



Universidad Autónoma de Querétaro
 Facultad de Ingeniería
 Doctorado en Ingeniería

**“Análisis multiescala morfológico y compensación en iluminación en imágenes digitales”
 TESIS**

Que como parte de los requisitos para obtener el grado de

Doctor en Ingeniería

Presenta:

Edgar Alejandro Rivas Araiza

Dirigido por:

Jorge Domingo Mendiola Santibáñez

SINODALES

Jorge Domingo Mendiola Santibáñez
 Presidente

Gilberto Herrera Ruiz
 Secretario

Pedro Daniel Alaniz Lumbreras
 Vocal

Iván Terol Villalobos
 Propietario

Damián Vargas Vázquez
 Propietario

René de Jesús Romero Troncoso
 Suplente

Rodrigo Castañeda Miranda
 Suplente

Dr. Gilberto Herrera Ruiz
 Director de la Facultad

Firma

Firma

Firma

Firma

Firma

Firma

Firma

Dr. Luis Gerardo Hernández Sandoval
 Director de Investigación y
 Posgrado

Centro Universitario
 Querétaro, Qro.
 Mayo 2008
 México

“Análisis multiescala morfológico y compensación en
iluminación en imágenes digitales ”

Edgar Alejandro Rivas Araiza

RESUMEN

En este trabajo de tesis se desarrollan dos herramientas de visión por computadora cuyo comportamiento esta basado principalmente en la teoría de la morfología matemática; por un lado se propone una técnica para el análisis multiescala de imágenes y por otra parte se introduce un operador de contraste de dos primitivas que realiza la normalización y mejora de contraste en imágenes. Distintos estudios realizados muestran que el análisis en el espacio de escala es de gran utilidad en un amplio rango de aplicaciones de visión por computadora, que incluyen el reconocimiento de patrones, codificación, segmentación de imágenes, etc. Con esta motivación se propuso una técnica de análisis multiescala basado en transformaciones morfológicas por reconstrucción las cuales tienen propiedades importantes, en comparación con las transformaciones normales , tales como, preservación de contornos, mayor inmunidad al ruido y además, a partir de ellas es posible realizar la segmentación de la imagen.

El operador de contraste emplea la noción de fondo de imagen y un modelo de percepción visual humana (Ley de Weber) con lo cual se consigue la normalización y mejora de contraste en imágenes con mala iluminación. En este trabajo se mostraran las propiedades más importantes de las propuestas y se ilustrara su desempeño a partir de varios ejemplos.

(**Palabras clave:** Ley de Weber, espacio de escala, transformación con reconstrucción, morfología matemática)

Abstract

In this work two computer vision approaches are presented, these are mainly based on the mathematical morphology theory; first a technique for scale space analysis is proposed, on the other hand a new transformation for image normalization and contrast enhancement is introduced. Various works reported in the literature show that scale space analysis is of great importance in a wide range of computer vision tasks, such as, pattern recognition, coding, segmentation, and so on. This fact yielded to propose a scale space approach based on morphological reconstruction transformation which have very advantageous properties, as, contour preservation, major noise immunity and offering a way to perform segmentation of the image.

The contrast operator uses the background notion and a human visual perception model (Weber Law), allowing the normalization and contrast enhancement in images with poor illumination. In this work the most important properties will be presented and the performance of both approaches will be illustrated with various examples.

(**Keywords:** Weber Law, scale space, reconstruction transformation, mathematical morphology)

Dedicatorias

Dedico este trabajo, como una promesa de continuar trabajando fuerte en el futuro, a las siguientes personas por quienes considero valió la pena tanto sacrificio y esfuerzo.

A Dios, quien me ha cuidado y me ha concedido estar con personas excepcionales en todo el trayecto del doctorado.

A mis padres, porque con su motivación, apoyo y ejemplo me han llevado hasta aquí.

A mi hermano Gerardo, por contagiarme desde chavo la pasión por la ingeniería.

A mi hermana Rocio, siempre atenta y echándome porras.

A mi hermano Erik, el celebre “Enigman”. Siempre has estado presente para enseñarme algo y ayudarme, dentro y fuera de la escuela.

A Sara una persona muy especial que llego a mi vida, quien con su compañía trajo un mundo nuevo.

Agradecimientos

Agradezco:

A mi director de tesis el Dr. Jorge Domingo Mendiola Santibáñez, quien con sus conocimientos y experiencia me transmitió las habilidades necesarias para lograr completar este trabajo de investigación. Además le agradezco ampliamente por su paciencia y comprensión, por su valiosa ayuda en los momentos difíciles.

Al Dr. Gilberto Herrera Ruiz, su constante motivación y apoyo me impulsaron a trabajar siempre con compromiso. Y por recomendarme trabajar con el Dr. Jorge Domingo Mendiola, quien llegaría a ser mi director de tesis.

Al Dr. Daniel Alanis Lumbreras, por todos sus consejos que me sirvieron para no distraerme en las ultimas etapas del doctorado.

Al Dr. Iván Terol Villalobos, sus acertadas observaciones permitieron mejorar considerablemente el trabajo de tesis.

A mis sinodales Dr. Damian Vargas, Dr. René Romero, Dr. Rodrigo Castañeda por sus revisiones, consejos y por ayudarme a expandir mi visión de los alcances de la investigación.

Al Dr. Mario Rodriguez, un gran profesor y excelente persona. Las actividades realizadas para sus materias fueron experiencias muy enriquecedoras.

A los profesores del Doctorado Alexis Herrería, Moises Acosta, Rafael Leyva, Alejandro Lozano y la maestra Adriana Medellin por ayudarme en mi preparación.

A mis camaradas de toda la vida Mario Trejo Perea y Gabriel Ríos Moreno, con quienes he perdido la cuenta de las largas horas de trabajo, desveladas y días sin comer para completar los trabajos.

A Manuel Toledano Ayala, por contar siempre con tu apoyo desinteresado y tu valiosa amistad.

A CONACyT por la beca otorgada durante la realización del doctorado en la Universidad Autónoma de Querétaro, sin la cual no hubiera sido posible este trabajo.

A la Universidad Autónoma de Querétaro por el apoyo prestado durante mi estancia en el doctorado.

Índice general

1. Introducción	1
1.1. Organización de la tesis	4
1.2. Material y equipo utilizado	5
2. Revisión de literatura	6
2.1. Análisis en el espacio de escala	6
2.2. Realce de contraste y normalización de la imagen	9
I Fundamentos Teóricos de Morfología Matemática	12
3. Morfología Matemática	13
3.1. Morfología matemática	13
3.2. Transformaciones morfológicas básicas	14
3.2.1. Nociones sobre teoría de conjuntos	14
3.2.2. Propiedades de base de las transformaciones morfológicas	16
3.2.3. El elemento estructural	17
3.2.4. Caso binario de la erosión y dilatación morfológicas	19
3.2.5. Caso niveles de gris de la erosión y dilatación morfológicas	20
3.2.6. Gradiente morfológico	22
3.2.7. Filtros morfológicos básicos. Apertura y cerradura morfológicas	23
3.3. Filtros por reconstrucción	26
3.4. Filtrado en el espacio de escala	31
3.4.1. Filtrado en el espacio de escala lineal	31
3.4.2. Fingerprints	34
3.4.3. Espacio escala dilatación-erosión morfológico	35
3.4.4. Morfología dependiente de la escala	36
II Espacio de escala apertura-cerradura por reconstrucción	

y cerradura-apertura por reconstrucción	43
3.5. Espacio de escala de la apertura-cerradura por reconstrucción	44
3.5.1. Apertura-cerradura por reconstrucción multiescala	44
3.5.2. Fidelidad de las transformaciones del espacio de escala Gaus- siano y morfológico	44
3.5.3. Fingerprint de transformaciones conexas	49

III Realce de contraste y normalización de iluminación en imágenes mal iluminadas **56**

4. Compensación en cambios de iluminación	57
4.1. Fondo de la imagen y normalización	57
4.2. Ley de Weber	57
4.3. Noción de fondo de la imagen	58
4.4. Mapeo de contraste basado en el fondo de la imagen	60
4.5. Experimentos y comparación de resultados	64
5. Conclusiones	70
5.1. Resumen	70
5.2. Discusión espacio escala apertura-cerradura por reconstrucción y cerradura- apertura por reconstrucción	70
5.3. Discusión realce de contraste en imágenes con iluminación deficiente . .	71
5.4. Trabajos futuros	72
5.4.1. Propuesta de aplicación para el reconocimiento de rostros	73

Índice de figuras

3.1. Diferentes elementos estructurales.	19
3.2. Punto de vista geométrico del dilatado y erosionado binario	21
3.3. Diferentes elementos estructurales.	24
3.4. Punto de vista geométrico de la apertura y cerradura morfológicas . . .	25
3.5. Cerradura por reconstrucción.	27
3.6. Cerradura por reconstrucción.	30
3.7. Espacio de escala Gaussiano de la imagen “Lena” para un rango de escalas de $\sigma = 1, 2, 3, 4, 5$ y 6 pixeles.	31
3.8. Kernel de convolución Gaussiano 2D.	32
3.9. Fingerprints en el espacio de escala Gaussiano: gráfica de los cruces por cero de una señal 1-D.	35
3.10. Diferentes niveles en la representación del espacio de escala de una ima- gen bidimensional (a) erosionada a niveles de escala $\sigma = 3, 5$ y 7 se muestra en recuadros (b), (c) y (d) respectivamente.	37
3.11. Diferentes niveles en la representación del espacio de escala de una ima- gen bidimensional (a) dilatada a niveles de escala $\sigma = 3, 5$ y 7 se muestra en recuadros (b), (c) y (d) respectivamente.	38
3.12. Mínimos locales de las imágenes de la Fig. 3.10.	39
3.13. Máximos locales de las imágenes de la Fig. 3.10.	40
3.14. Fingerprints en el espacio de escala dilatación-erosión morfológico de una señal 1D aleatoria. (a) Fingerprints y (b) Representación multies- cala de la señal	41
3.15. Fingerprint reducido de una señal aleatoria 1D en el espacio de escala morfológico. (a) Fingerprint reducido y (b) Representación multiescala de la señal	42
3.16. Cerradura por reconstrucción.	45

3.17. Gradiente morfológico de las imágenes de la Fig. 3.16 obtenido a partir de la definición dada por la Ec. 3.24. Gradiente de: a) Imagen original, b) filtrado Gaussiano con $\sigma = 8$, c) dilatado morfológico con $\mu = 6$ y d) apertura morfológica por reconstrucción con $\mu = 6$	46
3.18. Fingerprints multiescala para una señal 1D; (a) señal de datos; (b) fingerprints completos apertura-cerradura por reconstrucción; (c) fingerprint reducidos apertura-cerradura por reconstrucción; (d) fingerprints completos cerradura-apertura por reconstrucción; (e) fingerprints reducidos cerradura-apertura por reconstrucción; (f) fingerprints reducidos dilatación-erosión.	51
3.19. Fingerprint reducidos multiescala de dos imágenes faciales; (a1)-(b1) imágenes originales; (a2) y (b2) imágenes de entrada; (a3)-(a13) y (b3)-(b13) fingerprints reducidos apertura-cerradura por reconstrucción con $\mu = 1, 5, 10, 15, 20, 25, - 1 , - 2 , - 4 , - 6 y - 8 $; (c1) – (c2) una representación 3D de los fingerprint reducidos anteriores.	55
4.1. Imagen original f junto con el fondo de la imagen desigual en líneas punteadas	59
4.2. Respuesta del operador de contraste a cambios en el background	60
4.3. Detección del fondo de la imagen para regiones claras(b_1) y oscuras(b_2). El criterio de fondo de la imagen corresponde a τ	62
4.4. Evolucion de los minimos regionales de una imagen filtrada con la cerradura por reconstrucción $\tilde{\varphi}_\mu$, a medida que μ se incrementa tomando los valores de 1,4,8,15, y 21	63
4.5. Diferentes casos de pose e iluminación para un individuo. (a-c) Imágenes de prueba; (d) Resultados obtenidos cuando $\mu = 21$; (e-f) Imágenes de salida para $\mu = 15$	65
4.6. Normalización y mejora del contraste usando $\tilde{\varphi}_{\mu=10}$ para la detección del fondo de la imagen	66
4.7. Aplicación de varios operadores de contraste. (a) Imagen original, (b) Mapeo morfológico de contraste a nivel zona plana, (c) Ecuilización del histograma, (d) Operador de contraste de dos primitivas propuesto en este trabajo	67
4.8. Distribucion de los niveles de gris de las imagenes de la fig. 4.7(a), (b), (c) y (d) respectivamente.	68

5.1. Proceso para reconocimiento de patrones a partir de los fingerprints de la imagen y coordenadas polares. (a) Imagen bajo prueba; (b) Rostro segmentado; (c), (d) y (e) fingerprints reducidos apertura-cerradura con valores de escala $\sigma = 3, 4$ y 5 74

Índice de cuadros

1.1. Niveles de análisis en visión por computadora	1
--	---

Capítulo 1

Introducción

La visión es uno de los mecanismos sensoriales de percepción más importantes en el ser humano; desde hace varias décadas los investigadores han enfocado su interés en el estudio del funcionamiento del sistema visual humano, buscando encontrar un modelo de percepción visual que permita reproducir las características de la visión en una máquina. El objetivo de la visión por computadora es hacer posible un sistema que pueda interpretar el entorno visual de manera automática bajo cualquier condición de operación, esto es, duplicar (o en algunos casos compensar) el desempeño de la percepción visual humana.

Las aplicaciones que se pueden derivar de un sistema de visión por computadora son muy amplias; por ejemplo, en robótica, en la inspección automática de partes industriales, en aplicaciones de medicina, radar, sonar, sensado remoto vía satélite y en transmisión y almacenamiento de imágenes. De acuerdo a Levine (1985) se divide el campo de la visión por computadora en los niveles de análisis mostrados en el cuadro 1.1.

Nivel	Descripción
M+3	Interpretación de escenas 3D
M+2	Descripción de escenas 3D
M+1	Descripción de escenas 2D
6 a M	Agregación de alto nivel y correspondencia de modelos
5	Descubrimiento de relaciones estructurales
4	Clasificación de características
3	Segmentación de imágenes y detección de características
2	Preprocesado y restauración
1	Caracterización de sensores
0	Escenas

Cuadro 1.1: Niveles de análisis en visión por computadora

Los niveles del 0 al 3 son conocidos como procesamiento de imágenes o visión de bajo nivel (*low-level vision*), y los niveles superiores como interpretación de imágenes o visión de alto nivel (*high-level vision*); el nivel 4 también es conocido como reconocimiento de patrones. El procesamiento de la imagen (visión a bajo nivel) involucra transformar la imagen en una representación más adecuada para operaciones posteriores de reconocimiento de patrones y percepción de la imagen (visión a alto nivel). Una aplicación de visión por computadora requiere que se realicen varias operaciones de alto y bajo nivel, actualmente en la literatura se han propuesto una gran cantidad de técnicas para el procesado de la imagen y el número continúa en aumento.

La necesidad de seguir desarrollando nuevas estrategias de procesado es debido a que en la mayoría de los casos los requerimientos varían de una aplicación a otra. Si tomamos en cuenta que la captura de una imagen depende de tres factores (el contenido de la escena, la iluminación incidente en la escena y las características del elemento de captura) entonces los objetos en la imagen obtenida están expuestos a una gran variabilidad. Para un sistema de visión por computadora es complicado hacer frente a estas variaciones; por tanto es práctica común, para reducir la complejidad del sistema, imponer las condiciones bajo las cuales debe operar. Esto implica una solución desde el punto de vista *ad hoc* en el contexto de la aplicación considerada en cada caso.

La flexibilidad del sistema se puede mejorar de dos maneras: la primera es mediante una etapa de preprocesado donde se normalicen las propiedades geométricas (pose, alineación, rotación, etc.) y fotométricas (iluminación, contraste, etc.) y, segundo desarrollando algoritmos robustos frente a las variaciones.

Una herramienta de análisis de alto nivel es la representación en el espacio de escala introducida por Witkin (1983), esta se ha convertido en una metodología universal para una amplia variedad de aplicaciones de visión por computadora, tales como segmentación Petrovic et al. (2004); Vachier (2001); Mukhopadhyay (2003), reconocimiento de patrones Jackway et al. (1994) y codificación de imágenes Martens (1990); Xin (2002). Esta metodología está basada en la observación que los objetos del mundo real y sus proyecciones existen como entidades significativas solo en un determinado rango de escalas. Definiendo un factor de escala como un parámetro, una imagen puede descomponerse en una familia de versiones gradualmente simplificadas de ella misma.

El parámetro de escala controla el monto de suavizado, así entre más grande sea, la información presente en escalas más finas es suprimida. De esta manera los rasgos que contribuyen más significativamente a la apariencia de la imagen se mantienen conforme

se aumenta la escala, estos rasgos predominantes pueden usarse como descriptores de las características de la imagen.

Las representaciones en el espacio de la escala pueden formarse sobre la base de muchos principios diferentes. Solo es necesario ser capaz de obtener una descripción de la estructura de las imagen a través de la escala.

A partir de los filtros morfológicos por reconstrucción se puede construir un operador que cumpla con el criterio de análisis multiescala con una mejora considerable en el desempeño con respecto a las operaciones morfológicas básicas; uno de los objetivos de este trabajo de tesis es proponer una herramienta de análisis multiescala basado en transformaciones morfológicas por reconstrucción. La propuesta de este trabajo no se enfoca solamente a la obtención de una representación multiescala de la imagen, sino también en la obtención puntos característicos en cada escala para extraer los fingerprints. Se demostrará como el operador propuesto cumple con el criterio de análisis multiescala al definir los puntos característicos a partir de los *extremos regionales*, las propiedades del operador se presentaran en el contexto multiescala. Mientras que desde el punto de vista de las transformaciones usadas, se probará que se obtienen mejores resultados al usar filtros por reconstrucción; algunas ventajas evidentes son: menor sensibilidad al ruido, preservación de los contornos de la imagen y evitan la introducción de nuevas características.

Como se ha mencionado anteriormente, las características de la escena y la iluminación determinan la calidad de la captura de la imagen. Las variaciones en la intensidad y dirección de la iluminación afectan significativamente la apariencia de la imagen. Como segunda propuesta de esta tesis, se introduce un nuevo operador para normalización del contraste y compensación en imágenes con iluminación deficiente; la transformación propuesta incorpora un modelo de percepción visual humano definido a partir de la *Ley de Weber*. El operador realiza la normalización a partir de un criterio de contraste de dos primitivas, de esta manera se tiene una modificación del contraste para regiones claras y oscuras. El criterio de contraste utiliza la noción de *fondo de la imagen*, el fondo de la imagen es extraído a partir de transformaciones morfológicas de apertura y cerradura por reconstrucción.

En resumen, en este trabajo de tesis se introducen dos herramientas de procesamiento de imagen, una de ellas considerada dentro de la visión a bajo nivel (realce de contraste y normalización) y la otra cae en la categoría de operaciones de alto nivel (análisis multiescala), ambas herramientas basan su funcionamiento en la teoría de

la morfología matemática. Las herramientas se pueden usar de manera independiente en etapas de preprocesado o de segmentación y reconocimiento de patrones, también pueden emplearse en combinación para mejorar el desempeño e incrementar la robustez en diversas aplicaciones de visión por computadora. En las secciones siguientes se presentara el estado del arte del análisis multiescala y realce de contraste, de acuerdo a la literatura citada se mostrara la importancia y validez de lo que se propone en este trabajo.

1.1. Organización de la tesis

El contenido temático de esta tesis esta organizado de la siguiente manera:

Capítulo I Contiene una breve introducción para dar un panorama general de la problemática que se desea abordar.

Capítulo II Presenta de manera resumida la información que existe en la literatura sobre investigaciones realizadas acerca del análisis multiescala y mejoramiento de contraste. Se comienza describiendo la metodología de análisis de espacio de escala, el cual puede formarse teniendo como base distintas metodologías; se discutirán las ventajas, desventajas y prospectivas de dichas metodologías. además en esta sección, se presenta una reseña de la problemática que se presenta en las imágenes digitales debidas a condiciones de iluminación deficientes y las técnicas comúnmente usadas para compensar o reducir el efecto de la iluminación.

Capítulo III Este capítulo corresponde a la metodologia, aqui es donde se presentan las propuestas del análisis multiescala y el algoritmo para el mejoramiento de contraste en imágenes con mala iluminación. Este capitulo se dividió en tres partes, la primera de ellas trata sobre las definiciones, notación y fundamentos teóricos para el entendimiento del trabajo presentado; la segunda parte de este capitulo presenta el desarrollo y las propiedades de las transformaciones multiescala propuesta propuestas; mientras que en la tercera parte se desarrolla el operador de contraste de imágenes con iluminación deficiente. En cada caso, el desempeño de las transformaciones es presentado de manera individual con ayuda de varios ejemplos.

Capítulo IV Finalmente en esta sección se analiza y se concluye sobre los resultados obtenidos; además de presentar una prospectiva sobre trabajos futuros a desarrollarse.

1.2. Material y equipo utilizado

La plataforma de desarrollo para los algoritmos de procesamiento de imágenes se realizara con ayuda del programa Matlab versión 7.0 y el toolbox de procesamiento de imágenes que ya viene incluido. El toolbox proporciona funciones para manipular, procesar y mostrar imágenes digitales, en particular las funciones para calcular el erosionado, dilatado y reconstrucción fueron usadas para programar los algoritmos de las transformaciones propuestas.

De esta manera se salva tiempo de desarrollo, aunque a consecuencia se tiene un tiempo de ejecución más largo; esto es debido a que no se controla el uso de la memoria o los recursos del sistema. En la practica, las imágenes de salida se obtienen en unos cuantos segundos lo cual no representa un inconveniente al momento de verificar el funcionamiento de las transformaciones. Pero en aquellos casos en que los algoritmos deban ejecutarse en tiempo real se podría optar por transportar los algoritmos a algún lenguaje de alto nivel como Visual C, que permita una mejor administración de los recursos del sistema y de esta manera obtener una reducción considerable en el tiempo de respuesta.

Para realizar los experimentos se creo una base de datos de imágenes de distintos objetos, algunas fueron obtenidas de internet mientras que otras fueron tomadas con ayuda de una camara digital SONY cyber-shot®. Para los experimentos con las transformaciones multiescala los requerimientos de las imágenes no requieren consideraciones especiales, mientras que para los algoritmos de mejora de contraste fue necesario contar con una base de datos especifica que considere condiciones de iluminación variables.

Para la realización de los experimentos de mejora de contraste se hizo uso de una base de datos que contiene imágenes de rostros tomadas bajo condiciones de iluminación deficiente, las cuales se obtuvieron de la base de datos desarrollada por Peter Belhumeur y Athos Gheorgiades, en el centro de visión y control computacional de la universidad de Yale (<http://cvc.yale.edu/projects//yalefacesB/yalefacesB.html>).

Esta base de datos, originalmente creada para probar algoritmos de detección y reconocimiento de rostros, contiene 5760 imágenes tomadas con una sola fuente de iluminación de 10 individuos vistos bajo 576 condiciones de observación (9 poses x 64 condiciones de iluminación).

Capítulo 2

Revisión de literatura

Debido a que en este trabajo de investigación se concentra en el desarrollo de dos herramientas distintas, pero ligadas en el campo de visión por computadora, este capítulo presentara el estado del arte de cada una de ellas. En la primera sección de este capítulo trataremos lo concerniente al estado del arte del análisis en el espacio de escala, mientras que en la segunda sección se abarca el tema de mejoramiento de contraste.

2.1. Análisis en el espacio de escala

Las técnicas multiescala para el análisis de señales están motivadas por estudios psicofísicos que han demostrado que el sistema de visión humano procesa y analiza la información presente en las imágenes a diferentes resoluciones. En muchas aplicaciones de visión por computadora, los objetos de interés que se necesitan medir, segmentar o reconocer en una imagen pertenecen a una escala, y el resto de los objetos, cuya información desea ser descartada, pertenecen a otra escala. En algunos casos, sin embargo, no existe un valor de escala a partir del cual se puedan separar las características de los objetos, y la información de interés esta presente en varias escalas; por lo cual, se necesita extraer la información a partir de información desde varias escalas.

Para tales situaciones, se han desarrollado las metodologías de análisis multiescala, donde se derivan una serie representaciones más gruesas de la imagen. De esta manera el reconocimiento o segmentación usaran el conjunto completo de representaciones a varias escalas no solo de la imagen original. Un estudio anterior donde la noción de escala fue incorporado se puede encontrar en Marr (1980), en el se utilizó una combinación de filtros Gaussianos y Laplacianos para derivar la familia de representaciones de la imagen original; el objetivo del estudio era detectar los contornos a varias escalas

por medio del operador Laplaciano cuando el ancho del Gaussiano varia.

Mas tarde, el concepto de espacio de escala fue formalizado por Witkin (1983), la idea básica detrás de este enfoque es obtener una pila de representaciones simplificadas a partir de la señal original f . Las versiones filtradas tienen la característica de presentar menos detalles finos conforme el parámetro de escala se incrementa satisfaciendo el concepto de *causalidad*, i.e., cada característica en escalas más gruesas es causada por las que se encuentran en escalas más finas. En adición a la propiedad de causalidad, se necesita cumplir con la propiedad de *preservación de contornos* ya que los contornos son una de las características visuales más importantes.

Se puede llegar a una representación multiescala adecuada si se define una transformación desde una escala más fina a una más gruesa. El grado de libertad en la elección de la transformación se ve reducida al imponer un conjunto de propiedades que deben ser satisfechas:

- **Invariancia espacial** \Rightarrow invariancia a la translación.
- **Isotropía** \Rightarrow invariancia a la rotación.
- **Invariancia bajo los cambios de iluminación** \Rightarrow la transformación debe conmutar con una anamorfosis de la luminancia.

Además de los siguientes requerimientos en el efecto de la transformación:

-La transformación debe ser efectivamente una simplificación de la imagen. Como esta no sera reversible: algo de información se tiene que perder de una escala a la siguiente.

-Causalidad: La información presente en escalas más gruesas solo puede ser causada por lo que ocurra en escalas más finas.

-No debe crear nuevas características a escalas más gruesas: el requerimiento más frecuente es que no debe crear nuevos extremos locales o regionales.

Las representaciones en el espacio de la escala pueden formarse sobre la base de muchos principios diferentes. Solo es necesario ser capaz de obtener una descripción de la estructura de las imagen a través de la escala. Es posible derivar la pila de imágenes a partir de diferentes operadores de escala, en la literatura se pueden encontrar

diversas metodologías, pero de manera general se pueden clasificar como:

Espacio de escala lineal: Una familia de imágenes derivadas que dan solución a la ecuación de difusión de calor, Koenderink (1984) demostró que el único operador lineal que satisface la ecuación de difusión es el filtro Gaussiano. La imagen es suavizada convolucionando la imagen con un kernel Gaussiano donde la desviación estándar σ corresponde al factor de escala según Witkin (1983); Yuille (1986). Una desventaja del espacio de escala lineal es que el filtrado Gaussiano puede difuminar o recorrer las características de la imagen lo cual fue observado por Maragos (1989); Vachier (2001).

Espacio escala no-lineal: Se construye un espacio de escala a partir de operadores no lineales, tales como, filtros por promedio de acuerdo a Bangham et al. (1996b) y filtros morfológicos, como los presentados por Bosworth (2003); Chen (1989); Jackway (1996); Park (1996); Haralick et al. (1987); manteniendo las propiedades principales del espacio de escala lineal.

Se puede obtener un espacio de escala no-lineal usando operadores morfológicos donde el elemento estructural empleado esta en función del parámetro de escala; se pueden usar varios operadores morfológicos para obtener una representación en el espacio de la escala adecuado; en la literatura, estudios previos se han enfocado en aquellos creados por medio de la dilatación o erosión Jackway et al. (1994); Jackway (1996), apertura y cerradura Chen (1989); Park (1996), y filtros alternados Bangham et al. (1996a). Mas recientemente, se realizo un estudio para obtener una representación en el espacio de la escala usando una clase particular de filtros por reconstrucción llamados *levelings* Meyer (1999, 2000); Vachier (2001); aunque los trabajos anteriores usan morfología dependiente de la escala, no todos distinguen características de la imagen a través del espacio de escala.

Jackway et al. (1994); Jackway (1996) definió los *fingerprints* en base los *extremos locales* de la señal obtenida en la representación multiescala; del estudio de varias funciones estructurales aplicadas en el operador de erosión y dilatación, demostró que dicha representación multiescala cumple las propiedades de *causalidad* y evita la aparición de nuevos contornos en la imagen a medida que esta es procesada a distintas escalas. En el estudio de Jackway (1996) se presento la utilidad del concepto de fingerprints a partir del espacio de escala morfológico erosión-dilatación para el reconocimiento de patrones; en otros trabajos reportados en la literatura que emplean dicha metodología destacan aquellos realizados por Kotropoulos (1997); Kotropoulos et al. (2000); Raducanu (2000), donde se proponen técnicas para la verificación y localiza-

ción de rostros humanos en imágenes.

El inconveniente de usar los extremos locales como características de escala, es su alta sensibilidad al ruido, como menciona Bosworth (2003); además que los operadores de erosión y dilatación requieren un elemento estructural especial para evitar el corrimiento en las características y la aparición de nuevos contornos en la imagen procesada. En el marco de la morfología matemática se ha popularizado el uso de las transformaciones por reconstrucción (Salember (1995)), esta clase especial de filtros modifican los máximos y mínimos de la imagen sin introducir nueva información.

Los estudios anteriores sirvieron como motivación para realizar la propuesta de un nuevo espacio de escala construido con transformaciones apertura y cerradura por reconstrucción, estas transformaciones primero remueven los rasgos brillantes u oscuros de la imagen o aquellos que no encajen en el elemento estructural dado por el valor de escala, para después restaurar las componentes que no fueron removidos completamente usando información de la imagen original. Como resultado no aparecen problemas tales como creación de nuevos contornos o desplazamiento de bordes.

2.2. Realce de contraste y normalización de la imagen

Uno de los problemas fundamentales en el procesamiento de imágenes es provocado por los cambios de iluminación; las variaciones en la dirección y la intensidad de la iluminación, son factores que modifican significativamente la apariencia de los objetos en una imagen digital.

La iluminación ambiental puede variar a lo largo del día en ambientes interiores y exteriores. Debido a la forma tridimensional de los objetos una fuente de iluminación puede generar sombras que acentúan o disminuyen ciertos rasgos de la imagen. Mas aún, distintas condiciones de iluminación pueden producir representaciones desiguales de un mismo objeto, dichas variaciones son indeseables ya que dificultan por ejemplo el proceso de segmentación Klinker et al. (1990), el reconocimiento de patrones Swain (1991), y seguimiento Jiang (2003) entre otras.

Las variaciones en la imagen debidas a la iluminación deficiente afecta el contraste de la imagen, la modificación en el contraste es evidente al presentarse variaciones bruscas entre regiones oscuras mal iluminadas y claras sobreiluminadas. Mediante un

realce adecuado en las imágenes se puede directa e indirectamente corregir el efecto de la iluminación. A continuación se describen algunas técnicas de realce de contraste encontradas en la literatura.

El problema del realce o mejoramiento de contraste en imágenes digitales se puede abordar desde distintas metodologías, una de ellas, por ejemplo, es la morfología matemática. Los primeros trabajos en esta área fueron realizados por Meyer y Serra Meyer (1989), quienes introdujeron la noción de mapeos de contraste. Dichos operadores consisten, de acuerdo a cierto criterio de proximidad, en seleccionar en cada punto de la imagen analizada, un nuevo nivel de gris entre dos patrones (primitivas) Meyer (1989).

Otros trabajos basados en la noción de mapeos de contraste fueron realizados por Terol-Villalobos (1996, 1998b,a, 2001) y Mendiola-Santibanez (2002b,a). En Terol-Villalobos (1996), se propone una clase especial de mapeos de contraste llamados filtros morfológicos por pendiente, posteriormente se extendió el estudio resultando en una familia de filtros morfológicos por pendiente secuenciales Terol-Villalobos (1998b,a) y filtros por pendiente conexos Terol-Villalobos (2001). Las transformaciones anteriores, además de mejorar el contraste, permiten llevar a cabo la segmentación de la imagen tratada.

Por otro lado, en Mendiola-Santibanez (2002b) y Mendiola-Santibanez (2002a) se realizó un estudio de realce del contraste basado en la noción de zona plana. El estudio demostró que los operadores de contraste morfológicos sobre partición atenúan la degradación sobre la imagen de salida y evitan la generación de nuevos contornos. También, han sido reportado estudios basados en criterios multiescala sobre el contraste dentro de la morfología matemática Toet (1992); Mukhopadhyay (2000); Kasperek (2001). En el trabajo realizado por Mukhopadhyay y Chanda Mukhopadhyay (2000), se definió un esquema para realce de contraste local basado en una transformación tophat morfológico multiescala. Mientras que en Kasperek (2001), se implementó un sistema de procesamiento en tiempo real para su aplicación en el realce de imágenes angiocardiógráficas basado en el trabajo realizado por Mukhopadhyay.

A pesar de que el contraste morfológico ha sido estudiado, no existe una metodología desde el punto de vista morfología matemática que permita normalizar y realzar el contraste al mismo tiempo.

Otros trabajos que abordan el problema de variaciones de la iluminación, que emplean diferentes enfoques al de la morfología matemática, de manera general, se pueden clasificar en dos categorías: metodologías basadas en modelo y pre-procesado de la imagen.

En las metodologías basadas en modelo, se parte del cálculo de un *modelo* a partir de un conjunto de imágenes de entrenamiento del mismo objeto, capturadas en condiciones de iluminación variable. Aquí, el modelo Lambertiano Horn (1986) es empleado para describir las imágenes. Dentro de esta categoría se encuentran los conos de iluminación (*illumination cone*) Belhumeur et al. (1997); Georghiadis et al. (2001), los armónicos esféricos (*spherical harmonics*) Basri (2003). Mientras que para la metodología conocida como pre-procesado, las imágenes son pre-procesadas usando alguna técnica con la finalidad de ser normalizarlas y que aparezcan estables bajo diferentes condiciones de iluminación. Una de las técnicas de procesamiento de imágenes más comunes para realce de regiones oscuras es aquella que utiliza funciones no lineales. Las funciones empleadas son logaritmos o de potencias, las cuales trabajan a nivel píxel Jain (1989).

Por otra parte, un método que trabaja en el dominio de la frecuencia es el filtrado homomórfico Short and Kittler (2004). Existen también técnicas basadas en análisis estadístico de los datos; por ejemplo las técnicas de ecualización de histograma global y local son comúnmente usados. Durante el procesamiento de ecualización de histograma se re-ordena la distribución de los niveles de gris de una imagen para obtener un histograma uniformemente distribuido Gonzalez (1992). Sin embargo, la principal desventaja de la ecualización del histograma es que las propiedades globales de la imagen no pueden ser aplicadas apropiadamente en un contexto local Sherrier (1987). De hecho, las modificaciones inducidas por la ecualización del histograma trata todas las regiones de la imagen de igual manera y, en consecuencia, frecuentemente conduce a un desempeño local pobre en términos de preservación de detalles. De esta forma, aún cuando son varios los algoritmos reportados para compensar los cambios en la iluminación, dependiendo de la aplicación final a la que se destinen algunos pueden no ser los más adecuados. En el caso de las metodologías basadas en modelos, estos por lo general, suelen ser mas complejos y demandar una capacidad de cálculo elevada que en determinado caso pudiera ser prohibitiva (e.g. en sistemas en tiempo real o dispositivos embebidos). Mientras otras son de implementación más sencilla, tienen dificultad para preservar ciertas características de la imagen las cuales pueden ser fundamentales en etapas de procesamiento posteriores, tales como, la preservación de contornos o de extremos regionales.

Parte I

Fundamentos Teóricos de Morfología Matemática

Capítulo 3

Morfología Matemática

3.1. Morfología matemática

La morfología matemática (MM) es una rama del análisis y procesamiento de imágenes que utiliza conceptos de álgebra (teoría de conjuntos, retículos completos) y geometría (translación, distancia, convexidad). Los orígenes de la morfología matemática vienen de los trabajos de Matheron (1975) y Serra (1988), ambos investigadores pertenecientes a la escuela de Mines en Fontainebleau, quienes trabajaron en problemas de petrografía (estudia la estructura, composición y origen de las rocas) y mineralogía. El objetivo inicial de sus fundadores fue caracterizar propiedades mecánicas y físicas de ciertos materiales (secciones de roca, cerámica policristalina), tales como porosidad, en los cuales se examinaba la estructura geométrica de la muestra; siendo formalizados estos trabajos a partir de conjuntos aleatorios cerrados. Posteriormente se presentaron nuevas transformaciones morfológicas así como su formalización.

La MM como una técnica de procesamiento de imágenes, ha sido usada exitosamente en varios campos: citología, dermatología, control de tráfico vehicular, geografía, metalografía, robótica, medicina y muchos otros. No obstante, la morfología matemática no son solo una serie de métodos, es también una teoría sólida, cuya evolución esta fuertemente conectada a problemas concretos.

En las siguientes secciones se presentan los conceptos y definiciones mas empleadas en morfología matemática que servirán mas adelante para establecer las propiedades de las transformaciones propuestas en este trabajo.

3.2. Transformaciones morfológicas básicas

3.2.1. Nociones sobre teoría de conjuntos

Los operadores en MM primero fueron explicados desde el punto de vista de la teoría de conjuntos, sin embargo Serra (1988) y Heijmans (1990) demostraron que las operaciones morfológicas podían formularse a partir de la teoría de retículos completos (complete lattices). Este nuevo marco de trabajo permitió introducir a la morfología matemática en términos algebraicos.

A continuación se presentan algunas definiciones sobre conjuntos, se denotaran los conjuntos con letras mayúsculas (X, Y, Z, \dots) y los elementos que contienen con minúsculas (p, q, r, \dots).

Definición 3.2.1 *Dos conjuntos son iguales si están formados por los mismos elementos:*

$$X = Y \Leftrightarrow (p \in X \Rightarrow p \in Y \quad y \quad p \in Y \Rightarrow p \in X) \quad (3.1)$$

La igualdad de conjunto es reflexiva, simétrica y transitiva. El conjunto vacío sera denotado por \emptyset

Definición 3.2.2 *X es subconjunto de Y si todos los elementos de X pertenecen a Y :*

$$X \subseteq Y \Leftrightarrow (p \in X \Rightarrow p \in Y) \quad (3.2)$$

La inclusión es reflexiva, antisimétrica y transitiva.

Definición 3.2.3 *La intersección de dos conjuntos X y Y es el conjunto de los elementos que pertenecen a ambos conjuntos:*

$$X \cap Y = (p \mid p \in X \quad y \quad p \in Y) \quad (3.3)$$

La intersección es conmutativa, asociativa e idempotente. Esta ultima propiedad es importante en morfología y significa que $X \cap X = X$.

Definición 3.2.4 *La unión de dos conjuntos se constituye por los elementos que pertenecen a uno o al otro.*

$$X \cup Y = (p \mid p \in X \quad o \quad p \in Y) \quad (3.4)$$

Al igual que la intersección, la unión de conjuntos es conmutativa, asociativa e idempotente.

Definición 3.2.5 *La diferencia entre conjuntos X e Y , la componen los elementos que pertenecen a X pero que no están incluidos en Y .*

$$X \setminus Y = (p \mid p \in X \text{ y } p \notin Y) \quad (3.5)$$

Esta operación no es en general conmutativa. Una variante de diferencia se conoce como diferencia simétrica de conjuntos, definida por los elementos que pertenecen a uno y a otro conjunto, pero nunca a ambos. La diferencia simétrica es conmutativa y asociativa.

Definición 3.2.6 *La complementación de un subconjunto X , perteneciente a un conjunto Y (conjunto de referencia) se define como:*

$$X^C = (p \mid p \notin X \text{ y } p \in Y) \quad (3.6)$$

Definición 3.2.7 *Dado un conjunto no vacío X , una relación binaria ' \leq ' en X es un orden parcial si cumple las siguientes propiedades:*

- $x \leq x$ (reflexiva).
- $x \leq y, \quad y \leq x$ implica que $x = y$ (antisimétrica).
- $x \leq y, \quad y \leq z$ implica que $x \leq z$ (transitiva).

Para cualquier $x, y, z \in X$. Un conjunto con una relación de este tipo será un conjunto que presenta orden parcial y se denotará como (X, \leq) . El conjunto será totalmente ordenado si todos los elementos que componen son comparables, es decir: $x \leq y$ ó $y \leq x$ para cualquier par $(x, y) \in X$.

Definición 3.2.8 *Sea (X, \leq) un conjunto ordenado y $S \subset X$, un conjunto no vacío de X :*

- Un elemento $x \in S$, es el menor elemento de S (mínimo) si $x \leq y$, para todo $y \in S$.
- Un elemento $y \in S$, es el mayor elemento de S (máximo) si $x \leq y$, para todo $x \in S$.
- Un elemento $x \in X$, es cota inferior de S si $x \leq y$, para todo $y \in S$.
- Un elemento $y \in Y$, es cota superior de S si $x \leq y$, para todo $x \in S$.

- Un elemento $x \in X$, es extremo inferior o ínfimo de S si y solo si es cota inferior de S y para toda cota inferior i de S se verifica que $i \leq x$. (Es la mayor de las cotas inferiores). Si este elemento existe es único y se denota por \wedge .
- Un elemento $y \in X$, es extremo superior o supremo de S si y solo si es cota superior de S y para toda cota superior s de S se verifica que $y \leq s$. (Es la menor de las cotas superiores). Si este elemento existe es único y se denota por \vee .

Definición 3.2.9 Un conjunto ordenado (X, \leq) es un retículo completo si cada subconjunto $Y \subseteq X$ poseen un ínfimo $\wedge Y \in X$ y un supremo $\vee Y \in X$.

Definición 3.2.10 Sean X e Y dos retículos completos. La relación f es una anamorfosis si y solo si f es una biyección que conserva el ínfimo y el supremo:

$$f(\wedge\{x_i \mid i \in I\}) = \wedge\{f(x_i) \mid i \in I\} \quad (3.7)$$

$$f(\vee\{x_i \mid i \in I\}) = \vee\{f(x_i) \mid i \in I\} \quad (3.8)$$

para cualquier familia $\{x_i \mid i \in I\}$ en X , donde I es un conjunto de índices.

El concepto de retículo completo es la base para la formulación de la morfología matemática. Los operadores morfológicos de base deben conservar el orden presente en la estructura de retículo, deben ser crecientes. Un operador ψ , en un retículo completo X , es creciente si:

$$x \leq y \Rightarrow \psi(x) \leq \psi(y) \quad (3.9)$$

3.2.2. Propiedades de base de las transformaciones morfológicas

Toda operación morfológica es el resultado de una o más operaciones de conjuntos (unión, intersección, complementación, etc) haciendo intervenir dos conjuntos X , Y , ambos subconjuntos de un conjunto de espacio Z . De los dos subconjuntos, Y recibe el nombre de elemento estructurante que, para operar con X , se desplazara a través del espacio Z .

Las operaciones morfológicas, transformaciones de conjuntos $\psi(X)$, satisfacen, entre otras, las siguientes propiedades:

Invariabilidad a translación

$$\psi(X_p) = (\psi(X))_p \quad (3.10)$$

donde p es el factor de translación del conjunto.

Compatibilidad con las homotecias

Supongamos que λX es una homotecia de un conjunto de puntos X , por tanto, las coordenadas de cada punto del conjunto se multiplican por alguna constante positiva λ . Esto es equivalente a cambiar de escala con respecto a algún origen. Si ψ no depende de la escala, es invariante al cambio de escalas.

$$\psi(\lambda X) = \lambda\psi(x) \quad (3.11)$$

Conocimiento local

La transformación morfológica ψ posee el principio de conocimiento local si para cualquier conjunto de puntos M , subconjunto del dominio N , la transformación del conjunto X restringido al dominio de M , y después restringido al dominio N , es equivalente a aplicar la transformación $\psi(X)$ y restringir el resultado en M :

$$\psi(X \cap N) \cap M = \psi(X) \cap M \quad (3.12)$$

Continuidad

En forma resumida, este principio afirma que la transformación morfológica ψ no exhibe ningún cambio abrupto. La noción de continuidad depende de la noción de vecindad, es decir de la topología.

Dependiendo del tipo de operación morfológica se cumplirán otra serie de propiedades como la idempotencia, extensividad, antiextensividad, etc.

3.2.3. El elemento estructural

Las transformaciones en morfología matemática utilizan un conjunto geométrico conocido como elemento estructural el cual posee forma, tamaño definido a priori y un origen. El elemento estructural se traslada sobre la imagen bajo estudio con la finalidad de determinar el conjuntos de puntos (respecto al origen del elemento estructural) que

intervendrán en cada operación que se efectúe con las componentes de la imagen.

La forma que adquiere el elemento estructural puede ser variada, por ejemplo el caso bidimensional, puede ser un disco, un cuadrado, una línea recta entre otros. Sin embargo no se limita su forma al caso bidimensional, en tridimensional se definen diferentes clases de poliedros por ejemplo el cubo, el prisma hexagonal, etc.

La ubicación del origen del elemento estructural juega un papel fundamental al efectuarse las transformaciones correspondientes en la imagen, ya que a partir de la ubicación de dicho punto, se obtendrá una forma precisa de la imagen procesada.

Se dice que un elemento estructural es *isotrópico*, si el conjunto geométrico que lo constituye posee características idénticas en todas las direcciones que es posible considerar dentro del espacio en que se localiza.

En este trabajo se considera un círculo de radio variable mayor que cero como el elemento estructural con el cual se aplicarán los algoritmos morfológicos. Este elemento estructural es simétrico lo que garantiza que las propiedades de algunas transformaciones se cumplan sin dificultad. Al elemento estructural con origen en p se le denota regularmente como B_p , para simplificar la notación se considera que $B_p = B$. La figura (3.1) muestra algunas de las formas del elemento estructural, junto con su centro u origen (p).

La talla del elemento estructural que se denota como μ donde μ es un factor de escala ó tamaño. La talla del elemento estructural es un escalar que determina el tamaño de la estructura geométrica de B .

La talla del elemento estructural involucra a la suma de Minkowski. Consideremos la suma de Minkowski entre dos conjuntos X y A Serra (1982).

Definición 3.2.11 (Suma y resta de Minkowski) Sea E un conjunto y $\wp(E)$ el conjunto de todos los subconjuntos de E . Sea $E = R^n$ o $E = Z^n$. La adición y sustracción de Minkowski para subconjuntos X, A de E esta dada por:

$$X \oplus A = \bigcup_{a \in A} X_a \quad \text{y} \quad X \ominus A = \bigcap_{a \in A} X_{-a}$$

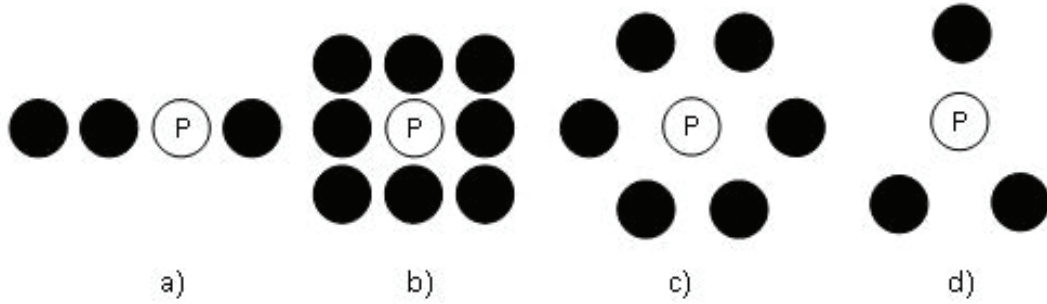


Figura 3.1: Diferentes elementos estructurales. (a) línea horizontal como elemento estructural, con origen en el punto p, (b) cuadrado como elemento estructural con origen en el punto p, (c) círculo como elemento estructural, con origen en el punto p, (d) triángulo como elemento estructural, con origen en el punto p.

donde

$$X_a = \{x + a : x \in X\}$$

Es la traslación de X por el vector $a \in E$, $x + a$ es la suma de x y a ; $-a$ es la reflexión de a .

Los símbolos \oplus y \ominus representan la suma y substracción de Minkowski respectivamente.

Empleando la definición 3.2.11, la talla de un elemento estructural se obtiene mediante la siguiente expresión:

$$\mu B = B \oplus \underbrace{\check{B} \dots \oplus \check{B}}_{\mu-1 \text{ sumas}} \quad \text{con} \quad \check{B} = \{-b : b \in B\}$$

3.2.4. Caso binario de la erosión y dilatación morfológicas

La erosión y la dilatación morfológicas son definidas de una manera general a partir de funciones estructurales. En el caso de imágenes binarias, la dilatación y el erosionado morfológicos se definen a continuación.

Definición 3.2.12 Sean $X, B \subseteq E$; se define la dilatación del conjunto X por el elemento estructural μB , denotado por $\delta_{\mu B}(X)$, como el operador sobre $\wp(E)$ dado por:

$$\delta_{\mu B}(X) = \{x | \mu \check{B}_x \cap X \neq \emptyset\} = \bigcup_{b \in \mu B} X_b \quad (3.13)$$

Definición 3.2.13 Sean $X, B \subseteq E$; se define la erosión del conjunto X por el elemento estructural μB , denotado por $\varepsilon_{\mu B}(X)$, como el operador sobre $\wp(E)$ dado

por:

$$\varepsilon_{\mu B}(X) = \{x \mid \mu B_x \subseteq X\} = \bigcap_{b \in \mu B} X_{-b} \quad (3.14)$$

Donde \check{B} representa el transpuesto de B(elemento estructural) respecto del origen.

El dilatado y el erosionado morfológico cumplen ser transformaciones crecientes, es decir, para todo conjunto X e Y, tal que $X \subset Y$ tenemos que:

$$\varepsilon_{\mu B}(X) \subset \varepsilon_{\mu B}(Y) \quad y \quad \delta_{\mu B}(X) \subset \delta_{\mu B}(Y)$$

La dilatación es una transformación extensiva y la erosión es antiextensiva; es decir, para todo X :

$$\varepsilon_{\mu B}(X) \subseteq X \quad y \quad X \subseteq \delta_{\mu B}(X)$$

Desde un punto de vista geométrico, la dilatación morfológica de un conjunto X por μB , es el lugar definido por el complemento del origen x de los trasladados de μB cuando dicho conjunto μB_x no toca al conjunto X. Mientras que la erosión morfológica es el lugar de centros de los trasladados de μB_x , cuando este se encuentra completamente al interior de X. La Fig. 3.2 muestra estas nociones desde un punto de vista geométrico.

De aquí en adelante se considera que las siguientes expresiones son iguales con la finalidad de simplificar la notación

$$\varepsilon_{\mu B} = \varepsilon_{\mu}, \quad \delta_{\mu B} = \delta_{\mu}$$

3.2.5. Caso niveles de gris de la erosión y dilatación morfológicas

El erosionado $\varepsilon_{\mu B}$ y dilatado $\delta_{\mu B}$ morfológicos en niveles de gris se definen como:

$$\varepsilon_{\mu B}(f(x)) = \bigwedge \{f(y); y \in \mu \check{B}_x\} \quad (3.15)$$

El dual de la ecuación 3.15 es el dilatado morfológico,

$$\delta_{\mu B}(f(x)) = \bigvee \{f(y); y \in \mu \check{B}_x\} \quad (3.16)$$

La dilatación sobre funciones es extensiva, mientras que la erosión es antiextensiva, es decir,

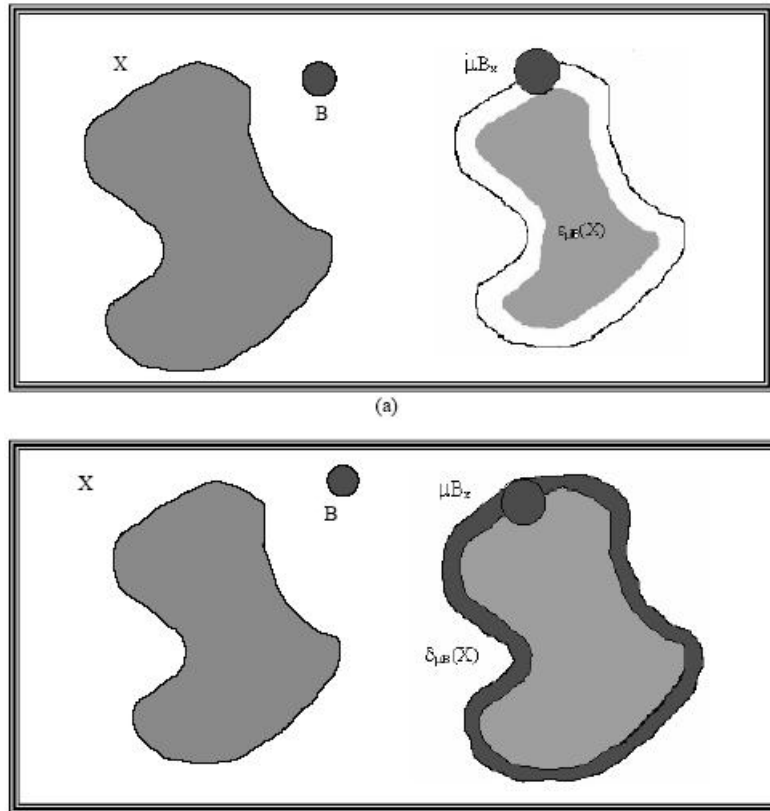


Figura 3.2: Punto de vista geométrico del dilatado y erosionado binario. Estas transformaciones se interpretan desde un punto de vista geométrico (a) el erosionado morfológico de un conjunto X por μB es el lugar de centros de los trasladados de μB_x , cuando éste se encuentra completamente al interior de X , (b) el dilatado de un conjunto X por μB , es el lugar definido por el complemento del origen x de los trasladados de B cuando dicho conjunto μB_x no toca el conjunto X .

$$f(x) \leq \delta_{\mu B}(f)(x) \quad y \quad f(x) \geq \varepsilon_{\mu B}(f)(x)$$

Y cumplen ser transformaciones crecientes, es decir:

$$\forall f, g, \text{ con } g \leq f \quad \varepsilon_{\mu B}(g)(x) \leq \varepsilon_{\mu B}(f)(x) \quad y \quad \delta_{\mu B}(g)(x) \leq \delta_{\mu B}(f)(x)$$

3.2.6. Gradiente morfológico

El residuo de dos operaciones o transformaciones ψ y ζ es su diferencia, en el caso de conjuntos, definimos esta diferencia como:

$$\rho_{\psi\zeta}(X) = \psi(X) \setminus \zeta(X) \quad (3.17)$$

en el caso de funciones tenemos:

$$\rho_{\psi\zeta}(f) = \psi(f) - \zeta(f) \quad (3.18)$$

El primer residuo de operaciones que se puede definir en morfología matemática es el gradiente morfológico, siendo la primera aproximación de la segmentación morfológica. El residuo gradiente, conocido en morfología como *gradiente de Beucher* es la diferencia entre la dilatación y una erosión; una dilatación y la imagen original; o una diferencia entre la imagen original y su erosión.

El primero de los gradientes a definir se conoce como gradiente por erosión y es la diferencia entre el conjunto o imagen original y la erosión por un elemento estructurante μB . En el caso de conjuntos:

$$\rho_{\mu B}^-(X) = X \setminus \varepsilon_{\mu B}(X) \quad (3.19)$$

En el caso de funciones este gradiente lo podemos definir como:

$$\rho_{\mu B}^-(f) = f - \varepsilon_{\mu B}(f) \quad (3.20)$$

El gradiente por dilatación se define como la diferencia entre una dilatación por el elemento estructural μB y la identidad (imagen original):

$$\rho_{\mu B}^+(X) = \delta_{\mu B}(X) \setminus (X) \quad (3.21)$$

en el caso de conjuntos. Cuando el gradiente se define para funciones tenemos:

$$\rho_{\mu B}^+(f) = \delta_{\mu B}(f) - f \quad (3.22)$$

El gradiente simétrico se define como una diferencia entre la dilatación y la erosión de una imagen por un elemento estructurante μB , para el caso de conjuntos y funciones este gradiente simétrico se define como sigue:

$$\rho_{\mu}B(X) = \delta_{\mu B}(X) \setminus \varepsilon_{\mu B}(X) \quad (3.23)$$

$$\rho_{\mu B}(f) = \delta_{\mu B}(f) - \varepsilon_{\mu B}(f) \quad (3.24)$$

La elección de un gradiente por erosión o dilatación depende de la geometría y la luminancia o intensidad de los objetos a destacar en la imagen. En la figura 3.3 se pueden observar las diferencias entre gradientes para imágenes de grises. Cabe destacar como el gradiente simétrico intensifica la detección de bordes. El gradiente morfológico de Beucher es invariante a la complementación $\rho = \rho C$. Asimismo, el gradiente por erosión y el gradiente por dilatación son operaciones complementarias entre si. Se puede comprobar también como el gradiente simétrico es la suma de los gradientes por erosión y dilatación. El gradiente morfológico puede usarse en conjunción con la umbralización para realizar una detección de contornos en imágenes de escala de grises.

3.2.7. Filtros morfológicos básicos. Apertura y cerradura morfológicas

A partir de las transformaciones básicas, erosión y dilatación morfológicas, es posible construir filtros más complejos, un filtro morfológico se define como:

Definición 3.2.14 *Un mapeo $\psi : \wp(E) \rightarrow \wp(E)$ es un filtro morfológico si y solamente si es creciente e idempotente.*

Las definiciones formales de la apertura y cerradura morfológicas están dadas tanto para el caso binario como para el caso numérico. En el caso binario, la apertura y la cerradura morfológicas se definen como:

$$\gamma_{\mu B}(X) = \delta_{\mu \check{B}} \varepsilon_{\mu B}(X) = (X \ominus \mu B) \oplus \mu \check{B} \quad (3.25)$$

y

$$\varphi_{\mu B}(X) = \varepsilon_{\mu \check{B}} \delta_{\mu B}(X) = (X \oplus \mu B) \ominus \mu \check{B}$$

La Fig. 3.4 muestra la interpretación geométrica de la apertura y la cerradura morfológicas. La cerradura de X por B será el complemento del área barrida por B

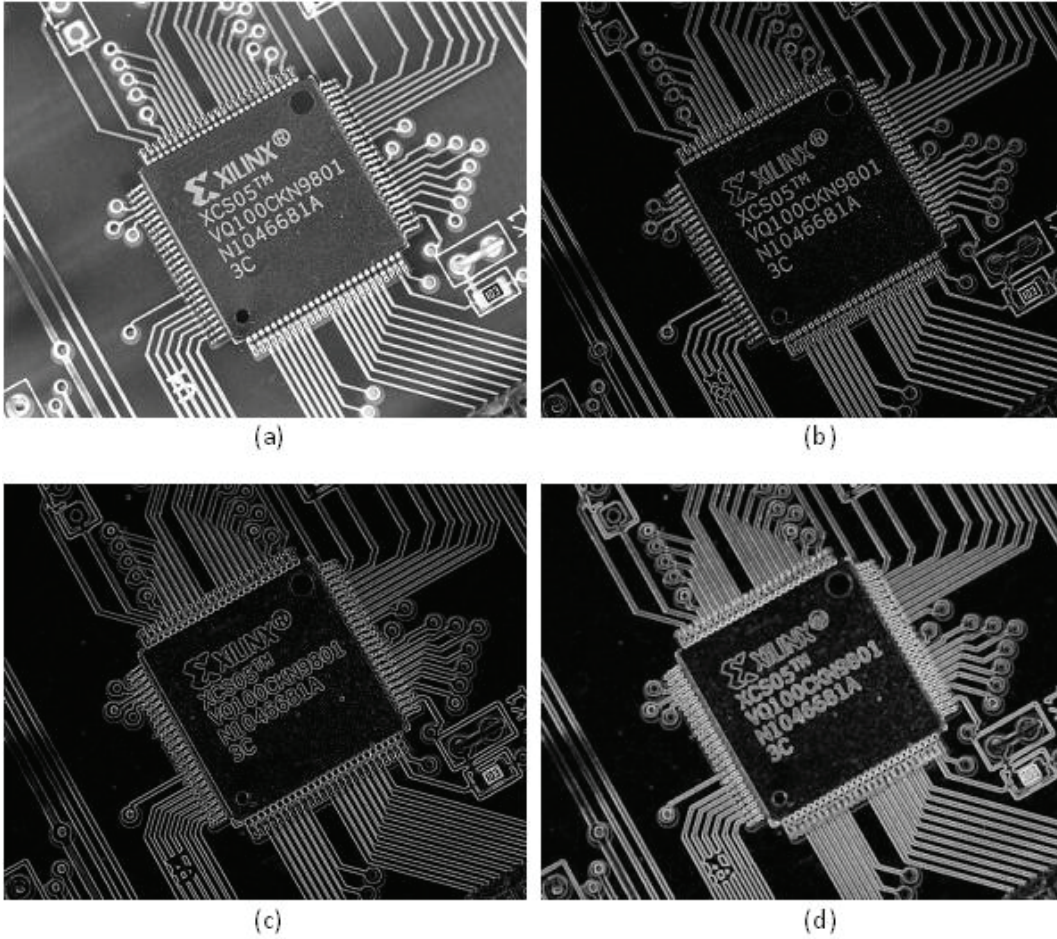


Figura 3.3: Gradientes morfológicos de una imagen. (a) imagen original, (b) gradiente morfológico por erosión, (c) gradiente morfológico por dilatación, (d) gradiente morfológico simétrico.

cuando éste se encuentre al exterior de X . Mientras que la apertura de X por B será el área de barrida por B cuando éste se encuentre completamente al interior de X .

En el caso numérico:

$$\gamma_{\mu B}(f)(x) = \delta_{\mu \check{B}}(\varepsilon_{\mu B}(f))(x) \quad (3.26)$$

y

$$\varphi_{\mu B}(f)(x) = \varepsilon_{\mu \check{B}}(\delta_{\mu B}(f))(x)$$

Las siguientes propiedades son verificadas por estas transformaciones (caso de conjuntos). Sean X, Y dos conjuntos.

(a) La $\gamma_{\mu B}$ es antiextensiva, mientras que la $\varphi_{\mu B}$ es extensiva.

$$\gamma_{\mu B} \subseteq X \quad y \quad X \subset \varphi_{\mu B}$$

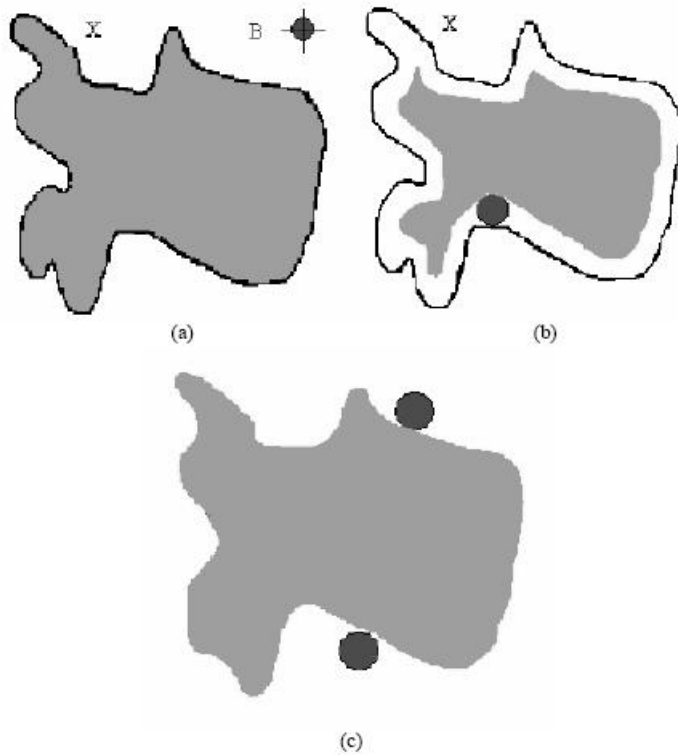


Figura 3.4: Punto de vista geométrico de la apertura y cerradura morfológicas. Se realiza una interpretación desde un punto de vista geométrico. (a) El conjunto X y el elemento estructural B , (b) el abierto de X por μB será el área barrida por μB cuando éste se encuentre completamente al interior de X , (c) el cerrado de X por μB será el complemento del área barrida por μB cuando éste se encuentra completamente al exterior de X .

(b) Son transformaciones crecientes

$$X \subset Y \Rightarrow \gamma_{\mu B}(X) \subset \gamma_{\mu B}(Y) \quad y \quad X \subset Y \Rightarrow \varphi_{\mu B}(X) \subset \varphi_{\mu B}(Y)$$

(c) Son transformaciones idempotentes

$$\gamma_{\mu B} \cdot \gamma_{\mu B}(X) = \gamma_{\mu B}(X) \quad y \quad \varphi_{\mu B} \cdot \varphi_{\mu B}(X) = \varphi_{\mu B}(X)$$

En el caso de funciones.

(a) La $\gamma_{\mu B}$ es antiextensiva, mientras que la $\varphi_{\mu B}$ es extensiva.

$$\gamma_{\mu B}(f) \leq f \quad y \quad f \leq \varphi_{\mu B}(f)$$

(b) Son transformaciones crecientes

$$\forall f, g \text{ con } g \leq f \text{ se cumple que } \gamma_{\mu B}(g) \leq \gamma_{\mu B}(f) \quad y \quad \varphi_{\mu B}(g) \leq \varphi_{\mu B}(f)$$

(c) Son transformaciones idempotentes

$$\gamma_{\mu B} \cdot \gamma_{\mu B}(f) = \gamma_{\mu B}(f) \quad y \quad \varphi_{\mu B} \cdot \varphi_{\mu B}(f) = \varphi_{\mu B}(f)$$

3.3. Filtros por reconstrucción

Para algunas aplicaciones de análisis de imágenes es conveniente restringir el campo de acción de una transformación a ciertas regiones de interés. Esta idea dió origen a una nueva clase de transformaciones conocidas como geodésicas. A diferencia de transformaciones morfológicas que actúan sobre toda la imagen, las geodésicas sólo actúan sobre parte de la imagen, i.e. sobre un subconjunto denominado máscara geodésica. Las transformaciones por reconstrucción o geodésicas son filtros que permiten modificar los mínimos y máximos de la imagen sin cambiar considerablemente la estructura de las demás componentes (Vincent (1993); Salembier (1995)). Las transformaciones por reconstrucción utilizan transformaciones geodésicas, las cuales son iteradas hasta la estabilidad, de tal manera que las componentes se reconstruyen fielmente de acuerdo al criterio creciente. a continuación se definen estas transformaciones para el caso binario.

(i) Sea $X \subseteq Y$, el erosionado geodésico de tamaño 1 del conjunto Y sobre el conjunto X se define como:

$$\varepsilon_X^1(Y) = \varepsilon_B(Y) \bigvee X \quad (3.27)$$

(ii) Sea $Y \subseteq X$, el dilatado geodésico de tamaño 1 del conjunto Y dentro del conjunto X se define como:

$$\delta_X^1(Y) = \delta_B(Y) \bigwedge X \quad (3.28)$$

La dilatación y la erosión por reconstrucción para el caso binario se definen como:

$$R(X, Y) = \underbrace{\delta_X^1 \delta_X^1 \dots \delta_X^1(Y)}_{\text{hasta la estabilidad}} = \lim_{n \rightarrow \infty} \delta_X^n(Y) \quad (3.29)$$

$$R^*(X, Y) = \underbrace{\varepsilon_X^1 \varepsilon_X^1 \dots \varepsilon_X^1(Y)}_{\text{hasta la estabilidad}} = \lim_{n \rightarrow \infty} \varepsilon_X^n(Y) \quad (3.30)$$

De las ecuaciones 3.29 y 3.30, si Y es igual a la erosión o dilatación morfológicas entonces se obtiene la apertura y el cerradura por reconstrucción respectivamente. La apertura y cerradura por reconstrucción se definen como:

$$\tilde{\gamma}_{\mu B} = \underbrace{\delta_X^1 \delta_X^1 \dots \delta_X^1(\varepsilon_{\mu B}(X))}_{\text{hasta la estabilidad}} = \lim_{n \rightarrow \infty} \delta_X^n(\varepsilon_{\mu B}(X)) \quad (3.31)$$

$$\tilde{\varphi}_{\mu B} = \underbrace{\varepsilon_X^1 \varepsilon_X^1 \dots \varepsilon_X^1(\delta_{\mu B}(X))}_{\text{hasta la estabilidad}} = \lim_{n \rightarrow \infty} \varepsilon_X^n(\delta_{\mu B}(X)) \quad (3.32)$$

$$\varepsilon_f^1(g) = f \bigvee \varepsilon_{\mu B}(g) \quad \text{con } g \geq f \quad (3.33)$$

$$\delta_f^1(g) = f \bigwedge \delta_{\mu B}(g) \quad \text{con } g \leq f \quad (3.34)$$

La dilatación y la erosión por reconstrucción para el caso numérico se definen como:

$$R(f, g) = \underbrace{\delta_f^1 \delta_f^1 \dots \delta_f^1(g)}_{\text{hasta la estabilidad}} = \lim_{n \rightarrow \infty} \delta_f^n(g) \quad (3.35)$$

$$R^*(f, g) = \underbrace{\varepsilon_f^1 \varepsilon_f^1 \dots \varepsilon_f^1(g)}_{\text{hasta la estabilidad}} = \lim_{n \rightarrow \infty} \varepsilon_f^n(g) \quad (3.36)$$

De las ecuaciones 3.35 y 3.36, si g es igual a la erosión o dilatación morfológicas, entonces la apertura y el cerradura por reconstrucción se definen como:

$$\tilde{\gamma}_{\mu B} = \underbrace{\delta_f^1 \delta_f^1 \dots \delta_f^1(\varepsilon_{\mu B}(f))}_{\text{hasta la estabilidad}} = \lim_{n \rightarrow \infty} \delta_f^n(\varepsilon_{\mu B}(f)) \quad (3.37)$$

$$\tilde{\varphi}_{\mu B} = \underbrace{\varepsilon_f^1 \varepsilon_f^1 \dots \varepsilon_f^1(\delta_{\mu B}(f))}_{\text{hasta la estabilidad}} = \lim_{n \rightarrow \infty} \varepsilon_f^n(\delta_{\mu B}(f)) \quad (3.38)$$

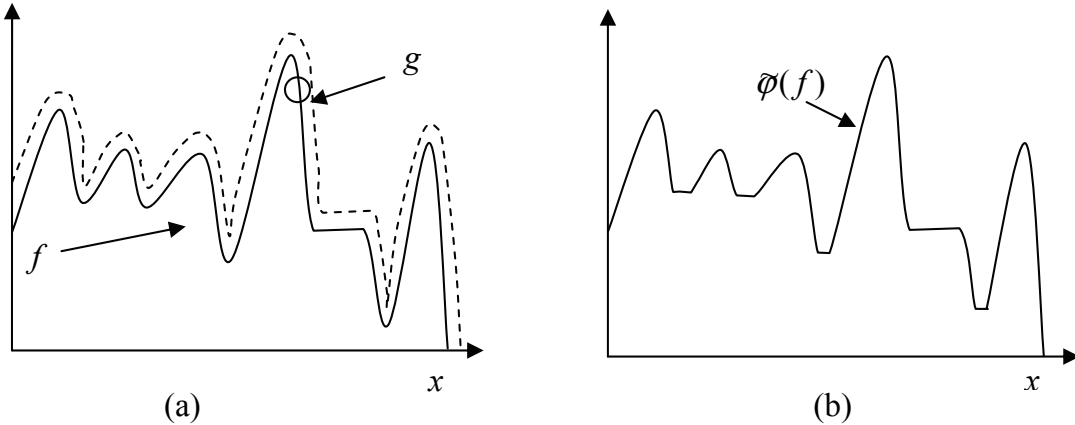


Figura 3.5: (a) Imagen original f y el marcador $g = \delta(f)$, (b) cerradura por reconstrucción.

La apertura y cerradura por reconstrucción fusionan los máximos y mínimos de la imagen a medida que se incrementa el tamaño del elemento estructural.

En la figura 3.5 se muestra el efecto de la cerradura por reconstrucción. La figura

3.5 muestra la función original f , así como el marcador g obtenido por medio de un dilatado morfológico. En la figura 3.5 (b) se muestra el resultado de la cerradura por reconstrucción $\bar{\varphi}_\mu$. Notesé como determinadas regiones asociadas con los mínimos regionales de la imagen original son fusionadas; mientras otras permanecen sin cambio. En el caso de la apertura por reconstrucción, los máximos son extendidos; mientras que los mínimos son extendidos en el caso de la cerradura por reconstrucción. En ambos casos los valores extremos se fusionan; pero en ningún caso se crean nuevos contornos. En este trabajo, los mínimos y máximos de la imagen se definen como sigue (Vincent 1993, Vachier 1995).

Definición 3.3.1 *Un mínimo regional M_{in} (máximo regional M_{ax} respectivamente) en una imagen f en niveles de gris, es un conjunto de píxeles conexos con un cierto valor h (región a una altura h), tal que cada píxel en la vecindad de M_{in} (M_{ax} respectivamente), tienen un valor más alto (más bajo respectivamente).*

Determinar el máximo regional en una imagen numérica (en escala de grises) es relativamente sencillo y se han propuesto varios algoritmos en la literatura, algunos de los cuales pueden consultarse en el trabajo de Vincent (1990). Uno de los métodos mas eficientes utiliza la reconstrucción numérica y esta basado en la siguiente proposición:

Definición 3.3.2 *La imagen binaria $M(I)$ de el máximo regional de I esta dado por:*

$$M(I) = I - \rho_I(I - 1) \quad (3.39)$$

Definición 3.3.3 *La reconstrucción de X con el marcador $Y \subseteq X$, se obtiene iterando la dilatación geodésica elemental de Y dentro de X hasta que la estabilidad es alcanzada. En otras palabras:*

$$\rho_X(Y) = \bigcup_{n \geq 1} \delta_X^{(n)}(Y). \quad (3.40)$$

Este algoritmo es la base para uno de los algoritmos mas simples para la evaluación de la reconstrucción geodésica en imágenes binarias o en escala de grises. Hay que hacer notar que las definiciones arriba citadas pueden “invertirse” para extraer los mínimos regionales de la imagen numérica.

Las transformaciones por reconstrucción no modifican los contornos de la imagen procesada; es decir, son transformaciones que preservan características de la imagen

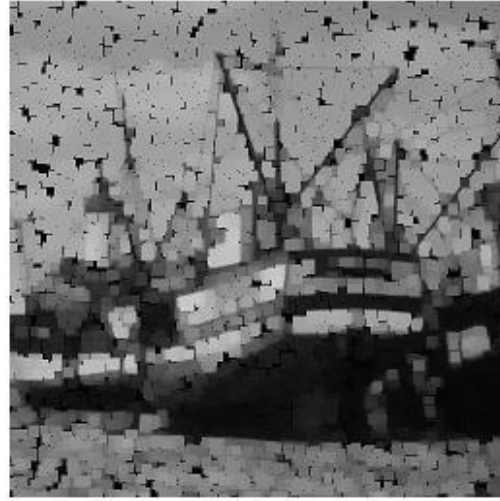
original; mientras que aquellas componentes que no cumplen con cierto criterio creciente dado por el tamaño del elemento estructural son eliminadas por completo.

La aplicación de la apertura por reconstrucción se ilustra en las figuras 3.6(b) y (c). En la figura se compara el desempeño entre transformaciones por reconstrucción (apertura por reconstrucción) y filtros morfológicos normales (apertura morfológica), a la imagen original se le ha agregado una componente de ruido del tipo “sal y pimienta”.

Se observa que la imagen de salida obtenida a partir de la apertura morfológica presenta mayor influencia al ruido de la imagen original, que para el caso de la apertura por reconstrucción.



(a)



(b)



(c)

Figura 3.6: (a) Imagen original con ruido agregado, (b) Cerradura morfológica y (c) cerradura por reconstrucción. En ambos filtros se utilizó un elemento estructural de talla $\mu = 3$.

3.4. Filtrado en el espacio de escala

Una propiedad inherente de los objetos del mundo real es que solo existen como entidades significativas sobre un cierto rango de escalas. Un ejemplo sencillo es el concepto de una rama de un árbol, el cual solo tiene sentido solo a cierta escala, digamos, de algunos centímetros y a lo mucho a unos pocos metros. Sería ilógico discutir el concepto de árbol a nivel nanométrico o kilométrico, a estos niveles sería más relevante hablar sobre las moléculas que forman las hojas de los arboles, y del bosque donde crece el árbol, respectivamente. El filtrado en el espacio de escala es una manera de asociar descriptores de una señal a través de múltiples escalas, esta técnica proporciona un medio para “.. gestionar la ambigüedad de la escala de una manera organizada y natural” Witkin (1983).

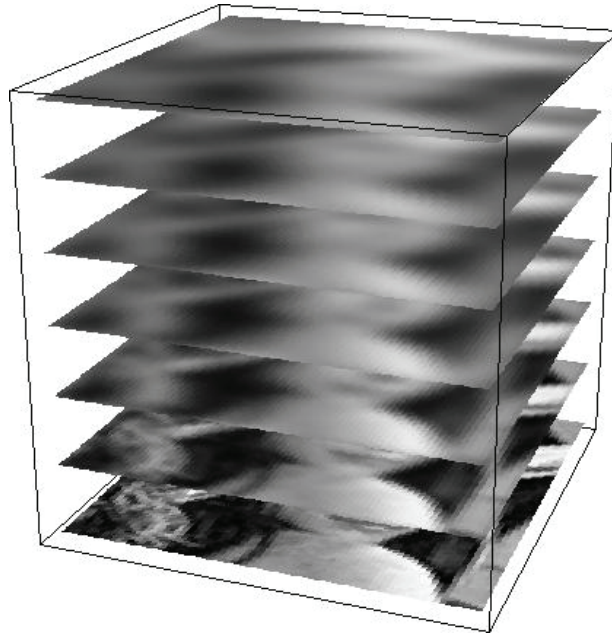


Figura 3.7: Espacio de escala Gaussiano de la imagen “Lena” para un rango de escalas de $\sigma = 1, 2, 3, 4, 5$ y 6 pixeles.

3.4.1. Filtrado en el espacio de escala lineal

Anteriormente se introdujo el filtrado en el espacio de escala, sin embargo a continuación se presenta desde un punto de vista mas técnico y riguroso. El concepto de espacio de escala fue nombrado, formalizado y llevado al análisis de imágenes por Witkin (1983). El argumento principal detrás de este marco de trabajo es que si no hay información previa disponible sobre cuales son las escalas adecuadas para un conjunto de datos determinado, entonces es razonable analizar los datos de entrada a múltiples escalas. Esto requiere que la señal de datos original sea apilada junto a una familia de

señales derivadas, en la cual las estructuras más finas de la señal son sucesivamente removidas, en la Fig. 3.7 se ilustra dicha idea, a escalas más bajas (en la parte inferior de la pila) la representación en el espacio de escala tiende al valor de la imagen original, mientras que a escalas más altas (parte superior) se van difuminando los rasgos de la imagen.

Si consideramos una señal N dimensional $f : \mathbb{R}^N \rightarrow \mathbb{R}$ y un kernel de suavizado $g : \mathbb{R}^N \times \mathbb{R}_+ \setminus \{0\} \rightarrow \mathbb{R}$, entonces la señal suavizada a escala σ esta dada por $F : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$

$$F(x, \sigma) = f(x) * g(x, \sigma) \quad (3.41)$$

donde $*$ denota la operación de convolución, F es una función en el espacio $(n + 1)$ -dimensional llamado “espacio de escala” y es conocido como la *imagen del espacio de escala* (también llamada *representación multiescala*) de la señal. A partir de la función Gaussiana se define el kernel de convolución como sigue:

$$g(x; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (3.42)$$

Para el caso de las señales 2D (imágenes), considerando la separabilidad de la función Gaussiana, la convolución 1D se aplica dos veces, primero en los renglones y luego en las columnas de la imagen, la forma para Kernel de convolución Gaussiano 2D se muestra en la Fig. 3.8. El espacio de escala, en vista apilada, obtenido con el operador de convolución Gaussiano 2D para la imagen “Lena” se muestra en la Fig. 3.7.

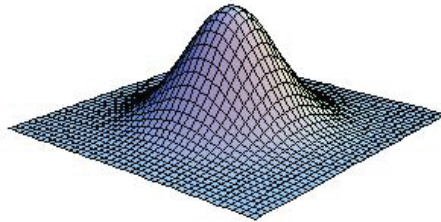


Figura 3.8: Kernel de convolución Gaussiano 2D.

además la varianza σ de este kernel es referido como el *parámetro de escala*, se usa la referencia “escala gruesa” cuando el valor de escala es grande y de manera similar se usa el término “escala fina” cuando el valor de σ es más pequeño. La idea central en el trabajo de Witkin (1983) es que los rasgos más importantes de la señal deben persistir a través de escalas relativamente gruesas aun cuando su localización pueda distorsionarse en el proceso de filtrado.

Para que una representación multiescala sea aceptada como un espacio de escala, debe cumplir ciertas propiedades. Estas propiedades son intuitivas, y están basadas en el concepto de simplificación sucesiva de la señal conforme se incrementa el valor de escala. La primera de estas propiedades es la *fidelidad*. Esta propiedad se refiere simplemente al requerimiento de que conforme el parámetro de escala se aproxima a cero, el espacio de escala de la señal se aproxima a la señal original. Este requerimiento puede escribirse como:

$$\lim_{\sigma \rightarrow 0} F(x, \sigma) = f(x) \quad (3.43)$$

Otra propiedad requerida es la *causalidad*, esto es, cada nivel del espacio de escala $F(\sigma_1)$ depende únicamente de $F(\sigma_2)$ si $\sigma_1 > \sigma_2$. De este modo, las escalas más gruesas son derivadas de las escalas más finas, pero no a la viceversa.

La propiedad de *monotonidad* de las características de la señal establece que el número de características de la señal debe decrementarse monotonamente conforme el parámetro de escala aumenta. En la literatura, el término de “causalidad” se usa genéricamente para describir la propiedad de monotonidad Koenderink (1984). De acuerdo a esta propiedad, las características solo pueden desaparecer pero no puede generarse nueva información conforme la escala aumenta. La definición de características de la escala depende del procedimiento de filtrado usado. Si $\#[F(x)]$ denota el número de características en una señal, entonces de acuerdo al criterio de monotonidad se requiere que:

$$\#[F(x, \sigma_1)] \leq \#[F(x, \sigma_2)] \quad \text{para todo } \sigma_1 > \sigma_2 > 0 \quad (3.44)$$

Una característica o rasgo puede ser un punto aislado, o uno de un conjunto de puntos, que representan información útil acerca de la señal. Los rasgos deben representar alguna simplificación o reducción de información que recolectan la información esencial de la representación del espacio de escala. Extrayendo estos rasgos a diferentes escalas y siguiéndolas a través del espacio de escala, las tareas de procesamiento de señal se simplifican y se hacen más eficientes.

Algunas tareas típicas que utilizan extracción de características son video seguimiento, problema de correspondencia en imágenes (como en visión estereo), segmentación de la imagen, etc. Algunas definiciones comunes de características incluyen los cruces por cero en la segunda derivada o Laplaciano, extremos locales de la señal, o contornos definidos por un detector de contornos particular.

además de la propiedad de monotonicidad, las características en el espacio de escala requiere se exhiba una propiedad de *continuidad*. No solo el numero de características debe disminuir al incrementar la escala, sino que también estas características que permanecen a escalas más altas debe corresponder con características encontradas a escalas más bajas.

Las propiedades de fidelidad, causalidad, monotonicidad y continuidad son algunos de los requerimientos más importantes para que una representación multiescala, obtenida a partir de un filtrado específico, sea considerado convenientemente como un espacio de escala.

3.4.2. Fingerprints

Como se explicó anteriormente, en la representación multiescala de una señal, se debe definir los puntos característicos de importancia del espacio de escala a partir de los cuales se analizan las propiedades multiescala del filtro. Se ha demostrado que el único filtro que satisface la propiedad de causalidad para señales 1D, si se consideran los cruces por cero de la representación multiescala como puntos característicos, es el operador de convolución Gaussiano de acuerdo a Babaud et al. (1986). Si \mathbf{Z} denota el conjunto de puntos de cruces por cero a determinada escala, esto es:

$$\mathbf{Z}_f(\sigma) = \{x : f(x) * g(x, \sigma) \text{ es un cruce por cero}\}. \quad (3.45)$$

La propiedad de causalidad Yuille (1985) para señales 1-D se puede establecer como sigue:

Definición 3.4.1 *Causalidad de los cruces por cero en señales 1-D: Si $\text{Card}[\mathbf{Z}]$ denota la cardinalidad de el conjunto \mathbf{Z} , entonces la causalidad para los cruces por cero se cumple, para todo, $0 \leq \sigma_1 \leq \sigma_2$*

$$\text{Card}[\mathbf{Z}_f(\sigma_1)] \geq \text{Card}[\mathbf{Z}_f(\sigma_2)] \quad (3.46)$$

Si graficamos $\mathbf{Z}(\sigma)$ contra σ , obtendremos un conjunto de líneas serpenteadas en el espacio de escala, las cuales se les dio el nombre de “*fingerprints*” de la señal por Yuille (1985). Se le dio este nombre debido al parecido de la forma de las líneas a los lazos y espirales de las huellas digitales humanas (ver Figura 3.9). Los fingerprint sirven como un auxiliar en la visualización de la correspondencia y posible movimiento de las características conforme se incrementa la escala. Frecuentemente, como en el caso de

las huellas digitales humanas, se puede extraer información útil para identificar una señal u objeto.

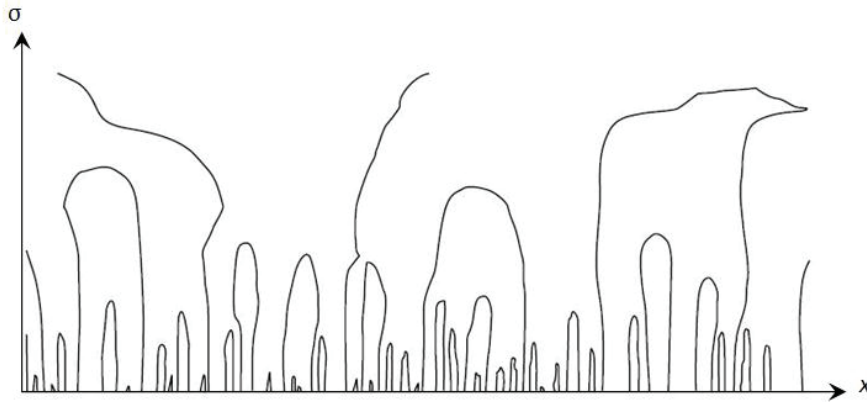


Figura 3.9: Fingerprints en el espacio de escala Gaussiano: gráfica de los cruces por cero de una señal 1-D.

3.4.3. Espacio escala dilatación-erosión morfológico

Las operaciones morfológicas, permiten modificar la estructura de una señal de manera ordenada, en consecuencia varios autores las han encontrado aplicables como filtros de suavizado para el espacio de escala. Uno de los primeros trabajos donde se propuso el empleo de operadores morfológicos para obtener una representación multiescala fue el presentado por Chen (1989), más tarde otros autores decidieron continuar con el estudio de las propiedades multiescala de los operadores morfológicos Jang (1991); Nacken (1994); Park (1996); Jackway et al. (1994); Jackway (1996, 1998).

Aunque el trabajo de Chen (1989) fue de los primeros en el espacio de escala morfológico, fue Jackway et al. (1994); Jackway (1996) quien desarrolló una teoría sólida al definir las propiedades multiescala de un operador dilatación-erosión morfológico, en sus estudios define los puntos característicos a partir de los extremos regionales obtenidos de la señal filtrada. A partir de esta metodología de análisis multiescala se encontraron soluciones a problemas de reconocimiento de patrones en mapas de relieve (Jackway (1996)) y tareas de localización y reconocimiento de rostros en los trabajos de Raducanu (2000); Kotropoulos et al. (2000). Debido a que el operador multiescala que se propone en este trabajo de tesis basa su funcionamiento en este tipo de análisis multiescala, a continuación se presentan los fundamentos teóricos y definiciones de dicho análisis.

3.4.4. Morfología dependiente de la escala

Las operaciones de erosión y dilatación morfológica (ecuaciones 3.15 y 3.16 respectivamente) se pueden hacer dependientes de un factor de escala por medio del uso de una *función estructural escalada*, $g_\sigma : G_\sigma \subset \mathbb{R}^n \rightarrow \mathbb{R}$,

$$g_\sigma(x) = |\sigma|g(|\sigma|^{-1}x) \quad x \in G_\sigma, \sigma \neq 0 \quad (3.47)$$

donde σ es el parámetro de escala y $G_\sigma = \{x : |\sigma|^{-1}x \in G\}$. Esta forma de escalar una función es compatible con la idea de una esfera escalada en morfología Euclidiana, con el radio como el parámetro de escala, el cual puede escribirse como:

$$g_\sigma(x) = |\sigma| \left(\sqrt{1 - \|\sigma|^{-1}x\|} - 1 \right) \quad \|x\| \leq |\sigma|, \sigma \neq 0, \quad (3.48)$$

con $G_\sigma = \{t : \|x\| \leq |\sigma|\}$ y $\|\cdot\|$ representa la norma Euclidiana.

Con la función estructural escalada (3.48) se juntan las operaciones de dilatación y erosión para formar una sola operación multiescala que unifica las operaciones morfológicas como sigue:

Definición 3.4.2 Dilatación-Erosión Multiescala. La dilatación-erosión multiescala de un señal $f(x)$ por la función estructural escalada $g_\sigma(x)$ se denota por $f \otimes g_\sigma$, y se define por:

$$(f \otimes g_\sigma)(x) = \begin{cases} (f \oplus g_\sigma)(x) & \text{si } \sigma > 0; \\ f(x) & \text{si } \sigma = 0; \\ (f \ominus g_\sigma)(x) & \text{si } \sigma < 0. \end{cases} \quad (3.49)$$

Esto es, para escalas positivas realizamos una dilatación, para escalas negativas una erosión. Con este método la escala puede ser negativa, esto es, con $|\sigma|$ se obtiene la noción intuitiva de la escala. Por otro lado las escalas positivas y negativas extraen diferentes aspectos de la forma de la señal de tal manera que ambas son necesarias para obtener una representación completa.

El efecto del filtrado de las operaciones morfológicas conforme varia el valor de escala (talla del elemento estructural) se puede apreciar en las Figs. 3.10 y 3.11 donde se presenta la representación en el espacio de escala para la erosión y dilatación, respectivamente.

Como puede observarse, para el caso de la erosión, en escalas más finas la estructura de la imagen tiende a la señal original, mientras que para escalas más gruesa la imagen se torna más oscura, puesto que la erosión pretende minimizar el valor de la

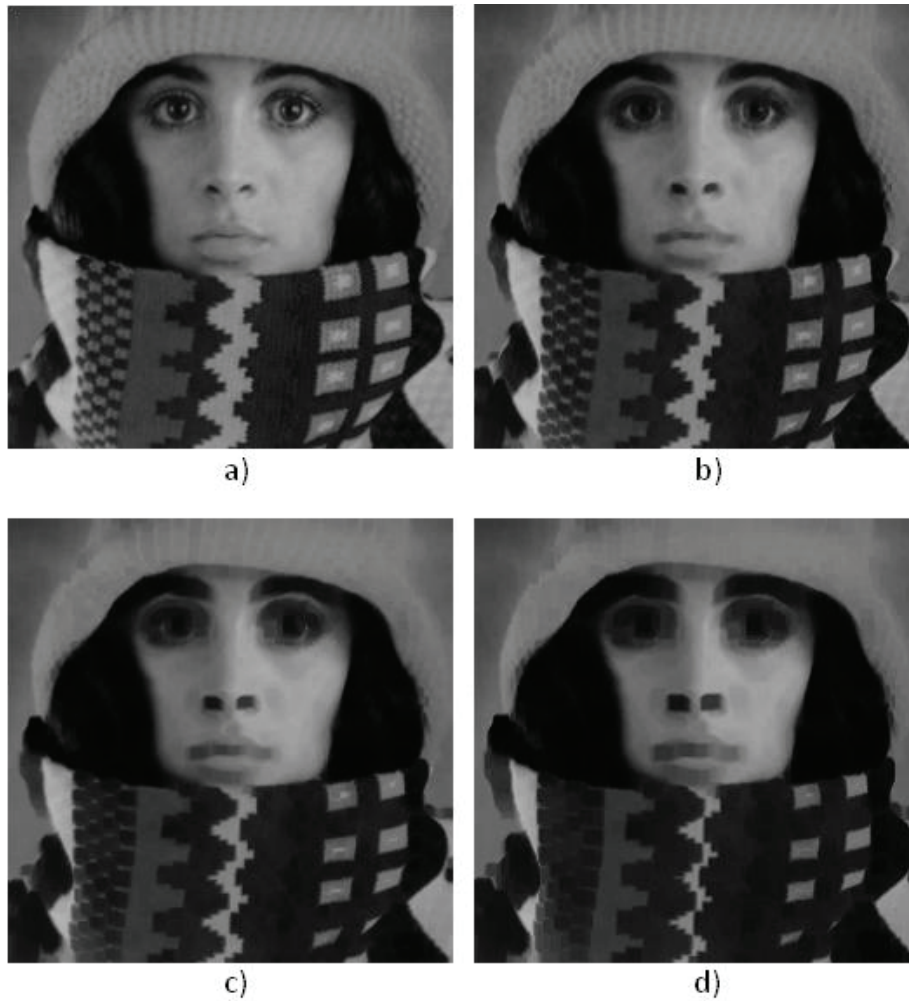


Figura 3.10: Diferentes niveles en la representación del espacio de escala de una imagen bidimensional (a) erosionada a niveles de escala $\sigma = 3, 5$ y 7 se muestra en recuadros (b), (c) y (d) respectivamente.

señal. Para el caso de la dilatación morfológica ocurre un fenómeno similar para las regiones más claras de la imagen.

Para el caso de imágenes 2D, siguiendo la formulación del espacio de escala, se puede considerar la dilatación-erosión multiescala como una función del espacio de escala \mathbb{R}^3 , dando la imagen del espacio de escala morfológico $F : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$ definido por:

$$F(x, y, \sigma) = (f \circledast g_\sigma)(x, y). \quad (3.50)$$

Esta formulación teórica contiene información acerca de la función en su forma original, y después de la dilatación y erosión en todas las escalas. En este espacio de escala se definieron los puntos característicos de la señal como el conjunto de los puntos extremos (máximos y los mínimos) locales, en el caso de imágenes 2D están definidos

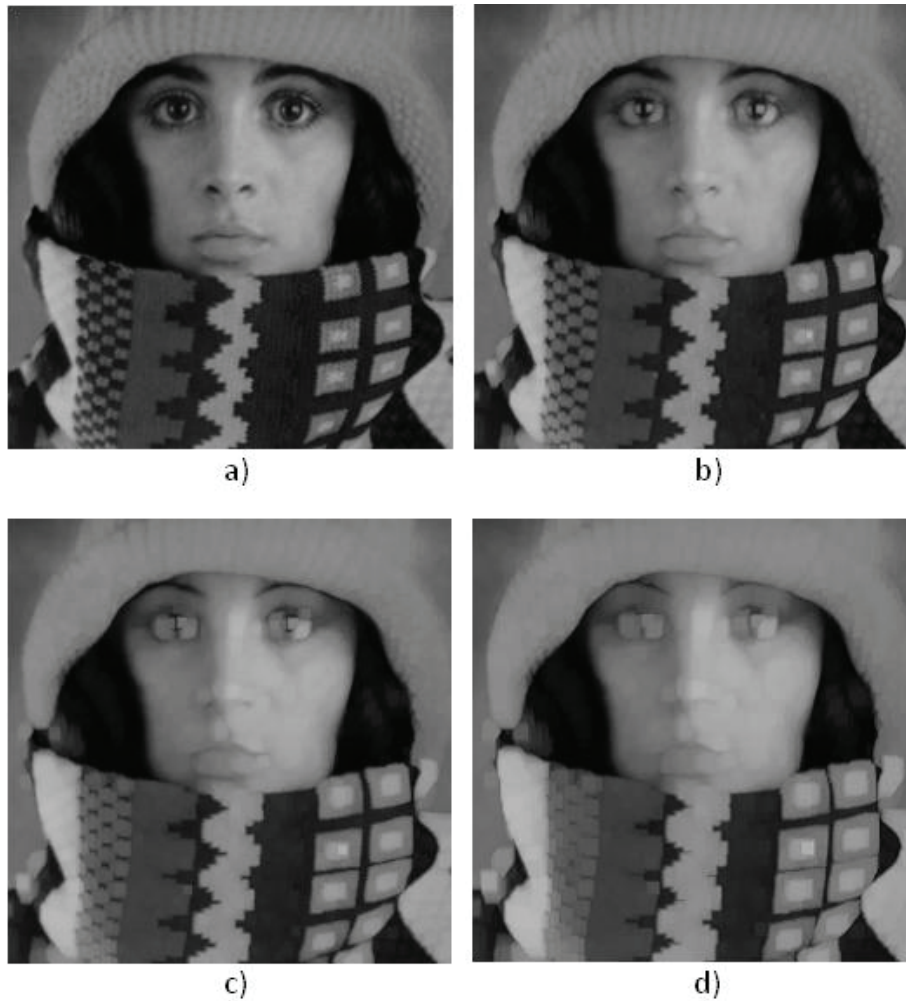


Figura 3.11: Diferentes niveles en la representación del espacio de escala de una imagen bidimensional (a) dilatada a niveles de escala $\sigma = 3, 5$ y 7 se muestra en recuadros (b), (c) y (d) respectivamente.

como:

Definición 3.4.3 *Extremos locales.* Un píxel ρ de una imagen I es un máximo (mínimo) local para una rejilla G si y solo si su valor $I(\rho)$ es mayor (menor) o igual a alguno de sus vecinos.

Es importante resaltar que un máximo *regional* (definición 3.3.1) no debe confundirse con un máximo local. Todos los pixeles que pertenecen a un máximo regional son máximos locales, pero el caso inverso no se cumple Vincent (1993).

Tomando en cuenta la definición anterior, se establecen los siguientes conjuntos de puntos de un extremo local de una función:

$$E_{max}(f) = \{(x, y) : f(x, y) \text{ es un máximo local}\} \quad (3.51)$$

$$E_{min}(f) = \{(x, y) : f(x, y) \text{ es un } \acute{m}inimo \text{ local}\} \quad (3.52)$$

Para observar como se comportan los extremos locales (maximos y mınimos) en la representaci3n multiescala, en la Fig. 3.13 se muestran los mınimos locales de las imagenes erosionadas a escalas $\sigma = 3, 5, 7$ mostradas en las Figs. 3.12 y 3.13.

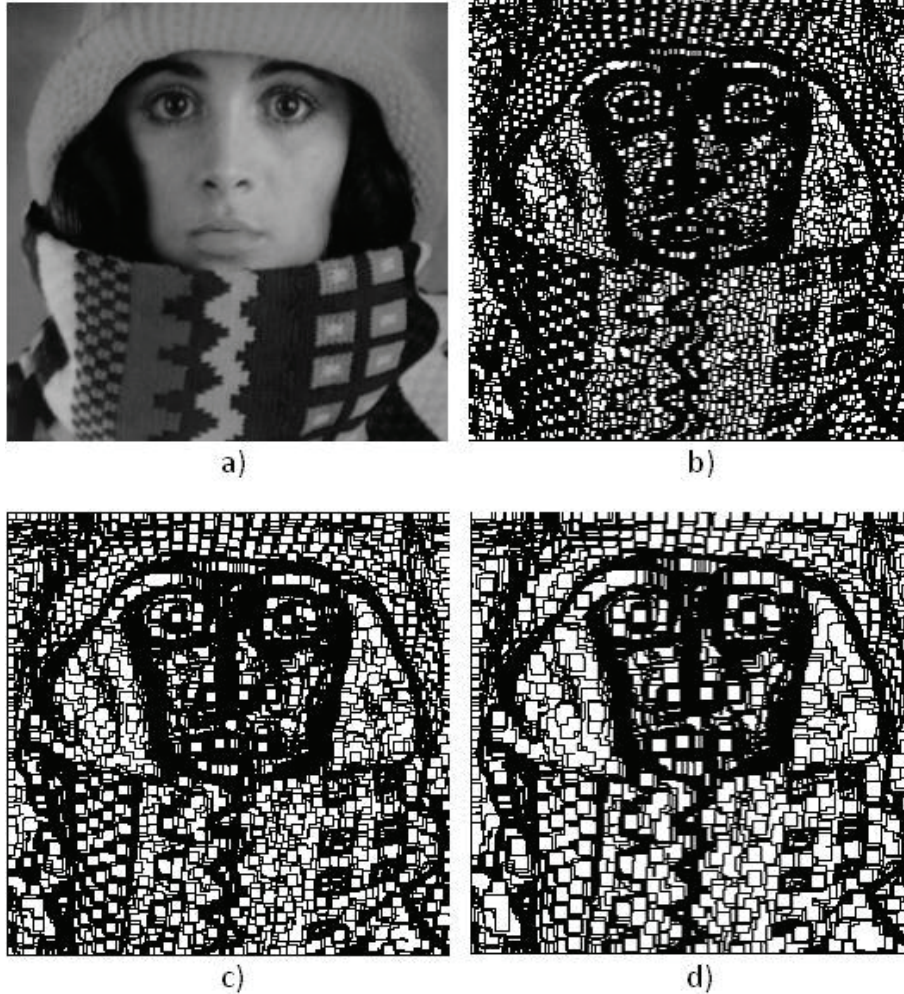


Figura 3.12: Mınimos locales de las imagenes de la Fig. 3.10.

En analogıa con la ecuaci3n (3.45), se define el siguiente conjunto de puntos dependientes de la escala:

$$E^{\circledast}(\sigma) = E_{max}(f \circledast g_{\sigma}) \cup E_{min}(f \circledast g_{\sigma}) \quad \sigma \in \mathbb{R} \quad (3.53)$$

De esta manera, el fingerprint para el espacio de escala dilataci3n-erosi3n se obtienen graficando este conjunto de puntos contra el valor de escala, en la figura 3.14 se muestra el fingerprint para una seal 1-D aleatoria.

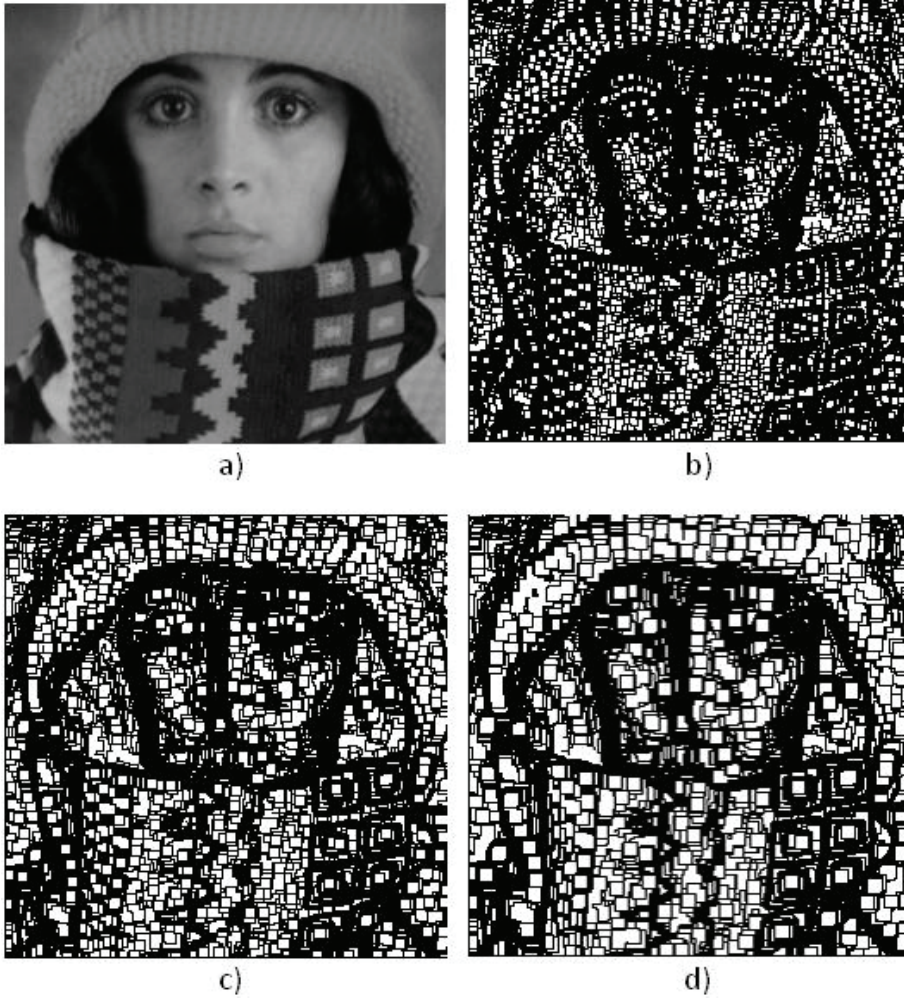


Figura 3.13: Máximos locales de las imágenes de la Fig. 3.10.

Para las escalas positivas las líneas del fingerprint correspondientes al máximo local viajan verticalmente en el diagrama del fingerprint, y de modo similar para los mínimos locales en escalas negativas. Esta observación lleva a la introducción del *fingerprint reducido*.

Definición 3.4.4 *Fingerprint reducido para el espacio de escala dilatación-erosión.* El fingerprint reducido de la representación multiescala dilatación-erosión morfológica está definido como:

$$E_r^{\otimes}(\sigma) = \begin{cases} E_{max}(f \oplus g_{\sigma}) & \text{si } \sigma > 0; \\ E_{max}(f) \cup E_{min}(f) & \text{si } \sigma = 0; \\ E_{min}(f \ominus g_{\sigma}) & \text{si } \sigma < 0. \end{cases} \quad (3.54)$$

Una de las principales contribuciones del artículo de Jackway (1996) fue demostrar la propiedad de causalidad de dilatación-erosión morfológica multiescala de los

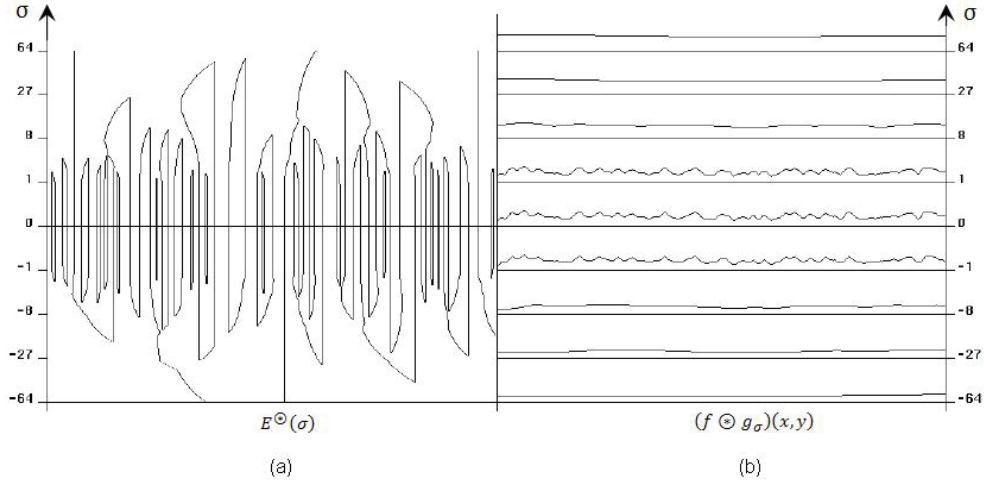


Figura 3.14: Fingerprints en el espacio de escala dilatación-erosión morfológico de una señal 1D aleatoria. (a) Fingerprints y (b) Representación multiescala de la señal

fingerprints definidos por los extremos locales. Este teorema establece que:

Teorema 3.4.1 Propiedad de monotonicidad del espacio de escala Si $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ denota una función acotada, $g_\sigma : G \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una función estructural escalada con un solo máximo en el origen, y el conjunto de puntos, $E_{max}(f) = \{x : f \text{ es máximo local}\}$, y, $E_{min}(f) = \{x : f \text{ es mínimo local}\}$, denotan los extremos locales de f . Entonces para cualquier $\sigma_1 < \sigma_2 < \sigma_3 < \sigma_4$,

$$E_{min}(f \circledast g_{\sigma_1}) \subseteq E_{min}(f \circledast g_{\sigma_2}) \subseteq E_{min}(f), \quad (3.55)$$

$$E_{max}(f \circledast g_{\sigma_4}) \subseteq E_{max}(f \circledast g_{\sigma_3}) \subseteq E_{max}(f), \quad (3.56)$$

El fingerprint reducido para la misma señal aleatoria 1D usada en la figura 3.14(b) se muestra en la figura 3.15.

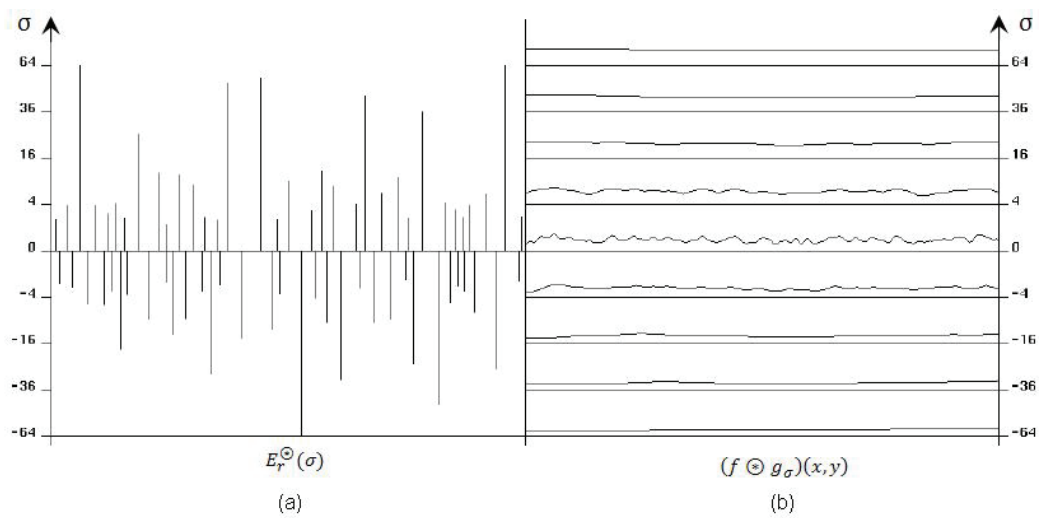


Figura 3.15: Fingerprint reducido de una señal aleatoria 1D en el espacio de escala morfológico. (a) Fingerprint reducido y (b) Representación multiescala de la señal

Parte II

Espacio de escala
apertura-cerradura por
reconstrucción y
cerradura-apertura por
reconstrucción

3.5. Espacio de escala de la apertura-cerradura por reconstrucción

3.5.1. Apertura-cerradura por reconstrucción multiescala

En Jackway (1996) se introdujeron dos transformaciones multiescala empleando una clase especial de funciones estructurales. Aunque ambas transformaciones evitan la introducción de nuevos puntos extremos, tienen la desventaja de ser sensibles a contornos ruidosos. Para evitar esta desventaja, se propone usar las transformaciones de apertura y cerradura por reconstrucción como primitivas. El uso de estas transformaciones morfológicas conexas nos permite construir un espacio de escala con las siguientes características: (i) las imágenes o señales de salida son mas fieles a la señal original; (ii) las transformaciones en el espacio de escala tienden a retener mejor los bordes y rasgos a través de la escala; (iii) las transformaciones en el espacio de escala son menos sensible a contornos ruidosos.

3.5.2. Fidelidad de las transformaciones del espacio de escala Gaussiano y morfológico

El operador de convolución empleado en el espacio de escala Gaussiano difumina los rasgos de la imagen al variar la escala resultando en una copia menos fiel de la imagen original. Por otro lado, el espacio morfológico dilatación-erosión multiescala actúa sobre la intensidad de niveles de gris de la imagen; para escalas positivas, se tendrá una imagen de salida más iluminada al aplicarse la operación de dilatación mientras que para escalas negativas, la imagen de salida sera más oscura.

En ambos casos, los rasgos de la imagen procesada presentan cambios dramáticos en la estructura de la imagen respecto a la original, una transformación adecuada para el análisis multiescala seria aquella que remueve ciertas características de la imagen de acuerdo al criterio de escala, mientras al mismo tiempo preserva fielmente los rasgos que no se ven afectados a dicha escala.

Los filtros por reconstrucción, en particular los filtros de apertura y cerradura por reconstrucción remueven rasgos claros y oscuros de la imagen de tamaño menor que el elemento estructural, sin alterar significativamente el resto de las componentes de la imagen.

En la figura 3.16 se presenta como se modifican los rasgos de una imagen Fig. 3.16(a), al ser procesada mediante un filtro Gaussiano de varianza $\sigma = 8$ Fig. 3.16(b); el resultado de aplicar un operador de dilatación morfológica con elemento estructural μ de talla 6 se muestra en Fig. 3.16(c) ; y al aplicar la transformación de apertura morfológica por reconstrucción con elemento estructural de talla $\mu = 6$ en la Fig. 3.16(d).

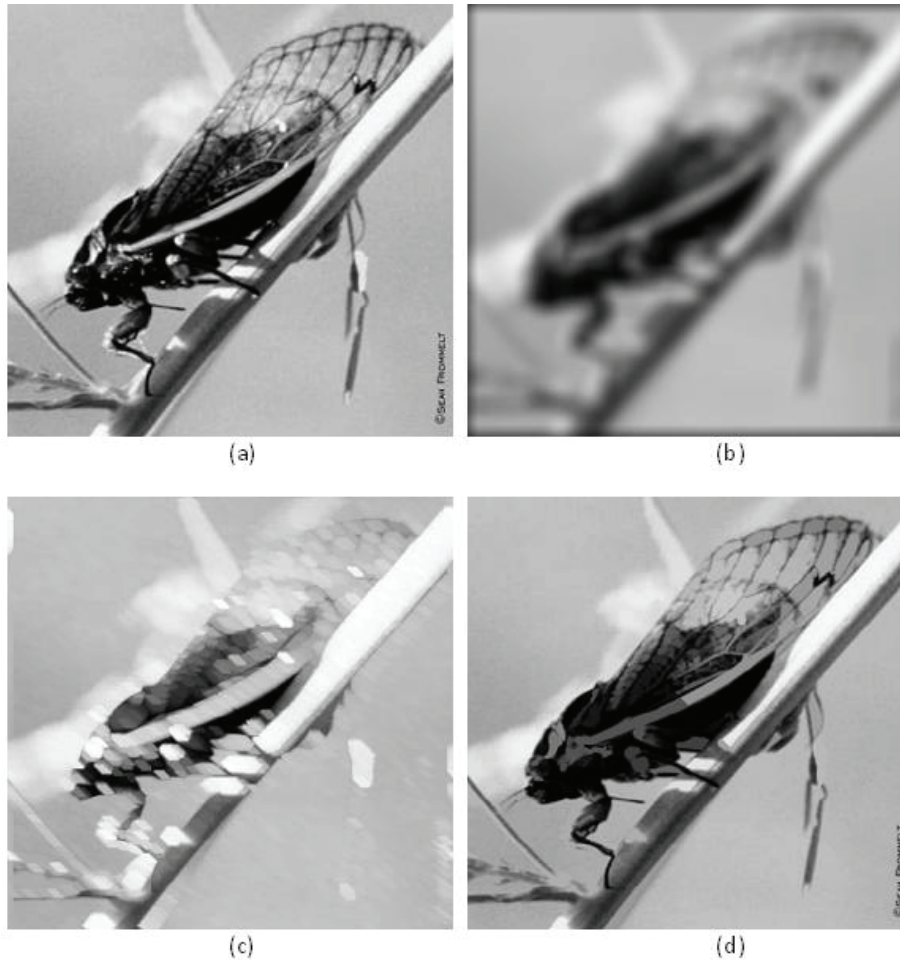


Figura 3.16: Fidelidad de los distintos operadores empleados en el espacio de escala: a) Imagen original, b) filtrado Gaussiano con $\sigma = 8$, c) dilatado morfológico con $\mu = 6$ y d) apertura morfológica por reconstrucción con $\mu = 6$.

Los bordes de una imagen constituyen una de las características mas importantes pues proporcionan información sobre la forma y delimitan los objetos presentes en la misma. A partir de los contornos de los objetos es posible realizar tareas de segmentación, reconocimiento de patrones, identificación de piezas, etc.

Por este motivo es deseable que la transformación empleada para el filtrado de la señal en el análisis multiescala preserve los bordes de los objetos en la imagen, además

de evitar la aparición de nuevos contornos en la representación de imagen de escala. El siguiente ejemplo ilustra el desempeño del filtro Gaussiano, la dilatación morfológica y la apertura morfológica por reconstrucción. La Fig. 3.17 muestra los bordes para las imágenes de la Fig. 3.16, en este ejemplo se utilizó la definición de gradiente morfológico (Ec. 3.24).

En el ejemplo de la Fig. 3.17 se obtuvieron las imágenes gradiente para el caso de la imagen original (Fig. 3.17(a)), la imagen procesada por un filtro Gaussiano de varianza $\sigma = 8$ (Fig. 3.17(b)) y para el caso de la dilatación y apertura morfológica por reconstrucción podemos ver las imágenes del gradiente morfológico en las Figs. 3.17(c) y 3.17(d) respectivamente.

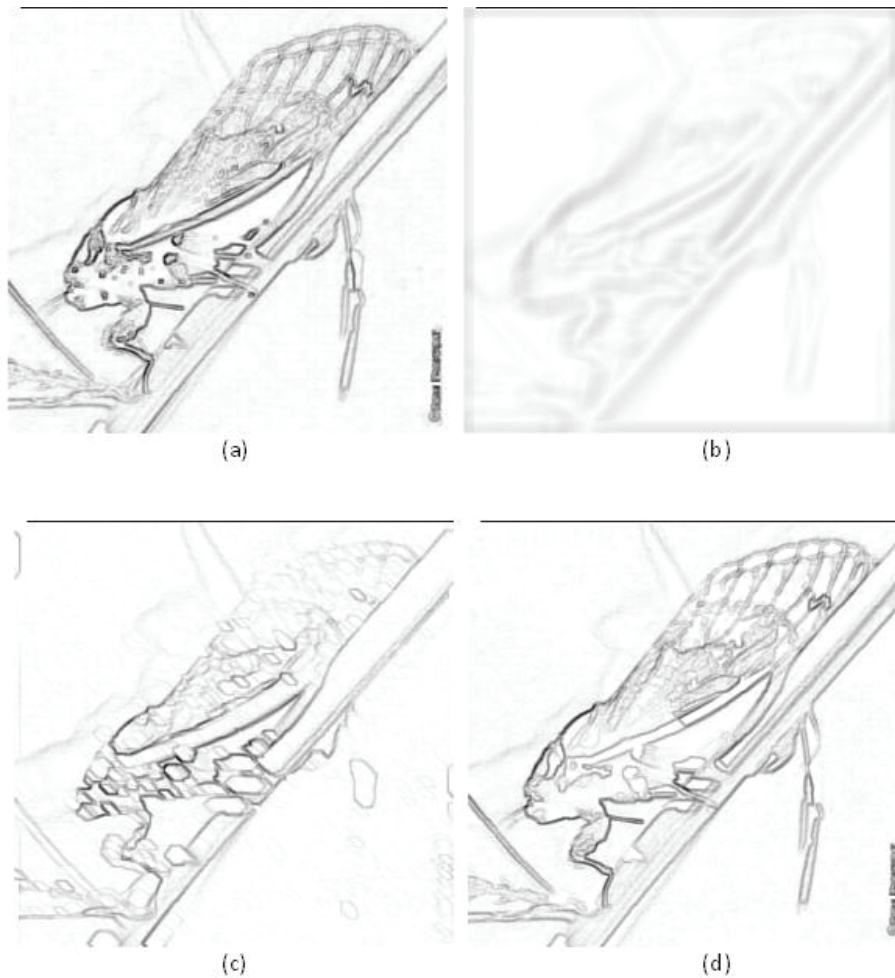


Figura 3.17: Gradiente morfológico de las imágenes de la Fig. 3.16 obtenido a partir de la definición dada por la Ec. 3.24. Gradiente de: a) Imagen original, b) filtrado Gaussiano con $\sigma = 8$, c) dilatado morfológico con $\mu = 6$ y d) apertura morfológica por reconstrucción con $\mu = 6$.

Es importante hacer notar, que en todos los casos debe evitarse introducir nuevos contornos de información a la imagen procesada. Para el caso de los contornos que se muestran en las imágenes de las Figs. 3.17(b), (c) y (d) son versiones simplificadas de los contornos de la imagen original (Fig. 3.17(a)). En este ejemplo puede verse como a partir de la apertura por reconstrucción, los contornos de la imagen procesada mantienen una estructura mas fiel al de la imagen original.

Como en todo proceso de captura de imagen esta sujeto a ruido de algún tipo, es importante considerar la tolerancia al ruido en el operador multiescala. El ruido no se puede predecir acertadamente debido a su naturaleza aleatoria, y tampoco puede ser medido precisamente en una imagen, dado que la contribución del ruido a los niveles de gris no es indistinguible de los píxeles de datos. Sin embargo, el ruido puede ser caracterizado a veces por su efecto en la imagen.

Como propuesta de este trabajo, se define la apertura-cerradura por reconstrucción multiescala $\chi_{\mu}^{\tilde{\gamma}, \tilde{\varphi}}(f)$ como:

$$\chi_{\mu}^{\tilde{\gamma}, \tilde{\varphi}}(f)(x) = \begin{cases} \tilde{\gamma}_{\mu}(f)(\mathbf{x}) & \mu > 0, \\ f(\mathbf{x}) & \mu = 0, \\ \tilde{\varphi}_{|\mu|}(f)(\mathbf{x}) & \mu < 0. \end{cases} \quad (3.57)$$

Como puede observarse, para escalas positivas, se realiza la apertura por reconstrucción, y para escalas “negativas”, la cerradura por reconstrucción con un elemento estructural de tamaño $|\mu|$. Debido a que μ toma valores negativos ($\mu \in \mathbb{Z}$) se considera el valor absoluto de μ (las escalas negativas son usadas para obtener una notación apropiada cuando se desea mostrar la representación multiescala). Por otro lado, de la ecuación 3.57, también se puede aplicar la cerradura por reconstrucción para valores positivos de μ , y la apertura por reconstrucción para escalas “negativas”. De esta manera se puede desarrollar otro operador multiescala conexo; tal operador es definido como la transformación cerradura-apertura por reconstrucción multiescala y se expresa como sigue:

$$\chi_{\mu}^{\tilde{\varphi}, \tilde{\gamma}}(f)(x) = \begin{cases} \tilde{\varphi}_{\mu}(f)(\mathbf{x}) & \mu > 0, \\ f(\mathbf{x}) & \mu = 0, \\ \tilde{\gamma}_{|\mu|}(f)(\mathbf{x}) & \mu < 0. \end{cases} \quad (3.58)$$

El análisis en esta sección solo considera la ecuación (3.57), la razón de esto es que la transformación (3.58) es una transformación dual del operador propuesto en

la ecuación (3.57). Por lo tanto se pueden obtener comentarios y resultados similares para la ultima transformación.

La apertura-cerradura por reconstrucción multiescala permite la detección de regiones claras y oscuras de talla creciente conforme se incrementa. Dado que la apertura por reconstrucción es anti-extensiva y la cerradura por reconstrucción es extensiva, entonces $\chi_{\mu>0}^{\tilde{\gamma},\tilde{\varphi}}(f) \leq f \leq \chi_{\mu<0}^{\tilde{\gamma},\tilde{\varphi}}(f)$. Las siguientes relaciones de orden se satisfacen para $\chi_{\mu}^{\tilde{\gamma},\tilde{\varphi}}(f)$.

Propiedad 3.5.1 *Sea D_f el dominio de la función f , la apertura-cerradura por reconstrucción multiescala $\chi_{\mu}^{\tilde{\gamma},\tilde{\varphi}}(f)$ cumple lo siguiente: (i) $\chi_{\mu=0}^{\tilde{\gamma},\tilde{\varphi}}(f)(\mathbf{x}) = f(\mathbf{x}) \forall \mathbf{x} \in D_f$, (ii) $\chi_{\mu \rightarrow +\infty}^{\tilde{\gamma},\tilde{\varphi}}(f)(\mathbf{x}) = \inf\{f(\mathbf{x})\} \forall \mathbf{x} \in D_f$, (iii) $\chi_{\mu \rightarrow -\infty}^{\tilde{\gamma},\tilde{\varphi}}(f)(\mathbf{x}) = \sup\{f(\mathbf{x})\} \forall \mathbf{x} \in D_f$, (iv) para escalas positivas con $0 < \mu_1 < \mu_2$, $\chi_{\mu_1}^{\tilde{\gamma},\tilde{\varphi}}(f)(\mathbf{x}) \geq \chi_{\mu_2}^{\tilde{\gamma},\tilde{\varphi}}(f)(\mathbf{x})$, (v) para escalas negativas con $0 < |\mu_3| < |\mu_4|$, $\chi_{|\mu_3|}^{\tilde{\gamma},\tilde{\varphi}}(f)(\mathbf{x}) \leq \chi_{|\mu_4|}^{\tilde{\gamma},\tilde{\varphi}}(f)(\mathbf{x})$*

Por otro lado, la apertura y cerradura por reconstrucción se definieron como levelings en un marco de trabajo digital Meyer (1998). Los levelings tienen la propiedad de preservar los contornos alargando los extremos. Cuando se aplican la transformación leveling, se observa un comportamiento interesante en los puntos extremos de la imagen el cual consiste en que cualquier máximo regional (mínimo, respectivamente) de la imagen procesada contiene un máximo regional (mínimo, respectivamente) de la imagen original. Considerando esta propiedad de preservación de extremos, sea $E_{max}(f) = \{\mathbf{x} : \mathbf{x} \in M_{ax} \text{ donde } M_{ax} \text{ es un máximo regional de } f\}$ y $E_{min}(f) = \{\mathbf{x} : \mathbf{x} \in M_{in} \text{ donde } M_{in} \text{ es un mínimo regional de } f\}$ el conjunto de puntos en el dominio de la definición representando los máximos y mínimos regionales anidados, respectivamente. Por lo tanto, cuando el parámetro de escala μ se incrementa, tenemos que:

para escalas positivas con $0 < \mu_1 < \mu_2$,

$$E_{max}(f) \subseteq E_{max}(\chi_{\mu_1}^{\tilde{\gamma},\tilde{\varphi}}(f)) \subseteq E_{max}(\chi_{\mu_2}^{\tilde{\gamma},\tilde{\varphi}}(f)); \quad (3.59)$$

para escalas negativas donde $0 < |\mu_3| < |\mu_4|$

$$E_{min}(f) \subseteq E_{min}(\chi_{|\mu_3|}^{\tilde{\gamma},\tilde{\varphi}}(f)) \subseteq E_{min}(\chi_{|\mu_4|}^{\tilde{\gamma},\tilde{\varphi}}(f)); \quad (3.60)$$

De las expresiones (3.59) ((3.60), respectivamente), notamos que el numero de máximos (mínimos, respectivamente) regionales disminuye conforme μ se incrementa.

3.5.3. Fingerprint de transformaciones conexas

Como se mencionó en la sección 3.4.2, el fingerprint de una señal es un gráfico del conjunto de puntos, contra escala, de la posición de las características de la señal. En analogía con la definición dada en Jackway (1996), definimos el siguiente conjunto de puntos dependientes de la escala;

Definición 3.5.1 (*Fingerprint del espacio de escala apertura-cerradura por reconstrucción multiescala*) *El fingerprint del espacio de escala apertura-cerradura por reconstrucción multiescala es un gráfico, contra valor de escala, del siguiente conjunto de puntos dependientes de la escala: $E_{\otimes}^{\tilde{\gamma}, \tilde{\varphi}}(\mu) = E_{max}(\chi_{\mu}^{\tilde{\gamma}, \tilde{\varphi}}(f)) \cup E_{min}(\chi_{\mu}^{\tilde{\gamma}, \tilde{\varphi}}(f))$.*

Es importante resaltar que para las transformaciones multiescala a a partir de operaciones por reconstrucción propuestas en este trabajo, los rasgos característicos de la señal corresponden a los máximos y mínimos **regionales** como se menciona en la definición 3.3.1 y no debe confundirse con los máximos y mínimos **locales** definidos por 3.4.3.

Mientras que el fingerprint reducido de una imagen se presenta como sigue:

Definición 3.5.2 (*Fingerprint reducido del espacio de escala apertura-cerradura por reconstrucción multiescala*) *El fingerprint reducido del espacio de escala apertura-cerradura por reconstrucción multiescala es un gráfico, contra valor de escala, del siguiente conjunto de puntos dependientes de la escala:*

$$E_r^{\tilde{\gamma}, \tilde{\varphi}}(\mu) = \begin{cases} E_{max}(\tilde{\gamma}_{\mu}(f)) & \mu > 0, \\ E_{max}(f) \cup E_{min}(f) & \mu = 0, \\ E_{min}(\tilde{\varphi}_{|\mu|}(f)) & \mu < 0. \end{cases} \quad (3.61)$$

Lo fingerprints para la transformación cerradura-apertura por reconstrucción se pueden encontrar remplazando en la ecuación (3.61), la cerradura por reconstrucción para escalas positivas y la apertura por reconstrucción para el caso de escalas negativas; esto es llevado a cabo como sigue:

$$E_r^{\tilde{\varphi}, \tilde{\gamma}}(\mu) = \begin{cases} E_{max}(\tilde{\varphi}_{\mu}(f)) & \mu > 0, \\ E_{max}(f) \cup E_{min}(f) & \mu = 0, \\ E_{min}(\tilde{\gamma}_{|\mu|}(f)) & \mu < 0. \end{cases} \quad (3.62)$$

Los fingerprint propuestos en la ecuación (3.61) tienen algunas diferencias importantes con respecto a los fingerprint reducidos del espacio de escala cerradura-apertura

por reconstrucción multiescala introducidos en Jackway (1996). Los fingerprint obtenidos usando las ecuaciones (3.57) y (3.61) expanden su tamaño conforme μ se incrementa. Debido a esta característica, los fingerprints pertenecen a regiones segmentadas que cambian conforme la talla del elemento estructural se modifica.

Por otro lado, el espacio de escala apertura-cerradura por reconstrucción permite el uso de elementos estructurales planos y convexos, los cuales no son recomendados para los enfoques basados en la dilatación-erosión multiescala y la cerradura-apertura multiescala propuesta en Jackway (1992, 1995). Por otro lado, de acuerdo con las ecuaciones (3.37) y (3.38), el elemento estructural permite la obtención del marcador g , y el proceso de reconstrucción depende únicamente de las regiones alcanzadas por tal marcador g ; por lo tanto si el marcador g de la función cambia, se pueden detectar distintas regiones.

Finalmente, la cerradura-apertura morfológica multiescala proporcionada en Jackway (1992) usa la cerradura morfológica para escalas positivas y la apertura para escalas negativas de manera similar a la ecuación (3.58), en la cual la cerradura por reconstrucción es el operador empleado para valores positivos de escala ($\mu > 0$) y la apertura por reconstrucción es empleada para escalas negativas. El efecto de intercambiar la apertura por la cerradura y viceversa permite la detección de los fingerprints como zonas planas localizadas en la imagen procesada.

Sin embargo, si la apertura por reconstrucción fuera aplicada para escalas negativas y la cerradura por reconstrucción para escalas positivas, los fingerprint completos y los reducidos serían similares a los reportados en Jackway (1992, 1995). No obstante, los fingerprints completos y reducidos obtenidos de la ecuación (3.58) serán menos sensibles a los contornos ruidosos, esto constituye una diferencia importante de aquellos propuestos por Jackway (1992, 1995).

Los fingerprint de la imagen obtenidos a partir del análisis multiescala, como ya se mencionó, sirven como descriptores de la imagen que pueden emplearse para tareas específicas de reconocimiento de patrones, codificación de imágenes e identificación de objetos, por mencionar algunos ejemplos.

Para dar una idea de la importancia de las transformaciones propuestas se mostrarán varios ejemplos de aplicación en señales unidimensionales y en imágenes en escala de grises, se demostrara como en ambos casos se obtienen resultados con mayor inmunidad al ruido que para el caso de la dilatación-erosión morfológica multiescala.

El primer ejemplo considera una señal 1D de la cual son extraídos los fingerprints. En este ejemplo se obtienen los fingerprints completos y reducidos del análisis de escala apertura-cerradura por reconstrucción y cerradura-apertura por reconstrucción, seguido de una comparación de estos resultados con aquellos obtenidos a partir de la metodología propuesta por Jackway (1995, 1996, 1998).

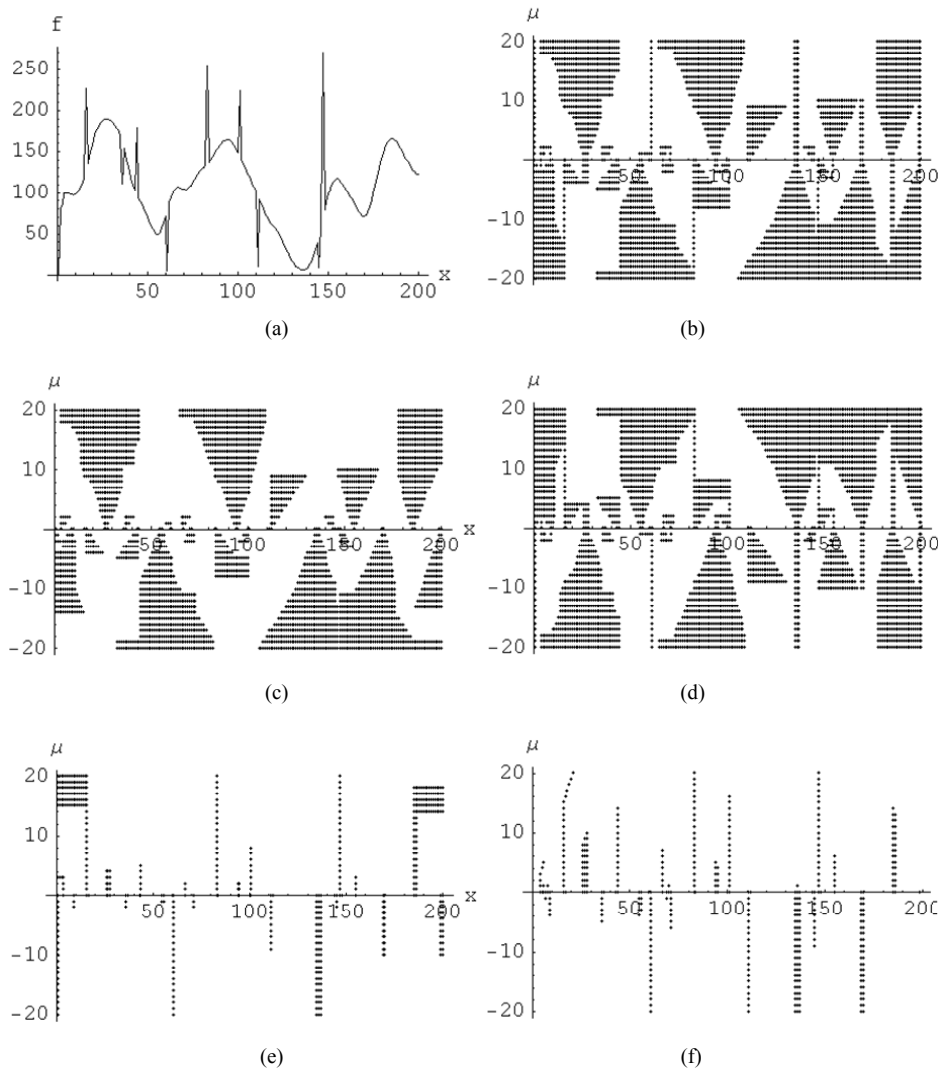


Figura 3.18: Fingerprints multiescala para una señal 1D; (a) señal de datos; (b) fingerprints completos apertura-cerradura por reconstrucción; (c) fingerprint reducidos apertura-cerradura por reconstrucción; (d) fingerprints completos cerradura-apertura por reconstrucción; (e) fingerprints reducidos cerradura-apertura por reconstrucción; (f) fingerprints reducidos dilatación-erosión.

La Fig. 3.18 (a) muestra una señal unidimensional ruidosa, el efecto del ruido gráficamente se puede apreciar como los picos ubicados en los valles y crestas de la señal.

Esta señal será procesada a diferentes escalas para obtener su representación multiescala usando las herramientas de análisis multiescala propuestas en este trabajo.

Los fingerprints completos y reducidos para el caso de la representación multiescala obtenida a partir del operador apertura-cerradura por reconstrucción se presentan en las Figs. 3.18 (b) y (c). Estas representaciones fueron obtenidas a partir de las Ecs. (3.57), (3.61) y la definición 3.5.2 respectivamente. La representación multiescala fue obtenida para el rango de valores de escala $\mu \in [-20, 20]$ (para valores negativos de μ , se debe emplear $|\mu|$).

Note en las Figs. 3.18 (b) y (c) que los extremos regionales de la señal original no son desplazados y se evita la creación de nuevos extremos regionales a través de la escala. Esta observación es importante y refleja las propiedades de causalidad y de preservación de extremos de la transformación apertura-cerradura por reconstrucción.

En contraste con los fingerprints completos adquiridos en otras metodologías Jackway (1992, 1995); Park (1996), los fingerprints completos obtenidos aquí corresponden a zonas planas de tamaño creciente. Este hecho se ilustra en las Figs. 3.18 (b), (c) y (d); mientras que en la Fig. 3.18 (e), se obtienen los fingerprint reducidos de la cerradura-apertura por reconstrucción. Solo estos fingerprint reducidos de cerradura-apertura por reconstrucción serán comparados con aquellos propuestos en Jackway (1992, 1996).

La razón de esto es que los fingerprint reducidos de cerradura-apertura por reconstrucción entregan una representación gráfica idéntica a la de los fingerprint reducidos obtenidos a partir de la dilatación-erosión morfológica multiescala como se demuestra en Jackway (1995).

De esta manera, los fingerprint reducidos son calculados y presentados en la Fig. 3.18 (f). Observe que los fingerprint reducidos en las Figs. 3.18 (e) y (f) tienen un comportamiento similar; sin embargo, hay algunas diferencias importantes entre ellas.

El operador multiescala proporcionado en Jackway (1992, 1996) usa la dilatación y erosión morfológica como primitivas; por esta razón los fingerprint son sensibles a los valores extremos en ciertas vecindades. En Fig. 3.18 (f) la propiedad de causalidad no se cumple; especialmente, el ruido impulsivo o de alta frecuencia afecta la representación de los fingerprint de la dilatación-erosión morfológica multiescala.

Por otro lado, si se lleva a cabo una etapa de preprocesamiento para suprimir el ruido, existe el riesgo de modificar la información en la señal original. Para los fingerprint reducidos cerradura-apertura por reconstrucción, los puntos con cambios abruptos en el valor de la intensidad (puntos ruidosos) serán fusionados con los máximos o mínimos regionales cuando se aplica la apertura o cerradura.

Por otro lado, es importante mencionar que Park y Lee (1996), probaron que las propiedades de causalidad para los filtros de apertura y cerradura morfológica se cumple para el caso de señales unidimensionales; lo cual no es el caso para las señales bidimensionales (2D), donde las propiedades de causalidad no se cumplen debido a la división de los contornos de las características que pueden ocurrir a altas escalas Park (1996); Bosworth (2003).

Por lo tanto, la sensibilidad al ruido y la pérdida de causalidad en señales de dimensión mayor constituyen desventajas indeseables las cuales se deben evitar para desarrollar una detección adecuada de los fingerprint. Las propuestas dadas en este trabajo permiten lo siguiente: un tratamiento en el espacio de escala basado en transformaciones conexas, y (ii) reducción del ruido conforme se incrementa el valor de escala.

Como resultado, los problemas mencionados anteriormente se logran evitar; de esta manera, las transformaciones de cerradura-apertura o la apertura-cerradura por reconstrucción ofrecen mejores resultados en comparación con resultados previos reportados en la literatura, para el caso de las señales 1D y 2D.

En el siguiente ejemplo el objetivo es presentar los fingerprint reducidos detectados en dos imágenes distintas que contienen rostros humanos, las cuales se presentan en las Figs. 3.19 (a1) y (b1). Ya que nuestro principal interés está centrado en las características presentes en el área del rostro, regiones indeseadas como el fondo de la imagen y el cabello facial son removidos. Las imágenes con esta característica se muestran en las Figs. 3.19 (a2) y (b2).

Posteriormente, se obtiene la representación multiescala de acuerdo al filtrado de la señal proporcionado por la Ec. (3.57) para extraer los fingerprint reducidos definidos por la Ec. (3.61).

Para el caso de la imagen de la Fig. 3.19 (a2), los fingerprint reducidos corresponden a las imágenes mostradas en las Figs. 3.19 (a3)-(13) (regiones blancas) donde se

detectan los máximos regionales en las Figs. 3.19 (a3)-(a8), y los mínimos regionales se muestran en las Figs. 3.19 (a9)-(a13).

Para la imagen de la Fig. 3.19 (b2) los máximos regionales corresponden a las imágenes en las Figs. 3.19 (b3)-(b8), mientras que en las Figs. 3.19, se obtienen los mínimos regionales. En las Figs. 3.19 (c1) y (c2), se proporciona una representación tridimensional de los fingerprint.

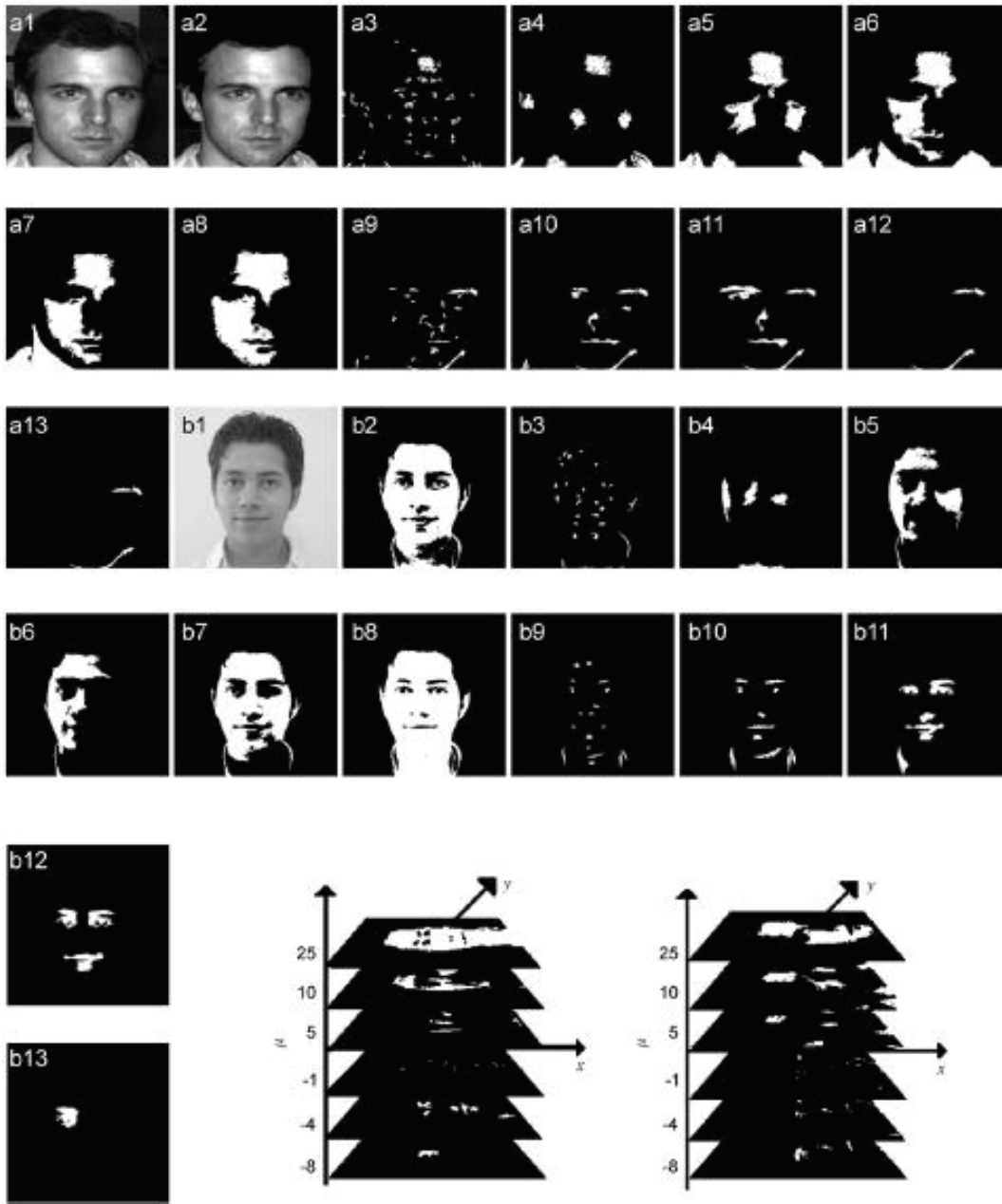


Figura 3.19: Fingerprint reducidos multiescala de dos imágenes faciales; (a1)-(b1) imágenes originales; (a2) y (b2) imágenes de entrada; (a3)-(a13) y (b3)-(b13) fingerprints reducidos apertura-cerradura por reconstrucción con $\mu = 1, 5, 10, 15, 20, 25, |-1|, |-2|, |-4|, |-6|, |-8|$; (c1) – (c2) una representación 3D de los fingerprint reducidos anteriores.

Parte III

Realce de contraste y normalización de iluminación en imágenes mal iluminadas

Capítulo 4

Compensación en cambios de iluminación

4.1. Fondo de la imagen y normalización

La iluminación es uno de los factores más significativos que afectan la apariencia de una imagen. Este hecho frecuentemente conduce a una variación del contraste debido a la presencia de sombras y regiones extremadamente claras; de tal manera que se encuentran transiciones importantes de luminancia al pasar de una zona oscura a una clara o viceversa.

En el tratamiento de las imágenes, frecuentemente es necesario segmentar y detectar características de interés que pueden ser utilizadas para un análisis posterior. Por ello, es necesario ajustar la apariencia de la imagen, mejorando las condiciones de iluminación; y de esta manera poder realizar la correcta detección o segmentación de las regiones bajo estudio.

En esta sección presentamos una estrategia para realizar una compensación en los cambios de iluminación mediante la normalización de la imagen estudiada; i.e, se evitarán los cambios bruscos de luminancia mediante la distribución de los niveles de gris; y para ello se utilizará la noción de la ley de Weber.

4.2. Ley de Weber

Como portadoras de información, todas las imágenes eventualmente son percibidas e interpretadas por el sistema de visión humano. De esta manera, el entendimiento del proceso de percepción visual es importante como auxiliar para el desarrollo de

algoritmos en el area de procesamiento de imágenes. La *ley de Weber* es un modelo de percepción visual, la cual indica que la razón entre la magnitud del estímulo necesario para percibir un cambio en la sensación apenas notable ΔL , y la luminancia de fondo L , es una constante; i.e :

$$\Delta L/L = const. \quad (4.1)$$

La ley de Weber se puede expresar por medio de una expresión logarítmica Jain (1989), tal como:

$$C = a_1 + a_2 \log L \quad L > 0 \quad (4.2)$$

Donde C es el contraste, L la luminancia, mientras que a_1 y a_2 son constantes relacionadas con contrastes iniciales; siendo a_1 el fondo de la imagen.

Para el cálculo de la constante a_1 , en este trabajo será utilizada la cerradura por reconstrucción, se puede consultar la sección 3.3 donde se presenta un breve respaldo teórico sobre transformaciones morfológicas por reconstrucción.

4.3. Noción de fondo de la imagen

De acuerdo a la ley de Weber, si se modifica el *fondo de la imagen* (background en inglés), entonces existirá un cambio en el contraste percibido. De esta manera decimos que existen dos planos, uno de ellos es el fondo de la imagen, mientras que el otro plano consiste del resto de las componentes.

A partir de estos conceptos, concluimos que la detección del fondo de la imagen es fundamental para mejorar el contraste; sin embargo cuando existen regiones con poca o mala iluminación, el fondo de la imagen se ve alterado; gráficamente se representa esta situación en la Fig. 4.1. Mientras que una imagen bien iluminada presenta un fondo uniforme en la mayoría de los casos dependiendo del tipo imagen.

De este modo, una imagen ideal f es alterada por un fondo de imagen desigual r tal que la imagen observada será $I = f + r$. Recuperar la imagen original f , desde la imagen observada I , no es sencillo debido a que r no es uniforme. De este modo, una estrategia para efectuar la normalización de la imagen es por medio de la mejora del contraste basado en la noción del fondo de la imagen. Partiendo de esta idea, y considerando que una aproximación a la luminancia esta dado por el nivel de gris de los pixeles de la imagen, la ecuación (4.2) es re-escrita de la siguiente manera,

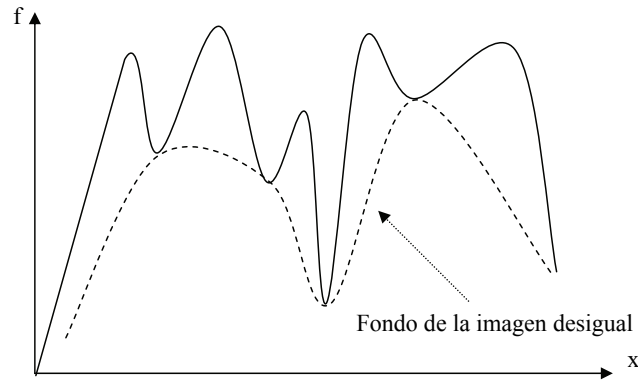


Figura 4.1: Imagen original f junto con el fondo de la imagen desigual en líneas punteadas

$$C^* = a_2 \text{Log}(f + 1) + b \quad (4.3)$$

Donde $a_1 = b$; mientras que la adición de la unidad es considerada para evitar la indeterminación del logaritmo. De acuerdo a la ley de Weber, las constantes a_2 y b son condiciones iniciales del contraste, siendo b el fondo de la imagen. Asumiendo que se cuenta con un valor adecuado para el fondo de la imagen, el valor de la constante a_2 se determina como sigue; en nuestro caso el valor máximo de nivel de gris denotado como $maxint$, es 255, por lo que,

$$a_2 = \frac{maxint - b}{\text{Log}(maxint + 1)} = \frac{255 - b}{\text{Log}(255 + 1)} \quad (4.4)$$

Algunas imágenes de salida donde se ilustra el desempeño de la transformación expresada en la ecuación (4.4) se presentan en la Fig. 4.2. En la Fig. 4.2, el contraste fue modificado al variar el valor del fondo de la imagen. A medida que el valor del fondo de la imagen se incrementa, la imagen tiende a mostrarse mas iluminada, esto es debido a que el fondo de la imagen es un efecto aditivo. Formalmente tenemos que:

$$\lim_{b \rightarrow maxint} C^* = maxint \quad (4.5)$$

Es importante seleccionar un valor adecuado del fondo de la imagen para el mejoramiento del contraste. El inconveniente de usar un fondo de la imagen constante para toda la imagen es que todos los pixeles se procesan en la misma medida. La imagen de la Fig. 4.2 fue mejorada usando el concepto de fondo de la imagen; sin embargo, es una imagen compleja debido a que está formada por componentes que varían su contraste de acuerdo a las regiones vecinas.

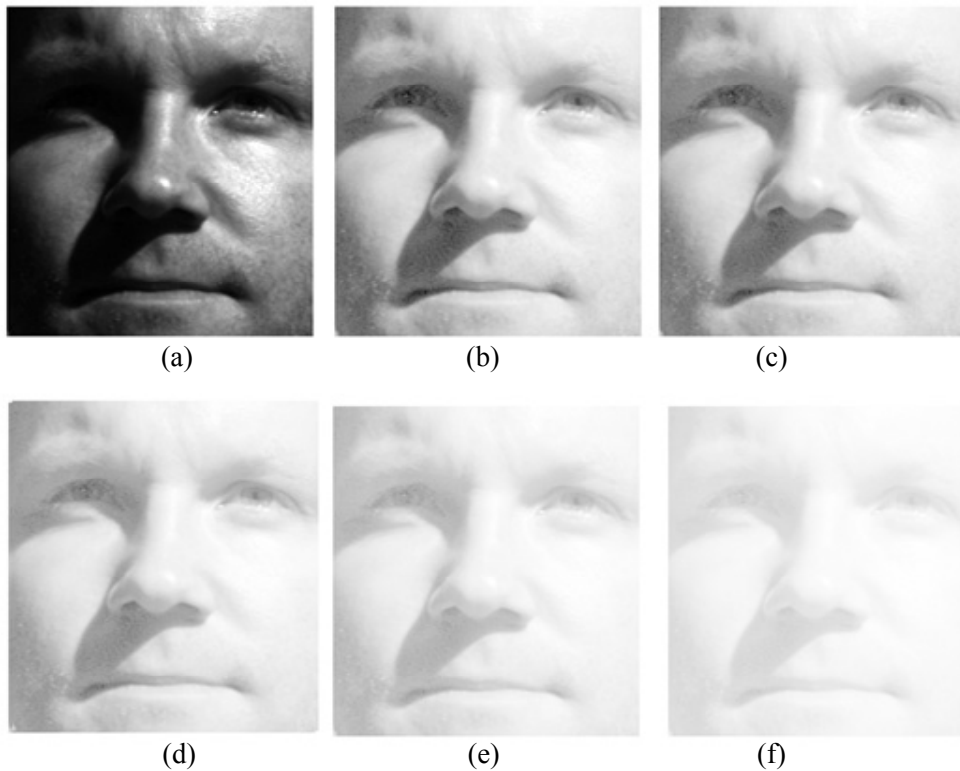


Figura 4.2: (a) Imagen original, (b-f) Imágenes procesadas con $b = 20, 50, 100, 150$ y 200 respectivamente

Para imágenes con cambios bruscos de iluminación, la detección del fondo de la imagen se puede simplificar al detectar uno para las regiones oscuras y otro para las regiones claras, esta propuesta es presentada en la siguiente sección.

4.4. Mapeo de contraste basado en el fondo de la imagen

Considerando los resultados mostrados en la Fig. 4.2, observamos que una estimación adecuada del fondo de la imagen originará un mejoramiento en el contraste. En imágenes reales el valor del fondo de la imagen varía tal y como se ilustra en la Fig. 4.1. Sin embargo, simplificando el problema de detección del fondo de la imagen, se propone detectarlo solo en ciertos intervalos, de tal manera que sea posible modificar a las regiones oscuras (sombras) con un cierto fondo de la imagen; mientras que las regiones claras serán modificadas con otro acorde para este propósito; esta situación se ilustra en la Fig. 4.3.

De acuerdo a lo planteado, esto da origen a un mapeo de contraste de dos primitivas basado en el fondo de la imagen; y cada primitiva será aplicada bajo un criterio de fondo de la imagen obtenido a partir de b_1 y b_2 .

Se define el siguiente operador:

$$\kappa_{b_1, b_2}(x) = \begin{cases} k_1 \log(f(x) + 1) + b_1 & f(x) > \tau \\ k_2 \log(f(x) + 1) + b_2 & f(x) \leq \tau \end{cases} \quad (4.6)$$

donde

$$\tau = \frac{(b_1 + b_2)}{2} \quad (4.7)$$

Las constantes k_1 y k_2 se seleccionan tal que el rango de la imagen de salida esté dentro del límite del valor máximo de nivel de gris de la imagen, tal como se vio en la sección 4.1. Los valores b_1 y b_2 asociados al fondo de la imagen, se obtendrán a partir de un análisis de los mínimos regionales; en donde la intensidad de los pixeles correspondientes a el mínimo regional más grande se asociará a b_1 , mientras que el valor del mínimo más pequeño se asociará a b_2 .

Formalmente tenemos que, sea $\wp_{min}(f)$ ($\wp_{max}(f)$), el conjunto de todos los mínimos regionales (máximos regionales respectivamente) de f , el mínimo regional más grande es determinado como, $\bigvee \wp_{min}(f)$, mientras que el mínimo regional más pequeño corresponde a $\bigwedge \wp_{min}(f)$.

Para la determinación de los valores extremos de la imagen, los filtros por reconstrucción juegan un papel importante, esto se debe a que la cerradura por reconstrucción ayuda a eliminar mínimos regionales que son ocasionados por el ruido, además de que son fusionados generando zonas planas, las cuales son las componentes conexas mas largas donde la señal es constante segun Salembier (1995). Los valores de b_1 y b_2 se obtienen como:

$$b_1 = \bigvee \wp_{min}(\tilde{\varphi}_\mu(f)) \quad y \quad b_2 = \bigwedge \wp_{min}(\tilde{\varphi}_\mu(f)) \quad (4.8)$$

Por otra parte, el operador κ_{b_1, b_2} modifica a la imagen original bajo un criterio de fondo de la imagen τ , el cual es el promedio entre los valores del fondo de la imagen para regiones claras y oscuras respectivamente. Si $b_1 \approx b_2$, entonces $k_1 \approx k_2$ y $\kappa_{b_1, b_2} = k_1 \log(f(x) + 1) + b_1$.

Debido a que τ es un promedio, es una cantidad sensible a valores extremos, por

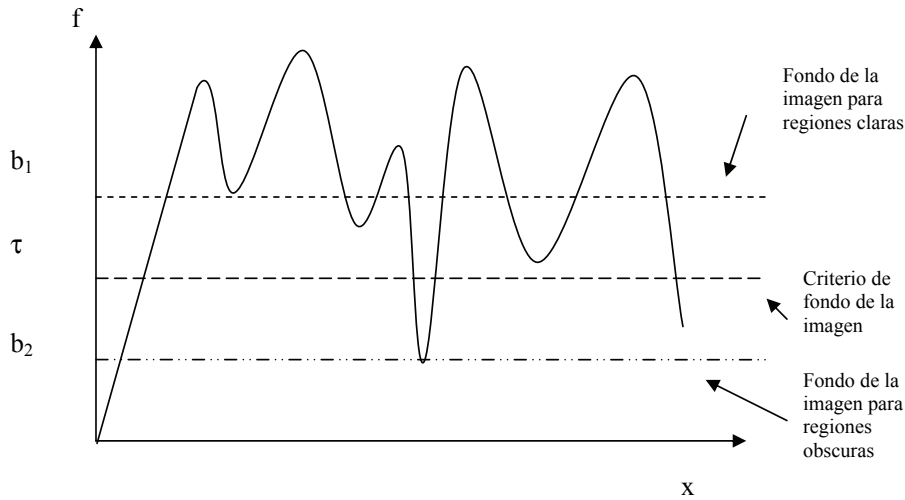


Figura 4.3: Detección del fondo de la imagen para regiones claras(b_1) y oscuras(b_2). El criterio de fondo de la imagen corresponde a τ

lo que se obtendrán mejores resultados cuando los mínimos ruidosos sean eliminados o fusionados mediante una etapa de pre-procesamiento, en nuestro caso, se emplea la cerradura por reconstrucción para evitar este problema. De este modo, es necesario aplicar la cerradura por reconstrucción y a continuación, los mínimos regionales de la imagen son obtenidos. El comportamiento de los mínimos de una imagen, cuando se aplica la cerradura por reconstrucción $\tilde{\varphi}_\mu$ con cierto tamaño de elemento estructural μ , se presenta en la Fig. 4.4.

Por otro lado, observe que la relación de orden $f \leq \kappa_{b_1, b_2}$ se verifica; i.e, es una transformación extensiva; esto se debe a que el logaritmo afecta mayormente a las regiones oscuras, mientras las regiones claras son menos modificadas. Además, las estructuras de la imagen se preservan,; mientras que la adición del fondo de la imagen se interpreta como un corrimiento de la función en el eje vertical.

El operador (4.6), actúa bajo un criterio de fondo de la imagen dado en la ecuación (4.7), y es posible clasificar los puntos del dominio de definición (denotado con S) de f en dos conjuntos:

- El conjunto S_{b_1, b_2} compuesto por regiones de alto contraste, donde para todo punto $x \in S_{b_1, b_2}$, $f(x) > \tau$ para $\kappa_{b_1, b_2}(f)$
- El conjunto S_{b_1, b_2}^c compuesto por regiones de bajo contraste (el conjunto complemento de S_{b_1, b_2} , donde para todo punto S_{b_1, b_2}^c , $f(x) \leq \tau$ para $\kappa_{b_1, b_2}(f)$)

La transformación definida en la ecuación (4.6) modifica a los pixeles claros u oscuros de manera similar; sólo son escalados con diferentes parámetros. También nótese que la

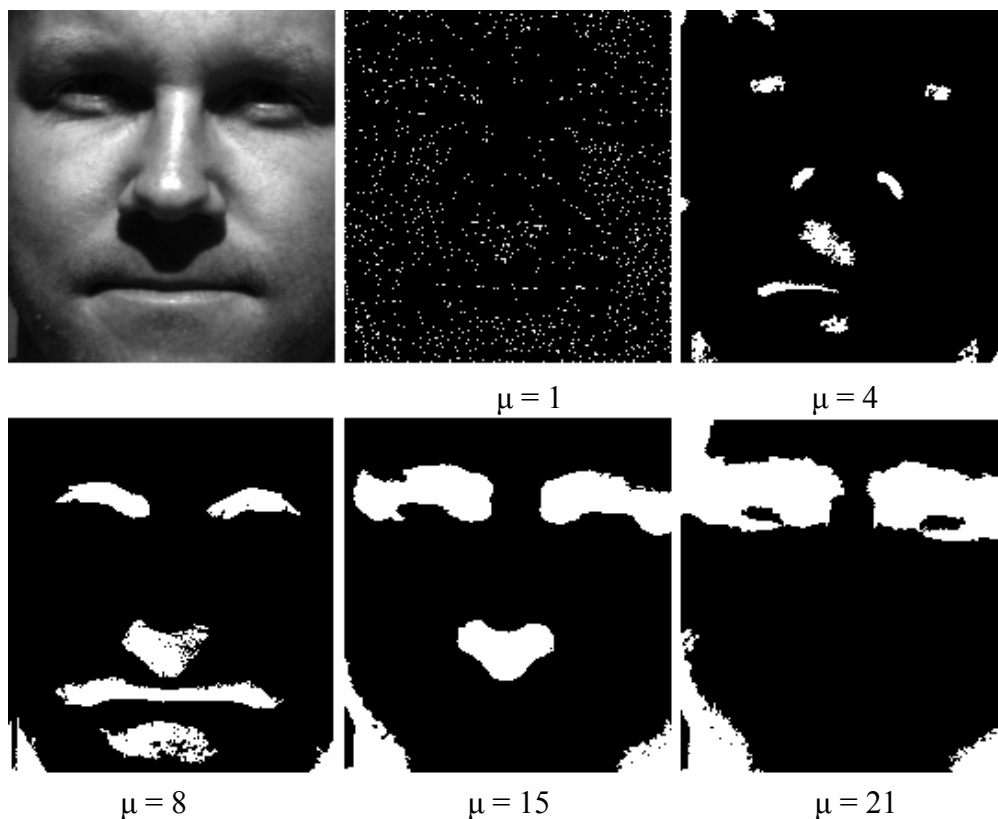


Figura 4.4: Evolucion de los minimos regionales de una imagen filtrada con la cerradura por reconstrucción $\tilde{\varphi}_\mu$, a medida que μ se incrementa tomando los valores de 1,4,8,15, y 21

función logaritmo afectará mayormente a las zonas oscuras, mientras las zonas claras son menos modificadas; este comportamiento permite obtener cambios suaves entre las regiones separadas por los contornos de la imagen; en otras palabras, los cambios bruscos de luminancia son corregidos al aplicar el operador κ_{b_1, b_2} .

La composición de mapeos de contraste utilizando la ecuación (4.6), dará como resultado imágenes que se aclaran en cada iteración, llegando a un límite impuesto por el valor del máximo nivel de intensidad de la imagen (*maxint*, que en nuestro caso es 255). Este comportamiento es explicado debido a las constantes aditivas del fondo de la imagen b_1 y b_2 presentes en la ecuación (4.6).

A continuación, se presentan algunos ejemplos orientados a imágenes con mala iluminación; las imágenes utilizadas corresponden a imágenes faciales las cuales son útiles para mostrar el desempeño de la propuesta de este artículo; sin embargo, el operador propuesto puede ser aplicado a otro tipo de imágenes.

4.5. Experimentos y comparación de resultados

Para ilustrar el desempeño de la propuesta de este trabajo, se presenta a continuación el procesamiento de imágenes del rostro con mala iluminación. Las imágenes cuyo contraste fue mejorado, se obtuvieron de la base de datos desarrollada por Peter Belhumeur y Athos Georghiades (2001), en el centro de visión y control computacional de la Universidad de Yale (<http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html>). Esta base de datos se creó con la finalidad de ser un recurso para evaluar sistemas de reconocimiento de rostros en condiciones de pose e iluminación variable. La base de datos contiene 5670 imágenes tomadas con una fuente de luz individual a 10 sujetos desde 576 puntos de vista distintos (9 poses de 64 condiciones de iluminación distintas).

Para la mejora del contraste, se aplicó el siguiente algoritmo:

Algoritmo 1

- (i) *La imagen original es procesada por medio de la cerradura por reconstrucción con la finalidad de fusionar los mínimos de la imagen y de esta manera evitar la presencia de ruido.*
- (ii) *Detectar el nivel de intensidad del mínimo de altitud mas grande; este valor es asignado a b_1 . Mientras que el valor del nivel de intensidad del mínimo de altitud más pequeña es asignado a b_2 de acuerdo a la Ec. (4.8).*
- (iii) *Se obtiene el criterio de fondo de la imagen definido en la ecuación (4.7); posteriormente se aplica la transformación definida en la ecuación (4.6) para obtener a la imagen de salida.*

En la Fig. 4.5 se presentan los resultados obtenidos al aplicar el algoritmo a un conjunto de imágenes de un mismo individuo bajo condiciones de poca iluminación. En la Fig. 4.5(d), los valores de b_1 y b_2 , se obtuvieron para $\mu = 21$; mientras que para las imágenes en las Figs. 4.5(e) y 4.5(f), el valor de $\mu = 15$ fue utilizado. Los valores de μ aplicados a las imágenes en la Fig.(4.5), fueron obtenidos de manera empírica; i.e, de acuerdo a el comportamiento de los mínimos regionales que se presentan en la Fig.(4.4), se observó que valores de μ en el intervalo de $[15,21]$, permiten detectar a mínimos ubicados en ojos, nariz y boca, lo cual no sucede con valores de μ menores que 15 y mayores que 25. A continuación se presenta otro ejemplo; en el cual se considera un grupo de individuos con condiciones similares de iluminación y pose.



Figura 4.5: Diferentes casos de pose e iluminación para un individuo. (a-c) Imágenes de prueba; (d) Resultados obtenidos cuando $\mu = 21$; (e-f) Imágenes de salida para $\mu = 15$.

Los rostros que se muestran en las Figs. 4.6(a-e), corresponden a una pose frontal tomadas con una dirección en la iluminación de 45° (imágenes de clase yaleB**_P00A+000E+45). La dirección de la iluminación reduce la presencia de sombras; sin embargo la intensidad de la luz no es suficiente para resaltar los rasgos faciales. El resultado de aplicar el operador propuesto en este trabajo, da como resultado a las imágenes de salida presentadas en las Figs.4.6(f-j). Note que el contraste es mejorado evitándose cambios bruscos de iluminación, además de que aparecen rasgos faciales que no se podían observarse en las imágenes originales. Para las imágenes procesadas, se utilizó un valor de $\mu = 10$, el cual se obtuvo de manera empírica de manera similar que en el ejemplo anterior.

A diferencia de otros algoritmos, como por ejemplo, técnicas de ecualización del histograma, mapeos de contraste morfológicos, entre otros; la metodología presentada en este trabajo es particularmente adecuada en imágenes cuya principal característica

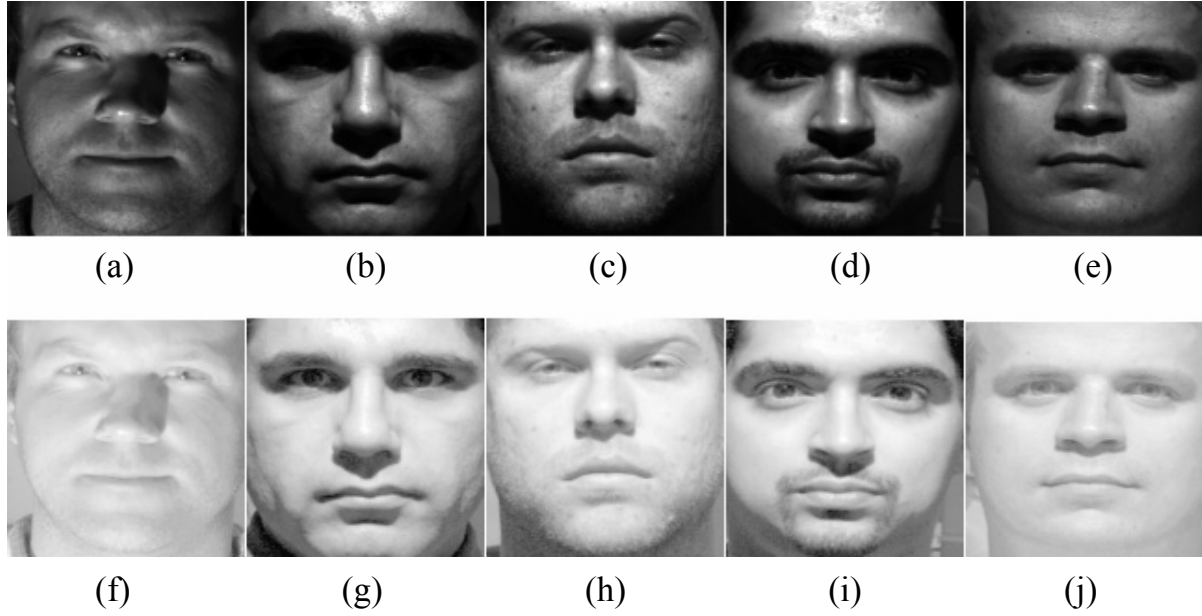


Figura 4.6: Normalización y mejora del contraste usando $\tilde{\varphi}_{\mu=10}$ para la detección del fondo de la imagen

es la presencia de sombras. La Fig. 4.7 muestra algunas imágenes de salida procesadas mediante el histograma de ecualización, mapeos morfológicos de contraste de tres estados a nivel zona plana, y la transformación propuesta en este trabajo. En la Fig. 4.7(b), se obtiene una imagen de salida procesada mediante un mapeo morfológico de contraste de tres estados a nivel zona plana Mendiola-Santibanez (2002a), el operador aplicado es el siguiente:

$$W_{\mu_1, \mu_2, \beta, \alpha}^3(f, P_f)(x) = \begin{cases} \varphi_{\mu_1}(f, P_f)(x) & 0 \leq \rho(x) < \beta \\ (f, P_f)(x) & \beta \leq \rho(x) < \alpha \\ \gamma_{\mu_2}(f, P_f)(x) & \alpha \leq \rho(x) \leq 1 \end{cases} \quad (4.9)$$

con,

$$\rho(x) = \frac{\varphi_{\mu_1}(f, P_f)(x) - (f, P_f)(x)}{\varphi_{\mu_1}(f, P_f)(x) - \gamma_{\mu_2}(f, P_f)(x)} \quad (4.10)$$

El operador $W_{\mu_1, \mu_2, \beta, \alpha}^3(f, P_f)(x)$ fue aplicado con $\mu_1 = \mu_2 = 5$, $\alpha = 0,47$ y $\beta = 0,263$, mientras que en la Fig. 4.7(c), se presenta la imagen resultante después de aplicar el histograma de ecualización. La Fig. 4.7(d) ilustra el comportamiento de la propuesta de este trabajo. Note como en la Fig. 4.7(d), aparecen con mayor claridad las características del rostro, además de que se evitan transiciones bruscas de iluminación, lo cual no ocurre con las imágenes en las Figs.4.7(b) y 4.7(c).

La desventaja de aplicar un mapeo de contraste de tres estados a nivel zona plana a imágenes con mala iluminación es que no existe forma de incrementar el nivel de gris de los píxeles de la imagen. Por ejemplo, el rango de niveles de gris de la imagen en la Fig. 4.7(a) se encuentra en el intervalo de 0-245. Si un mapeo de contraste morfológico a nivel zona plana es aplicado sobre una imagen, los niveles de gris de la imagen de salida permanecerán dentro de este rango.

Por otro lado, en la Fig. 4.7(c), el histograma de ecualización asigna el mejor valor de intensidad a cada píxel, considerando la información global de la imagen. Aunque en este caso, algunos de los píxeles incrementan su valor, no es suficiente para evitar cambios importantes de intensidad en la imagen de salida(4.6).

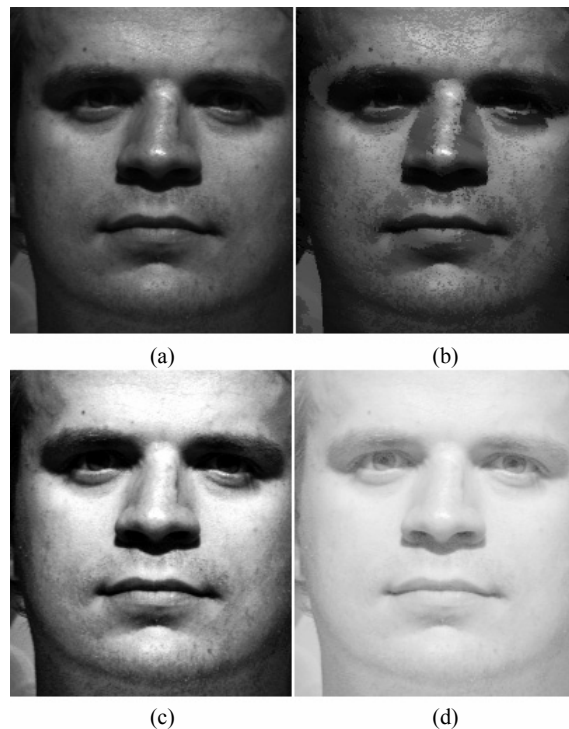


Figura 4.7: Aplicación de varios operadores de contraste. (a) Imagen original, (b) Mapeo morfológico de contraste a nivel zona plana, (c) Ecualización del histograma, (d) Operador de contraste de dos primitivas propuesto en este trabajo

Respecto a la imagen de la Fig.4.7(d), la mejora del contraste obedece a dos situaciones, la primera de ellas es el efecto aditivo del fondo de la imagen, el cual permite incrementar los valores de los píxeles de la imagen procesada, y en segundo lugar, la función logaritmo que distribuye los niveles de intensidad de una manera que se trata de evitar cambios bruscos de intensidad entre las diferentes regiones de la imagen. Es

claro que al aplicar nuestra transformación a imágenes donde existan grandes regiones con sombras, no será posible eliminarlas completamente; sin embargo la imagen de salida tendrá mejoras significativas, como se muestra en el ejemplo de las Figs. 4.7 (b), (c) y (d).

En la Fig. 4.8 se presenta la distribución de niveles de gris para las imágenes original y procesadas de la Fig. 4.7, en la Fig. 4.8(a) (imagen bajo prueba) se aprecia como la mayoría de los pixeles se concentran en zonas donde los niveles de gris son bajo valor (100 en este ejemplo). La Fig. 4.8(b) muestra el histograma para la imagen procesada con el operador de contraste de tres primitivas dado en la Ec. 4.9 de la Fig. 4.7(b), en este caso solo se evitan cambios bruscos en los niveles de gris, pero estos no incrementan su valor, por lo cual la imagen sigue dando una apariencia oscura.

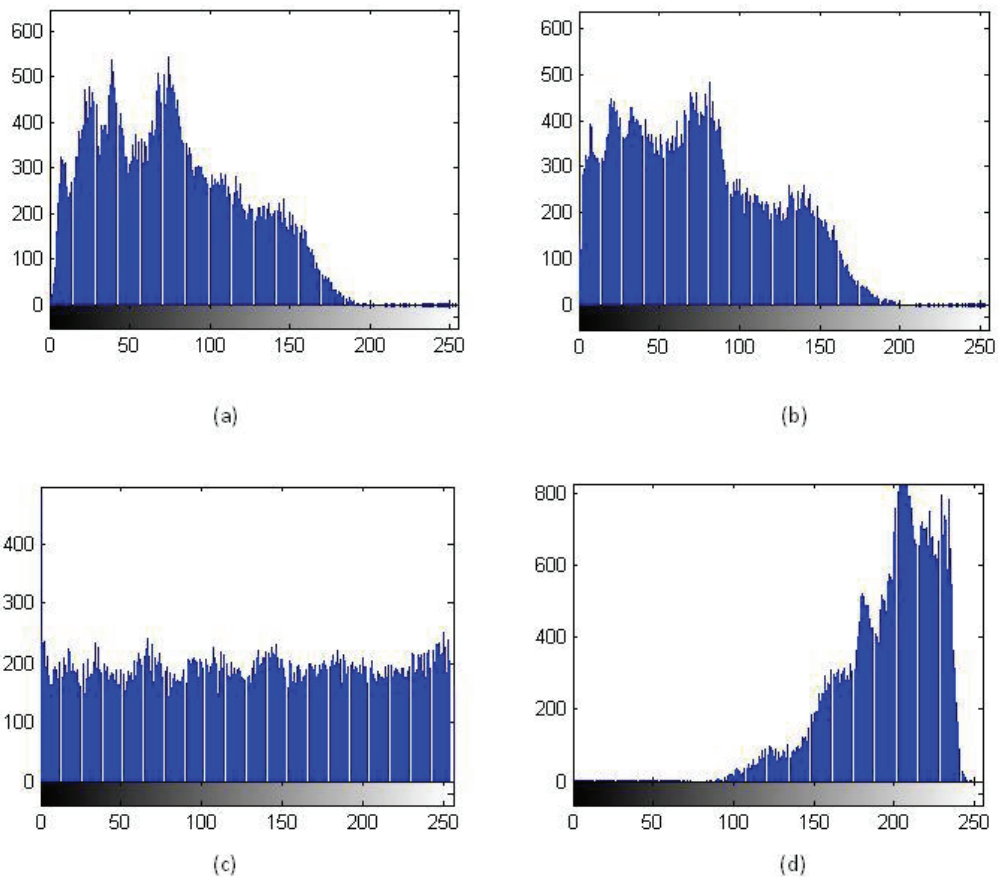


Figura 4.8: Distribución de los niveles de gris de las imágenes de la fig. 4.7(a), (b), (c) y (d) respectivamente.

En la Fig. 4.8(c) se presenta el histograma, de la imagen de la Fig. 4.7 (c) donde la distribución global de los niveles de gris es uniforme, como consecuencia pixeles se

ven incrementados en su nivel de gris pero no evitan los cambios bruscos de intensidad entre regiones vecinas de la imagen.

Por ultimo el histograma de la imagen procesada con el operador de contraste de dos primitivas (Fig. 4.7(d)) se muestra en la Fig. 4.8(d), debido al efecto aditivo del operador de contraste la intensidad de los pixeles se incrementa; el tratamiento por regiones a partir de la Ley de Weber ayuda a reducir los cambios bruscos de intensidad.

Un punto importante a mencionar es que hasta el momento se ha procedido al análisis de imágenes con iluminación deficiente y no se ha considerado el caso para las imágenes tomadas bajo condiciones adecuadas de iluminación. Por lo cual antes de aplicar el operador de contraste es necesario obtener información sobre las características de la imagen, por ejemplo se podría lograr por medio de datos obtenidos a partir del histograma, obteniendo una relación del porcentaje de pixeles totales de la imagen que se encuentran por debajo de cierto valor de umbral de intensidad.

Capítulo 5

Conclusiones

5.1. Resumen

Este trabajo de tesis está principalmente enfocado en presentar el desarrollo y caracterización de nuevas herramientas de procesamiento de imágenes; una de ellas corresponde a una técnica de análisis en el espacio de escala morfológico construido a partir de transformaciones morfológicas por reconstrucción y un operador de contraste de dos primitivas cuya finalidad es normalizar los niveles de gris en imágenes mal iluminadas. Cada una de ellas fue caracterizada para determinar sus propiedades básicas más importantes, los resultados obtenidos en ambos casos demuestran su buen funcionamiento. A continuación se discuten los puntos más sobresalientes que fueron observados, se divide la discusión en dos partes.

5.2. Discusión espacio escala apertura-cerradura por reconstrucción y cerradura-apertura por reconstrucción

Los resultados experimentales demostraron que a partir de la transformación cerradura-apertura por reconstrucción es posible obtener una representación de los fingerprint de una señal 1D similar a los obtenidos empleando otras metodologías (Jackway et al., 1994; Jackway, 1996). Los beneficios de utilizar dicha transformación se obtienen a partir de la mayor inmunidad al ruido que resulta de usar transformaciones por reconstrucción como primitivas en la transformada multiescala.

Las transformaciones apertura y cerradura por reconstrucción son filtros morfológicos que remueven características de la imagen de acuerdo a un criterio de tamaño de

elemento estructural y valor de nivel de gris en la imagen. Debido a que las transformaciones por reconstrucción emplean información de la imagen original para construir la imagen de salida, se conservan fielmente las características en la imagen de salida. Lo anterior es importante pues garantiza que al variar el parámetro de escala, los rasgos de importancia en la imagen no son difuminados o recorridos; i.e., los contornos de la imagen conservaran su posición y forma en un amplio rango de escalas.

Por otro lado la transformación apertura-cerradura por reconstrucción tiene tres propiedades importantes que la diferencian de aquellas encontradas en la literatura: (i) el operador apertura-cerradura por reconstrucción extiende los extremos regionales (máximos y mínimos) generando zonas planas, las cuales corresponden a regiones segmentadas de la imagen o señal original; (ii) ambos operadores de apertura-cerradura y cerradura-apertura por reconstrucción son menos sensibles a contornos ruidosos en comparación con otros métodos basados en la dilatación o erosión morfológica; y (iii) la generación de nuevos contornos es evitada ya que las transformaciones propuestas usan operadores conexos como primitivas.

5.3. Discusión realce de contraste en imágenes con iluminación deficiente

Las técnicas de realce de contraste aparte de mejorar la apariencia global de las imágenes, tienen como propósito final hacerlas más adecuadas para una aplicación o proceso determinado, como por ejemplo, segmentación, detección de contornos, reconocimiento de patrones, entre otros.

Aún cuando se han propuesto en la literatura diversas técnicas para realzar el contraste, el mejoramiento del mismo en imágenes afectadas por una iluminación desigual sigue siendo un problema de estudio, esto es debido por una parte a la naturaleza compleja de los objetos; y a los efectos producidos sobre tales objetos, principalmente como sombras o regiones muy claras. Las zonas afectadas pueden variar considerablemente modificando impredeciblemente la apariencia de las imágenes. En este trabajo se propuso un nuevo operador de contraste basado en la ley de Weber. Los parámetros involucrados en la transformación propuesta son determinados a partir de información de la imagen original.

La transformación propuesta tiene la característica de evitar los cambios bruscos

de iluminación. La forma en que los píxeles incrementan su valor de intensidad está en función del fondo de la imagen y la función logaritmo. El fondo de la imagen fue detectado a partir de los mínimos de la imagen original con ayuda de la cerradura por reconstrucción; mientras que la función logaritmo, además de incrementar la intensidad de los píxeles, también asigna valores que evitan los cambios bruscos de intensidad.

Finalmente, aun cuando las mejoras en el contraste en imágenes fuertemente influenciadas por las sombras no es completa, permite una mejora considerable como puede observarse en los experimentos presentados.

5.4. Trabajos futuros

Hay varias direcciones para las investigaciones en el futuro; una aplicación de visión por computadora que ha tomado auge corresponde a la detección y reconocimiento de rostros, ambas transformaciones presentadas en este trabajo de tesis pueden usarse como parte de un sistema de este tipo.

La influencia de la iluminación en la apariencia de los rostros puede compensarse empleando el operador de contraste de dos primitivas, facilitando con ello la etapa de segmentación del rostro la cual es fuertemente influenciada por la iluminación. El análisis multiescala morfológico para detección de rostros ya tiene sus orígenes en los trabajos reportados por Raducanu (2000); Kotropoulos et al. (2000), por lo que sería una ruta para una nueva investigación estudiar las consideraciones necesarias y ventajas de usar las transformaciones multiescala propuestas para obtener un sistema de reconocimiento de rostros robusto a variaciones debidas a iluminación deficiente.

Desde el punto de vista de la implementación, sería interesante proponer una plataforma de hardware y para el cálculo de las transformaciones; esto con la finalidad de obtener un sistema embebido que pueda operar en tiempo real sin necesidad de una computadora. En la actualidad los sistemas embebidos basados en FPGA's (Field Programmable Gate Array) ofrecen una alternativa poderosa para implementar algoritmos complejos en tiempo real; una característica adicional de estos dispositivos es que son reconfigurables y pueden fácilmente ser adaptados con nuevos algoritmos.

5.4.1. Propuesta de aplicación para el reconocimiento de rostros

Como se ha mencionado anteriormente, una aplicación interesante donde podrían emplearse el análisis multiescala es en un sistema automático de reconocimiento de rostros. Como no existe una metodología general desde el punto de vista de la morfología matemática para el reconocimiento de patrones, en esta sección se tratara de describir una posible solución al problema basado en el conocimiento de la estructura de la imagen y en algunos trabajos reportados en la literatura.

El reconocimiento de rostros es una tarea compleja, las teorías que se han venido manejando en la literatura apuntan a las redes neuronales, las maquinas de vector de soporte (support vector machines o SVM) y métodos basados en relaciones geométricas de los rasgos faciales.

En todo caso es necesario un operador que permita reducir la cantidad de información que debe ser procesada, en este sentido se han empleado el análisis de los componentes principales (PCA), contornos de la imagen o rasgos hallados a partir de la transformada wavelet.

Los fingerprint de la imagen obtenidos a partir de la representación multiescala pueden usarse como descriptores de la imagen, esto es, como puntos característicos de dimensionalidad reducida. De esta manera es posible analizar un conjunto reducido de puntos de la imagen en lugar de evaluar información de cada uno de los pixeles de la imagen.

El reconocimiento de rostros puede ser visto como un problema de reconocimiento de patrones, a partir de una base de datos de objetos en escenas, el objetivo es reconocer los objetos a partir de la base de datos. Extrayendo los fingerprint de cada objeto de la base de datos, el problema de reconocimiento de rostros llega a ser un problema de emparejamiento de patrones, el cual es solucionado en el “espacio de los fingerprints” por medio de un procedimiento de hipótesis y verificación.

Para cada objeto de la base de datos se puede seleccionar un rasgo dominante presente en un nivel de escala lo suficientemente grande, usando este rasgo dominante como origen obtenemos los fingerprints en el espacio de escala morfológico del objeto en coordenadas polares. De esta manera cada rasgo de importancia en la imagen toma la forma de una lista de valores de (r, θ, σ) .

Esta lista es ordenada de acuerdo al valor de escala σ y almacenada en la base de datos. Hay que notar que, bajo cambios aumento, rotación y translación, la relación proporcional r/σ sigue siendo una constante y puede usarse para emparejar rasgos en la escena. Las coordenadas polares son usadas para atender los cambios de rotación de los objetos en la escena.

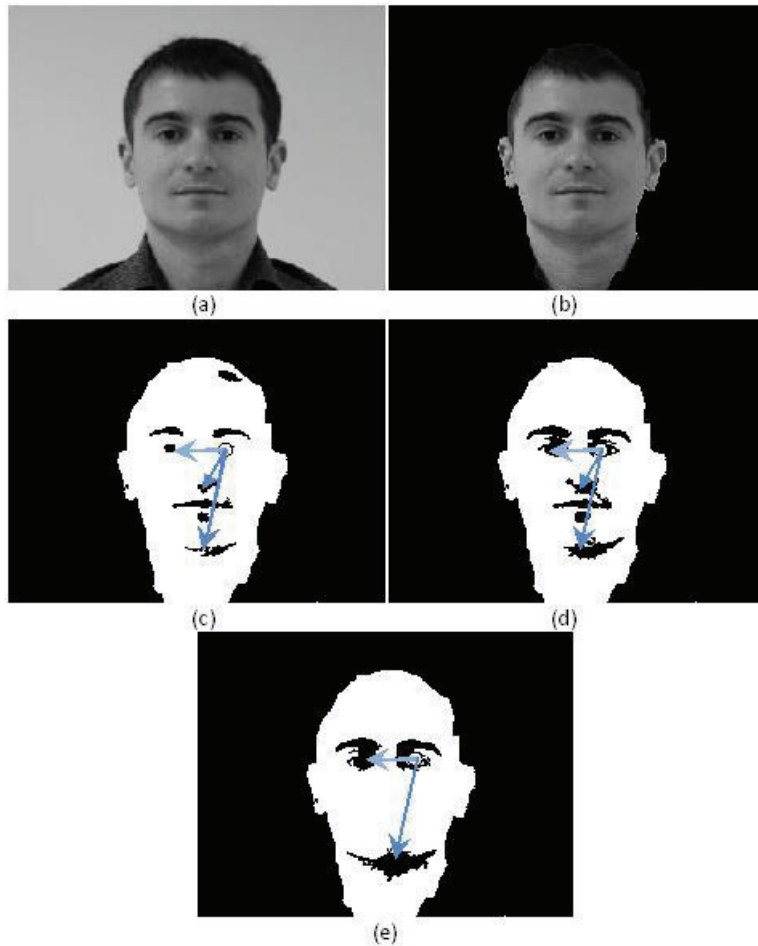


Figura 5.1: Proceso para reconocimiento de patrones a partir de los fingerprints de la imagen y coordenadas polares. (a) Imagen bajo prueba; (b) Rostro segmentado; (c), (d) y (e) fingerprints reducidos apertura-cerradura con valores de escala $\sigma = 3, 4$ y 5 .

Los fingerprint morfológicos multiescala reducidos de la imagen bajo prueba son hallados en coordenadas cartesianas y ordenados por valor de escala. Para plantear una hipótesis, se encuentra una lista de rasgos en la escena con valores de escala grandes, estos rasgos corresponderán al origen del objeto. Cada uno de los rasgos con un valor cercano a la relación r/σ de un rasgo genera una hipótesis para la próxima etapa. Cada hipótesis esta en la forma: **objeto:** m ; en la **posición:** (x, y) ; con **ángulo:** ϕ ; a

un **ampliación:** M.

Cada hipótesis es verificada o rechazada de acuerdo a si, bajo esta hipótesis, una proporción suficiente de los rasgos en el objeto coinciden con los rasgos en la escena y entonces son consistentes con la hipótesis. Este proceso se hecho en orden de escalas gruesas a escalas finas. Este método para reconocimiento de patrones es relativamente sencillo, una referencia a esta metodología se puede encontrar en el trabajo reportado por Jackway et al. (1994).

Para ilustrar mejor la idea consideraremos las imágenes de la Fig. 5.1, una imagen de la vista frontal de un rostro se muestra en la Fig. 5.1 (a), como solo se requiere analizas las características del rostro el cabello y cuello de la camisa son eliminados tal como aparece en la Fig. 5.1 (b). A continuación se extraen los fingerprints reducidos para valores de escala $\sigma = -3, -4$ y -5 (los cuales corresponden a los mínimos regionales de la imagen de la imagen original tras un cierre con tamaño de elemento estructural dado) como se ilustra en las Figs. 5.1 (c), (d) y (e) respectivamente.

Si consideramos los centros de masa (o centroide) de las regiones comprendidas en los mínimos regionales como los rasgos característicos; entonces el primer paso es seleccionar un rasgo dominante, en este ejemplo se considera el centro de masa del fingerprint donde es detectado el ojo derecho de la persona, este punto es considerado como origen a partir del cual se obtiene una lista de valores (r, θ, σ) en escalas mas finas ($\sigma = -3$ y -4). A partir de estos puntos se obtienen las hipótesis o relaciones estructurales que permitirán obtener un grado de correspondencia de patrones contenidos en una base de datos, con lo cual se obtiene información sobre la identidad de la persona.

Bibliografía

- Babaud, J., Witkin, A., Baudin, M., and Duda, R. (1986). Uniqueness of the gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):26–33.
- Bangham, J., Ling, P., and Harvey, R. (1996a). Scale-space from nonlinear filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):520–527.
- Bangham, J., Ling, P., and Young, R. (1996b). Multiscale recursive medians, scale-space, and transforms with applications to image processing. *IEEE Transactions on Image Processing*, 5(6):1043–1048.
- Basri, R. y Jacobs, D. W. (2003). Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):218–233.
- Belhumeur, P., Hespanha, J., and Kriegman, D. (1997). Eigenfaces vs fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720.
- Bosworth, J. H. y Acton, S. T. (2003). Morphological scale-space in image processing. *Digital Signal Processing*, 13(2):338–367.
- Chen, M.-H. y Yan, P.-F. (1989). A multiscale approach based on morphological filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):694–700.
- Georghiades, A. S., Belhumeur, P. M., and Kriegman, D. (2001). From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):630–660.
- Gonzalez, R.C. y Woods, R. (1992). *Digital Image Processing*. Prentice Hall.
- Haralick, R., Lin, C., and Lee, J.S.J. y Zhuang, X. (1987). Multi-resolution morphology. In *ICCV87*, pages 516–520.

- Heijmans, H. y Ronse, C. (1990). The algebraic basis of mathematical morphology-part i: Dilations and erosions. *Computer Vision, Graphics and Image Processing*, 50:245–295.
- Horn, B. K. (1986). *Robot Vision*. MIT. Press.
- Jackway, P. y Deriche, M. (1996). Scale-space properties of the multiscale morphological dilation-erosion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):38–51.
- Jackway, P. (1992). Scale-space properties of the multiscale morphological closing-opening filter. pages 278–281.
- Jackway, P. (1995). On dimensionality in multiscale morphological scale-space with elliptical poweroid structuring functions. *Journal Visual Comm. Image Representation*, 6(2):189–195.
- Jackway, P. (1998). On the scale-space theorem of chen and yan. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):351–352.
- Jackway, P., Boles, W., and Deriche, M. (1994). Morphological scale-space fingerprints and their use in object recognition in range images. *IEEE International Conference on Acoustics, Speech and Signal Processing, 1994. ICASSP-94*, 5.
- Jain, A. K. (1989). *Fundamentals of digital image processing*. Prentice Hall International Inc.
- Jang, B.K. y Chin, R. (1991). Shape analysis using morphological scale space. *Proc. 25th Ann. Conf. Information Sciences and Systems*, pages 1–4.
- Jiang, H. y Drew, M. (2003). Tracking objects with shadows. In *Proc. SPIE Image and Video Communications and Processing, 2003*.
- Kaspersek, J. (2001). Real time morphological image contrast enhancement in virtex fpga. In *Field-Programmable Logic and Applications*, pages 430–440. Springer-Verlag.
- Klinker, G., Shafer, S., and Kanade, T. (1990). A physical approach to color image understanding. *International Journal Computer Vision IJCV*, 4(1):7–38.
- Koenderink, J. J. (1984). The structure of images. *Biological Cybernetics*, (50):363–370.

- Kotropoulos, C., Tefas, A., and Pitas, I. (2000). Frontal face authentication using morphological elastic graph matching. *IEEE Transactions on Image Processing*, 9(4):555–560.
- Kotropoulos, C. y Pitas, I. (1997). Face authentication based on morphological grid matching. In *International Conference on Image Processing, ICIP'97*, pages Vol. 1: 105–108.
- Levine, M. D. (1985). *Vision in Man and Machine*. McGraw-Hill.
- Maragos, P. (1989). Pattern spectrum and multiscale shape representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(11):701–716.
- Marr, D. y Hildret, E. (1980). Theory of edge detection. *Proceedings of the Royal Society*, 207:187–217.
- Martens, J.-B. (1990). Application of scale space to image coding. *IEEE Transactions on Communications*, 38(9):1585–1591.
- Matheron, G. (1975). *Randoms Sets and Integral Geometry*. John Wiley and Sons, New York.
- Mendiola-Santibanez, J.D. y Terol-Villalobos, I. (2002a). Mapeos de contraste morfologicos sobre particiones basados en la nocion de zona plana. *Computacion y Sistemas*, (6):25–37.
- Mendiola-Santibanez, J.D. y Terol-Villalobos, I. (2002b). Morphological contrast enhancement using connected transformations. In *Proceedings of SPIE, 44667*, pages 365–376.
- Meyer, F. (1998). *From connected operations to levelings*. Kluwer.
- Meyer, F. y Serra, J. (1989). Activiy mappings. *Signal Processing*, 16:303–317.
- Meyer, F. y Maragos, P. (1999). Morphological scale-space representation with levelings. In *Scale-Space'99*, pages 187–198. Springer-Verlag.
- Meyer, F. y Maragos, P. (2000). Nonlinear scale-space representation with morphological levelings. *Journl of Visual Communication and Image Representation JVCIR*, 11(2):245–265.
- Mukhopadhyay, S. y Chanda, B. (2000). A multiscale approach to local contrast enhancement. *Signal Processing*, 80(4):685–696.

- Mukhopadhyay, S. y Chanda, B. (2003). Multiscale morphology segmentation of gray-scale images. *IEEE Transactions on Image Processing*, 12(5):533–549.
- Nacken, P. (1994). Openings can introduce zero crossings in boundary curvature. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6):656–658.
- Park, K.-R. y Lee, C.-N. (1996). Scale-space using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11):1121–1126.
- Petrovic, A., Divorrra, O., and Vandergheynst, P. (2004). Multiresolution segmentation of natural images: from linear to nonlinear scale-space representations. *IEEE Transactions on Image Processing*, 13(8).
- Raducanu, B. y Grana, M. (2000). Face localization on the morphological multi-scale fingerprint. In *International Conference on Pattern Recognition, ICPR2000*, pages Vol. 2: 925–928.
- Salembier, P. y Serra, J. (1995). Flat zones filtering, connected operators and filters by reconstruction. *IEEE Transactions on Image Processing*, 3(8):1153–1160.
- Serra, J. (1982). *Mathematical Morphology vol. I*. Academic Press, London.
- Serra, J. (1988). *Image Analysis and Mathematical Morphology*. Academic Press, London.
- Sherrier, R.H. y Johnson, G. (1987). Regionally adaptive histogram equalization of the chest. *IEEE Transactions on Medical Imaging*, 6:1–7.
- Short, J. and Kittler, J. y Messer, K. (2004). A comparison of photometric normalization algorithms for face verification. *IEEE International Conference Automatic Face and Gesture Recognition*, pages 254–259.
- Swain, M.J. y Ballard, D. (1991). Color indexing. *International Journal Computer Vision*, 7(1):11–32.
- Terol-Villalobos, I. (1996). Nonincreasing filters using morphological gradient criteria. *Optical Engineering*, 35:3172–3182.
- Terol-Villalobos, I. (1998a). *Toggle mappings and some related transformations. A study of contrast enhancement*.
- Terol-Villalobos, I. (2001). *Morphological Image Enhancement and Segmentation*, volume 118. Academic Press.

- Terol-Villalobos, I.R. y Cruz-Mandujano, J. (1998b). Contrast enhancement and image segmentation using a class of nonincreasing morphological filters. *Journal on Electronics Imaging*, 7:641–654.
- Toet, A. (1992). Multiscale contrast enhancement with application to image fusion. *Optical Engineering*, 32(5):1026–1031.
- Vachier, C. (2001). Morphological scale-space analysis and feature extraction. *Proceedings International Conference on Image Processing, 2001*, 3(5):676–679.
- Vincent, L. (1990). *Algorithmes Morphologiques a Base de Files d’Attente et de Lacets: Extension aux Graphes*. PhD thesis, Ecole des Mines, Paris.
- Vincent, L. (1993). Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. *IEEE Transactions on Image Processing*, 2(2):176–201.
- Witkin, A. (1983). Scale-space filtering. *Proc. IJCAI Carlsruhe, W. Germany*, pages 1019–1022.
- Xin, L. (2002). Low bit rate image coding in the scale space. *Proceedings Data Compression Conference, 2002, DCC 2002*, pages 33–42.
- Yuille, A. y Poggio, T. (1986). Scaling theorems for zero-crossings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):15–25.
- Yuille, A.L. y Poggio, T. (1985). Fingerprint theorems for zero crossings. *Journal of the Optical Society of America*, 2(5):683–692.