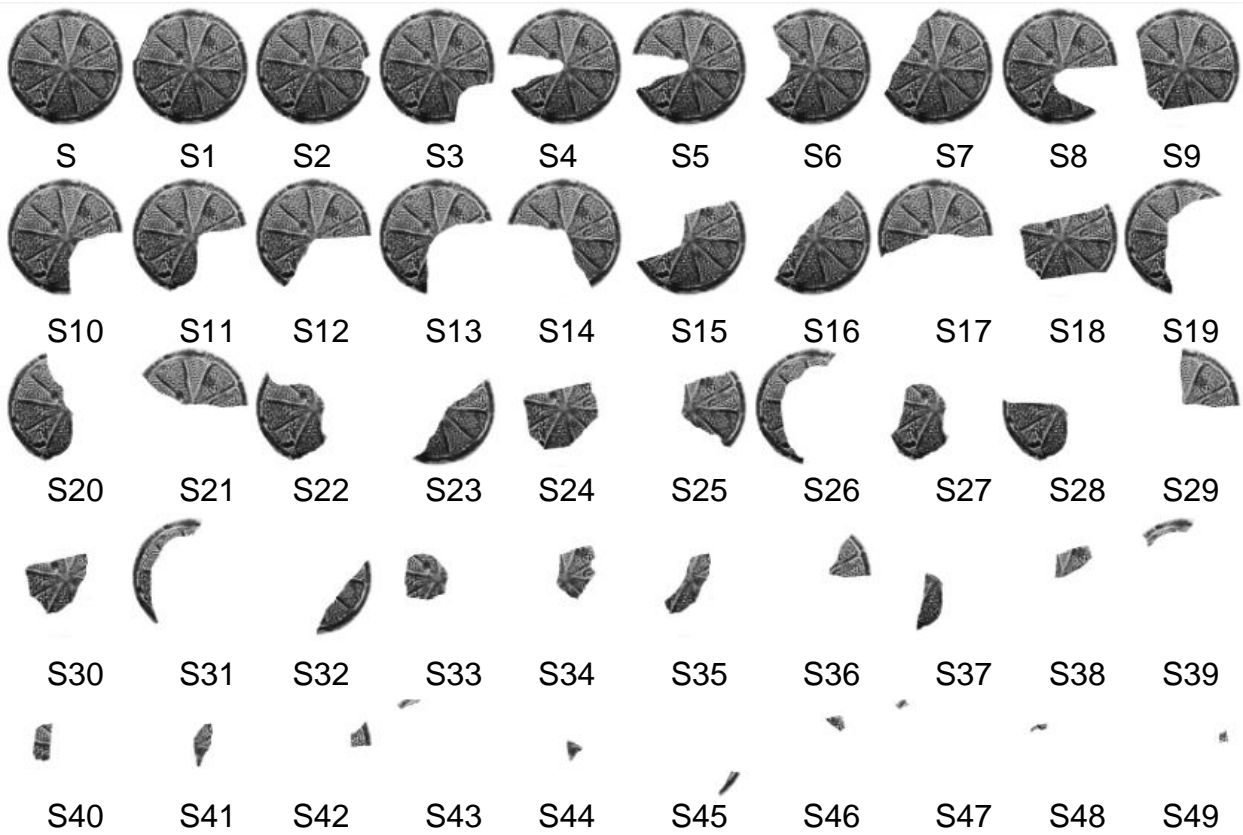
Diatomea fósil P y sus 49 fragmentos.



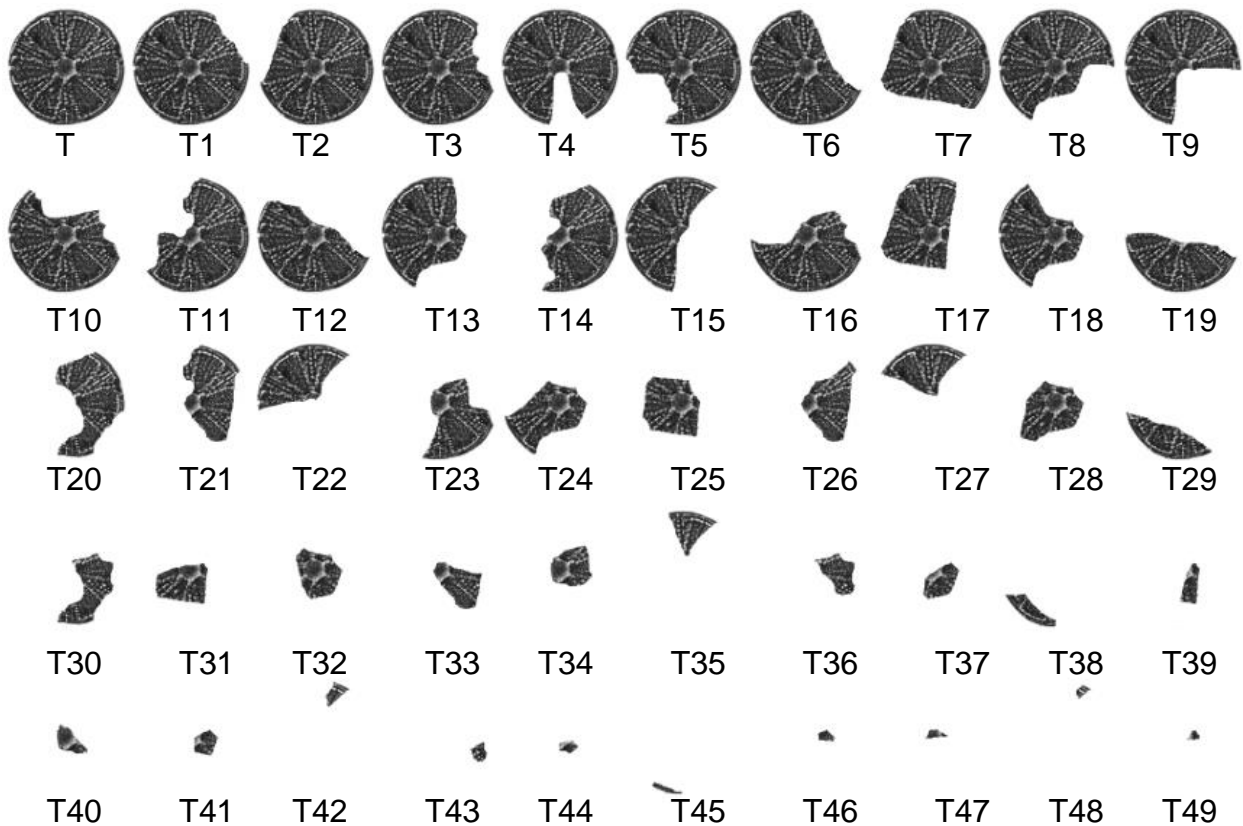
Diatomea fósil Q y sus 49 fragmentos.



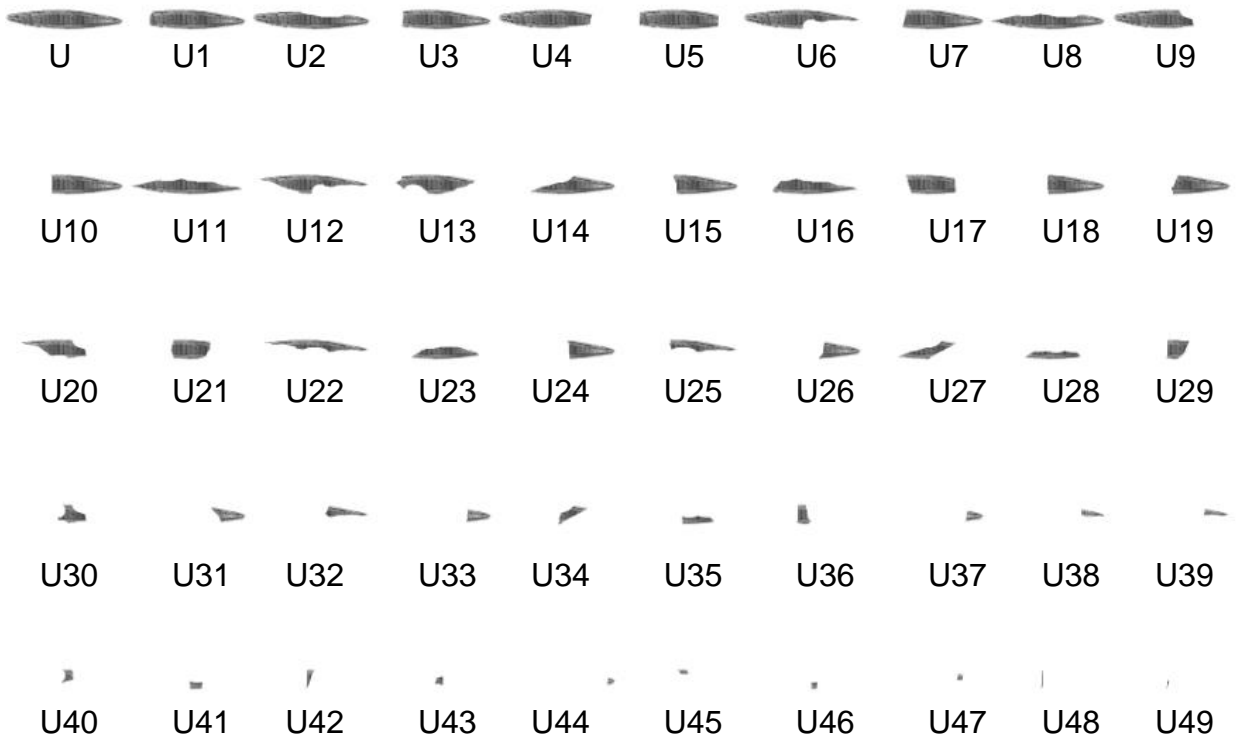
Diatomea fósil R y sus 49 fragmentos.



Diatomea fósil S y sus 49 fragmentos.



Diatomea fósil T y sus 49 fragmentos.



Diatomea fósil U y sus 49 fragmentos.