



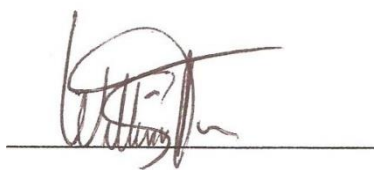
UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICAS

SEGMENTACIÓN NO SUPERVISADA DE IMÁGENES SATELITALES BASADA EN EL MÉTODO DE CRECIMIENTO DE REGIONES

Trabajo Especial de Grado
presentado ante la **ilustre**
Universidad Central de
Venezuela por la **Br. Katherine
Alexander Gomez Sanchez**
para optar al título de
Licenciado en Matemática
Tutor: Wuilian Torres

Caracas, Venezuela
Marzo 2015

Nosotros, los abajo firmantes, designados por la Universidad Central de Venezuela como integrantes del Jurado Examinador del Trabajo Especial de Grado titulado "Segmentación no supervisada de imágenes satelitales basada en el método de crecimiento de regiones", presentado por la Br. Katherine Alexander Gómez Sánchez, titular de la Cédula de Identidad V-19.195.447, certificamos que este trabajo cumple con los requisitos exigidos por nuestra Magna Casa de Estudios para optar al título de Licenciado en Matemática.

A handwritten signature in dark ink, appearing to be 'Torres Wuilian', written over a horizontal line.

Prof. Torres Wuilian
Tutor

A handwritten signature in dark ink, appearing to be 'Martín Miguel', written over a horizontal line.

Prof. Martín Miguel
Jurado

A handwritten signature in dark ink, appearing to be 'Figueroa Giovanni', written over a horizontal line.

Prof. Figueroa Giovanni
Jurado

Agradecimientos

Este trabajo representa el inicio de una vida profesional que si bien es incierta, mi convicción es hacer con todos los conocimientos adquiridos durante toda la carrera una práctica ética.

Gracias a mi familia, pues, a pesar de que vivimos en un mundo de trabajo, desde el punto de vista de Hanna Arendt, en un mundo de “fabricación de la pura variedad inacabable de cosas” donde el tiempo es valioso para la adquisición de la mayor cantidad posible de conocimientos con el fin de ser competente en una vida profesional y poder entrar en el mundo artificial que hemos construido con la excusa de una vida confortable. Su apoyo y comprensión ha sido invaluable en mi formación, Muchas Gracias.

Gracias a mis profesores y en especial a mi tutor Wuilian Torres, por el acompañamiento, la paciencia y los conocimientos brindados.

Y gracias a mis amigos y en general a todas las personas que directa o indirectamente influyeron en mí para que este camino haya sido tan alegre, especial y motivador.

Resumen

La segmentación de imágenes satelitales se refiere al proceso de dividir una imagen multiespectral proveniente de sensores a bordo de los satélites, en grupos de píxeles que tengan una propiedad común. En este trabajo el método utilizado para la segmentación es el de crecimiento de regiones, que consiste en la obtención automática de un conjunto de píxeles llamados semillas. Estas semillas son representativas de las distintas regiones en la imagen, a partir de ellas se van incorporando de manera iterativa los píxeles vecinos que cumplan con una propiedad definida. Como procesamiento preliminar se proponen y evalúan distintos filtros con el fin de conseguir el conjunto de semillas más apropiado al momento de aplicar las dos técnicas de segmentación utilizadas en este trabajo. Estas técnicas son: Transformada Watershed y la técnica basada en autómatas celulares Growcut. La primera de ellas es ampliamente conocida y ha sido presentada en muchos estudios anteriores, sirve de referencia para evaluar los resultados con los diferentes filtros propuestos. La segunda técnica, permite incorporar al proceso de segmentación diversas propiedades de las semillas con el fin de obtener un mejor criterio para el crecimiento de regiones. La segmentación se realizó sobre cuatro imágenes de diferentes zonas de Venezuela con regiones variadas, para ser evaluados los resultados por un especialista en interpretación de imágenes.

Palabras claves: imágenes satelitales, segmentación de imágenes, autómatas celulares, filtros morfológicos.

Contenido

INTRODUCCIÓN	7
CAPÍTULO 1. CONCEPTOS BÁSICOS DEL PROCESAMIENTO DE IMÁGENES	10
1.1 PROCESAMIENTO DIGITAL, IMAGEN DIGITAL, PÍXEL.....	10
1.2 VECINDAD.....	10
1.3 IMAGEN BINARIA	12
1.4 IMAGEN EN ESCALA DE GRISES.....	12
1.5 IMAGEN A COLOR	13
1.6 IMAGEN MULTIDIMENSIONAL	14
CAPÍTULO 2. IMÁGENES SATELITALES	15
2.1 TELEDETECCIÓN O PERCEPCIÓN REMOTA	15
2.2 ELEMENTOS DEL PROCESO DE TELEDETECCIÓN	16
2.3 LA ENERGÍA ELECTROMAGNÉTICA.....	16
2.4 EL ESPECTRO ELECTROMAGNÉTICO.....	17
2.5 REFLECTANCIA DE LAS SUPERFICIES TERRESTRES	18
2.6 RESOLUCIÓN DE LOS SENSORES REMOTOS.....	19
<i>Resolución Espacial</i>	19
<i>Resolución Espectral</i>	20
<i>Resolución Radiométrica</i>	20
<i>Resolución Temporal</i>	21
2.7 TIPOS DE IMÁGENES DE TELEDETECCIÓN	22
2.8 COMBINACIONES DE COLOR	23
2.9 TÉCNICAS DE PREPROCESAMIENTO DE IMÁGENES.....	26
2.9.1 <i>Análisis de Componentes Principales</i>	26
2.9.2 <i>Realce de Imágenes</i>	39
CAPÍTULO 3. MORFOLOGÍA MATEMÁTICA.....	41
3.1 NOCIONES SOBRE TEORÍA DE CONJUNTOS.....	41
3.1.1 <i>Propiedades básicas de las transformaciones morfológicas</i>	43
3.2 TRANSFORMACIONES MORFOLÓGICAS ELEMENTALES	46
3.2.1 <i>Erosión</i>	46
3.2.2 <i>Dilatación</i>	48
3.2.3 <i>Propiedades de las operaciones básicas de erosión y dilatación</i>	50
3.3 GRADIENTE MORFOLÓGICO.....	51
3.3.1 <i>Gradiente Multiescala</i>	53
3.4 APERTURA Y CIERRE	54
3.4.1 <i>Apertura morfológica</i>	55
3.4.2 <i>Cierre morfológico</i>	56
3.4.3 <i>Propiedades de la apertura y del cierre morfológico</i>	57
3.5 FILTRADO MORFOLÓGICO.....	58
3.6 MORFOLOGÍA PARA IMÁGENES MULTICANAL	60
CAPÍTULO 4. TRANSFORMACIONES GEODÉSICAS.....	61

4.1 DISTANCIA GEODÉSICA.....	61
4.2 DILATACIONES Y EROSIONES GEODÉSICAS	62
4.3 RECONSTRUCCIÓN GEODÉSICA	64
4.3.1 Apertura y cierre por reconstrucción	65
4.3.2 Reconstrucción auto-dual y nivelación	66
4.4 OPERADORES BASADOS EN RECONSTRUCCIÓN	69
4.4.1 Extremos regionales.....	69
4.4.2 Imposición de mínimos	70
4.4.2 Extendidos y h-extremos.....	71
CAPÍTULO 5. SEGMENTACIÓN	73
5.1 DETECCIÓN DE PUNTOS	74
5.2 DETECCIÓN DE LÍNEAS	74
5.3 DETECCIÓN DE BORDES.....	75
5.4 MÉTODOS DE SEGMENTACIÓN	76
5.4.1 Métodos basados en píxeles.....	77
5.4.2 Métodos basados en bordes.....	80
5.4.3 Métodos basados en crecimiento de regiones.....	80
5.4.4 Métodos basados en división y fusión de regiones.....	82
5.5 SEGMENTACIÓN MORFOLÓGICA	82
5.5.1 Esqueleto geodésico por zonas de influencia.....	83
5.5.2 La transformada Watershed.....	83
5.5.3. Segmentación controlada por marcadores.....	84
5.5.4 Autómatas Celulares.....	86
CAPÍTULO 6. RESULTADOS	90
6.1 IMAGEN DE ENTRADA	90
6.2 FILTRADO DE LA IMAGEN.....	91
6.2.1 Filtro Morfológico Alterno Secuencial.....	92
6.2.2 Reconstrucción Auto-Dual.....	93
6.2.3 Gradiente Multiescala.....	95
6.2.4 Apertura y cierre por reconstrucción	96
6.3 GRADIENTE MÁXIMO	98
7.4 OBTENCIÓN DE SEMILLAS	98
6.5 MÉTODOS DE SEGMENTACIÓN	99
6.5.1 Aplicación de la Transformada Watershed.....	100
6.5.2 Aplicación del Growcut	101
CAPÍTULO 7. EVALUACIÓN DE RESULTADOS	111
CONCLUSIONES.....	117
APÉNDICE. MATLAB	118
BIBLIOGRAFÍA	121

Introducción

Existen numerosos satélites de percepción remota orbitando alrededor de nuestro planeta, en particular, Venezuela posee su satélite propio de observación llamado Satélite Francisco de Miranda (VRSS1: Venezuelan Remote Sensing Satellite 1). Estos satélites proporcionan cuantiosa información para el estudio del territorio nacional, particularmente en el ámbito de gestión ambiental y sistemas productivos. Propiamente en los sistemas productivos se desarrollan proyectos de producción agrícola y estimación de cosechas, donde el análisis de las imágenes proporcionadas por los satélites son de gran ayuda en el monitoreo de la producción, la actualización de la frontera agrícola, el dimensionamiento de fenómenos que perturban la producción, así como la atención y mitigación de los efectos de contingencias climatológicas como sequías e inundaciones.

En el centro de Procesamiento Digital de Imágenes de la Fundación Instituto de Ingeniería de Caracas, Venezuela, los especialistas en interpretación de imágenes satelitales han tenido una importante experiencia con un nuevo paradigma denominado Análisis de Imágenes basado en Objetos (OBIA: Object Based Image Analysis) (Hay G., 2006). En las técnicas tradicionales, el monitoreo de la evolución temporal de la vegetación en zonas extensas conformadas por parcelas de tamaño, forma y estado fenológico variable, se hace por interpretación visual de las imágenes o utilizando un algoritmo de clasificación. La interpretación visual conlleva una gran dedicación de tiempo, donde se introducen los errores de subjetividad propios a la experiencia del especialista. Los métodos tradicionales de clasificación requieren del usuario una estimación preliminar del número de clases a considerar para su caracterización estadística. Se busca con esta investigación, una metodología que ubique a través de las herramientas que proporciona la Morfología Matemática representantes ideales de las distintas regiones presentes en la imagen, para posteriormente, caracterizar espectralmente en base a cálculos estadísticos cada una de ellas y obtener una segmentación que optimice el trabajo de los especialistas.

En este sentido, las imágenes satelitales son imágenes multicanal, o multidimensionales, y en su conjunto permiten caracterizar propiedades físicas de las zonas presentes en la escena. Es por esto, que se requiere de la implementación de programas que sean capaces de segmentar de manera automática la región bajo estudio. La segmentación a través del análisis de características de dichas regiones tiene como fin la clasificación y estudio temporal, por ejemplo, de cultivos como el arroz, maíz, sorgo, caraotas, entre otros.

Recientemente (Hernandez, 1996) (Popovici, 2002), han propuesto una modalidad de autómatas celulares que permite enfrentar la n-dimensionalidad que poseen las imágenes captadas por los múltiples sensores a bordo de los satélites. Estos autómatas poseen ventajas ya que son representados en arreglos de celdas que pueden tener cualquier número de dimensiones. Otra ventaja importante al segmentar las imágenes es integrar en el procesamiento información espacial (topología) con la información digital de los píxeles, en las n-dimensiones.

Sin embargo la práctica de la segmentación se extiende a un gran número de campos, donde se requiera el análisis de imágenes para el estudio de objetos o regiones delimitadas o particionadas que posteriormente pueden ser clasificadas y estudiadas, de acuerdo con el objetivo del usuario.

Cada capítulo de este trabajo describe la metodología seguida para la segmentación de imágenes satelitales a través de dos métodos, la transformada Watershed y el método basado en autómatas celulares, Growcut. Los siete capítulos cubren el contenido del área de la siguiente manera.

Capítulo 1: Conceptos Básicos del Procesamiento de Imágenes. Este capítulo contiene las definiciones fundamentales para el procesamiento digital de imágenes. Estos conceptos permiten introducir la terminología y las bases para abordar el resto del trabajo.

Capítulo 2: Imágenes satelitales. En este capítulo se aborda la temática relacionada con la obtención y características de las imágenes satelitales. Así como también la técnica de procesamiento de análisis de componentes principales.

Capítulo 3: Morfología Matemática. En este capítulo se introducen los conceptos básicos de la morfología matemática, que son necesarios para definir posteriormente, los operadores morfológicos que permiten procesar las imágenes digitales. Se introducen los filtros: alternos, alternos secuenciales y el gradiente morfológico.

Capítulo 4: Transformaciones Geodésicas. Se definen una serie de operadores morfológicos denominados operadores geodésicos, se abordan distintas transformaciones basadas en reconstrucción morfológica, así como los operadores basados en reconstrucción.

Capítulo 5: Segmentación. Definiciones, herramientas, técnicas y métodos utilizados para la segmentación de imágenes son abordados en este capítulo.

Capítulo 6: Resultados. Este capítulo muestra los resultados obtenidos a través de la descripción de los distintos procesos y técnicas aplicadas para lograr la segmentación de las imágenes.

Capítulo 7: Evaluación de Resultados. Se analizan los resultados de las segmentaciones obtenidas sobre diferentes imágenes, y la evaluación por un especialista.

Apéndice: El apéndice es una breve descripción de la herramienta Matlab, utilizada en el desarrollo de este trabajo especial de grado.

Propuesta de Solución

Objetivo General

Proponer un algoritmo para la segmentación no supervisada de imágenes satelitales, basado en la técnica de crecimiento de regiones con el uso de autómatas celulares deterministas.

Objetivos Específicos

- Implementar diferentes filtros en el pre-procesamiento de las imágenes para la homogeneización de las regiones de interés.
- Determinar el conjunto de píxeles-semillas, que representen cada una de las regiones de interés del usuario.
- Verificar la distribución de las semillas utilizando la segmentación morfológica por medio de la transformada de Watershed.
- Proponer un método basado en análisis estadísticos para las etiquetas de las semillas.
- Implementar un autómata celular para la segmentación de imágenes satelitales.

Capítulo 1. Conceptos Básicos del Procesamiento de Imágenes

Esta sección presenta conceptos y terminologías significativos que deben ser manejados para comprender y expresar propiedades de las imágenes.

1.1 Procesamiento Digital, Imagen Digital, Píxel

Se puede definir el Procesamiento Digital de imágenes como la ciencia de la modificación de imágenes digitales por medio de un ordenador digital. Los cambios que se producen en las imágenes por lo general se realizan de forma automática y se basan en algoritmos diseñados cuidadosamente. El Procesamiento de imágenes digitales puede implicar numerosos procedimientos, incluyendo formateo y corrección de datos, para facilitar una mejor interpretación visual, o incluso clasificación automatizada de objetos realizadas totalmente por el computador.

Una imagen es una representación visual de un objeto, una persona, o una escena producida por un dispositivo óptico tal como un espejo, un sensor, o una cámara. Puede ser representada por una matriz de dos dimensiones, $f(x, y)$, de tamaño $M \times N$, donde x y y son coordenadas espaciales, e indican el número de filas (entre 1 y M) y el número de columnas (entre 1 y N) respectivamente. La amplitud de f para cualquier par de coordenadas (x, y) se llama intensidad, nivel digital o nivel de gris de la imagen en ese punto. Cuando todos los valores de amplitud de f son cantidades discretas finitas, se llama imagen digital.

Una imagen digital entonces, se compone de un número finito de elementos, cada uno de los cuales tiene una ubicación y un valor particular. Cada uno de estos elementos se denomina Píxel. Cada píxel está representado por uno o más valores numéricos, esto depende de la dimensión de la imagen. Los valores máximos y mínimos que una intensidad de píxel puede asumir variará dependiendo del tipo de datos y convención utilizada.

Hay una gran variedad de maneras para representar una imagen en formato digital, en este trabajo de grado se utiliza el formato en el nivel básico Ráster, que utiliza una o más matrices de píxeles de dos dimensiones para representar una imagen, en el caso específico del trabajo una imagen satelital.

1.2 Vecindad

Los píxeles que rodean a un píxel dado constituyen su vecindad, se puede interpretar como una matriz pequeña que contiene el píxel referencia, generalmente centrada alrededor de él. La mayoría de las vecindades utilizadas en algoritmos de procesamiento

de imagen son pequeños arreglos cuadrados con un número impar de píxeles, por ejemplo, la zona de 3×3 se muestra en la Figura 1.1.

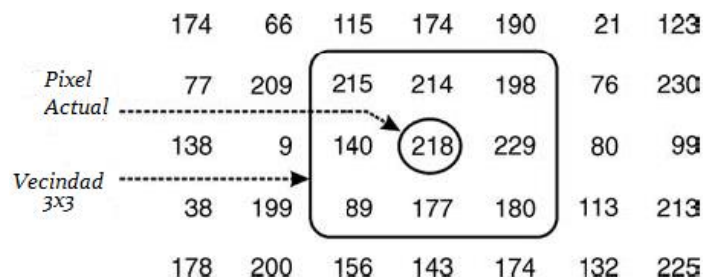


Figura 1.1. Vecindad de Píxeles.

En el contexto de la topología de la imagen, es común referirse a la vecindad-4 de un píxel como el conjunto de píxeles situado arriba, abajo, a la derecha, y hacia la izquierda del píxel de referencia (p), mientras que el conjunto de todos los vecinos inmediatos de p se refiere como su vecindad-8. Los píxeles que pertenecen a la vecindad-8, pero no a la vecindad-4, conforman la vecindad diagonal de p , la Figura 1.2 se muestra como representación.

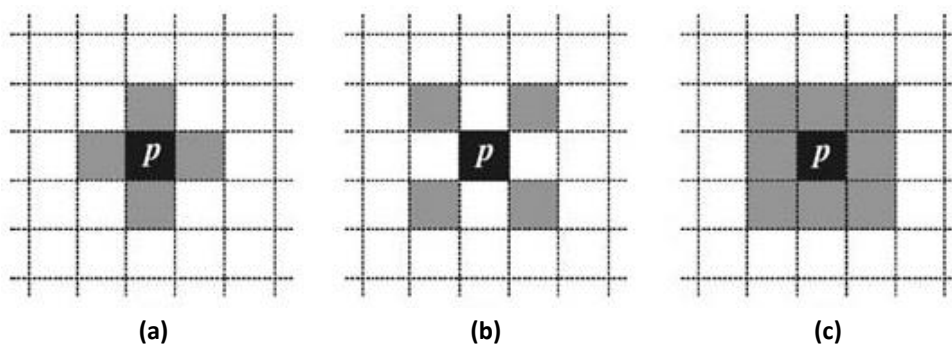


Figura 1.2 Vecindades del píxel p , desde una perspectiva topológica.
(a) vecindad-4, (b) vecindad diagonal, (c) vecindad-8.

Un conjunto de píxeles que están conectados el uno al otro se llama Componente conexo.

A continuación se presentan diferentes tipos de imágenes y su representación.

1.3 Imagen binaria

Las imágenes binarias se codifican como una matriz bidimensional (2D), por lo general por medio de 1 bit por píxel, donde un 0 por lo general significa "negro" y un 1 significa "blanco", aunque no hay un acuerdo universal sobre eso. La principal ventaja de esta representación, por lo general adecuada para imágenes que contienen gráficos simples, texto o imágenes de líneas es su pequeño tamaño (Marques, 2011). Un ejemplo de estas imágenes se observa en la figura 1.3.

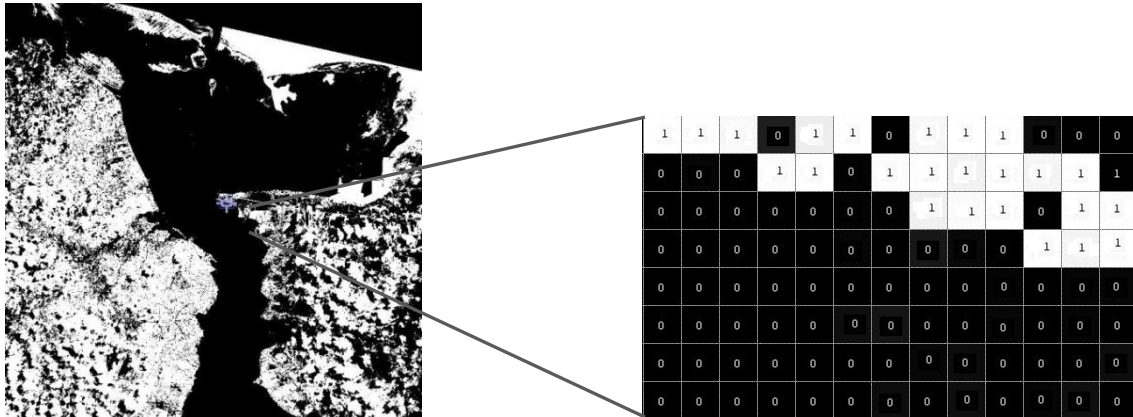


Figura 1.3 Imagen Binaria.

1.4 Imagen en escala de grises

Las imágenes en escala de grises (también conocidas como monocromática) se codifican como una matriz de píxeles 2D, con N bits por píxel, donde convencionalmente, el valor de píxel 0 corresponde a "negro", y el valor de píxel $2^N - 1$ pertenece al "blanco" y los valores intermedios indican distintas tonalidades de gris. El número total de niveles de gris es más grande que los requisitos del sistema visual humano que, en la mayoría de los casos, no pueden apreciar las variedades más allá de 64 niveles de gris, haciendo de este formato un buen compromiso en la calidad visual subjetiva. En la figura 1.4 podemos observar una imagen de 8 bits a escala de grises.

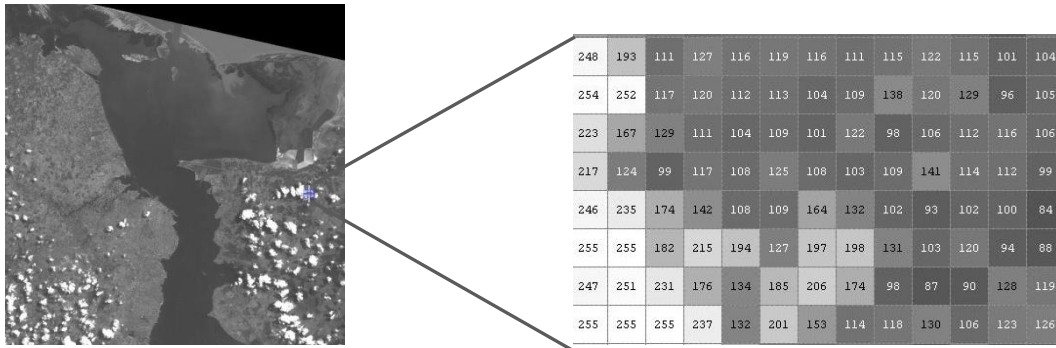


Figura 1.4. Imagen en escala de grises.

1.5 Imagen a color

La representación de imágenes en color es un poco más compleja y variada. La forma más común de almacenamiento del contenido de la imagen es la representación RGB, en la que cada píxel se representa usualmente por un número de N bits que contiene en sus componentes la cantidad de rojo (R), verde (G) y azul (B). Las imágenes a color se pueden representar mediante tres arreglos 2D del mismo tamaño, uno para cada canal de color: rojo (R), verde (G) y azul (B). Cada elemento de la matriz contiene un valor de N bits, que indica la cantidad de rojo, verde o azul de ese punto en una escala de $[0, 2^N - 1]$. La combinación de los tres valores de N bits permite $2^{N \times 3}$ colores posibles. Para ejemplificar esto, se presenta en la figura 1.5 una imagen a color de 8 bits, donde cada píxel contiene tres valores correspondientes a las componentes RGB, y cuyas cantidades están comprendidas entre 0 (negro) y 255 (blanco).

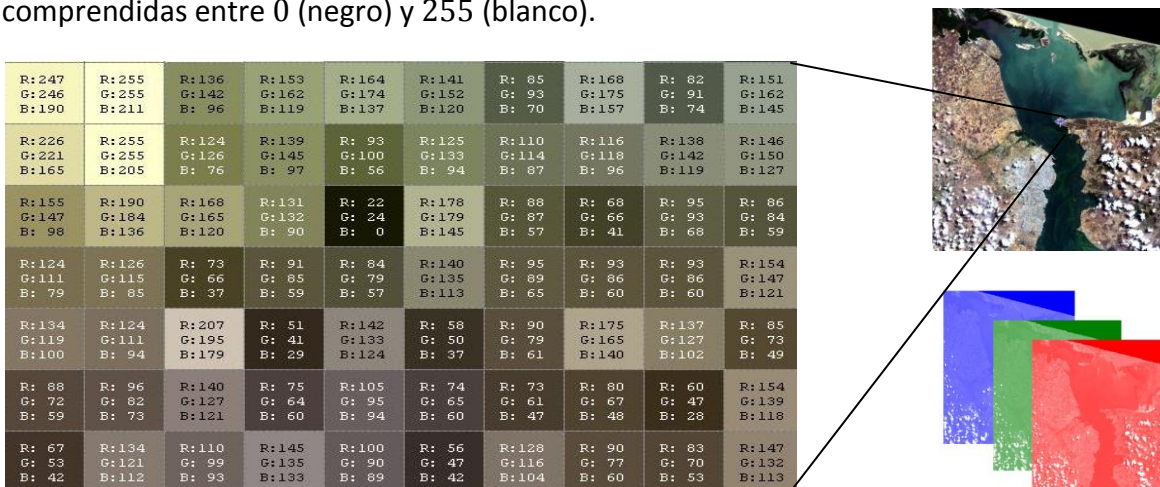


Figura 1.5. Representación de una Imagen a color.

1.6 Imagen Multidimensional

El procesamiento y el análisis de imágenes son vistos generalmente como operaciones en matrices bidimensionales de valores. Sin embargo, hay un número de campos en los que se debe analizar imágenes de más alta dimensionalidad. Muestra de ello son las imágenes médicas, imágenes biológicas y las imágenes satelitales.

Una imagen multidimensional permite almacenar dimensiones de datos por encima de una imagen simple 2D, por ejemplo, se puede extender una imagen simple para incluir dimensiones de tiempo, sonido y longitudes de ondas. En la figura 1.6 se muestra un ejemplo de representación de una imagen multidimensional por bandas.

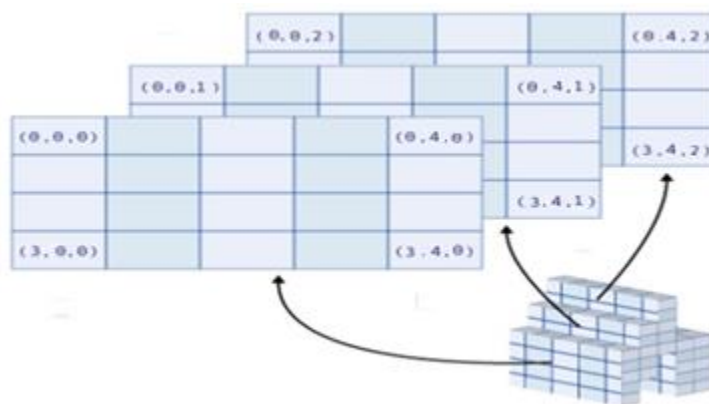


Figura 1.6. Representación de una imagen multidimensional.

Capítulo 2. Imágenes Satelitales

2.1 Teledetección o Percepción Remota

El término teledetección viene del francés *teledetection*, y percepción remota del inglés *remote sensing*. La teledetección es la ciencia de obtener información sobre un objeto sin tener que entrar en contacto directo con él (James, 2011). Otros autores, definen la teledetección como, procedimientos y técnicas que permiten evaluar características y propiedades físicas de un objeto o fenómeno de la superficie terrestre, sin estar en contacto directo con éste (Thomas, 1979). El proceso implica una interacción entre la radiación incidente y los objetos de interés. Generalmente los datos son recogidos en una imagen digital a través de sensores, los cuales captan la radiancia emitida o reflejada.

Un Sensor es un Instrumento que detecta radiación electromagnética y es capaz de convertirla en valores digitales para formar normalmente una imagen.

La teledetección como tecnología puede decirse que ha comenzado con la aparición de las primeras fotografías.

- La foto aérea - surgió en la década de 1840 con las fotos tomadas de globos.
- En la Primera Guerra Mundial, cámaras montadas en aviones proporcionaban vistas aéreas de áreas de superficie bastante grandes que resultó invaluable en reconocimiento militar.
- Desde entonces y hasta la década de 1960, la fotografía aérea mantuvo una única herramienta estándar para representar la superficie de la tierra.

El 23 de julio 1972 la NASA lanza el primer satélite de tecnología de recursos terrestres (ERTS-1). Los datos multiespectrales proporcionadas por los sensores a bordo llevaron a una mejor comprensión de la dinámica de las superficies cultivadas, los minerales, los suelos, el crecimiento urbano, y muchas otras características de la Tierra y los procesos. El nombre del satélite, y los que siguieron, pronto fue cambiado a Landsat, el más reciente denominado Landsat 8, fue lanzado el 11 de febrero de 2013.

El primer satélite de observación remota en Venezuela, se lanza el 28 de septiembre de 2012, y tiene como nombre Satélite Francisco de Miranda (VRSS-1). Entre los beneficios que aporta su utilización se encuentran: Dispone de una Base Cartográfica homogénea, precisa y actualizada; Seguimiento a los cambios en los cauces de los ríos y en los cuerpos de agua; Determinación en tiempo casi real de cualquier variación que se produzca en el territorio nacional; Realizar actualizaciones en cuanto a las variables uso y cobertura del territorio (ABAE, 2014).

2.2 Elementos del proceso de teledetección

Los elementos involucrados en un proceso de teledetección desde satélites se muestran en la Figura 2.1. El primer requerimiento supone disponer de una fuente de energía que ilumine o provea energía al objeto de interés (cultivo, bosque, mar, ciudad, etc.). El caso más habitual consiste en que esa fuente sea el Sol (A). La radiación solar, en su “viaje” hacia la Tierra, atraviesa e interacciona con la atmósfera (B). Una vez alcanza la superficie terrestre interactúa con los objetos que en ella se encuentran. La radiación reflejada dependerá de las características de esos objetos, permitiendo distinguir a unos de otros (C). Un sensor a bordo de un satélite recoge y graba esa radiación reflejada por la superficie terrestre y la propia atmósfera (D). La energía captada por el sensor se transmite a una estación de recepción y procesamiento donde los datos se convierten en imágenes digitales (E). La imagen procesada se interpreta, visualmente y/o digitalmente, para extraer información acerca de los objetos que fueron iluminados (F). El paso final del proceso de teledetección consiste en aplicar la información extraída de la imagen para conseguir un mejor conocimiento de la zona de estudio, revelando nuevas informaciones o ayudándonos a resolver un problema particular (G).

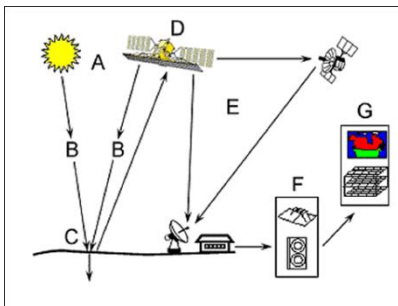


Figura 2.1. Elementos de un proceso de teledetección desde satélites.

Fuente: Fundamentals of Remote Sensing. Canada Centre for Remote Sensing, Remote Sensing Tutorial.

2.3 La energía electromagnética

La energía electromagnética se produce por la oscilación o aceleración de una carga eléctrica. Cualquier objeto natural o sintético que se encuentre a una temperatura superior a 0 °K genera energía electromagnética. La energía electromagnética se propaga a la velocidad de la luz y tiene como parámetros característicos la frecuencia de oscilación y la longitud de onda (Figura 2.2).

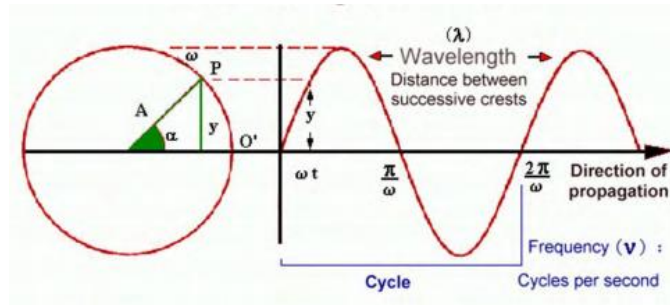


Figura 2.2. Ondas Electromagnéticas.

La longitud de onda (λ) y la frecuencia (f) tienen una relación inversa definida por la ecuación (1.1), donde c es la velocidad de la luz 300.000 km/s:

$$(1.1) \quad f = \frac{c}{\lambda}$$

La correspondencia entre frecuencia y longitud de onda puede observarse en la figura 2.2, la energía electromagnética de baja frecuencia es equivalente a ondas largas, mientras que la energía de frecuencia alta corresponde a ondas cortas.

2.4 El espectro electromagnético

Los ojos de los seres humanos se pueden considerar como sensores remotos ya que detectan la luz reflejada por los objetos de nuestro entorno. Sin embargo, la visión humana sólo es capaz de percibir una pequeña parte del espectro electromagnético, el visible. La luz visible es sólo una de las muchas formas de radiación electromagnética que existen. Así, las ondas de radio, el calor, los rayos ultravioleta o los rayos X son otras formas comunes. En teledetección, lo normal es caracterizar a las ondas electromagnéticas por su longitud de onda en micrómetros (μm) o nanómetros (nm), es decir, por la posición que ocupan dentro del espectro electromagnético. De esta forma quedan definidas varias regiones del espectro. Aunque por conveniencia se le asignan diferentes nombres a estas regiones (ultravioleta, visible, infrarrojo, microondas, etc.), no existen divisiones exactas entre unas y otras (Figura 2.3). Los sensores montados a bordo de los satélites de teledetección son capaces de detectar y grabar radiaciones de las regiones no visibles del espectro electromagnético, desde el ultravioleta hasta las microondas.

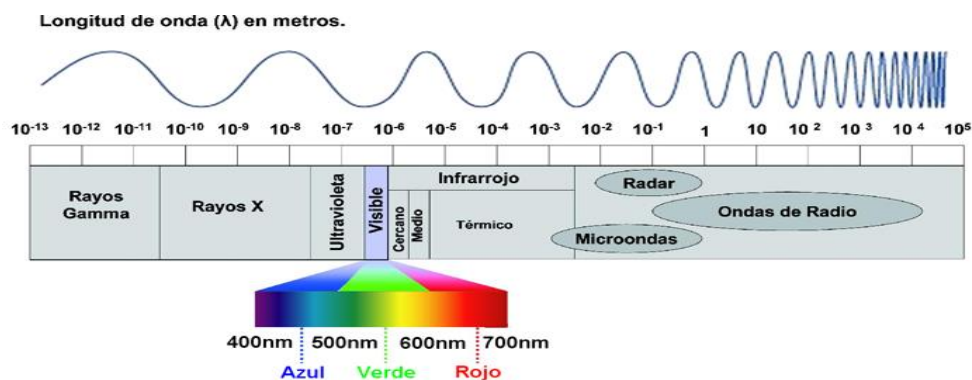


Figura 2.3. Espectro electromagnético.

Fuente: Página Web (Soc14).

El visible (VIS) es una pequeña región del espectro electromagnético que apenas abarca desde los 400 nm hasta los 700 nm. El color azul iría desde 400 hasta 500 nm, el verde desde 500 nm hasta 600 nm y el rojo de 600 nm a 700 nm. La energía ultravioleta (UV) se encuentra inmediatamente por debajo del color azul. Por encima del rojo se sitúa la región infrarroja (IR), que a su vez está dividida en tres categorías: IR cercano (NIR) (700 – 1300 nm), IR medio (SWIR) (1300 – 3000 nm) e IR térmico (TIR) (3000– 100000 nm). La porción de microondas se encuentra más allá del IR, a longitudes de onda mucho más grandes (1 mm – 1 m), que son las longitudes de onda más largas usadas en teledetección. De éstas, las más cortas tienen propiedades similares al IR térmico, mientras que las restantes son similares a las usadas en comunicaciones. En la actualidad existen discrepancias entre los científicos que usan las técnicas de teledetección en microondas y los encargados de establecer la asignación de frecuencias para las telecomunicaciones, ya que hay una gran competencia por determinar qué uso se le da a determinadas longitudes de onda dentro de esta región del espectro electromagnético.

La radiación electromagnética al interactuar con los objetos es modificada de manera que la energía resultante incorpora información que caracteriza al objeto y en consecuencia indirectamente permite su análisis. De acuerdo a su longitud de onda, una parte de la energía incidente en el objeto es dispersada en diferentes direcciones, otra parte es absorbida por la materia, otra es transmitida, es decir, cuando la radiación traspasa la superficie y otra parte de la energía es reflejada. En percepción remota la energía de interés para el análisis de los objetos en la superficie terrestre es la energía reflejada o reflectancia y la energía emitida o emitancia en el caso de mediciones de temperatura.

2.5 Reflectancia de las superficies terrestres

La reflectancia espectral es una característica de las superficies terrestres, algo fundamental en teledetección. Se define como la proporción de energía incidente que es reflejada por una superficie. Por lo tanto, es una magnitud adimensional que puede tomar valores entre 0 y 1 ó porcentajes entre 0 y 100%. Para una determinada superficie este

parámetro varía en función de la longitud de onda. Al gráfico de la reflectancia espectral frente a la longitud de onda se le denomina curva de reflectancia espectral o firma espectral (Figura 2.4). La configuración de estas curvas permite extraer las características espectrales de una superficie y tiene una gran influencia sobre la elección de la región espectral en la cual los datos de teledetección se deben adquirir para una aplicación particular. Así, por ejemplo, las curvas de reflectancia espectral para la vegetación casi siempre manifiestan los picos y valles que se muestran en color verde en la figura 2.4. Los valles en la región del visible vienen dados por los pigmentos en las hojas de las plantas.

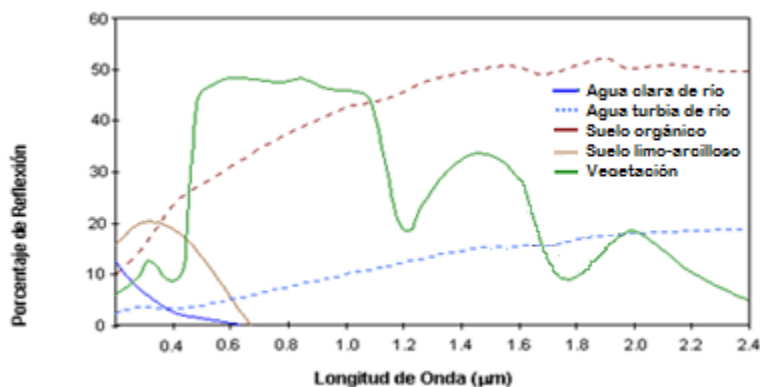


Figura 2.4. Curvas de Reflectancia Espectral.

2.6 Resolución de los sensores remotos

Los sensores instalados en los satélites de teledetección poseen una serie de particularidades que determinan las características de las imágenes que van a proporcionar. Estas características vienen definidas básicamente por diferentes tipos de resolución: Espacial, Espectral y Radiométricas.

Resolución Espacial

La resolución espacial es una medida de la distancia angular o lineal más pequeña que puede captar un sensor remoto de la superficie de la Tierra, y viene representada por un píxel. Un píxel es la unidad mínima que conforma una imagen digital (Figura 2.5).

466	515	487	404	414	486	536	554	551
402	465	538	489	434	506	517	464	507
344	396	493	540	500	470	527	514	496
331	346	398	488	548	488	503	522	507
326	342	363	409	514	544	516	526	541
325	327	339	352	462	536	539	510	538

Figura 2.5. Formato de una imagen digital.

La figura 2.5 representa el formato de una imagen digital de tamaño 6 x 9 píxeles. Cada píxel representa un área de la superficie terrestre. Los tonos de gris de cada píxel hacen referencia a distintos niveles de energía detectada.

La resolución espacial de un sensor se suele expresar en metros o metros/píxel. Son varios los factores que determinan la resolución espacial de un sensor remoto (distancia sensor-superficie terrestre, ángulo de visión y campo de visión instantáneo). Para el caso de los sensores a bordo de satélites estos factores son prácticamente fijos, por lo que la resolución espacial puede ser considerada constante, siempre y cuando el ángulo de visión no sea grande. Cuanto mayor sea la resolución espacial, es decir, menor superficie represente un píxel de la imagen, más pequeños serán los objetos que se pueden distinguir en la superficie y viceversa. A modo de ejemplo, una imagen con una resolución de 0,5 m/píxel permitirá distinguir objetos más pequeños que una imagen de 2 m/píxel. Para que un objeto homogéneo pueda ser detectado, su tamaño tiene que ser generalmente igual o más grande que la superficie de terreno que representa un píxel. Si el objeto es más pequeño puede que no sea detectado y el sensor grabará un promedio de todo lo que haya dentro. Sin embargo, algunas veces se detectan objetos muy pequeños porque su reflectancia domina dentro de la superficie del píxel.

Resolución Espectral

Como se comentó previamente, las distintas superficies responden de manera diferente a la radiación electromagnética. Esto significa que se puede obtener una firma espectral específica para cada superficie. Así, los diferentes tipos de superficie, naturales o no, se pueden identificar en base a sus firmas espectrales (Figura 2.4), pero será necesario que el espectro sea suficientemente detallado en términos de intervalos de longitud de onda y que cubra un rango espectral ancho. Los dispositivos de teledetección generalmente sólo muestrean el espectro electromagnético detectando la radiación en determinados intervalos de longitudes de onda. Por ejemplo, un sensor que es sensible a las longitudes de onda entre 0.4 y 0.5 μm detectaría la luz azul. Este intervalo se conoce con el nombre de banda espectral o canal de los datos de una imagen. Se define la resolución espectral de un sensor como el número y anchura de las bandas espectrales que puede discriminar. Un incremento en la resolución espectral resultará en un número mayor de canales o bandas espectrales. Sin embargo, esta resolución adicional supone también un costo en términos de volumen de datos e incremento del costo de procesamiento.

Resolución Radiométrica

La resolución radiométrica de los datos de teledetección se define como la cantidad mínima de energía requerida para incrementar el valor de un píxel en un nivel digital (ND). Asimismo, se define la amplitud o extensión radiométrica como el intervalo dinámico, o máximo número de niveles digitales, que pueden ser detectados por un sensor particular.

En los sensores más recientes lo habitual suele ser que los niveles vayan de 0 a 2047. En este caso hablaríamos de 11 bits de resolución radiométrica, ya que todos los valores de ese intervalo se pueden representar mediante 11 bits (dígitos binarios) en un sistema digital.

La resolución radiométrica en imágenes digitales es comparable al número de tonos de gris en una fotografía en blanco y negro, ya que ambos se relacionan con el contraste. El ojo humano solo es capaz de percibir aproximadamente 30 tonos de gris diferentes, lo que implica que normalmente la información visual en las imágenes digitales es menor a la que realmente contienen. Aunque la resolución radiométrica define el máximo número de niveles digitales detectables por un sensor, normalmente una imagen real no los contiene todos y además, no suele haber máximos y mínimos simultáneamente. En estos casos se pueden aplicar técnicas de tratamiento de imágenes para mejorar su apariencia visual, pero nunca la resolución radiométrica propia del sensor. La dispersión y absorción que provoca la atmósfera en la radiación que alcanza el sensor reducen el número de ND en las imágenes, especialmente en las longitudes de onda más cortas. A efectos visuales esto se traduciría en una pérdida de contraste. Existen procedimientos que permiten obtener medidas de reflectancia relativas a los objetos de la superficie eliminando o reduciendo el efecto de la atmósfera.

Resolución Temporal

La resolución temporal es el ciclo de repetición, o intervalo de tiempo, entre dos adquisiciones de imágenes sucesivas de una misma porción de la superficie y depende, en gran medida, de las características orbitales del satélite. Muchas veces también se la denomina periodo de revisita. Muchos satélites actuales tienen la capacidad de reorientar el sensor, lo que les permite aumentar su frecuencia de revisita para una zona determinada, muy importante en el seguimiento de desastres naturales o para detectar procesos que tienen poca perdurabilidad en el tiempo. La resolución temporal de un sensor depende principalmente de tres factores: capacidad de reorientación del sensor a ambos lados de la línea de paso del satélite, del ancho de barrido y de la latitud, ya que en el caso de órbitas cuasi polares, a mayor latitud, menor periodo de revisita.

La posibilidad de captar imágenes de una misma zona de la superficie terrestre en diferentes periodos de tiempo o épocas del año, es una de las características más importantes de los satélites de teledetección. Las características espectrales de una superficie terrestre pueden cambiar a lo largo del tiempo. Estos cambios pueden ser detectados con la adquisición y comparación de imágenes multitemporales.

En la tabla 2 se presentan las características de los satélites Francisco de Miranda, Landsat 7 y Landsat 8, referentes a las resoluciones antes mencionadas, así como también los sensores abordo de ellos.

Satélites	Sensores	RESOLUCIONES			
		Espacial (metros)	Espectral (N° de bandas)	Radiométrica (bits)	Temporal (días)
Miranda	MSS-2	16	4	10	12
	VRSS-1	10/Pan:2,5	5		57
Landsat 7	ETM+	30/Thermal:60/Pan:15	8	8	15
Landsat 8	OLI	30/Pan:15	9	16	16
	TIRS	100	2		

Tabla 2. Resoluciones y características de Satélites.

2.7 Tipos de imágenes de teledetección

El tipo de producto más común que suministran los satélites de teledetección es una imagen digital tipo ráster (ver figura 2.5), donde cada píxel tiene asignado uno o varios valores numéricos (niveles digitales) que hacen referencia a la energía media recibida dentro de una determinada banda espectral.

Los datos de una imagen digital se organizan en una matriz numérica de tres dimensiones (X,Y,Z) donde las dos primeras (X,Y) corresponden a las coordenadas de ubicación del píxel en la imagen, y la tercera (Z) indica la dimensión espectral. Un ejemplo se muestra en la figura 2.6.

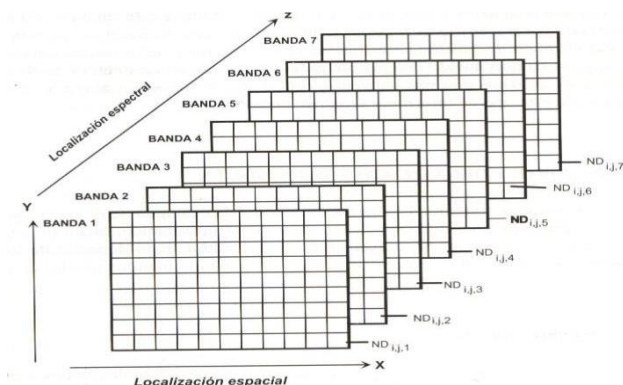


Figura 2.6. Representación de una imagen digital.

Teniendo esto en cuenta, se pueden adquirir los siguientes tipos de imágenes:

Imagen multispectral (MS). Imagen que asocia a cada píxel tantos niveles digitales como bandas espectrales sea capaz de detectar el sensor. A priori, es el tipo de producto más útil ya que nos proporciona, en cierto modo, la firma espectral de los distintos elementos

presentes en la imagen. Cuanto mayor sea el número de bandas que proporciona el sensor, mayor será la capacidad de análisis de los elementos presentes en la imagen.

Aparte de las multiespectrales también existen las denominadas **imágenes hiperespectrales**, menos habituales. Vienen caracterizadas por poseer información en un gran número de bandas. Se requieren para estudios de identificación y clasificación muy precisos, principalmente en estudios de prospección minera.

Imagen pancromática (PAN). Dispone de una sola banda espectral que abarca comúnmente gran parte del rango visible en el espectro electromagnético, obteniendo como resultado una imagen que habitualmente se representa en una escala de grises (imagen en blanco y negro). Como contrapartida, tienen la ventaja de poseer mayor resolución espacial que las multiespectrales que proporciona el mismo satélite. Es por ello que son muy interesantes para la detección de pequeños elementos de la superficie terrestre que no son distinguibles en la imagen multiespectral.

En aquellos satélites donde existe la posibilidad de obtener imágenes multiespectrales y pancromáticas de forma simultánea es habitual la opción de suministrar, ambas imágenes.

Imagen fusionada (PS). Este tipo de imagen se obtiene mediante la fusión de una imagen multiespectral con una pancromática. Las siglas PS provienen de *pan-sharpened*, su denominación en inglés. Básicamente, consiste en asignar a cada píxel de la imagen pancromática los valores procedentes de un algoritmo que combina la imagen pancromática con la multiespectral. El resultado final es una imagen multiespectral con la resolución espacial de la pancromática. El inconveniente de este tipo de imágenes es que se modifica a través de los algoritmos usados para la fusión, la información espectral original que es captada por los sensores, por lo que se suelen utilizar únicamente como herramientas de interpretación visual y no para análisis espectral. Esta fusión se encuentra dentro de la oferta de los distribuidores oficiales de los satélites capaces de obtener una imagen multiespectral y pancromática.

Imagen estéreo. En realidad se refiere a dos imágenes de una misma zona tomadas con ángulos de visión distintos. Muchos satélites tienen la capacidad de reorientar el sensor, lo que les permite tomar, en una o en sucesivas pasadas, este tipo de imágenes. Se suelen emplear para generar modelos de elevación del terreno.

2.8 Combinaciones de Color

Las imágenes satelitales suelen ser multiespectrales, es decir que son registradas simultáneamente en varias regiones del espectro electromagnético. Estas imágenes pueden ser estudiadas individualmente en escalas de grises o en imágenes coloreadas obtenidas a partir de las primeras. Estas últimas se generan según el modelo de color RGB (del inglés Red, Green, Blue). Este hace referencia a la composición del color

en términos de la intensidad de los colores primarios con los que se forma: el rojo, el verde y el azul. Es un modelo de color basado en la síntesis aditiva, es decir basado en la mezcla por adición de dichos primarios. Para indicar en qué proporción se mezcla cada color se asigna un valor a cada uno de los colores primarios.

Así por ejemplo, para una imagen con resolución radiométrica de 8 bits por píxel el rango de valores posibles para cada componente de color es de 0 a 255. En consecuencia a cada píxel de una combinación de color le corresponderá un punto dentro del cubo de color representado en la figura 2.7(a). La figura 2.7(b) es una extracción de un sector para mejor visualización interior.

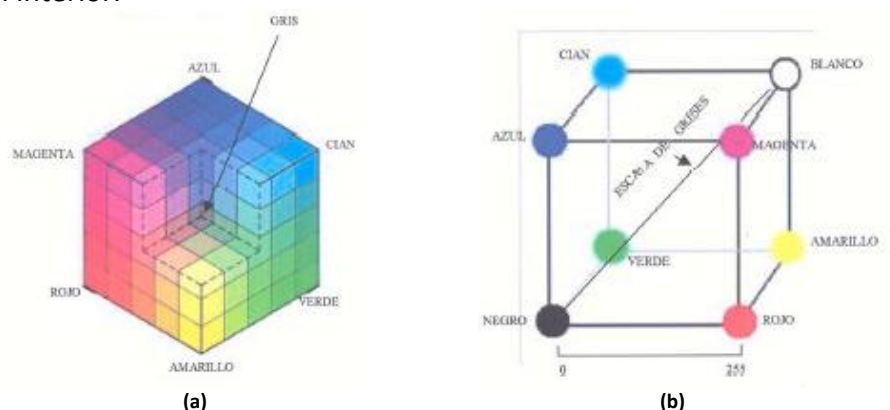


Figura 2.7. Representación de la combinación de colores

Por lo tanto, las coordenadas del rojo serán (255,0,0), del verde (0,255,0) y del azul (0,0,255). La ausencia de color, es decir el negro corresponde al punto (0,0,0). La combinación de dos colores a nivel 255 con un tercero a nivel 0 da lugar a tres colores intermedios: el amarillo (255,255,0), el cian (0,255,255) y el magenta (255,0,255). El blanco se forma con los tres colores primarios a su máximo nivel (255,255,255).

La escala de grises es la diagonal que une el blanco y el negro, corresponde a la misma cantidad de rojo, verde y azul. Para visualizar una imagen se hace la combinación de tres bandas de interés y se le asigna a cada una de ellas uno de los tres colores primarios. La primera corresponderá al color rojo, la segunda al verde, y la tercera al azul.

La imagen que se muestra entonces, se presenta en una combinación RGB correspondiente a las bandas seleccionadas y a la asignación de colores. Las bandas a seleccionar quedarán condicionadas, por una parte por las bandas ofrecidas por el sensor del satélite, y por otra a criterio de los especialistas. Sin embargo existen ciertas combinaciones que son de uso frecuente, dichas combinaciones son:

Color Natural [Rojo, Verde, Azul]

Involucra las tres longitudes de onda al cual es sensible el ojo humano (rojo, verde y azul) y se le asigna a cada una de ellas su verdadero color, resultando una combinación que se aproxima a los colores naturales de la escena. La vegetación aparece en diversos tonos de verde y los suelos en colores marrones o tostados. Además, las bandas visibles tienen buena penetración en los cuerpos de agua y esta combinación permite observar detalles en agua poco profundas (turbidez, corrientes, batimetría, plumas de sedimentos, etc.).

Falso Color Infrarrojo [Infrarrojo Cercano, Rojo, Verde]

Al asignar el rojo a la bandas Infrarrojo cercano resultará que todos los tonos de rojo, desde el rojo muy oscuro al rosado pálido corresponden a vegetación: los tonos más oscuros se relacionan con la presencia de vegetación arbórea densa, como bosques, montes, etc., pudiéndose identificar algunas especies como pinos y eucaliptos. Rojos claros indican cultivos y praderas vigorosas. Los rosados corresponden a vegetación poco densa como pasturas naturales. El celeste corresponde a suelo desnudo, seco o áreas rocosas. El azul-celeste puede corresponder a zonas urbanas. El azul oscuro a negro se relaciona con agua clara en cursos o cuerpos de agua. Si éstos presentan partículas en suspensión aparecerán con tonalidades azul-verdosas (cian). Los colores entre el verde y el verde azulado corresponden a parcelas aradas o suelos descubiertos con mayor o menor contenido de humedad.

Tierra/Agua [Infrarrojo Cercano, Infrarrojo, Rojo]

Al asignarle el color rojo a la infrarroja cercana, esta banda va a tener ciertas similitudes con la combinación falso color infrarrojo. Sin embargo, al dar más peso a la región infrarroja se ve realzada la diferencia de humedad en suelos y vegetales. Generalmente cuanto mayor es la humedad del suelo más oscuro aparecerá éste. En la figura 2.8, se presenta una subimagen Landsat 8, de cultivos en el estado Guárico, Venezuela. La imagen es sometida a las combinaciones de color antes mencionadas.

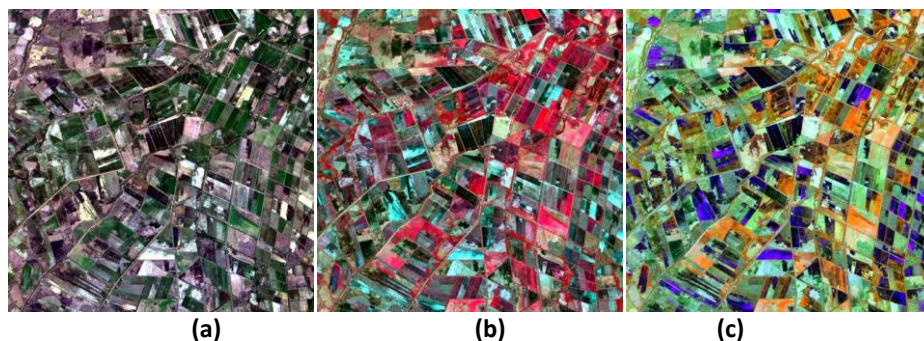


Figura 2.8. Combinaciones de bandas. (a) Color natural [4,3,2], (b) Falso color infrarrojo [5,4,3], (c) Tierra/agua [5,6,4].

Las combinaciones de color que son mencionadas anteriormente hacen referencia a las bandas del LANDSAT 8, pero los criterios son aplicables a otros satélites teniendo en cuenta las características espectrales de sus bandas. No todas las imágenes satelitales presentan la diversidad de bandas del LANDSAT 8, y en tales casos algunas de las combinaciones antes referidas no pueden realizarse con las escenas de dichos satélites, no obstante lo cual el analista puede intentar otras combinaciones que permitan realzar características de interés.

2.9 Técnicas de Preprocesamiento de Imágenes

En este apartado se abordan dos técnicas de procesamiento digital que permiten mejorar las imágenes con el fin de facilitar la extracción de información en cada uno de sus píxeles.

2.9.1 Análisis de Componentes Principales

Las imágenes satelitales presentan un volumen importante de información que se incrementa con la alta resolución espacial, el número de bits que representa cada píxel y la resolución espectral o número de bandas. Así cualquier procesamiento que se haga sobre la imagen conlleva a emplear mucho tiempo en la ejecución de los procesos llevados a cabo por la computadora. Para incrementar la eficiencia computacional y mejorar otras características que se explica más adelante, se propone el análisis de componentes principales como un procesamiento previo que se detalla a continuación.

Teniendo presente el carácter matricial de cualquier imagen digital, se pueden realizar sobre ella transformaciones y operaciones estadísticas. Por ejemplo, con los datos de una imagen digital se pueden calcular medidas de tendencia central y dispersión en cada banda, aumentar el contraste, cambiar su orientación numérica (rotación de la matriz), realizar combinaciones aritméticas entre bandas, sintetizar varias bandas reduciendo la información redundante (componentes principales) o discriminar grupos con niveles digitales homogéneos dentro de la matriz (clasificación).

El carácter digital de las imágenes y su forma vectorial permite generar nuevas imágenes aplicando transformaciones lineales a una o varias imágenes. Las nuevas imágenes representan una descripción alternativa de los datos, en la cual los nuevos niveles digitales de un píxel están relacionados con sus niveles digitales originales a través de una operación lineal. La imagen transformada puede destacar ciertas características que no era posible discernir en los datos originales o, alternativamente, preservar la información esencial contenida en la imagen en un número menor de dimensiones. Las transformaciones pueden llevarse a cabo para realizar un mejoramiento de la imagen o como un análisis previo a la aplicación de técnicas de clasificación.

El objetivo del Análisis de Componentes Principales (ACP) es resumir un grupo amplio de variables en un nuevo conjunto, más pequeño, sin perder una parte significativa de la información original. Para el usuario final de productos de teledetección, el objetivo del ACP es construir una o varias imágenes que incrementen su capacidad de diferenciar distintas coberturas. Es por ello que al realizar una composición color resulta interesante usar, en lugar de algunas bandas de la imagen, los componentes principales 1, 2 y 3 en la secuencia RGB respectivamente.

El ACP puede aplicarse como realce previo a la interpretación visual o como procesamiento anterior a la clasificación. En general, esta técnica incrementa la eficiencia computacional de la clasificación porque reduce la dimensionalidad de los datos.

Por otra parte, desde el punto de vista estadístico, el ACP facilita una primera interpretación sobre los ejes de variabilidad de la imagen, lo que permite identificar aquellos rasgos que aparecen en la mayoría de las bandas y aquellos otros que son específicos de algún grupo de ellas. Este trabajo se refiere a casos en los que interesa identificar la información común a la mayoría de las bandas, que está presente en los primeros componentes.

El ACP también es usado en aplicaciones multitemporales con el objeto de detectar cambios en distintas fechas. En este caso los primeros componentes resultantes del análisis no son los más interesantes ya que recogen información común a las distintas fechas. Los últimos componentes ofrecen la información no común (el cambio) que es lo que interesa en este contexto.

Algebraicamente, el ACP genera nuevas variables (componentes), mediante una combinación lineal de las p variables originales (bandas). Aunque se requieren los p componentes principales para reproducir la variabilidad total, muchas veces la mayor parte de ella está contenida en un número menor de componentes m . En ese caso, reemplazando las p bandas por los m componentes, se reduce la dimensionalidad del problema conservando casi la totalidad de la información.

En teledetección, la adquisición de imágenes en bandas adyacentes del espectro implica, con frecuencia, detectar información redundante (en apariencia las bandas de la imagen se visualizan de manera similar). Por ello, los niveles digitales de los píxeles de una banda pueden presentar una importante relación con los de otra, resultando una o más de una de ellas irrelevantes.

Aunque la imagen puede representarse como en una matriz tridimensional, para realizar el ACP se utiliza una matriz bidimensional. Estadísticamente los píxeles de cada banda representan las realizaciones de un evento, cada banda es considerada como una nueva

variable del evento, por ello las realizaciones de cada variable se presentan en una columna. Formalmente, los niveles digitales de los n píxeles de una imagen en p bandas pueden arreglarse en una matriz $X_{n \times p}$,

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

La i -ésima columna de la matriz contiene los n niveles digitales de la i -ésima banda.

Como el ACP es un análisis descriptivo no requiere que X tenga distribución normal multivariada. Si X tuviera esta distribución se podría realizar inferencia (Mardia, y otros, 1982). El estudio de la relación entre bandas, que es la base del ACP, puede realizarse de dos maneras:

- Con la matriz de varianza-covarianza Σ_X :

$$\Sigma_X = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{bmatrix}$$

Donde los elementos de la diagonal son las varianzas de los niveles digitales en cada banda:

$$\sigma_i^2 = \frac{1}{p} \sum_{k=1}^p [x_{ik} - E(x_i)]^2$$

Donde, $E(x_i) = \frac{1}{p} \sum_{k=1}^p x_{ik}$ (Esperanza) y los elementos fuera de la diagonal son las covarianzas entre los niveles digitales de dos bandas:

$$\sigma_{ij} = \frac{1}{p} \sum_{k=1}^p [x_{ik} - E(x_i)][x_{jk} - E(x_j)]$$

Como la covarianza entre la banda i y la j es la misma que entre la banda j y la i ($\sigma_{ij} = \sigma_{ji}$) la matriz Σ_X es simétrica.

- Con la matriz de correlación ρ_X :

$$\rho_X = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pp} \end{bmatrix}$$

Donde los elementos ρ_{ij} son los coeficientes de correlación lineal de Pearson:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$$

Donde:

σ_{ij} es la covarianza entre las bandas i, j .

σ_i es la desviación estándar de x_{ik} .

σ_j es la desviación estándar de la variable x_{jk} .

Los elementos de la diagonal son unos porque son las correlaciones de cada banda consigo misma. Como la correlación entre la banda i y la j es la misma que entre la banda j y la i ($\rho_{ij} = \rho_{ji}$) la matriz ρ_X es simétrica. Cuando hay relación lineal entre pares de bandas las correlaciones son cercanas a 1 ó a -1.

Cuando no hay relación entre las bandas, es decir, cuando son variables independientes, ambas matrices son diagonales (los elementos fuera de la diagonal son ceros, pues $\sigma_{ij} = 0$). En este caso cada banda aporta información diferente y por lo tanto el ACP sería innecesario (en teledetección esta situación es poco común). El objetivo del ACP es generar un nuevo sistema de coordenadas en el espacio multispectral en el cual los datos pueden ser representados sin correlación, de tal manera que la matriz de varianza-covarianza sea diagonal en el nuevo sistema de coordenadas. La figura 2.9 representa un sistema de coordenadas transformado donde los vectores tienen componentes no correlacionados.

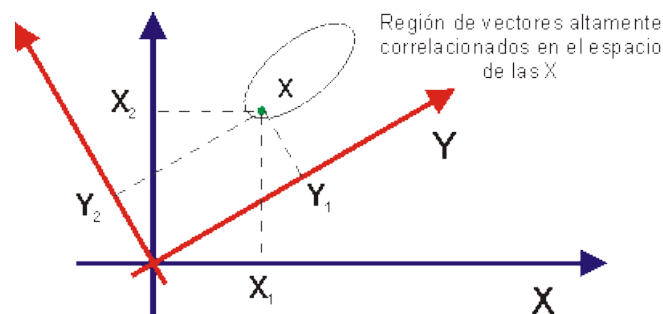


Figura 2.9. Ilustración de un sistema de coordenadas modificado en el cual los vectores tienen componentes no correlacionados.

Componentes principales obtenidos usando la matriz de varianza-covarianza

Como se dijo, los componentes principales son nuevas variables $Y_1, Y_2, \dots, Y_p$ que se obtienen como combinaciones lineales de las variables originales (bandas):

$$\begin{aligned} Y_1 &= a_1^t X = a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p \\ Y_2 &= a_2^t X = a_{12}X_1 + a_{22}X_2 + \dots + a_{p2}X_p \\ &\vdots \\ Y_p &= a_p^t X = a_{1p}X_1 + a_{2p}X_2 + \dots + a_{pp}X_p \end{aligned}$$

Como los Y_i son combinaciones lineales de los X_i , se tiene:

$$\text{Var}(Y_i) = a_i^t \Sigma a_i \quad i = 1, 2, \dots, p$$

$$\text{Cov}(Y_i, Y_k) = a_i^t \Sigma a_k = 0 \quad i \neq k$$

De todas las combinaciones lineales posibles, los componentes principales son aquellas que no están correlacionadas y tienen máxima varianza. Como la varianza se incrementa multiplicando el vector de coeficientes a_i por una constante, para que esta combinación lineal sea única es conveniente usar los vectores de coeficientes normalizados, es decir con longitud 1 ($a_i^t a_i = 1$).

Para demostrar esto se hace uso de los siguientes resultados:

Lema de Maximización. Sea B una matriz definida positiva y $d_{(px1)}$ un vector dado. Entonces, para un vector $x_{(px1)}$ distinto de cero se tiene,

$$(2.1) \quad \max_{x \neq 0} \frac{(x^t d)^2}{x^t B x} = d^t B^{-1} d$$

Con el máximo alcanzado cuando $x_{(px1)} = c B_{(pxp)}^{-1} d_{(px1)}$ para cualquier constante $c \neq 0$.

Demostración.

Por la extensión de la desigualdad de Cauchy-Schwarz, $(x^t d)^2 \leq (x^t B x) (d^t B^{-1} d)$. Como $x \neq 0$ y B es definida positiva, $x^t B x > 0$. Dividiendo ambos miembros de la desigualdad por el escalar positivo $x^t B x$ se obtiene el límite superior

$$\frac{(x^t d)^2}{x^t B x} \leq d^t B^{-1} d$$

Tomando el máximo sobre x obtenemos (2.1) ya que el límite se alcanza para

$$x = c B^{-1} d \blacksquare$$

Maximización de Formas Cuadráticas para Puntos en la Esfera Unitaria. Sea B una matriz definida positiva con autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ y vectores $e_1, e_2, \dots, e_p$ normalizados asociados. Entonces

$$(2.2) \quad \max_{x \neq 0} \frac{x^t B x}{x^t x} = \lambda_1$$

Alcanzado cuando $x = e_1$

Por otra parte,

$$(2.3) \quad \max_{x \perp e_1, \dots, e_k} \frac{x^t B x}{x^t x} = \lambda_{k+1}$$

Alcanzado cuando $x = e_{k+1}, k = 1, 2, \dots, p-1$

Demostración. Sean $P_{(p \times p)}$ la matriz ortogonal cuyas columnas son los autovectores $e_1, e_2, \dots, e_p$ y Λ la matriz diagonal con los autovalores $\lambda_1, \lambda_2, \dots, \lambda_p$ en la diagonal principal. Sean $B^{1/2} = P \Lambda^{1/2} P^t$ y $y_{(p \times 1)} = P^t_{(p \times p)} x_{(p \times 1)}$.

Por consiguiente, $x \neq 0$ implica $y \neq 0$. Así,

$$\frac{x^t B x}{x^t x} = \frac{x^t B^{1/2} B^{1/2} x}{x^t P P^t x} = \frac{x^t P \Lambda^{1/2} P^t P \Lambda^{1/2} P^t x}{y^t y} = \frac{y^t \Lambda y}{y^t y}$$

Como $P P^t = I_{(p \times p)}$

$$(2.4) \quad = \frac{\sum_{i=1}^p \lambda_i y_i^2}{\sum_{i=1}^p y_i^2} \leq \lambda_1 \frac{\sum_{i=1}^p y_i^2}{\sum_{i=1}^p y_i^2} = \lambda_1$$

Estableciendo $x = e_1$ se obtiene,

$$y = P^t e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \text{ pues } e_k' e_1 = \begin{cases} 1, & k = 1 \\ 0, & k \neq 1 \end{cases}$$

para la elección de este x , se tiene $y^t \Lambda y / y^t y = \lambda_1 / 1 = \lambda_1$ ó

$$\frac{e_1^t B e_1}{e_1^t e_1} = e_1^t B e_1 = \lambda_1$$

Ahora, $x = P y = y_1 e_1 + y_2 e_2 + \dots + y_p e_p$, así $x \perp e_1, \dots, e_k$ implica

$$0 = e_1^t x = y_1 e_1^t e_1 + y_2 e_1^t e_2 + y_p e_1^t e_p = y_i, \quad i \leq k$$

Por lo tanto, para x perpendicular a los primeros k autovectores e_i , el lado izquierdo de la desigualdad (2.4) se convierte en

$$\frac{x^t B x}{x^t x} = \frac{\sum_{k+1}^p \lambda_i y_i^2}{\sum_{k+1}^p y_i^2}$$

Tomando $y_{k+1} = 1, y_{k+2} = \dots = y_p = 0$ se obtiene el máximo afirmado ■

Teorema 2.1. Sea Σ la matriz de covarianza asociada al vector aleatorio $X' = [X_1, X_2, \dots, X_p]$. Sean $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$ los pares autovalores-autovectores de Σ , tal que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. Entonces los i -ésimos componentes principales están dado por

$$Y_i = e_i^t X = e_{i1} X_1 + e_{i2} X_2 + \dots + e_{ip} X_p, \quad i = 1, 2, \dots, p$$

Con esta elección,

$$Var(Y_i) = e_i^t \Sigma e_i = \lambda_i \quad i = 1, 2, \dots, p$$

$$Cov(Y_i, Y_k) = e_i^t \Sigma e_k = 0 \quad i \neq k$$

Si algunos λ_i coinciden, le corresponde el vector de coeficiente e_i , y por lo tanto los Y_i no son únicos.

Demostración. Si $B = \Sigma$, Por (2.2) se tiene que,

$$\max_{a \neq 0} \frac{a^t B a}{a^t a} = \lambda_1 \quad (\text{alcanzado cuando } a = e_1)$$

Pero $e_1^t e_1 = 1$ ya que los autovectores están normalizados. Así,

$$\max_{a \neq 0} \frac{a^t B a}{a^t a} = \lambda_1 = \frac{e_1^t \Sigma e_1}{e_1^t e_1} = e_1^t \Sigma e_1 = Var(Y_1)$$

Similarmente, usando (2.4), se obtiene

$$\max_{a \perp e_1, \dots, e_k} \frac{a^t \Sigma a}{a^t a} = \lambda_{k+1} \quad k = 1, 2, \dots, p-1$$

Eligiendo $a = e_{k+1}$, con $e_{k+1}^t e_i = 0$, para $i = 1, 2, \dots, k$ y $k = 1, 2, \dots, p-1$

$$\frac{e_{k+1}^t \Sigma e_{k+1}}{e_{k+1}^t e_{k+1}} = e_{k+1}^t \Sigma e_{k+1} = \text{Var}(Y_{k+1})$$

Pero $e_{k+1}^t (\Sigma e_{k+1}) = \lambda_{k+1} e_{k+1}^t e_{k+1} = \lambda_{k+1}$, así $\text{Var}(Y_{k+1}) = \lambda_{k+1}$.

Queda por demostrar que si e_i perpendicular a e_k se obtiene $\text{Cov}(Y_i, Y_k) = 0$. Ahora, los autovectores de Σ son ortogonales si todos los autovalores $\lambda_1, \lambda_2, \dots, \lambda_p$ son distintos. Si todos los autovalores no son distintos, los autovectores correspondientes a los autovalores comunes pueden ser elegidos de manera que sean ortogonales. Por lo tanto, para cualesquiera dos autovectores e_i y e_k , $e_i^t e_k = 0, i \neq k$. Como $\Sigma e_k = \lambda_k e_k$, se premultiplica por e_i^t y se obtiene

$$\text{Cov}(Y_i, Y_k) = e_i^t \Sigma e_k = e_i^t \lambda_k e_k = \lambda_k e_i^t e_k = 0$$

para todo $i \neq k$. Así queda demostrado ■

De esta manera:

- Primer componente principal = combinación lineal ($a_1^t X$) que maximiza

$$\text{Var}(a_1^t X) \text{ tal que } a_1^t a_1 = 1.$$

- Segundo componente principal = combinación lineal ($a_2^t X$) que maximiza

$$\text{Var}(a_2^t X) \text{ tal que } a_2^t a_2 = 1 \text{ y } \text{Cov}(a_1^t X, a_2^t X) = 0$$

⋮

- El i-ésimo componente principal = combinación lineal ($a_i^t X$) que maximiza

$$\text{Var}(a_i^t X) \text{ tal que } a_i^t a_i = 1 \text{ y}$$

$$\text{Cov}(a_i^t X, a_k^t X) = 0 \text{ para } k < i$$

Como en el caso de las bandas, la matriz de varianza-covarianza de los componentes (Σ_Y) es simétrica (por ser $\text{Cov}(Y_i, Y_k) = \text{Cov}(Y_k, Y_i)$, para cada $i \neq k$). Además como $\text{Cov}(Y_i, Y_k) = 0$ (para cada $i \neq k$), Σ_Y ésta es diagonal.

A partir de la demostración del Teorema 2.1 (Johnson, y otros, 1992) se tiene que

$$Y_i = e_i^t X = e_{i1} X_1 + e_{i2} X_2 + \dots + e_{ip} X_p, \quad i = 1, 2, \dots, p$$

donde los λ_i son los autovalores (con $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$) y los e_i son los autovectores de Σ_X .

Así,

$$Var(Y_i) = Var(e_i^t X) = e_i^t \Sigma_X e_i = \lambda_i \quad i = 1, 2, \dots, p$$

$$Cov(Y_i, Y_k) = e_i^t \Sigma_X e_k = 0 \quad i \neq k$$

Es decir,

$$\Sigma_Y = \begin{bmatrix} \lambda_1 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_p \end{bmatrix}$$

Si algunos autovalores son iguales, los autovectores también y por lo tanto los componentes no son únicos. Sin embargo, los autovectores correspondientes a autovalores iguales pueden elegirse de manera tal que sean ortogonales, y así los componentes son únicos. La matriz Σ_Y contiene en la diagonal las varianzas (autovalores) de los niveles digitales de los píxeles en las coordenadas transformadas. Los autovalores son no crecientes, entonces la máxima varianza (en general) está en la primera componente Y_1 , la subsiguiente en Y_2 , y así sucesivamente.

Geométricamente, las combinaciones lineales representan la selección de un nuevo sistema de coordenadas $Y_1, Y_2, \dots, Y_p$ obtenido por rotación del sistema original con coordenadas $X_1, X_2, \dots, X_p$. Los nuevos ejes representan las direcciones de máxima variabilidad y proveen una descripción más simple y parsimoniosa de la estructura de covarianza. Los autovalores expresan la longitud de cada uno de los ejes (componentes), mientras que los autovectores dan la dirección de los mismos.

Por otra parte, cada autovalor representa la proporción de información que retiene el componente principal asociado (lo cual es útil para decidir qué componentes son más interesantes), en tanto que el autovector indica la ponderación que debe aplicarse a cada una de las bandas para obtener el componente principal (equivalentes a los coeficientes de regresión en una transformación lineal estándar, siendo las bandas de la imagen las variables independientes y los componentes principales, las dependientes). El valor absoluto del elemento e_{ji} del autovector e_i indica el grado de contribución de la banda j al componente principal i definido por la transformación lineal (Chuvieco Salinero, 1996).

En la figura 2.10 se muestra la aplicación del método de análisis de componentes principales a una subimagen del sensor Operational Land Imager (OLI) abordo del satélite

Landsat 8, con 7 bandas espectrales. La imagen fue adquirida en marzo de 2014 y representa las adyacencias del Embalse Guárico en Calabozo, estado Guárico, Venezuela, que sirve como sistema de riego para una extensa zona arrocera, conocido como el Sistema de Riego del Río Guárico.

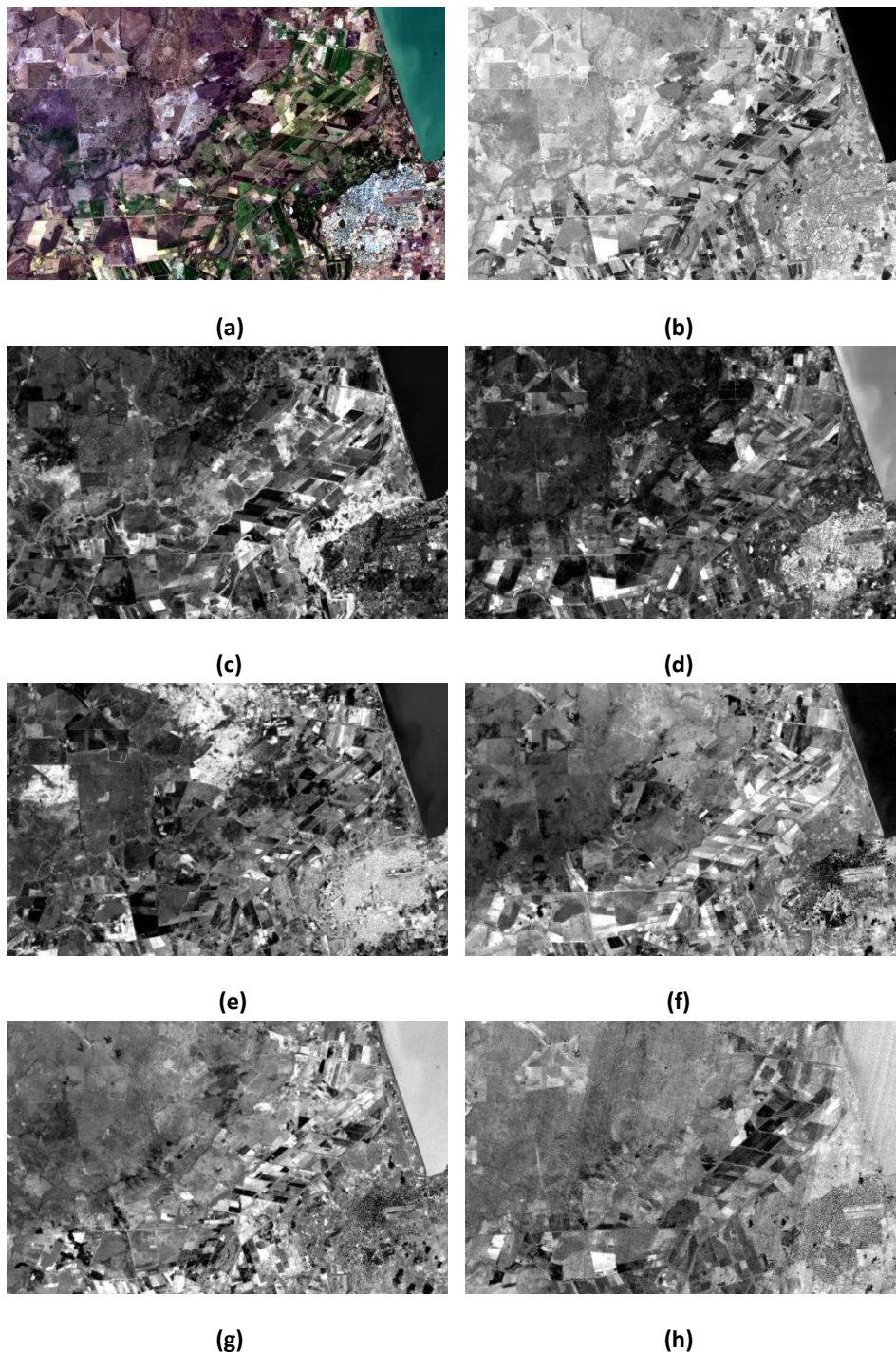


Figura 2.10. Análisis de componentes principales. (a) Imagen original visualizada según la combinación RGB, (b)-(h) Componentes principales 1°-7° respectivamente.

Por consiguiente se calcula su matriz de covarianza:

$$\sum_x = \begin{bmatrix} 140554 & 182351 & 244328 & 391117 & 224844 & 722354 & 574657 \\ 182351 & 244249 & 336033 & 543097 & 322317 & 974787 & 762336 \\ 244328 & 336033 & 514418 & 791952 & 588771 & 1249651 & 959681 \\ 391117 & 543097 & 791952 & 1389762 & 941703 & 2543118 & 1891171 \\ 224844 & 322317 & 58771 & 941703 & 6031084 & 5169019 & 2529554 \\ 722354 & 974787 & 1249651 & 2543118 & 5169019 & 10558166 & 6738325 \\ 574657 & 762336 & 959681 & 1891171 & 2529554 & 6738325 & 7800987 \end{bmatrix}$$

Así $\sum_x$ es una matriz simétrica de tamaño 7x7 donde los elementos de la diagonal son las varianzas de los niveles digitales en cada banda. En la tabla 3 aparecen los autovectores obtenidos, que representan los coeficientes de ponderación que han de aplicarse sobre los valores digitales de las bandas originales (filas) para obtener cada uno de los componentes resultantes (columnas) como combinación lineal ponderada de aquellas.

Bandas	Componentes Principales						
	CP1	CP2	CP3	CP4	CP5	CP6	CP7
B1	0,0544	-0,0768	0,2043	0,0709	-0,4941	-0,4883	-0,6796
B2	0,0737	-0,0986	0,2901	0,0303	-0,4781	-0,3781	0,7268
B3	0,0995	-0,0905	0,5068	6,4973	-0,3880	0,7528	-0,0883
B4	0,1907	-0,2143	0,6829	-0,2630	0,5765	-0,2188	-0,0446
B5	0,4168	0,8728	0,1740	0,17442	0,0437	-0,0441	0,0052
B6	0,7405	-0,1699	-0,3398	-0,5360	-0,1359	0,0394	-0,0091
B7	0,4726	-0,3737	-0,0721	0,7792	0,1549	0,0222	0,0112

Tabla 3. Autovectores de la matriz de covarianzas de la imagen original.

La interpretación de los nuevos componentes resulta complicada. En este caso, a la vista de los coeficientes mencionados (Tabla 3), observamos cómo para la obtención del primer componente se ponderan de forma más o menos similar las bandas originales, por lo que la imagen resultante contendrá principalmente información sobre la intensidad o brillo global de la imagen. Análogamente, en el segundo componente se aprecia una ponderación positiva muy superior para las bandas 5,6 y 7 (correspondientes a las bandas del infrarrojo) que para el resto, teniendo en cuenta que estas bandas están directamente relacionadas con un incremento acusado a la reflectancia de la hoja vegetal debido a su estructura morfológica interna, parece lógico asociar el segundo componente al vigor vegetativo y a la presencia de vegetación, lo cual se corrobora examinando la Figura 2.10(c), en las que las zonas de regadío aparecen con más brillo que el resto.

La tabla 4 muestra los autovalores asociados a cada nuevo componente. Dividiendo el autovalor de cada componente entre la suma de todos ellos se obtiene la proporción de varianza total contenida en los datos originales expresada por el componente en cuestión, lo cual proporciona un criterio de selección del número de componentes para utilizar en cualquier procedimiento de análisis posterior.

Componente	Autovalores	% Varianza	% Varianza Acumulada
1	24268367	80.5639	80.5639
2	4431841	14.7124	95.2763
3	1106198	3.6723	98.9485
4	219375	0.7283	99.6768
5	75218	0.2497	99.9265
6	20058	0.0666	99.9931
7	2078	0.0069	100.0000

Tabla 4. Autovalores asociados y porcentajes de varianza acumulada.

1.9.1.1 Aplicaciones del Análisis de Componentes Principales

A continuación se describen las principales aplicaciones del método de análisis de componentes principales en teledetección.

Reducción del volumen de datos

La principal utilidad del método consiste, en la reducción de la dimensionalidad de los datos, eliminando así la información redundante debida a la correlación entre bandas. Una vez seleccionados los componentes que conlleven la mayoría de la información pueden realizarse las mismas operaciones de análisis de datos (realce, clasificación) que se aplicaban a la totalidad de las bandas, consiguiendo resultados similares. Esta aplicación tiene un enorme potencial cuando se trabaja con un número elevado de bandas espectrales tomadas a intervalos de longitud de onda muy cortos, lo que ocasiona una gran correlación entre bandas, haciéndose necesario un proceso estadístico de selección.

Separación del ruido

Los valores digitales de las imágenes obtenidas mediante sensores a bordo de satélites se ven especialmente alterados debido tanto a la influencia de los fenómenos de dispersión atmosférica como a deficiencias, a veces muy sutiles, en el funcionamiento de los propios sensores. Al conjunto de alteraciones de la señal que se manifiesta en las imágenes se le denomina ruido.

Los últimos componentes obtenidos en el análisis de componentes principales, aquellos que contienen la mínima información de los datos, suelen ir asociados a ruido del sistema o a las alteraciones de la imagen debidas a la dispersión atmosférica, por lo que su eliminación contribuye directamente a la reducción del ruido, esto es, a la restauración y realce de las imágenes.

Detección de Cambios

La utilización de este método para la detección de cambios en el tiempo requiere la integración de dos o más imágenes multispectrales, de una misma zona y adquiridas en fechas distintas, en una misma matriz o archivo de datos sobre el cual se aplica el análisis de componentes principales. Suponiendo que las zonas modificadas sean cuantitativamente menos importantes, esto es, ocupen menos superficie que las zonas que no han evolucionado, los primeros componentes suelen retener la información relativa a las áreas comunes en todas las bandas, mientras que los componentes secundarios registrarán los cambios o evoluciones en el tiempo (características u objetos que sólo aparecen en una de las fechas).

Ampliación del contraste por decorrelación

Otra aplicación de este método al realce de imágenes se basa en la ampliación del contraste por decorrelación. Consiste en la conversión de los datos al espacio definido por los componentes principales de las bandas originales (este espacio se caracteriza por la inexistencia de correlación entre los componentes), seguido de la ecualización de los datos según los nuevos ejes, y por último la conversión de los datos al espacio inicial y la combinación de las bandas con los colores primarios RGB, por lo que la imagen mostrará un contraste muy superior.



Figura 2.11. Contraste por decorrelación. Componentes Principales [1 2 3].

La figura 2.11 representa una correlación entre las primeras tres componentes principales. Como se aprecia, la imagen adquiere unos colores un tanto exagerados, pocos reales. Sin embargo, éstos no se reasignan al azar, sino que mantienen su tonalidad inicial (la vegetación sigue siendo verde, el agua azul etc.) lo cual permite la asociación de cada uno de ellos a las mismas características físicas de la escena inicial. Este método constituye una interesante herramienta de apoyo a la interpretación, especialmente en ciertos tipos de aplicaciones en las que se pretende localizar e identificar objetos poco evidentes.

2.9.2 Realce de Imágenes

Una de las principales técnicas que componen el procesamiento digital de imágenes es la de realce o mejora de imágenes. Esta técnica comprende un conjunto de operaciones tendentes a mejorar la calidad visual de una imagen. Permiten realzar las características de brillo y contraste de una imagen, reducir su contenido de ruido, o agudizar o intensificar detalles presentes en ella. En esta técnica, tal como en otras operaciones de procesamiento digital de imágenes, interviene una imagen de entrada y una imagen de salida. La primera constituye la imagen cuyos datos serán procesados, es decir, sometidos al realce, y la segunda es la resultante de tal procesamiento (Aldalur B, 2002).

Las operaciones que componen la técnica de realce pueden dividirse en dos tipos:

- a) Operaciones de procesamiento puntual o píxel por píxel
- b) Operaciones de procesamiento por grupo de píxeles o sobre vecindades.

Las operaciones de primer tipo se conocen como técnicas de realce radiométrico, tienden a mejorar el contraste tonal en la imagen, es decir, la diferencia entre los valores más oscuros y más claros que se visualizan en un monitor. El nivel de gris de cada píxel en la imagen de entrada es modificado por un nuevo valor mediante operaciones matemáticas o relaciones lógicas. El valor resultante es ubicado en la imagen de salida en la misma posición espacial, es decir, en la misma posición (x, y) del arreglo rectangular de píxeles correspondiente a la imagen de entrada. Esto se realiza píxel por píxel, de manera individual, y los píxeles vecinos no tienen influencia. Ejemplo de estas operaciones son los filtros morfológicos que se abordan en el capítulo siguiente.

Las operaciones de segundo tipo o técnicas de realce geométrico, mejoran el contraste espacial en la imagen, es decir, la diferencia entre el valor digital de brillo de un determinado píxel y el de sus vecinos. Pretenden suavizar o reforzar estos contrastes espaciales de forma tal que, los valores de brillo en cada píxel de la imagen se asemejen o diferencien más de los correspondientes a los píxeles que los rodean. Este procesamiento opera sobre un grupo de píxeles de entrada que circundan a un píxel central. Los píxeles vecinos proveen información valiosa sobre las tendencias del brillo en el área bajo procesamiento (Richards, 1995). En la Figura 2.12 se muestra un cambio de forma del histograma reasignando los valores de los píxeles.

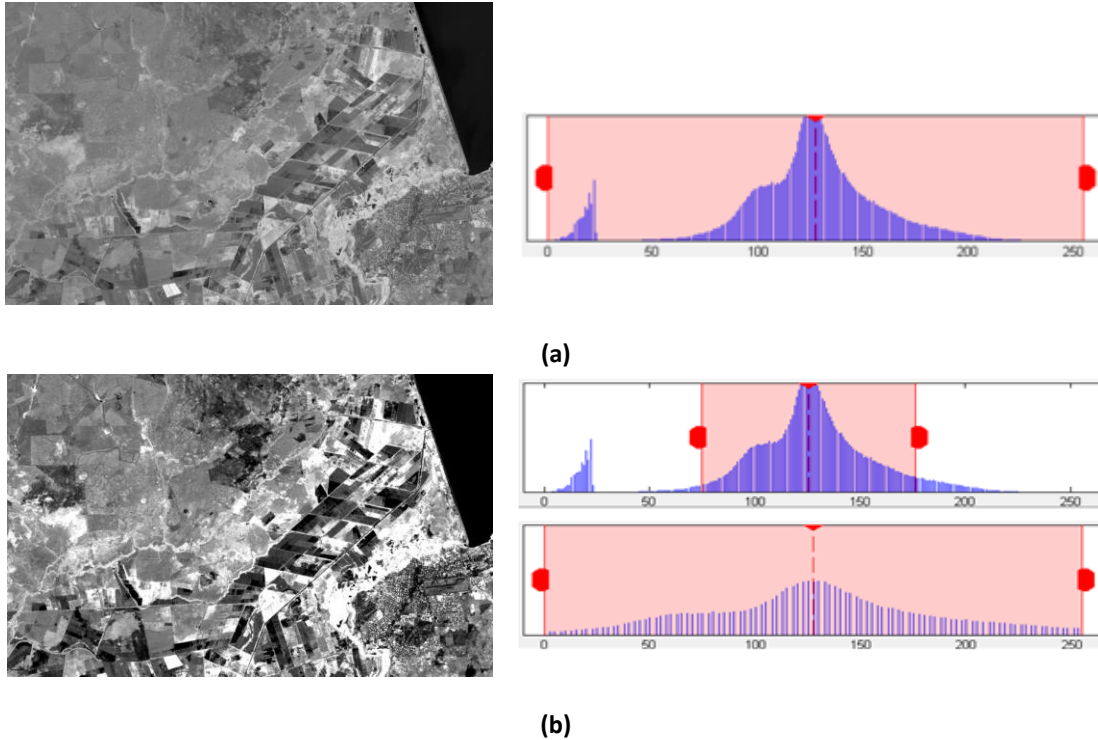


Figura 2.12. Realce de imagen. (a) Imagen e histograma original, (b) Imagen e histograma ajustados.

La figura 2.12 es un ejemplo de un estiramiento de contraste lineal, donde el rango de valores iniciales de los píxeles (75 a 176) es expandido hasta obtener el rango máximo de 0 a 255. Al expandir la escala de grises se logra realzar detalles sutiles de la imagen, difíciles de discriminar con la compresión inicial.

Puesto que estas operaciones permiten mejorar la calidad visual de una imagen, su aplicación como paso previo a otras técnicas de procesamiento digital es necesaria para explorar y detectar las distintas características presentes en la imagen.

Capítulo 3. Morfología Matemática

La base conceptual de la Morfología Matemática descansa en la "teoría de conjuntos" cuyos primeros trabajos se deben a Minkowsky (Minkowsky, 1897), (Minkowsky, 1901) y, Hadwiger, (Hadwiger, 1957), (Hadwiger, 1959). La continuación de estos trabajos de investigación, bajo el impulso y reformulación de Matheron y Serra, se darían posteriormente a conocer bajo la denominación de Morfología Matemática, como una técnica no lineal de tratamiento de señales.

Actualmente, el ámbito y alcance de los procesamiento morfológicos es tan amplio como el propio procesamiento de imágenes. Se pueden encontrar aplicaciones tales como la segmentación, restauración, detección de bordes, aumento de contraste, análisis de texturas, compresión, etc. En este capítulo se describen nociones básicas de la Morfología Matemática que serán útiles de repasar, ya que se hará referencia a ellas en el procesamiento para la segmentación. Las obras (Matheron, 1967), (Matheron, 1975), (Serra, 1981) y (Serra, 1988), contienen en detalle la definición de los conceptos básicos en morfología a tratar a continuación.

3.1 Nociones sobre teoría de conjuntos

Definición 3.1. Dos conjuntos son iguales si están formados por los mismos elementos:

$$X = Y \Leftrightarrow (p \in X \Rightarrow p \in Y \text{ y } p \in Y \Rightarrow p \in X)$$

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

Definición 3.2. X es subconjunto de Y si todos los elementos de X pertenecen a Y :

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

La inclusión es reflexiva, antisimétrica y transitiva.

Definición 3.3. La intersección de dos conjuntos X e Y es el conjunto de los elementos que pertenecen a ambos conjuntos:

$$X \cap Y = \{p \mid p \in X \text{ y } p \in Y\}$$

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

Definición 3.4. La unión de dos conjuntos se constituye por los elementos que pertenecen a uno o al otro:

$$X \cup Y = \{p \mid p \in X \text{ o } p \in Y\}$$

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

Definición 3.5. La diferencia entre conjuntos X e Y , la componen los elementos que pertenecen a X pero que no están incluidos en Y :

$$X \setminus Y = \{p \mid p \in X \text{ y } p \notin Y\}$$

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 u a otro conjunto, pero nunca a ambos. La diferencia simétrica es conmutativa y asociativa.

Definición 3.6. El complemento de un subconjunto X , perteneciente a un conjunto Y (conjunto de referencia) se define como:

$$X^c = \{p \mid p \notin X \text{ y } p \in Y\}$$

Definición 3.7. Sean $X, Y \subset R^2$ para todo elemento $x \in X$ e $y \in Y$, es posible hacer corresponder una suma algebraica $x + y$. De esta manera se forma un nuevo conjunto denominado adición de Minkowski y denotado por $X \oplus Y$:

$$X \oplus Y = \{x + y \mid x \in X, y \in Y\}$$

Definición 3.8. 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).
- $x \leq y, y \leq x$, implica que $x = y$ (antisimétrica).
- $x \leq y, y \leq z$, 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 un orden parcial y se denotará como $(X, \leq)$. El conjunto será totalmente ordenado si todos los elementos que lo componen son comparables, es decir: $x \leq y$ ó $y \leq x$, para cualquier par $(x, y) \in X$.

Definición 3.9. 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 $y \in X$, es cota superior de S si $x \leq y$, para todo $x \in S$.

- Un elemento $x \in X$, es ínfimo de S si y sólo 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 supremo de S si y sólo 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 3.10. Un conjunto ordenado $(X, \leq)$ es un retículo completo si todos los subconjuntos de X poseen un ínfimo y un supremo.

Definición 3.11. Sean X e Y dos retículos completos. La relación f es una anamorfosis si y sólo si f es una biyección que conserva el ínfimo y supremo:

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

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

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:

$$x \leq y \Rightarrow \psi(x) \leq \psi(y).$$

3.1.1 Propiedades básicas de las transformaciones morfológicas

Un elemento estructurante (ES) es una forma geométrica simple (círculo, rectángulo, línea, puntos, etc.), que se utiliza como elemento de referencia, está definido por una ventana deslizante, con la cual se compara cada objeto de la imagen, permitiendo discriminar aquellos que poseen propiedades geométricas comunes.

Toda operación morfológica es el resultado de una o más operaciones de conjuntos (unión, intersección, complementación, etc.) haciendo intervenir dos subconjuntos X, Y , de un conjunto espacio Z . El conjunto X es la imagen, que puede ser binaria o en escala de grises y el subconjunto Y , recibe el nombre de *elemento estructurante* (ES) que para operar con X , se desplazará a través del espacio Z .

El ES consiste en un patrón especificado, como las coordenadas de un número de puntos discretos en relación con algún origen. Normalmente se usan coordenadas cartesianas, así una manera conveniente de representar el elemento es como una matriz. La Figura 3.1 muestra un número diferente de ES de diversos tamaños. En cada caso el origen está marcado por un anillo alrededor de ese punto.

Un punto importante a tener en cuenta es que, aunque una matriz se utiliza para representar el elemento estructurante, no todos los puntos de la rejilla forman parte del ES, en la figura 3.1 son representados como ceros.

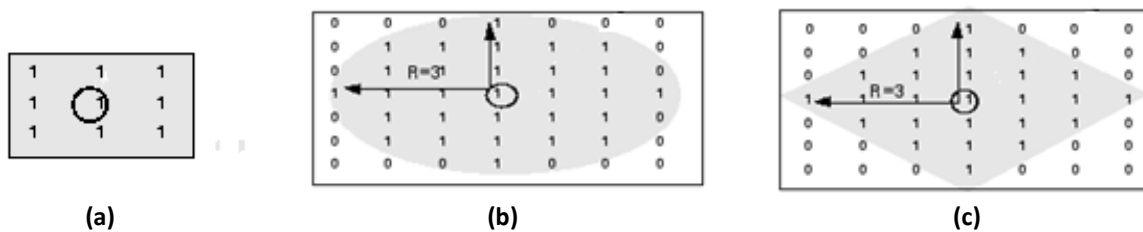


Figura 3.1. Elementos Estructurantes. (a) Cuadrado 3x3, (b) Disco 7x7 radio 3 , (c) Diamante 7x7 radio 3.

También debe considerarse que, el origen no necesariamente tiene que estar en el centro. Los ES que tienen un valor constante se denominan planos, convenientemente representado como uno. Los que tienen diferentes valores son denominados de volumen. En la figura 3.2 se ilustra un ES de volumen unidimensional donde el origen del elemento se sitúa en su centro y posee valor 3, frente a los valores 1 y 2 presentes en los demás puntos. Los ES de volumen modifican el efecto de las operaciones morfológicas, respecto a un elemento estructurante plano.

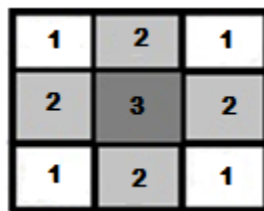


Figura 3.2. Elemento estructurante de volumen de tamaño 3x3.

Cuando se lleva a cabo una operación morfológica, se desplaza el origen del ES a cada posición del píxel de la imagen, a su vez, los puntos dentro del elemento estructurante se comparan con los valores del píxel imagen subyacente. Los detalles de esta comparación, y el efecto de los resultados dependen de qué operador morfológico está siendo utilizado.

El ES se elige de manera que sus características geométricas (forma, tamaño, orientación, etc.) se identifiquen con las características morfológicas de los objetos presentes en la imagen. Ejemplos básicos de ES utilizados en los resultados se muestran como imagen en la figura 3.3.

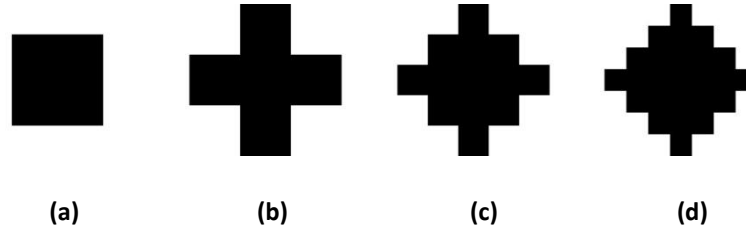


Figura 3.3. Formas básicas de elementos Estructurantes.
(a) Cuadrado Radio 3, (b) Diamante Radio 1, (c) Diamante Radio 2, (d) Diamante Radio 3.

El objetivo de las transformaciones morfológicas es la extracción de estructuras geométricas en los conjuntos sobre los que se opera, mediante la utilización del ES.

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

Invariabilidad a traslación

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

donde p es el factor de traslación 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 escala:

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

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 :

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

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 Transformaciones morfológicas elementales

3.2.1 Erosión

En teoría de retículos, una erosión es una operación que conmuta con el ínfimo. Dado un retículo completo X , una erosión es una función $\varepsilon : X \rightarrow X$ en la que:

$$\varepsilon(\bigwedge_{i \in I} x_i) = \bigwedge_{i \in I} \varepsilon(x_i)$$

donde I es cualquier conjunto de índices y $\{x_i\}$ es una colección arbitraria de valores $x_i \in X$.

La transformación de erosión es el resultado de comprobar si el elemento estructurante Y está totalmente incluido dentro del conjunto X . Cuando esto no ocurre, el resultado de la erosión es el conjunto vacío.

La erosión de un conjunto X por un elemento estructurante Y se define como el conjunto de puntos o elementos x , pertenecientes a X , de forma que cuando el elemento estructurante Y se traslada a ese punto, el elemento queda incluido en X :

$$\varepsilon_Y(X) = \{x / Y_x \subseteq X\}$$

La ecuación anterior puede reformularse en términos de una intersección de conjuntos trasladados. Las traslaciones vienen determinadas por el elemento estructurante Y :

$$\varepsilon_Y(X) = \bigcap_{s \in Y} X_{-s}$$

El efecto de una operación de erosión puede observarse en la figura 3.4, en la que un elemento estructurante Y , en forma de disco circular, hace desaparecer las estructuras de menor tamaño al elemento.



Figura 3.4. Erosión de X por el elemento estructurante Y .
Los elementos conectados del conjunto X más pequeños que Y son eliminados.

La última definición de erosión puede extenderse directamente al caso de las imágenes. La erosión de una imagen f por un elemento estructurante Y se denota por $\varepsilon_Y(f)$ y se define como el mínimo ($\wedge$) de las traslaciones de f por los elementos s de Y :

$$\varepsilon_Y(f) = \bigwedge_{s \in Y} (f_{-s})$$

A la imagen en escala de grises correspondiente a la banda dos de una imagen del sensor MSS-2 del satélite Francisco de Miranda, adquirida en marzo de 2013 donde se encuentra parte del Lago de Valencia, Venezuela, se aplica la erosión por un elemento estructurante cuadrado de tamaño 6x6 y se ilustra en la figura 3.5(b). Es posible observar como predominan los objetos oscuros en aquellas zonas donde el elemento estructurante está contenido, mientras que las zonas de menor tamaño al elemento desaparecen. En las imágenes la noción de ínfimo pasa a denotarse mínimo, en este caso, la erosión es el mínimo valor de la función imagen en la ventana definida por el elemento estructurante cuando su origen se sitúa en (x, y) .

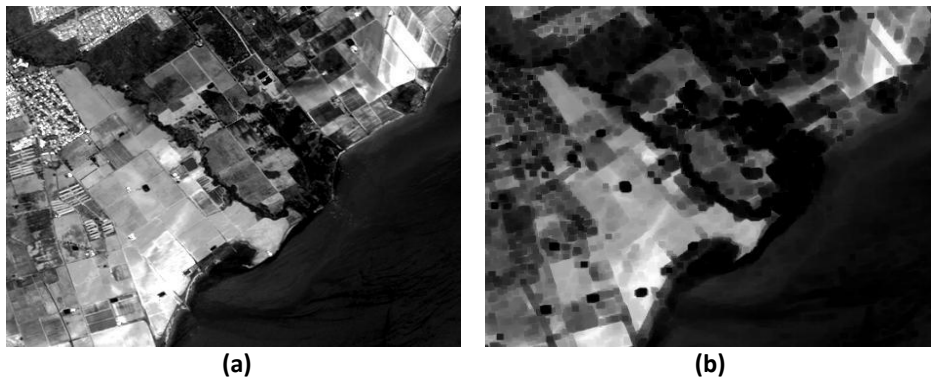


Figura 3.5. (a) Canal 2 de la Imagen original, (b) Erosión de la imagen.

Para imágenes bidimensionales la erosión puede ser definida como:

$$\varepsilon_Y(f)(x, y) = \min_{(s,t) \in Y} f(x + s, y + t)$$

El resultado de la erosión es una imagen de menor valor, es decir, una imagen más oscura, puesto que la erosión pretende minimizar el valor de cada nivel digital. La erosión de una imagen en escala de grises por un elemento estructurante de volumen se define como sigue:

$$\varepsilon_Y(f)(x) = \min_{s \in Y} \{f(x + s) - Y(s)\}$$

La operación de erosión, junto a la dilatación que será definida a continuación, son la base de cualquier transformación morfológica. Cualquier operador, transformación o algoritmo incluirá una erosión, una dilatación o ambas primitivas en su implementación. Sin necesidad de formar nuevos operadores es posible encontrar aplicaciones interesantes en las transformaciones básicas.

3.2.2 Dilatación

En teoría de retículos, un operador $\delta : X \rightarrow X$ se denomina dilatación en el caso que conmute con el supremo de una colección de valores.

$$\delta(\vee_{i \in I} x_i) = \vee_{i \in I} \delta(x_i)$$

donde I es cualquier conjunto de índices y $\{x_i\}$ es una colección arbitraria de valores $x_i \in X$.

El resultado de la dilatación es el conjunto de puntos origen del elemento estructurante Y tales que el elemento estructurante contiene algún elemento del conjunto X , cuando el elemento se desplaza por el espacio que contiene a ambos conjuntos:

$$\delta_Y(X) = \{x / Y_x \cap X \neq \emptyset\}$$

Esta última ecuación puede describirse como una unión de conjuntos trasladados. Las traslaciones vienen definidas por el dominio del elemento estructurante:

$$\delta_Y(X) = \bigcup_{s \in Y} X_{-s}$$

El efecto de una operación de dilatación puede observarse en la figura 3.6, en donde un elemento estructurante Y de forma de disco circular aumenta la definición del objeto X .



Figura 3.6. Dilatación de X por el elemento estructurante Y . El conjunto X aumenta su definición.

Se puede extender la última definición de dilatación a imágenes f , interpretando la dilatación como el máximo valor de las traslaciones de f (definidas por la forma del elemento estructurante) en cada punto de la imagen:

$$\delta_Y(f) = \bigvee_{s \in Y} (f_{-s})$$

En otras palabras, el valor de dilatación de un píxel (x, y) es el máximo valor de la imagen en la ventana de vecindad definida por el elemento estructurante cuando su origen se sitúa en (x, y) :

$$\delta_Y(f)(x, y) = \max_{(s, t) \in Y} f(x - s, y - t)$$

La dilatación de la imagen binaria de la figura 3.7(a) por un elemento estructurante cuadrado de tamaño 6x6 se ilustra en la figura 3.7(b). En la imagen, los objetos aumentan su definición.

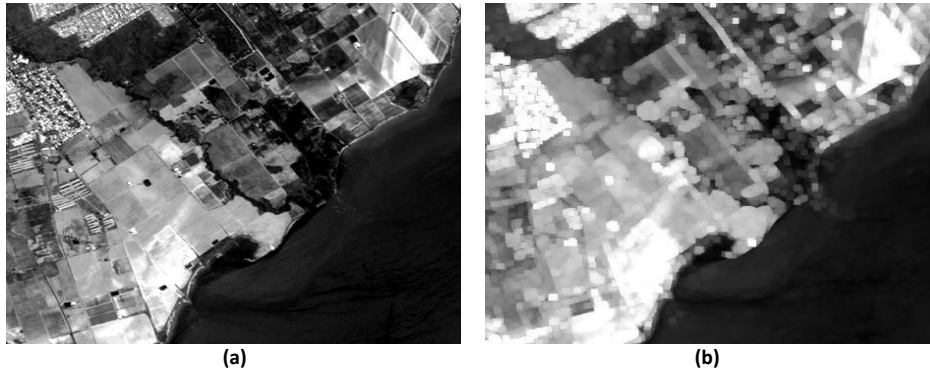


Figura 3.7. (a) Canal 2 de la imagen original, (b) Dilatación de la imagen.

El resultado de la dilatación imágenes es, generalmente, una imagen de mayor valor, es decir, una imagen más clara, puesto que la dilatación maximiza el valor de los ND. El efecto es una intensificación de la claridad de la imagen de acuerdo con el valor que posean los puntos del elemento estructurante:

$$\delta_Y(f)(x, y) = \max_{(s, t) \in Y} \{f(x - s, y - t) + Y(s, t)\}$$

3.2.3 Propiedades de las operaciones básicas de erosión y dilatación

Las operaciones morfológicas de erosión y dilatación cumplen con las propiedades de base de los operadores morfológicos definidas en el apartado 3.1.1. Además, presentan otra serie de particularidades:

Dualidad

La erosión y la dilatación son operaciones duales con respecto a la complementación. Esto quiere decir que, la erosión de una imagen f por medio de Y es la dilatación de su complemento f^c , usando el simétrico del elemento estructurante Y , denotado por $\hat{Y} = \{-y \mid y \in Y\}$, es decir:

$$\varepsilon_Y(f) = (\delta_{\hat{Y}}(f^c))^c$$

Crecientes

Las operaciones de erosión y dilatación son crecientes, respetan el orden presente en la estructura de retículo. Para dos imágenes f y g :

$$\begin{aligned} Si f \leq g &\Rightarrow \varepsilon(f) \leq \varepsilon(g) \\ Si f \leq g &\Rightarrow \delta(f) \leq \delta(g) \end{aligned}$$

Gracias a esta propiedad, las operaciones morfológicas pueden extenderse del caso binario a las imágenes en escala de grises, puesto que estas últimas pueden expresarse como una suma de imágenes binarias (una superposición de conjuntos apilados). Cada una de las imágenes binarias corresponde al conjunto de píxeles cuyo nivel de gris es superior a un umbral determinado. La imagen de grises se define como:

$$f = \sum_{t=1}^{t \max} CS_t(f)$$

Donde $CS_t(f)$ es la imagen binaria que contiene los píxeles con valor en f mayor o igual a t .

El que la erosión y la dilatación sean crecientes implica que las relaciones de orden entre los diferentes conjuntos apilados de una imagen sean respetadas por estos operadores. La erosión y la dilatación son, pues, invariantes a la descomposición por umbrales:

$$\begin{aligned} \varepsilon_Y &= \sum_{t=1}^{t \max} \varepsilon_Y(CS_t(f)) \\ \delta_Y &= \sum_{t=1}^{t \max} \delta_Y(CS_t(f)) \end{aligned}$$

Extensividad y antiextensividad

La operación de dilatación es una operación extensiva. Para la imagen f ,

$$f \leq \delta(f).$$

La erosión, por el contrario, es una operación antiextensiva:

$$\varepsilon(f) \leq f.$$

En general, para elementos estructurantes que contienen su origen se cumple que:

$$\varepsilon(f) \leq f \leq \delta(g)$$

3.3 Gradiente Morfológico

Un gradiente, es un operador basado en la primera derivada de una función continua que proporciona las variaciones locales con respecto a la variable, de forma que el valor de la derivada es mayor cuanto más rápida son estas variaciones. En el caso de funciones bidimensionales $f(x, y)$, la derivada es un vector que apunta en la dirección de la máxima variación de $f(x, y)$ y cuyo módulo es proporcional a dicha variación. Las componentes de éste miden la rapidez en que los valores de los píxeles cambian en la distancia y en las direcciones X y Y . Este vector se denomina gradiente y se define:

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{bmatrix}$$

En el caso bidimensional discreto, las distintas aproximaciones del operador gradiente se basan en diferencias entre los niveles de grises de la imagen. La derivada parcial $f_x(x, y)$ (gradiente de la fila $G_F(i, j)$) puede aproximarse por la diferencia de píxeles adyacentes de la misma fila.

$$\frac{\partial f(x, y)}{\partial x} \approx \nabla_x f(x, y) = f(x, y) - f(x - 1, y)$$

La discretización del vector gradiente en el eje Y ($G_C(i, j)$), será:

$$\frac{\partial f(x, y)}{\partial y} \approx \nabla_y f(x, y) = f(x, y) - f(x, y - 1)$$

Para la implementación y computación del gradiente se utilizan máscaras o filtros que representan o equivalen a dichas ecuaciones. El gradiente de la fila G_F y de la columna G_C en cada punto se obtiene mediante la convolución de la imagen con la máscara H_F y H_C , esto es:

$$\begin{aligned} G_F(i, j) &= F(i, j) \otimes H_F(i, j) \\ G_C(i, j) &= F(i, j) \otimes H_C(i, j) \end{aligned}$$

La magnitud y orientación del vector gradiente suele aproximarse por la expresión:

$$|G(i, j)| = \sqrt{G_F^2 + G_C^2} \approx |G_F(i, j)| + |G_C(i, j)|.$$

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

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

En el caso de funciones:

$$\rho_{\psi\zeta}(X) = \psi(X) - \zeta(X)$$

El Gradiente Morfológico, conocido en morfología como gradiente de Beucher(Beucher, 1977) es la diferencia entre una 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 Y . En el caso de conjuntos:

$$\rho_{\bar{Y}}(X) = X \setminus \varepsilon_Y(X)$$

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

$$\rho_{\bar{Y}}(f) = f - \varepsilon_Y(f)$$

El efecto del gradiente por erosión con un elemento estructurante cuadrado de tamaño 6x6 se ofrece en la figura 3.8(b).

El gradiente por dilatación se define como la diferencia entre una dilatación por elemento estructurante Y y la identidad (imagen original):

$$\rho_{Y^+}(X) = \delta_Y(X) \setminus X$$

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

$$\rho_{Y^+}(f) = \delta_Y(f) - f$$

El efecto de este mismo gradiente con un elemento estructurante cuadrado de tamaño 6x6 se visualiza en la figura 3.8(c).

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

$$\begin{aligned} \rho_Y(X) &= \delta_Y(X) \setminus \varepsilon_Y(X) \\ \rho_Y(f) &= \delta_Y(f) - \varepsilon_Y(f) \end{aligned}$$

Este gradiente acentúa los bordes, como se puede observar en la figura 3.8(d). El grosor de los bordes a obtener viene determinado por el tamaño del elemento estructurante aplicado, por lo que se puede decir que el grosor de los bordes detectados en los gradientes será proporcional al tamaño del ES utilizado.

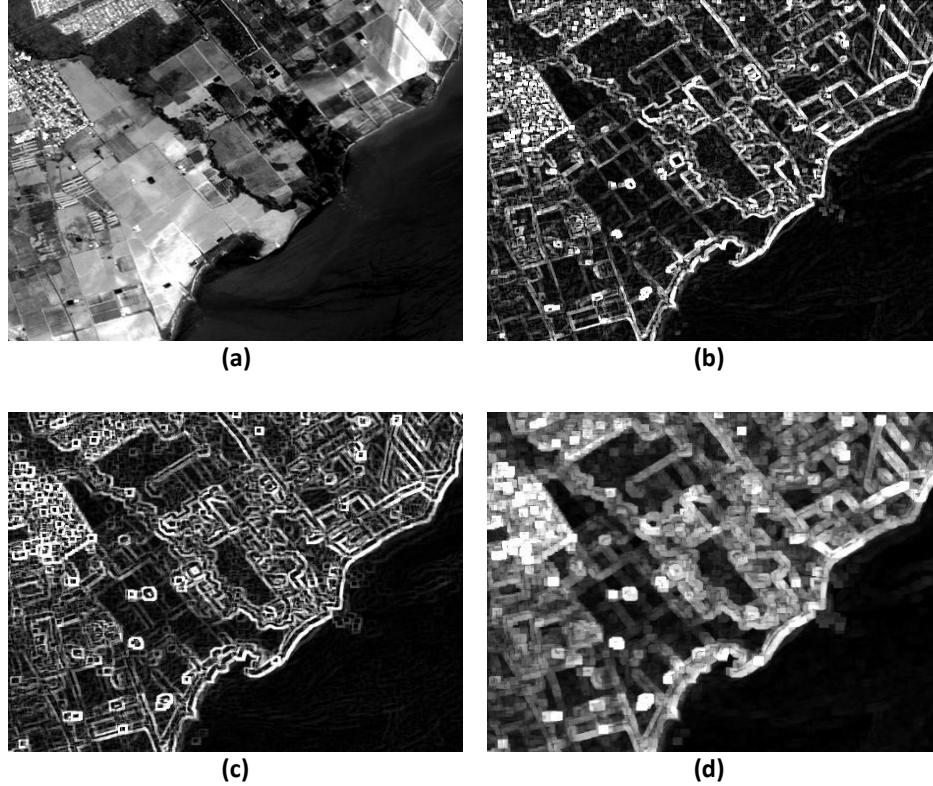


Figura 3.8. Gradientes morfológicos de una imagen, elemento estructurante 6x6.
 (a) Canal 2 de la Imagen original, (b) Gradiente por erosión,
 (c) Gradiente por dilatación, (d) Gradiente simétrico.

3.3.1 Gradiente Multiescala

La variación de los niveles de grises en la imagen muy bien puede ser dada por el gradiente morfológico. Como ya se ha mencionado un gradiente ayuda a la detección de bordes y evita el engrosamiento y la fusión de los mismos. Los bordes gruesos se pueden utilizar para detectar variaciones suaves entre las regiones vecinas. Por otra parte, cuando la distancia que separa dos límites de una región es menor que la anchura del elemento estructurante los bordes resultantes se fusionan. Ambos problemas se pueden evitar con el gradiente multiescala morfológico. La imagen de gradiente $\rho(f)$, se obtiene morfológicamente por la resta de la imagen erosionada, $\varepsilon(f)$ con su versión dilatada $\delta(f)$. El grosor de los bordes obtenidos a partir del gradiente de tamaño n puede ser reducido por una erosión con un elemento estructurante de tamaño $n - 1$. Un gradiente multiescala $\bar{\rho}(f)$, es el promedio de los gradientes morfológicos tomado para diferentes escalas del elemento estructurante Y_i , y se denota:

(3.1)

$$\bar{\rho}(f) = \frac{1}{n} \sum_{i=1}^n \left[\varepsilon_{Y_{i-1}} \left(\delta_{Y_i}(f) \right) - \varepsilon_{Y_i}(f) \right]$$

Donde Y_i es un elemento estructurante de tamaño $(2i + 1) \times (2i + 1)$.

En la figura 3.9 se observa la aplicación del gradiente multiescala al canal 2 de la imagen, con un elemento estructurante en forma de diamante con tamaño inicial 3x3 y final 13x13, es decir, $n = 6$.

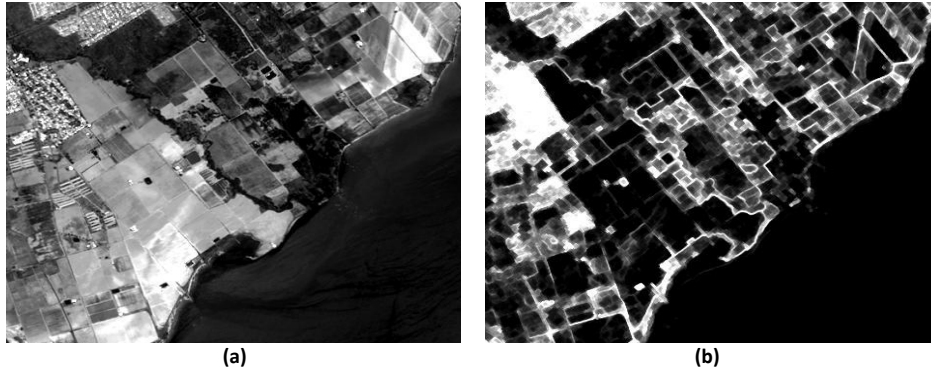


Figura 3.9. Gradiente multiescala, elemento estructurante inicial diamante 3x3.
(a) Canal 2 de la imagen original, (b) gradiente multiescala.

3.4 Apertura y cierre

Generalmente, en un retículo completo X , la dilatación $X \rightarrow \delta(X)$ y la erosión $X \rightarrow \varepsilon(X)$ son operaciones que no admiten inversa, por lo tanto, no hay manera de determinar el origen X desde las imágenes $\delta(X)$ o $\varepsilon(X)$. Sin embargo, es posible, mediante una adjunción de operadores básicos aproximarse a la forma original en base a la dualidad que poseen. Se puede intentar recuperar el elemento original X erosionado si, una vez efectuada la erosión, se realiza una dilatación, adjunción de operaciones conocida como apertura. Lo mismo ocurre si se efectúa una dilatación seguida de una erosión, proceso conocido como cierre.

En este apartado se detallarán las operaciones de apertura y cierre, conocidas en morfología como filtros básicos a partir de los cuales se desarrollan otros más complejos.

3.4.1 Apertura morfológica

La apertura de una imagen f por un elemento estructurante Y se denota por $\gamma_Y(f)$ y se define como la erosión de f por Y , seguida de la dilatación por el mismo elemento estructurante:

$$\gamma_Y(f) = \delta_Y(\varepsilon_Y(f))$$

La apertura de una imagen es independiente del origen del elemento estructurante, puesto que si la erosión se corresponde con una intersección de traslaciones, la dilatación que sigue es una unión de traslaciones en dirección opuesta.

En el caso general de conjuntos se puede definir la apertura como la unión de los elementos estructurantes Y que se encuentran totalmente dentro del conjunto X :

$$\gamma_Y(X) = \cup \{Y \mid Y \subseteq X\}$$

El efecto de una operación de apertura puede observarse en la figura 3.10, en la que un elemento estructurante Y , en forma de disco circular, provoca en la erosión la desaparición de una estructura que, en la operación de dilatación, no se puede recuperar.



Figura 3.10. Apertura morfológica del conjunto X por el elemento estructurante Y .
La apertura redondea las convexidades importantes.

La erosión elimina los elementos menores en tamaño al elemento estructurante. La imagen original no puede recuperarse totalmente. El tamaño y la forma de los elementos estructurantes empleados en la apertura deben ser acordes con las estructuras de la imagen que se desean eliminar. En ocasiones, tamaños elevados de elementos estructurantes eliminarán formas indeseables en una imagen pero afectarán al resto de estructuras. Tamaños reducidos serán óptimos cuando las imágenes contengan pequeños detalles. En la figura 3.11 se observa como la apertura de la imagen de grises, por un elemento estructurante de tamaño 3x3, elimina las formas claras de tamaño menor que el elemento.

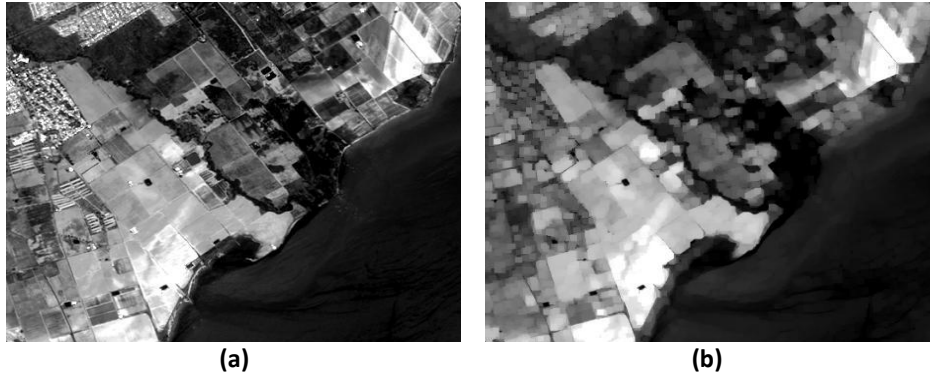


Figura 3.11. Filtrado de imagen. Eliminación de objetos claros mediante apertura morfológica.
(a) Canal 2 de la Imagen original, (b) Apertura morfológica.

3.4.2 Cierre morfológico

El cierre de una imagen f por un elemento estructurante Y se denota por $\varphi_Y(f)$ y se define como la dilatación de f por Y , seguida de la erosión por el mismo elemento estructurante:

$$\varphi_Y(f) = \varepsilon_Y(\delta_Y(f))$$

Por la propiedad de complementación de las operaciones básicas, se puede expresar el cierre como la intersección de todas las traslaciones del complemento del elemento estructurante, tal que éste contiene a X :

$$\varphi_Y(X) = \cap \{Y^c | X \subseteq Y^c\}$$

Al igual que la apertura, el cierre de una imagen es independiente del origen del elemento estructurante. El cierre de un conjunto X por un elemento estructurante con forma de disco se ilustra en la figura 3.12. El cierre es el espacio descrito por el elemento estructurante cuando es forzado a estar fuera de los conjuntos.



Figura 3.12. Apertura morfológica del conjunto X por el elemento estructurante Y .
El cierre redondea las concavidades importantes.

La dilatación une estructuras que la erosión no puede separar. En una imagen de niveles de grises el cierre morfológico ayuda a eliminar estructuras oscuras menores en tamaño al elemento estructurante. La dilatación maximiza los valores de forma que se atenúan los

objetos oscuros. La erosión minimiza la imagen y sólo los elementos no eliminados quedan presentes en la imagen final. En la figura 3.113 se presenta el efecto del cierre morfológico sobre una imagen de grises. Los objetos oscuros quedan sensiblemente atenuados.

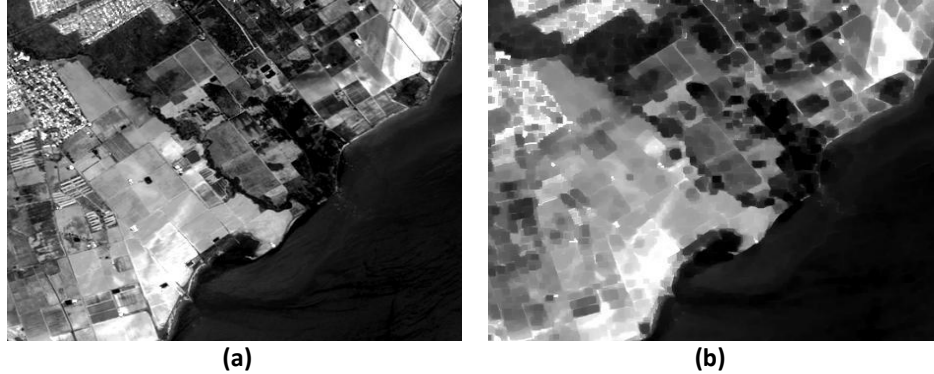


Figura 3.13. Filtrado de imagen. Eliminación de objetos oscuros mediante cierre morfológico.
(a) Canal 2 de la Imagen Original, (b) Cierre morfológico.

3.4.3 Propiedades de la apertura y del cierre morfológico

Dualidad

La apertura de una imagen es equivalente al complemento del cierre de la imagen complementada. Esto significa que la apertura y el cierre son operaciones duales con respecto a la complementación:

$$\gamma_Y(f) = (\varphi_Y(f^c))^c$$

Por la dualidad de la erosión y de la dilatación:

$$(\varphi_Y(f^c))^c = (\varepsilon_Y(\delta_Y(f^c)))^c = (\varepsilon_Y((\varepsilon_Y)^c(f)))^c = \delta_Y(\varepsilon_Y(f)) = \gamma_Y(f)$$

El empleo de una apertura o un cierre para eliminar objetos no deseados en una imagen dependerá del tipo de imagen y del tipo de objetos a eliminar. En presencia de objetos indeseados con igual tonalidad al resto de estructuras de la imagen, por ejemplo, ruido simétrico (claro y oscuro) habrá que considerar una composición de aperturas y cierres.

Relaciones de orden

La operación de cierre φ es una operación extensiva, mientras que la apertura morfológica γ es antiextensiva. Esto significa que se puede presentar la siguiente relación de orden entre la imagen original I y las transformaciones:

$$\gamma = \delta\varepsilon \leq I \leq \varepsilon\delta = \varphi$$

Crecientes e idempotentes

Las operaciones de apertura y cierre son transformaciones crecientes. Para dos imágenes f, g :

$$\begin{aligned} Si f \leq g &\Rightarrow \gamma(f) \leq \gamma(g) \\ Si f \leq g &\Rightarrow \varphi(f) \leq \varphi(g) \end{aligned}$$

Es posible comprobar que la erosión de una apertura es igual a la erosión del conjunto o imagen original. La misma propiedad se tiene para la dilatación y el cierre:

$$\begin{aligned}\varepsilon(\delta\varepsilon) &= \varepsilon \\ \delta(\varepsilon\delta) &= \delta\end{aligned}$$

Por composición de operaciones:

$$\begin{aligned}\delta\varepsilon(\delta\varepsilon) &= \delta\varepsilon \\ \varepsilon\delta(\varepsilon\delta) &= \varepsilon\delta\end{aligned}$$

que significa que las transformaciones de apertura y cierre son idempotentes, $\gamma = \gamma\gamma\gamma$, $\varphi = \varphi\varphi\varphi$. La propiedad de idempotencia es importante en el filtrado de imágenes, al asegurar que la imagen no sufrirá modificaciones por iteraciones de la transformación.

3.5 Filtrado morfológico

Según Serra (Serra, 1988) y Matheron (Matheron, 1967), un filtro morfológico es cualquier transformación no lineal, en un retículo completo, creciente e idempotente. El requerimiento de creciente es fundamental, pues asegura la conservación del orden del retículo después del filtrado. Los dos filtros morfológicos más elementales son la apertura y el cierre. Concretamente, la apertura es un filtro morfológico antiextensivo. El cierre es un filtro morfológico extensivo. Para la construcción de nuevos filtros morfológicos se parte de los básicos y mediante composición o combinación con otras operaciones se implementan los filtros alternos, los filtros alternos secuenciales, el operador de centro, de contraste, entre otros (Fernández, 1994).

Las aplicaciones básicas de los filtros morfológicos son la atenuación del ruido y la extracción selectiva de objetos en la imagen. Ambas aplicaciones son importantes en visión artificial, pues técnicas de interpretación o medida fallarán en presencia de ruido, mientras que la selección de estructuras y objetos es fundamental en los procesos de segmentación.

Filtros alternos

Sean τ y ϑ dos filtros morfológicos, con $\vartheta \leq \tau$. Por combinación de estos dos filtros se pueden generar cuatro filtros, crecientes e idempotentes: $\tau\vartheta$, $\vartheta\tau$, $\tau\vartheta\tau$, $\vartheta\tau\vartheta$, que satisfacen, entre otras, las siguientes propiedades:

1. $\vartheta \leq \vartheta\tau\vartheta \leq \vartheta\tau \leq \tau\vartheta\tau \leq \tau$
 $\tau\vartheta$
2. $\tau\vartheta\tau = \vartheta\tau \Leftrightarrow \vartheta\tau\vartheta = \tau\vartheta \Leftrightarrow \vartheta\tau \geq \tau\vartheta$
3. $\tau\vartheta\tau$ es el menor filtro mayor que $\tau\vartheta \vee \vartheta\tau$
4. $\vartheta\tau\vartheta$ es el mayor filtro menor que $\tau\vartheta \wedge \vartheta\tau$

Para el caso de una apertura arbitraria γ , y un cierre arbitrario φ , con un mismo elemento estructurante, se tiene que $\gamma\varphi, \varphi\gamma, \gamma\varphi\gamma, \varphi\gamma\varphi$, son filtros. La combinación de γ , y φ , etc, genera la secuencia: $\varphi \rightarrow \gamma\varphi \rightarrow \varphi\gamma\varphi \rightarrow \varphi\gamma \rightarrow \dots$

Es importante señalar que, salvo en el caso en el que se cambie el elemento estructurante, la composición de más de tres operadores no va a proporcionar un nuevo filtro, puesto que si $\tau\vartheta$ es idempotente, entonces $\tau\vartheta\tau\vartheta = \tau\vartheta$.

Filtros alternos secuenciales

El filtrado de una imagen contaminada por ruido simétrico ‘blanco y negro’ puede ser realizado con la combinación de una apertura y cierre o cierre y apertura. Sin embargo, cuando el nivel de ruido en la imagen es elevado, estos filtros alternos sencillos no resuelven el problema, pues el empleo de elementos estructurantes elevados deterioran el resto de estructuras de la imagen.

Una solución es la creación de los filtros alternos secuenciales. A partir de un elemento estructurante pequeño se encadenan combinaciones de aperturas y cierres o cierres y aperturas de elementos estructurantes crecientes. La aparición de los filtros alternos secuenciales en morfología matemática se debe a los trabajos realizados por S. Sternberg (Sternberg, 1986).

Considérese dos familias de operadores $\{\tau_i\}, \{\vartheta_i\}$ tales que $\{\tau_i\}$ crece con i , $\{\vartheta_i\}$ decrece con i y $\vartheta_1 \leq \tau_1$. Se tiene entonces que:

$$\dots \vartheta_n \leq \dots \leq \vartheta_2 \leq \vartheta_1 \leq \tau_1 \leq \tau_2 \leq \dots \leq \tau_n \dots$$

Las composiciones de filtros morfológicos presentes a continuación:

$$\begin{aligned} N_i &= \tau_i \vartheta_i \dots \tau_2 \vartheta_2 \tau_1 \vartheta_1 \\ M_i &= \vartheta_i \tau_i \dots \vartheta_2 \tau_2 \vartheta_1 \tau_1 \end{aligned}$$

son filtros alternos secuenciales y se forman con una composición encadenada de aperturas y cierres, realizadas con un tamaño creciente de elemento estructurante. Es necesario tener en cuenta la forma del elemento estructurante, pues si se trata de elementos cuadrados y el factor o índice del filtro es i , los elementos tendrán una altura y anchura de $2i + 1$. Por ejemplo, N_4 sería una composición de cierre seguido de apertura con elemento estructurante de tamaño 3×3 . Al resultado de estos dos filtros se aplicaría un cierre y una apertura realizada con elemento estructurante de tamaño 5×5 , así sucesivamente, hasta aplicar la última apertura con elemento estructurante de tamaño 9×9 .

Es fácilmente demostrable que los filtros alternos secuenciales (FAS) constituyen filtros morfológicos. Asimismo, satisfacen la siguiente ley de absorción:

$$i \leq j \Rightarrow FAS_j FAS_i = FAS_j \text{ y } FAS_i FAS_j \leq FAS_j$$

La eficacia de los filtros alternos secuenciales en la eliminación de ruido está ampliamente demostrada y constituye en numerosas aplicaciones un paso previo en la segmentación de la imagen. En la figura 3.14 se muestra un ejemplo que nos permite homogenizar las zonas presentes en la imagen de niveles de grises, empleando un filtro alternativo secuencial de nivel 3 que hace uso, inicialmente, de un elemento estructurante con forma de diamante de tamaño 3×3 .

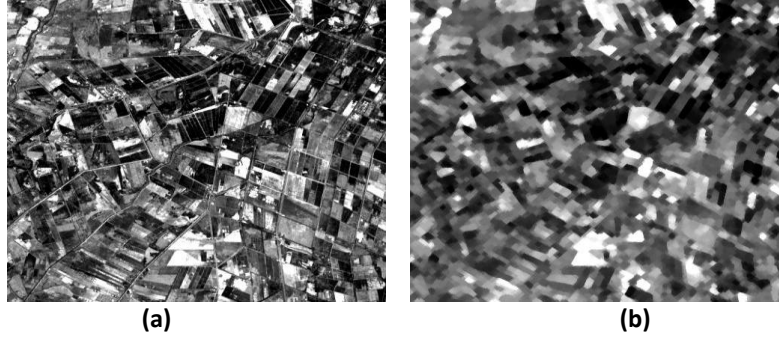


Figura 3.14. (a) Imagen original, (b) Filtrado alternado secuencial de tamaño 7 y patrón $\gamma\phi$.

3.6 Morfología para imágenes multicanal

Contrariamente a las imágenes con tonos de grises, las multicanal no producen escalares sino vectores de píxeles valorados. Desafortunadamente, no hay una manera satisfactoria para definir los valores máximos y mínimos entre dos vectores de más de una dimensión. Una solución parcial para este problema es considerar el ordenamiento marginal. Al hacerlo, cada canal de la imagen es procesada separadamente usando las definiciones presentadas para imágenes con tonos de grises:

$$\begin{aligned} \varepsilon_Y(f) &= [\varepsilon_{Y_1}(f_1), \varepsilon_{Y_2}(f_2), \dots, \varepsilon_{Y_m}(f_m)], \\ \delta_Y(f) &= [\delta_{Y_1}(f_1), \delta_{Y_2}(f_2), \dots, \delta_{Y_m}(f_m)], \end{aligned}$$

Donde $Y = (Y_1, Y_2, \dots, Y_m)$ representa un elemento estructurante multicanal.

Capítulo 4. Transformaciones Geodésicas

El concepto fundamental en el desarrollo geodésico es la conectividad, fuente de multitud de aplicaciones en procesamiento de imágenes. La conectividad de un conjunto puede ser definida como sigue: un conjunto es conexo si cada par de elementos pueden ser unidos por un camino definido por elementos pertenecientes al conjunto. En el caso discreto 2D se puede definir la conectividad como una vecindad local de píxeles adyacentes, en este sentido se tiene, por ejemplo, la conectividad-4 o la conectividad-8 como vecindades válidas entre puntos, como se muestra en la figura 4.1. El etiquetado de una imagen es una transformación directamente asociada con la noción de conectividad.



Figura 4.1. Diferencia entre conectividades (a) conectividad-4, (b) conectividad-8.

La noción de conectividad, que es natural en los conjuntos, no se puede extender automáticamente a funciones numéricas. Se pueden introducir definiciones alternativas que requieren el concepto de partición en el retículo de funciones. Este proceso consiste en dividir los valores de la función en particiones disjuntas de acuerdo a un criterio previamente establecido.

4.1 Distancia geodésica

En la distancia euclídea, tanto en versión continua como digital, todos los posibles obstáculos entre dos puntos son ignorados a la hora de calcular la distancia, es decir, la convexidad del conjunto puede existir o no y la distancia poseerá el mismo valor entre los puntos.

Definición 4.1. Sea X un conjunto no vacío. La distancia geodésica $d_x(p, q)$ entre dos puntos p y $q \in X$ es el mínimo de todos los caminos desde p hasta q formados por puntos intermedios pertenecientes al conjunto X . Si el conjunto es no conexo, como el del ejemplo presente en la figura 4.2, no existe un camino que una los puntos p y r , en este caso se adopta el criterio que la distancia geodésica que separa los puntos es infinita ∞ .



Figura 4.2. Distancia geodésica en el conjunto X .

La distancia geodésica satisface los tres axiomas de cualquier métrica:

$$\begin{aligned} d_X(p, q) &= d_X(q, p) \\ d_X(p, q) &= 0 \Leftrightarrow p = q \\ d_X(p, s) &\leq d_X(p, q) + d_X(q, s) \end{aligned}$$

La distancia geodésica entre un píxel p del conjunto X y un subconjunto A del conjunto X se define como la menor de las distancias geodésicas entre p y cualquier punto $a \in A$. En este caso, al conjunto A se le conoce como marcador y al conjunto X como máscara.

A partir de la distancia geodésica es posible definir el concepto de disco geodésico:

$$D_X^\lambda(p) = \{y : d_X(p, y) \leq \lambda\}$$

En donde λ , es el radio. Cuando este valor aumenta, los discos progresan como frente de ondas emitidos por la fuente p dentro del medio de propagación delimitado por el conjunto, X (figura 4.3). La fuente puede ser cualquier conjunto $A \subset X$.

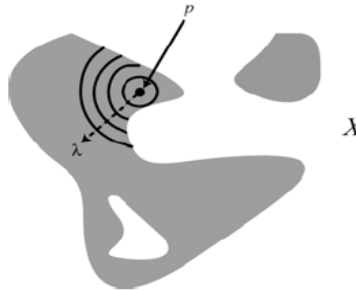


Figura 4.3. Disco geodésico de radio λ en la máscara X .

4.2 Dilataciones y erosiones geodésicas

La dilatación geodésica está íntimamente relacionada con el concepto de distancia geodésica. De hecho, la unión de los discos geodésicos con un radio λ corresponde con la dilatación geodésica de tamaño λ de A (marcador) en X (máscara). Por lo que se define como:

$$\delta_X^{(\lambda)}(A) = \cup \{D_X^\lambda(p), p \in A\}$$

La máscara actúa de límite de propagación de la dilatación del marcador. En el caso de funciones o imágenes, g denota el marcador y f la máscara, tal que $g \leq f$. La dilatación geodésica de tamaño 1 del marcador g con respecto a f se define como el mínimo entre la dilatación de g y la máscara f :

$$\delta_f^{(1)}(g) = \delta^{(1)}(g) \wedge f$$

Donde $\delta_f^{(0)}(g) = g$. Como f actúa de límite, es evidente que:

$$\delta_f^{(1)}(g) \leq f$$

Además, la dilatación geodésica es, al igual que la dilatación clásica, una operación creciente y extensiva. Un ejemplo de dilatación geodésica de señal 1D se proporciona en la figura 4.4. Se observa como la señal marcador es inferior a la señal máscara (figura 4.4(a)). El resultado de dilatar el marcador con un elemento estructurante de tamaño 3 se presenta en la figura 4.4(b). La dilatación geodésica del marcador anterior respecto a la máscara suprime los puntos de g mayores a f . (figura 4.4(c)). La máscara frena la dilatación clásica del marcador.

La dilatación geodésica de tamaño n de una imagen marcador g con respecto a la imagen máscara f se obtiene mediante iteración:

$$(4.1) \quad \delta_f^{(n)}(g) = \delta_f^{(1)}[\delta_f^{(n-1)}(g)]$$

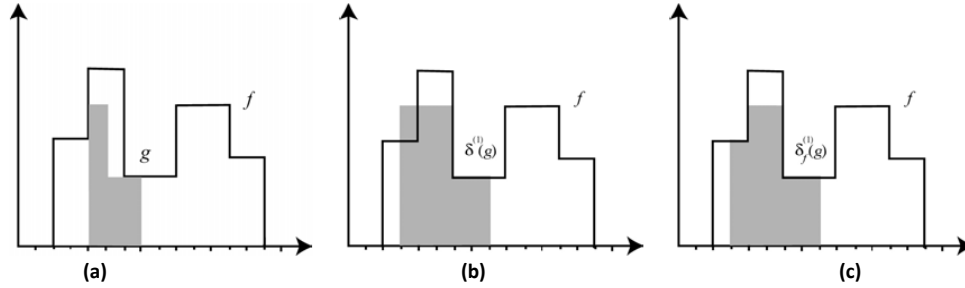


Figura 4.4. Dilatación geodésica de una señal marcador g con respecto a una señal máscara f . Comparación con la dilatación clásica, (a) Señales originales, (b) Dilatación clásica de g , (c) Dilatación geodésica de g respecto a f .

La erosión geodésica es la operación dual a la dilatación geodésica, con respecto a la complementación:

$$\varepsilon_f^{(1)}(g) = [\delta^{(1)}(g^c) \wedge f^c]^c = [(\varepsilon^{(1)}(g))^c \wedge f^c]^c = \varepsilon^{(1)}(g) \vee f$$

Donde $g \geq f$, es decir, el marcador es mayor que la máscara. La imagen marcador g se erosiona y seguidamente se calcula el máximo entre la función erosionada y la máscara. El efecto visual de este tipo de erosión es que la máscara retiene al marcador para que éste

no desaparezca (se limita la contracción del marcador). En este caso, la erosión geodésica es mayor o igual que la máscara, además, es una operación creciente y antiextensiva. La figura 4.5 muestra la diferencia entre una erosión clásica (figura 4.5(a)) y una erosión geodésica (figura 4.5(b)). En la erosión geodésica se puede observar el límite impuesto por la máscara a la desaparición del marcador.

La formación de erosiones geodésicas de tamaño mayor a 1 se realiza, como en la dilatación geodésica, mediante la iteración de erosiones geodésicas de tamaño creciente:

$$(4.2) \quad \varepsilon_f^{(n)}(g) = \varepsilon_f^{(1)}[\varepsilon_f^{(n-1)}(g)]$$

Una erosión geodésica de tamaño 0 no produce cambios en la imagen original,

$$\varepsilon_f^{(0)}(g) = g$$

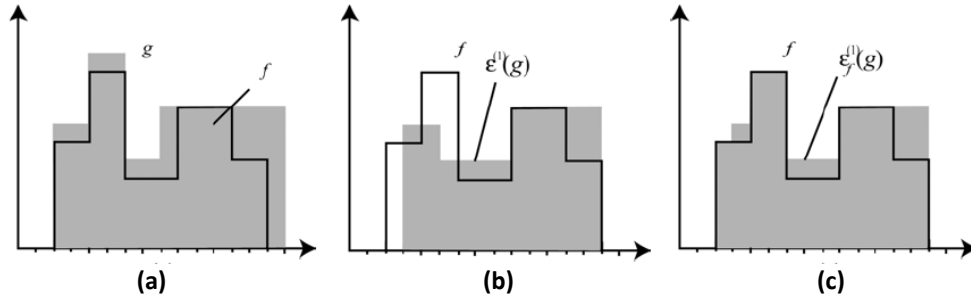


Figura 4.5. Erosión geodésica de una señal marcador g con respecto a una señal máscara f .
(a) Señales originales, (b) Erosión clásica de g , (c) Erosión geodésica de g respecto a f .

Nótese que las transformaciones geodésicas en escala de grises son posibles porque éstas son crecientes.

4.3 Reconstrucción geodésica

La erosión y la dilatación geodésica poseen la particularidad que cuando iteran hasta la estabilidad o idempotencia permiten la definición de algoritmos de reconstrucción morfológica. Tanto la erosión geodésica como la dilatación convergen a la idempotencia en un número finito de iteraciones.

La reconstrucción por dilatación de una imagen-máscara f desde una imagen marcador g , ambas con el mismo dominio y $g \leq f$, se define como la dilatación geodésica de g respecto a f hasta la idempotencia y se denota por $R_f^\delta(g)$:

$$R_f^\delta(g) = \delta_f^{(n)}(g)$$

Donde n es tal que $\delta_f^{(n)}(g) = \delta_f^{(n+1)}(g)$, es decir, la operación no varía.

Como f actúa de límite de propagación, $\delta_f^{(n)}(g) \leq f$, por lo que llegado el momento, la dilatación geodésica no produce ninguna variación de la señal. Es evidente que la reconstrucción es una operación idempotente. Además, la reconstrucción por dilatación es una operación antiextensiva, pues $R_f^\delta(g) \leq f$. Un ejemplo numérico de reconstrucción por dilatación de una señal unidimensional se presenta en la figura 4.6. Cada una de las dilataciones geodésicas que componen la reconstrucción se realiza a partir de la dilatación geodésica de la iteración anterior. De este modo, el marcador puede ir reduciendo progresivamente su intensidad en la propagación bajo la máscara.

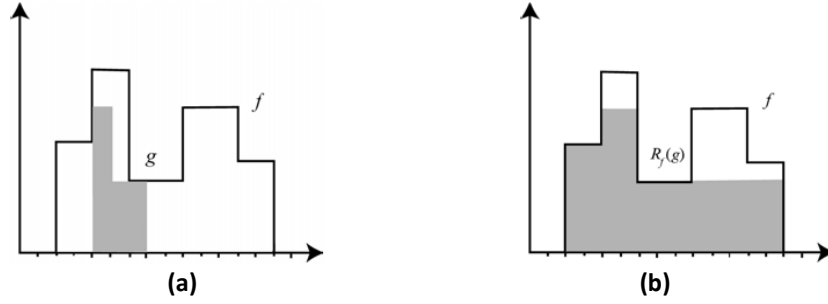


Figura 4.6. Reconstrucción geodésica por dilatación de la señal f desde la señal g .
La reconstrucción se obtiene por la dilatación geodésica de tamaño 7 del marcador g .

La reconstrucción geodésica permite, mediante el uso de marcadores, destacar en una imagen los objetos de interés, facilitando su extracción del resto de la escena, pues la simplificación en la imagen mejorará sensiblemente el proceso posterior de segmentación. Es de destacar que, en la reconstrucción, la definición del objeto sobre el que se sitúa el marcador es perfecta, siempre que éste no presente elevaciones de intensidad en su superficie.

Por dualidad, es posible introducir la operación de reconstrucción por erosión de una máscara f desde un marcador g , como la erosión geodésica de g con respecto a f hasta la estabilidad:

$$R_f^\varepsilon(g) = \varepsilon_f^{(n)}(g)$$

Siendo la máscara menor o igual que el marcador ($f \leq g$). Asimismo, la reconstrucción por erosión de f desde el marcador g es equivalente al complemento de la reconstrucción del complemento de f , desde el complemento de g :

$$R_f^\varepsilon(g) = [R_{f^c}^\delta(g^c)]^c$$

4.3.1 Apertura y cierre por reconstrucción

La apertura por reconstrucción de tamaño n de una imagen f se define como la reconstrucción por dilatación de f a partir de la erosión clásica de tamaño n de f :

$$\gamma_R^{(n)}(f) = R_f^\delta[\varepsilon^n(f)]$$

Contrariamente a la apertura clásica, la apertura por reconstrucción preserva perfectamente la forma de los objetos que no son eliminados por la erosión. Todos los objetos menores que el elemento estructurante son suprimidos, el resto permanece sin alteración en la imagen.

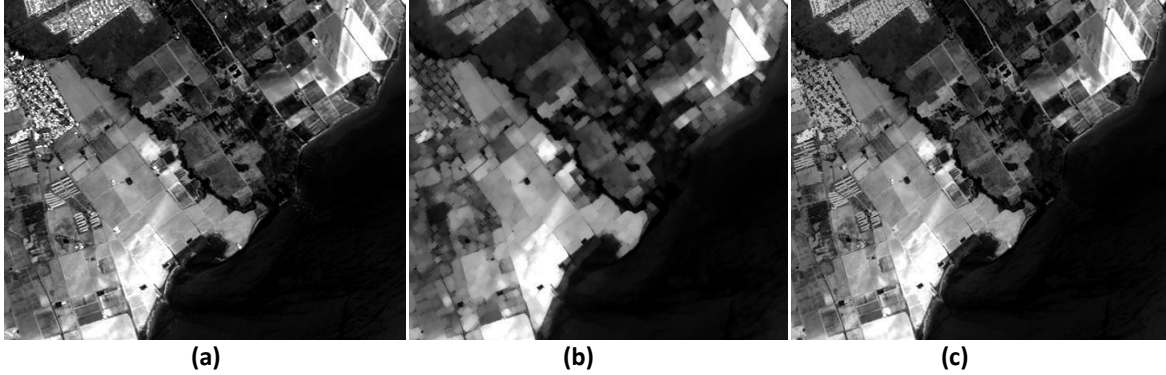


Figura 4.7. Operaciones en la banda azul. Elemento estructurante Diamante Radio 3.
(a) Original, (b) Apertura clásica, (c) Apertura por reconstrucción.

Cuando se hace una apertura (figura 4.7(b)), los objetos brillantes menores al elemento estructurante desaparecen y la imagen se vuelve más oscura. Con la reconstrucción por dilatación utilizando como marcador la imagen erosionada (figura 4.7(c)), los objetos claros menores al elemento estructurante desaparecen, pero la diferencia está en la preservación de las formas. La apertura por reconstrucción, elimina objetos menores en tamaño que el elemento estructurante.

El cierre por reconstrucción se define por dualidad y se corresponde con la erosión geodésica hasta la idempotencia de la imagen dilatada:

$$\varphi_R^{(n)}(f) = R_f^\varepsilon[\delta^{(n)}(f)]$$

Es posible observar la siguiente relación de orden entre los filtros de apertura y cierre geodésicos anteriores y los filtros clásicos:

$$\gamma \leq \gamma_R \leq I \leq \varphi_R \leq \varphi$$

4.3.2 Reconstrucción auto-dual y nivelación

Los operadores conexos son herramientas de filtrado que actúan por la fusión de regiones elementales llamadas zonas planas. Los operadores conexos no crean contornos nuevos ni modifican su posición. Por lo tanto, tienen muy buenas propiedades de conservación de

contornos y son capaces del reconocimiento de objetos tanto con el filtrado de bajo nivel como el de nivel superior.

En la práctica, los operadores conexos se obtienen teniendo en cuenta que la imagen marcador g es una transformación $\tau(f)$ de la imagen de entrada f que también desempeña el papel de imagen marcador. Como resultado, la mayoría de los operadores conexos obtenidos ψ por la reconstrucción se pueden escribir como

$$\psi(f) = \delta_f^{(n)}(\tau(f)) \text{ (Operador antiextensivo) ó}$$

$$\psi(f) = \varepsilon_f^{(n)}(\tau(f)) \text{ (Operador extensivo)}$$

Las reconstrucciones por apertura y cierre discutidos en el apartado anterior son antiextensivas o extensivas. Ellas permiten la simplificación de cualquiera de las componentes brillantes u oscuras de la imagen. Para algunas aplicaciones, este comportamiento es un inconveniente por lo que gustaría simplificar de manera simétrica, todos los componentes. Desde el punto de vista teórico, esto significa que el filtro tiene que ser auto-dual, es decir,

$$\psi(f) = \sim \psi(\sim f).$$

Con el objetivo de construir los operadores conexos auto-duales, el concepto de nivelación es propuesto mediante la adición de algunas restricciones en la definición de los operadores conexos.

Definición 4.2. El operador ψ es una nivelación si $\forall n, n'$ son los píxeles vecinos,

$$\psi(f)[n] > \psi(f)[n'] \Rightarrow f[n] \geq \psi(f)[n] \text{ y } \psi(f)[n'] \geq f[n'].$$

Esta definición no sólo establece que sí existe una transición en la imagen de salida, que ya estaba presente en la imagen original (equivalente a decir que el operador está conectado), sino también que, el sentido de variación de nivel de gris entre n y n' tiene que ser preservado y como segundo el intervalo de variación entre $\psi(f)[n]$ y $\psi(f)[n']$ debe contener el intervalo de variación entre $f[n]$ y $f[n']$.

En particular, la apertura o el cierre por reconstrucción es una nivelación. Si ψ_1, ψ_2 son nivelaciones, su composición $\psi_2\psi_1$ es también una nivelación. Finalmente, si $\{\psi_i\}$ son nivelaciones, el supremo $\bigvee_i \psi_i$ y el ínfimo $\bigwedge_i \psi_i$, son nivelaciones.

La técnica más popular para crear nivelaciones se basa en el proceso de reconstrucción auto-dual que se describe a continuación.

Definición 4.3. Si f y g son dos imágenes (llamadas imágenes máscara y marcador respectivamente), la reconstrucción auto-dual $R^\dagger(g|f)$ de g con respecto a f está dada por

$$g_k = \varepsilon_c(g_{k-1}) \vee [\delta_c(g_{k-1}) \wedge f] = \delta_c(g_{k-1}) \wedge [\varepsilon_c(g_{k-1}) \vee f] \text{ y}$$

$$R^\dagger(g|f) = \lim_{k \rightarrow \infty} g_k,$$

Donde $g_0 = g$, δ_c y ε_c son la dilatación y la erosión respectivamente con el elemento estructurante plano con conectividad definida.

De hecho, la reconstrucción auto-dual es la reconstrucción antiextensiva de (4.1) para los píxeles donde $g[n] < f[n]$, y la reconstrucción extensiva de (4.2) para los píxeles donde $f[n] < g[n]$. Un ejemplo de reconstrucción en una dimensión se muestra en la figura 4.8.

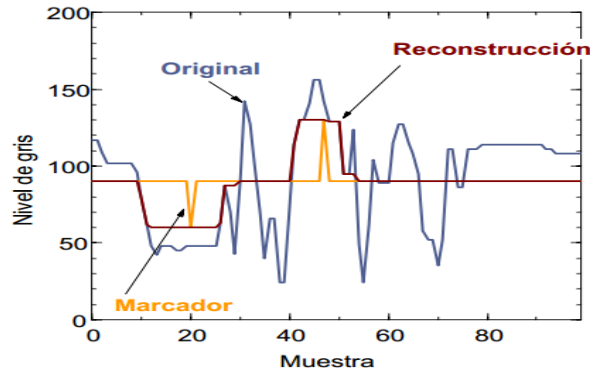


Figura 4.8. Reconstrucción Geodésica.

En la práctica, la reconstrucción auto-dual se utiliza para restaurar la información de contorno después de un proceso de filtrado inicial. En otras palabras, la reconstrucción permite la creación de una versión conexa $R^\dagger(\psi(f)|f)$ de cualquier filtro $\psi(f)$.

Un ejemplo típico de filtro inicial $\psi(f)$ es un filtro alternado secuencial

$$\psi(f) = \varphi_{h_k} \gamma_{h_k} \varphi_{h_{k-1}} \gamma_{h_{k-1}} \dots \varphi_{h_1} \gamma_{h_1} (f),$$

Donde φ_{h_k} y γ_{h_k} son cierres y aperturas respectivamente con un elemento estructurante h_k .

Los resultados de la reconstrucción están fuertemente influenciados por la elección correcta de los conjuntos máscara y marcador. La imagen bajo estudio es usualmente el elemento máscara. Se establece que un buen marcador será aquel que proceda de:

- Conocimiento del resultado esperado.
- Conocimiento de los objetos presentes en las imágenes.
- Transformaciones básicas de la imagen máscara.
- Interacción con el usuario: marcadores definidos manualmente.

4.4 Operadores basados en reconstrucción

Las técnicas de reconstrucción morfológica son la base de numerosas transformaciones de alto nivel que van desde, procedimientos simples para eliminar objetos conectados al borde de una imagen, hasta los más sofisticados como mediciones de profundidad a los extremos de una imagen.

4.4.1 Extremos regionales

Los mínimos y los máximos son características importantes de una imagen porque con frecuencia, ellos marcan objetos relevantes: los mínimos para objetos oscuros y los máximos para los objetos brillantes. En morfología, el término mínimo se utiliza en el sentido de mínimo regional, es decir, un mínimo cuyo alcance no se limita necesariamente a un píxel único. Un ejemplo de mínimo regional está dado en la figura 4.9.

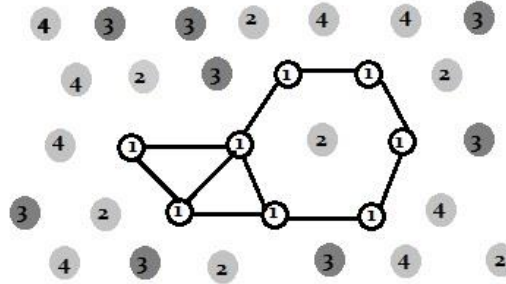


Figura 4.9. Mínimo regional de nivel 1.

Un mínimo regional M de una imagen f de elevación t es un componente conexo de píxeles con valor t cuyo entorno está formado por píxeles que tienen un valor estrictamente mayor que t . Similarmente, un máximo regional M de una imagen f de elevación t es un componente conexo de píxeles con valor t cuyo entorno está formado por píxeles que tienen un valor estrictamente menor que t .

En la detección de máximos regionales se emplea una imagen marcador resultado de desplazar hacia abajo la función original f por una constante c , en este caso, $c=1$. Los máximos regionales son el residuo entre la imagen original y la reconstrucción por dilatación de f desde la imagen desplazada:

$$RMAX(f) = f - R_f^\delta(f - 1)$$

Si el tipo de datos de la imagen no admite valores negativos, la siguiente definición equivalente debe ser considerada:

$$RMAX(f) = f + 1 - R_{f+1}^\delta(f)$$

Análogamente, para obtener los mínimos regionales se emplea la reconstrucción por erosión a partir de un marcador que eleva la señal en un nivel:

$$RMIN(f) = R_f^\varepsilon(f + 1) - f$$

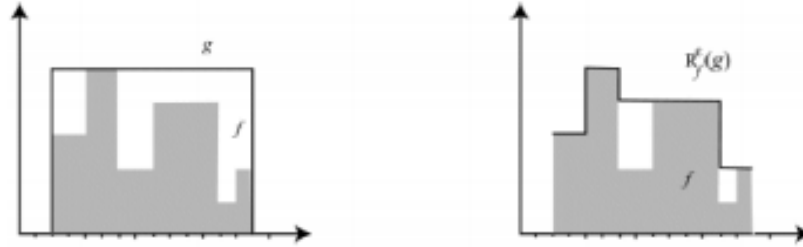


Figura 4.10. Detección de Mínimos por reconstrucción.

4.4.2 Imposición de mínimos

La técnica de imposición de mínimos se refiere a la filtración de los mínimos de la imagen. Requiere de un conjunto de marcadores en zonas oscuras relevantes. Ellos pueden ser obtenidos usando algunas transformaciones apropiadas de la imagen de entrada o, alternativamente, definidos manualmente por el usuario. La imagen marcadora f_m para cada píxel p , se define de la manera siguiente:

$$f_m(p) = \begin{cases} 0, & \text{Si } p \text{ pertenece a un marcador} \\ t_{max}, & \text{en otro caso} \end{cases}$$

La imposición de mínimos de una imagen f se presenta de dos maneras. La primera, se calcula el mínimo puntual entre la imagen de entrada y la imagen marcadora: $f \wedge f_m$. Al hacerlo, los mínimos se crean en los lugares correspondientes a los marcadores (si es que no existen aún) y nos aseguramos de que la imagen resultante es menor o igual a la imagen marcadora. Por otra parte, dos mínimos distintos a imponer podrían caer dentro de un mínimo de f con nivel 0, por lo que es mejor considerar $(f + 1) \wedge f_m$. La segunda manera consiste en una reconstrucción morfológica por erosión de la imagen marcadora f_m con respecto a $(f + 1) \wedge f_m$:

$$R_{(f+1) \wedge f_m}^\varepsilon(f_m).$$

Un ejemplo de imposición de mínimos por marcadores se muestra en la figura 4.11(a). La figura 4.11(b) representa la eliminación de mínimos por medio de una reconstrucción por erosión con un elemento estructurante en forma de círculo y de radio ρ .

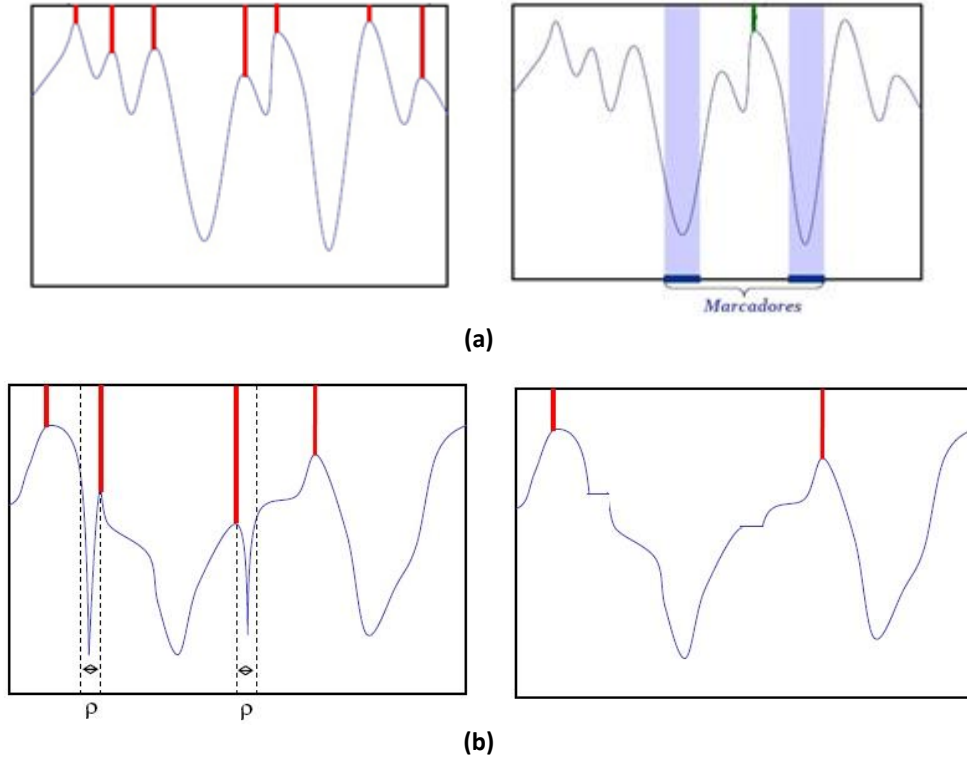


Figura 4.11. Imposición de Mínimos. (a) Manual, (b) Por reconstrucción.

4.4.2 Extendidos y h-extremos

La transformación *h-extremos* proporciona una herramienta para filtrar los extremos de la imagen utilizando un criterio de contraste. Más precisamente, la transformación de *h – máximos* suprime todos los máximos cuya profundidad es inferior o igual a un umbral dado de nivel h , es decir, cuya diferencia con sus vecinos sea inferior a h . Esto se consigue mediante la reconstrucción por dilatación de $f - h$ respecto a f :

$$Hmax_h(f) = R_f^\delta(f - h)$$

Similarmente se define la transformación *h – mínimos*, la cual permite suprimir mínimos regionales cuya diferencia con sus vecinos sea inferior a h :

$$Hmin_h(f) = R_f^\varepsilon(f + h)$$

La figura 4.12 es un ejemplo de la transformación de *h – mínimos*.

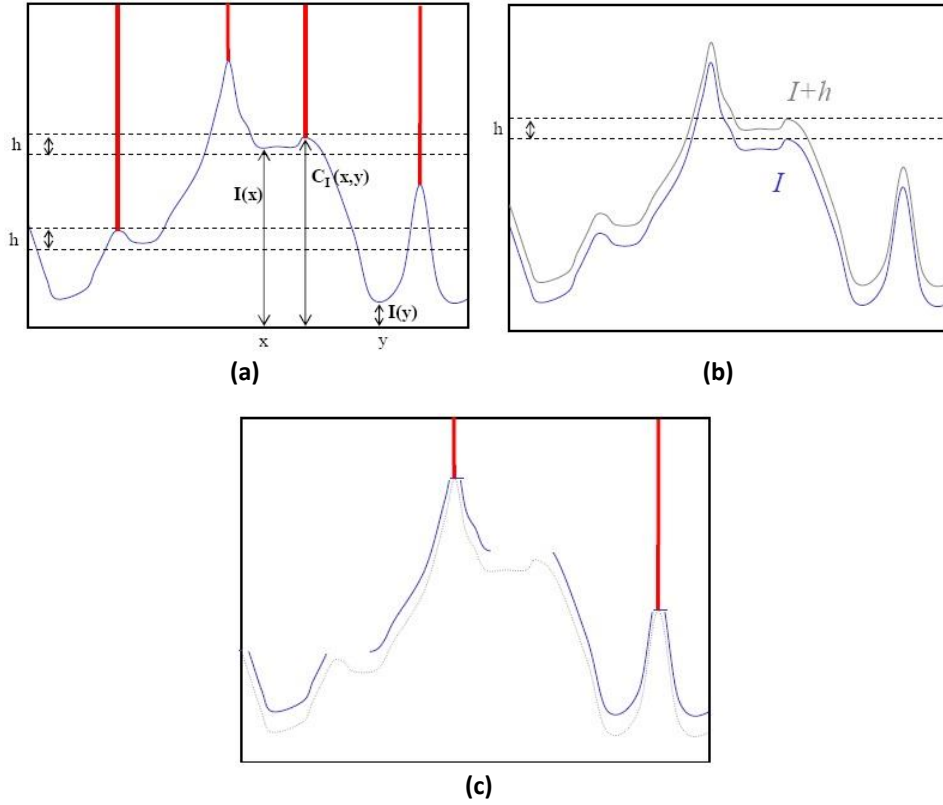


Figura 4.12. Eliminación de Mínimos regionales de profundidad h .

Los mínimos extendidos $EMIN$ son definidos como los mínimos regionales de la correspondiente transformación h – mínimos, los máximos extendidos $EMAX$ se definen por la dualidad:

$$EMAX_h(f) = RMAX[Hmax_h(f)],$$

$$EMIN_h(f) = RMIN[HMIN_h(f)].$$

Estos extremos son extendidos en el sentido de que si un extremo original es todavía perteneciente a un extremo de la imagen reconstruida, su extensión espacial es o bien invariable o extendido debido al proceso de reconstrucción. El número total de extremos extendidos sólo puede decrecer cuando incrementa el valor de h .

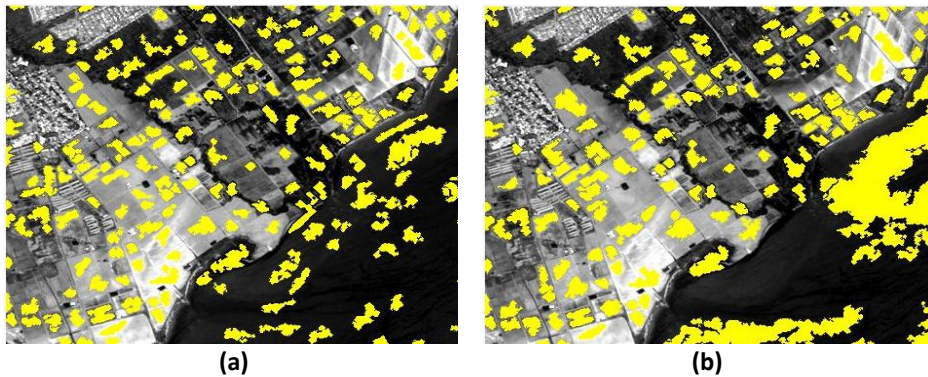


Figura 4.13. Ejemplo de Mínimos Extendidos.

(a) Mínimos Regionales Originales, (b) Mínimos Extendidos $h=2$.

Capítulo 5. Segmentación

La segmentación de una imagen se puede entender como su partición en diferentes regiones disjuntas, teniendo cada una de éstas propiedades que le son específicas. Es un paso clave de interpretación cuantitativa de los datos presentes en la imagen.

En términos matemáticos, la segmentación de una imagen f es una partición de su dominio definido D_f en n conjuntos disjuntos no vacíos $X_1, X_2, \dots, X_n$, llamados segmentos tales que:

$$D_f = \bigcup_{i=1}^n X_i$$

Usualmente, una imagen que ha sido segmentada se representa como una imagen etiquetada en la que a los píxeles de cada segmento X_i reciben un código único que los representa.

Los descriptores de las regiones constituyen un conocimiento previo sobre los objetos de la imagen que permitirá agrupar los píxeles. Este conocimiento se refiere a las características tales como la forma, tamaño, orientación, distribución de los niveles de grises, y textura. Idealmente, deben permitir que se discriminen dos objetos diferentes de la imagen. Desafortunadamente, las características que permiten el reconocimiento de todos los objetos de la imagen rara vez se disponen en la aplicación práctica. Además, el valor de una de ellas suele ser afectado por parámetros ajenos a los objetos presentes en la imagen como son la iluminación, la resolución del sensor, el ángulo de visión, el nivel de ruido, etc. Debido a esto, no existe un algoritmo general que produzca una segmentación óptima para todas las imágenes. Una hipótesis aplicada en muchas técnicas de segmentación consiste en la definición de objetos de la imagen como regiones que cumplan alguna propiedad de uniformidad tal como la distribución de niveles de grises homogénea.

En lo que sigue, se concentran técnicas de segmentación basadas en la transformación directa de la definición del dominio de la imagen. Existen dos enfoques importantes en función de si el núcleo de los segmentos está creciendo (técnica de crecimiento de regiones), o el de sus límites (técnica de detección de bordes).

Los algoritmos de segmentación de imágenes se basan en alguna de las tres propiedades siguientes:

Discontinuidad en los tonos de gris de los píxeles de un entorno, que permite detectar puntos aislados, líneas y aristas (bordes).

Similitud en los tonos de gris de los píxeles de un entorno, que permite construir regiones por división y fusión, por crecimiento o por umbralización.

Conectividad de los píxeles desempeña un papel importante en la segmentación de imágenes.

Como herramienta previa al estudio de los métodos, se muestran las técnicas de detección de puntos, rectas, bordes y contornos.

5.1 Detección de puntos

Un punto aislado de una imagen tiene un tono de gris que difiere significativamente de los tonos de gris de sus píxeles vecinos, es decir, de los ocho píxeles de su entorno 3×3. Un punto se detecta en la posición en la que está centrada la máscara si $|R| \geq T$, donde T es un umbral no negativo y R es la suma de productos de los coeficientes con los niveles de gris contenidos en la región en la que se superpone la máscara.

5.2 Detección de Líneas

Una línea es una secuencia de píxeles en la que dos píxeles consecutivos están conectados, es decir, son vecinos en un entorno 3×3 de alguno de ellos. Cada píxel se puede conectar con alguno de sus 8 píxeles vecinos, y por lo tanto, vamos a tener sólo 4 direcciones (tramos lineales) posibles: horizontal, vertical, de 45° y de -45°. En consecuencia, cualquier curva digital va a venir dada por una línea constituida sólo por tramos lineales de cualquiera de estos cuatro tipos. Así, para la detección de líneas podemos utilizar una máscara 3×3 que se irá moviendo por los píxeles de la imagen. Los píxeles que forman parte de una línea tendrán respuestas extremas. Por ejemplo, para detectar una raya o segmento horizontal (línea recta), cada píxel del segmento tiene un tono de gris que difiere del tono de gris de los píxeles vecinos superiores e inferiores. Por ello, podemos utilizar la siguiente máscara para detectar una línea recta horizontal:

$$M_1 = \begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}$$

Para detectar rayas verticales, de 45° y de -45°, podemos utilizar, respectivamente las máscaras siguientes:

$$M_2 = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}, M_3 = \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix} \text{ y } M_4 = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Obsérvese que la dirección preferida de cada máscara está ponderada con un coeficiente mayor que las demás direcciones posibles.

Se aplicarán las cuatro máscaras anteriores a todos los píxeles de la imagen representada en la figura 5.1(a). Sea $g_k(i, j)$ el resultado de aplicar la máscara M_k centrada en el píxel (i, j) . Fijado un valor umbral $T > 0$, diremos que el píxel (i, j) constituye una parte horizontal de una línea si,

$$|g_1(i,j)| > |g_k(i,j)|, \quad \forall k \neq 1, \text{ y } |g_1(i,j)| > T.$$

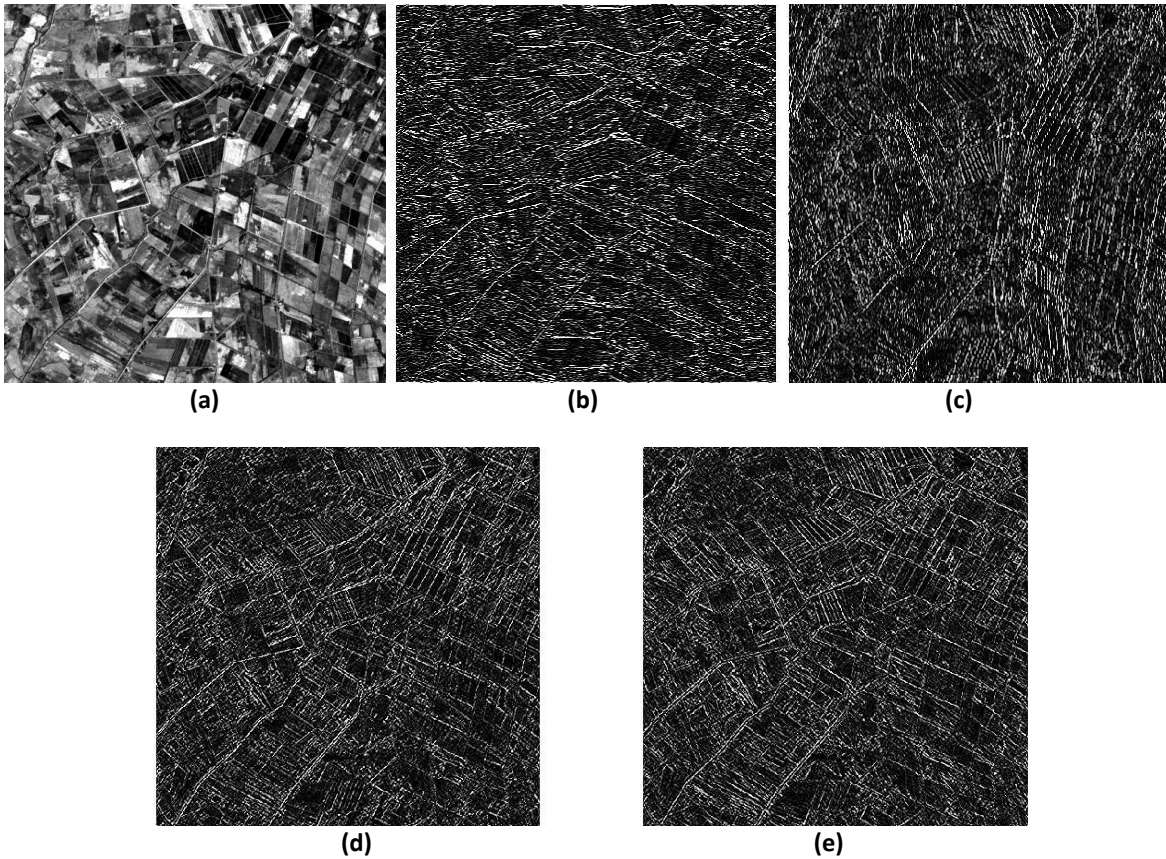


Figura 5.1. Detección de líneas. (a) Canal Rojo, subimagen de Calabozo, (b), (c), (d), (e) Filtros con las máscaras M_1 , M_2 , M_3 , M_4 respectivamente.

5.3 Detección de bordes

Un borde o arista es la frontera entre dos regiones cuyos tonos de gris difieren significativamente o tienen propiedades diferentes, como ocurre en el caso de texturas. Si queremos detectar los bordes tenemos que poner énfasis en los cambios bruscos de los niveles de gris de píxeles vecinos y suprimir aquellas áreas con valores de gris constantes. Los operadores derivada serán útiles para realizar dichas tareas.

Las técnicas de segmentación basadas en los bordes son adecuadas cuando las regiones son suficientemente homogéneas de manera que la transición entre regiones se pueda realizar en base a los tonos de gris. En caso contrario es más adecuado utilizar las técnicas de segmentación basadas en regiones.

Los métodos basados en el gradiente son adecuados cuando la transición en los niveles de gris de los píxeles del entorno es brusca, similar a una función escalón (paso). Sin embargo, estos operadores son pocos sensibles a los cambios graduales en los niveles de

gris, que corresponden a funciones tipo rampa. En estos casos es más adecuado aplicar operadores basados en las derivadas de segundo orden, como, por ejemplo, el operador Laplaciano, definido a partir de la laplaciana de una función:

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}$$

En el caso, donde las imágenes son digitales (discretas), este operador se obtiene aplicando consecutivamente los operadores gradiente en cada dirección.

Es decir,

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Que nos dice que la imagen final es la suma de una imagen de bordes verticales con una imagen de bordes horizontales.



Figura 5.2. Detección de bordes mediante el filtro de Laplace.

5.4 Métodos de segmentación

Los métodos de segmentación se pueden agrupar en cuatro clases diferentes:

- a) Métodos basados en píxeles, que a su vez pueden ser:
 - locales (basadas en las propiedades de los píxeles y su entorno)
 - globales (basadas en la información global obtenida, por ejemplo, con el histograma de la imagen).
- b) Métodos basados en bordes.
- c) Métodos basados en regiones, que utilizan las nociones de homogeneidad y proximidad geométrica, como las técnicas de crecimiento, fusión o división.
- d) Métodos basados en modelos.

5.4.1 Métodos basados en píxeles

Consideremos la imagen digital $\{f(m, n): (m, n) \in D \subset \mathbb{Z} \times \mathbb{Z}\}$ donde el dominio

$$D = \{(m, n): m = 0, 1, \dots, M - 1; n = 0, 1, \dots, N - 1\}$$

Se ha dividido en k regiones disjuntas $R_1, \dots, R_k$, de manera que para un cierto predicado P que asegura la homogeneidad de la región se verifica:

$$P(R_i) = \text{VERDADERO} \text{ para } i = 1, 2, \dots, k$$

$$P(R_i \cup R_j) = \text{FALSO} \text{ para } i \neq j$$

Esto expresa que la unión de regiones no puede darnos una región homogénea. Dichas regiones deben cubrir el dominio de la imagen, D , es decir, $D = \bigcup_i R_i$, y cada píxel sólo puede pertenecer a una región ($R_i \cap R_j = \emptyset$, para $i \neq j$). Al predicado lógico P definido sobre el conjunto de las partes del dominio de la imagen, D , lo llamaremos *regla de segmentación*.

5.4.1.1 Segmentación basada en el histograma

El histograma de una imagen digital con niveles de gris en el rango $[0, L - 1]$ es una función discreta:

$$p(r_k) = \frac{n_k}{n},$$

Donde r_k es el k -ésimo nivel de gris, n_k es el número de píxeles de la imagen con ese nivel de gris y n el número total de píxeles de la imagen. En forma general podemos decir que $p(r_k)$ da una idea acerca de la probabilidad de que aparezca el nivel de gris r_k . La representación gráfica de esta función para todos los valores de k proporciona una descripción global de la apariencia de la imagen. Aunque el histograma no indica nada específico sobre el contenido de la imagen, su perfil proporciona sin duda información muy útil sobre la posibilidad de mejorar la imagen, entendiendo por mejora el destacar características de interés de la misma.

En la segmentación basada en histogramas se define un predicado lógico P de la forma $P(R, \mathbf{x}, \mathbf{t})$, donde R es una región del dominio, $\mathbf{x}$ es un vector de características de un píxel de la imagen (nivel de gris, textura, etc.) y $\mathbf{t}$ es un conjunto de parámetros (umbrales). Así, una regla de segmentación simple es:

$$P(R, f, \mathbf{t}): t_1 < f(m, n) < t_2$$

Que expresa que la región R está formada por los píxeles cuyo nivel de gris o intensidad luminosa está en el intervalo (t_1, t_2) .

El uso del histograma consiste en retener aquellos píxeles que se encuentren en un cierto entorno de niveles de gris del histograma definido por el usuario. Permite al usuario analizar la imagen, extraer información relevante y sacar conclusiones a partir de esta interacción.

El histograma es la base de numerosas técnicas de procesamiento de la imagen en el dominio espacial. La manipulación del histograma puede utilizarse eficazmente para el realce de la imagen, como se muestra a continuación. La figura 5.6(a) corresponde al canal rojo de una imagen proveniente del satélite Landsat 8, con fecha de captura de enero de 2014, se observa el Lago de Tacarigua y sus adyacencias junto con su respectivo histograma. En la figura 5.6(b) y figura 5.6(c), se muestra el filtrado en dos intervalos diferentes.

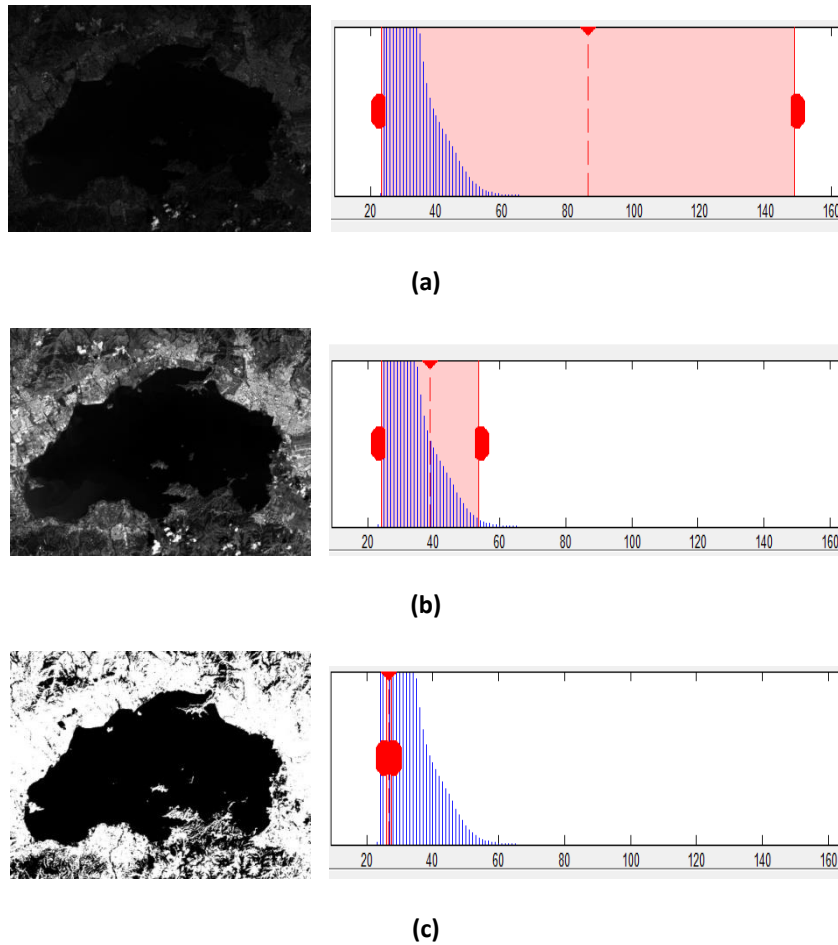


Figura 5.6. (a) Canal Rojo Histograma original, (b) Eliminación del 2% de datos a los extremos, (c) Parámetro de delimitación $r_k = 28$.

El problema más sencillo de segmentación se presenta cuando la imagen está formada por un sólo objeto que tiene intensidad luminosa homogénea sobre un fondo con un nivel de intensidad diferente. En este caso la imagen se puede segmentar en dos regiones utilizando el operador umbral con parámetro t . La elección del valor umbral t se puede hacer a partir del histograma. Si el fondo tiene también intensidad luminosa homogénea, entonces el histograma es bimodal y el umbral t que se debe tomar es el que corresponde al mínimo local que está entre los dos máximos del histograma. En ocasiones no se distinguen bien las dos zonas modales (los máximos) y se debe aplicar una técnica de variación espacial del umbral que consiste en dividir la imagen en bloques cuadrados (ventanas) a los que se les calcula el histograma y el valor umbral correspondiente (cuando no es bimodal se determina entonces el umbral por interpolación de los valores umbrales de las ventanas de su entorno).

Cuando el histograma es muy irregular, es difícil calcular este mínimo local, en este caso es conveniente suavizarlo (utilizando un filtro).

Si la imagen contiene N objetos (regiones homogéneas) se puede utilizar una umbralización múltiple, es decir,

$$P(R_i) = \text{VERDADERO}, \text{ siempre y cuando } R_i = \{(m, n): t_{i-1} \leq f(m, n) \leq t_i\}$$

Los valores umbrales se pueden obtener también del histograma calculando los $N - 1$ mínimos locales entre los N picos consecutivos del histograma. A veces estos picos o crestas no se distinguen bien, debido a los píxeles que están en las transiciones de los contornos. Por ello, es conveniente excluir a estos píxeles (bordes locales) en el cálculo del histograma. En la figura 5.7 se muestra una imagen de la luna. A partir del histograma se observa que los píxeles del intervalo $[75, 255]$ corresponden a la luna y el resto al fondo de la imagen.

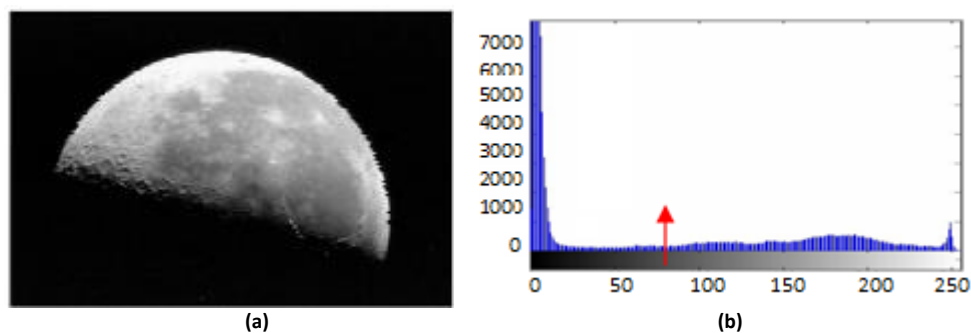


Figura 5.7. (a) Imagen de la Luna, (b) Histograma.

5.4.2 Métodos basados en bordes

El conjunto de píxeles que están en la frontera raramente caracterizan completamente el contorno de un objeto, debido a varias causas, el ruido o una iluminación no uniforme, que introducen discontinuidades falsas en los tonos de gris. En la segmentación basada en píxeles se suele producir un sesgo en el tamaño de los objetos segmentados, es decir, el tamaño de los objetos va a depender del valor umbral seleccionado, debido a que el tono de gris de los bordes de un objeto puede cambiar gradualmente desde el fondo hasta el valor del objeto. Para evitar esto se pueden utilizar los métodos basados en las derivadas. Estos métodos son secuenciales, van siguiendo el contorno de los objetos utilizando las posiciones de los bordes que vienen determinadas por los valores extremos del gradiente o por los cruces con el cero (si utilizamos operadores de segundo orden).

5.4.3 Métodos basados en crecimiento de regiones.

Las técnicas basadas en píxeles aseguran la homogeneidad de las regiones pero no garantizan que sean conexas. En la segmentación de imágenes se debe considerar también la información suministrada por los píxeles del entorno para conseguir así regiones homogéneas y conexas, pues los píxeles de un mismo entorno suelen tener propiedades estadísticas similares y pertenecer a una misma región.

Una segmentación basada en regiones de una imagen digital I consiste en realizar una partición de la imagen en k regiones con las siguientes propiedades:

- a) Las regiones obtenidas en la partición, $R_1, R_2, \dots, R_k$, deben ser disjuntas

$$R_i \cap R_j = \emptyset, \quad i \neq j$$

- b) Su unión debe ser la imagen completa

$$\bigcup_{i=1}^k R_i = I$$

- c) Cada región R_i tiene que ser conexa, es decir, todos sus píxeles conectados
d) Se debe verificar que:

$$P(R_i) = \text{VERDADERO} \text{ y } P(R_i \cup R_j) = \text{FALSO},$$

Para regiones adyacentes cualesquiera, R_i y R_j , siendo P el predicado que nos proporciona el test de homogeneidad de la región.

El crecimiento de las regiones es un procedimiento que consiste en ir formando grupos (regiones) por incorporación sucesiva de píxeles de la imagen a los grupos según algún

criterio predefinido. El procedimiento más sencillo de crecimiento de regiones comienza con un conjunto de píxeles, llamados semillas, que representan las distintas regiones de la imagen (al menos una semilla por región). A partir de las semillas se van incorporando nuevos píxeles a las regiones utilizando un mecanismo de crecimiento que detecta, en cada etapa k y para cada región $R_i(k)$, los píxeles aun no clasificados que pertenecen a un entorno predefinido de algún píxel del contorno de la región $R_i(k)$. Para cada píxel (x, y) , detectado de esta forma, se comprueba si cumple la regla de homogeneidad, es decir, si la nueva región $D_i(k) \cup (x, y)$ sigue siendo homogénea, en cuyo caso se consigue ampliar la región añadiéndole dicho píxel. El píxel se elige del entorno de un píxel del contorno de la región para garantizar de alguna manera la conectividad de la región.

Cuando las imágenes son monocromáticas el análisis de la región se hace utilizando un conjunto de descriptores como el tono de gris medio, textura, momentos, propiedades espaciales, etc. La regla de homogeneidad obedecerá a algún criterio de similitud de los píxeles de la misma región y se basa en dichos descriptores.

La solución dada por esta técnica depende en gran medida de las semillas elegidas. Por ello es conveniente elegir bien dichas semillas. Otra cuestión importante es la formulación de una regla de parada. El procedimiento de parada termina cuando no quedan más píxeles para que sean asignados a una región pues los criterios de similitud suelen ser de naturaleza local y no se modifican durante el proceso de crecimiento, es decir, no tiene en cuenta la historia del proceso. También se pueden utilizar criterios de similitud dinámicos que tienen en cuenta conceptos como el tamaño o la forma de la región y la representatividad de un píxel con respecto a los píxeles de la región.

Cuando se elige más de una semilla por región es necesario unir o fusionar algunas de las regiones obtenidas. Está claro que las regiones a unir deben de ser adyacentes y tener propiedades estadísticas similares, como pueden ser la intensidad (nivel de gris) media de la región y su desviación típica.

La fusión de regiones se puede incorporar en el mecanismo de crecimiento y conseguir así un algoritmo combinado de crecimiento y fusión. Para ello se necesita de una regla de fusión, de la forma $P(R_i \cup (r, s))$, que especifica cuando un píxel (r, s) se asigna a una de sus regiones R_i adyacentes:

$$P(R_i \cup (r, s)) = \begin{cases} VERDADERO & \text{si } |f(r, s) - m_i| \leq T_i(r, s) \\ Falso & \text{en otro caso} \end{cases}$$

Donde m_i es el valor medio de la intensidad luminosa de la región R_i , y $T_i(r, s)$ es el valor umbral utilizado para dicha región R_i .

El crecimiento de la región depende en gran medida del valor umbral T seleccionado. Si T es pequeño es más difícil que se produzcan fusiones y se generan muchas regiones pequeñas, en cambio, si T es grande se favorece la fusión y así se generan pocas regiones pero grandes.

5.4.4 Métodos basados en división y fusión de regiones.

Este método parte de la imagen completa que se divide iterativamente en subregiones. Cada subimagen obtenida que no sea homogénea se vuelve a dividir y así sucesivamente. Así, por ejemplo, se puede comenzar con una imagen inicial. Esta se divide en cuatro subimágenes, llamadas cuadrantes, y cada subimagen obtenida que no sea homogénea se vuelve a dividir en otras cuatro subimágenes; el proceso se continúa sucesivamente hasta que todas ellas sean homogéneas, es decir, cumplan las condiciones a), b), c) y d) establecidas anteriormente para que constituyan una segmentación. Sólo se necesita aplicar la regla de homogeneidad en cada cuadrante. Si el predicado P es FALSO para algún cuadrante entonces se subdivide en 4 subcuadrantes y así sucesivamente. Dichos métodos se llaman de partición o división. Si la imagen tiene una dimensión $N \times N$ donde N es potencia de 2, es decir, $N = 2^n$, el procedimiento recursivo se puede representar mediante un árbol cuaternario (cada vértice tiene 4 descendientes) como se muestra en la figura 5.8, donde el nodo raíz corresponde a la imagen completa y sus descendientes son cada uno de los 4 cuadrantes en que se divide, y así sucesivamente.

Las técnicas basadas en división tienen el inconveniente de que pueden generar regiones que sean adyacentes y con idénticas propiedades (homogéneas), que deberían estar fundidas en una sola región. Este inconveniente sugiere que se utilicen técnicas mixtas, es decir, algoritmos de partición (división) y fusión.

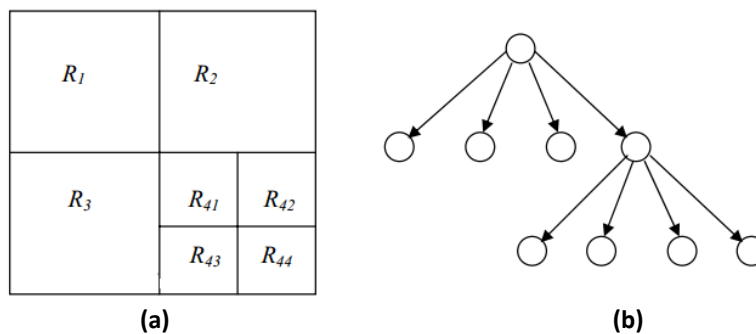


Figura 5.8. (a) Imagen dividida. (b) Árbol cuaternario correspondiente.

5.5 Segmentación morfológica

La segmentación morfológica es un método de crecimiento de regiones que combina operaciones morfológicas, tales como, mínimos extendidos, gradientes morfológicos, líneas divisorias de agua (Watershed), entre otras, para segmentar imágenes en escala de grises.

5.5.1 Esqueleto geodésico por zonas de influencia

La definición de esqueleto geodésico por zona de influencia (SKIZ), procede directamente del concepto de distancia geodésica.

Definición 4.2. Sea X un conjunto compuesto de N elementos desconectados B_i definido en el dominio D_f , de una imagen en escala de grises f . La zona de influencia geodésica de una componente conexa B_i en X es el conjunto de puntos del dominio de definición de f cuya distancia geodésica hasta B_i es más pequeña que la distancia geodésica a los otros componentes de X :

$$IZ_X(B_i) = \{p \in D_f \mid \forall i \neq j, d_X(p, B_i) < d_X(p, B_j)\}$$

El conjunto $IZ_X(B)$ es la unión de las zonas de influencia geodésica de las componentes conexas B_i :

$$IZ_X(B) = \bigcup_{i=1}^N IZ_X(B_i)$$

El complemento del conjunto $IZ_X(B)$ define el esqueleto por zona de influencia geodésica SKIZ:

$$SKIZ_X(B) = X \setminus IZ_X(B)$$

Que presenta todos los puntos equidistantes (en distancia geodésica) a dos componentes del conjunto X . En la figura 5.14 se muestra un ejemplo de SKIZ.

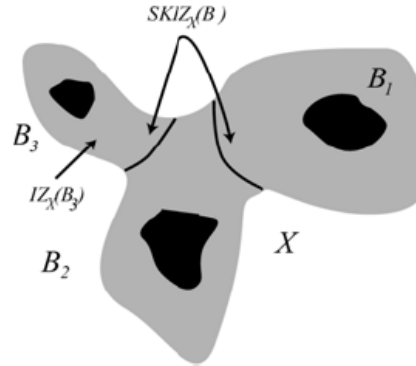


Figura 5.14. SKIZ del conjunto X constituido a partir de las zonas de influencia geodésica IZ .

El SKIZ es la base de la definición de la transformada Watershed.

5.5.2 La transformada Watershed

Las cuencas de una imagen en escala de grises pueden ser consideradas como la zona de influencia de los mínimos regionales de la imagen. En este sentido, las líneas Watershed no son más que un esqueleto por zona de influencia de los mínimos de la imagen:

$$WS(f) = SKIZ_f[RMIN(f)]$$

El significado de la operación de Watershed se encuentra en su notación francesa como la *ligne de partage des eaux* (línea de separación de las aguas). Esta transformación consiste en simular la "inundación numérica" de la imagen (vista como un mapa topográfico) a partir de los mínimos regionales. El agua es introducida desde los mínimos regionales, de manera que en el proceso de inundación hace que las cuencas contiguas se unan. Las líneas de unión representan las fronteras de las regiones homogéneas, éstas son las que constituyen el esqueleto geodésico. El resultado es un conjunto de contornos que identifican las regiones de la imagen.

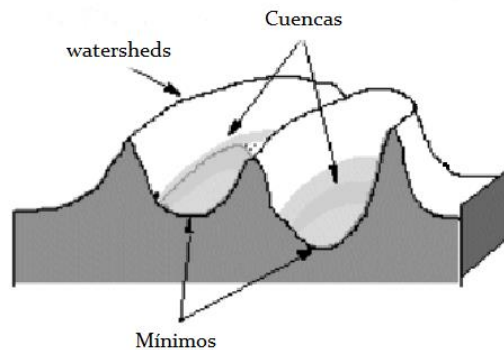


Figura 5.15. Mínimos, cuencas y Watershed en la representación topográfica de una imagen.

El proceso de segmentación morfológica suele partir de la imagen de gradiente, aplicándose a esta última la transformación Watershed, pues se pretende identificar el contorno de los objetos y no su zona de influencia. Los mínimos proceden en este caso de la imagen gradiente. El resultado en imágenes reales suele conducir al fenómeno de la sobresegmentación, porque la presencia de mínimos en esta imagen es bastante alta. Para evitar este inconveniente se suele procesar la imagen con anterioridad y así reducir la cantidad de mínimos que representan una misma zona u objeto. La obtención de éstos mínimos será clave para unos buenos resultados.

5.5.3. Segmentación controlada por marcadores

La idea básica detrás de la segmentación controlada por marcadores es transformar la imagen de entrada de tal manera que las líneas divisorias de la imagen transformada correspondan a bordes significativos de los objetos. La imagen transformada es llamada la función de segmentación. En la práctica, un cálculo directo de las líneas divisorias de la función de segmentación produce una sobre-segmentación que es debido a la presencia de mínimos falsos. En consecuencia, la función de segmentación debe ser filtrada antes de calcular sus líneas divisorias a fin de eliminar todos los mínimos irrelevantes. Cualquier técnica de filtrado puede ser considerada. Sin embargo, la técnica de imposición de mínimos descrita en la sección 4.4.2 es uno de los filtros más apropiados en la mayoría de las aplicaciones. Esta técnica requiere la determinación de una función marcador que

permita marcar los objetos relevantes de la imagen y su fondo. Los marcadores correspondientes se utilizan entonces como el conjunto de mínimos a imponer en la función de segmentación.

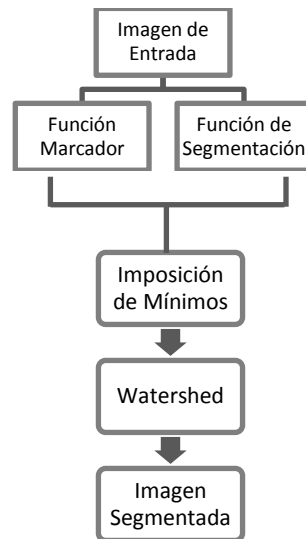


Figura 5.16. Paradigma Morfológico para la segmentación de imágenes.

El esquema de este método se resume en la figura 5.16. Los objetos marcados son extraídos de la imagen utilizando algunos detectores de características. La elección apropiada de detectores de características se basa en un conocimiento a priori acerca de las propiedades de un objeto de la imagen. Características comunes comprenden extremos de la imagen, zonas planas (*i. e.* componentes conexas de píxeles con valores de niveles de grises constantes), zonas de texturas homogéneas, etc. En algunas aplicaciones, los marcadores pueden ser definidos manualmente. Es necesario un marcador por región ya que habrá correspondencia uno a uno entre los marcadores y los segmentos en la partición final. Sin embargo, si la clase de cada objeto marcado es conocida, varios marcadores de la misma clase pueden ser considerados para cada objeto de la imagen. El tamaño de un marcador puede variar de un píxel único a un gran componente conexo de píxeles. En el tratamiento de imágenes con ruido, grandes marcadores se desempeñan mejor que los pequeños (Soille, 2003).

La determinación de la función de segmentación se basa en un modelo para la definición de la frontera de un objeto. Por ejemplo, si los objetos de la imagen se definen como regiones con valores constantes en escala de grises, un operador gradiente morfológico es una herramienta que puede ser usado como marcador, ya que mejorará los bordes de los objetos. Si los objetos de la imagen son regiones con textura homogénea, los operadores que destacan la transición entre dos texturas deben ser considerados.

Los marcadores de objetos son entonces usados como el conjunto de marcadores para imponer en la función de segmentación. Finalmente, las fronteras de los objetos son obtenidas con el cálculo de las líneas divisorias de la función de segmentación filtrada.

5.5.4 Autómatas Celulares

En las investigaciones realizadas por Rueda y Torres (Rueda, Torres, 2013) se define un autómatas celular como un modelo dinámico discreto, representado en un arreglo regular de celdas. Cada una de estas celdas se encuentra en uno de un número finito de estados posibles en un paso de evolución t . El arreglo de celdas puede ser multidimensional. Por cada celda del autómatas, se define un conjunto de celdas adyacentes llamado vecindad. Un estado inicial para el autómatas, es seleccionado asignando un valor escalar o una tupla de valores para cada celda. Se actualiza el estado del autómatas formando una nueva generación de celdas en cada paso de evolución t , de acuerdo a alguna regla fija de evolución. Esta regla de evolución es una función matemática, que es evaluada en cada celda y determina su nuevo estado en términos de su estado actual y los estados de sus celdas vecinas. Típicamente, la regla de evolución se aplica simultáneamente en todo el arreglo de celdas y es invariable a lo largo del tiempo, aunque también pueden considerarse variaciones probabilísticas a las reglas y existen además los autómatas celulares asíncronos, con secciones de independencia temporal (Wolfram, 1983).

Los autómatas celulares fueron primero estudiados por John Von Neumann en 1947, cuando desarrollaba modelos de máquinas capaces de duplicarse a sí mismas, a través de una sucesión de pasos discretos. Los primeros autómatas de Von Neumann fueron modelados en dos dimensiones, con reglas complejas derivadas de la interpretación finita de ecuaciones diferenciales; poseían 29 estados posibles para cada celda. En la década de 1960, Stanislaw Ulam, antiguo colaborador de Von Neumann, que había hecho numerosas sugerencias a sus estudios originales, implementó simulaciones basadas en autómatas celulares y observó que mediante el uso de reglas simples de evolución, el autómatas era capaz de producir comportamientos complejos. Este fenómeno es conocido en las ciencias físicas como emergencia: la aparición de comportamientos complejos a partir de la interacción de agentes con comportamiento individual simple. Stanislaw Ulam argumentó la relevancia de este comportamiento en los autómatas celulares para el estudio de las ciencias naturales, particularmente la biología, señalando la variedad de procesos que operan en base a unidades elementales con estados dependientes del estado de otros elementos que se encuentren en su vecindad (Von Neumann, 1966).

En la década de 1980, Stephen Wolfram populariza el estudio de los autómatas celulares como mecanismos para el modelado de sistemas complejos, a través de journal *Complex Systems*, y son famosos sus estudios postulando que *“cualquier sistema que satisfaga ecuaciones diferenciales puede ser aproximado como un autómatas celular introduciendo diferencias finitas y variables discretas”* (Wolfram, 1983).

En el 2004, Matthew Cooke, investigador de Wolfram Research, publicó una demostración de la conjetura original de Wolfram acerca de la capacidad de una regla de autómatas celulares 1D para funcionar como máquina de cómputo universal, haciendo al autómata programable y capaz de calcular cualquier función computable (Cook, 2004).

En el año 2002 fue publicado un compendio de las investigaciones de Wolfram, acerca de los autómatas celulares llamado “*A New Kind of Science*” (Wolfram, 2002). Los autómatas han sido utilizados para modelar una amplia gama de sistemas dinámicos en variados dominios de aplicaciones. Una de las aplicaciones la podemos encontrar en el procesamiento de imágenes, para realzar bordes y eliminar ruido (Popovici, 2002), segmentar y etiquetar imágenes (Vezhnevets, 2005), (Kauffmann, 2009), (Kim, 2010).

Los resultados de las investigaciones de Rueda y Torres (Rueda, Torres, 2013) son una continuación de la línea de investigación en autómatas celulares deterministas para la segmentación de imágenes, iniciada por Vezhnevets (Vezhnevets, 2005) y continuada por Kim (Kim, 2010).

Formalmente, un autómata celular es un algoritmo discreto en espacio y tiempo que opera en un retículo de elementos $p \in P \subseteq \mathbb{Z}^n$. Es una tripleta $A = (S, N_g, \delta)$, donde S es un conjunto de estados no vacío que puede ser asignado a cada celda, N_g es el sistema de vecindad que afecta a todas las celdas y $\delta: S^{N_g} \rightarrow S$ es la función de transición usada para formar una nueva generación de celdas en cada paso t . La función δ es usada para calcular la transición del estado de una celda p en un paso de evolución t hacia el paso de evolución $t + 1$ en base a los estados de las celdas en la vecindad de p .

La vecindad (N) de un píxel p se define en $\mathbb{R}^2$ como:

$$N = B(p, r) \cap \mathbb{Z}^2$$

Donde $B(p, r)$ es una bola centrada en el píxel p y radio $r \geq 1$. Los sistemas de vecindad N_g usados para la segmentación de imágenes son:

- Vecindad Von Neumann

$$N(p) = \left\{ q \in \mathbb{Z}^n : \|p - q\|_1 := \sum_{i=1}^n |p_i - q_i| = 1 \right\};$$

- Vecindad Moore

$$N(p) = \left\{ q \in \mathbb{Z}^n : \|p - q\|_\infty := \max_{i=1, \dots, n} |p_i - q_i| = 1 \right\};$$

Para ambos casos, la figura 5.17 muestra la celda p a actualizar en verde, y las celdas azules corresponden a la vecindad.

Para tareas de segmentación de imágenes en 2D, el autómata tiene el mismo tamaño en dimensiones x, y que la imagen, y en 3D, la misma dimensión x, y, z que el volumen.

La única variación formal en las reglas de evolución para los autómatas entre el caso 2D y el caso 3D, son los sistemas de vecindad considerados. La figura 5.17(a) muestra el sistema de vecindad Von Neumann utilizados en volúmenes, por su parte la figura 5.17(b) muestra el sistema de vecindad Moore.

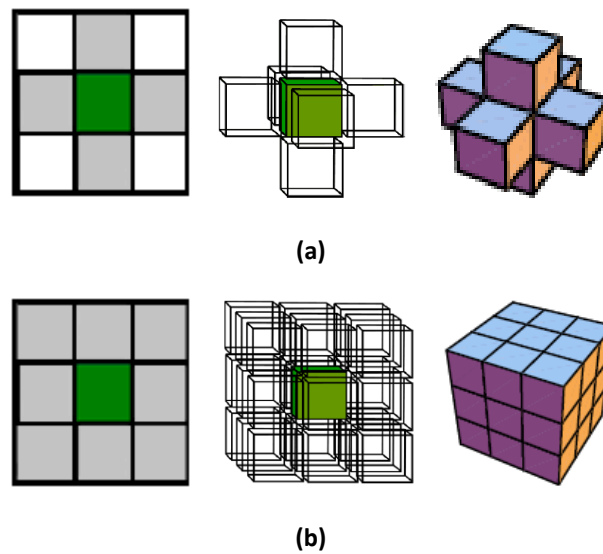


Figura 5.17. Vecindad de celdas. (a) Von Neumann 2D, 3D, (b) Moore 2D, 3D.

Método Growcut

Es una técnica basada en un autómata celular donde se busca, a partir del estado inicial de los píxeles semillas, asignar etiquetas a los píxeles restantes de la imagen. El proceso de etiquetado es una técnica iterativa. El Método de Growcut posee propiedades interesantes, es extensible (permite la construcción de nuevas familias de algoritmos de segmentación con propiedades específicas), interactiva (la segmentación es refinada con cada iteración, el usuario puede observar la evolución y refinar la segmentación), el algoritmo es sencillo tanto en la comprensión como en la aplicación, además permite la ejecución en paralelo (Rueda, 2013).

Para la imagen, el estado de una celda (p), se define como $S = (l_p, \theta_p, \vec{C}_p)$ donde l_p es la etiqueta de la celda actual, $\theta_p \in [0,1]$ es la fortaleza de la celda y $\vec{C}_p$ es el vector con los niveles digitales correspondientes a cada banda del píxel p .

Para una imagen de tamaño $k \times m$, el espacio celular (P) es definido por un arreglo de tamaño $k \times m$. Así hay una correspondencia uno a uno entre un píxel de la imagen y una celda del autómata, por lo tanto p es usado para representar un píxel o una celda.

Antes de aplicar el autómata para la segmentación, se determinan las semillas y por consiguiente su etiqueta y su fortaleza. El estado inicial $\forall p \in P$ es

$$l_p = a, \text{ donde } a \in \mathbb{N}$$

$$\theta_p = b, \text{ donde } b \in [0,1]$$

$$\vec{C}_p = (p_1, p_2, \dots, p_M)$$

Donde p_i es la intensidad del píxel p en las bandas i y M corresponde al número de bandas que posee la imagen.

Capítulo 6. Resultados

La herramienta utilizada para la obtención de los resultados tratados en este capítulo es Matlab 2013b (Ver Apéndice).

Se presenta a continuación los dos procedimientos utilizados para la segmentación de imágenes; el primero por medio de la transformada Watershed; el segundo por el método basado en autómatas celulares, para proceder a evaluar los resultados y hacer una comparación entre las segmentaciones obtenidas. Se esquematiza el procedimiento de la siguiente manera:



Figura 6.1. Proceso para la Segmentación.

Con el uso de imágenes satelitales provenientes de los satélites Francisco de Miranda, Landsat 7 y Landsat 8, en las que se observan zonas de cultivos, zonas pobladas, masas de agua, entre otras, se evalúa la calidad de la segmentación en base a la cantidad de segmentos que arroja, y a las características similares que tenga cada zona segmentada, debe cumplirse un criterio de homogeneidad para cada objeto segmentado tanto en la técnica de Watershed como en el crecimiento de regiones.

El esquema presentado en la figura 6.1 da noción del proceso seguido en este trabajo para la segmentación de las distintas imágenes satelitales. Consiste primeramente en el filtrado de una imagen que se denomina imagen de entrada, posterior a ello, se aplica el gradiente multiescala el cual genera una imagen que pone en evidencia las fronteras entre las regiones presentes en la imagen de entrada, luego se aplica la función de máximo gradiente y se procede a emplear la función para la obtención de las semillas, para hacer uso posteriormente de uno de los dos métodos propuestos, con el fin de conseguir una imagen resultado correspondiente a la imagen segmentada.

Se explica a continuación cada elemento del proceso de manera detallada.

6.1 Imagen de entrada

En el proceso de segmentación la imagen de entrada que suministra el usuario, es de dos tipos como lo muestra la figura 6.2. Es posible, como primera opción, la segmentación de la imagen satelital original, y como segunda hacer un preprocesamiento en base al análisis de componentes principales, que permite reducir entre otras cosas la dimensión de la imagen original. Cuando la imagen de entrada corresponda a un análisis de componentes

principales se trabaja con las tres primeras componentes resultantes que deben poseer un rango de información entre el 90 y 99 por ciento.



Figura 6.2. Tipos de imágenes a procesar.

6.2 Filtrado de la imagen

Las imágenes procedentes de los satélites pueden presentar artefactos (calibración de sensores, atmósfera, ruido, etc.) que dificultan la identificación de los objetos presentes en la superficie de la tierra, es por esto, que se hace necesario aplicar previamente filtros, con el objeto de eliminar ruido o detalles pequeños de poco interés para el usuario, atenuar zonas de totalidad uniforme, resaltar bordes, entre otras. Esto permite mejorar la identificación posterior de los objetos presentes en la imagen, en general adecuarla a para la segmentación reduciendo en cierta medida la sobre-segmentación, y enfatizar cierta información o conseguir un efecto especial en ella. En este trabajo se aplican los siguientes filtros; filtro morfológico alterno secuencial apertura-cierre; reconstrucción dual; gradiente multiescala; Apertura y cierre por Reconstrucción, con la intención de obtener la mayor homogenización de las zonas sin perder información relevante, así como evaluar los resultados de la unificación y/o separación de las regiones. En la figura 6.3 se muestra los diferentes tipos de filtros que se aplican dentro del proceso.

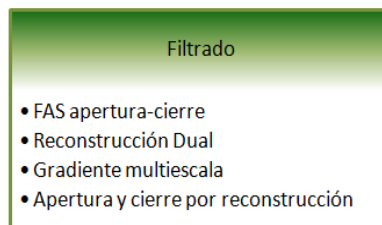


Figura 6.3. Tipos de filtros.

La aplicación de los primeros dos filtros se hacen directamente a la imagen de entrada para posteriormente aplicar el gradiente multiescala. Mientras que el otro método es aplicar directamente a la imagen de entrada el gradiente multiescala para proceder a filtrar por medio de apertura y cierre por reconstrucción. Se muestran en la figura 6.4 los dos pasos de filtrado.

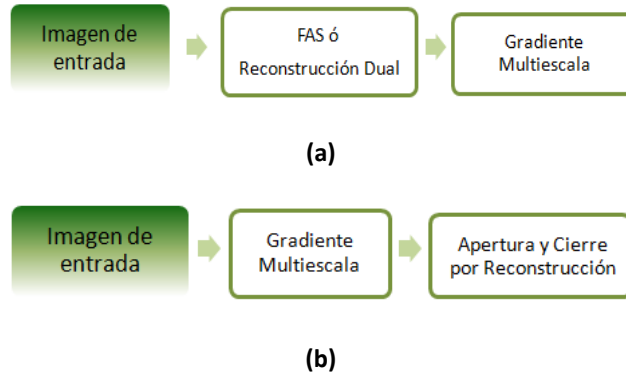
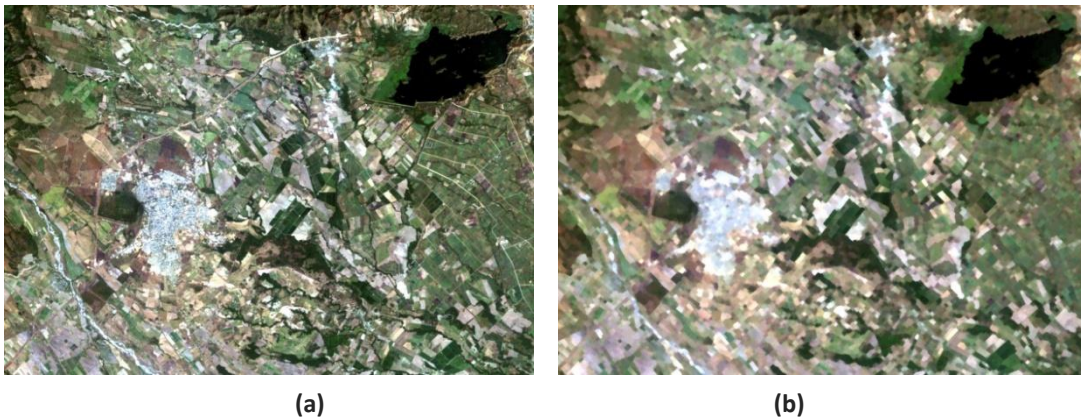


Figura 6.4. Métodos de filtrado. (a) Proceso del Filtro alternativo secuencial y de la reconstrucción dual, (b) Proceso del filtro apertura y cierre por reconstrucción.

6.2.1 Filtro Morfológico Alternativo Secuencial

Se hace uso de un filtro alternativo secuencial (FAS), definido por la secuencia de una composición de cierre y apertura $\varphi_i(\gamma_i(f))$, con $i = 3$, donde el elemento estructurante es de tamaño creciente, e inicialmente es un diamante de radio 1. Éste filtro permite reducir la variabilidad de una imagen, y conservar los bordes de las zonas o elementos que se encuentran presentes, siempre y cuando sean mayores al elemento estructurante.

Para mostrar el efecto de éste y otros filtros, se tomará como imagen de entrada una subimagen multicanal de tamaño 850x1221x6 proveniente del satélite Landsat 7, correspondiente a la zona de Acarigua Estado Portuguesa. La figura 6.5 presenta la aplicación del filtro secuencial, y un acercamiento para apreciar mejor los resultados. El principal efecto de este filtro es el suavizado de las imágenes reduciendo la textura de los objetos sin perder el borde.



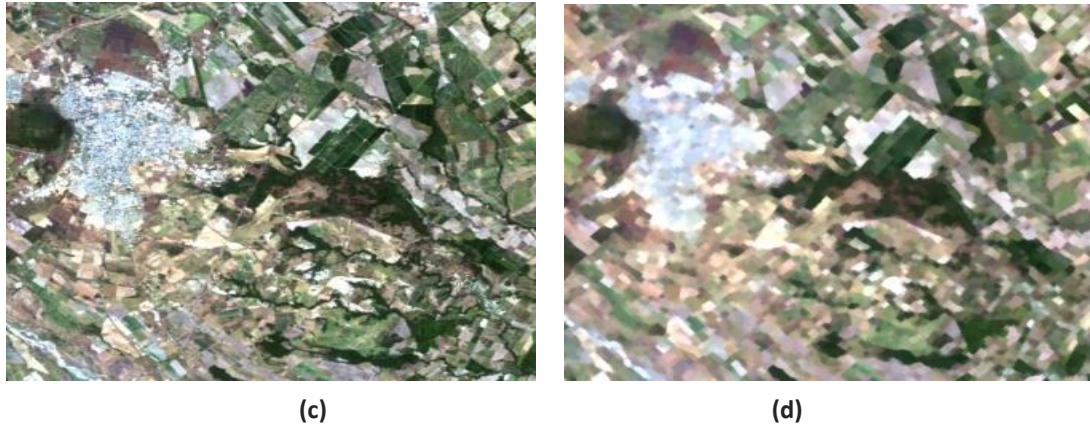


Figura 6.5. Filtro morfológico alterno secuencial Apertura-Cierre. (a) Imagen Original combinación [R G B], (b) FAS de 3 iteraciones, (c) Zoom de la imagen original, (d) Zoom del filtro.

6.2.2 Reconstrucción Auto-Dual

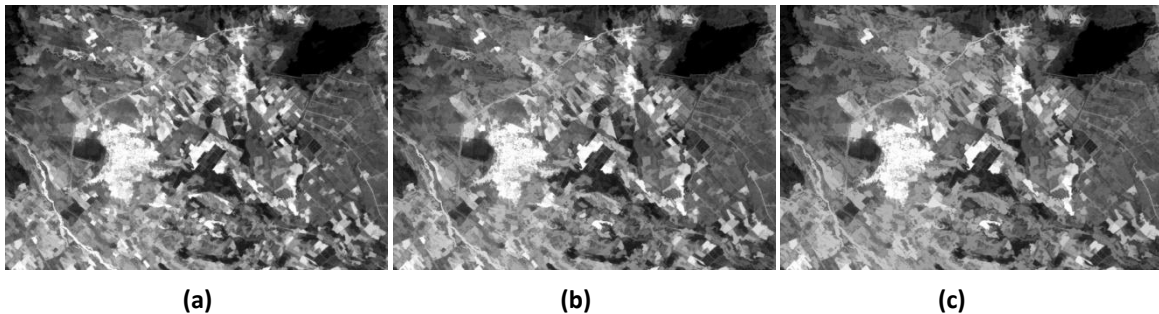
Como se menciona en el apartado 4.3.2 la reconstrucción autodual actúa sobre las componentes oscuras y claras a la vez, además de ser un filtro que restaura la información de los contornos después de un proceso de filtrado inicial $\psi(f)$, en este caso específico el filtro inicial es un filtro alterno secuencial como el descrito en el apartado anterior, que consta de la composición de un cierre con una apertura. La representación de la reconstrucción autodual definida está dada de la siguiente manera:

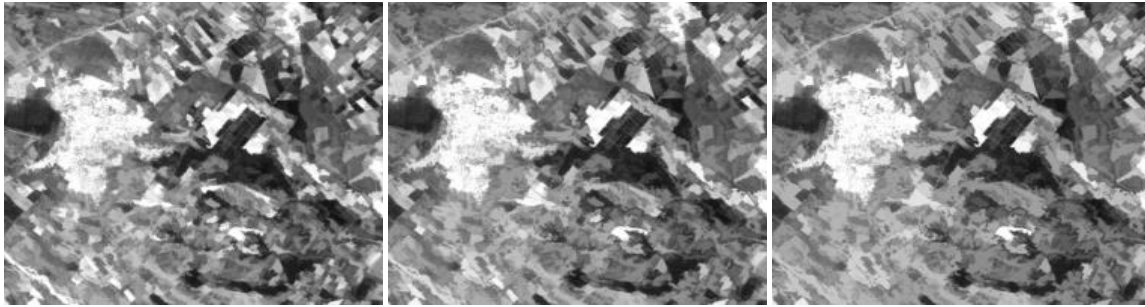
$$R^\dagger(\psi(f)|f)$$

Donde:

$\psi(f) = \varphi_i \gamma_i(f)$, y f es la imagen de entrada.

Se aplica a la imagen de la figura 6.5(a) la reconstrucción anterior con un número de iteraciones $i = 3, 5, 7$ donde el elemento estructurante tiene forma de diamante y tamaño inicial radio 1. Se muestra el resultado en la figura 6.6 de la primera banda para constatar los cambios entre las componentes claras y oscuras.





(d)

Figura 6.6. Reconstrucción dual. (a) 3 iteraciones, (b) 5 iteraciones, (c) 7 iteraciones, (d) Zoom de las reconstrucciones.

Se puede dificultar observar los cambios en las imágenes anteriores, por lo que se presenta en la figura 6.7 una gráfica que permite relacionar los niveles digitales de los píxeles presentes en la fila 310 de cada una de ellas, para evaluar las diferencias que se producen en los niveles digitales. Por su parte se muestra la visualización a color de la reconstrucción autodual mencionada de la imagen en la figura 6.8. La figura 6.8(a) muestra una línea roja correspondiente a los píxeles graficados en la figura 6.7.

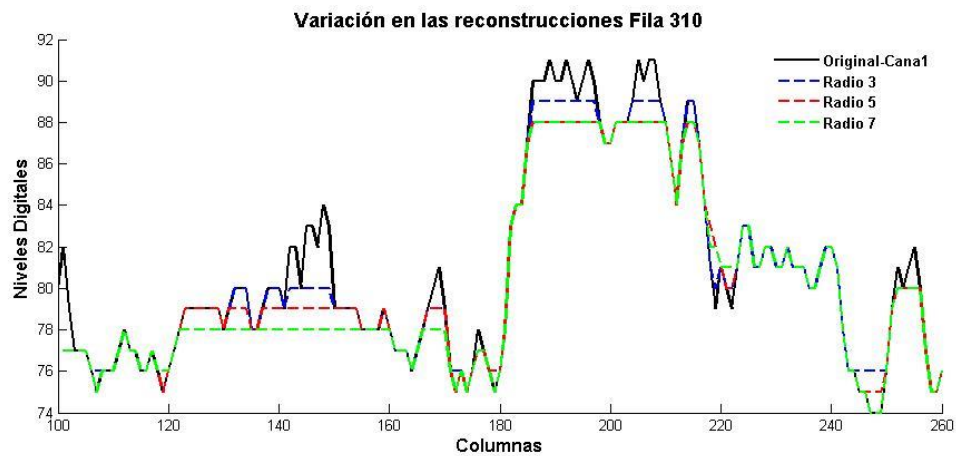


Figura 6.7. Gráfica de la Reconstrucción en la fila 310.



(a)

(b)

(c)

Figura 6.8. Reconstrucción dual a color. (a) FAS de 3 iteraciones, (b) FAS de 5 iteraciones, (c) FAS de 7 iteraciones.

6.2.3 Gradiente Multiescala

Los elementos u objetos presentes en las imágenes satelitales se caracterizan por poseer niveles digitales más o menos homogéneos, delimitados por un borde. El gradiente morfológico, permite discriminar las zonas homogéneas con un gradiente muy pequeño. Para reducir la cantidad excesiva de zonas homogéneas correspondientes a elementos presentes de tamaño reducido, se utilizó el gradiente multiescala (ME) descrito en la ecuación 3.1, con 5 iteraciones partiendo de un elemento estructurante con forma de diamante y radio 1 que es dilatado en cada iteración. Con el método de la figura 6.4(a) se aplicó el gradiente multiescala a la banda 1 de la figura 6.5(a), sus resultados se presenta en la figura 6.9.

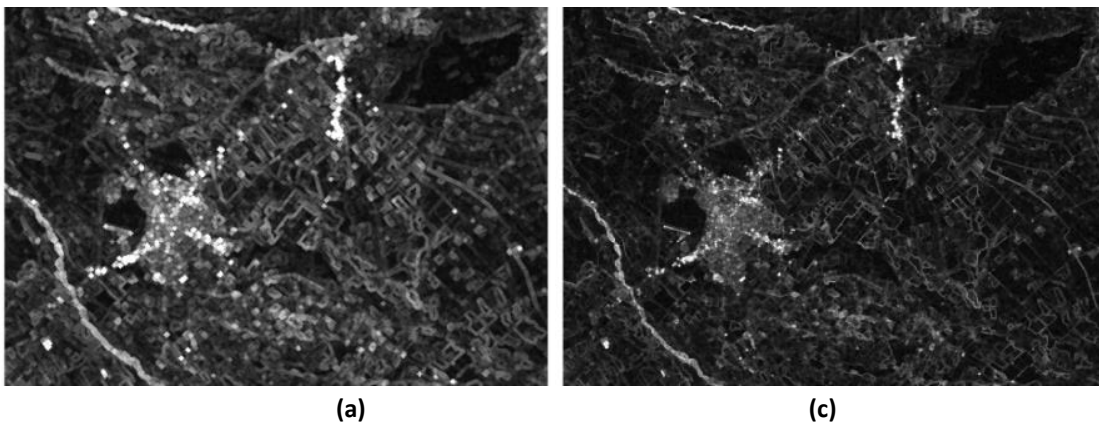


Figura 6.9. Ejemplos de Gradientes de un canal. (a) Gradiente simétrico $\delta - \varepsilon$, ES con forma de diamante 5x5, (b) Gradiente multiescala de 5 iteraciones.

Para las imágenes multicanal se aplica el gradiente multiescala a cada una de sus bandas, por lo tanto la imagen gradiente tiene las mismas dimensiones que la imagen multicanal de entrada. Al aplicar el gradiente multiescala de las primeras 6 bandas se obtiene una nueva imagen de seis bandas, como se muestra en la figura 6.10.

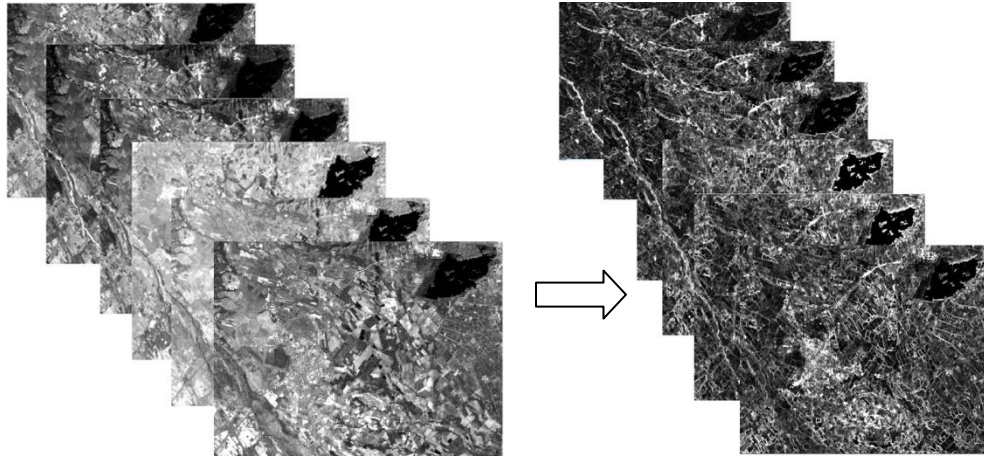


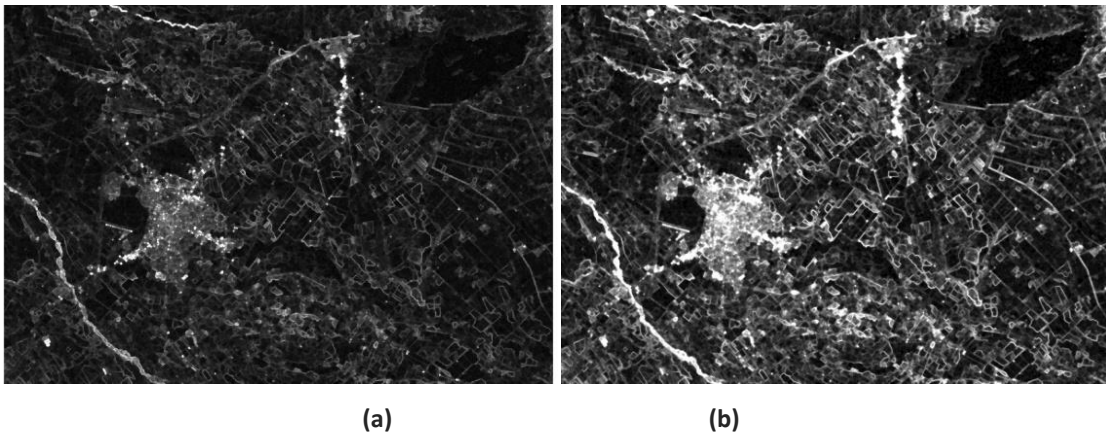
Figura 6.10. Gradiente multiescala para una imagen multicanal.

6.2.4 Apertura y cierre por reconstrucción

El filtro de apertura y cierre por reconstrucción se aplica una vez obtenida la imagen de gradiente multiescala. Los bordes del gradiente indican la separación entre las zonas homogéneas, por lo que es posible suavizar esas regiones para eliminar algunos mínimos. Es por esto, que se aplica un filtro de apertura por reconstrucción y posteriormente un cierre por reconstrucción que se denota de la siguiente manera:

$$\varphi_R(\gamma_R(f))$$

Donde, $\gamma_R = R_f^\delta[\varepsilon(\bar{\rho}(f))]$, es decir, la apertura por reconstrucción del gradiente multiescala de la imagen de entrada.



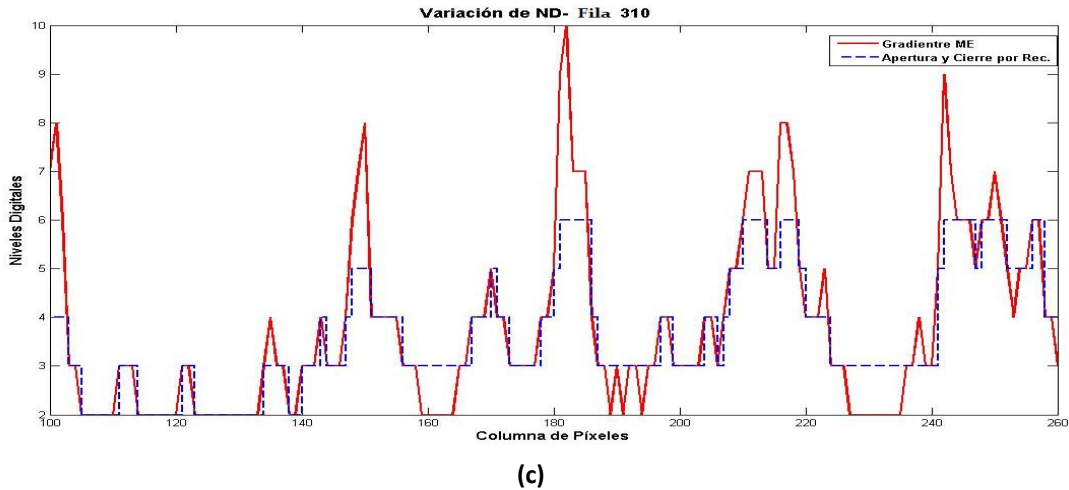


Figura 6.11. Filtrado del Gradiente Multiescala. (a) Gradiente ME 5 iteraciones, (b) Apertura y cierre por reconstrucción, (c) Gráfica de variación entre los filtros.

La figura 6.11 muestra la aplicación del filtro apertura y cierre por reconstrucción al gradiente multiescala, con un elemento en forma de diamante y 5 iteraciones. A través de la gráfica podemos observar cómo se homogenizan los niveles digitales de la columna 310 de ambas imágenes, permite eliminar así una cantidad considerable de mínimos, que es de gran ayuda para el siguiente paso del proceso de segmentación, referente a la obtención de las semillas. Como cada mínimo se asocia a un segmento, la selección de los más relevantes contribuye a disminuir la sobre-segmentación.

Para terminar de contrastar el proceso de filtrado de la imagen de entrada de dimensiones 850x1221x6, vamos a considerar en la tabla 6 los tiempos empleados en el cómputo de cada filtro, incluyendo el filtro gradiente multiescala. Estos cálculos se hicieron en un computador con procesador Intel Xeon E5405 (2.00 GHz), y memoria de 8GB de RAM.

FILTROS	Tiempo (segundos)
Alternativo Secuencial $\phi\gamma \rightarrow$ Gradiente ME	1.45
Reconstrucción Dual $\rightarrow$ Gradiente ME	53.74
Apertura y Cierre por Reconstrucción del Gradiente ME	4.41

Tabla 6. Tiempo de Cálculo por tipo de Filtros.

6.3 Gradiente Máximo

El proceso anterior al gradiente máximo es el filtro de gradiente multiescala, que provee los bordes relevantes en cada una de las bandas de la imagen de entrada. Al introducir la función de máximo gradiente se pretende representar en una sola matriz 2D todos los bordes detectados a lo largo de la dimensión Z . Así, a todo píxel de la nueva imagen gradiente máximo (B) de posición (x, y) se le asigna el máximo nivel digital que posee en cada una de las bandas del gradiente multiescala. Es decir

$$P_B^{(x,y)} = \max_{1 \leq i \leq n} P_{b_i}^{(x,y)}$$

Donde n es el número de bandas.

En la figura 6.12 se muestra la aplicación de la función a una imagen gradiente multiescala de 6 bandas.

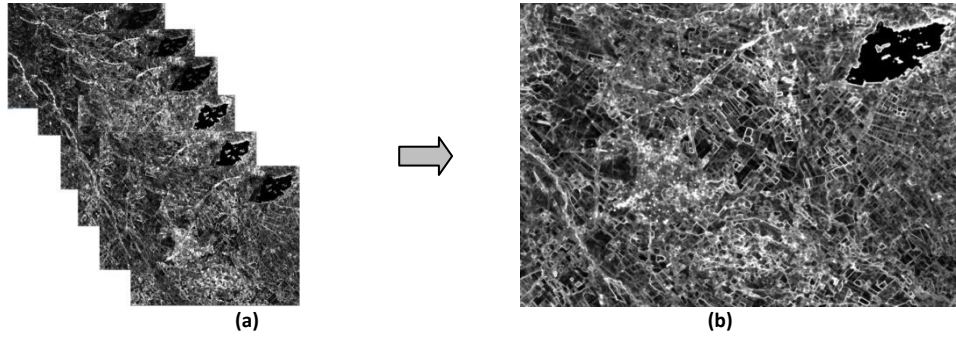


Figura 6.12. Máximo gradiente de una imagen. (a) Gradiente ME, (b) Máximo Gradiente.

7.4 Obtención de semillas

Una vez obtenidas las regiones o zonas delimitadas es necesario representar cada una de ellas, a estos representantes se les denominan semillas y para hallarlas se hace uso de la función de mínimos regionales, definida en el apartado 4.4.1.

Los mínimos regionales de la imagen de máximo gradiente resultan ser buenos representantes aunque pueden hallarse varios mínimos en una misma región, lo ideal sería encontrar una semilla por zona. La figura 6.13 muestra los niveles digitales de los píxeles ubicados en la fila 315 y entre las columnas 1000 y 1100, se evidencia en ella los mínimos resultantes.



Figura 6.13. Gráfica de los niveles digitales del máximo gradiente.

Los mínimos regionales hallados en la imagen gradiente máximo de cada filtro se aprecia en la figura 6.14. Las imágenes muestran la imposición de mínimos a la imagen gradiente. En la práctica cada mínimo genera una región, si hay demasiados mínimos contribuyen a la sobre-segmentación de la imagen.

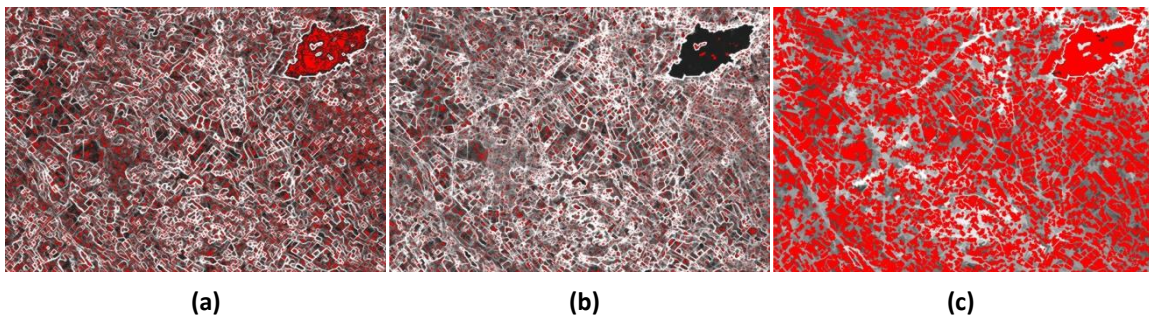


Figura 6.14. Mínimos regionales impuestos a una imagen de gradiente máximo. (a) Gradiente máximo de un FAS, (b) Gradiente máximo de una Reconstrucción Autodual, (c) Gradiente máximo de una Apertura y cierre por reconstrucción.

6.5 Métodos de segmentación

Existen muchos paradigmas para la segmentación de imágenes, sin embargo el objeto de estudio de este trabajo es el método de crecimiento de regiones por medio de la técnica de la transformada Watershed (líneas divisorias de agua) y la aplicación de autómatas celulares.

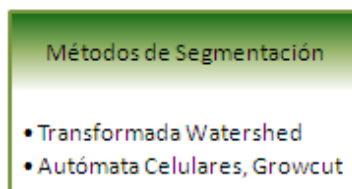
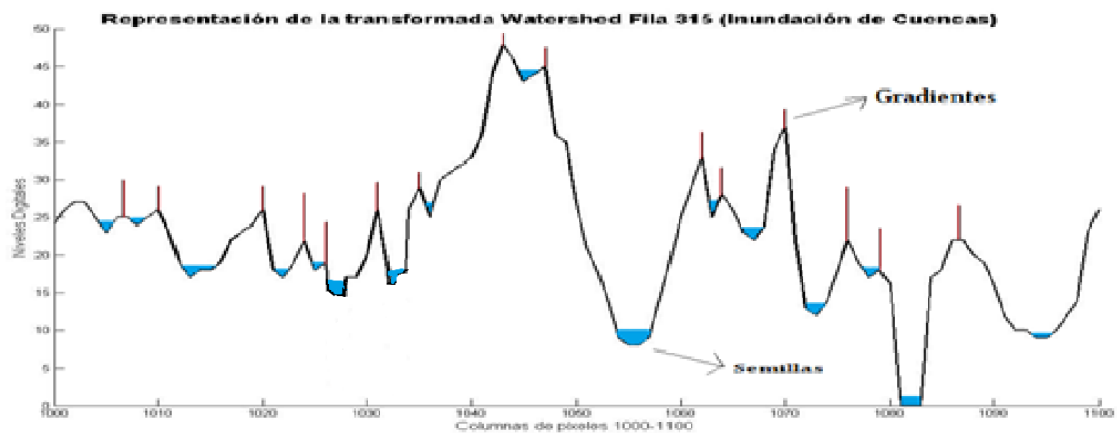


Figura 6.15. Métodos aplicados de Segmentación.

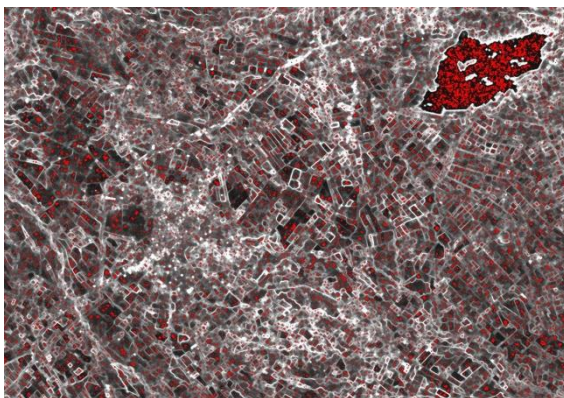
Una vez obtenidas las semillas, se procede a la elección de uno de los dos métodos propuestos para la segmentación, a continuación se explican los parámetros y funciones usados en cada técnica.

6.5.1 Aplicación de la Transformada Watershed

Para aplicar la transformada Watershed, como se menciona en la sección 5.5.2 se requiere de la definición de un esqueleto por zona de influencia, y de las cuencas. En este caso el esqueleto está dado por el gradiente máximo y las cuencas están definidas por la obtención de las semillas que corresponden a los mínimos regionales. Se toma como parámetro en todas las aplicaciones de la transformada Watershed la conectividad-8. En la figura 6.16 se muestra el resultado de la aplicación de la transformada Watershed de la figura 6.14(b).



(a)



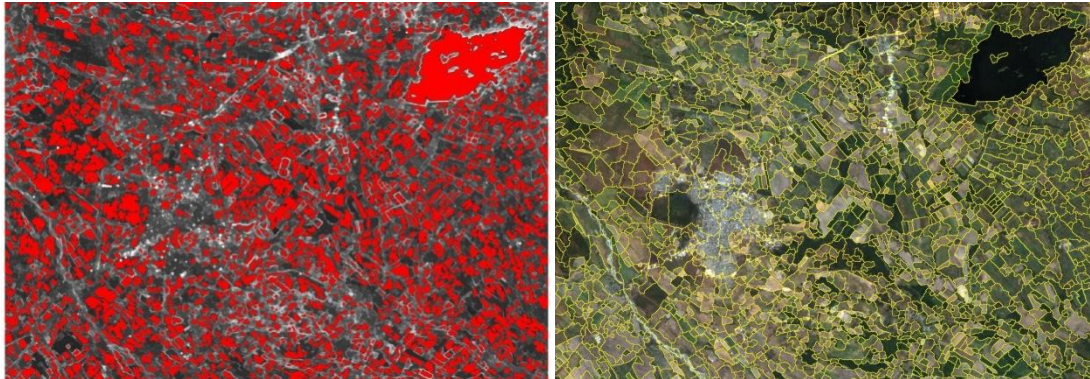
(b)



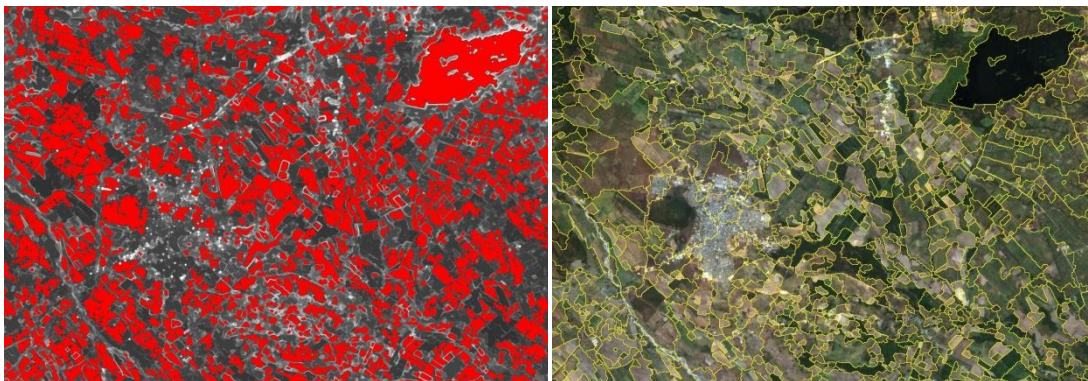
(c)

Figura 6.16 Transformada Watershed del gradiente máximo. (a) Representación de la inundación de cuencas, (b) Imposición de mínimos a la imagen gradiente máximo, (c) Superposición de la segmentación obtenida a la imagen original.

Por cada semilla en la imagen se genera un segmento, y se produce en la mayoría de los casos la sobre-segmentación como lo muestra la figura 6.16(c), una reducción de éste conjunto se da con la ayuda de la función de mínimos extendidos. Una vez obtenido el nuevo conjunto de semillas se emplea la función de imposición de mínimos para aplicar la transformada nuevamente. Para mostrar un ejemplo, se presenta en la figura 6.17 los resultados de la transformación Watershed aplicando la función de mínimos extendidos para distintos valores de h .



(a)



(b)

Figura 6.17. Uso de la función de Mínimos Extendidos. (a) Resultados para $h=5$, (b) Resultados para $h=8$.

6.5.2 Aplicación del Growcut

En el método de crecimiento de regiones sin supervisión se incorpora una estructura basada en un autómata celular que utiliza características de la imagen, tal como los valores de intensidad de los píxeles, para deducir las reglas del estado de transición. Las etiquetas iniciales y el número de píxeles semillas se derivan aleatoriamente desde el espacio de los enteros positivos. Antes de aplicar el autómata, se declaran las semillas y

por consiguiente se determinan su etiqueta y fortaleza. Es decir, se requiere en los parámetros del autómata un estado inicial para las semillas $S = (l_p, \theta_p, \vec{C}_p)$ donde, en cada una de las iteraciones, la fortaleza (θ_p) y la etiqueta (l_p) de las celdas se actualizan considerando, la etiqueta y la fuerza de ataque de las celdas vecinas.

La fortaleza de una celda es una función basada en las características de la imagen y se utiliza para definir la función de transición para la actualización de las etiquetas de las celdas en cada paso de tiempo. Para el estado inicial, la fortaleza de las semillas es $\theta = 1$, y para las celdas que no son parte de las semillas el valor de la fortaleza es $\theta = 0$. Las celdas semillas deben conquistar las celdas vecinas asignándole una fortaleza siguiendo las reglas del autómata.

A continuación se explica el proceso de etiquetado de las semillas, no sin antes aclarar que las celdas (píxeles) que no pertenecen al grupo de las semillas no tienen ningún impacto en el proceso de etiquetado.

Etiquetado

Se propone establecer los valores de las etiquetas a través del estudio de las firmas espectrales. El método consiste en generar N grupos de acuerdo a un factor de brillo obtenido por la suma de los canales, posteriormente se analiza cada grupo mediante estadísticas marginales por banda y se determinan los valores de las etiquetas. La principal ventaja del método es que permite relacionar las etiquetas de las semillas a pesar de su distancia espacial. Se explica a continuación el método de forma detallada.

Sea P un espacio celular definido por un arreglo de dimensiones $m \times n \times k$ y $\bar{S} = \{p^{(1)}, p^{(2)}, \dots, p^{(r)}\}$ el conjunto de los píxeles semillas, tal que $r \in \mathbb{N}$, se construye el vector

$$\bar{S}_{\Sigma} = (s^1, s^2, \dots, s^r)$$

Donde, $s^i = p_{b_1}^{(i)} + p_{b_2}^{(i)} + \dots + p_{b_k}^{(i)}$, $1 \leq i \leq r$. Es decir, la suma de intensidades en cada una de las bandas de la semilla p^i .

Como primer orden de clasificación se agrupan los píxeles por la suma de los niveles digitales que poseen en cada una de las bandas (s^i). Se calculan los percentiles al 2% y 98% de los valores del vector suma $\bar{S}_{\Sigma}$, y se divide su rango en $N = 10$ grupos iguales. Luego, se establece como primer grupo de clasificación los datos por debajo del percentil al 2% y como último grupo los datos por encima del percentil al 98%, con el objetivo de darle menos peso a las colas de la distribución.

A modo de ejemplo se presenta la figura 6.18, que muestra la clasificación de primer orden para una imagen de resolución radiométrica de 8 bits, el vector $\bar{S}_2$ en este caso varía entre 0 y 255x3.

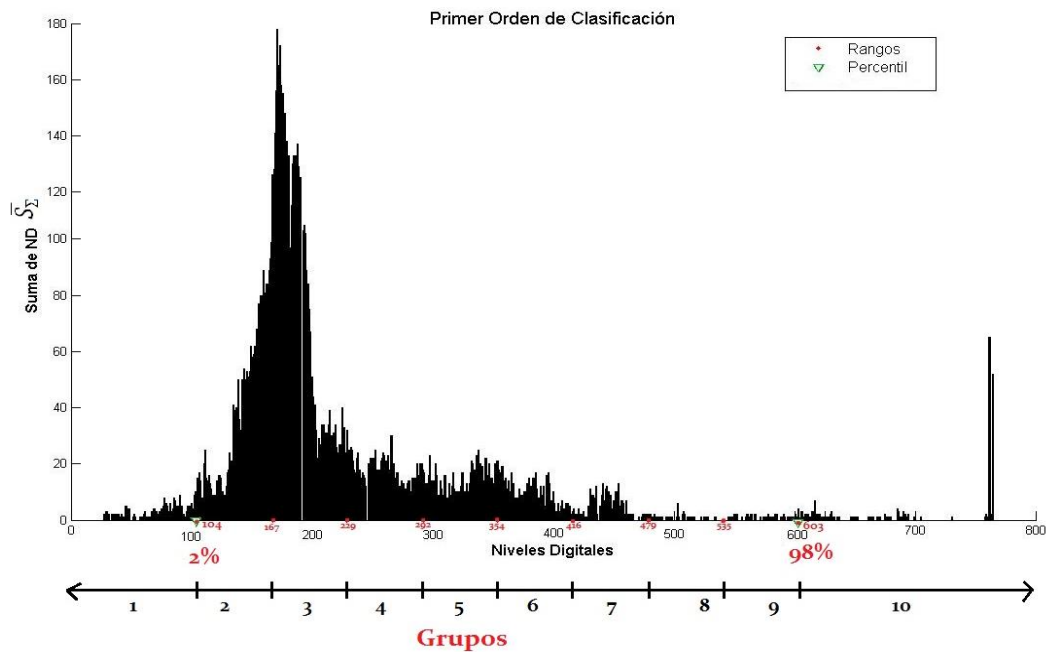


Figura 6.18. Primer orden de clasificación.

El segundo orden de clasificación es el estudio de los 10 grupos por separado, ya que, a pesar de que coincidan los s^i , puede que la distribución de sus valores en las bandas sea distinta. Es decir, existen distintos $\vec{C}_{p^x}$ (vector con sus niveles digitales correspondientes en cada banda) tal que la suma por canales (s^x) coinciden. Como ejemplo se muestra en la figura 6.19 la representación gráfica de la distribución de las semillas pertenecientes al grupo 1 de la figura 6.18, cuyas vectores de suma coinciden en $s^x = 70$.

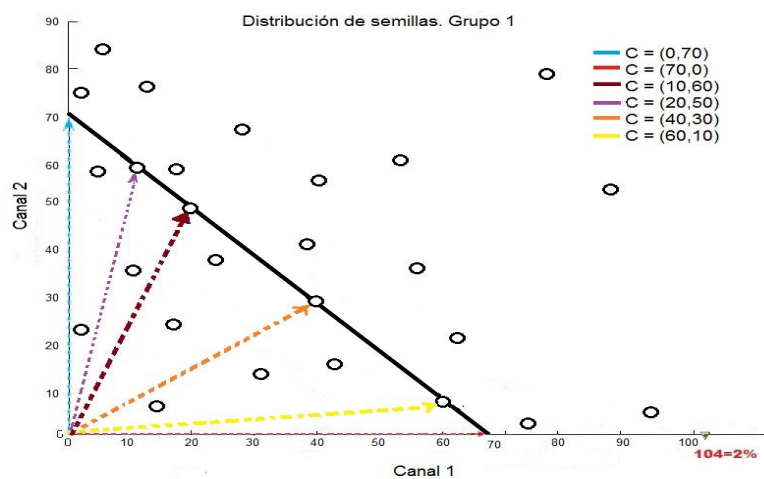


Figura 6.19. Distribución de semillas. Representación de los vectores $\vec{C}$ en $s^x = 70$.

La determinación de los grupos se hace mediante análisis de las estadísticas marginales por cada banda. Se identifican los máximos y mínimos de los histogramas por cada canal y se identifican los grupos que reciben etiquetas distintas.

En la figura 6.20(a) se muestra la distribución de píxeles y la proyección de la distribución de dos bandas. Se busca discriminar, por el peso de los valores en cada canal. De este modo, se hallan la cantidad de máximos y mínimos en la distribución por banda, si se encuentra un mínimo entre dos máximos se establece un umbral (μ_j) y se generan nuevas etiqueta (e_t) para los píxeles mayores y menores a los distintos umbrales. Un ejemplo de ello se aprecia en la figura 6.20(b). Si resulta que la distribución de los píxeles en una banda es de un solo máximo entonces solo se genera una etiqueta.

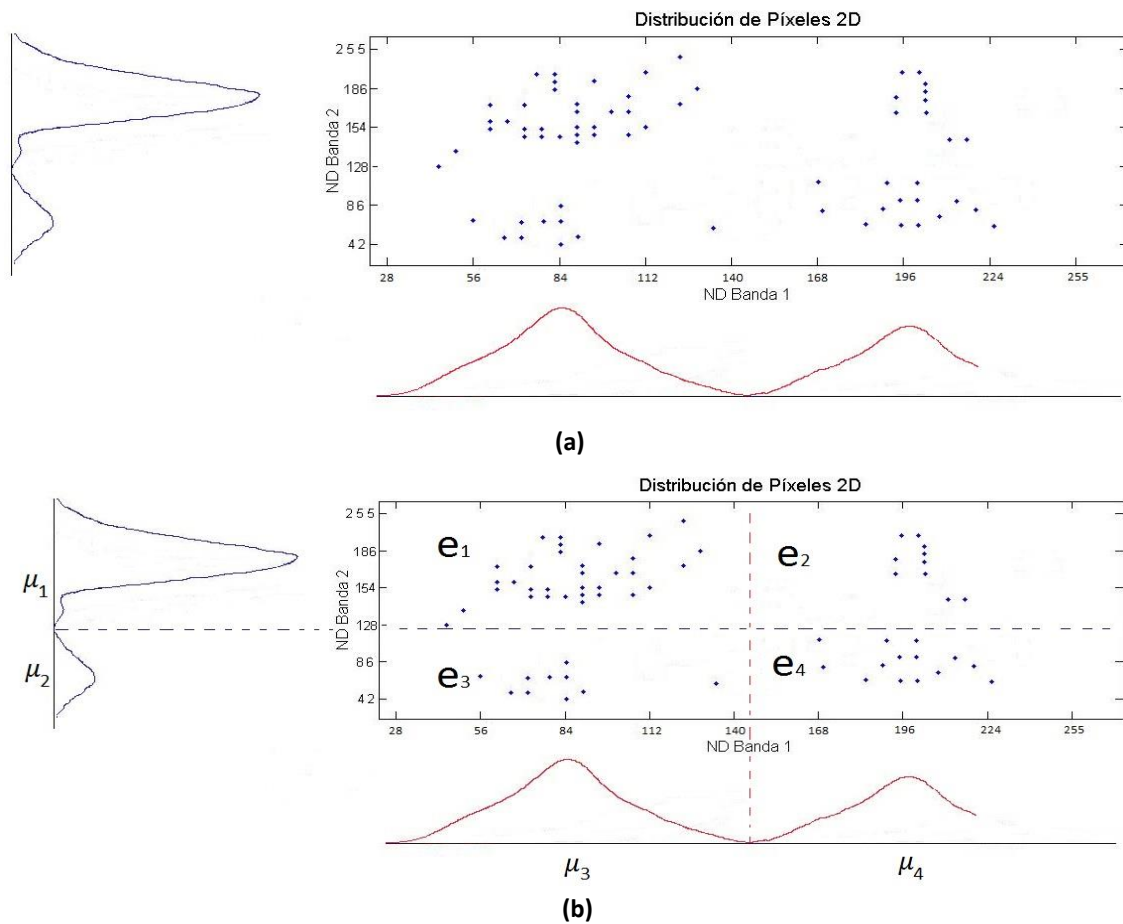


Figura 6.20. Proceso de etiquetado, segundo orden.

Después de obtener el vector $e = (e_1, e_2, \dots, e_j)$ de etiquetas en el segundo orden de clasificación, se analizan las firmas espectrales de los píxeles perteneciente a cada una de ellas en los dos pasos siguientes:

1. Se descartarán las semillas donde las firmas espectrales estén distanciadas en un cierto umbral ρ_1 , de la moda de los datos, con el fin de obtener una representación más o menos uniforme de cada etiqueta. En este sentido, se halla el vector moda de la siguiente manera:

Sea X la matriz de semillas donde las filas están definidas por sus firmas espectrales en cada banda,

$$X = \begin{bmatrix} p_{b_1}^{(1)} & p_{b_2}^{(1)} & \cdots & p_{b_k}^{(1)} \\ p_{b_1}^{(2)} & p_{b_2}^{(2)} & \cdots & p_{b_k}^{(2)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{b_1}^{(r)} & p_{b_2}^{(r)} & \cdots & p_{b_k}^{(r)} \end{bmatrix}_{r \times n}$$

Se denota $M = (m_1, m_2, \dots, m_k)$, como el vector cuyas componentes son la moda (M_o) de las columnas de la matriz X , es decir, el valor con mayor frecuencia en cada banda. Por lo tanto

$$\begin{aligned} m_1 &= M_o(p_{b_1}^{(1)}, p_{b_1}^{(2)}, \dots, p_{b_1}^{(r)}) \\ m_2 &= M_o(p_{b_2}^{(1)}, p_{b_2}^{(2)}, \dots, p_{b_2}^{(r)}) \\ &\vdots \\ m_k &= M_o(p_{b_k}^{(1)}, p_{b_k}^{(2)}, \dots, p_{b_k}^{(r)}) \end{aligned}$$

Una vez que se obtiene el vector M , se calcula el vector (D) correspondiente a la distancia euclidiana entre el vector M y cada firma espectral de las semillas.

$$D = (d_{M,p^{(1)}}, d_{M,p^{(2)}}, \dots, d_{M,p^{(r)}})$$

Donde $d_{M,p^{(i)}}$, es la distancia euclidiana entre M y la semilla p^i para $i = 1, \dots, r$.

Luego se establece un umbral:

$$\rho_1 = \bar{x} + 0,7\sigma$$

Tal que $\bar{x}$ es la media aritmética de D , y σ su desviación estándar. En el caso donde $d_{M,p^{(i)}} > \rho$, se elimina la etiqueta del píxel $p^{(i)}$ y deja de pertenecer al conjunto de las semillas. De esta manera los píxeles pertenecientes a una etiqueta son más o menos uniformes con respecto a las firmas espectrales por banda. En la gráfica de la figura 6.21 se muestra distancia entre los píxeles de una banda y se muestra las etiquetas a ser eliminadas por encontrarse fuera del umbral.

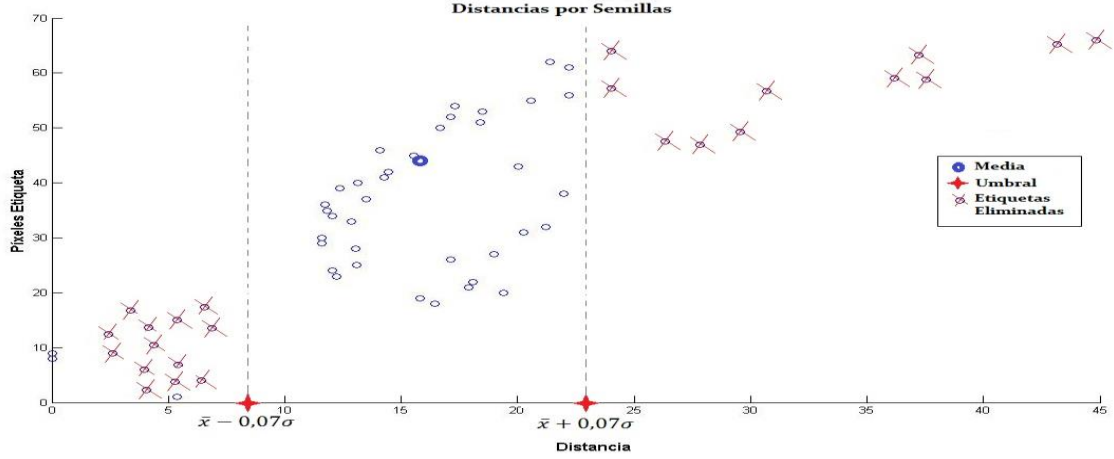


Figura 6.21. Eliminación de semillas alejadas de la moda.

2. Se igualarán entre las semillas más cercanas el valor de las etiquetas donde las firmas espectrales estén próximas en un cierto umbral ρ_2 . Una vez modificado el vector e en el primer paso, se consideran las etiquetas donde la distancia entre las firmas espectrales de los píxeles pertenecientes a ellas sean muy cercanas. Se calcula la distancia euclidiana entre dos etiquetas adyacentes:

$$\bar{D} = (d_{e_1, e_2}, d_{e_2, e_3}, \dots, d_{e_{t-1}, e_t})$$

Donde

$$d_{e_1, e_2} = \sqrt{(p_{b_1}^{(1)} - p_{b_1}^{(2)})^2 + (p_{b_2}^{(1)} - p_{b_2}^{(2)})^2 + \dots + (p_{b_k}^{(1)} - p_{b_k}^{(2)})^2}$$

$$\vdots$$

$$d_{e_{t-1}, e_t} = \sqrt{(p_{b_1}^{(t-1)} - p_{b_1}^{(t)})^2 + (p_{b_2}^{(t-1)} - p_{b_2}^{(t)})^2 + \dots + (p_{b_k}^{(t-1)} - p_{b_k}^{(t)})^2}$$

Y se establece el umbral:

$$\rho_2 = 0,01 \max(\vec{C}_{p^x}) \sqrt{k}$$

Las etiquetas donde la distancia $d_{e_i, e_{i+1}} < \rho_2$ obtendrán el mismo valor, de manera que estas semillas pasan a representar un mismo tipo de región.

Una vez que todas las semillas poseen su estado inicial $S = (l_p, \theta_p, \vec{C}_p)$, se procede a establecer la función de transición $\delta: S^{N_g} \rightarrow S$ del autómata celular que es usada para

formar una nueva generación de celdas en cada paso t , la vecindad N_g empleada es la vecindad-4, se explica a continuación el proceso del mismo:

Ataque

Este método se inicia desde los píxeles semilla y trata de etiquetar todos los demás píxeles de la imagen. Cuando la fuerza de la celda actual (p) es menor que la fuerza de ataque de la celda vecina (q), entonces la fuerza de ataque y la etiqueta de q se establecen como la fuerza y la etiqueta de p . Atacar la fuerza de p depende de la fuerza θ_p y la distancia entre los vectores $\vec{C_p}$ y $\vec{C_q}$. Las etiquetas y la fuerza de las celdas se actualizan, en el paso $t + 1$, utilizando la Regla de evolución del Autómata siguiente:

La etiqueta y la fuerza de p en el paso t se denota como l_p^t y θ_p^t respectivamente.

1. *for* $p \in P$ //para todo píxel
2. $l_p^{t+1} = l_p^t$; $\theta_p^{t+1} = \theta_p^t$ //se copia el estado previo
3. *for* $q \in N_g(p)$ //para todo píxel vecino
4. *if* $g(||\vec{C_p} - \vec{C_q}||_2) \cdot \theta_q^t > \theta_p^t$
5. $l_p^{t+1} = l_q^t$;
6. $\theta_p^{t+1} = g(||\vec{C_p} - \vec{C_q}||_2) \cdot \theta_q^t$
7. *end if*
8. *end for*
9. *end for*

Donde g es una función delimitada al conjunto $[0,1]$ y definida como

$$g(x) = 1 - \frac{x}{\max ||\vec{C}||_2}$$

A mayor cercanía de las firmas espectrales entre las celdas vecinas, mayor es la fuerza de ataque de una celda q a su vecina defensora p . Si la fuerza de ataque $g(||\vec{C_p} - \vec{C_q}||_2) \cdot \theta_q$, es mayor que la fuerza de la celda defensora θ_p , la celda defensora p es conquistada y su estado para el paso de evolución $t + 1$ se actualiza con la etiqueta de la celda atacante q . La nueva fortaleza de p es la fuerza del ataque con que fue conquistada.

Este cálculo continúa hasta que el autómata converge a una configuración estable, es decir, cuando no existen cambios en el estado de las celdas.

Para lograr segmentar la imagen, debe haber al menos dos celdas con etiquetas distintas, y fuerzas asignadas antes que el autómata empiece su evolución ■

Una vez detallado los procedimientos propuestos para obtener la segmentación de una imagen se presentan algunos resultados para posteriormente ser evaluados.

En la técnica de segmentación por medio de autómatas celulares se debe tener en cuenta el método para la obtención de semillas y la definición de los estados iniciales de cada una de ellas, que en este estudio se realiza por medio del análisis estadístico de las firmas espectrales, así como el establecimiento de la regla de transición que es elemental en el proceso de crecimiento de las regiones. La figura 6.22 muestra los resultados al aplicar esta técnica en una subimagen que corresponde a la zona del lago de Maracaibo de dimensiones 934x971x7, resolución radiométrica de 16 bits, y proveniente del satélite Landsat 8.

La figura 6.22(b) representa en escala de grises las fortalezas finales de las celdas, el blanco para referir los píxeles de fortaleza 1, y el negro para aquellos que poseen fortaleza en el intervalo $[0.93,1)$, se evidencia que los píxeles ubicados en la frontera entre dos regiones distintas poseen fortaleza cercana a 1 (representado en grises), mientras que los píxeles dentro de las regiones poseen la máxima fortaleza, esto se debe a la distancia con respecto a las semillas y al reajuste de su valor en el proceso de invasión. En la figura 6.22(c) se puede observar la preclasificación obtenida con respecto a las etiquetas, generadas independientes de la distancia espacial, es decir, se aprecia que las zonas que poseen el mismo color simbolizan un mismo tipo de objeto.

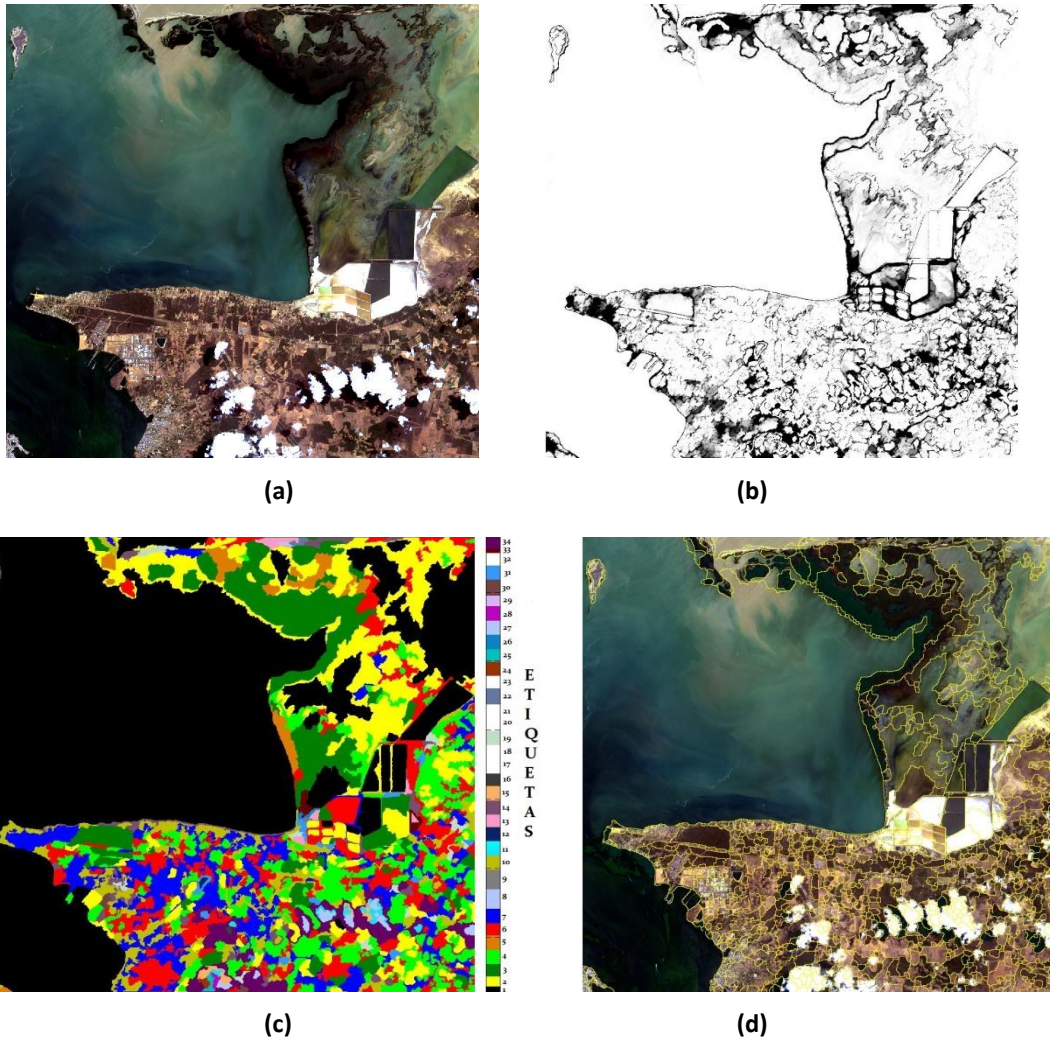


Figura 6.22. Resultados de la Segmentación por autómatas celulares. (a) Imagen Original, (b) Representación de la fortaleza, (c) Regiones y etiquetas obtenidas, (d) Superposición de bordes de las regiones en la imagen.

La transformada Watershed es una herramienta morfológica donde la elección correcta de las semillas es fundamental para evitar la sobre-segmentación, por lo que en el presente trabajo se aplicaron los procedimientos antes descritos donde los resultados obtenidos permiten evaluar y elegir el mejor filtro a través de las segmentaciones conseguidas por la técnica de la transformada Watershed. A continuación se muestra en la figura 6.23 los resultados de aplicar los procedimientos en una subimagen proveniente del satélite Francisco de Miranda de dimensiones 502x506x4, y resolución radiométrica de 10 bits. La visualización de las segmentaciones se hace por medio de la combinación de bandas [4, 3, 2] referente al falso color infrarrojo del satélite Miranda.

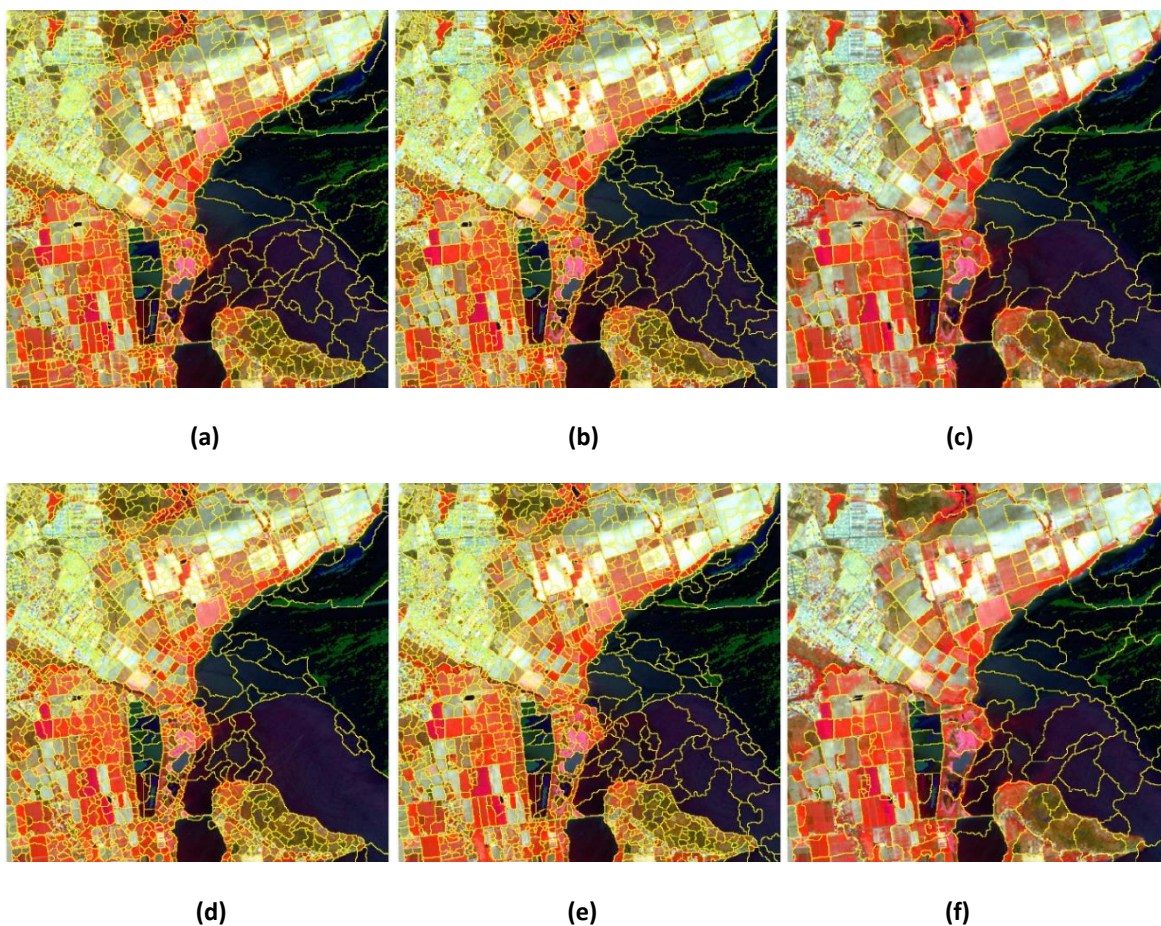


Figura 6.23. Resultados de las Segmentaciones por medio de la técnica de la Transformada Watershed. (a) Imagen original mediante FAS, (b) Imagen original mediante una Reconstrucción Dual, (c) Imagen original mediante una Reconstrucción por apertura y cierre, (d) ACP mediante FAS, (e) ACP mediante una Reconstrucción Dual, (f) ACP mediante una Reconstrucción por apertura y cierre.

La evaluación de los resultados por un especialista en la generación de mapas temáticos a partir de imágenes satelitales permite la escogencia del mejor procedimiento y método de segmentación es el objeto del capítulo siguiente.

Capítulo 7. Evaluación de Resultados

Los procedimientos propuestos se aplicaron y analizaron a 4 imágenes satelitales, de diferentes zonas de Venezuela donde se aprecian regiones de distintos tipos como; parcelas que representan estados fenológicos diferenciados; masas de agua; zonas montañosas y boscosas, entre otras. Las imágenes evaluadas poseen distintas resoluciones espaciales, espectrales y radiométricas con el fin de evaluar el comportamiento de las segmentaciones en imágenes con píxeles de características distintas. La figura 7.1 muestra las imágenes de prueba y sus particulares.



Subimagen: Landsat 7. Dim 635x860x6. Resolución 8 bits

(a)



Subimagen: Landsat 8. Dim 500x500x7. Resolución 16 bits

(b)



Subimagen: Landsat 8. Dim 934x971x7. Resolución 16 bits

(c)



Subimagen: Francisco de Miranda. Dim 502x506x4. Resolución 10 bits

(d)

Figura 7.1. Imágenes Satelitales. (a) Acarigua Estado Portuguesa, (b) Calabozo Estado Guárico, (c) Lado de Maracaibo Estado Zulia, (d) Lago de Tacarigua Estado Carabobo.

Las evaluaciones de los procedimientos se hacen con la ayuda del Geógrafo Wladimir Barrios, especialista en interpretación de imágenes satelitales del Centro de Procesamiento Digital de Imágenes en la Fundación Instituto de Ingeniería para Investigación y Desarrollo Tecnológico de Caracas, Venezuela.

Se presentan en la tabla 7 las evaluaciones de los procesos descritos en el capítulo anterior, aplicados a cada imagen de prueba, nombradas por la zona territorial a la que corresponden. Se mencionan los filtros empleados, donde, F1 simboliza la segmentación por medio del filtro alternado secuencial; F2 el filtro por reconstrucción dual; y F3 referente al filtro apertura y cierre por reconstrucción. Además se distingue la imagen de entrada suministrada.

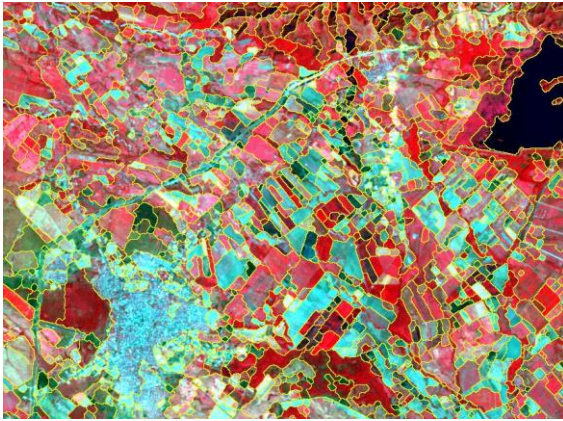
Los criterios determinados por el especialista para la evaluación se basan fundamentalmente en: la cantidad de segmentos obtenidos y la separación de las regiones con características visuales similares que permitirán la clasificación orientada a objetos posteriormente.

La escala numérica a utilizar es de tres valores, tal que, representan una segmentación: 1-Deficiente, 2-Buena, 3-Muy Buena.

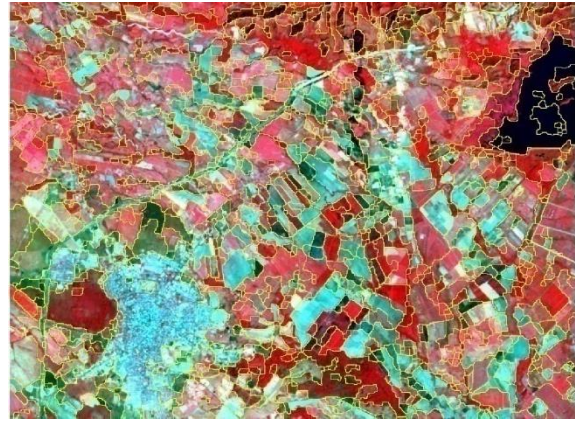
	Transformada Watershed						Autómatas Celulares					
	Imagen Original			Análisis de Componentes Principales			Imagen Original			Análisis de Componentes Principales		
	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
Acarigua	3	2	1	2	2	1	2	2	2	2	3	2
Calabozo	1	1	3	1	1	1	3	2	2	2	2	2
Lago de Maracaibo	1	1	3	1	1	1	2	2	2	2	2	3
Lago de Tacarigua	1	3	1	1	1	2	2	2	1	2	2	3

Tabla 7. Escala de Evaluación de las segmentaciones resultantes.

Se muestra en la figura 7.2 las segmentaciones con mayor puntuación de cada una de las imágenes de prueba. La visualización se hace por medio de la combinación referente al falso color infrarrojo.



(a)



(b)



(c)



(d)



(e)



(f)

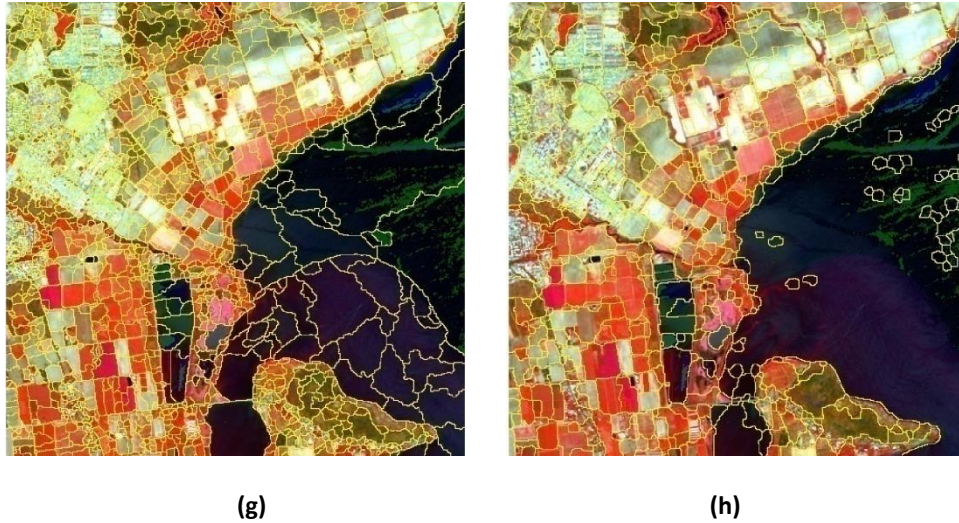


Figura 7.2. Segmentaciones de puntuación alta. (a) Imagen Original Acarigua, F1 por transformada Watershed, (b) ACP de Acarigua, F2 por autómata celular, (c) Imagen Original Calabozo, F3 por transformada Watershed, (d) Imagen Original Calabozo, F1 por autómata celular, (e) Imagen Original Lago de Maracaibo, F3 por transformada Watershed, (f) Imagen Original Lago de Maracaibo, F3 por autómata celular, (g) Imagen Original Lago de Tacarigua, F2 por transformada Watershed, (h) Imagen Original Lago de Tacarigua, F3 por autómata celular.

La puntuación de las segmentaciones permite discriminar entre las técnicas y los filtros propuestos. A través de las evaluaciones se escoge el procedimiento que más se adapte a las distintas características de una imagen, sin embargo, se debe mencionar que los parámetros que influyen en la segmentación como la cantidad de iteraciones de los filtros, la forma y el tamaño del elemento estructurante inicial, son factores determinantes que se relacionan estrechamente con la escala de trabajo en la que se realiza la interpretación y más aún con las zonas de interés del especialista. Se muestra en la figura 7.3, la representación gráfica de los datos de la tabla 7, con el fin de observar las puntuaciones obtenidas por filtros en cada uno de los métodos propuestos.

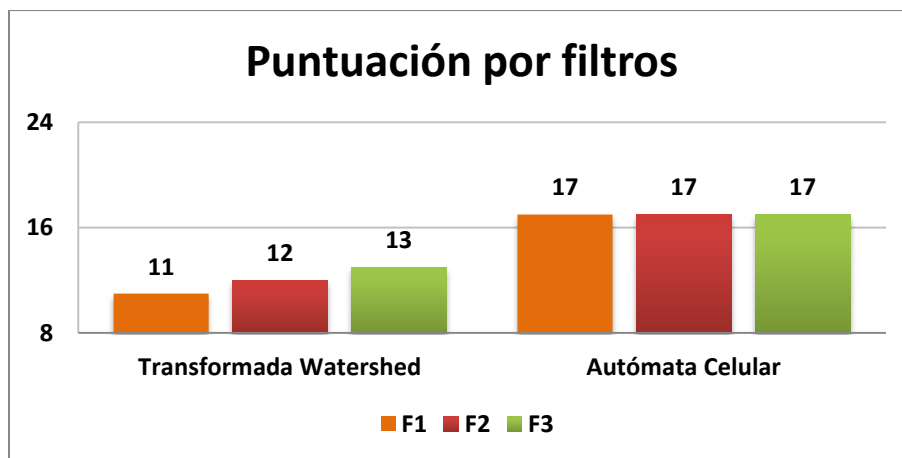


Figura 7.3. Representación gráfica del puntaje obtenido por filtro en cada una de las técnicas.

La figura anterior permite observar que las segmentaciones logradas utilizando el método Growcut por los distintos filtros obtienen la misma puntuación, con el análisis de la tabla 7 y la figura 7.3 se podría decir que, los filtros se comportan de manera similar a través de esta técnica. Sin embargo, cuando se analizan las puntuaciones de los filtros por el método de la transformada Watershed se presentan pequeñas variaciones en la que F3 correspondiente al filtro apertura y cierre por reconstrucción es la transformación que consigue mayor puntuación, lo que permite establecer el mismo, como el filtro con mejores resultados en las segmentaciones.

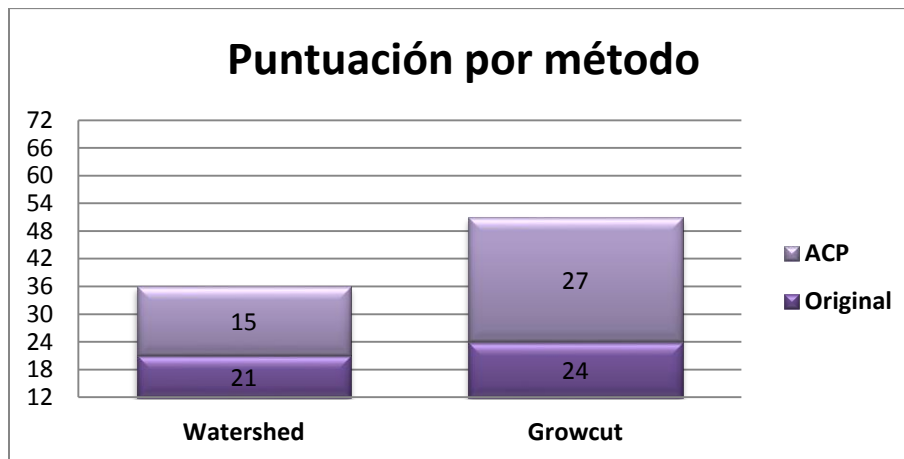


Figura 7.4. Representación gráfica del puntaje obtenido por método.

La figura 7.4 indica que la técnica basada en un autómata celular, es el método que posee mayor calificación (51 puntos) en la evaluación de las segmentaciones, de esta manera se puede establecer como la mejor técnica que se adapta a las distintas características presentes en las imágenes de prueba.

Una vez que se obtiene la técnica de autómata celular como el mejor método para la segmentación, se analiza el puntaje total por el tipo de imagen de entrada que se le suministra al proceso, donde, el análisis de componentes principales es la imagen que ofrece la mayor calificación, además de tener la ventaja en cuanto a la reducción de la dimensionalidad de la imagen original que permite la reducción del tiempo de cálculo del proceso. Se debe tomar en cuenta que en el método de la transformada Watershed la puntuación permite establecer que resulta ser más factible suministrar la imagen original para obtener una mejor segmentación. Por lo tanto, en base a la evaluación se puede escoger como el mejor proceso de segmentación el que se representa en la figura 7.5.

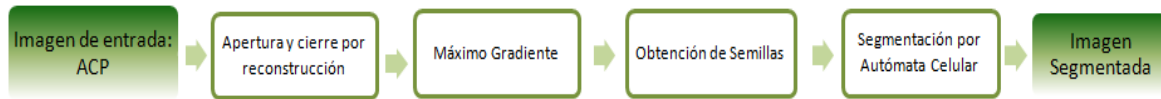


Figura 7.5. Proceso Destacado para la segmentación de imágenes.

Por consiguiente, se considera que el proceso esquematizado en la figura 7.5 posee ventajas respecto a aquel que se aplica con la transformada Watershed, por cuanto elimina la sobre-segmentación que este genera conservando la delimitación de las regiones similares. El parámetro esencial que se debe considerar para su buena aplicación, corresponde a la resolución espacial que posee la imagen de entrada.

Conclusiones

Se presenta una técnica de crecimiento de regiones para la segmentación de imágenes sin supervisión mediante autómatas celulares, donde las semillas y su estado inicial se determinan automáticamente por el algoritmo. Se comprueba que la apertura y cierre por reconstrucción y el gradiente multiescala son eficientes transformaciones que permiten eliminar el ruido, homogenizar y preservar los bordes de las distintas regiones para posteriormente identificar sus representantes por medio de las funciones empleadas en la obtención de las semillas. La determinación de las etiquetas de las semillas basadas en análisis estadísticos resulta ser un buen criterio que involucra sus firmas espectrales y permite una preclasificación automática de las distintas regiones a pesar de que estén espacialmente distanciadas.

El preprocesamiento de la imagen original a través del análisis de componentes principales arroja buenos resultados, lo que permite establecer una manera de optimizar los tiempo en la aplicación de los filtros y en general al proceso de segmentación debido a la reducción de la dimensionalidad de los datos, teniendo la ventaja de no perder información relevante.

El método propuesto ha sido comparado con la segmentación por medio de la transformada Watershed y resulta ventajoso por las siguientes razones:

Disminución de sobresegmentación; beneficio en la aplicación del análisis de componentes principales; preclasificación de las regiones presentes en la imagen; uso de la firma espectral como descriptor de las semillas; la complejidad computacional es independiente a la dimensión de la imagen (permite el análisis por bloques y no píxel a píxel).

A través del método por autómatas celulares se define una única regla de transición de estados determinista que permite que la segmentación y la caracterización de las regiones se puedan realizar en un solo paso. Los parámetros de ésta técnica hacen de él, un sistema ampliamente adaptable a distintas imágenes.

En estudios posteriores se sugiere la aplicación de éste proceso a imágenes multicanal de distintos tipos para verificar la precisión de las segmentaciones, así como también la propuesta de nuevos parámetros de tipo topológico que permitan mejorar el proceso de etiquetado hasta llegar a una segmentación más precisa.

A partir de los resultados mencionados, se puede concluir que el método Growcut permite obtener una segmentación de buena calidad, siempre y cuando se elijan los indicadores adecuados como la forma y el tamaño del elemento estructurante para el filtrado.

Apéndice. MATLAB

MATLAB[®] es un lenguaje de alto nivel y un entorno interactivo para el cálculo numérico, visualización y programación. Usando MATLAB, se pueden analizar datos, desarrollar algoritmos y crear modelos y aplicaciones. El lenguaje, las herramientas y funciones matemáticas integradas permiten explorar múltiples enfoques y llegar a una solución más rápida que con hojas de cálculo o lenguajes de programación tradicionales, como C / C++ o Java[®]. Puede utilizar MATLAB para una gama de aplicaciones, incluyendo el procesamiento de señales y comunicaciones, procesamiento de imágenes y de vídeo, sistemas de control, prueba y medición, finanzas computacionales, y la biología computacional(The Mathworks, 1994-2015).

Image Processing Toolbox[™] ofrece un conjunto completo de referencias de algoritmos, funciones y aplicaciones para el procesamiento de imágenes, análisis, visualización y desarrollo de algoritmos. Puede realizar análisis de imágenes, segmentación de imágenes, mejora de imagen, reducción de ruido, transformaciones geométricas, y registro de imágenes. Muchas de las funciones de Toolbox soportan procesadores multinúcleo, GPU, y la generación de código C. Admite diversos tipos de imágenes y contiene funciones de visualización y aplicaciones que permiten explorar las imágenes y vídeos, examinar una región de píxeles, ajustar el color y el contraste, crear contornos o histogramas y manipular las regiones de interés (ROI). Es compatible con los flujos de trabajo para el procesamiento, visualización, y la navegación de grandes imágenes .

Características principales

- Análisis de imágenes, lo cual incluye segmentación, morfología, estadística y medición.
- Mejora de imágenes, filtrado y enfoque de imágenes borrosas.
- Transformaciones geométricas y métodos de registro de imágenes basados en intensidad.
- Flujos de trabajo para imágenes grandes, que incluyen procesamiento por bloques, creación de mosaicos y visualización con varias resoluciones.
- Aplicaciones de visualización, como Image Viewer y Video Viewer.
- Características compatibles con procesadores multinúcleo y GPUs, así como soporte para la generación de código C.

Image Processing Toolbox soporta imágenes y videos generados por una amplia gama de dispositivos, tales como: webcams; cámaras digitales; y sensores aerotransportados;

dispositivos de generación de imágenes médicas; microscopios; telescopios y otros instrumentos científicos. Las diversas características y aplicaciones permiten visualizar, analizar y procesar estas imágenes en muchos tipos de datos.

MATLAB[®] soporta formatos de datos e imágenes como los siguientes:

- AVI
- JPEG
- JPEG-2000
- FITS
- HDF
- HDF-EOS
- M4V
- MOV
- MP4
- PNG
- TIFF
- ASCII
- Archivos binarios
- Microsoft[®] Excel[®]

También es compatible con los formatos de imágenes multibanda BIP y BIL, utilizados por LANDSAT.

Image Processing Toolbox es compatible con diversos formatos de archivo de imagen especializados. Para imágenes médicas, soporta archivos DICOM, incluidos los metadatos asociados, así como los formatos ANALYZE 7.5 y INTERFILE. Puede leer imágenes geoespaciales en archivos NITF e imágenes de alto rango dinámico en archivos HDR.

Image Processing Toolbox proporciona una serie de aplicaciones de procesamiento de imágenes que permite explorar y descubrir diferentes enfoques algorítmicos. Gracias a la aplicación Color Thresholder, se puede segmentar una imagen en función de los distintos espacios de color. La aplicación Image Viewer permite ubicar y manipular interactivamente ROI, lo que incluye puntos, líneas, rectángulos, polígonos, elipses y formas a mano libre. También es posible ver información sobre los píxeles, realizar acercamientos, ajustar el contraste, y medir distancias. Puede realizar estas tareas mediante programación y utilizar las funciones individuales para crear interfaces personalizadas.

Mejora de la imagen. Permite aumentar la proporción de señal a ruido y acentuar las características de las imágenes modificando los colores o las intensidades de una imagen. Incluye rutinas de filtrado especializadas y una función de filtrado multidimensional generalizado, ofrece opciones de relleno de contornos múltiples, y realiza convolución y correlación.

Mediante los filtros y las funciones predefinidas se puede:

- Filtrar con operadores morfológicos: Permite mejorar el contraste, eliminar ruido y regiones delgadas, o realizar esqueletización en regiones. Funciones morfológicas incluidas en Image Processing Toolbox: Dilatación, Erosión, Apertura y Cierre.
- Enfocar imágenes borrosas y aumentar la nitidez.
- Eliminar el ruido mediante el filtrado lineal, de mediana, o adaptativo.
- Realizar ecualización de histogramas.
- Ajustar el contraste.

Image Processing Toolbox ofrece un conjunto de algoritmos y funciones de visualización estándar de referencia para tareas de análisis de imágenes tales como análisis estadístico, extracción de funciones y medición de propiedades. Proporciona diversas transformaciones de imágenes, incluyendo la Transformada de Hough, Radon, FFT, DCT, y proyecciones fan-beam. También hay transformaciones disponibles en Wavelet Toolbox™. Las funciones estadísticas permiten analizar las características generales de una imagen mediante: Cálculo de la media o la desviación estándar.

Los algoritmos de segmentación de imágenes determinan los límites de una región presente en la imagen. Se puede explorar diferentes enfoques para la segmentación de imágenes, entre los que se encuentran métodos progresivos, la creación automática de umbrales, métodos basados en bordes y métodos morfológicos como la transformada de Watershed.

Transformaciones geométricas. Soporta operaciones simples, tales como el cambio de tamaño, la rotación y el recorte, así como transformaciones geométricas bidimensionales más complejas, tales como las afines y las proyectivas. También proporciona un marco flexible para la creación y aplicación de transformaciones geométricas personalizadas y métodos de interpolación para arreglos N-dimensionales.

Gracias al uso de Image Processing Toolbox junto con MATLAB Coder™, Vision HDL Toolbox™, y HDL Coder™, se puede generar código C, C++ y HDL directamente desde MATLAB. Muchas funciones de procesamiento de imágenes soportan la generación de código, lo que permite ejecutar algoritmos de procesamiento de imágenes en hardware para PC, FPGAs y ASICs. Esto permite desarrollar sistemas de generación de imágenes para los ámbitos médico, aeroespacial y de defensa.

Bibliografía

ABAE [En línea] // Agencia Bolivariana para Actividades Espaciales. - 2014. - www.abae.gob.ve.

Aldalur B Santamaria M Realce de imágenes: filtrado espacial // Revista de Teledetección. - 2002. - págs. 31-42.

Beucher S. Meyer, F Méthodes d'analyse des contrastes á l'analyseur de textures. Tech Report CMM N-625. - Paris, 1977.

Chuvieco Salinero E Fundamentos de Teledetección Espacial. - Madrid : RIALP, 1996.

Cook Mathew Universality in Elementary Celullar Automata. Complex Systems. - 2004. - Vol. 15.

Fernández J. Desarrollo de técnicas y operadores morfológicos: aplicación al procesamiento y análisis de imágenes en cardiología. - 1994.

Hadwiger H. Normale köer im euklidischen raum und ihre topologischen und metrischen eigenschaften. - 1959.

Hadwiger H. Vorlesungen über Inhalt, Oberfläche und Isoperimetrie. 1957.

Hay G. Castilla G. Object-Based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (Swot). - 2006.

Hernandez G. and Herrmann, H Cellular automata for elementary image enhancement // Graphical Model and Image Processing. - 1996.

James B. Campbell y Randolph H. Wynne Introduction to Remote Sensing. The Guilford Press, 2011.

Johnson R. y Wichern D. Applied multivariate statistical analysis. Nueva Jersey : Academic Press, 1992.

Kauffmann C. Piche, N Seeded ND Medical Image Segmentation by Cellular Automaton on GPU. - International Journal of Computer Assisted Radiology and Surgery, 2009. - Vol. 5.

Kim Edward. Shen, Tian. Huang, Xiaolei A Parallel Cellular Automata with Label Priors for Interactive Brain Tumor Segmentation. - IEEE 23rd International Symposium on Computer-Based Medical Systems Conference Publications, 2010.

Mardia K.V., Kent J.T. y Bibby J.M. Multivariate analysis. - 1982. - pág. 521.

Marques Oge Practical Image and Video Processing Using. - New Jersey : John Wiley & Sons, Inc, 2011.

- Matheron G.** Eléments pour une Théorie des Milieux Poreux. - Masson, Paris, 1967.
- Matheron G.** Random Sets and Integral Geometry. - New York, 1975.
- Minkowsky** Allgemeine Lehrsätze über konvexe Polyeder. - 1897.
- Minkowsky** Über die Begriffe Länge, Oberfläche und Volumen. - 1901.
- Popovici A, Popovici, D** Cellular Automata in Image Processing. - Department of Computer Science and Mathematics, University of West Timisoara, 2002.
- Richards J. A.** Remote Sensing Digital Image Analysis - An Introduction. - Berlin : Springer-Verlag, 1995.
- Rueda T. A.** Segmentación multicanal de tumores cerebrales en imagenología de resonancia magnética con autómatas celulares deterministas. - Caracas, 2013.
- Rueda T. A., Torres W.** Autómatas Celulares para la Segmentación y Clasificación de Imágenes Multiespectrales. V Jornadas Nacionales de Geomática, 2013.
- Serra J.** Image Analysis and Mathematical Morphology. Academic Press, 1981.
- Serra J.** Image Analysis and Mathematical Morphology. Theoretical Advances. Academic Press, 1988.
- Sociedad Española de Astronomía. - 5 de Diciembre de 2014. - <http://www.sea-astronomia.es>.
- Soille Pierre** Morphological Image Analysis: Principles and Applications. - Paris : Springer, 2003.
- Sternberg S.** Grayscale Morphology, Computer Graphics and Image Processing. 1986.
- The Mathworks Inc** www.mathworks.com [En línea]. - 1994-2015. - 12 de Marzo de 2015. - <http://www.mathworks.com/products/image>.
- Thomas M. Lillesand, Ralph W. Kiefer** Remote Sensing and Image Interpretation. - California : Wiley, 1979.
- Vezhnevets Vladimir. Konouchine, Vadim** GrowCut -Interactive Multi-Label N-D Image Segmentation by Cellular Automata. Proceedings of Graphicon. - 2005.
- Von Neumann John** Theory of Self-Reproducing Automata. University of Illinois Press, 1966.
- Wolfram Stephen** A New kind of Science. - Estados Unidos : Wolfram Medios, 2002.
- Wolfram Stephen** Statistical Mechanics of Cellular Automata. Reviews of Modern Physics. The Institute for Advanced Study, Princeton, 1983. - Vol. 55.