

UNIVERSIDAD DE MÁLAGA
ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA
INGENIERO EN INFORMÁTICA

DETECCIÓN DE CAMBIOS EN IMÁGENES DE SATÉLITE
DE ALTA RESOLUCIÓN

Realizado por
D. JOSÉ ANTELO GÁLVEZ

Dirigido por
Dr. D. JAVIER GONZÁLEZ JIMÉNEZ

Departamento
INGENIERÍA DE SISTEMAS Y AUTOMÁTICA

MÁLAGA, Septiembre de 2013

Agradecimientos

A mis padres, por su infinita paciencia y apoyo incondicional. Por confiar siempre en mí. Por dárme todo y porque sin su ayuda, llegar hasta el final hubiera sido imposible.

A Aida por su importantísima y desinteresada ayuda en la corrección de esta memoria. Por ver siempre lo mejor de mí y estar siempre ahí en los momentos más difíciles.

A Javier, Goyo y Vicente por iniciarme en la Teledetección y la Detección de Cambios, temas de este proyecto, porque sin sus conocimientos y consejos el camino recorrido hubiera sido prácticamente intransitable.

A Juanjo, Dani y resto de compañeros de los que aprendí mucho durante tantas horas de estudios y exámenes superados.

Y al resto de mi familia y amigos que me han apoyado y animado a llegar hasta aquí.

A todos. GRACIAS.

Índice de tablas

<i>Tabla 2-1. Clasificación del espectro electromagnético.....</i>	<i>27</i>
<i>Tabla 2-2. Ventanas atmosféricas.</i>	<i>31</i>
<i>Tabla 2-3. Características del sensor Quickbird y los sensores Wordlview.....</i>	<i>46</i>
<i>Tabla 2-4. Características de los sensores Ikonos, Geoeye 1 y 2, Eros B y Cartosat 2.</i>	<i>47</i>
<i>Tabla 3-1. Coeficientes métodos lineales de ajuste radiométrico.</i>	<i>57</i>
<i>Tabla 3-2. ECM niveles de gris en experimento 1 tras normalización radiométrica.....</i>	<i>67</i>
<i>Tabla 3-3. ECM niveles de gris en experimento 2 tras normalización radiométrica.....</i>	<i>68</i>
<i>Tabla 4-1. Parámetros experimentos de corrección geométrica.....</i>	<i>104</i>
<i>Tabla 4-2. Datos adquisición imágenes de Rincón de la Victoria.</i>	<i>105</i>
<i>Tabla 4-3. Datos adquisición imágenes de Huelva.</i>	<i>106</i>
<i>Tabla 4-4. Datos de error en los experimentos de corrección geométrica.</i>	<i>113</i>
<i>Tabla 4-5. Datos comparativos de las técnicas de detección geométrica.</i>	<i>114</i>
<i>Tabla 4-6. Ventajas e inconvenientes de técnicas de detección y matching de puntos de control.....</i>	<i>119</i>
<i>Tabla 6-1. Comparativa de técnicas de umbralización automática.</i>	<i>160</i>

Índice de figuras

Figura 2-1. Componentes de un sistema de teledetección.....	22
Figura 2-2. Onda electromagnética.....	25
Figura 2-3. Espectro electromagnético.....	26
Figura 2-4. Interacciones de la radiación electromagnética con la atmósfera.....	29
Figura 2-5. Absorción de la radiación en la atmósfera.....	31
Figura 2-6. Tipos de reflectores.....	32
Figura 2-7. Transmisión de la radiación.....	33
Figura 2-8. Absorción de la radiación.....	33
Figura 2-9. Tipos de sensores.....	35
Figura 2-10. Resolución espacial.....	37
Figura 2-11. Salida obtenida para diferentes resoluciones.....	38
Figura 2-12. Resolución radiométrica.....	39
Figura 2-13. Respuesta espectral de las bandas de Worldview 2.....	42
Figura 2-14. Respuesta espectral de las bandas de Ikonos.....	43
Figura 2-15. Taxonomía de las causas de los cambios radiométricos en imág. de alta resolución.....	49
Figura 2-16. Posibles estructuras de sistema de detección de cambios.....	50
Figura 3-1. Histogramas de imagen de referencia e imagen a corregir.....	57
Figura 3-2. Ajustes radiométricos lineales.....	58
Figura 3-3. Resultados ajustes radiométricos lineales.....	59
Figura 3-4. Resultado de la especificación de histograma.....	63
Figura 3-5. Histogramas de imagen de referencia y de imagen con histograma especificado.....	63
Figura 3-6. División de la imagen para procesamiento adaptativo.....	64
Figura 3-7. Resultado de una corrección radiométrica adaptativa bilineal.....	65
Figura 3-8. Imágenes de Huelva para corregir radiométricamente.....	66
Figura 4-1. Ventanas usadas por el operador de Moravec.....	73
Figura 4-2. Ventana gaussiana usada en operador de Harris.....	74
Figura 4-3. Regiones de los autovalores de la matriz M. Harris.....	76
Figura 4-4. Detalle de una imagen de satélite a diferentes escalas.....	79
Figura 4-5. Operadores Gaussiana, DoG y LoG.....	79
Figura 4-6. Formación de la pirámide Gaussiana de imágenes.....	80
Figura 4-7. Selección de extremos por comparación con sus 26 píxeles vecinos.....	81
Figura 4-8. Cálculo de orientaciones de los puntos característicos SIFT.....	83
Figura 4-9. Cálculo del descriptor de una característica SIFT.....	84
Figura 4-10. Efecto de un pequeño desplazamiento en el cálculo del descriptor.....	84
Figura 4-11. Pesos de un píxel en el cálculo de histogramas del descriptor.....	85
Figura 4-12. Pesos en la asignación de orientaciones a un determinado bin del histograma.....	85
Figura 4-13. Suma de píxeles de una región a partir de imagen integral.....	86
Figura 4-14. Filtros aproximados de derivadas parciales de segundo orden.....	87
Figura 4-15. Uso diferente de pirámides en SIFT y SURF.....	88

Figura 4-16. Filtros D_{yy} y D_{xy} para dos escalas sucesivas (9x9 y 15x15).	88
Figura 4-17. Cálculo de orientación dominante de características SURF.	90
Figura 4-18. Cálculo de descriptores SURF.	90
Figura 4-19. Ejemplos de descriptores para cada una de las regiones en un descriptor SURF.	90
Figura 4-20. Búsqueda de una ventana por correlación cruzada normalizada.	92
Figura 4-21. Ejemplo de triangulación de Delaunay.	102
Figura 4-22. Ajuste de imagen esclava mediante CPs.	102
Figura 4-23. Tipos de remuestreo de imágenes.	103
Figura 4-24. Imágenes de zona de Nerja.	104
Figura 4-25. Imágenes de zona de Rincón de la Victoria.	105
Figura 4-26. Imágenes de zona de Huelva.	105
Figura 4-27. Porc. de puntos candidatos seleccionados como emparejamientos tras eliminar outliers.	112
Figura 4-28. Porcentaje de outliers respecto al total de correspondencias.	112
Figura 4-29. Diferencias en la posición del Sol en R. Victoria (arriba) y Huelva (abajo).	115
Figura 4-30. Error cuadrático medio de los niveles de gris tras la corrección geométrica.	117
Figura 4-31. Error RMS ajuste polinomial hasta orden 4.	118
Figura 5-1. Tipos de sombra.	123
Figura 5-2. Componente C3 de una imagen.	125
Figura 5-3. Semillas detectadas en sombras.	126
Figura 5-4. Sombras detectadas en una imagen RGB.	128
Figura 5-5. Proceso de detección de sombras.	128
Figura 5-6. Sombras usando el espacio de color HSI (izquierda) e YIQ (derecha).	129
Figura 5-7. Uso de restauración de sombras I.	130
Figura 5-8. Uso de restauración de sombras II.	131
Figura 5-9. Imagen en color verdadero, en falso color y su índice NDVI.	135
Figura 5-10. Representación de los espacios RGB, HSV y LAB.	139
Figura 6-1. Vector de cambios en técnica CVA para imágenes de dos bandas.	144
Figura 6-2. Proceso de detección de cambios por post-clasificación.	147
Figura 6-3. Asignación de bandas para análisis visual de cambios.	149
Figura 6-4. Ventanas de la función interactiva de selección de valores umbrales.	151
Figura 6-5. Selección de nivel umbral para histograma unimodal.	154
Figura 6-6. Iteraciones en la umbralización con histéresis.	157
Figura 6-7. Histograma de imagen de cambios por método de diferencias.	158
Figura 6-8. Ejemplos de cambio en λ y ν . Tomado de (Metternicht, 1999).	159
Figura 6-9. Umbralización multinivel por probabilidad de cambio.	159
Figura 6-10. Curvas ROC.	163
Figura 6-11. Curvas ROC usando imágenes RGB, HSV, LAB e imágenes en niveles de gris.	163
Figura 6-12. Cambios pequeños y ruido.	164
Figura 6-13. Venta para la corrección geométrica local.	165
Figura 6-14. Consenso con vecinos en corrección geométrica local.	166
Figura 6-15. Cambios debidos a un elemento mal alineado.	167
Figura 6-16. Cambios debidos a elementos ligeramente mal alineados.	167
Figura 6-17. Cambios debido a sombras.	168
Figura 6-18. Ángulo de adquisición de las imágenes.	168
Figura 6-19. Efectos del ángulo de adquisición en los cambios.	169
Figura 6-20. Falsos negativos por píxeles con color muy similar.	169
Figura 6-21. Representación del vector de cambios en el espacio de cambio radiométrico.	172
Figura 6-22. Ejemplo de clasificación por dirección del vector de cambios.	173
Figura 6-23. Ejemplo de división en parcelas para la detección de cambios.	176
Figura A-1. Menú de opciones de correcciones radiométricas.	185
Figura A-2. Menú de opciones de correcciones geométricas.	186

<i>Figura A-3. Puntos de control distribuidos homogéneamente.</i>	187
<i>Figura A-4. Selección de métodos de extracción y matching de características.</i>	188
<i>Figura A-5. Ventana de parámetros de los métodos de corrección geométrica.</i>	189
<i>Figura A-6. Ventanas del detector simple de sombras para el espacio de color c1c2c3.</i>	190
<i>Figura A-7. Captura de ventanas del detector de sombras.</i>	191
<i>Figura A-8. Selección de método de detección de cambios.</i>	193
<i>Figura A-9. Menú de opciones de transformación.</i>	193
<i>Figura A-10. Menú de opciones de umbralización.</i>	194
<i>Figura A-11. Selección de umbral con histéresis.</i>	195
<i>Figura A-12. Opciones de filtrado y parámetros para corrección geométrica local.</i>	196
<i>Figura A-13. Herramientas para la definición (izq.) y clasificación manual (der.) de cambios.</i>	198
<i>Figura A-14. Asistente para la detección de cambios.</i>	199

Tabla de contenidos

Capítulo 1. Introducción	15
1.1 Antecedentes.....	15
1.2 Objetivos.....	16
1.3 Estructura de la memoria del proyecto	17
1.4 Elementos hardware y software usados	18
1.5 ¿Por qué ENVI y OpenCV?	18
Capítulo 2. Introducción a la teledetección espacial	20
2.1 Componentes de un sistema de teledetección	20
2.2 Ventajas de la teledetección espacial.....	22
2.3 Aplicaciones de la teledetección.....	23
2.4 Radiación electromagnética	25
2.5 Interacciones con la atmósfera	28
2.6 Interacciones con la superficie terrestre	31
2.7 Satélites de Alta Resolución	33
2.8 Principales satélites de alta resolución espacial	40
2.9 Detección de cambios	48
Capítulo 3. Correcciones radiométricas	52
3.1 Métodos lineales de corrección radiométrica relativa	53
3.2 Especificación del histograma (EH)	60
3.3 Normalización adaptativa bilineal (AB)	64
3.4 Experimento con imágenes con cambios	65
3.5 Resultados y conclusiones	66
3.6 Futuras mejoras.....	69
Capítulo 4. Correcciones geométricas	70

4.1	Detección automática de características	71
4.2	Matching o emparejamiento de características	91
4.3	Estimación del modelo de transformación	99
4.4	Remuestreo de la imagen y transformación	102
4.5	Comparación de las técnicas y conclusiones	103
4.6	Futuras mejoras.....	120
Capítulo 5. Extracción de sombras y otras capas de información.....		122
5.1	Detección de sombras	123
5.2	Restauración de sombras.....	129
5.3	Detección de nubes	131
5.4	Índices de vegetación	134
5.5	Otras capas de información.....	135
Capítulo 6. Detección de cambios		140
6.1	Técnicas de detección de cambios.....	141
6.2	Umbralización.....	150
6.3	Espacios de color en la detección de cambios	161
6.4	Filtrado de cambios	164
6.5	Clasificación de cambios	170
6.6	Futuras mejoras.....	175
Capítulo 7. Conclusiones y resumen.....		177
Apéndice A. Detalles de implementación e instalación del software.....		184
A.1	Funciones para la corrección radiométrica.....	184
A.2	Funciones para la corrección geométrica	185
A.3	Funciones para detección de sombras	189
A.4	Funciones para la restauración de sombras	192
A.5	Funciones para la detección de cambios.....	192
A.6	Funciones para el filtrado y clasificación de cambios	195
A.7	Asistente para la detección de cambios	198
A.8	Instalación del software	199
Apéndice B. Listado de funciones.....		201
Referencias.....		209

Capítulo 1

Introducción

Vivimos en un mundo globalizado en el que las actividades en un lugar del planeta se ven reflejadas en el otro extremo. Esto ha hecho que sea necesario el estudio constante de la superficie terrestre y los cambios que produce la actividad humana en ella, usando para ello nuevos tipos de información como la proporcionada por los satélites.

Los satélites, a través de los sensores instalados en ellos, nos ofrecen información muy variada que nos permiten estudiar los distintos elementos de la superficie o la atmósfera terrestre, generalmente, en forma de imágenes que recogen la radiación reflejada o emitida por el planeta y los gases de la atmósfera. Las imágenes más comunes son aquellas que recogen la información del espectro visible, pues con un tratamiento sencillo, pueden ser directamente utilizadas ya que nos muestran los diferentes elementos tal como los veríamos si estuviéramos en órbita sobre la Tierra.

La ciencia que emplea estas imágenes como fuente de información se conoce como Teledetección y, dentro de ésta, se conoce como Detección de Cambios al estudio, técnicas y procedimientos usados para identificar y clasificar los distintos cambios que se producen en el planeta. En este área se centra este proyecto y, para ello, se hará uso de imágenes en el espectro visible (con algunas referencias al infrarrojo cercano) y de alta resolución espacial.

1.1 Antecedentes

Existen multitud de herramientas software para el tratamiento de imágenes de satélite y datos espaciales: unas especializadas en lo que se conoce como imágenes raster y otras en el uso de capas o imágenes vectoriales. Software como ERDAS IMAGINE o ENVI, se centran en el uso de imágenes raster frente a las vectoriales. Por otro lado, los Sistemas de Información Geográfica (SIGs) se centran en el uso de las imágenes vectoriales aunque pueden trabajar también con imágenes raster usando diferentes extensiones software.

En la detección de cambios, ya que la fuente de información suele provenir de imágenes aéreas, es preferible usar software de teledetección especializado en tratamiento de imágenes raster en vez de software GIS, aunque ambos pueden trabajar indistintamente con ambos tipos de imágenes.

Entre el software de teledetección y software GIS más conocido encontramos las siguientes aplicaciones:

- ERDAS IMAGINE
- IDL/ENVI
- ArcGIS con la extensión Image Analysis.
- GRASS
- gvSIG
- IDRISI
- eCognition

Todo este software y algunas aplicaciones más nos ofrecen múltiples herramientas que permiten trabajar con las imágenes de satélites y procesarlas con el fin de obtener un mapa de cambios que pueda utilizarse en estudios posteriores, como análisis de cambios de uso del suelo, detección de cambios urbanísticos, elaboración de mapas de riesgo, etc.

Al no existir un método que sirva para detectar todos y cada uno de los posibles tipos de cambio, cada estudio requiere el uso de diferentes herramientas y técnicas. Son tantas las posibilidades existentes que un usuario inexperto en el uso de software de teledetección podría perderse entre las técnicas y herramientas disponibles.

Además, el uso de imágenes de satélite lleva implícito ciertos problemas que son difíciles de corregir y que deben ser tenidos en cuenta para una óptima detección de cambios. Todo esto hace que tampoco existan herramientas totalmente automáticas y que, en muchos casos, se requiera de la pericia del analista para determinar los mejores parámetros para la detección de los cambios.

De lo comentado anteriormente surge la idea de este proyecto fin de carrera, donde se intenta recopilar las herramientas mínimas necesarias para una buena detección de cambios y elaborar un sistema que guíe a los usuarios noveles en el proceso, que tenga en cuenta los inconvenientes habituales en el procesamiento de imágenes de satélite y que permita la selección de parámetros de forma interactiva en aquellos procesos donde la automatización sea más compleja.

1.2 Objetivos

El objetivo principal de este trabajo es el estudio y recopilación de las técnicas y procedimientos necesarios para realizar una detección de cambios en imágenes de alta resolución, incluyendo las fases de pre-procesamiento que se realizan previas a

la detección, y su integración en alguna de las aplicaciones de teledetección y GIS existentes.

Asimismo, nos centraremos en algunos de los problemas que suelen surgir en la detección de cambios como son: 1) una imprecisa corrección geométrica, presentando una técnica de ajuste geométrico local basada en funciones de similitud como la Información Mutua; 2) la presencia de sombras en las imágenes, mostrando técnicas de detección basada en crecimiento de regiones y presentando un método interactivo alternativo, así como un algoritmo para su restauración; 3) la dificultad de determinar un valor umbral para la clasificación de los píxeles como cambios o no, estudiando diferentes algoritmos automáticos y presentando una herramienta interactiva para la selección de un valor umbral en imágenes en niveles de gris; y, por último, 4) la influencia del color en la detección de cambios.

Como objetivos secundarios, se persigue que este proyecto sirva como base para estudios posteriores en la detección de cambios a través de la integración de diferentes funciones de tratamiento de imágenes en el software comercial ENVI, pensando en la modularidad y reutilización de código, y permitiendo el uso de estas funciones en el diseño de nuevas técnicas y algoritmos. Igualmente, se pretende que esta memoria sirva como manual de referencia y punto de partida para todos aquellos con breves nociones sobre imágenes digitales, y sin conocimientos de teledetección y detección de cambios.

1.3 Estructura de la memoria del proyecto

La memoria de este proyecto se estructura en siete capítulos en los que haremos un recorrido por todas las fases de cualquier estudio de detección de cambios. Este primer capítulo presenta los objetivos, los medios utilizados e introduce los diferentes capítulos de la memoria.

El capítulo número dos nos introduce en la teledetección espacial y en los satélites de alta resolución. En él se tratan los conceptos más importantes de la ciencia de la teledetección que ayudarán a entender los capítulos posteriores.

Los capítulos tres, cuatro y cinco, tratan las fases iniciales de un proceso de detección de cambios. En el tercer capítulo, *Correcciones Radiométricas*, se tratará el proceso de normalización del color de las imágenes, proceso importante y necesario cuando la detección de cambios se basa en cambios de color de los diferentes píxeles. En el capítulo cuatro, llamado *Correcciones Geométricas*, se verá en qué consiste el proceso de registro de imágenes necesario para ajustar la posición de las imágenes y asegurar que los píxeles con iguales coordenadas representan el mismo punto de la superficie terrestre. Dedicaremos el capítulo cinco, *Extracción de sombras y otras capas de información*, a tratar la extracción de las distintas capas de información que permitirán simplificar el proceso de detección de cambios, como máscaras de fondo (sin información), nubes, vegetación o sombras.

El capítulo seis es el que se centra en la *Detección de cambios*. En él se estudia el proceso de obtención de cambios una vez las imágenes han sido pre-procesadas, tratando las distintas tareas que se llevan a cabo, como son la obtención de la imagen de cambios, la binarización o umbralización de esta imagen, y el posterior filtrado de cambios y su clasificación.

En un último capítulo resumiremos todo lo estudiado durante la elaboración del proyecto y recopilaremos las conclusiones obtenidas en cada una de las fases.

1.4 Elementos hardware y software usados

Para la elaboración de este proyecto fin de carrera se han usado los siguientes elementos hardware y software:

Hardware:

- Ordenador personal con las siguientes características:
 - o CPU Intel Core 2 Quad 2.5 GHz
 - o Memoria RAM DDR2 4 GB
 - o Monitor 22"
 - o Tarjeta gráfica NVIDIA GeForce 7100 / NVIDIA nForce 630i
 - o SO: Windows 7
 - o Disco duro 1 TB

Software:

- IDL 7.0
- ENVI 4.5
- Librería OpenCV 2.1
- Entorno de desarrollo IDL: Eclipse
- Entorno de desarrollo DLLs: Microsoft Visual Studio 2008

1.5 ¿Por qué ENVI y OpenCV?

Para la integración de las distintas herramientas se ha optado por usar el software IDL/ENVI de entre los programas existentes para tratamiento de imágenes de satélite. ENVI es un conjunto de funciones y herramientas escritas en el lenguaje IDL que está orientado al tratamiento de imágenes de teledetección. El uso de IDL/ENVI tiene multitud de ventajas que comentaremos a continuación.

Al estar ENVI escrito en IDL, permite añadir nuevas funcionalidades simplemente escribiendo nuevas funciones y añadiéndolas al menú del programa. Además, IDL está orientado a vectores, es decir, está ideado para trabajar rápidamente con grandes cantidades de información (como lo es la contenida en las imágenes) sin necesidad de acceder individualmente al valor contenido en cada píxel.

IDL y ENVI ofrecen varios widgets que hacen fácil la elaboración de nuevas interfaces para trabajar con las distintas funciones. Otra ventaja es la de extensiones de la funcionalidad a través del uso de librerías escritas en otros lenguajes. Puesto que ENVI está pensado para trabajar con imágenes muy grandes en extensión (como suele ocurrir con las imágenes de satélite), incorpora funciones para procesamiento por bloques, lo que ayuda a evitar problemas por falta de memoria. Todo esto y el hecho de que el conjunto IDL/ENVI sea usado cada día más en el ámbito científico, hacen que sea el mejor candidato para la elaboración de este proyecto.

Como inconveniente, tenemos que IDL es un lenguaje interpretado y, como tal, la ejecución de las distintas funciones es más lenta que la del código compilado. Por este motivo y con el objetivo también de reutilizar código existente, se ha optado por desarrollar dos librerías en lenguaje C y usar la librería OpenCV para aquellas tareas que exigen más recursos en su ejecución.

Capítulo 2

Introducción a la teledetección espacial

Las imágenes nos dan información sobre la posición, tamaño y relación entre los objetos representados en las mismas. Es al intentar procesar estas imágenes con medios informáticos cuando reparamos en la enorme eficiencia de nuestro cerebro para interpretarlas.

Las imágenes empleadas en teledetección son captadas por sensores instalados, generalmente, en satélites. Este tipo de imágenes, difíciles de captar por otros medios, nos permiten estudiar la superficie terrestre y los cambios que se producen en ella: cambios de longitud, área, color, brillo, incluso en algunos casos la profundidad y altura de los distintos elementos. La teledetección es la ciencia que estudia este tipo de imágenes con el fin de extraer información útil para una amplia variedad de aplicaciones. Una posible definición de teledetección es la siguiente: *Teledetección* es una disciplina que busca obtener información de la superficie terrestre y las superficies de agua a través de imágenes aéreas usando la radiación reflejada o emitida por la superficie de la Tierra de una o más regiones del espectro electromagnético.

La teledetección nos permite conocer mejor el medio que nos rodea, favoreciendo el estudio de los múltiples procesos que afectan al planeta. La instalación de sensores en plataformas orbitando en el espacio permite el estudio de zonas extensas difíciles o imposibles de observar a simple vista desde la Tierra. Además, estos sensores permiten captar otros tipos de información no percibida por el ojo humano y muy útiles para el estudio de la superficie terrestre y la atmósfera.

2.1 Componentes de un sistema de teledetección

Un sistema de teledetección consta de los siguientes elementos:

1. Fuente de energía. Es el origen de la radiación electromagnética que capta el sensor. Hablamos de *teledetección pasiva* si la fuente de energía es externa al sensor que capta la radiación y de *teledetección activa* en caso

de que la radiación captada sea previamente emitida por el sensor. La fuente de energía más importante es el Sol compuesta por distintos tipos de radiación, incluyendo la región visible del espectro (en la que se encuentran la luz azul, verde y roja) así como otros tipos de radiación (como los infrarrojos) que quedan fuera del rango de la visión humana. Esta radiación solar debe atravesar la atmósfera para alcanzar la superficie terrestre. En este recorrido, parte de la radiación es absorbida y otra parte es dispersada por la atmósfera.

2. Superficie terrestre. La radiación emitida por la fuente de energía es reflejada y se une a la radiación emitida por las distintas masas de vegetación, suelos, agua o construcciones humanas, de acuerdo con sus características físicas.
3. Sistema sensor, compuesto por el sensor y la plataforma que lo alberga. Se encarga de captar la energía procedente de la superficie terrestre, codificarla, y almacenarla o enviarla directamente al sistema de recepción. En su recorrido, la radiación emitida por el Sol ha sido de nuevo en parte absorbida y dispersada por la atmósfera en su viaje de vuelta. Parte de la radiación recibida por el sensor, proviene también de la energía dispersada por la atmósfera sin haber alcanzado la superficie terrestre. Por tanto, la radiación registrada por el sensor podría no corresponder exactamente a la radiación reflejada por la superficie terrestre ni a los colores observados en ella. De ahí que las imágenes captadas por los sensores no sean interpretables por medio de una simple observación.
4. Sistema de recepción-comercialización. Es donde se recibe la información transmitida por la plataforma del sensor, se graba en un formato apropiado, se le aplican distintas correcciones y se distribuye a los interesados. Hay dos tipos de corrección de los datos.
 - La *corrección geométrica* es aplicada para disminuir o eliminar errores en las posiciones relativas de los píxeles debidos a:
 - i. Geometría de visión del sensor.
 - ii. Variaciones del terreno.
 - La *corrección radiométrica* considera las variaciones en las intensidades o niveles digitales (NDs) de los píxeles, que no son causadas por el objeto o la escena que está siendo captada, debidas a:
 - i. Diferencia de sensibilidad o mal funcionamiento de los detectores.
 - ii. Efectos topográficos.
 - iii. Efectos atmosféricos.
5. Intérprete o analista. Convierte los datos en información temática de interés ya sea visual o digitalmente, de cara a facilitar la evaluación del problema en estudio. El intérprete es un experto con conocimientos de la información ofrecida por este tipo de imágenes y es el encargado de elaborar mapas,

imágenes o cualquier otro tipo de información fácilmente interpretable por usuarios no expertos en teledetección.

6. Usuario final, encargado de analizar el documento fruto de la interpretación, así como de dictaminar sobre las consecuencias que de él se deriven.

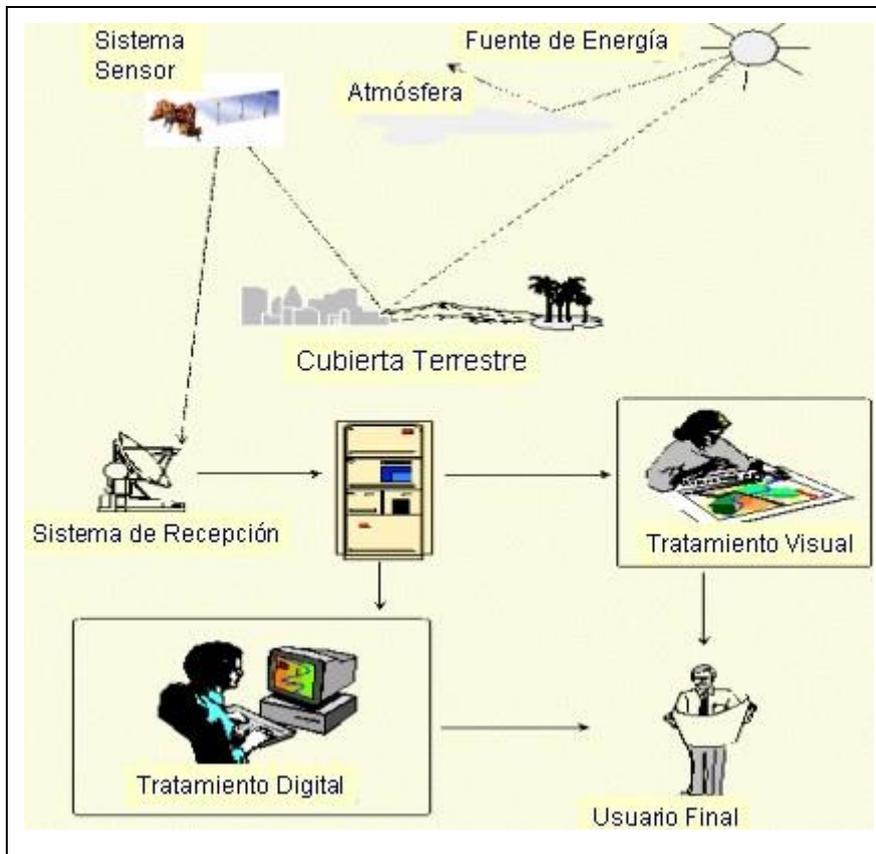


Figura 2-1. Componentes de un sistema de teledetección.

2.2 Ventajas de la teledetección espacial

La teledetección desde satélite tiene numerosas aplicaciones debido a las ventajas que ofrece frente a otros medios de observación más convencionales, como la fotografía aérea o los trabajos de campo. Las características de un sistema de teledetección espacial son:

- **Cobertura global y exhaustiva de la superficie terrestre.** La teledetección espacial es una fuente de información homogénea (un único sensor en la misma plataforma puede captar toda la información) y exhaustiva (cubre todo el planeta). Esto permite su uso en estudios globales tales como los procesos de deterioro de la capa de ozono o el calentamiento de la Tierra.
- **Amplia perspectiva y observación multiescala.** Una escena de satélite abarca grandes extensiones de suelo dependiendo de la resolución espacial y altura del mismo, obteniendo desde imágenes de precisiones por debajo

del metro y extensiones de algunos kilómetros cuadrados, a imágenes de menor precisión con áreas de varios millones de kilómetros cuadrados.

- **Adquisición de datos en regiones no visibles del espectro electromagnético.** Los sensores de los satélites permiten obtener imágenes sobre tipos de energía no accesibles al ojo humano o la fotografía convencional, como son los infrarrojos medio y térmico, o las microondas. Son bandas del espectro que son muy útiles en estudios medioambientales.
- **Periodicidad en la adquisición de imágenes.** Los satélites en su recorrido en órbita visitan cualquier zona de la tierra cada cierto período. Esto permite obtener las llamadas *imágenes multitemporales* que son imágenes de la misma área de estudio captadas en diferentes instantes de tiempo. Esta característica es muy útil en el estudio de la evolución de procesos naturales, como la desertización, los desastres naturales, incendios e inundaciones, y para la detección de cambios, tema de este proyecto. El disponer de distintas imágenes de la misma zona también permite la elaboración de *modelos de elevación del terreno* o incluso crear representaciones 3D de la superficie terrestre.
- **Rápida obtención de las imágenes y en formato digital.** La información captada por los sensores se almacena en formato digital y puede ser enviada rápidamente a las estaciones receptoras si se encuentra dentro del radio de alcance de las antenas de las mismas. Esta posibilidad de adquirir imágenes de una forma casi inmediata es muy útil en la gestión de desastres naturales. Al encontrarse en formato digital permiten su interpretación y estudio con medios y tecnologías informáticas.

2.3 Aplicaciones de la teledetección

Algunas de las muchas aplicaciones de la teledetección se listan a continuación:

- **Agricultura:**
 - Seguimiento de cultivos.
 - Inventariado y obtención de estadísticas sobre los cultivos.
 - Análisis de estado del suelo.
 - Informes de daños ante desastres naturales.
 - Cálculo de índices de vegetación.
 - Análisis del vigor de los cultivos.
- **Bosques:**
 - Inventariado de árboles y bosques.
 - Análisis de la salud de los bosques.
 - Creación de mapas de clasificación.
 - Detección de cambios.

- Mapas de riesgo de incendios.
 - Mapas de daños por incendio.
- **Geología:**
 - Mapas de placas tectónicas.
 - Mapas de clasificación del suelo.
- **Cartografía:**
 - Mapas cartográficos.
 - Mapas catastrales.
 - Mapas de carreteras e infraestructuras.
 - Modelos digitales de elevación.
 - Elaboración de mapas temáticos de cambios.
- **Medio ambiente:**
 - Mapas de uso del suelo.
 - Mapas hidrológicos.
 - Mapas de hábitats de animales salvajes.
 - Mapas de contaminación del suelo.
 - Análisis de contaminación del aire y de emisiones.
 - Mapas de condición de las superficies de agua.
 - Análisis de humedales.
 - Análisis de procesos de desertificación.
 - Identificación y control de canteras.
 - Identificación y control de vertederos.
 - Control de erosión de las costas.
 - Batimetría (mapas de relieve del suelo marino).
 - Estudio de las corrientes oceánicas.
 - Mapas de temperatura de la superficie del mar.
- **Análisis y gestión de riesgos.**
 - Cálculo de la extensión de zonas afectadas en caso de inundaciones en 2D/3D.
 - Mapas de permeabilidad.
 - Mapas de temperatura de los volcanes.
 - Mapas de movimientos y deslizamientos de tierras.
 - Mapas de daños por terremotos.
 - Mapas de vertidos de aceites y combustibles.
- **Defensa y seguridad.**
 - Mapas para la elaboración de planes de emergencia en catástrofes.
 - Mapas de criminalidad y accidentes.
- **Telecomunicaciones**
- **Visualización**
 - Elaboración de vistas panorámicas.
 - Elaboración de modelos en 3 dimensiones de la superficie terrestre.
 - Vuelos virtuales.

2.4 Radiación electromagnética

A excepción de los cuerpos en el cero absoluto, todos los objetos emiten *radiación electromagnética*, y algunos también reflejan la radiación emitida por otros objetos. Todas las aplicaciones de la teledetección se basan en el registro de esta radiación emitida o reflejada, y en el conocimiento del comportamiento de ésta al atravesar la atmósfera y de cómo interactúa con los distintos elementos (masas de agua, vegetación, estructuras, rocas, suelo, etc.) de la superficie terrestre.

El tipo más familiar de radiación electromagnética es la *luz visible*, que corresponde sólo a una pequeña, pero muy importante, porción del espectro electromagnético. La energía electromagnética es generada por distintos mecanismos, entre ellos, las reacciones nucleares producidas en el Sol. Esta radiación viaja por el espacio sin sufrir cambios, atraviesa la atmósfera y alcanza la superficie de la Tierra. Parte de esta radiación es reflejada y puede ser captada por sensores para formar imágenes. Es de esta forma como se captan las fotografías y podemos ver los objetos que nos rodean. Otra parte de la radiación emitida por el Sol es absorbida por la superficie terrestre y re-emitida en forma de energía térmica. Esta energía térmica también puede ser captada para obtener otro tipo de imágenes. Otra forma de radiación empleada en la teledetección es la generada directamente por el hombre como ocurre en la obtención de imágenes *radar*.

En la Figura 2-3 se muestra un gráfico con la clasificación del *espectro electromagnético* y las longitudes de onda de cada tipo de radiación.

Una *onda electromagnética* tiene cuatro propiedades que la describen:

- Longitud de onda (λ): es la distancia que recorre el pulso mientras una partícula del medio que recorre la onda realiza una oscilación completa.
- Frecuencia (f): es el número de oscilaciones que efectúa cualquier partícula del medio de propagación de la onda en un segundo.
- Período (T): es el tiempo que tarda una oscilación de la onda.
- Amplitud (A): es el valor máximo o mínimo alcanzado por la onda.

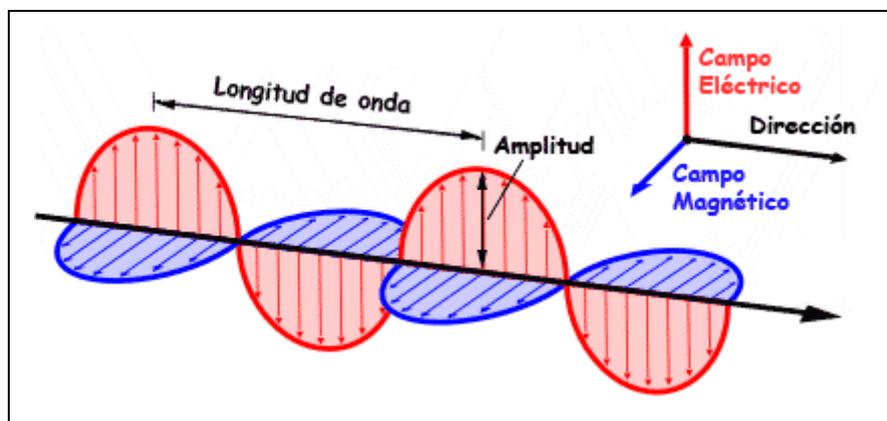


Figura 2-2. Onda electromagnética.

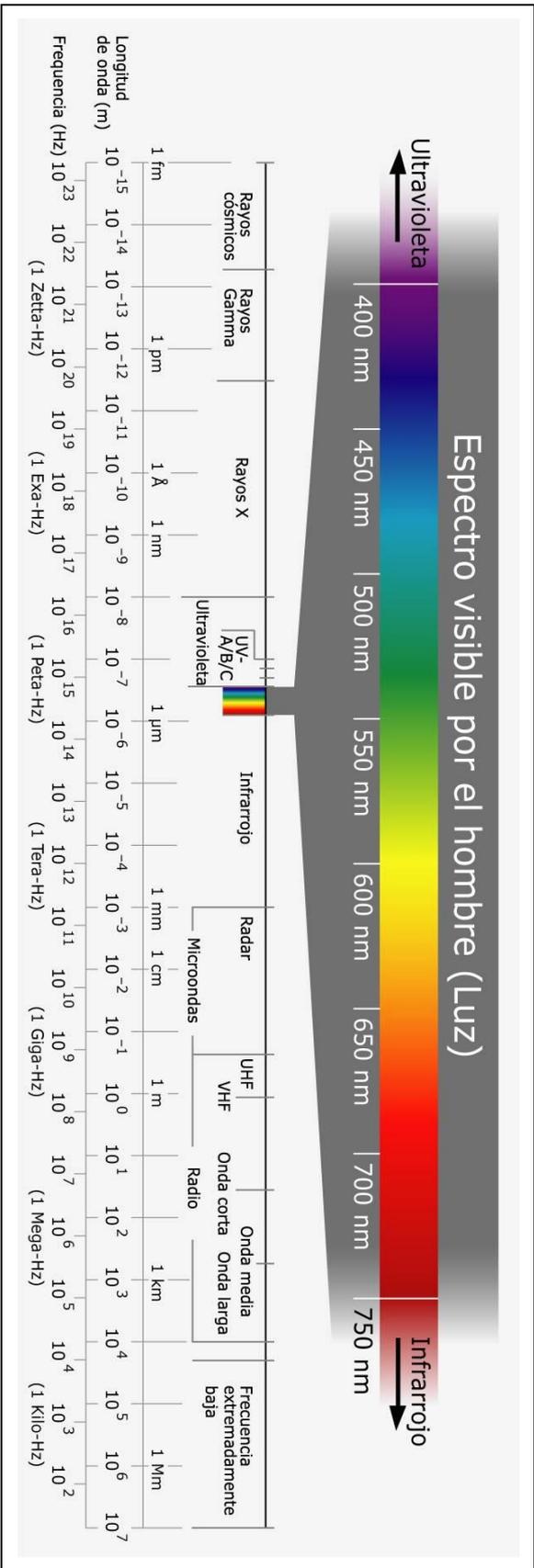


Figura 2-3. Espectro electromagnético.

Como hemos visto, el espectro electromagnético puede dividirse en distintos tipos de radiación en función de su longitud de onda. Es difícil establecer divisiones claras (excepto para el rango de la luz visible), de ahí que se puedan encontrar distintas propuestas de clasificación, aunque sólo varían en el rango de longitudes de onda. Algunas clasificaciones son la de la Figura 2-3 mostrada anteriormente y la de la Tabla 2-1 al final de esta sección.

Dos categorías importantes no mostradas en la tabla son el *espectro óptico*, de 0.30 a 15 μm , que corresponde a las longitudes de ondas que pueden ser reflejadas y refractadas con lentes y espejos; y el *espectro reflexivo* que abarca de los 0.38 a los 3.0 μm aproximadamente y define el rango del espectro solar que es reflejado y usado directamente en la teledetección.

Espectro	Límites
Rayos gamma	< 0.03 nm
Rayos X	0.03-300 nm
Radiación ultravioleta	0.30-0.38 μm
Luz visible	0.38-0.72 μm
Radiación infrarrojo	
Infrarrojo cercano	0.72-1.30 μm
Infrarrojo medio	1.30-7.00 μm
Infrarrojo lejano	7.00-1.000 m (1 mm)
Radiación microondas	1 mm-30 cm
Radio	30 cm

Tabla 2-1. Clasificación del espectro electromagnético.

2.4.1 Espectro ultravioleta

El espectro ultravioleta corresponde a longitudes de onda inferiores a la luz violeta del espectro visible. Tiene la propiedad de inducir fluorescencia (radiación visible) en ciertos materiales, de ahí que puede ser utilizada en teledetección. No obstante, la radiación ultravioleta se ve muy afectada por la atmósfera de la Tierra y por tanto no es usada con frecuencia.

2.4.2 Espectro Visible

Los límites del espectro visible son definidos por la sensibilidad del sistema visual humano. Las propiedades ópticas de la radiación visible fueron por primera vez investigados por Isaac Newton quién con sus experimento descubrió que la luz visible podía ser dividida en tres segmentos. Estos tres segmentos se conocen como *primarios aditivos*, definidos por los siguientes rangos de longitudes de onda: 0.4 a 0.5 μm (azul), 0.5 a 0.6 μm (verde) y 0.6 a 0.7 μm (rojo). La combinación de los primarios aditivos en iguales proporciones da lugar a la luz blanca. El *color* de un objeto se debe

al color de la luz que refleja. Decimos que un objeto es rojo cuando refleja la luz roja y absorbe las demás.

Por otro lado, existen también los llamados *primarios substractivos*. Al igual que con los primarios aditivos se combinan para obtener los colores, pero en este caso los colores se restan, es decir, si con los primarios aditivos combinados en iguales proporciones obteníamos el blanco, con los primarios sustractivos obtenemos el negro. Son usados por ejemplo, en la impresión en papel, donde se combinan para restar al blanco luminosidad y así poder obtener los distintos colores. Los primarios substractivos absorben un tercio del espectro visible (y reflejan por tanto los otros dos). Son primarios substractivos el amarillo (absorbe el azul, y refleja el rojo y verde), el magenta (absorbe el verde, y refleja el rojo y azul) y el cian (absorbe el rojo, y refleja el azul y verde).

2.4.3 Infrarrojos

Corresponde a longitudes de onda mayores de la luz roja del espectro visible. Fueron descubiertos en 1800 por el astrónomo William Herschel. Es un segmento del espectro electromagnético mucho más amplio que el espectro visible. Debido a su amplitud los infrarrojos tienen distintas propiedades, de ahí que se subdividan en varias categorías. Las dos primeras, conocidas como *infrarrojo cercano* e *infrarrojo medio*, son las más cercanas a la radiación visible y se comportan de una manera similar respecto a los sistemas ópticos. Además, el infrarrojo cercano puede ser captado con películas, filtros y sensores similares a los empleados para la luz visible.

Por otro lado, encontramos el *infrarrojo lejano* con longitudes de ondas lejanas al espectro visible y que se extiende hasta las conocidas *microondas*. Esta radiación es muy diferente al resto de infrarrojos y a la luz visible. Los infrarrojos cercanos se corresponden a radiación del sol reflejada por la superficie de la Tierra, mientras que los infrarrojos lejanos, corresponden a radiación emitida por la Tierra. Se corresponden con el calor o energía térmica emitida por la superficie terrestre. Por este motivo, el infrarrojo lejano se conoce a veces como *infrarrojo emitido*.

2.4.4 Microondas

Es la radiación de mayor longitud de onda usada en teledetección. Las longitudes de onda más cortas se comportan de una manera similar al infrarrojo lejano, mientras que las más largas se solapan con las conocidas ondas de *radio*. Fueron estudiadas por los físicos James Clerk Maxwell y Heinrich Hertz.

2.5 Interacciones con la atmósfera

Toda la radiación usada en teledetección debe pasar a través de la atmósfera. Los efectos de distorsión de la radiación electromagnética son mayores en el caso de sensores instalados en satélites que aquellos instalados en aviones ya que la radiación debe atravesar dos veces la atmósfera. Esto puede suponer una drástica

reducción de la calidad de las imágenes y datos generados en el sensor. La niebla, contaminación, polvo y otras partículas en la atmósfera afectan al color y brillo de las imágenes. Incluso con cielos despejados, la radiación se ve afectada por ciertos fenómenos físicos que se detallan a continuación y son representados en la imagen siguiente:

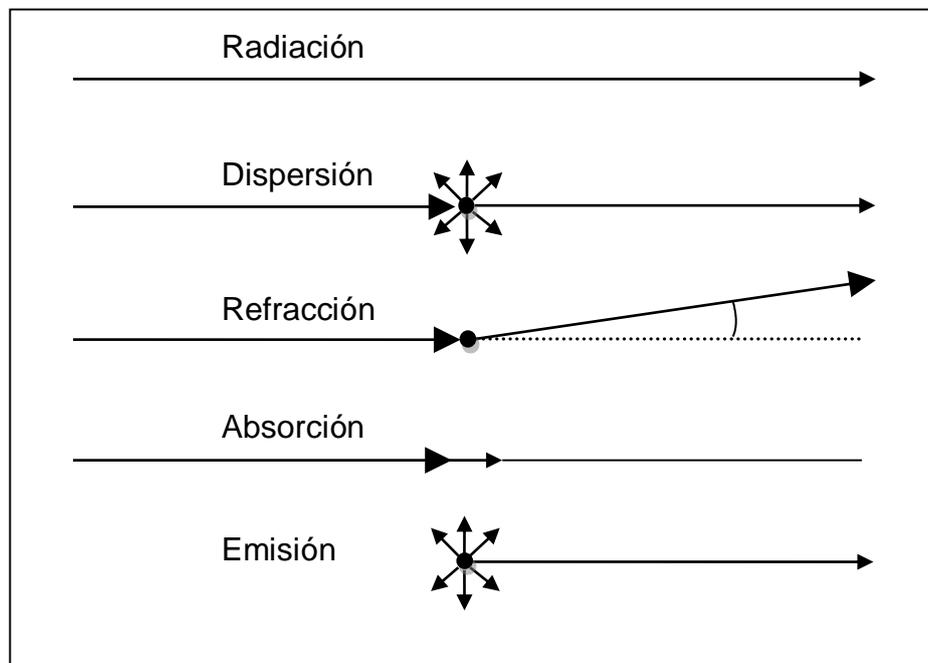


Figura 2-4. Interacciones de la radiación electromagnética con la atmósfera.

2.5.1 Dispersión

Es la redirección de la energía electromagnética por partículas suspendidas en la atmósfera o por grandes moléculas de los gases atmosféricos. La dispersión depende del tamaño de estas partículas, su abundancia, la longitud de onda de la radiación y la profundidad de la atmósfera a través de la cual viaja la radiación. Este efecto provoca que parte de la radiación del Sol sea reflejada al espacio y otra a la superficie de la Tierra.

Lord J. W. S. Rayleigh demostró que en una atmósfera perfectamente limpia, consistente sólo en gases atmosféricos, la dispersión de la luz es mayor cuanto menor es la longitud de onda. La luz azul por ejemplo, es dispersada unas 4 veces más que la luz roja y la radiación ultravioleta, unas 16 veces más. La *dispersión Rayleigh* ocurre cuando las partículas atmosféricas tienen un diámetro pequeño respecto a la longitud de onda de la radiación. La ley de Rayleigh establece que la luz dispersada es inversamente proporcional a la cuarta potencia de la longitud de onda de la radiación. La dispersión de este tipo es la causa del color azul del cielo y de los colores rojo y naranja de los atardeceres.

Otro tipo de dispersión habitual es la *dispersión Mie* causada por grandes partículas presentes en la atmósfera como polvo, polen, humo y gotas de agua. Estas partículas son varias veces más grandes que las responsables de la dispersión

Rayleigh y tienen un diámetro similar a las longitudes de onda de la radiación dispersa. Sus efectos son dependientes de la longitud de onda pero en menor grado que en la dispersión Rayleigh.

Hablamos de *dispersión no selectiva* cuando es provocada por grandes partículas (con diámetros superiores a las longitudes de onda de la radiación dispersada). En este caso la dispersión afecta por igual a las distintas longitudes de ondas. Es la responsable del color blanco de las nubes donde las grandes gotas de agua dispersan toda la radiación que reciben.

2.5.2 Refracción

La *refracción* es el cambio en la dirección de propagación de un fenómeno ondulatorio como la luz, cuando atraviesa la superficie de separación entre dos medios. Las diferentes densidades de las distintas capas de la atmósfera causadas por distintos valores de humedad, temperatura y composición, da lugar a que la radiación electromagnética se desvíe en su trayectoria.

2.5.3 Absorción

La *absorción* de la radiación se da cuando la atmósfera impide o atenúa fuertemente la radiación electromagnética que la atraviesa. La atmósfera se comporta como un filtro selectivo a distintas longitudes de onda, de tal forma que en algunas bandas del espectro elimina prácticamente cualquier posibilidad de observación remota. Los principales causantes de esta absorción son:

- Oxígeno (O₂), que filtra las radiaciones ultravioleta por debajo de 0.1 μm, así como pequeños sectores en el infrarrojo térmico y las microondas.
- Ozono (O₃), responsable de la eliminación de la energía ultravioleta (inferior a 0.3 μm), así como un sector de las microondas (en torno a los 27 mm).
- Vapor de agua (H₂O), con una fuerte absorción en el infrarrojo térmico (> 15 μm), con importantes efectos en el infrarrojo medio, entre 2.5 y 4.5 μm.
- Aerosoles atmosféricos, que principalmente absorben en el espectro visible.

Como consecuencia de esta absorción, la observación espacial se reduce a determinadas bandas del espectro, conocidas como *ventanas atmosféricas*, en donde la transmitancia de la atmósfera es suficientemente alta. Distintos autores proponen distintos rangos para las principales ventanas atmosféricas. Un posible listado de estas ventanas podría ser el mostrado en la Tabla 2-2.

Longitudes de onda	Banda
0.35 – 1.35 μm	Espectro visible, Infrarrojo cercano
1.5 – 1.8 μm	Infrarrojo medio (reflejado)

2.0 – 2.4 μm	Infrarrojo medio (reflejado)
2.9 – 4.2 μm	Infrarrojo medio (emitido)
4.5 – 5.5 μm	Infrarrojo medio (emitido)
8 – 14 μm	Infrarrojo lejano
> 20 μm	Microondas

Tabla 2-2. Ventanas atmosféricas.

Estas ventanas atmosféricas son idóneas para realizar procesos de teledetección, por lo que el diseño de los sensores espaciales tiende a ajustarse a estas bandas, evitando interferencias extrañas al fenómeno que pretenden observarse, salvo la presencia de nubes, que no pueden evitarse y absorben en todo el espectro óptico. Si se pretende por el contrario observar la atmósfera en lugar de la superficie terrestre, los sectores espectrales más convenientes son precisamente aquellos en donde la absorción atmosférica es alta. Bajo este enfoque, lo que se pretende es derivar a partir de teledetección estimaciones de la concentración de los distintos gases en la atmósfera, analizando sus espectros de absorción. La Figura 2-5 representa la absorción de la radiación en la atmósfera y muestra las principales ventanas para la teledetección.

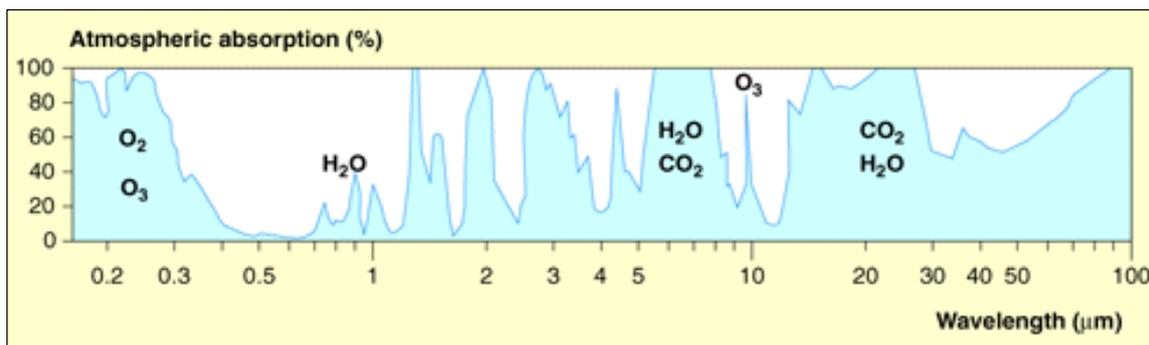


Figura 2-5. Absorción de la radiación en la atmósfera.

2.5.4 Emisión

Al igual que cualquier cuerpo por encima del cero absoluto, la atmósfera emite energía, por lo que ese parámetro debe considerarse para separarlo de la radiación emitida por el suelo. Este tipo de fenómenos es importante al trabajar con infrarrojo térmico. No afecta a aquellos sensores que trabajan únicamente con radiación reflejada.

2.6 Interacciones con la superficie terrestre

Cuando la energía electromagnética alcanza la superficie terrestre, es reflejada, absorbida o transmitida. Estos procesos dependen de la naturaleza de la superficie, de la longitud de onda de la energía y del ángulo de iluminación.

Todo en la naturaleza tiene su propia distribución de radiación reflejada, emitida y absorbida. Esta distribución se conoce como *firma espectral* y permite distinguir un elemento de otro en la superficie terrestre, así como ofrecernos información sobre su forma, tamaño, y otras propiedades físicas y químicas. Pueden existir objetos diferentes con una firma espectral similar y por tanto no sería suficiente su estudio para distinguirlos. Las interacciones que determinan la firma espectral de los distintos elementos de la superficie terrestre son:

2.6.1 Reflexión

La *reflexión* se da cuando un rayo de luz incide en una superficie no transparente. La reflexión depende del tamaño de las irregularidades de la superficie (rugosidad) en relación con la longitud de onda de la radiación incidente. La relación entre la radiación incidente y la reflejada por una superficie se conoce como *reflectancia*.

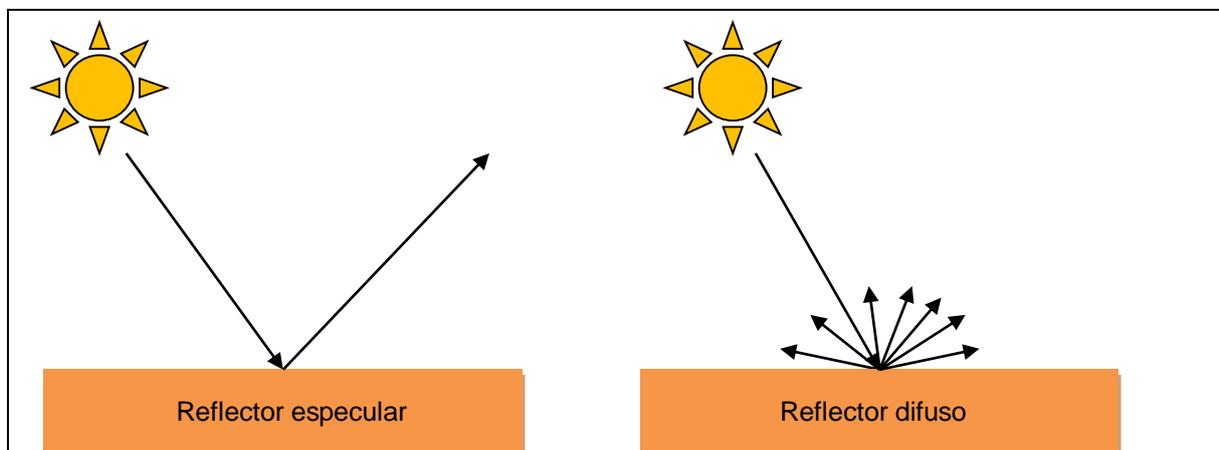


Figura 2-6. Tipos de reflectores.

Si la superficie es suave se da la llamada *reflexión especular* en la que toda o casi toda la radiación incidente es redirigida en una misma dirección. En este caso el ángulo de incidencia es igual al ángulo de reflexión. Para la radiación visible, la reflexión especular se da en espejos, metal con superficie suave y masas de agua en calma.

Si la superficie es rugosa, actúa como un reflector *difuso* o *isotrópico*. La energía es dispersada más o menos equitativamente en todas las direcciones. La mayoría de las superficies naturales actúan como reflectores difusos. Un reflector difuso perfecto (conocido como *superficie Lambertiana*) tendría igual brillo cuando es observado desde cualquier ángulo.

2.6.2 Transmisión

Hablamos de *transmisión* cuando la radiación pasa a través de una sustancia sin experimentar una absorción significativa. Se conoce como *transmitancia* a la relación entre la radiación incidente y la radiación transmitida. Un ejemplo de

transmisión se da en las hojas que son generalmente opacas a la radiación visible pero transmiten cantidades significativas de radiación en el infrarrojo.

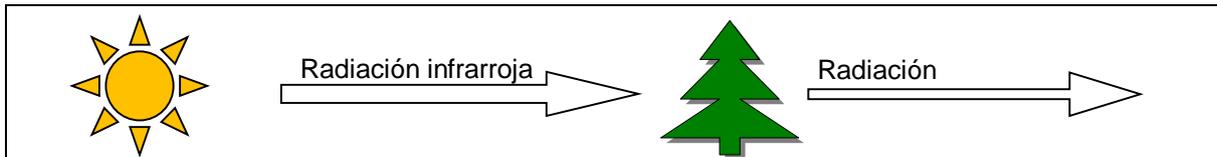


Figura 2-7. Transmisión de la radiación.

2.6.3 Absorción

Tal como ocurría en la atmósfera, en la superficie terrestre también se da la *absorción* de la radiación. Es la responsable del color que percibimos. La relación entre la radiación incidente y la radiación absorbida se conoce como *absortancia*.

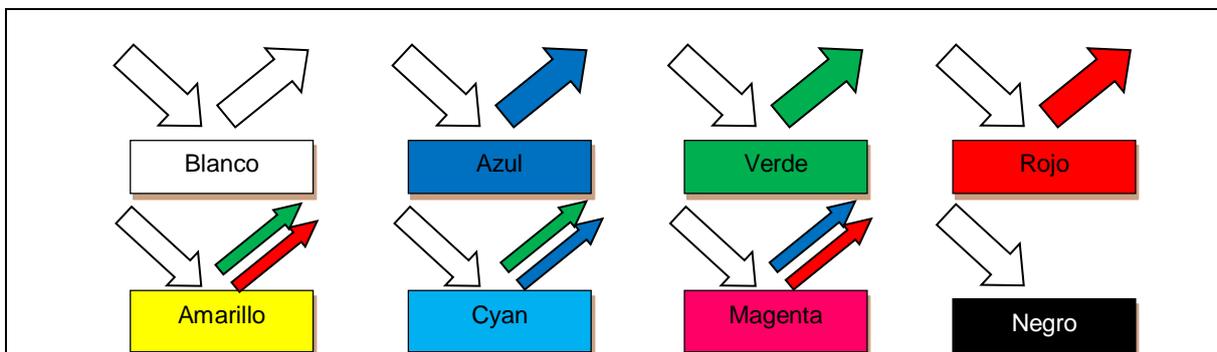


Figura 2-8. Absorción de la radiación.

2.6.4 Emisión

La superficie terrestre, como cualquier otro cuerpo (y como hemos visto también la atmósfera), emite radiación electromagnética. A este proceso se conoce como *emisión*. A la radiación emitida por la superficie terrestre hay que sumar la radiación incidente absorbida en el proceso comentado en el apartado anterior. Esta radiación re-emitida tiene una longitud de onda mayor que la radiación original. La relación entre la radiación emitida y la radiación emitida por un cuerpo negro a la misma temperatura se conoce como *emitancia*.

2.7 Satélites de Alta Resolución

Las imágenes obtenidas por los satélites de percepción remota o teledetección ofrecen una perspectiva única de la Tierra, sus recursos y el impacto que sobre ella ejercen los seres humanos. La teledetección por satélite ha demostrado ser una fuente rentable de valiosa información para numerosas aplicaciones SIG.

En este nuevo siglo podemos contar con docenas de satélites de teledetección en órbita alrededor de la Tierra, cada uno de ellos dedicado a captar imágenes de tipo muy específico. El valor práctico y la multiplicidad de aplicaciones de las imágenes se incrementan a medida que se lanzan nuevos satélites. Cada vez se dispone de más

imágenes en una cantidad creciente de tamaños de escena, resoluciones espectrales, frecuencias de paso y detalles espaciales.

Las ventajas de las imágenes de satélite son las siguientes:

- **Digital:** Casi todas las imágenes procedentes de satélites se adquieren digitalmente. Esto implica que no es necesario efectuar conversiones de datos, escaneos o digitalizaciones. Con una preparación mínima las imágenes quedan listas para ser cargadas directamente en un SIG. Dada su naturaleza digital, estas se procesan manipulan y realzan para extraer de ellas detalles e información que otras fuentes no detectan.
- **Rápido:** Debido a que los satélites se encuentran en órbitas muy altas y en plataformas estables, pueden recopilar información de enormes áreas en tan solo unos cuantos segundos.
- **Económico:** Para zonas extensas, las imágenes de satélite resultan más económicas que la fotografía aérea y los levantamientos cartográficos o topográficos.
- **Global:** Los satélites no están limitados por fronteras políticas ni geográficas y se encuentran en órbitas que permiten sobrevolar todas las zonas del planeta.
- **Actualizado:** Se puede disponer de una imagen de satélite unas semanas después de su toma anterior.
- **Sinóptico:** Los satélites captan en una sola imagen detalles de la cubierta del suelo, carreteras, infraestructuras, etc. Con esto puede estudiarse la relación entre los diferentes elementos observados y en las diferentes escalas de interés.
- **Preciso:** Dado que una imagen de satélite es capturada sin intervención humana, la información que contiene es una representación objetiva e imparcial de los objetos y detalles de la superficie terrestre.
- **Flexible:** El tratamiento y la extracción de información de las imágenes de satélite pueden ser tan complicados o sencillos como se desee. Según la aplicación pueden hacerse desde interpretaciones visuales hasta clasificaciones automáticas expertas.

2.7.1 Tipos de satélite

Entre las variadas formas de clasificar los sensores remotos, una de las más habituales considera la forma de recibir la energía procedente de las distintas cubiertas de la superficie terrestre. Se habla de sensores *pasivos* cuando se limitan a recibir la energía procedente de un foco exterior a ellos. Si el sensor es quien emite la

energía que una vez reflejada es recibida por el mismo, se conoce como sensores *activos*. Los sensores activos son más flexibles al no depender tanto como los primeros de las condiciones exteriores al sistema sensor-Tierra.

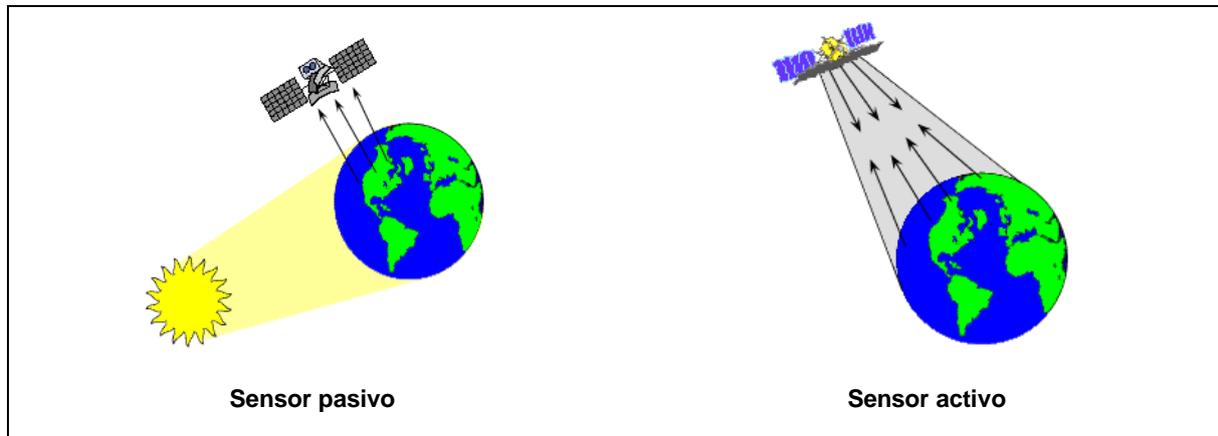


Figura 2-9. Tipos de sensores.

Otra forma de clasificar los sensores es atendiendo a su procedimiento para grabar la energía recibida. Los sensores pasivos pueden clasificarse en sensores fotográficos, óptico-electrónicos y de antena. Los primeros se basan en las mismas técnicas aplicadas en la fotografía convencional. El resto aunque poco conocidos se utilizan en sistemas como cámaras de vídeo, equipos de barrido o scanners en exploración médica, y radiómetros de microondas en radiodifusión. En cuanto a los sensores activos, el equipo más conocido es el *radar*, bastante utilizado en aplicaciones militares así como en control del tráfico aéreo. El radar trabaja en la región de las microondas. Por otra parte, el *líd*ar, opera con luz polarizada (láser) en el espectro óptico.

Los satélites pueden operar en diferentes clases de órbitas. Las órbitas más comunes para satélites de observación de la Tierra son las geostacionarias y las polares, pero algunos instrumentos también vuelan en órbitas inclinadas.

Una *órbita geostacionaria* es aquella en la que el satélite siempre está en la misma posición con respecto a la Tierra. El satélite orbita a una altura de aproximadamente 35790 Km porque esto hace que el período orbital (la duración de una órbita) sea igual al período de rotación de la Tierra (23h 56m 4.09s). Un satélite estacionario aporta a la teledetección la ventaja de que siempre ve la Tierra desde la misma perspectiva, lo que significa que puede registrar la misma imagen a breves intervalos. Esto es particularmente útil para observar las condiciones meteorológicas. Un inconveniente de las órbitas geostacionarias es la gran distancia a la Tierra, que reduce la *resolución espacial* que se puede lograr. Existen varios satélites meteorológicos distribuidos regularmente sobre la órbita geostacionaria, cubriendo todo el mundo y proporcionando una visión global.

Los satélites que vuelan en órbitas polares proporcionan una visión más completa de la Tierra, girando con una inclinación (ángulo entre el plano ecuatorial y

el plano de la órbita del satélite) cerca de la polar (con una inclinación de 90°). Orbitando a una altura de 700 a 800 Km, estos satélites cubren de mejor forma las partes del mundo más difíciles de cubrir in situ. Estos satélites operan en una órbita sincronizada con el Sol. El satélite pasa cada día por el ecuador y cada latitud a la misma hora solar local durante todas las estaciones del año. Esta característica permite la recolección regular de datos en horas consistentes así como las comparaciones a largo plazo. El plano orbital en una órbita sincronizada con el Sol debe también rotar aproximadamente un grado al día para mantenerse su inclinación con respecto a la Tierra.

Las *órbitas inclinadas* están entre las dos anteriores. Pueden estar determinadas por la región de la Tierra que es de mayor interés o por la latitud del lugar de lanzamiento. La altura de la órbita de estos satélites generalmente es del orden de unos cientos de kilómetros por lo que el período orbital es del orden de unas cuantas horas. Sin embargo, estos satélites no están sincronizados con el Sol y por tanto ven cada lugar de la Tierra a diferentes horas.

2.7.2 Resolución de un sensor

Podemos definir la resolución de un sensor como la capacidad para discernir objetos o detalles en las imágenes captadas. Un satélite o sistema sensor de mayor resolución que otro permitirá distinguir más fácilmente unos objetos de otros además de permitir una mayor precisión en su localización. Una imagen de satélite se caracteriza por las siguientes modalidades de resolución:

Resolución espacial

Este concepto designa al objeto más pequeño que se puede distinguir en una imagen. Está determinada por el tamaño del píxel de la imagen captada y se mide en metros sobre el terreno. Depende de la altura del sensor con respecto a la Tierra, el ángulo de visión, la velocidad de escaneado y las características ópticas del sensor. En los sensores de antena, su resolución depende del radio de apertura, de la altura de la plataforma y de la longitud de onda a la que trabajen. Cuanto mayor sea el radio, y menores la altitud y la longitud de onda, mayor será la resolución.

La resolución espacial, en la interpretación de las imágenes, marca el nivel de detalle que ésta ofrece. Sólo serán identificables sobre la imagen elementos que superen el tamaño de varios píxeles, aunque podrían detectarse elementos de tamaño similar o incluso más pequeño al del píxel cuando exista suficiente contraste radiométrico. La resolución espacial está en estrecha relación con la escala de trabajo y con la fiabilidad finalmente obtenida en la interpretación. Cuanto menor sea el tamaño del píxel, menor también la probabilidad de que éste sea un compuesto de dos o más cubiertas del terreno diferenciadas. Un píxel mixto estará definido por una señal intermedia a las distintas cubiertas que lo componen.

Por otro lado, se ha comprobado que en algunas cubiertas como las urbanas, el incremento de resolución espacial puede afectar negativamente a la interpretación de la imagen al incrementarse la heterogeneidad de elementos presentes en ella. De todas formas, una alta resolución espacial facilitaría la interpretación de la imagen en un análisis visual.

En la detección de cambios, una mayor resolución espacial permite detectar cambios en la superficie terrestre de menor área al ser identificables elementos que no serían visibles a resoluciones menores. Para estudios de cambios urbanos, son necesarias imágenes de alta resolución donde sean observables los distintos edificios y elementos de las ciudades, no siendo válidas otras de resoluciones inferiores que tan sólo podrían permitir el estudio de los límites del crecimiento urbano. Las imágenes de muy alta resolución permiten detectar la construcción de nuevos edificios, la apertura de nuevos viales, movimientos de tierra y un largo etcétera de información relevante para la planificación del desarrollo de las ciudades.

En la Figura 2-10 se muestra una zona de Huelva en el año 2007 a distintas resoluciones espaciales. Podemos observar que en la de mayor resolución se pueden apreciar más detalles como los vehículos u objetos de pequeños tamaño en la terrazas. En la Figura 2-11 podemos ver como quedarían representadas en píxeles la planta de un edificio a distintos tamaños de píxeles.

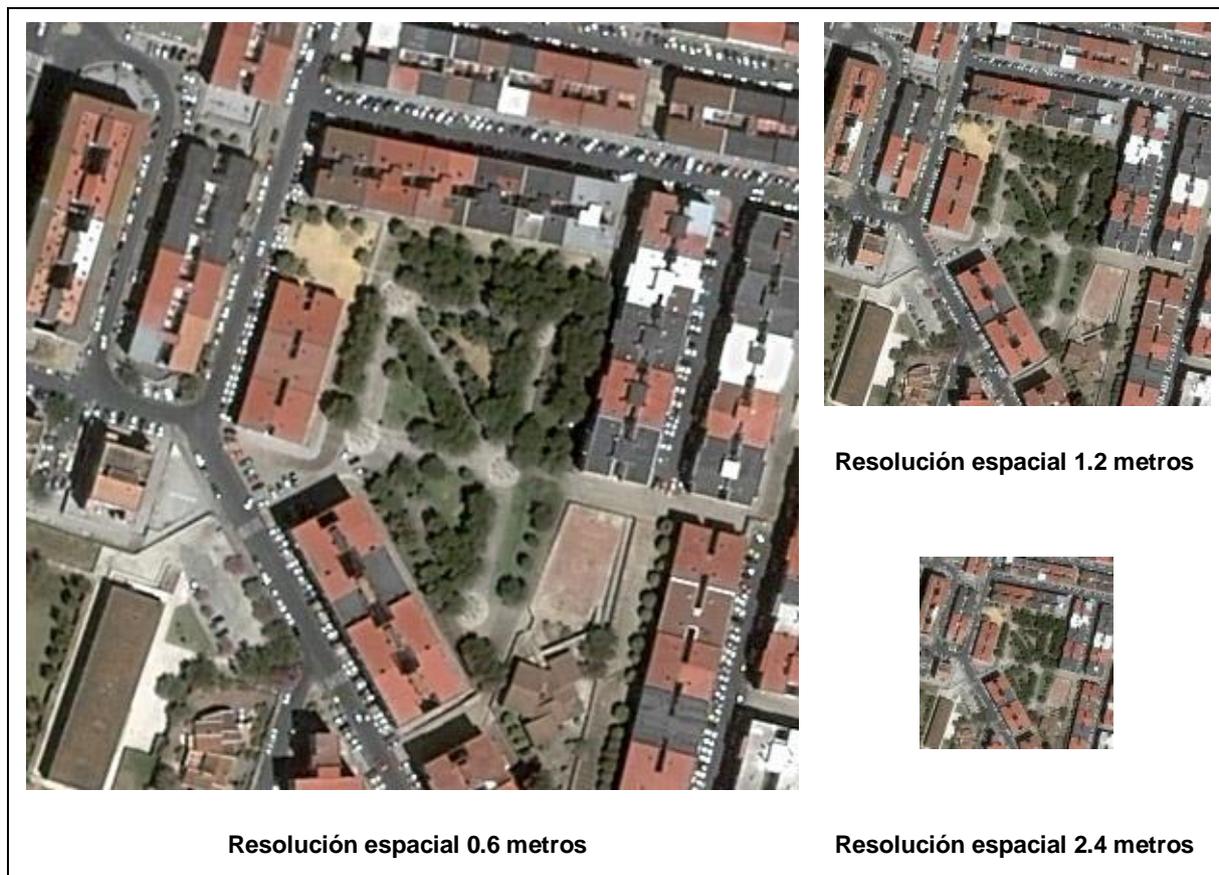


Figura 2-10. Resolución espacial.

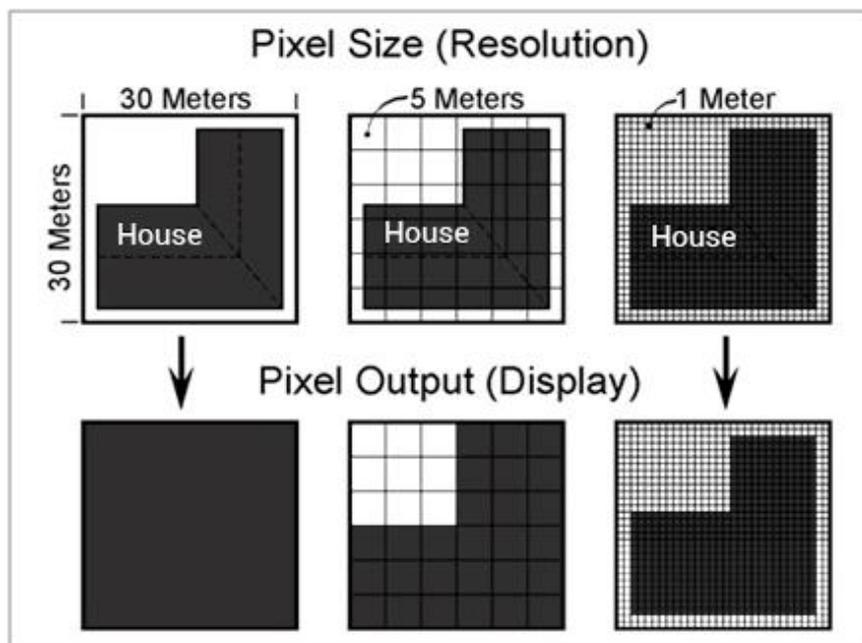


Figura 2-11. Salida obtenida para diferentes resoluciones.

Resolución espectral

Indica el número y anchura de las bandas espectrales que puede discriminar el sensor. Un sensor será tanto más idóneo cuanto mayor número de bandas proporcione, ya que facilita la caracterización espectral de las distintas cubiertas. A la vez, conviene que esas bandas sean suficientemente estrechas, con objeto de recoger la señal sobre regiones coherentes del espectro. Bandas muy amplias, supone registrar un valor promedio que puede encubrir la diferenciación espectral entre cubiertas de interés. La elección del número, anchura y localización de las bandas que incluye el sensor está estrechamente relacionada con los objetivos de su diseño. Para un sensor dedicado a fines meteorológicos, por ejemplo el Meteosat, basta situar una sola banda en el visible, puesto que no existen diferencias cromáticas en las nubes. Por el contrario, será conveniente que incluya una o dos bandas en el infrarrojo térmico (que le permitan conocer la temperatura de dichas nubes) y otra en el infrarrojo medio (para el control de la humedad en la atmósfera).

En la detección de cambios, bastaría con una banda en el espectro visible que permita obtener las diferencias de brillo en imágenes tomadas en diferentes instantes de tiempo. Un cambio en las propiedades de un objeto de la superficie terrestre se refleja en un cambio significativo en el brillo y, por tanto, en un cambio en la radiación captada por el sensor. La inclusión de más de una banda facilitaría el análisis visual de los cambios y la clasificación de los distintos objetos. Esta clasificación del suelo podría ser útil para detectar cambios en la naturaleza del mismo permitiendo, por ejemplo, detectar zonas de vegetación que han pasado a ser suelo urbano.

Resolución radiométrica

Se refiere a la sensibilidad del sensor, esto es, a su capacidad para detectar variaciones en la radiación que recibe. En el caso de los sistemas fotográficos, la resolución radiométrica del sensor se indica por el número de niveles de gris recogido en la película. Para los óptico-electrónicos, la imagen habitualmente se presenta en formato digital, por lo que la resolución radiométrica suele identificarse con el rango de valores que codifica el sensor. Esta resolución se expresa por el número de bits que precisa cada píxel de la imagen para ser almacenado. En la Figura 2-12 podemos ver una imagen con diferentes resoluciones radiométricas, es decir, con distinto número de niveles de gris.

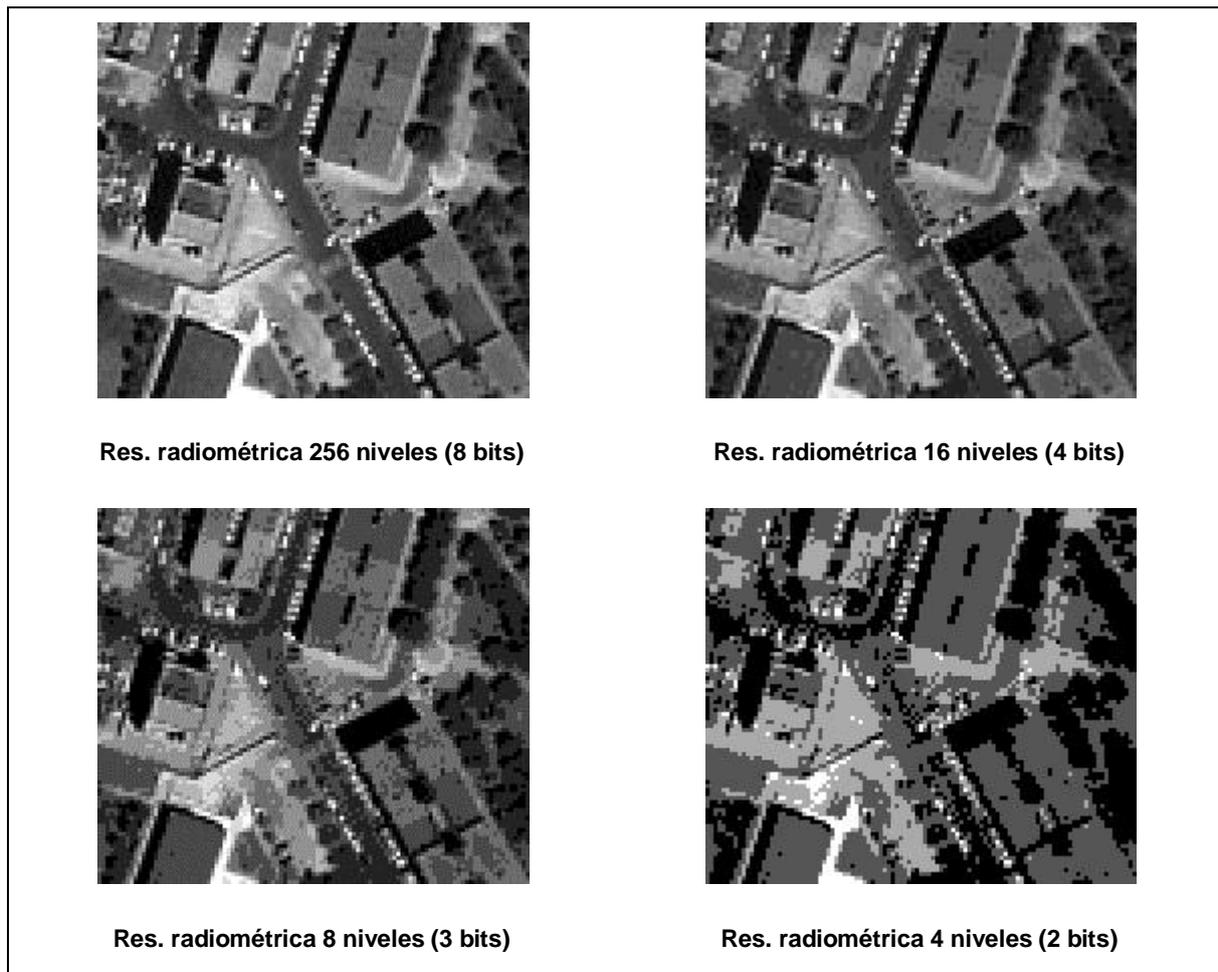


Figura 2-12. Resolución radiométrica.

La resolución radiométrica resulta de notable ayuda en la interpretación de imágenes, especialmente cuando realizamos análisis digital. El número de niveles de gris que son capaces de distinguir nuestros ojos no excede de 64, mientras que las tonalidades de color no superan las 200000. Algunos autores estiman que basta con 16 niveles por color para garantizar un análisis visual adecuado, sin embargo, cuando la interpretación es digital, el ordenador aprovecha todo el rango disponible, con lo que una mayor resolución radiométrica puede permitir discriminar objetos con niveles muy similares de reflectividad que no serían separables con otros sensores menos sensibles. Esta es la razón por la que para una buena detección de cambios es

necesaria una buena resolución radiométrica. Una mayor resolución permitirá diferenciar más cambios de distintas intensidades y reducirá por tanto el número de falsos positivos (se etiqueta un objeto como cambio cuando no lo es) y negativos (donde un cambio no es detectado como tal).

Resolución temporal

La *resolución temporal* alude a la frecuencia de cobertura que proporciona el sensor, es decir, a la periodicidad con la que éste adquiere imágenes de la misma porción de la superficie terrestre. El ciclo de cobertura está en función de las características orbitales de la plataforma (altura, velocidad, inclinación), así como del diseño del sensor, principalmente del ángulo total de apertura. Conviene tener en cuenta que la resolución temporal efectiva dependerá también de las condiciones atmosféricas, ya que con sensores ópticos o térmicos no podemos observar áreas cubiertas por nubes, por lo que en algunas zonas la cadencia de imágenes útiles puede ser notablemente más baja que el ciclo de adquisición del sensor.

La cadencia temporal de los sistemas espaciales varía de acuerdo a los objetivos fijados para el sensor. Los satélites meteorológicos por ejemplo, deben obtener imágenes en períodos cortos de tiempo ya que se dedican a observar un fenómeno muy dinámico. Para la detección de cambios esta resolución no es tan importante como las otras ya que, para poder llevarla a cabo, basta con tomar imágenes de la misma zona con una diferencia temporal de unos meses.

2.8 Principales satélites de alta resolución espacial

En este apartado, trataremos algunos de los satélites de mayor resolución espacial utilizados para adquirir imágenes empleadas habitualmente en la detección de cambios.

2.8.1 Quickbird

El *Quickbird* es el primero de una constelación de satélites que la empresa DigitalGlobe está desarrollando para ofrecer imágenes de la Tierra de alta resolución y precisión. Hasta el lanzamiento del WorldView 1 era el satélite comercial de mayor resolución espacial con un tamaño del píxel de 60 cm y el único con una resolución inferior al metro. El Quickbird es el segundo con su mismo nombre ya que el primero fracasó en su lanzamiento al no alcanzar el cohete que lo transportaba, la órbita deseada.

Fue lanzado el 18 de Octubre de 2001. Recoge imágenes durante las horas de sol. Está diseñado para cubrir grandes áreas con gran eficacia y precisión. El Quickbird puede adquirir anualmente imágenes de hasta 75 millones de metros cuadrados de la superficie terrestre.

Se encuentra en una órbita polar sincronizada con el sol con una inclinación de 98° respecto al ecuador a una altura de 450 Km. De esta forma puede tomar imágenes de la Tierra a la misma hora en todas las estaciones del año. El tiempo en recorrer la órbita completa es de 93.4 minutos. Cruza el ecuador a las 10:30 a.m. (hora local). Las especificaciones del satélite se recogen en la Tabla 2-3.

2.8.2 Worldview 1

El *Worldview 1* ofrece una muy alta resolución espacial con un tamaño del píxel de 50 cm. Tiene un único sensor pancromático y es capaz de captar hasta 750000 km² por día con una resolución máxima.

Fue lanzado el 18 de Septiembre de 2007 con éxito. Recorre una órbita polar sincronizada con el Sol a una altura de 496 Km y con una duración de 94.6 minutos. Cruza el ecuador a las 10:30 a.m. (hora local). Las especificaciones del satélite se recogen en la Tabla 2-3.

2.8.3 Worldview 2

Su lanzamiento se realizó el 8 de Octubre de 2009 y junto al *Worldview 1* y el *Quickbird* forma parte de la constelación de satélites de la empresa DigitalGlobe. Es el satélite comercial de mayor resolución espacial aunque para la distribución de sus imágenes son remuestreadas a 50 cm/píxel ya que resoluciones mayores solo están disponibles para el gobierno de los Estados Unidos. Facilita imágenes pancromáticas y multiespectrales con 8 bandas disponibles, las conocidas azul, verde, rojo e infrarrojo y 4 bandas nuevas. En las nuevas bandas encontramos una por debajo del azul, una por encima del rojo, un nuevo infrarrojo y una banda para el color amarillo. La respuesta espectral de las distintas bandas se muestra en la Figura 2-13. Recorre una órbita polar sincronizada con el Sol a una altura de 770 kilómetros y con un período de 100 minutos. Cruza el ecuador a las 10:30 a.m. (hora local). En la Tabla 2-3 se muestran las especificaciones del *Worldview 2*.

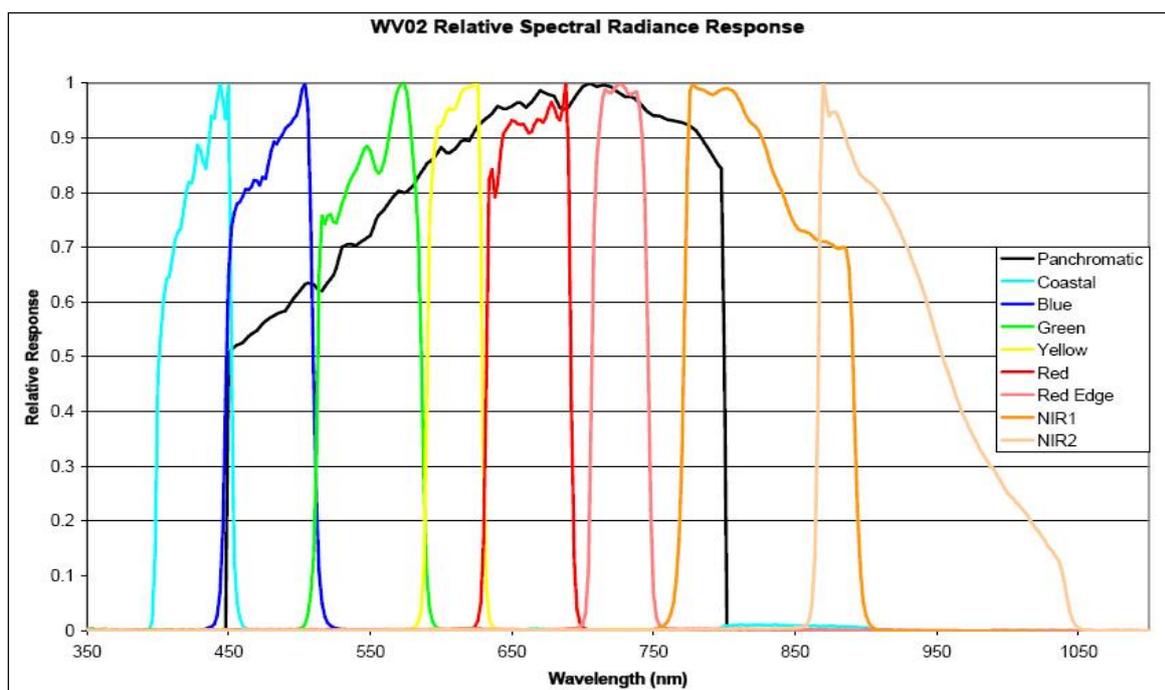


Figura 2-13. Respuesta espectral de las bandas de Worldview 2.

2.8.4 Worldview 3

El *Worldview 3* ya está en desarrollo y con un lanzamiento previsto para mediados del año 2014, formará parte de la constelación de satélites de DigitalGlobe. Adquirirá imágenes en una órbita a 617 km de altura. Además de la banda pancromática y las 8 bandas multispectrales, añadirá 8 bandas en el infrarrojo. Capaz de adquirir imágenes para una extensión de 680000 km² al día. Sus especificaciones se listan en la Tabla 2-3.

2.8.5 Ikonos

Fue el satélite comercial de mayor resolución espacial (1 metro) hasta la llegada del Quickbird. Pertenece a la empresa norteamericana Space Imaging que en 2006 fue adquirida por ORBIMAGE. Fue lanzado el 24 de Septiembre de 1999 y desde entonces da 14.7 vueltas a la Tierra diarias. Su lanzamiento estaba previsto para el año 2000 y fue adelantado por el fracaso del lanzamiento en 1998 de su predecesor, el Ikonos 1, y pasó a tomar su nombre que inicialmente era Ikonos 2.

Adquiere datos en blanco y negro, y multispectrales de la superficie terrestre a una razón de 2000 km² por minuto recorriendo una órbita sincronizada con el Sol a una altura de 681 kilómetros con una inclinación respecto al ecuador de 98.1 grados. Tarda 98 minutos en recorrer una órbita completa. Al paso por el ecuador del satélite son las 10:30 hora local. A continuación se muestra la respuesta espectral del sensor multispectral y las especificaciones se listan en la Tabla 2-4.

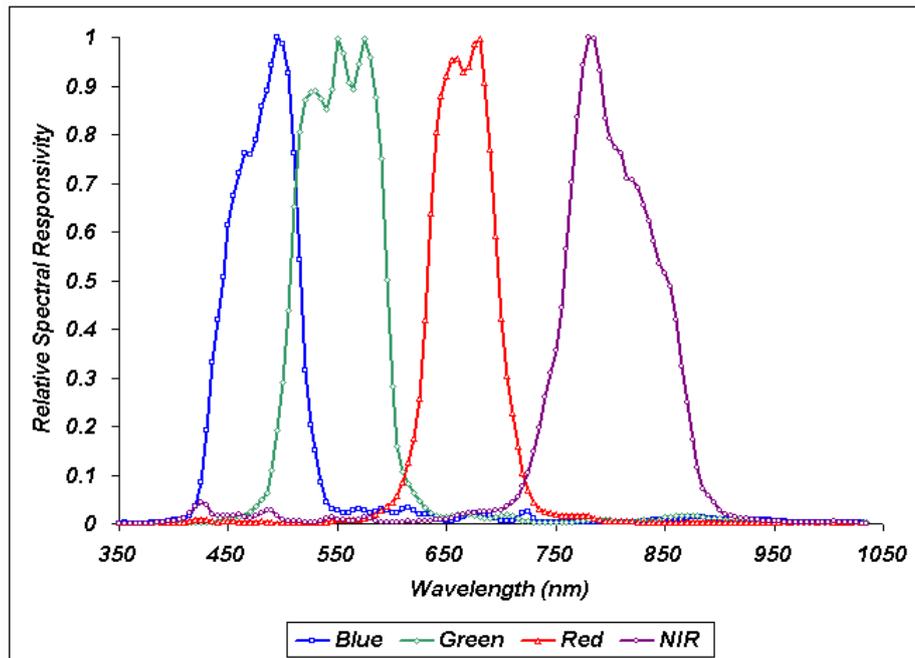


Figura 2-14. Respuesta espectral de las bandas de Ikonos.

2.8.6 Orbview 3

Es otro de los satélites de la constelación de satélites comerciales de la empresa GeoEye. Fue lanzado el día 26 de Junio de 2003. Capturaba unos 21000 km² en una pasada de 10 minutos. Recorría una órbita polar a una altura de 470 kilómetros. Con período de revisita inferior a 3 días, adquiría imágenes pancromáticas con 1 m/píxel de resolución e imágenes multispectrales a 4 m/píxel. Dejó de estar en servicio en abril de 2007 y en 2011 se procedió a su retirada de forma controlada en una zona extensa del Océano Pacífico.

2.8.7 GeoEye 1

GeoEye 1 desarrollado por GeoEye (constituida por la adquisición en 2006 de Space Imaging por parte de ORBIMAGE) fue lanzado el 6 de Septiembre de 2008. Es el sensor de mayor resolución espacial en la actualidad, ofreciendo imágenes pancromáticas de 0.41 metros por píxel. Permite localizar un objeto en la superficie de la Tierra con un error inferior a 3 metros respecto su verdadera localización. Permite la recolección de 700000 km² de imágenes pancromáticas y 350000 km² de imágenes a color cada día. Recorre una órbita polar sincronizada con el Sol que permite adquirir imágenes de cualquier parte de la tierra a la misma hora (10:30 a.m.), con una inclinación de 98° a una altura de 684 kilómetros. El recorrido en la órbita dura 98 minutos a una velocidad de 7.5 km/seg. El resto de especificaciones se lista en la Tabla 2-4.

2.8.8 GeoEye 2

GeoEye ya tiene en desarrollo el sensor *GeoEye 2* cuyo lanzamiento está previsto para el año 2013. Se convertirá en el sensor de mayor resolución espacial

ofreciendo imágenes pancromática de 0.34 m/píxel de resolución y multiespectrales de 1.36 m/píxel. Permite adquirir imágenes cubriendo un área de 600000 km² de imágenes pancromáticas y a color cada día. Recorrerá una órbita polar sincronizada con el Sol atravesando el ecuador a las 10:30 a.m. hora local, a una altura de 681 kilómetros. Tendrá un período de revisita de 3 días. Sus especificaciones se listan en la Tabla 2-4.

2.8.9 Eros B

Eros B es el segundo satélite de la constelación de satélites israelíes planificada por la empresa ImageSat International. Fue lanzado el 25 de Abril de 2006. Supera la resolución espacial de su predecesor *Eros A* que fue lanzado el día 5 de Diciembre de 2000. Ambos recorren una órbita polar heliosincrónica a una altura de 500 kilómetros. Dan la vuelta a la Tierra unas 15 veces cada día. La posibilidad de adquisición hasta un ángulo de 45° permite captar imágenes de cualquier lugar de la Tierra dos o tres veces por semana. La Tabla 2-4 muestra el resto de especificaciones de esta satélite.

2.8.10 Cartosat-2

Es el duodécimo satélite del programa de satélites de teledetección indio (IRS). Fue lanzado el 10 de Enero de 2007. Recorre una órbita polar heliosincrónica a una altura de 635 kilómetros. La órbita tiene un inclinación de 97.92 grados respecto el ecuador. El satélite recorre 14 órbitas cada día con una duración de 97.4 minutos cada una. Dispone de un único sensor pancromático para captar imágenes de muy alta resolución. El resto de especificaciones las podemos encontrar en la Tabla 2-4.

2.8.11 Otros satélites de muy alta resolución

- Pleiades 1A y 1B: El primero lanzado en diciembre de 2011 y el segundo en diciembre de 2012. Toman imágenes pancromáticas y multiespectrales con una resolución de 0.7 metros y 2.8 metros respectivamente en una órbita polar sincronizada con el Sol a una altura de 694 kilómetros. Dependen del Centro de Estudios Espaciales de Francia.
- Kompsat-2: Fue lanzado con éxito el 28 de Julio de 2006. KOMPSAT-2 proporciona imágenes en blanco y negro (Pan) a 1 m de resolución, así como imágenes en color (MS) compuestas de 4 bandas en el espectro visible (azul, verde, rojo) y el infrarrojo cercano con 4 m de resolución. La adquisición simultánea de imágenes Pan y MS permite la creación de imágenes fusionadas con 1 m de resolución. Es un satélite surcoreano.
- Resurs-DK: El primero de esta constelación de satélites ruso fue lanzado el 15 de Junio de 2006. Tomas imágenes de la Tierra con una resolución espacial de 1 metro en pancromática y de 3 metros en multiespectral (3 bandas).

	Quickbird	Worldview 1	Worldview 2	Worldview 3
Ancho de la franja de adquisición	16.5 km en nadir	17.6 Km en el nadir	16.4 Km en el nadir	1.2 Km en el nadir
Capacidad de almacenaje	128 Gbits (57 escenas)	2200 Gbits 331 Gbits en una órbita.	2200 Gbits 524 Gbits en cada órbita	2199 Gbits
Rango dinámico	11 bits/píxel	11 bits/píxel	11 bits/píxel	11 bits/píxel
Período de revisita	3 a 7 días dependiendo de latitud en 60cm de resolución	1.7 días a 1 m. de res. 5.4 días con un ángulo máximo de 20° y 51cm de res.	1.1 días a 1 m. de res. 3.7 días con un ángulo máximo de 20° y 52cm de res.	Menos de un día a 1 m. de resolución.
Resolución espacial	Pancromática	0.61 m en nadir 0.72 m a 25° del nadir	0.50 m en el nadir 0.55 m a 20° del nadir	0.46 m en el nadir 0.52 m a 20° del nadir
	Multiespectral	2.44 m en el nadir 2.88 m a 25° del nadir		1.8 m en el nadir 2.4 m a 20° del nadir
	SWIR			3.70 m en el nadir 4.10 m a 20° del nadir
Resolución espectral	Pancromática	450 - 900 nanómetros	400 - 900 nanómetros	450 - 800 nanómetros
	Multiespectral	Azul: 450 – 520 nanómetros Verde: 520 – 600 nanómetros Rojo: 630 – 690 nanómetros NIR: 760 – 900 nanómetros		Azul: 450 – 510 nanómetros Verde: 510 – 580 nanómetros Rojo: 630 – 690 nanómetros NIR: 770 – 895 nanómetros Costas: 400 – 450 nanómetros Amarillo: 585 – 625 nm Límite del rojo: 705 – 745 nm NIR2: 860 – 1040 nanómetros
	SWIR			8 bandas en el rango 1192 – 2365 nm

Tabla 2-3. Características del sensor Quickbird y los sensores Wordview.

	Ikonos	Geoeye 1	Geoeye 2	Eros B	Cartosat-2
Ancho de la franja de adquisición	11.3 Km en el nadir 13.8 Km a 26° en el nadir	15.2 Km en el nadir	14.5 Km en el nadir	7 Km en el nadir	9.6 Km en el nadir
Capacidad de almacenaje		1 Tbits	3.2 Tbits		64 Gbits
Rango dinámico	11 bits/píxel	11 bits/píxel	11 bits/píxel	10 bits/píxel	10 bits/píxel

Período de revisita		Revisita de aproximadamente 3 días dependiendo de la latitud a 1 metro de resolución	8.3 días con un ángulo máximo de 10° y resolución de 42 cm 2.8 días con un ángulo máximo de 28° y resolución de 50 cm 2.1 días con un ángulo máximo de 35° y resolución de 59 cm	Revisita cada 3 días.	Dos o tres veces en semana en función del ángulo de adquisición hasta 45° de inclinación.	Revisita cada 4 días.
	Resolución espacial	Pancromática	0.82 metros en el nadir 1.0 m en 26° respecto al nadir	0.41 metros en el nadir	0.34 m en el nadir	0.7 m en el nadir
Multiespectral		3.2 metros en el nadir 4.0 m en 26° respecto al nadir	1.65 metros en el nadir	1.36 m en el nadir		
Resolución espectral	Pancromática	526 – 929 nanómetros	450 – 800 nanómetros	450 – 800 nanómetros	500 – 900 nanómetros	500 – 850 nanómetros
	Multiespectral	Azul: 445 – 516 nanómetros Verde: 506 – 595 nanómetros Rojo: 632 – 698 nanómetros NIR: 757 – 853 nanómetros	Azul: 450 – 510 nanómetros Verde: 510 – 580 nanómetros Rojo: 655 – 690 nanómetros NIR: 780 – 920 nanómetros	Azul: 450 – 510 nanómetros Verde: 510 – 580 nanómetros Rojo: 655 – 690 nanómetros NIR: 780 – 920 nanómetros		

Tabla 2-4. Características de los sensores Ikonos, Geoeye 1 y 2, Eros B y Cartosat 2.

2.9 Detección de cambios

La detección de cambios es una de las áreas de mayor importancia de la teledetección espacial. Trata el estudio de los cambios que se han producido en la superficie terrestre a partir de una serie de imágenes captadas en diferentes instantes de tiempo. Estos cambios pueden ser de muy distinta naturaleza, por ejemplo los producidos por desastres naturales, o los cambios producidos por la actividad humana como la construcción de edificios y carreteras.

Generalmente, cualquier proceso de detección de cambios, requiere un procesamiento previo de las imágenes a emplear en el estudio. Este suele constar de una corrección radiométrica, en la que se unifica el color de las imágenes para hacerlas comparables; y una corrección geométrica en la que se alinean las imágenes para que los píxeles con iguales coordenadas representen la misma zona de la superficie terrestre. Estos dos procesos se estudiarán y describirán de forma detallada en los siguientes capítulos.

Tras las correcciones previas, se pasa a la fase en la que los cambios son localizados en las imágenes. Es tan diverso el conjunto de tipos de cambios que pueden aparecer en las imágenes de alta y muy alta resolución, que se hace indispensable seguir un marco de trabajo como el propuesto por (Bruzzone & Bovolo, 2013), que guíe el proceso de diseño de un sistema de detección de cambios adecuado para cada estudio concreto. Básicamente, se basa en tres componentes principales: 1) Identificación de los posibles tipos de cambios radiométricos ocurridos en las imágenes; 2) Definición de un sistema de identificación y localización de los cambios de interés; y 3) Dotar a los cambios de un significado que permita su correcta clasificación.

Para la identificación de los posibles tipos de cambios, es necesario definir una clasificación de todos los posibles cambios que pueden ocurrir entre dos imágenes de satélite de alta resolución. Los autores proponen una primera división entre los cambios producidos por las condiciones de adquisición (Ω_{Acq}) y los cambios producidos sobre el terreno (Ω_{Grd}). Los cambios por las condiciones de adquisición se clasificarían en los producidos por las condiciones atmosféricas (Ω_{Atm}), por ejemplo por presencia de nubes o variaciones en la transmitancia de la atmósfera; y los producidos por diferencias en el sistema de adquisición (Ω_{Sys}), por ejemplo por sensores diferentes, cambios en la iluminación debidos a las estaciones y por diferentes ángulos de adquisición. En cuanto a los cambios producidos en el terreno, se pueden clasificar en aquellos debidos al estado fenológico de la vegetación (Ω_{Veg}); cambios en las condiciones medioambientales (Ω_{Env}), como puede ser presencia de nieve, los desplazamientos de glaciares o por desertización; cambios producidos por desastres naturales (Ω_{Dis}); y los cambios producidos por la actividad antrópica (Ω_{Ant}), como la tala de árboles y construcciones de edificios y carreteras.

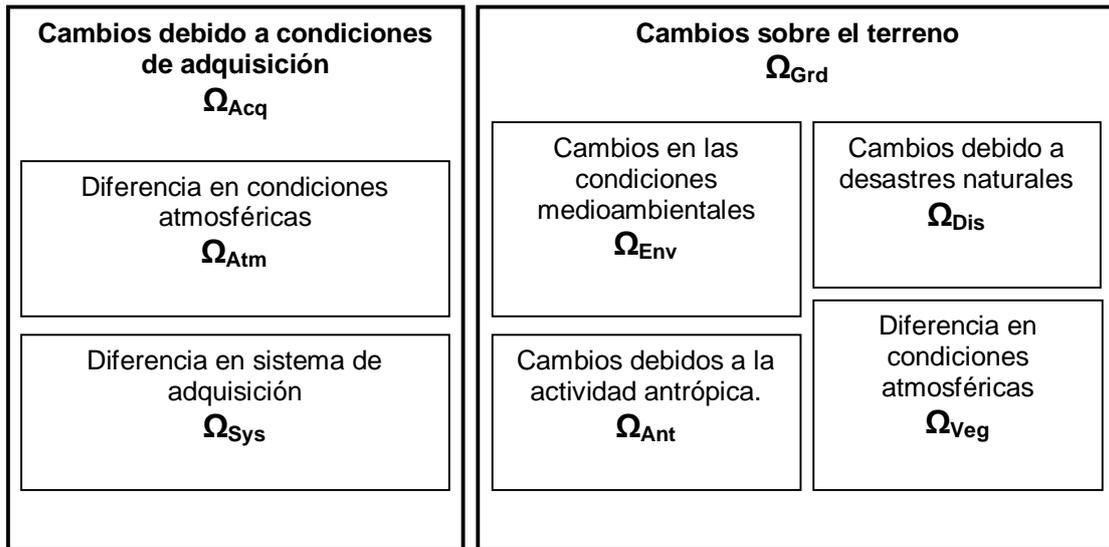


Figura 2-15. Taxonomía de las causas de los cambios radiométricos en imág. de alta resolución.

Una vez definidas las posibles causas de cambios, es fácil identificar en cada estudio, los tipos de cambios que se han producido e identificar entre todos ellos, cuáles son de interés para el usuario final.

Después de determinar qué tipos de cambios pueden darse y cuáles son de interés, se procede a diseñar el sistema de detección que se usará, que puede tener dos estructuras posibles: 1) se detectan todos los cambios radiométricos ocurridos en las imágenes y posteriormente se eliminan aquellos que no son de interés, o 2) se detectan cada uno de los tipos de cambios de interés y se fusionan en un mapa de cambios final. En este proyecto seguiremos la primera idea, proporcionando herramientas para detectar todo tipo de cambio radiométrico y herramientas para eliminar o reducir en la medida de lo posible los cambios que no sean de interés.

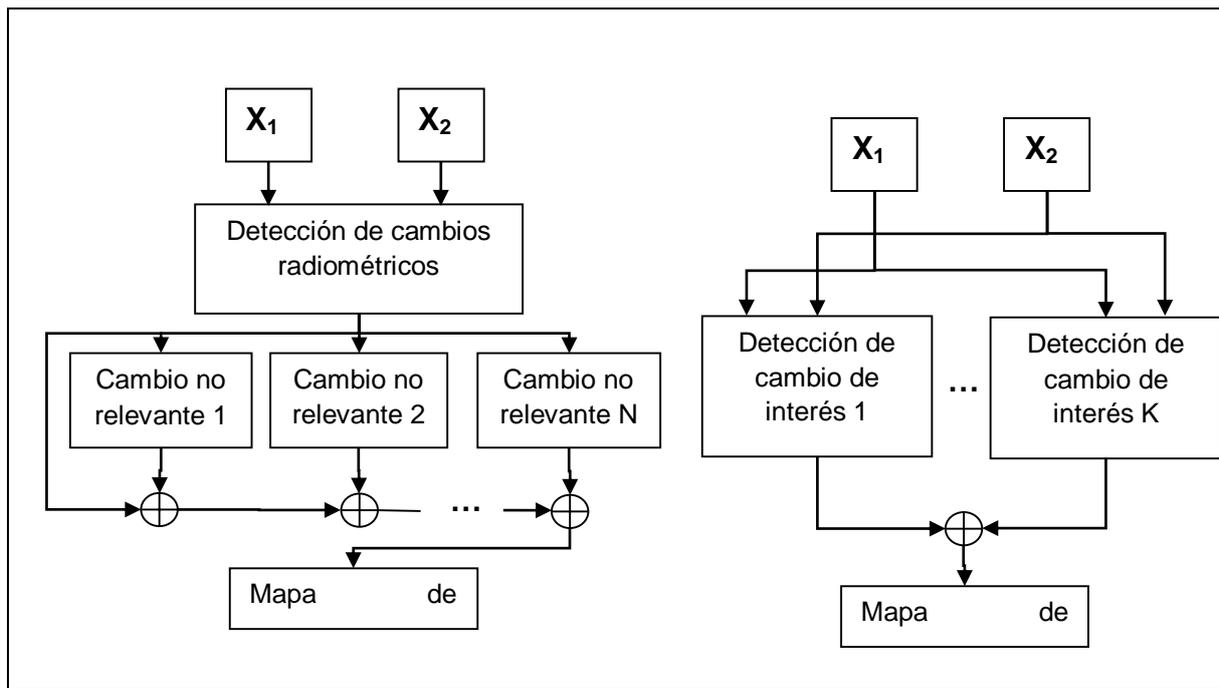


Figura 2-16. Posibles estructuras de sistema de detección de cambios.

Las propiedades radiométricas de un objeto pueden haber cambiado incluso sin haber cambiado la semántica y las condiciones de adquisición, y, por otro lado, varios cambios con significado diferente pueden suponer un mismo cambio radiométrico. Esto hace que la obtención del mapa de cambios radiométricos a veces no sea suficiente y sea necesaria una tercera fase en la que se dé a cada cambio un significado que permita compararlos.

Se trataría de dotar a cada cambio de un vector de características que los describa desde tres niveles diferentes: nivel de píxel, de primitiva y de objeto. El nivel de píxel considera el valor radiométrico de los píxeles del cambio y otros índices como los de vegetación o sombras. El nivel de primitiva se asocia con características geométricas o estadísticas que pueden extraerse de las imágenes como líneas o regiones obtenidas por segmentación o algún tipo de filtro. Y por último, el nivel de objeto trata de dar un significado al cambio combinando varias primitivas. La detección basada en parcelas descrita en el apartado de futuras mejoras del Capítulo 6, es un ejemplo de implementación de esta tercera fase del marco de trabajo.

En los próximos capítulos trataremos distintas técnicas para preparar las imágenes, obtener mapas de cambios, extraer algunos tipos de cambios comunes como las sombras, eliminar cambios no deseados y clasificarlos de forma no supervisada.

Capítulo 3

Correcciones radiométricas

La radiación captada por el sensor se ve afectada por diversas causas ajenas a los cambios producidos en la superficie terrestre. La calibración del sensor, sus parámetros de operación, la elevación del sol, las condiciones atmosféricas y la topografía del terreno son las más importantes. Todas ellas, excepto la topografía, cambian en el tiempo, de ahí que un mismo objeto en la superficie terrestre no sea captado con el mismo brillo en distintos instantes de tiempo y que, por tanto, sea más difícil diferenciar los cambios debidos a distintas condiciones en la adquisición de las imágenes, de los cambios que se han producido en la superficie estudiada. La corrección radiométrica es necesaria para disminuir estas distorsiones en la radiación observada y poder detectar los cambios adecuadamente.

Algunos autores denominan a la corrección radiométrica como *normalización radiométrica* ya que este proceso unifica el color de las distintas imágenes en un mismo sistema métrico y las hace comparables. Hay dos tipos de corrección radiométrica: absoluta y relativa.

Para la corrección radiométrica *absoluta* es necesario conocer los parámetros del sensor, los parámetros de comportamiento de la radiación electromagnética a través de la atmósfera y otros datos que son difíciles de obtener tras la adquisición de la imagen desde el satélite. Este tipo de correcciones, generalmente, son realizadas por el operador del satélite en las estaciones de recepción de datos antes de ser enviados al usuario final.

A la hora de realizar una detección de cambios no es indispensable una corrección absoluta y es suficiente con una corrección radiométrica *relativa* para obtener buenos resultados. En una corrección relativa no es necesario el conocimiento de las condiciones atmosféricas en el momento en el que el satélite adquirió la imagen. En este proceso, se toma una imagen como referencia y se ajustan los valores radiométricos de la otra para aproximarlos a los de la imagen de referencia, disminuyendo las diferencias donde no debería haberlas y haciendo que las imágenes

parezcan haber sido captadas con las mismas condiciones atmosféricas y de iluminación. Este tipo de corrección no elimina errores debido a las condiciones de la atmósfera en el momento de la adquisición.

A continuación, vamos a estudiar distintos métodos de corrección radiométrica relativa, indicando sus ventajas e inconvenientes, e intentar llegar a la conclusión de cuál es el método más adecuado para la detección de cambios.

3.1 Métodos lineales de corrección radiométrica relativa

En este tipo de métodos, el problema de la diferencia radiométrica entre dos imágenes se resuelve mediante la corrección lineal de los histogramas, modificando el contraste y el brillo. Estas correcciones responden a transformaciones lineales del tipo:

$$U_k = a_k X_k + b_k$$

donde U_k es la banda k de la imagen normalizada, X_k es la banda k de la imagen a ajustar antes de la normalización, a y b son los coeficientes de normalización y k es el número de banda de la imagen. Llamaremos Y_k a la banda k de la imagen de referencia. La normalización o corrección puede dividirse en dos pasos:

1. Seleccionar las áreas o conjuntos de píxeles a emplear en la normalización.
2. Determinar los coeficientes de normalización.

Varios autores han propuesto distintos métodos para seleccionar conjuntos ideales de normalización de los cuales estimar los coeficientes de transformación. En general, para todos los métodos de normalización, un conjunto de píxeles ideal debe cumplir los siguientes criterios:

- Las áreas usadas en el cálculo de los coeficientes deberían tener aproximadamente la misma elevación de tal forma que la atmósfera tenga un espesor similar en cada una.
- Los conjuntos o áreas empleadas no deberían contener grandes cantidades de vegetación porque su reflectancia espectral cambia en el tiempo.
- Las áreas de normalización deben ser relativamente planas para que cambios en el ángulo del Sol produzca un cambio proporcional en la iluminación de todas ellas.
- Las áreas empleadas no deben contener patrones o elementos que cambien en el tiempo y que puedan ser observados al visualizar las imágenes.
- Los conjuntos de píxeles seleccionados deben contener un amplio rango de niveles de gris para que el modelo de regresión empleado sea fiable.

En las siguientes secciones se presentan distintos métodos propuestos y evaluados en (Mateu & Ruiz, 1999) para la normalización de imágenes.

3.1.1 Corrección de bruma (Haze Correction - HZ)

Este tipo de corrección es muy simple y asume que los píxeles con reflectancia cero en ambas imágenes (la de referencia y la de ajuste) deberían tener ambas el mismo DN (*número digital*). Se basa en la corrección de un desplazamiento en el valor de los valores digitales, con una ganancia (término a_k en la ecuación) de 1. Los coeficientes empleados por tanto en este tipo de corrección son:

$$a_k = 1, \quad b_k = y_{k_{min}} - x_{k_{min}}$$

donde $x_{k_{min}}$ e $y_{k_{min}}$ son los valores de bruma en la banda k en las imágenes a ajustar y de referencia respectivamente. Los valores de bruma se corresponden con los valores digitales (DN) que aíslan el 0.1% de los píxeles más oscuros de la imagen. Se usan estos valores porque tomar simplemente el menor DN de las imágenes daría lugar a correcciones de pocos niveles, dejando la imagen casi inalterada.

3.1.2 Corrección mínimo – máximo (MM)

Este método normaliza la imagen de tal forma que tenga el mismo DN mínimo y máximo que la imagen de referencia en todas las bandas. Los coeficientes de normalización usados en este método son:

$$a_k = \frac{y_{k_{max}} - y_{k_{min}}}{x_{k_{max}} - x_{k_{min}}}, \quad b_k = y_{k_{min}} - a_k x_{k_{min}}$$

donde $x_{k_{min}}, x_{k_{max}}, y_{k_{min}}, y_{k_{max}}$ son los valores digitales mínimos y máximos de la banda k para las dos imágenes a corregir (la imagen a normalizar y la de referencia respectivamente). Los valores mínimos y máximos para las dos imágenes son seleccionados tal que aíslan el 0.1% de los píxeles más oscuros y el 0.1 % de los píxeles más brillantes.

3.1.3 Corrección por la media y desviación típica (MS)

Este método corrige la imagen de tal forma que tenga la misma media y desviación típica que la imagen de referencia en todas las bandas. Suponiendo que \bar{x}_k e \bar{y}_k son las medias y s_{x_k} e s_{y_k} son las desviaciones típicas de las imágenes a ajustar y de referencia respectivamente, los coeficientes de normalización son obtenidos de:

$$a_k = \frac{s_{y_k}}{s_{x_k}}, \quad b_k = \bar{y}_k - a_k \bar{x}_k$$

3.1.4 Método de regresión simple (SR)

En este método se emplea el ajuste por mínimos cuadrados intentando hacer mínima las diferencias entre ambas imágenes. Se trata de encontrar los parámetros que hacen mínima la expresión siguiente:

$$Q = \sum_{escena} (y_k - a_k x_k - b_k)^2$$

donde el sumatorio se realiza en la escena completa. Los coeficientes que resuelven la ecuación anterior son los siguientes:

$$a_k = \frac{s_{x_k y_k}}{s_{x_k x_k}}, \quad b_k = \bar{y}_k - a_k \bar{x}_k$$

donde

$$s_{x_k x_k} = \frac{1}{|escena|} \sum_{escena} (x_k - \bar{x}_k)^2$$

y

$$s_{x_k y_k} = \frac{1}{|escena|} \sum_{escena} (x_k - \bar{x}_k)(y_k - \bar{y}_k)$$

siendo *escena* el conjunto de píxeles de la imagen y $|escena|$ el número de píxeles en la escena.

3.1.5 Corrección por conjuntos oscuros y brillantes de píxeles (DB)

Este método en vez de usar los valores digitales mínimos y máximos propone usar los promedios de los conjuntos de píxeles más oscuros y brillantes de las imágenes. Fue usado por primera vez para la corrección de imágenes Landsat (Hall, Strebel, Nickeson, & Goetz, 1991). Los conjuntos de píxeles oscuros y brillantes fueron obtenidos a partir de la transformación KT también conocida como transformación *Tasseled Cap* (Kauth & Thomas, 1976). Esta transformación permite obtener, entre otras, imágenes del verdor y brillo de las imágenes originales, de donde se obtienen los píxeles oscuros y claros a partir de expresiones del tipo:

$$\begin{aligned} \text{Píxeles oscuros} &= \{\text{verdor} \leq t_1 \text{ y brillo} \leq t_2\} \\ \text{Píxeles claros} &= \{\text{verdor} \leq t_1 \text{ y brillo} \geq t_2\} \end{aligned}$$

Los coeficientes de la recta de normalización vienen determinados por los parámetros $\bar{x}_k^{(d)}$, $\bar{x}_k^{(b)}$, $\bar{y}_k^{(d)}$, $\bar{y}_k^{(b)}$ que se corresponden con las medias de los píxeles oscuros (d) y brillantes (b) de las dos imágenes y son obtenidos a través de las expresiones:

$$a_k = \frac{\bar{y}_k^{(b)} - \bar{y}_k^{(d)}}{\bar{x}_k^{(b)} - \bar{x}_k^{(d)}}$$

$$b_k = \bar{y}_k^{(d)} - a_k \bar{x}_k^{(d)}$$

Teóricamente, los píxeles oscuros corresponden a aguas de gran profundidad y los píxeles brillantes representan objetos de alta reflectividad, como edificios de

hormigón y suelo rocoso. Los conjuntos de píxeles seleccionados pueden ser diferentes para las dos imágenes a normalizar si la cobertura terrestre ha cambiado.

Las expresiones de verdor y brillo a través de la transformación Tasseled Cap no son válidas para imágenes de satélites de alta resolución, como las de los satélites Quickbird e Ikonos, debido a la diferencia en la respuesta espectral entre estos y el satélite Landsat para el que se diseñó esta transformación. En (Horne, 2003) se desarrollan las expresiones adecuadas para el satélite Ikonos:

$$\begin{aligned} \text{brillo} &= 0.326x_{\text{blue}} + 0.509x_{\text{green}} + 0.560x_{\text{red}} + 0.567x_{\text{nir}} \\ \text{verdor} &= -0.311x_{\text{blue}} - 0.356x_{\text{green}} - 0.325x_{\text{red}} + 0.819x_{\text{nir}} \end{aligned} \quad (3.1)$$

En (Hong & Zhang, 2008) se propone usar (3.1) también para imágenes Quickbird ya que tiene una misma resolución radiométrica y una respuesta espectral similar a la del sensor del satélite Ikonos. Como valores umbrales t_1 y t_2 proponen los siguientes: para Ikonos, un valor de $t_1 = 1$ y $t_2 = 300$ para el conjunto de píxeles oscuros y un valor de $t_1 = 1$ y $t_2 = 300$ para el conjunto de píxeles claros; para Quickbird proponen un valor de $t_1 = 1$ y $t_2 = 190$ para los píxeles oscuros y $t_1 = 1$ y $t_2 = 220$ para el conjunto claro.

3.1.6 Método de los pseudoinvariantes (PI)

Este método supone la existencia de objetos pseudoinvariantes que no han experimentado cambios de una fecha a otra en términos de reflectividad. Algunas de estos objetos pseudoinvariantes podrían ser carreteras, áreas urbanas y polígonos industriales. Las características pseudoinvariantes son extraídas del análisis de la proporción del infrarrojo respecto al rojo de las imágenes de análisis y referencia, con el objetivo de identificar píxeles de poca vegetación verde; y de aplicar un valor umbral al infrarrojo para eliminar píxeles del agua.

Suponiendo que las medias y desviaciones típicas de los conjuntos pseudoinvariantes son para las dos fechas: $\bar{x}_k^{(pi)}$, $\bar{y}_k^{(pi)}$, $s_{x_k}^{(pi)}$ y $s_{y_k}^{(pi)}$, los coeficientes de normalización serían:

$$\begin{aligned} a_k &= \frac{s_{y_k}^{(pi)}}{s_{x_k}^{(pi)}} \\ b_k &= \bar{y}_k^{(pi)} - a_k \bar{x}_k^{(pi)} \end{aligned}$$

El conjunto pseudoinvariante estaría definido por la expresión:

$$PI = \left\{ \frac{NIR}{Rojo} \leq t_1 \text{ y } NIR > t_2 \right\}$$

donde NIR es la banda de infrarrojo cercano y Rojo es la banda del color rojo. (Hong & Zhang, 2008) propone valores de $t_1 = 2$ y $t_2 = 80$ para imágenes Ikonos, y valores de $t_1 = 2$ y $t_2 = 140$ para imágenes Quickbird.

3.1.7 Prueba de los métodos de ajuste lineal propuestos

En la detección de cambios se usan imágenes con un número variable de bandas, desde imágenes pancromáticas a imágenes multiespectrales de 4 bandas (como las de Quickbird o Ikonos) o incluso más. Para las pruebas de los diferentes algoritmos usaremos imágenes en color natural (bandas roja, verde y azul) del satélite Quickbird. Aunque se han implementado e integrado todos los métodos de corrección comentados en los apartados anteriores, no serán estudiados en este apartado los métodos DB y PIF por no disponer de la necesaria banda de infrarrojo.

La imagen que debemos ajustar en la prueba está formada a partir de la imagen de referencia modificando su histograma en cada una de las tres bandas de color. Usamos esta imagen sintética para evitar que un mal ajuste espacial (errores en la corrección geométrica) influya en los resultados. De esta forma evitamos también que existan cambios entre las dos imágenes que afecten a la corrección. En imágenes reales en las que se han producido cambios, la corrección radiométrica será mejor cuanto menor sea el porcentaje de píxeles que cambian de una fecha a otra. Los histogramas de ambas imágenes son los mostrados en la Figura 3-1. Los coeficientes de ajuste obtenidos de aplicar los métodos HZ, MM, MS, y SR se listan en la Tabla 3-1 para las tres bandas.

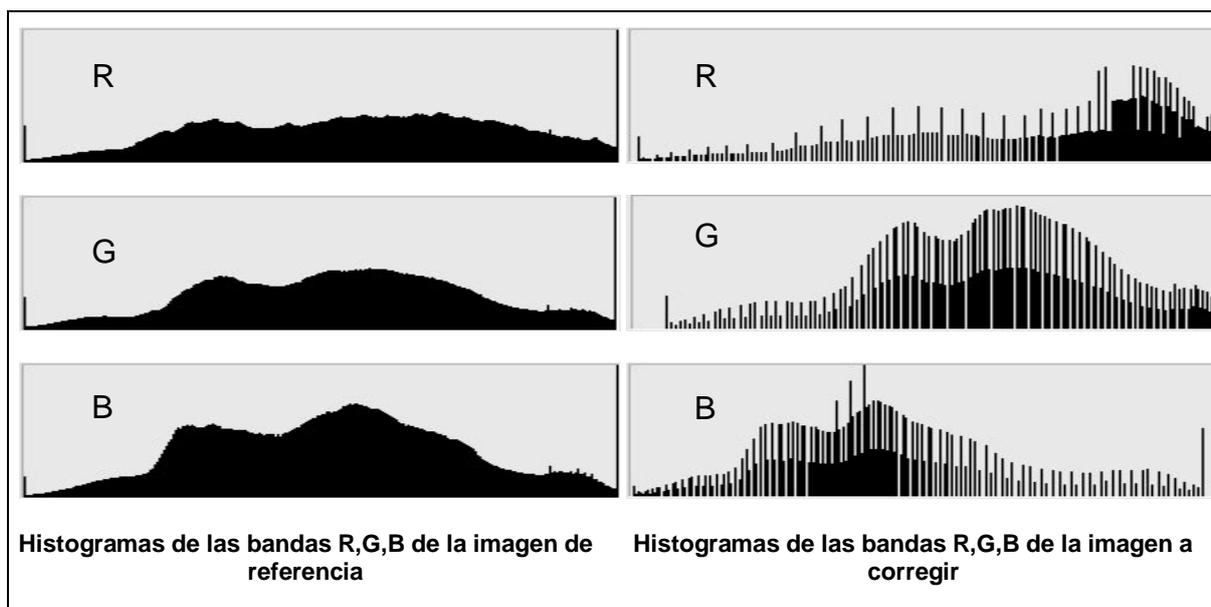


Figura 3-1. Histogramas de imagen de referencia e imagen a corregir.

	Banda R		Banda G		Banda B	
	a	b	a	b	a	b
HZ	1	-3	1	-14	1	-1
MM	1.0119	-3.0358	1.0672	-14.9411	1.0325	-1.0325
MS	1.0460	-48.4545	1.1776	-46.4654	1.1288	11.6860
SR	0.9948	-39.0139	1.1724	-45.6479	1.1024	14.5311

Tabla 3-1. Coeficientes métodos lineales de ajuste radiométrico.

Con estos parámetros podemos representar la función de la relación entre las dos imágenes (en azul) y las rectas de ajuste obtenidas en cada método (Figura 3-2). En el eje X se representa los niveles de gris de la imagen a ajustar y en el eje Y los niveles de gris de los mismos píxeles en la imagen de referencia. Si ambas imágenes fueran iguales la función sería lineal pasando por los puntos (0,0) y (255,255). Aún así la función es una curva y no una nube de puntos como se podría esperar ya que la imagen a ajustar se basa en una transformación monótona creciente de los niveles de gris de la imagen de referencia.

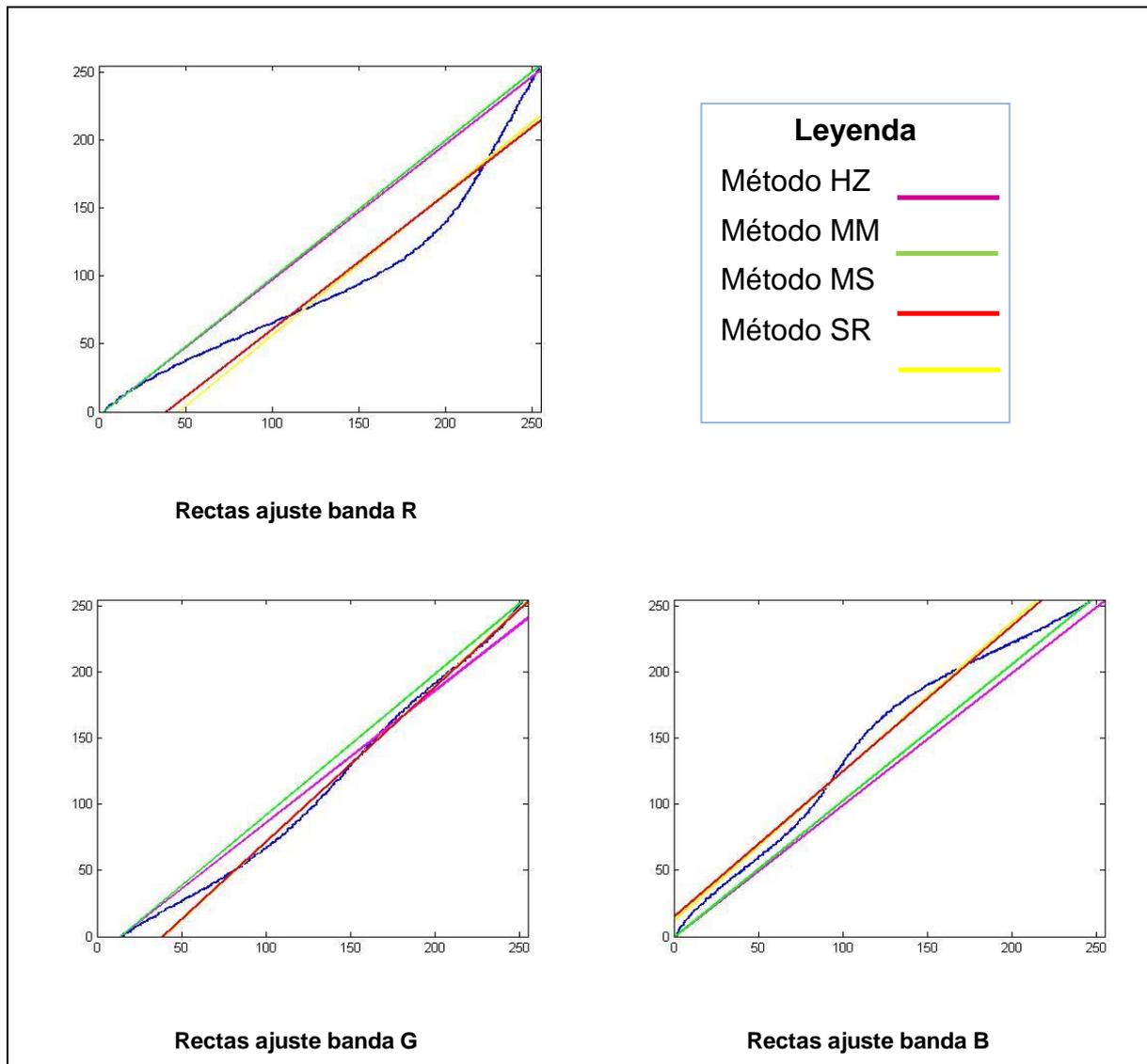


Figura 3-2. Ajustes radiométricos lineales.



Imagen de referencia



Imagen a corregir



Método Haze Correction



Método Minimum-Maximum



Método Mean - Standard Deviation



Método Simple Regression

Figura 3-3. Resultados ajustes radiométricos lineales.

Podemos ver como el resultado obtenido por los métodos MS y SR es muy parecido, llegando casi a solaparse las rectas de ajuste en alguna banda. También se observa claramente cómo estas rectas representan mejor la relación entre las imágenes que las rectas de los métodos HZ y MM.

El resultado de aplicar los distintos métodos de corrección a las imágenes se pueden observar en la Figura 3-3. Como era de esperar, los métodos HZ y MM presentan malos resultados a diferencia de los métodos MS y SR que, para las imágenes usadas, obtienen un resultado similar y en apariencia bueno. Recordemos que ambas imágenes son coincidentes (al derivar una de la otra) y por tanto carecen de errores de desplazamiento y de cambios que puedan distorsionar los coeficientes de ajuste resultantes. El desempeño de estos métodos será más bueno cuanto más lineal sea la relación entre los píxeles de las dos imágenes.

En la siguiente sección veremos dos métodos no lineales de corrección radiométrica y los aplicaremos a las imágenes ya presentadas para, a continuación, analizar el error cometido en cada uno de los métodos usados y llegar a conclusiones sobre cuál es el método más adecuado.

3.2 Especificación del histograma (EH)

Recordemos que el histograma de una imagen digital con niveles de gris en el rango $[0, L-1]$ es una función discreta $h(r_k) = n_k$, donde r_k es el k -ésimo nivel de gris y n_k es el número de píxeles en la imagen con un nivel de gris r_k .

Para entender el funcionamiento de la especificación de histograma es necesario conocer con anterioridad el procedimiento de igualación (o ecualización) de histograma. Sea la siguiente función discreta:

$$p_r(r_k) = \frac{n_k}{n}$$

donde n es el número total de píxeles de la imagen, y n_k y r_k son los comentados anteriormente, podríamos decir que $p_r(r_k)$ es una estimación de la probabilidad de ocurrencia del nivel de gris r_k en la imagen. Sea r los niveles de gris de la imagen original normalizados en el rango $[0, 1]$, la igualación del histograma consistirá en encontrar una función T tal como $s = T(r)$ que transforme cada nivel de gris r de la imagen original en el nivel de gris s de tal forma que la densidad de probabilidades de s sea uniforme para toda la amplitud de niveles de gris posible. La transformación T debe cumplir además que no sea multivaluada (para que exista la función inversa), que sea monótona creciente (para que no se inviertan píxeles) y que $0 \leq T(r) \leq 1$ para $0 \leq r \leq 1$ para que la salida de la función esté dentro del rango de niveles de gris permitido.

Suponiendo que las variables aleatorias r y s son continuas vamos a demostrar que la función T que buscamos se corresponde con la función de distribución de la variable r :

$$F(r_k) = \int_0^{r_k} p_f(w)dw = P(r \leq r_k)$$

donde w es la variable auxiliar de integración. Tal como hemos supuesto, $s = F(r)$ y por tanto:

$$F(s_k) = P(s \leq s_k) = P(F(r) \leq s_k)$$

Si F es continua y monótona creciente, existe F^{-1} y por tanto:

$$F(r) \leq s_k \Rightarrow r \leq F^{-1}(s_k)$$

Por lo que llegaríamos a:

$$F(s_k) = P(F(r) \leq s_k) = P(r \leq F^{-1}(s_k)) = F(F^{-1}(s_k)) = s_k$$

que sólo es posible si la variable aleatoria s tiene una función de densidad de probabilidad uniforme en el rango $[0, 1]$ y quedando demostrado que la función de distribución es la transformación que estábamos buscando.

Trasladando esta función al ámbito discreto la función T quedaría como sigue:

$$s_k = T(r_k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{n}, \quad k = 0, 1, \dots, L - 1$$

donde, recordemos, los niveles de gris r_k y s_k están normalizados. Una vez aplicada la función de transformación se debe redondear los valores s_k a los niveles de gris más próximos s_k^* convirtiendo los niveles originales r_k en los niveles s_k^* en la imagen con histograma igualado.

La especificación del histograma consiste en transformar el histograma de una imagen para que se aproxime a la forma de un histograma determinado. Por este motivo, éste es uno de los métodos empleados en la corrección radiométrica relativa ya que nos permite ajustar el histograma de una imagen para que adopte la forma del histograma de una imagen de referencia, aproximando por tanto el color en ambas imágenes y disminuyendo las diferencias.

Al igual que en la igualación de histograma, al ser los niveles de gris una variable discreta, el resultado obtenido es una aproximación. Sean r y z los niveles de gris de las imágenes a normalizar y la de salida con histograma especificado respectivamente (que debe corresponderse con el histograma de la imagen de referencia), suponiendo que ambas son variables aleatorias continuas y que $p_r(r_k)$ y

$p_z(z_k)$ son sus funciones de densidad de probabilidad, tenemos las siguientes transformaciones de igualación del histograma:

$$s = T(r) = \int_0^r p_r(w)dw$$

$$v = G(z) = \int_0^z p_z(w)dw$$

Ya que las funciones de transformación anteriores dan lugar a imágenes con función de densidad de probabilidades homogéneas y por tanto iguales, se puede obtener la imagen de histograma especificado aplicando la transformación G^{-1} , a la imagen de histograma igualado que se obtiene de aplicar la transformación T a la imagen a normalizar, en la forma que indica la siguiente expresión:

$$z = G^{-1}(s) = G^{-1}(T(r))$$

Trasladando el procedimiento anterior al caso discreto la especificación de histograma consistiría en lo siguiente:

1. Igualar el histograma de la imagen a normalizar:

$$s_k = T(r_k) = \sum_{j=0}^k p_r(r_k)$$

2. Obtener la función de transformación $G(z_k)$ igualando el histograma especificado $p_z(z_k)$ (el de la imagen de referencia).

$$v_k = G(z_k) = \sum_{j=0}^k p_z(z_k)$$

3. Aplicar la función de transformación inversa G^{-1} a los niveles de gris s_k de la imagen obtenida en el paso 1.

$$z_k = G^{-1}(s_k) = G^{-1}(T(r_k))$$

La especificación de histograma es muy usada en la teledetección como método de corrección radiométrica por su sencillez y rapidez, sobre todo para análisis visual de imágenes. El procedimiento completo se ha extraído de (González, 2000) donde se describe con ejemplos numéricos.

A continuación, vamos a aplicar la especificación de histograma a las imágenes presentadas en el apartado de métodos de corrección lineales. El resultado junto a la imagen de referencia se muestran en la siguiente Figura:



Figura 3-4. Resultado de la especificación de histograma.

En la Figura 3-5 podemos observar los histogramas de la imagen de referencia y los de la imagen tras la normalización.

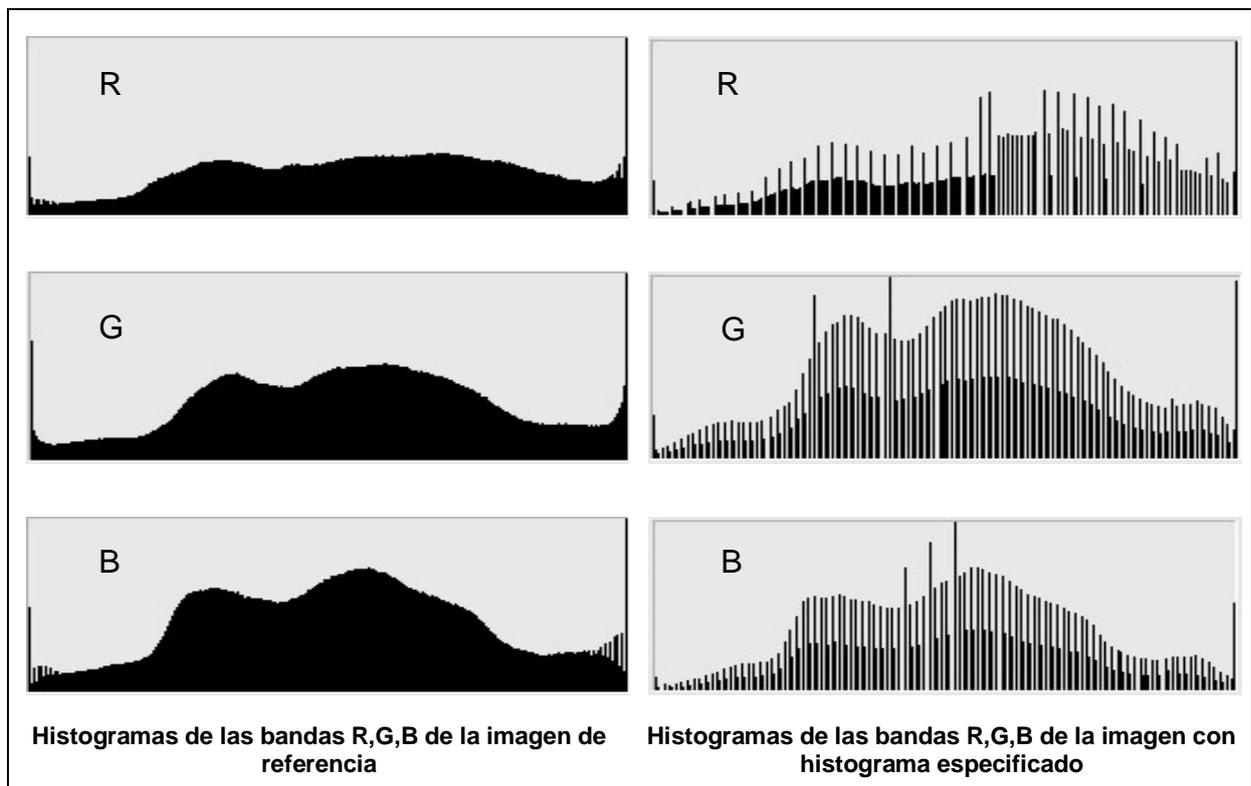


Figura 3-5. Histogramas de imagen de referencia y de imagen con histograma especificado.

Podemos comprobar cómo tras la especificación los histogramas tienen una forma muy similar. En el apartado 3.5 compararemos los resultados de este método con los obtenidos en los métodos lineales a través del error cuadrático medio. También

estudiaremos por qué los resultados de la especificación de histograma son aparentemente tan buenos.

3.3 Normalización adaptativa bilineal (AB)

Los métodos lineales de corrección, así como la especificación del histograma, se ven afectados por el tamaño de las imágenes a corregir y por los cambios que se hayan producido en ellas.

Si hay cambios en las imágenes, cuanto más grande sean éstas, menos representativas serán las medidas estadísticas empleadas en la corrección y por tanto peor su resultado. De manera similar, cuanto mayor sea el porcentaje de píxeles que han cambiado, mayor distorsión habrá en las estadísticas y más diferentes serán los histogramas entre sí, aunque este problema se puede atenuar aplicando máscaras de cambios conocidos como sombras, saturaciones o cualquier otra zona de cambio que pueda ser determinada antes de realizar el proceso de detección de cambios, evitando así, que se tengan en cuenta los píxeles enmascarados en el cálculo de las estadísticas.

Para disminuir en la medida de lo posible los problemas de la corrección radiométrica comentados anteriormente, se presenta aquí el método de normalización adaptativa bilineal (Prat, 1991). Consiste en dividir la imagen en ventanas conectadas por puntos de malla. Sobre cada uno de estos puntos de malla se calculan los parámetros de la función lineal de normalización por la media y desviación típica (el que hemos llamado método MS) para un determinado vecindario de los mismos. Los parámetros a y b (véase el apartado de métodos lineales) del resto de puntos en la ventana son obtenidos por interpolación bilineal de los parámetros de los cuatro puntos de malla que lo rodean. De esta forma, cada píxel se transforma mediante su propia función lineal, a diferencia de los métodos lineales donde una misma función se aplicaba a todos los píxeles de la imagen. La siguiente figura explica los conceptos tratados:

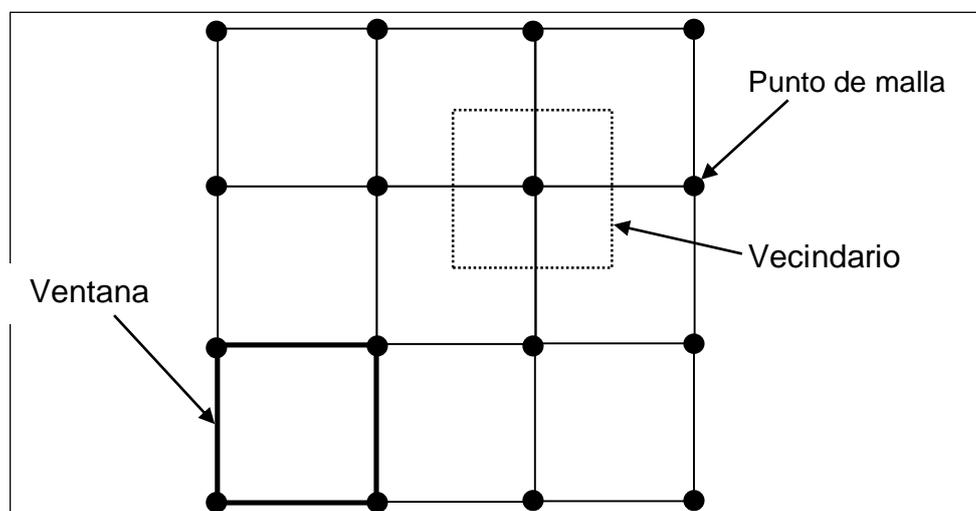


Figura 3-6. División de la imagen para procesamiento adaptativo.

El resultado de la corrección para un tamaño de ventana de 127 x 127 se muestra en la Figura 3-7.

Desde el punto de vista visual, el método de corrección adaptativo bilineal es adecuado para la detección de cambios. Estudiaremos el error cuadrático medio para esta corrección y compararemos los resultados para distintos tamaños de ventana en el apartado 3.5 de este capítulo.

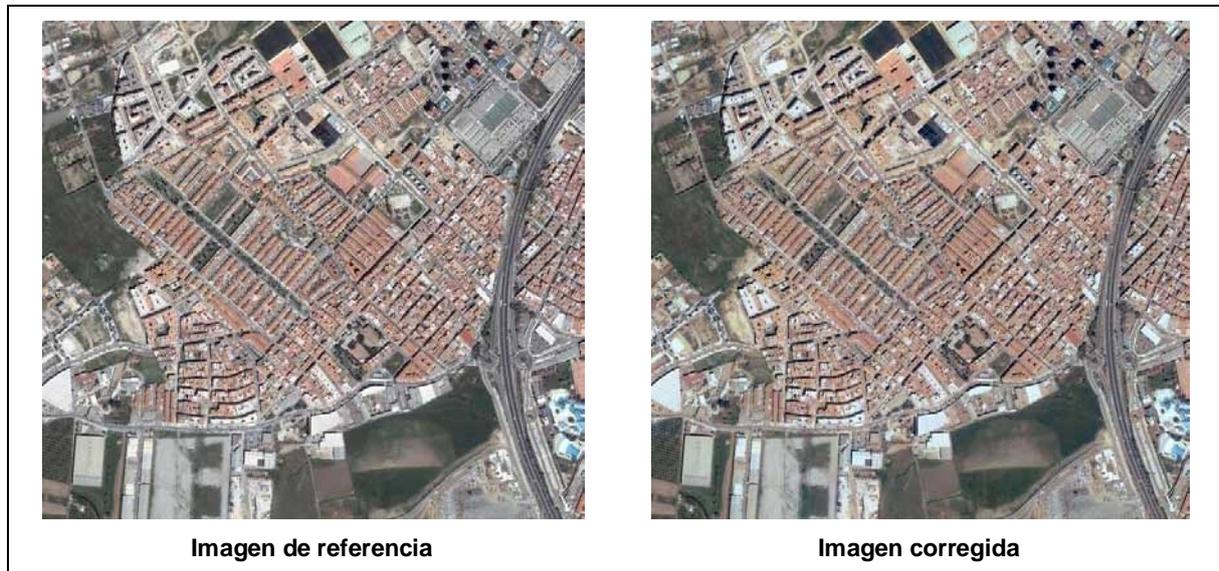


Figura 3-7. Resultado de una corrección radiométrica adaptativa bilineal.

3.4 Experimento con imágenes con cambios

Anteriormente en este capítulo, se han estudiado los distintos métodos de corrección radiométrica con una par de imágenes en la que no se han producido cambios en los diferentes elementos que aparecen en ella, tan sólo en el color de los píxeles. En este apartado, vamos a realizar la corrección radiométrica de dos imágenes adquiridas en distintas fechas con el propósito de comprobar los efectos de un posible pequeño error de registro, así como la influencia de los cambios en los resultados. Las imágenes que se van a normalizar se muestran en la Figura 3-8.

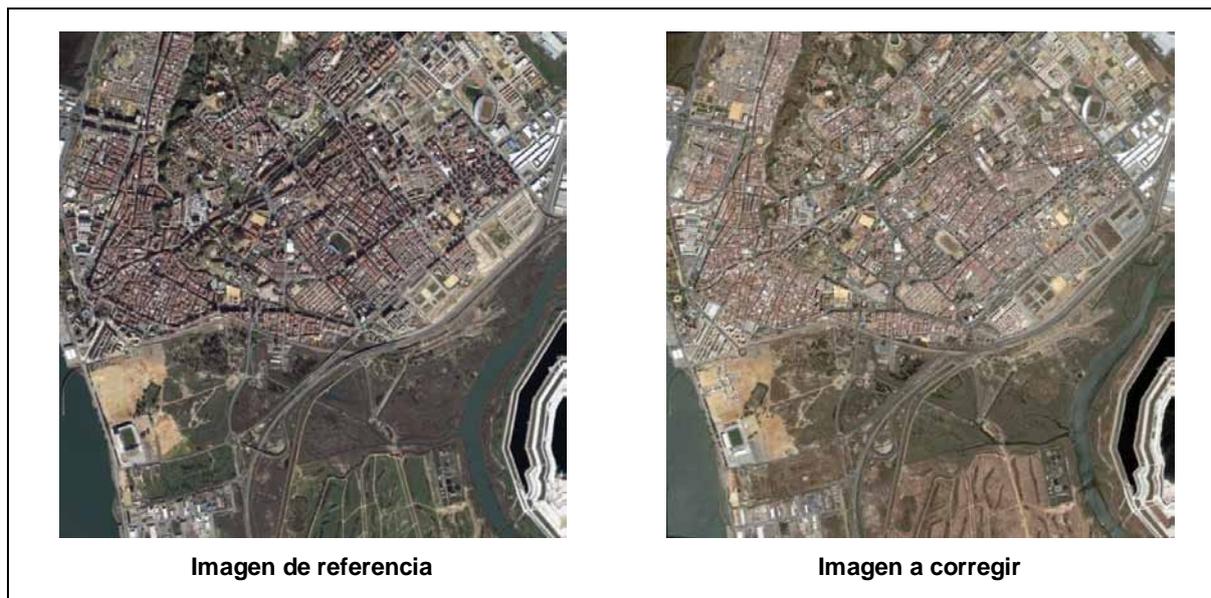


Figura 3-8. Imágenes de Huelva para corregir radiométricamente.

Son imágenes del satélite Quickbird, de una zona de la ciudad de Huelva, adquiridas en Diciembre de 2005 y Mayo de 2007, con una resolución de 1.2 metros/píxel. La imagen de referencia se ve bastante afectada por las sombras en lugares de edificios altos debidos a la poca elevación del Sol en el momento de la captura. Debido a la diferencia temporal entre las dos imágenes se han producido algunos cambios en la zona, tanto urbanos como naturales.

3.5 Resultados y conclusiones

Recordemos que para probar los distintos métodos planteados se han realizado dos experimentos. En el primero, obtuvimos una imagen modificando el histograma de otra para evitar que los resultados se vean influenciados por errores de registro o cambios en las imágenes. En el segundo, se han usado dos imágenes de satélite *ortorrectificadas* eliminando todo lo posible errores de localización de los distintos elementos de la escena, pero se han producido cambios (nuevas construcciones, cambios estacionales, sombras, etc.) debidos a las diferentes condiciones de iluminación y ser adquiridas con más de un año de diferencia.

Para comparar los distintos métodos se calcula el error cuadrático medio según la ecuación siguiente:

$$E = \sqrt{\frac{1}{n * m} \sum_{i=1}^n \sum_{j=1}^m (f_a(x_i, y_j) - f_r(x_i, y_j))^2}$$

donde $f_a(x_i, y_j)$ y $f_r(x_i, y_j)$ son niveles de gris del píxel (i, j) en la imagen a ajustar y de referencia respectivamente, y los valores n y m , el número de columnas y filas de las imágenes. Los resultados se muestran en la Tabla 3-2 y la Tabla 3-3. En el

segundo caso, los valores se han obtenido eliminando de la normalización los píxeles de sombra que no son comunes a las dos imágenes. Se muestra el error E para la imagen completa en la columna de la izquierda de cada banda y para la imagen en la que se han enmascarado los píxeles de sombra en la columna de la derecha. En ambas tablas se muestra en la cabecera de columna el error E antes de la normalización.

Como era de esperar, en los métodos lineales, los métodos más simples HZ y MM obtienen malos resultados para ambos experimentos demostrando ser ineficientes para un buen ajuste. Los métodos MS y SR mejoran estos resultados aunque a veces no son suficientes por la, generalmente, relación no lineal entre las dos imágenes.

Podemos apreciar también el excelente resultado del método de especificación de histograma en el primer experimento, pero de aquí no podemos concluir que sea el mejor de los métodos propuestos. Este error tan bajo se debe a que estamos corrigiendo dos imágenes idénticas, en las que tan sólo hemos modificado el histograma de una de ellas sin introducir saturaciones de ningún tipo y por tanto sin perder información. Al aplicar este método, se invierten las modificaciones utilizadas y da como resultado una imagen muy parecida a la original. Podemos ver como en el segundo caso, el resultado no es tan bueno y presenta un mayor error que el método de regresión simple SR. Esto demuestra la gran influencia de los cambios en la forma del histograma y que, dependiendo del porcentaje de píxeles de cambio o la presencia de sombras, la especificación de histograma puede ser no muy adecuada para la corrección radiométrica, aunque en muchos casos puede ser suficiente para una posterior detección de cambios.

	B. Rojo	B. Verde	B. Azul
	44,31	21,23	28,52
HZ	41,63	10,99	29,42
MM	43,12	15,65	26,52
MS	18,04	4,47	10,05
SR	18,49	4,52	10,40
EH	0,92	0,78	0,73
AB255	17,01	4,27	9,28
AB127	16,27	4,19	8,80
AB63	15,37	4,02	8,19
AB33	13,98	3,81	7,36

Tabla 3-2. ECM niveles de gris en experimento 1 tras normalización radiométrica.

El método adaptativo bilineal es el que ha obtenido los mejores resultados. Se puede observar cómo el error es menor cuanto menor es la ventana empleada. No se debe caer en la idea de reducir al máximo posible la ventana ya que, aunque el error

cuadrático medio se reduce, la imagen se ve afectada muchísimo por el ruido y el resultado de la normalización es una imagen con iluminación no uniforme. Esto se debe a que el conjunto de píxeles usado en cada ventana deja de ser adecuado para la corrección. Recordemos que no de los criterios que se deben cumplir para una correcta normalización es que el conjunto de píxeles usado contenga un amplio rango de niveles de gris.

	B. Rojo		B. Verde		B. Azul	
	52,45		48,07		42,56	
HZ	45,43	45,43	42,44	42,44	40,74	40,74
MM	43,22	43,22	39,51	39,51	36,87	36,87
MS	44,69	42,05	40,86	38,25	37,47	35,67
SR	42,16	39,06	38,59	35,58	35,59	33,11
EH	45,98	42,78	41,70	38,75	38,23	36,06
AB255	44,20	40,51	40,66	37,11	36,89	31,25
AB127	43,46	39,79	40,01	36,50	36,28	32,69
AB63	42,34	38,79	39,02	35,59	35,37	33,63
AB33	40,61	37,38	37,39	34,21	33,81	34,24

Tabla 3-3. ECM niveles de gris en experimento 2 tras normalización radiométrica.

En el segundo experimento, el método SR tiene un error menor que el método AB para unos cuantos tamaños de ventana. Pero SR tiene el inconveniente de que requiere de una precisa corrección geométrica previa, pues obtiene los coeficientes de transformación usando la covarianza de los píxeles de las imágenes a normalizar y de referencia. Asimismo, si representamos la relación entre los píxeles de las imágenes del segundo experimento para cada banda, obtenemos nubes de puntos que para nada tienen una forma lineal. Esto hace que la función propuesta por SR, aunque reduzca el error cuadrático medio, reduzca también el contraste de las imágenes perdiendo información radiométrica y calidad visual.

Resumiendo, los métodos HZ y MM no son adecuados para la corrección radiométrica. El método MS y el de especificación de histograma (EH) han mostrado ser sensibles a los cambios dando lugar a valores más altos de error que otros procedimientos, aunque también han demostrado ser efectivos reduciendo las diferencias radiométricas existentes (al reducir el valor de error inicial). Por tanto, estos métodos sólo deberían usarse en imágenes con un porcentaje de píxeles de cambio pequeño, enmascarando píxeles de cambios a priori en el cálculo de los coeficientes o usarlos de forma adaptativa, aplicando el método a pequeñas ventanas de la imagen.

Los métodos que mejores resultados han proporcionado en los dos experimentos son SR y AB. Generalmente, es recomendable usar el método adaptativo ya que aplicará distintas transformaciones a los distintos y no la misma para todos como hace el método SR. Esto último a veces provoca que la imagen a

corregir pérdida contraste tras la transformación por no ser lineal la relación entre los píxeles de las imágenes.

3.6 Futuras mejoras

Como se ha podido comprobar en las pruebas realizadas, el método SR obtenía un error cuadrático medio menor que el MS. Esto nos lleva a pensar que una posible mejora del método adaptativo bilineal consistiría en usar el método SR en el procesamiento de cada ventana, aunque sería necesaria una corrección geométrica previa.

Se ha comentado también que para obtener mejores resultados es necesario identificar a priori píxeles de cambio y excluirlos de los cálculos de los parámetros mediante máscaras. Una forma de automatizar esta obtención de cambios a priori consistiría en elaborar un mapa de correlación entre las dos imágenes. Una vez obtenido este mapa se aplicaría un umbral para obtener aquellos píxeles que serán determinados como “no-cambio” y que podrían usarse en el cálculo de los parámetros. El inconveniente de este método es que sólo es aplicable con una corrección geométrica previa y si el registro tiene un error muy pequeño.

Hemos observado también que, generalmente, la relación entre dos imágenes no es lineal, por tanto, sería útil el estudio de nuevos métodos no lineales. Recordemos que en los métodos lineales la transformación venía dada por:

$$u_k = a_k * x_k + b_k$$

donde x_k y u_k son los valores digitales del píxel antes y después de la transformación, y a_k y b_k los coeficientes de normalización. Se trata entonces de usar una función no lineal f tal que:

$$u_k = f(x_k)$$

Para obtener f se usaría una red neuronal que aprenda la relación entre los píxeles de las dos imágenes. Asimismo, la selección del conjunto de píxeles de entrenamiento se podría llevar a cabo usando el método del mapa de correlación propuesto anteriormente.

Capítulo 4

Correcciones geométricas

La corrección geométrica es una de las operaciones de pre-procesamiento necesarias para una precisa detección de cambios. Las imágenes empleadas en la detección de cambios, generalmente, están tomadas en diferentes fechas y pueden estar tomadas desde diferentes puntos de vista, incluso con diferentes sensores. Este tipo de corrección consiste en asignar a cada píxel unas coordenadas, con el fin de que los píxeles que tengan las mismas coordenadas, representen la misma área de la superficie terrestre.

Para los métodos de detección de cambios empleados en este proyecto, no es estrictamente necesario asignar un sistema de coordenadas a cada una de las imágenes, sino que es suficiente con hacer corresponder los píxeles de cada una para que puedan ser comparados.

Al proceso de asignar un sistema de coordenadas a una imagen se conoce como *georreferenciación*. Si lo que realizamos es una alineación de las imágenes en las que se va a detectar cambios, el proceso se conoce como *registro de imágenes*. Las imágenes deben registrarse dos a dos. La imagen usada como referencia se conoce como imagen maestra y la que ajustamos en el proceso, imagen esclava. Si la imagen maestra está georreferenciada antes de realizar el registro de las imágenes, la imagen esclava también lo estará al finalizar el registro.

Sería imposible diseñar un método universal aplicable a todas las tareas de registro debido a la diversidad de imágenes y degradaciones a las que pueden verse afectadas. En el registro debe tenerse en cuenta, además de la deformación geométrica entre las imágenes, las posibles deformaciones radiométricas existentes en ellas. En algunos casos y dependiendo de la técnica de registro empleada, para obtener un buen resultado a veces será necesario realizar un ajuste radiométrico global previo, como la especificación de histograma o alguna de las normalizaciones radiométricas vistas en el Capítulo 3.

Generalmente, el proceso de registro de imágenes consiste en cuatro pasos:

a) Detección de *características*. Las características son objetos distintivos como esquinas, bordes, intersecciones de líneas, etc., que pueden ser seleccionados manual o incluso automáticamente. A estos puntos distintivos se les conoce también como *puntos de control* (CPs).

b) Emparejamiento de características (o *matching*). En este paso se establece la correspondencia entre los puntos de control detectados en las imágenes maestra y esclava. Para ello, se utilizan *descriptores* o alguna medida de similitud entre las características.

c) Estimación del modelo de transformación. En este paso son estimados el tipo y los parámetros de la función que permitirá el alineamiento de la imagen esclava a la imagen maestra.

d) Remuestreo de la imagen y transformación. En este último paso se aplica la función de transformación calculada en el paso anterior y se aplican técnicas de interpolación para aquellos píxeles con coordenadas no enteras.

La implementación de cada paso tiene sus problemas típicos. El primero de ellos es decidir qué tipo de características son las adecuadas para las imágenes que deseamos registrar. Estas características deben ser objetos distintivos que se extiendan por las imágenes y que se puedan detectar fácilmente, a veces, buscando que éstos sean interpretables físicamente. Los conjuntos de características detectadas en ambas imágenes deben tener bastantes elementos en común, incluso en situaciones donde las imágenes no cubren exactamente la misma escena o cuando hay objetos ocultos u otros cambios inesperados como los que pretendemos obtener al final del proceso de detección de cambios. En el proceso de emparejamiento, podemos encontrar parejas de características con diferencias debidas a cambios de iluminación o al funcionamiento del sensor o sensores que adquirieron las imágenes. Estas diferencias deben ser tenidas en cuenta a la hora de obtener los descriptores o medidas de similitud que se emplearán. Los descriptores y características deben ser suficientemente estables para que no se vean afectados por variaciones inesperadas y ruido. El algoritmo de *matching* debe ser suficientemente robusto y eficiente, y no verse afectado por las características sin correspondencia en la otra imagen. Otro aspecto a tener en cuenta es el tipo de función de transformación, la cual debe seleccionarse acorde a la información a priori que se disponga sobre el proceso de adquisición de las imágenes y las deformaciones esperadas. Si no se dispone de esta información, el modelo debe ser lo suficientemente flexible para tratar todas las deformaciones posibles. Finalmente, una técnica de interpolación adecuada debe ser seleccionada teniendo en cuenta la precisión de la misma y el coste computacional.

4.1 Detección automática de características

En los primeros trabajos de registro de imágenes, los puntos de control se seleccionaban a mano por un experto *fotointérprete*. Posteriormente, surgieron

trabajos que trataban de obtener características en las imágenes de forma automática. Estos métodos se podían clasificar en dos tipos: métodos basados en área y métodos basados en características.

Los métodos basados en área se centran en el *matching* o emparejamiento y, por tanto, no se detectan características en la imagen esclava, aunque pueden modificarse para que sí las usen y así poder reducir el coste computacional como veremos más adelante.

Estos métodos consisten en extraer estructuras destacadas en las imágenes: regiones (bosques, lagos, campos, etc.), líneas (líneas costeras, carreteras, ríos, etc.) o puntos (esquinas, intersecciones de líneas, etc.). Estas estructuras deben extenderse por toda la imagen y ser estables a diferentes cambios en el tiempo (geométricos, condiciones radiométricas, presencia de ruido y cambios en la escena) para que puedan ser localizadas en ambas imágenes.

Para este proyecto se han implementado en IDL/Envi funciones de extracción de puntos (Harris y Minimum eigenvalues) y los de extracción de regiones o *blobs* (SIFT y SURF) más conocidos.

4.1.1 Harris

El detector de Harris fue desarrollado por Chris Harris y Mike Stephens en 1988 (Harris & Stephens, 1988) para ayudar a los investigadores a construir interpretaciones del entorno de un robot a partir de una secuencia de imágenes. Trataban de obtener correspondencias entre puntos en imágenes consecutivas (*frames*) de una secuencia captada por una cámara móvil.

Este detector se basa en el operador de Moravec, intentando corregir sus limitaciones. El operador de Moravec se presentó en 1977 (Moravec, 1977). En este trabajo se define el concepto de *puntos de interés* como regiones distintivas en las imágenes y concluye que son estos puntos los que deberían usarse para encontrar correspondencias entre *frames* consecutivos.

El operador de Moravec es considerado un detector de esquinas ya que define los puntos de interés como aquellos donde la variación de intensidad de un píxel es alta en todas las direcciones. Para medir esta variación V en un píxel P de una imagen propone colocar una ventana cuadrada (3×3 , 5×5 o 7×7 píxeles) centrada en P y desplazarla un píxel en las ocho direcciones posibles (horizontal, vertical y cuatro diagonales). La variación V para un desplazamiento dado se obtiene sumando los cuadrados de las diferencias de los píxeles que se corresponden en las dos ventanas, como se muestra en la Figura 4-1.

El cuadrado rojo representa la ventana original y el azul la ventana desplazada un píxel en la diagonal. La ecuación general de V en el píxel (x, y) para un desplazamiento (u, v) sería:

$$V_{u,v}(x,y) = \sum_w [I(x_i, y_i) - I(x_i + u, y_i + v)]^2$$

donde $I(x, y)$ es la intensidad del píxel (x, y) y (x_i, y_i) es cada uno de los puntos de la ventana que se desplaza en una dirección (u, v) del conjunto de posibles direcciones $\{(1,0), (1,1), (0,1), (-1,1), (-1,0), (-1,-1), (0,-1), (1,-1)\}$.

La variación de intensidad C en el píxel (x,y) será la mínima variación V calculada en las ocho direcciones:

$$C(x,y) = \min(V_{u,v}(x,y))$$

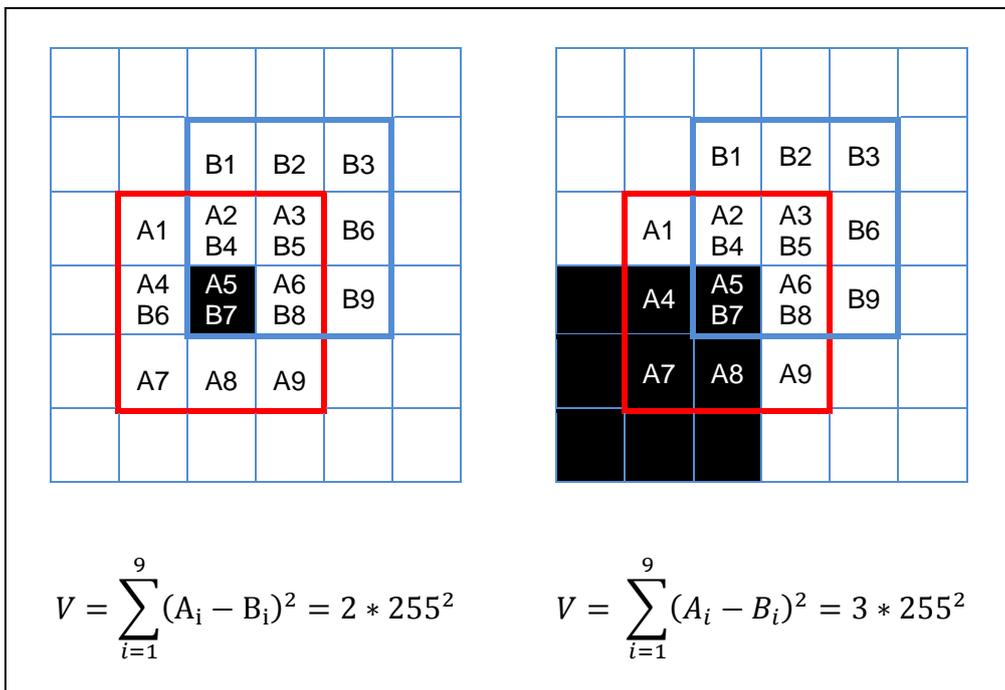


Figura 4-1. Ventanas usadas por el operador de Moravec.

El cálculo de C da lugar a un mapa de valores cuyos máximos se corresponden con las esquinas de la imagen. Para obtener estos puntos de interés bastaría con aplicar un algoritmo de supresión de no-máximos que elimine aquellos píxeles que no son máximos locales. Esto sería así en el caso de una imagen binaria, pero para una imagen en niveles de gris, el mapa de valores C , puede dar lugar a máximos locales de poca variación de intensidad y que por tanto no se corresponden con puntos de interés. Por este motivo, es necesario asegurar que los máximos locales superan un umbral th antes de marcarlos como puntos de interés.

El operador de Moravec tiene los siguientes inconvenientes y fueron los que motivaron el desarrollo del detector de Harris:

- La respuesta es anisotrópica, es decir, depende de la dirección en la que se calcule, ya que sólo tiene en cuenta desplazamientos discretos en las ocho principales direcciones. Además, no es invariante a rotaciones.
- Es ruidosa por usar una ventana cuadrada y binaria y, en consecuencia, la distancia del píxel central a los bordes de la ventana no es la misma en todas las direcciones.
- Es muy sensible al ruido. Un píxel ruidoso aislado dará lugar a variaciones en todas las direcciones y, por tanto, que sea seleccionado como punto de interés.
- Otro inconveniente es la alta respuesta del operador a algunos ejes que no están en ninguna de las ocho direcciones principales, o a las imperfecciones en los mismos, como las producidas por ruido.

El operador de Harris surge como posible solución a estos problemas. Con el fin de hacer el operador menos ruidoso se usa una ventana gaussiana dando mayor peso al píxel central. Aunque la ventana es cuadrada se puede considerar prácticamente circular ya que el valor en los extremos es casi cero.

$$w_i = \frac{e^{-(u^2+v^2)}}{2\sigma^2}$$

w1 0.04	w2 0.12	w3 0.04
w4 0.12	w5 0.36	w6 0.12
w7 0.04	w8 0.12	w9 0.04

Figura 4-2. Ventana gaussiana usada en operador de Harris.

De esta forma, la ecuación para calcular la variación de intensidades queda así:

$$V_{u,v}(x, y) = \sum_w w_i [I(x_i, y_i) - I(x_i + u, y_i + v)]^2 \quad (4.1)$$

Con la finalidad de resolver el problema de la respuesta anisotrópica, los autores proponen una nueva función que obtenga las variaciones de intensidad en todas las direcciones (u, v) (véase 3.2). La nueva función surge de aproximar el término $I(x_i + u, y_i + v)$ por su expansión en series de Taylor hasta el término de primer orden y sustituyendo en la función original:

$$I(x_i + u, y_i + v) \approx I(x_i, y_i) + u \frac{\partial I(x_i, y_i)}{\partial x} + v \frac{\partial I(x_i, y_i)}{\partial y}$$

$$V_{u,v}(x, y) = \sum_w w_i \left[u \frac{\partial I(x_i, y_i)}{\partial x} + v \frac{\partial I(x_i, y_i)}{\partial y} \right]^2$$

$$V_{u,v}(x, y) = \sum_w w_i \left[u^2 \left(\frac{\partial I(x_i, y_i)}{\partial x} \right)^2 + 2uv \frac{\partial I(x_i, y_i)}{\partial x} \frac{\partial I(x_i, y_i)}{\partial y} + v^2 \left(\frac{\partial I(x_i, y_i)}{\partial y} \right)^2 \right]$$

$$V_{u,v}(x, y) = Au^2 + 2Cuv + Bv^2$$

donde $A = \left(\frac{\partial I}{\partial x} \right)^2 \otimes w$, $B = \left(\frac{\partial I}{\partial y} \right)^2 \otimes w$, $C = \left(\frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \right) \otimes w$, y \otimes es el operador de convolución y el gradiente de la imagen se aproxima por las siguientes convoluciones:

$$\frac{\partial I}{\partial x} \approx I \otimes (-1, 0, 1)$$

$$\frac{\partial I}{\partial y} \approx I \otimes (-1, 0, 1)^T$$

Harris y Stephen observaron que la función de cálculo de variación de intensidades se puede expresar en forma matricial y que sus autovalores eran proporcionales a las principales curvaturas de la imagen. De esta forma, dependiendo del valor de estos autovalores, se puede determinar si un determinado píxel es de una región interior, un borde, o una esquina.

$$V_{u,v}(x, y) = Au^2 + 2Cuv + Bv^2$$

$$V_{u,v}(x, y) = [u \ v] \begin{bmatrix} A & C \\ C & B \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \quad (4.2)$$

$$V_{u,v}(x, y) = [u \ v] M \begin{bmatrix} u \\ v \end{bmatrix}$$

Sean λ_1 y λ_2 los autovalores de la matriz M, el plano descrito por los mismos puede ser dividido en tres regiones, diferenciando regiones interiores, bordes o esquinas tal como muestra la Figura siguiente:

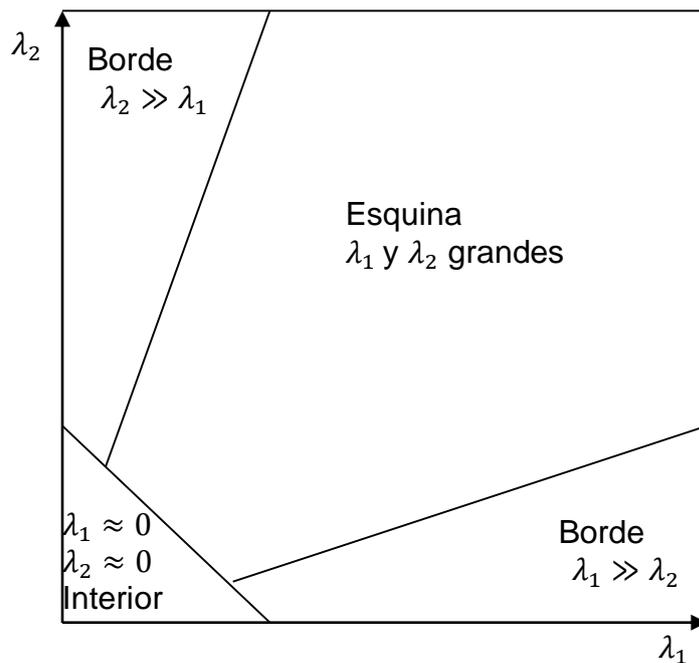


Figura 4-3. Regiones de los autovalores de la matriz M. Harris.

En el operador de Moravec se obtenía un mapa de valores en el que las esquinas correspondían a los máximos locales. Harris y Stephens en su trabajo proponen algo similar, evitando así la necesidad de calcular los autovalores de la matriz M.

$$C(x, y) = \det(M) - k (\text{trace}(M))^2$$

$$\det(M) = \lambda_1 \lambda_2 = AB - C^2$$

$$\text{trace}(M) = \lambda_1 + \lambda_2 = A + B$$

$$k = \text{constante}$$

De esta forma, $C(x, y)$ es negativo para píxeles de borde y positivo para píxeles interiores y esquinas. Para separar las esquinas del resto de píxeles es necesario aplicar un umbral al mapa de valores determinado por la función $C(x, y)$. Los autores proponen un valor de la constante k entre 0.04 y 0.06.

Recopilando, para obtener puntos de interés con Harris: 1º) obtenemos la matriz de autocorrelación M para cada píxel; 2º) obtenemos un mapa de valores con la función $C(x, y)$; 3º) aplicamos un umbral al mapa de valores obtenido para poner a cero todos los píxeles que pertenecen a regiones interiores o bordes; y por último, 4º) aplicamos un algoritmo de supresión de no-máximos para obtener los máximos locales del mapa de valores que se corresponden con las esquinas de la imagen.

4.1.2 Minimum eigenvalues (mínimos autovalores)

Con este método propuesto en (Shi & Tomasi, 1994) las características o puntos más fuertes para emparejar en dos imágenes, son obtenidos directamente de los autovalores de la matriz hessiana.

Harris proponía aplicar la fórmula $C(x, y) = \det(M) - k (\text{trace}(M))^2$ para evitar el cálculo de los autovalores de la matriz M. Recordemos, que aquellos puntos de interés que son interesantes para la corrección geométrica, son aquellos con autovalores grandes, que representan esquinas, ruido u otra textura de fácil seguimiento. En este método, se obtiene el mínimo autovalor para cada píxel en la matriz M y se seleccionan aquellos que superan un umbral determinado, es decir, sean λ_1 y λ_2 los autovalores de la matriz y λ un valor umbral, se seleccionan los píxeles tales que:

$$\min(\lambda_1, \lambda_2) > \lambda$$

Para ello recordemos, los autovalores de la matriz se obtienen a partir de la ecuación:

$$|M - \lambda I| = 0$$

Y el menor autovalor por tanto por:

$$\lambda_m = \frac{A + B - \sqrt{(A - B)^2 + 4C^2}}{2}$$

donde A, B, y C son los elementos de la matriz M (véase 4.2).

4.1.3 SIFT

Las características SIFT (Scale Invariant Feature Transform) fueron desarrolladas por D.G. Lowe (Lowe, 2004) para ayudar en los procesos de correspondencia de imágenes y reconocimiento de objetos. Son invariantes a transformaciones de escala y rotación, y parcialmente invariantes a cambios en la iluminación y punto de vista. A diferencia del detector de Harris, que localizaba esquinas en la imagen, el detector SIFT localiza regiones o *blobs*.

La obtención de estas características se realiza a través de distintas fases:

1. Detección de extremos en el espacio-escala.
2. Localización de puntos característicos.
3. Asignación de orientaciones.
4. Cálculo de descriptores.

Detección de extremos en el espacio-escala:

Mientras que el detector de Harris localiza puntos característicos en una escala predeterminada, el detector SIFT amplía su espacio de búsqueda en el conocido como

espacio-escala, añadiendo invariancia a escala. Al trabajar en una única escala, el operador de Harris sólo localiza objetos de interés en un nivel de detalle concreto, pero estos objetos pueden tener diferentes tamaños y estar a diferentes distancias del sensor que capta las imágenes.

Como no es posible a priori obtener las escalas útiles de una imagen, es necesario analizar todas las posibles a partir de su representación espacio-escala. Esta representación consiste en una familia de imágenes suavizadas a diferentes niveles de detalle (definido por un parámetro de escala). Una imagen $I(x, y)$ genera la familia:

$$L(x, y; \sigma) = g(x, y; \sigma) \otimes I(x, y)$$

donde $g(x, y; \sigma)$ es el kernel de suavizado en la escala y \otimes es el operador de convolución. En el espacio-escala, las estructuras a bajo nivel de detalle, deben resultar de la simplificación de las estructuras a alto nivel de detalle sin introducir nuevos elementos cuando se sube de escala. Por tanto, $g(x, y; \sigma)$ debe ser una Gaussiana como demuestran (Koenderink, 1984) y (Lindeberg, 1994).

$$g(x, y; \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Para $\sigma = 0$, $g(x, y; 0)$ es una función impulso $\delta(x, y)$, de manera que $L(x, y; 0) = I(x, y)$ es la imagen original. Además, $\sqrt{\sigma}$ es la desviación estándar de $g(x, y; \sigma)$ de manera que $L(x, y; \sigma)$ no contiene detalles menores a $\sqrt{\sigma}$. La Figura 4-4 muestra una imagen en diferentes escalas.

Debido a que los detalles pequeños desaparecen cuando la escala aumenta, el detector de características SIFT usa pirámides de imágenes en las que se reduce la resolución de las mismas a escalas más altas.





Figura 4-4. Detalle de una imagen de satélite a diferentes escalas.

Otra operación empleada en la localización de características es el Laplaciano de la función Gaussiana (LoG). El Laplaciano está basado en la segunda derivada y es muy sensible al ruido, de ahí que suele ir acompañado de una función de suavizado para atenuarlo, en este caso la función Gaussiana. Al convolucionar el LoG con una imagen, los cruces por cero permiten detectar bordes, y los extremos, la detección de manchas o blobs cuyo tamaño depende de la escala del LoG. En (Mikolajczyk, 2002) el autor demostró que los extremos obtenidos por el operador LoG, correspondían a características más estables en la imagen que las obtenidas por otras funciones como el gradiente o el operador de Harris.

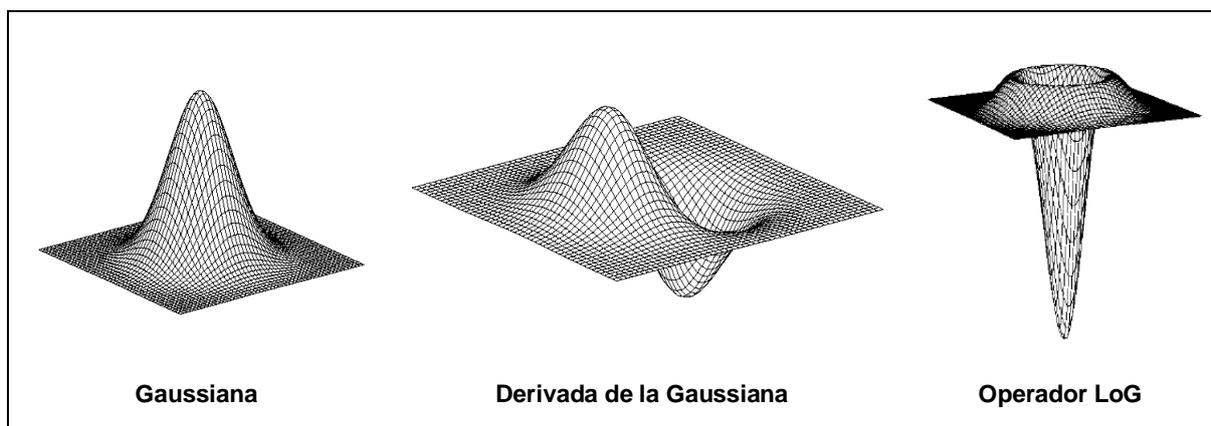


Figura 4-5. Operadores Gaussiana, DoG y LoG.

Para obtener la pirámide de imágenes suavizadas o pirámide gaussiana, se calculan varias escalas entre σ y 2σ (a lo que se conoce como *octava*), todas con la misma resolución. A partir de la escala 2σ , se divide a la mitad la resolución y se usa esta resolución reducida para la siguiente octava. En cuanto al valor de σ , Lowe, propone $\sigma = 1.6$, ya que aunque la repetibilidad aumenta cuando se incrementa σ , también disminuye la eficiencia del detector.

Una vez obtenidas todas las escalas, se obtiene el LoG de cada una de ellas. El LoG se puede aproximar por una Diferencia de Gaussianas (DoG) que es más eficiente desde el punto de vista del coste computacional, pues se obtiene con simples sustracciones, y las imágenes suavizadas necesarias para la operación, se han obtenido previamente al elaborar la pirámide de imágenes ya comentada.

Representaremos matemáticamente estas imágenes con la función $D(x, y, \sigma\sigma)$. La Figura 4-6 muestra el proceso de creación de las pirámides gaussiana y laplaciana.

Las octavas están divididas en s intervalos. Para ello se generan $s + 3$ imágenes de las que se obtienen posteriormente $s + 2$ diferencias de Gaussianas. Éstas se emplean en la fase de detección de extremos tomándolas en grupos de 3 imágenes consecutivas como veremos a continuación. Lowe propone en su trabajo un valor de $s = 3$ que es el que mayor porcentaje de repetibilidad ofrece en sus experimentos. Si aumentamos este valor, la repetibilidad (aunque se podría esperar lo contrario) disminuye debido a que aunque se detectan muchos más extremos, en promedio son más inestables y por tanto no repetibles.

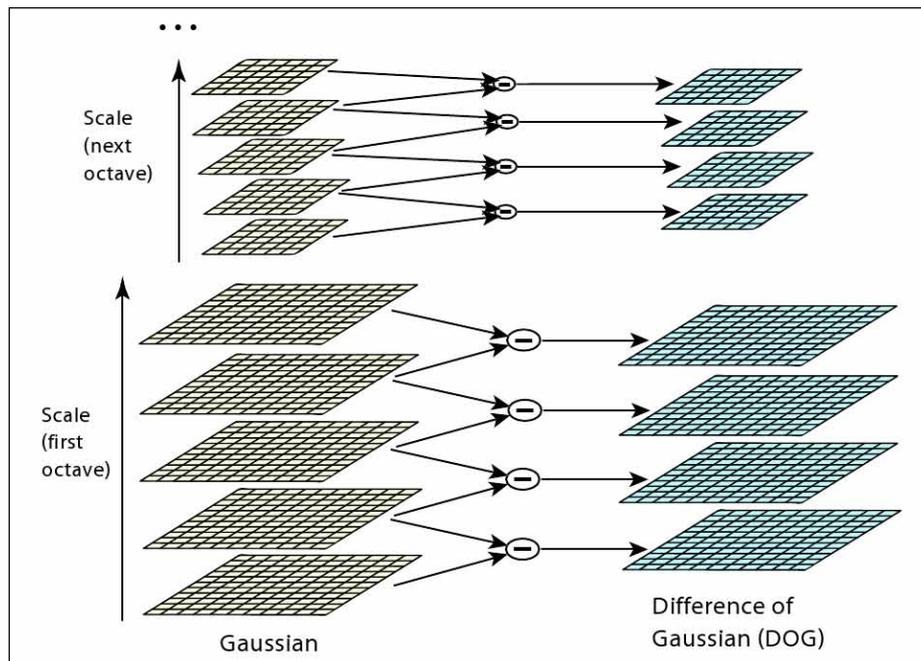


Figura 4-6. Formación de la pirámide Gaussiana de imágenes.

Después de generar las pirámides necesarias, se procede a detectar los extremos locales en las diferencias de Gaussianas, comparando cada píxel con sus 8 vecinos en la misma escala y con los 9 píxeles en cada una de las escalas vecinas. Un extremo será seleccionado si es mayor o menor que sus 26 vecinos.

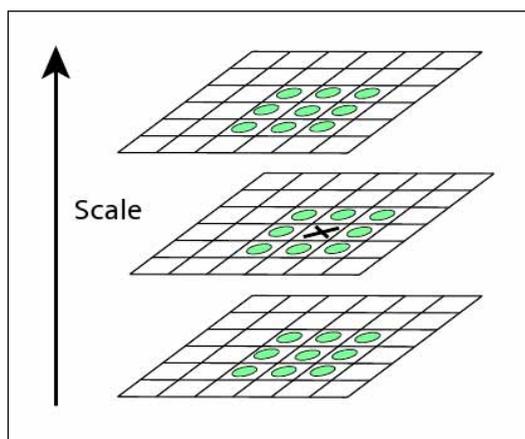


Figura 4-7. Selección de extremos por comparación con sus 26 píxeles vecinos.

Localización de puntos característicos

En versiones anteriores del detector SIFT (Lowe, 1999), se utilizaban directamente los extremos seleccionados en la fase anterior. En trabajos posteriores, se precisa a nivel sub-píxel la localización de estas características. Para ello, se interpola la función $D(x, y, \sigma)$ con su expansión de Taylor hasta el término de segundo orden:

$$D(x, y, \sigma) = D + \frac{\partial D^T}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \mathbf{x}$$

donde D y sus derivadas son evaluadas en el extremo que se quiere precisar (x_1, y_1, σ_1) y $\mathbf{x} = (x - x_1, y - y_1, \sigma - \sigma_1)$ es el desplazamiento desde ese punto. Si derivamos esta función con respecto a \mathbf{x} y la igualamos a cero, obtendremos el desplazamiento necesario para alcanzar el extremo con precisión sub-píxel, $\hat{\mathbf{x}}$:

$$\hat{\mathbf{x}} = -\frac{\partial^2 D^{-1}}{\partial \mathbf{x}^2} \frac{\partial D}{\partial \mathbf{x}}$$

Si el desplazamiento $\hat{\mathbf{x}}$ es mayor que 0.5 en cualquier dimensión, significa que el extremo está localizado más cerca de otro píxel. En este caso se vuelve a realizar la operación de interpolación centrada en este nuevo píxel, hasta alcanzar el desplazamiento que se sumará para obtener la localización definitiva. Tanto la derivada como la Hessiana de D se pueden aproximar usando sustracciones entre los vecinos del extremo que se quiere precisar.

El siguiente paso consiste en desechar aquellos puntos candidatos que tienen bajo contraste. Para esto se obtiene el valor de la función siguiente en cada extremo calculado previamente:

$$D(\hat{\mathbf{x}}) = D + \frac{1}{2} \frac{\partial D^T}{\partial \mathbf{x}} \hat{\mathbf{x}}$$

Lowe propone en su trabajo que todos los extremos con un valor de $|D(\hat{x})| < 0.03$ deben ser descartados. En este proyecto se usa la implementación de Rob Hess (Hess, 2010) que descarta todos los extremos con valor menor a 0.04, suponiendo que el valor del píxel está comprendido en el intervalo $[0,1]$.

También, es necesario rechazar los puntos sobre bordes por su localización inestable. Esto se hace usando las curvaturas de $D(x, y, \sigma)$ dadas por las segundas derivadas y aplicando un criterio similar al de Harris para esquinas. Las dos curvaturas principales de D se pueden obtener calculando la matriz hessiana 2x2 en la escala y localización del punto candidato. Como se hizo anteriormente, las derivadas se pueden obtener con diferencias entre los píxeles vecinos.

$$\mathcal{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

La razón entre las dos curvaturas principales debe ser menor a 10, lo que indica que el punto no pertenece a un borde. Cualquier extremo que no cumpla esta restricción será también rechazado. Para comprobar esta restricción basta con que se cumpla la siguiente desigualdad:

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r + 1)^2}{r}$$

donde $Tr(H)$ es la traza de H , $Det(H)$ es el determinante de H y r toma el valor $r = 10$.

Asignación de orientaciones

En esta fase, a todo punto de interés se le asigna una orientación que se corresponde con la orientación dominante del gradiente en su vecindad. Así, para cada uno de ellos, se toma la imagen $L(x, y; \sigma_0)$ donde σ_0 es la escala en la que se identificó dicho punto, y se calculan la magnitud y la orientación del gradiente a cada uno de los píxeles vecinos:

$$m(x, y) = \sqrt{[L(x + 1, y) - L(x - 1, y)]^2 + [L(x, y + 1) - L(x, y - 1)]^2}$$

$$\theta(x, y) = \tan^{-1} \left[\frac{L(x, y + 1) - L(x, y - 1)}{L(x + 1, y) - L(x - 1, y)} \right]$$

Con estos valores se construye un histograma de orientaciones de 36 bins, que cubra los 360 grados, ponderadas por la magnitud del gradiente y por una gaussiana con $\sigma = 1.5 \sigma_0$. La orientación del punto quedará determinada por el máximo del histograma, pero para precisarlo aún más, se interpola una parábola entre los 3 puntos más cercanos al máximo (véase Figura 4-8).

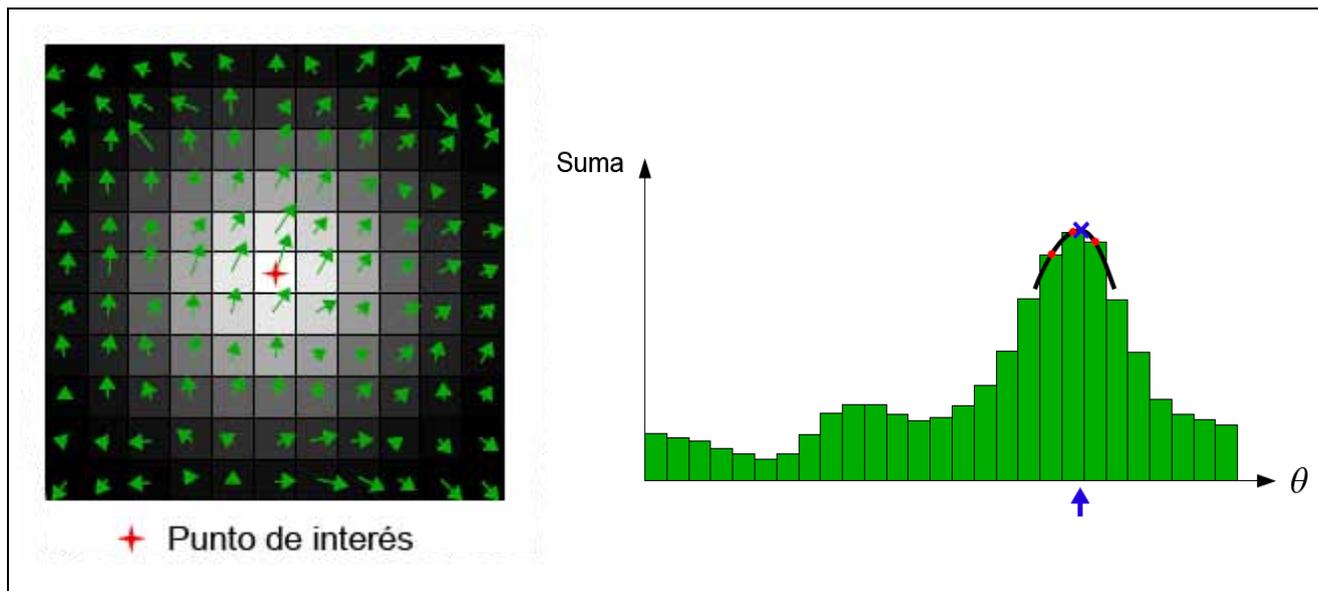


Figura 4-8. Cálculo de orientaciones de los puntos característicos SIFT.

Si existen otros picos mayores al 80% del máximo, se genera un punto independiente para cada pico con la orientación correspondiente, dando lugar a puntos de interés con misma localización pero distintas orientaciones.

Cálculo de descriptores

Hasta aquí se han obtenido puntos de interés en distintas escalas y se les ha asignado una orientación dominante, asegurando la invariancia a escala y rotación. Esta última fase se trata de generar descriptores que proporcionen, a estos puntos de interés, invariancia a otros parámetros como cambios en la iluminación o el punto de vista.

Con el fin de obtener estos descriptores, se calculan las magnitudes y orientaciones del gradiente de los píxeles vecinos del punto de interés, usando la imagen suavizada $L(x, y; \sigma)$ correspondiente a la escala en la que se detectó el extremo. Para que la invariancia a rotación sea efectiva, las coordenadas del descriptor y las orientaciones del gradiente son rotadas respecto a la orientación calculada en el paso anterior. De nuevo se aplica un peso a cada magnitud mediante una gaussiana con un valor $\sigma = \frac{1}{2}w$, donde w es el ancho de la ventana usada para el descriptor.

El descriptor consiste en dividir el vecindario del píxel en 4x4 regiones de 4x4 píxeles y generar un histograma de orientaciones de 8 bins para cada región. Aunque la ventana del descriptor podría tener otro tamaño, dividir en 4x4 regiones es lo que mejores resultados ofrece. Por tanto, en este caso, el parámetro σ de la gaussiana empleada debe tomar el valor 2.

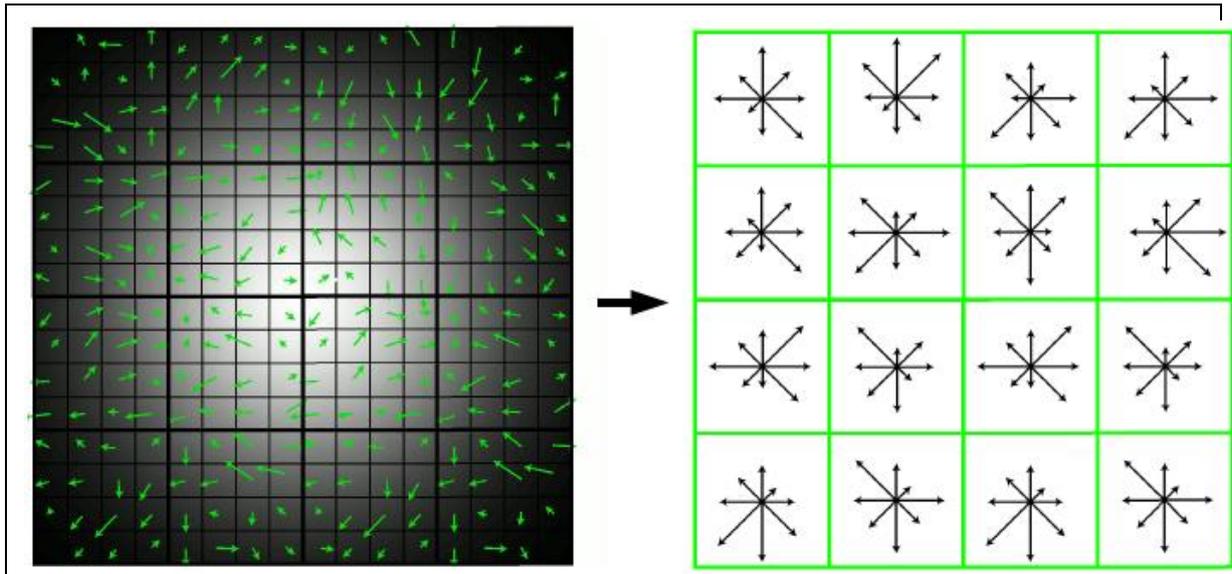


Figura 4-9. Cálculo del descriptor de una característica SIFT.

Con un pequeño desplazamiento espacial, la contribución de un píxel puede pasar de una región a otra, lo que provoca cambios repentinos del descriptor. Lo mismo ocurre con una pequeña rotación.

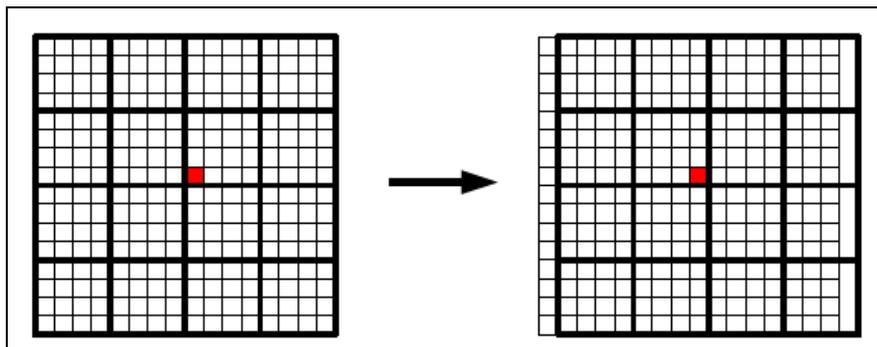


Figura 4-10. Efecto de un pequeño desplazamiento en el cálculo del descriptor.

A fin de evitar este inconveniente, un píxel contribuye a todos sus vecinos multiplicando su valor por un peso $1 - d$, donde d es la distancia al centro de la región y cada una de ellas tiene un ancho de 1 unidad.

Lo mismo ocurre con las orientaciones, a las que se le aplica un peso en función de la proximidad a los bins vecinos.

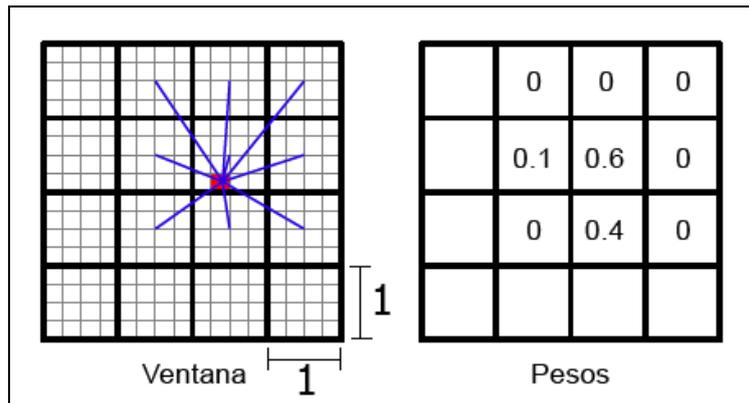


Figura 4-11. Pesos de un píxel en el cálculo de histogramas del descriptor.

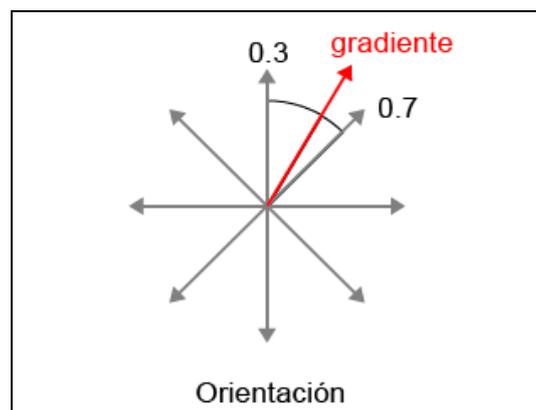


Figura 4-12. Pesos en la asignación de orientaciones a un determinado bin del histograma.

Con todo lo explicado, nos queda un descriptor de $4 \times 4 \times 8 = 128$ elementos para cada punto de interés. Se genera un vector con estos elementos que es entonces modificado para reducir los efectos de cambios en la iluminación (contraste y brillo). Por un lado, un cambio en el contraste en el que cada píxel es multiplicado por una constante, multiplicará los gradientes por la misma constante. A fin de obtener invariancia a cambios en el contraste, se normaliza la magnitud del vector a la unidad. Por otro lado, un cambio en el brillo en el que a cada píxel se suma una constante, no afectará al gradiente que se calcula mediante diferencias entre los píxeles.

En casos de cambios de iluminación no uniforme, las magnitudes de los gradientes se verán afectadas, pero es menos probable que afecte a sus orientaciones. Con el propósito de evitar la influencia de grandes gradientes, se aplica un umbral a las magnitudes impidiendo que superen el valor de 0.2 y volviendo a normalizar la magnitud del vector a la unidad.

Resumen

El detector SIFT se basa en localizar, a diferentes escalas y con precisión sub-píxel, el centro de las manchas o blobs más estables en la imagen. Posteriormente, a cada una de estas características se le asigna la orientación correspondiente al gradiente dominante en su vecindario. Y por último, se genera un descriptor de 128

características que permita identificarlo y ayude en los procesos de correspondencia de imágenes y reconocimiento de objetos.

El descriptor está normalizado en escala porque se trabaja en la escala significativa del punto. La normalización en rotación se obtiene restando la orientación dominante, calculada en la fase 3, a todos los ángulos del histograma calculado en la fase 4. También se obtiene una cierta tolerancia a desplazamientos locales, provocados por cambios de punto de vista, ya que cada histograma abarca varios píxeles. Por último, una tolerancia a cambios de iluminación es obtenida normalizando la magnitud del vector descriptor.

4.1.4 SURF

Este método de detección de características (Bay, Ess, Tuytelaars, & Gool, 2008) trata de mejorar el tiempo de computación de los blobs así como el de comparación de descriptores. Se basa en el uso de imágenes integrales y en las fortalezas de detectores anteriores como SIFT.

La imagen integral $I_{\Sigma}(x)$ en el punto $x = (x, y)$ se define como la suma de las intensidades de todos los píxeles de la imagen I contenidos en una región rectangular comprendida entre el origen y x :

$$I_{\Sigma}(x) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j)$$

Una vez la imagen integral ha sido calculada, basta con realizar tres operaciones de suma y resta para obtener la suma de las intensidades de todos los píxeles en un área rectangular cualquiera de la imagen, tal como muestra la siguiente Figura.

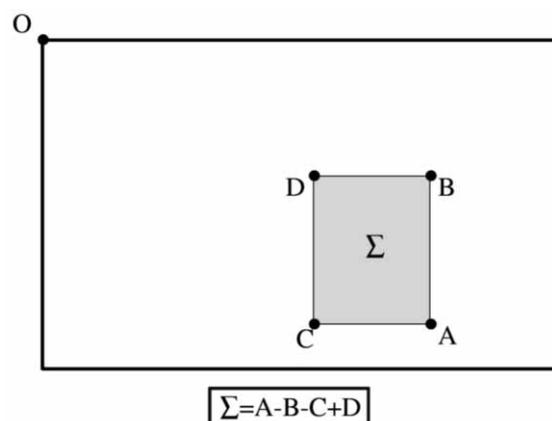


Figura 4-13. Suma de píxeles de una región a partir de imagen integral.

El detector SURF se basa en la matriz Hessiana para localizar blobs (aquellos donde el determinante es máximo) y para seleccionar la escala de cada característica. Esta matriz se define como:

$$\mathcal{H}(\mathbf{x}, \sigma) = \begin{bmatrix} D_{xx}(\mathbf{x}, \sigma) & D_{xy}(\mathbf{x}, \sigma) \\ D_{xy}(\mathbf{x}, \sigma) & D_{yy}(\mathbf{x}, \sigma) \end{bmatrix}$$

donde $D_{xx}(\mathbf{x}, \sigma)$ es la convolución de la Gaussiana de segundo orden $\frac{\partial^2}{\partial x^2} g(\sigma)$ con la imagen en un punto $\mathbf{x} = (x, y)$, y de manera similar para $D_{xy}, D_{yx},$ y D_{yy} . Si recordamos, en el detector SIFT se usaban aproximaciones de la LoG para obtener la Hessiana. En la caso del detector SURF, para conseguir una mayor rapidez de cálculo, los autores del detector proponen usar nuevos filtros cuadrados que aproximen las derivadas de segundo orden y sean fáciles de obtener usando la imagen integral.

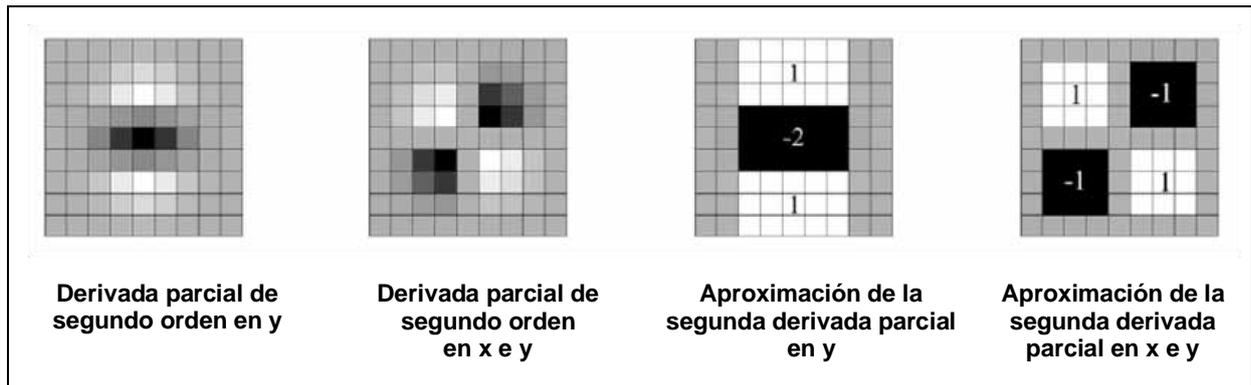


Figura 4-14. Filtros aproximados de derivadas parciales de segundo orden.

En la imagen anterior se pueden observar: a la izquierda, los filtros gaussianos de segundo orden discretos y de tamaño 9x9; a la derecha, las aproximaciones usadas por los autores del detector SURF en su método, quienes demostraron que se obtienen resultados similares o incluso mejores en rendimiento. Para un filtro 9x9, se requerían 81 accesos a array y operaciones con los filtros originales. Con las nuevas aproximaciones y haciendo uso de imágenes integrales, tan sólo son necesarios 8 accesos y operaciones. Incluso cuando el tamaño del filtro aumenta, mientras que el coste computacional de las Laplacianas originales aumentaba, para las nuevas aproximaciones permanece constante e independiente del tamaño.

Para corregir las diferencias en respuesta de los filtros originales y las aproximaciones, se usa un peso w de tal forma que:

$$\det(\mathcal{H}_{aprox}) = D_{xx}D_{yy} - (wD_{xy})^2$$

En el trabajo original se propone un valor de $w = 0.9$. Los valores máximos obtenidos por el determinante anterior en el espacio-escala, nos determinará los puntos de interés de la imagen. Al igual que SIFT, en el detector SURF es necesario previamente construir el espacio-escala de la imagen. Pero a diferencia del primero, donde era necesario construir una pirámide de imágenes convolucionando la imagen de entrada con un filtro gaussiano y submuestreando en cada escala, en SURF basta con aplicar a la imagen original filtros cada vez de mayor tamaño. De nuevo es

evidente la reducción del coste computacional al no tener que crear costosas pirámides de imágenes.

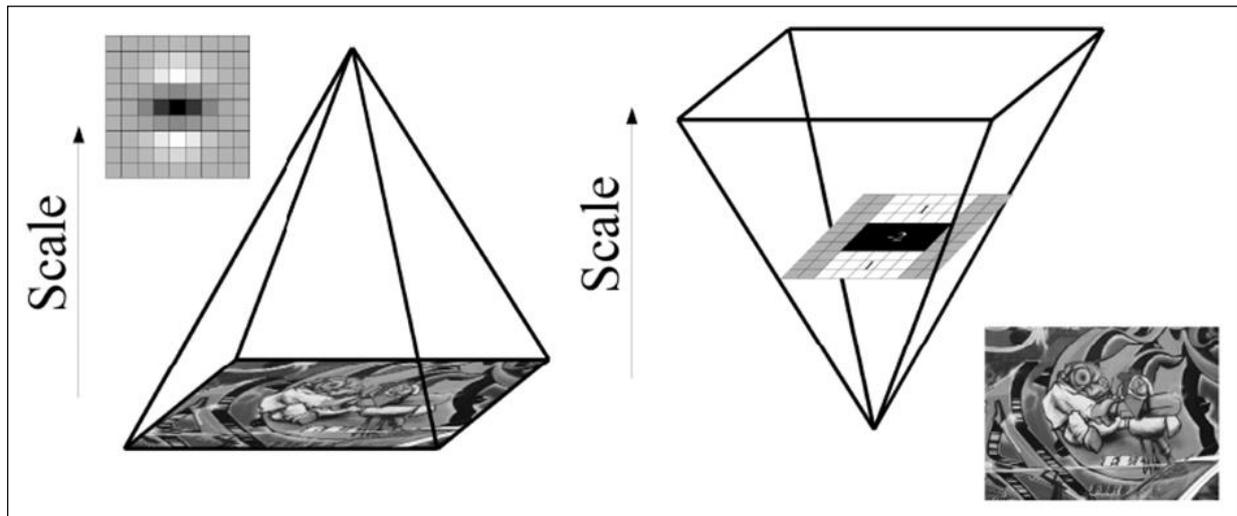


Figura 4-15. Uso diferente de pirámides en SIFT y SURF.

El espacio-escala se divide en n octavas en las cuales la imagen es filtrada a diferentes escalas entre σ y 2σ . En SURF, la menor escala se obtiene de la salida de aplicar los filtros 9x9 vistos anteriormente. Estos filtros corresponden a una Gaussiana real con $\sigma_0 = 1.2$. Las siguientes escalas son obtenidas aumentando el tamaño de los filtros manteniéndose siempre el ratio entre ambos parámetros:

$$\sigma_{aprox} = \text{Ancho del filtro} \cdot \frac{\text{Escala filtro base}}{\text{Ancho filtro base}} = \text{Ancho del filtro} \cdot \frac{1.2}{9}$$

Con el fin de mantener la estructura de los filtros y dejar un píxel central, éstos deben aumentar en al menos dos píxeles por cada lóbulo negativo y positivo (ver Figura 4-16), lo que implica un incremento mínimo de 6 píxeles. Por lo tanto, en la primera octava, la segunda escala se obtiene aplicando un filtro de 15x15, con una escala $\sigma = 15 \cdot \frac{1.2}{9} = 2$. La siguientes serían aplicando filtros de 21x21 y 27x27.

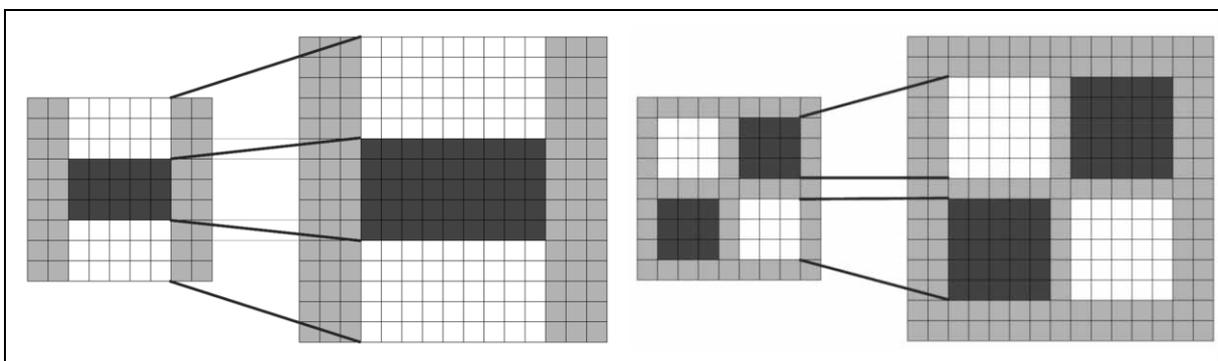


Figura 4-16. Filtros D_{yy} y D_{xy} para dos escalas sucesivas (9x9 y 15x15).

Con el filtro 27x27 se obtendría una escala de $\sigma = 3.6$, mayor que el doble de la escala inicial, pero es necesario ya que la búsqueda de extremos se realiza entre

los vecinos de la misma escala y los de las escalas contiguas y, por lo tanto, la primera y última escala de la octava no pueden usarse para extraer extremos.

Una vez creado el espacio-escala, se procede a localizar los blobs característicos de manera similar al método SIFT. En primer lugar, se descartan aquellos puntos cuyo determinante de la hessiana no supera un determinado umbral. Después de la umbralización, se localizan los máximos comparando la respuesta de cada píxel con sus 26 vecinos (8 de la misma escala, 9 de cada escala contigua). Por último, se realiza una interpolación mediante series de Taylor para afinar a nivel sub-píxel la localización de estos máximos.

Cuando ya se han localizado los puntos de mayor respuesta, se procede a generar sus descriptores. El descriptor SURF describe la distribución de intensidades de los píxeles vecinos de cada punto de interés de una manera similar a SIFT pero, en este caso, se usan imágenes integrales y filtros conocidos como wavelets de Haar para incrementar la robustez y reducir el coste de computación. Los wavelets de Haar son filtros simples que pueden usarse para encontrar gradientes en las direcciones x e y .

La extracción del descriptor puede ser dividida en dos tareas. En primer lugar, a cada punto de interés se le asigna una orientación basada en la información extraída de una región circular alrededor del mismo y, a continuación, se construye una región cuadrada alineada en la orientación seleccionada de la que se extrae el descriptor SURF.

Con la finalidad de obtener la orientación dominante de un punto de interés (y de esa forma asegurar la invariancia a rotación), primero se calcula la respuesta del wavelet de Haar de tamaño $4\sigma_i$ para los píxeles dentro de un radio $6\sigma_i$ del punto detectado, donde σ_i es la escala donde se localizó el mismo. Los vectores obtenidos son ponderados por una Gaussiana de $\sigma = 2\sigma_i$ y, posteriormente, representados como puntos en un plano con sus componentes x e y en los ejes de abscisas y ordenadas respectivamente. Seguidamente, se rota un segmento de círculo cubriendo un ángulo de $\frac{\pi}{3}$ alrededor del origen (sección gris en la imagen de la Figura 4-17) y, en cada posición, las componentes x e y son sumadas y usadas para obtener un nuevo vector. La orientación dominante vendrá dada por la dirección del vector de mayor longitud.

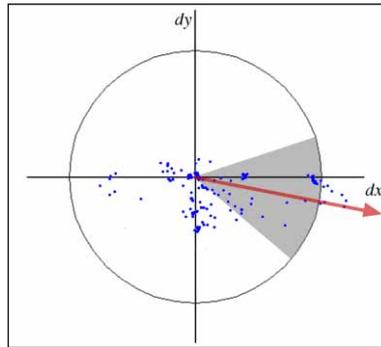


Figura 4-17. Cálculo de orientación dominante de características SURF.

Para extraer el descriptor se crea una región cuadrada de tamaño 20σ con la orientación seleccionada en el paso anterior y alrededor del punto de interés. Esta región es dividida en 4×4 sub-regiones y a cada una de ellas se aplica un wavelet de Haar en 5×5 puntos distribuidos regularmente por la misma. Si llamamos dx y dy al resultado del wavelet de Haar en la dirección horizontal y vertical respectivamente (respecto a la orientación dominante seleccionada), se obtiene por cada subregión las siguientes 4 características:

$$[\sum dx, \sum dy, \sum |dx|, \sum |dy|]$$

lo que nos proporcionan un descriptor de $4 \times 4 \times 4 = 64$ características, que es invariante a rotación, escala, brillo y, después de normalizar a la unidad, al contraste. Se puede obtener un vector extendido de 128 características si el cálculo de dx y $|dx|$ se divide para $dy < 0$ y $dy \geq 0$ y lo mismo para el cálculo de dy y $|dy|$ respecto al signo de dx .

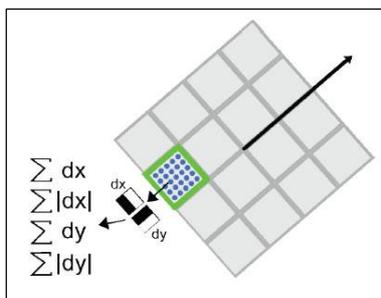


Figura 4-18. Cálculo de descriptores SURF.

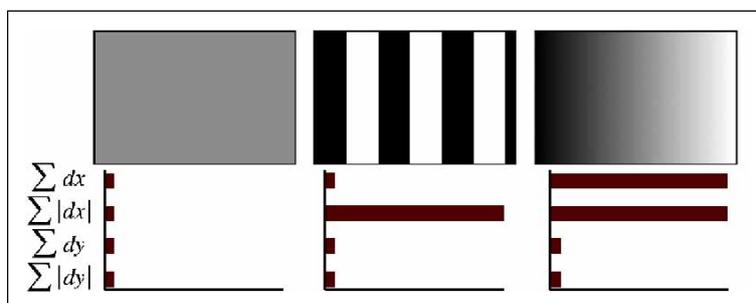


Figura 4-19. Ejemplos de descriptores para cada una de las regiones en un descriptor SURF.

4.2 Matching o emparejamiento de características

Las características detectadas en un par de imágenes maestra-esclava, pueden ser emparejadas teniendo en cuenta o una descripción simbólica, o los valores de intensidad y distribución espacial de los píxeles vecinos.

Recordemos que había dos tipos de métodos de registro: basados en área y en características. En los basados en área no es necesario detectar características previamente aunque sí podrían facilitar la búsqueda de correspondencias. Se basan en ir desplazando una ventana de la imagen maestra de tamaño predefinido (o incluso la imagen entera) sobre la imagen esclava hasta encontrar el máximo para una función de similitud. Los píxeles centrales de las ventanas para las que se alcanza el máximo constituirán un par de puntos de control.

Dos métodos de este tipo han sido implementados en este proyecto: correlación cruzada normalizada (NCC) e información mutua (MI).

La correlación cruzada normalizada trata de encontrar el máximo de la función:

$$NCC(m, n) = \frac{\sum_{x,y} [f(x, y) - \hat{f}(x, y)][w(m + x, n + y) - \hat{w}]}{\sqrt{\sum_{x,y} [f(x, y) - \hat{f}(x, y)]^2 \sum_{x,y} [w(m + x, n + y) - \hat{w}]^2}}$$

donde f es el fragmento de imagen maestra que se pretende encontrar en la imagen esclava, \hat{f} es la media de las intensidades de los píxeles de la imagen f , w es el fragmento de imagen esclava bajo la ventana en estudio y \hat{w} es la media de las intensidades de los píxeles de la imagen w . Este método puede usarse con precisión sub-píxel siempre que la imagen w sea interpolada adecuadamente. El movimiento de la ventana no tiene por qué estar limitado a una traslación, pudiéndose añadir invariancia a rotación y escalado, eso sí, con un gran aumento del coste computacional debido a la ampliación del espacio de búsqueda. En este proyecto, se ha optado por reducir este espacio asumiendo que pequeñas ventanas en la imagen maestra sólo han sufrido una traslación en la imagen esclava.

Podemos ver un ejemplo de la correlación cruzada normalizada en la Figura 4-20. A la izquierda, vemos la imagen completa, a la derecha, la ventana que se pretende localizar y el resultado de aplicar la correlación cruzada normalizada. Como se puede observar, la función presenta un máximo cuando la ventana y la imagen coinciden.

Además del método de búsqueda de correspondencias comentado y para reducir el tiempo de ejecución, se ha implementado el *matching* por correlación entre pares de puntos característicos. Éste consiste en obtener un conjunto de puntos característicos en las imágenes maestra y esclava y, posteriormente, para cada par de puntos, seleccionar como puntos de control aquellos que ofrecen un máximo para la función de correlación. Al seleccionar primero puntos característicos también se

evita obtener correspondencias erróneas en zonas donde la intensidad de los píxeles es homogénea.

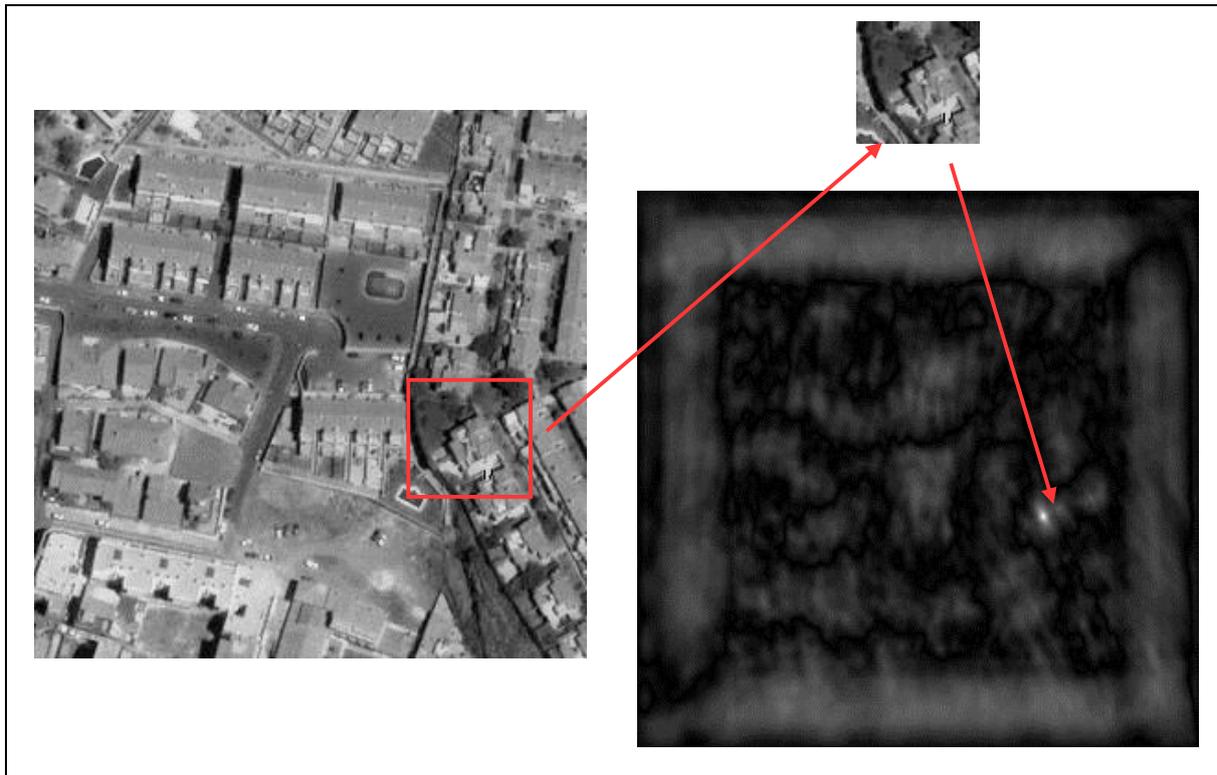


Figura 4-20. Búsqueda de una ventana por correlación cruzada normalizada.

La correlación cruzada depende de la intensidad de los píxeles de las imágenes, así que una corrección radiométrica podría ser necesaria antes del registro para normalizarlas y reducir las diferencias. La información mutua o mutual information (MI) soluciona este inconveniente ya que se basa en la dependencia estadística entre los dos conjuntos de datos y no se ve afectada por los cambios de intensidad de los píxeles de una imagen a otra. Este método puede ser de gran utilidad en las imágenes de satélite, donde es habitual encontrar cambios de intensidad o color en los píxeles debidos a cambios en la iluminación y a cambios estacionales.

La MI de dos variables aleatorias A y B se obtiene de:

$$MI(A, B) = H(A) + H(B) - H(A, B)$$

donde $H(A)$ y $H(B)$ son las entropías de A y B , y $H(A, B)$ es la entropía conjunta. Si consideramos que A y B son imágenes, ambas estarán corregistradas cuando $MI(A, B)$ es máxima. Las entropías se obtienen con las siguientes expresiones:

$$H(A) = \sum_a -P_A(a) \log P_A(a)$$

$$H(B) = \sum_b -P_B(b) \log P_B(b)$$

$$H(A, B) = \sum_{a,b} -P_{A,B}(a, b) \log P_{A,B}(a, b)$$

donde $P_A(a)$ y $P_B(b)$ son las funciones de masa de probabilidad, y $P_{A,B}(a, b)$ es la función de masa de probabilidad conjunta, obtenidas por las siguientes expresiones:

$$P_{A,B}(a, b) = \frac{h(a, b)}{\sum_{a,b} h(a, b)}$$

$$P_A(a) = \sum_b P_{A,B}(a, b)$$

$$P_B(b) = \sum_a P_{A,B}(a, b)$$

donde h es el histograma conjunto del par de imágenes, una matriz con la forma:

$$h = \begin{bmatrix} h(0,0) & h(0,1) & \dots & h(0, N-1) \\ h(1,0) & h(1,1) & \dots & h(1, N-1) \\ \dots & \dots & \ddots & \dots \\ h(M-1,0) & h(M-1,1) & \dots & h(M-1, N-1) \end{bmatrix}$$

Los valores $h(a, b)$, $a \in [0, M-1]$, $b \in [0, N-1]$, es el número de pares de píxeles que tienen valor de intensidad a en la primera imagen y b en la segunda. Se trataría entonces, como en el caso de la correlación cruzada, de detectar puntos característicos en las imágenes maestra y esclava, y emparejar aquellos que ofrezcan un máximo de la función MI usando como parámetros ventanas centradas en dichos puntos. En (Maes, Collignon, Vandermeulen, Marchal, & Suetens, 1997) los autores exponen que MI es sólo una de las medidas de dependencia estadística, y proponen usar otra medida conocida como *coeficiente de correlación de entropía* (ECC) que deriva de MI y ofrece mejores resultados. Es esta medida la que usaremos en este proyecto y su ecuación es la siguiente:

$$NMI = \frac{H(A) + H(B)}{H(A, B)}$$

$$ECC = 2 - \left(\frac{2}{NMI}\right)$$

El ECC será usado en otra fase del proyecto: la detección de cambios. Una vez detectados los cambios, se procede a un ajuste local para eliminar posibles errores de registro. Volveremos a hablar de ello en un capítulo posterior.

Además de estos dos métodos basados en áreas se han implementado dos métodos de *matching* basados en características: SIFT y SURF.

Recordemos que para cada región SIFT o SURF se obtiene un vector de 128 características. El proceso de emparejamiento consistiría entonces en seleccionar para cada blob en la imagen maestra el correspondiente en la imagen esclava de tal

forma que sus vectores de características estén a la menor distancia posible. Calcular la distancia para cada par posible supone un alto coste computacional, de ahí que las distintas implementaciones de SIFT y SURF propongan diferentes algoritmos para acelerar el proceso.

El más habitual es el uso de *árboles-kd* o *kd-tree*. Los vectores de características son clasificados en árboles de este tipo los cuales permiten acelerar la búsqueda de la correspondencia más cercana (Beis & Lowe, 2003). En este proyecto se ha usado la implementación SIFT de (Hess, 2010) que incluye la implementación de los árboles-kd y el algoritmo de búsqueda de vecinos más cercanos que detallaremos a continuación.

En primer lugar, se crea un árbol-kd con las características detectadas en la imagen esclava. Sea N el número de blobs detectados con vectores de características en R^k (generalmente, con $k=128$). Ahora, se divide el espacio de datos en la dimensión i que presenta una mayor varianza, usando como valor de corte el valor de la mediana m en esa dimensión. Se crea un nodo que almacene los valores i y m . Los blobs con un valor en la dimensión i menor que m quedarán en el lado izquierdo del árbol, en caso contrario se almacenarán en el lado derecho. De esta forma, se procede iterativamente por cada rama del árbol hasta clasificar todos los vectores de características, dando lugar a un árbol binario balanceado con profundidad $d = \log_2 N$.

Una vez creado el árbol-kd es fácil encontrar los vecinos más próximos a cada vector de características de la imagen maestra siguiendo el procedimiento a continuación. Primero se recorre el árbol en profundidad decidiendo en cada nodo que lado del árbol tomar basándonos en los valores almacenados en el nodo (dimensión i , valor de la mediana m). Una vez alcanzado el nodo que contendría al vector consultado, se procede con el resto de ramas no visitadas, pudiendo aquellas cuyos nodos están más alejados que el nodo hoja ya visitado. Esta operación puede ser muy costosa cuando se trabaja con puntos o vectores de muchas dimensiones como es el caso, por este motivo, se determina un valor máximo de nodos hoja a visitar. Con el fin de asegurar que los vecinos más próximos encontrados no se vean afectados por la distribución de los puntos y la estructura del árbol, durante el recorrido se va manteniendo una cola con prioridad donde se almacenan los nodos no visitados, ordenados por la distancia al vector de características cuyos vecinos más próximos se está buscando. Durante la búsqueda, cuando se hace una decisión en cada nodo interno en una dirección, se añade un elemento a la cola con prioridad con la opción no tomada. Cuando un nodo hoja es alcanzado, se extrae la cabeza de la cola y se procede a su examen de la misma forma. La búsqueda finaliza cuando el número máximo de hojas visitadas es alcanzado o todo el árbol es analizado.

Otra opción para encontrar los emparejamientos de las características SIFT y SURF es usar la librería FLANN (Muja, 2011) Fast Library for Approximate Nearest Neighbors basada en su artículo (Muja & Lowe, 2009) que, como su nombre indica,

nos permite buscar los descriptores más cercanos de manera rápida y aproximada y se basa en los árboles-kd comentados anteriormente.

La precisión de la búsqueda de los descriptores más cercanos depende del número de nodos hoja que se visitan durante la misma. Pero aumentar el número máximo de hojas visitadas provoca un aumento del coste de computación y una disminución de rendimiento. La librería FLANN hace uso de las ideas de (Silpa-Anan & Hartley, 2008) quienes proponen buscar en paralelo en N árboles (4 en nuestro caso) construidos introduciendo algo de aleatoriedad. En la creación de un árbol-kd se seleccionaba la dimensión del descriptor con mayor varianza y se dividían en dos grupos a partir la mediana en esa dimensión. En este caso, y teniendo en cuenta que varias de las dimensiones de los descriptores pueden tener una varianza similar, en cada división se escoge aleatoriamente entre las 5 dimensiones con mayor varianza y no necesariamente la de mayor valor. De esta forma, la búsqueda consistirá en encontrar el descriptor más próximo en 4 árboles en paralelo. Pero esta vez, la cola con prioridad donde se almacena los nodos no visitados durante la búsqueda será común para todos los árboles. Así, si se llega a un nodo hoja se continuará buscando en el nodo y árbol que presente una menor distancia en la cola con prioridad.

Usando FLANN se reduce significativamente el tiempo de búsqueda de emparejamientos pero los resultados obtenidos pueden diferir en cada ejecución debido a la aleatoriedad introducida en la creación de los árboles k-dimensionales usados.

Además de la comparación de características SIFT y SURF, y la implementación de la Correlación Cruzada Normalizada (NCC) e Información Mutua (MI), se ha integrado el algoritmo de seguimiento de características Lucas-Kanade (Lucas & Kanade, 1981) basado en pirámides. El seguimiento de características consiste en obtener el movimiento aparente de los distintos elementos (o los distintos niveles de gris) en una sucesión de imágenes. Mediante esta idea podemos buscar pares de características si consideramos que la imagen maestra y esclava son dos fotogramas o instantáneas de una sucesión de imágenes. Consistiría en detectar características con algunos de los métodos citados en el apartado 4.1 y calcular el movimiento sufrido por cada uno de ellos en la imagen esclava. A este movimiento aparente de los niveles de gris entre una imagen y la otra se conoce como *flujo óptico*. De los distintos algoritmos existentes para la estimación del flujo óptico se ha usado el propuesto por Lucas y Kanade implementado en la librería OpenCV.

El algoritmo LK-Tracking está basado en gradientes espacio-temporales. Se parte de la hipótesis de que los niveles de gris permanecen constantes ante movimientos espaciales en un tiempo dado. Dicha hipótesis da lugar a la ecuación general de flujo óptico:

$$I(x, y, t) = I(x + dx, y + dy, t + dt)$$

donde $I(x, y, t)$ corresponde a la intensidad en niveles de gris del píxel (x, y) de la imagen I en el tiempo t . Expandiendo en series de Taylor sobre el punto (x, y, t) :

$$I(x, y, t) = I(x, y, t) + dx \frac{\partial I}{\partial x} + dy \frac{\partial I}{\partial y} + dt \frac{\partial I}{\partial t} + \varepsilon$$

donde ε contiene la información de las derivadas de orden superior. Si se asume que ε es despreciable, la ecuación de flujo óptico puede reescribirse como:

$$I_x u + I_y v + I_t = 0$$

donde (u, v) , con $u = dx/dt$ y $v = dy/dt$, corresponde al vector de flujo óptico y, I_x y I_y son las derivadas parciales horizontal y vertical de la imagen respectivamente. No existe una única solución para la ecuación anterior y se trata de aplicar algunas restricciones para estimarla. El método de Lucas y Kanade asume que el flujo óptico es constante sobre una región. Sea R una región de la imagen I , y (u, v) su vector de flujo óptico asociado, entonces la ecuación anterior se cumple para cada píxel de la región, es decir:

$$I_x(p_i)u + I_y(p_i)v = -I_t(p_i) \quad \forall p_i \in R$$

Organizando el conjunto de ecuaciones en forma matricial obtenemos:

$$\underbrace{\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_y(p_2) \\ \vdots & \vdots \\ I_x(p_N) & I_y(p_N) \end{bmatrix}}_A \underbrace{\begin{bmatrix} u \\ v \end{bmatrix}}_d = - \underbrace{\begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \vdots \\ I_t(p_N) \end{bmatrix}}_b$$

donde la matriz A contiene las derivadas espaciales de la imagen, el vector d corresponden al vector de flujo óptico (u, v) y el vector b contiene las derivadas temporales de la imagen. Pre-multiplicando por la traspuesta de A se tiene:

$$A^T A d = A^T b$$

El vector de flujo óptico se deriva entonces de:

$$d = (A^T A)^{-1} A^T b$$

El cálculo del flujo óptico implica la inversión de la matriz:

$$A^T A = \begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_y I_x & \sum I_y I_y \end{bmatrix}$$

por lo cual la solución existe si la matriz $A^T A$ es invertible y bien condicionada. Shi y Tomasi (Shi & Tomasi, 1994) definen las propiedades que debe cumplir una región para que el flujo óptico sea estimado apropiadamente utilizando la técnica de LK. Sean

λ_1 y λ_2 los autovalores de la matriz $A^T A$ para cierta región R de la imagen, entonces se debe cumplir que $\min(\lambda_1, \lambda_2) > \lambda_{\min} \in \mathbb{R}^+$, lo que garantiza que $A^T A$ es invertible y la región no es ruidosa, y $\lambda_1/\lambda_2 < \tau$, lo que garantiza que $A^T A$ está bien condicionada y no se presentan bordes en una sola dirección.

En la práctica, existen ciertos factores que pueden inducir errores en la estimación. Entre ellos se encuentran la variación temporal de los niveles de gris sobre la región, desplazamientos grandes de la región entre las imágenes e incoherencia de movimiento. El primero es independiente de la técnica de LK pero los otros dos factores pueden ser controlados seleccionando un tamaño apropiado para la región R . Una implementación piramidal de la técnica de LK es propuesta en (Bouguet, 1999) en la cual el flujo óptico es calculado recursivamente sobre versiones de diferentes escalas de las imágenes. En principio el flujo es estimado sobre imágenes en una escala baja para permitir grandes desplazamientos, posteriormente, la escala se reduce para realizar una estimación más precisa y evitar inconsistencias de movimiento. El tamaño de la región se mantiene fijo sobre todas las escalas.

Al final del procedimiento de *matching* tendremos un conjunto de correspondencias entre las imágenes maestra y esclava pero muchas de ellas serán erróneas por diferentes motivos. Para añadir robustez al procedimiento y asegurar que las correspondencias son adecuadas, se realizan dos procedimientos para eliminar lo que se conoce como *outliers*: filtrado por distancia y algoritmo RANSAC. En algunos casos un tercer procedimiento de filtrado estadístico puede mejorar el ajuste como veremos más adelante.

En cualquier transformación sufrida por dos imágenes de satélite georreferenciadas, podemos suponer que todo punto en la imagen maestra tiene su par correspondiente en la imagen esclava dentro de un vecindario de radio r . Podemos emplear esta peculiaridad para marcar como *outliers*, o simplemente no seleccionar, aquellos pares de puntos cuya distancia sea mayor a este valor r . En las imágenes aéreas, podemos encontrar estructuras que generan texturas similares, por ejemplo, en urbanizaciones donde mucho de los edificios son construidos de forma idéntica. Esto da lugar a que, a veces, pares de puntos que no se corresponden en ambas imágenes sean seleccionados por tener una respuesta espectral similar. Aunque muchos de estos emparejamientos pueden ser eliminados por RANSAC posteriormente, se pueden descartar gran cantidad de ellos previamente si tenemos en cuenta la restricción de distancia comentada, lo que reduce drásticamente el tiempo de cómputo y facilita al algoritmo RANSAC llegar a un consenso.

RANSAC (Fischler & Bolles, 1981) del inglés *Random Sample Consensus Set*, es un algoritmo capaz de estimar los parámetros de un modelo matemático dado a partir de un conjunto de datos observados. Los pares de puntos o blobs localizados anteriormente deben seguir un modelo matemático conocido como *Matriz Fundamental* que se representa por la letra F y explicaremos más adelante. RANSAC destaca por su habilidad de lograr una estimación robusta de los parámetros de esta

matriz fundamental con un alto grado de precisión, aunque haya datos atípicos (*outliers*) entre el conjunto de datos.

RANSAC opera de la siguiente manera: Se selecciona aleatoriamente el número mínimo de correspondencias necesarias para estimar los parámetros de la matriz fundamental y se realiza la estimación de la misma a partir de esos datos. Todos los demás pares que se ajusten al modelo sin superar una distancia pasada como parámetro constituyen el Conjunto de Consenso. Si el porcentaje de elementos de este conjunto respecto al número total de correspondencias es mayor que el porcentaje pasado como parámetro al algoritmo, se estiman de nuevo los parámetros de la Matriz Fundamental usando todos los elementos en el Conjunto de Consenso y termina. Si el porcentaje es menor, se regresa al primer paso y se seleccionan aleatoriamente otro conjunto mínimo de elementos para proceder de la misma manera. Después de un número máximo de intentos, se selecciona el Conjunto de Consenso de mayor tamaño y la matriz es estimada de nuevo usando todos los elementos en ese conjunto. Todos los pares de puntos o blobs excluidos de la estimación de la matriz fundamental son marcados como *outliers* y eliminados.

El modelo matemático de la Matriz Fundamental procede de la visión estéreo donde una imagen es captada a la vez desde diferentes puntos de vista. Si consideramos que las imágenes maestra y esclava son la misma escena adquirida desde dos puntos distintos en el mismo plano (son captadas desde la misma altura), podremos aplicar las restricciones de la geometría epipolar a las mismas.

Se conoce como *geometría epipolar* a la geometría que relaciona un par de imágenes captadas por unas cámaras desde posiciones distintas. Esta geometría es independiente de la estructura de la escena, y sólo depende de la posición relativa y de los parámetros internos de los dispositivos de adquisición de las imágenes (en nuestro caso, el sensor del satélite o cámara aerotransportada). La matriz fundamental F encapsula esta geometría, tiene un tamaño 3×3 y rango 2. Si un punto de la escena en la imagen maestra se denomina \mathbf{x} y llamamos \mathbf{x}' a su correspondencia en la imagen esclava, se cumple la relación $\mathbf{x}'^T F \mathbf{x} = 0$. Puesto que F relaciona geoméricamente los puntos de dos imágenes de una misma escena, se puede emplear en el algoritmo RANSAC para eliminar aquellas correspondencias que cumplan dicha relación. En este proyecto se usa la función de estimación de la matriz fundamental usando el algoritmo RANSAC proporcionada por la librería OpenCV.

Aún sometiendo el conjunto de pares a un filtrado por distancia y al consenso por el algoritmo RANSAC, a veces quedan pares erróneos que pueden distorsionar mucho la transformación que se realice posteriormente. En estadística, uno de los métodos más simples para eliminar *outliers* en una distribución de datos se basa en uso de cuartiles. Sea X un conjunto de datos ordenado, Q_1 el primer cuartil (valor que deja a su izquierda el 25% de los datos de la distribución) y Q_3 el tercer cuartil (valor que deja a su derecha el 75% de los datos), se consideran como *outliers* o datos

atípicos aquellos que no superen un límite inferior lim_{inf} y aquellos superiores a un límite superior lim_{sup} donde:

$$\begin{aligned}lim_{inf} &= Q_1 - 1.5 * (Q_3 - Q_1) \\lim_{sup} &= Q_3 + 1.5 * (Q_3 - Q_1) \quad (4.3)\end{aligned}$$

Como veremos un poco más adelante, para cualquier transformación polinomial podemos obtener una estimación del error aportado por cada par de puntos característicos. Teniendo en cuenta esto, podemos eliminar aquellos pares que presenten un error mayor del límite superior dado por (4.3), que por norma general, corresponden a pares mal localizados.

Para concluir el apartado de *matching* de características, es preciso insistir en la importancia de que los pares de características detectados sean suficientes y estén repartidos de la manera más homogénea posible por toda la imagen. Si no se tiene en cuenta esto, en el caso de una posterior transformación por triángulos, perderíamos información de las imágenes y cambios de las zonas que queden fuera del polígono que encierra a todos los puntos de control. Por otro lado, en el caso de usar una transformación polinomial, cuanto más alto sea el grado de la transformación, mayor será la deformación a la que se somete la imagen esclava. Si los puntos no están distribuidos homogéneamente o no son suficientes la imagen sufrirá deformaciones que harán imposible la detección de cambios.

4.3 Estimación del modelo de transformación

Después de haber establecido una serie de correspondencias o puntos de control entre las dos imágenes, se debe construir una función que transforme la imagen esclava para superponerla a la imagen maestra. Esta función de transformación deberá estimarse teniendo en cuenta la deformación geométrica sufrida por la imagen esclava, el método empleado para la adquisición de las imágenes y la precisión requerida para el registro.

Las funciones de transformación pueden dividirse en dos grupos dependiendo de la cantidad de datos que empleen. Por un lado, las funciones globales usan todos los puntos de control para estimar unos parámetros válidos para toda la imagen. Por otro lado, las funciones locales tratan la imagen como una composición de fragmentos más pequeños y los parámetros dependen sólo de los puntos de control que contiene cada fragmento.

En este proyecto se usan las siguientes funciones proporcionadas por ENVI: dos funciones de tipo global llamadas RST y Polinomial, y una de tipo local conocida como triangulación de Delaunay. Todas ellas usan polinomios para la conversión de las coordenadas y para obtenerlos se calcula lo que se conoce como matriz de transformación. El tamaño de la matriz depende del orden de la transformación y sus

elementos se corresponden con los coeficientes de las ecuaciones polinómicas para las cuales se obtiene el mínimo error.

La transformación RST es la más sencilla y consiste en aplicar una rotación, un escalado y una traslación. Para esta transformación se necesitan tres o más pares de puntos de control y viene dada por las ecuaciones a continuación:

$$\begin{aligned}x_0 &= a_0 + a_1x + a_2y \\y_0 &= b_0 + b_1x + b_2y\end{aligned}$$

donde x e y son las coordenadas de referencia (en el sistema de coordenadas de la imagen maestra), x_0 e y_0 son las coordenadas rectificadas (las correspondientes en el sistema de coordenadas de la imagen esclava), y a_i y b_i los coeficientes de la transformación.

Esta transformación no corrige otro tipo de distorsiones diferentes a la rotación, escalado y traslación, para las cuales se debe usar un ajuste del segundo tipo, el que hemos llamado Polinomial.

Para una transformación de orden t es necesario el siguiente número de coeficientes:

$$2 \sum_{i=1}^{t+1} i$$

Las ecuaciones polinómicas para una transformación de orden t tienen la siguiente forma:

$$\begin{aligned}x_0 &= \sum_{i=0}^t \sum_{j=0}^i a_k x^{i-j} y^j \\y_0 &= \sum_{i=0}^t \sum_{j=0}^i b_k x^{i-j} y^j\end{aligned}$$

donde t es el orden del polinomio, a_k y b_k los coeficientes de transformación, y el subíndice k está determinado por la siguiente ecuación:

$$k = \frac{i \cdot i + j}{2} + j$$

Como mínimo, se necesita un número de puntos de control mayor que el grado del polinomio más uno al cuadrado:

$$\#CPs > (grado + 1)^2$$

En ENVI, una transformación polinómica de orden uno incluye un término xy para permitir transformaciones de tipo sesgo o inclinación además de las ya citadas rotación, escalado y traslación:

$$\begin{aligned}x_i &= a_0 + a_1x + a_2y + a_3xy \\y_i &= b_0 + a_1x + a_2y + a_3xy\end{aligned}$$

Cuando el ajuste polinomial no es adecuado se recurre a un ajuste local usando triangulación de Delaunay que consiste en dividir la imagen en triángulos y aplicar una transformación polinómica a cada uno. Los triángulos se construyen tomando puntos de control de tres en tres siguiendo un criterio determinado. Entre los métodos de triangulación, la de Delaunay es la más usada y puede ser construida mediante el criterio de la *circunferencia vacía*: toda circunferencia formada a partir de los tres puntos de cualquier triángulo no debe tener ningún otro punto adentro. Los triángulos definidos de esta manera son los más equiangulares posibles. La imagen Figura 4-21 muestra un ejemplo de red de triángulos formada por 31 puntos de control.

En caso de usar una transformación polinómica, una vez obtenidos los coeficientes de los polinomios, se puede obtener una medida del error cometido para esa transformación. Como vimos anteriormente, los polinomios son construidos de tal forma que minimicen el error cuadrático medio. Sean x_r e y_r las coordenadas obtenidas por la transformación para un punto (x, y) del sistema de referencia de la imagen maestra y x_i e y_i las coordenadas reales, el error cometido en ese punto puede obtenerse con la ecuación:

$$Error\ RMS = \sqrt{(x_r - x_i)^2 + (y_r - y_i)^2}$$

Será el error cuadrático medio dado por la ecuación siguiente:

$$Error\ RMS\ Total = \sqrt{\frac{1}{n} \sum_{i=0}^n (x_r - x_i)^2 + (y_r - y_i)^2} \quad (4.4)$$

donde n es el número total de puntos de control, el que nos dará una idea de cuan buena es la transformación para el conjunto de puntos de control seleccionado.



Figura 4-21. Ejemplo de triangulación de Delaunay.

4.4 Remuestreo de la imagen y transformación

Una vez seleccionada la función de transformación, se procede al ajuste de la imagen esclava. Al hacer esto, generalmente, los píxeles de la imagen esclava no quedarán superpuestos exactamente sobre píxeles de la imagen maestra sino que corresponderán a varios de ellos, de ahí que sea necesario un remuestreo de la imagen. Existen varios algoritmos de remuestreo siendo el vecino más próximo, la interpolación bilineal y la convolución cúbica los más habituales.

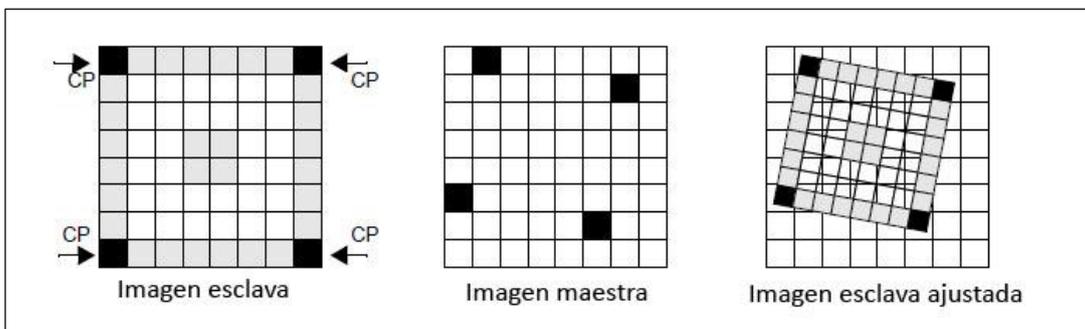


Figura 4-22. Ajuste de imagen esclava mediante CPs.

El vecino más cercano usa el valor del píxel más cercano para asignar el valor del píxel. No modifica los valores, por lo que la imagen seguirá siendo apropiada para un posterior análisis radiométrico. Es el método más fácil y rápido, y adecuado cuando se usa con imágenes de clasificación en la que un promedio de los píxeles no sería factible.

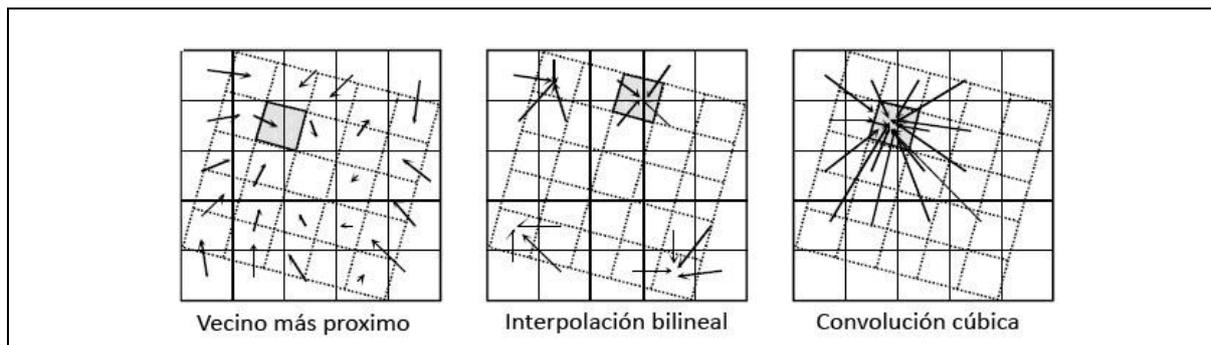


Figura 4-23. Tipos de remuestreo de imágenes.

En interpolación bilineal y convolución cúbica, a cada píxel de la imagen corregida, se le asigna como nivel digital la media de los niveles digitales de los 4 o 16 píxeles vecinos más cercanos respectivamente. Es una media ponderada en función de la distancia del píxel original al corregido. Introduce nuevos valores debido a la interpolación, son más lentos pero evita transiciones bruscas en la imagen al hacerlas más suaves. Ofrecen un mejor aspecto visual.

4.5 Comparación de las técnicas y conclusiones

A continuación, compararemos los distintos algoritmos de extracción y *matching* de características para la corrección geométrica, en diferentes experimentos y usando imágenes con distintas propiedades. El sistema está diseñado para seleccionar un método de detección de características independientemente del método de *matching*, permitiendo combinar las técnicas como se desee. En este apartado pondremos a prueba las combinaciones más habituales. Con el fin de simplificar usaremos para los distintos algoritmos y procedimientos la siguiente nomenclatura:

- HA – Esquinas de Harris.
- ME – Mínimos autovalores de la matriz Hessiana.
- NCC – *Matching* por Correlación Cruzada Normalizada.
- NCCS – *Matching* por NCC con ventana de búsqueda.
- MI – *Matching* por valor de la Información Mutua (ECC).
- MIS – *Matching* por MI con ventana de búsqueda.
- LK – Lucas-Kanade Tracking
- SIFT – Extracción y *matching* de características SIFT.
- SURF – Extracción y *matching* de características SURF.

HA/ME:	Máximo nº puntos	3000
	Distancia mínima	10
	Parámetro K	0.04
	Calidad	0.05
NCC:	Valor mínimo	variable
	Tamaño ventana correlación	128
	Desplazamiento máximo	64

NCCS/MIS:	Máximo nº puntos	1
	Ancho y alto tile	64
MI:	ECC mínimo	variable
	Tamaño ventana MI	128
	Desplazamiento máximo	64
LKT:	Tamaño ventana	128
	Niveles de la pirámide	variable
SIFT:	Parámetros por defecto propuestos en (Lowe, 2004)	
	Ratio máximo distancias dos vecinos más próximos	variable
SURF:	Octavas	3
	Layers	4
	Hessiana	500
	Ratio máximo distancias dos vecinos más próximos	variable

Tabla 4-1. Parámetros experimentos de corrección geométrica.

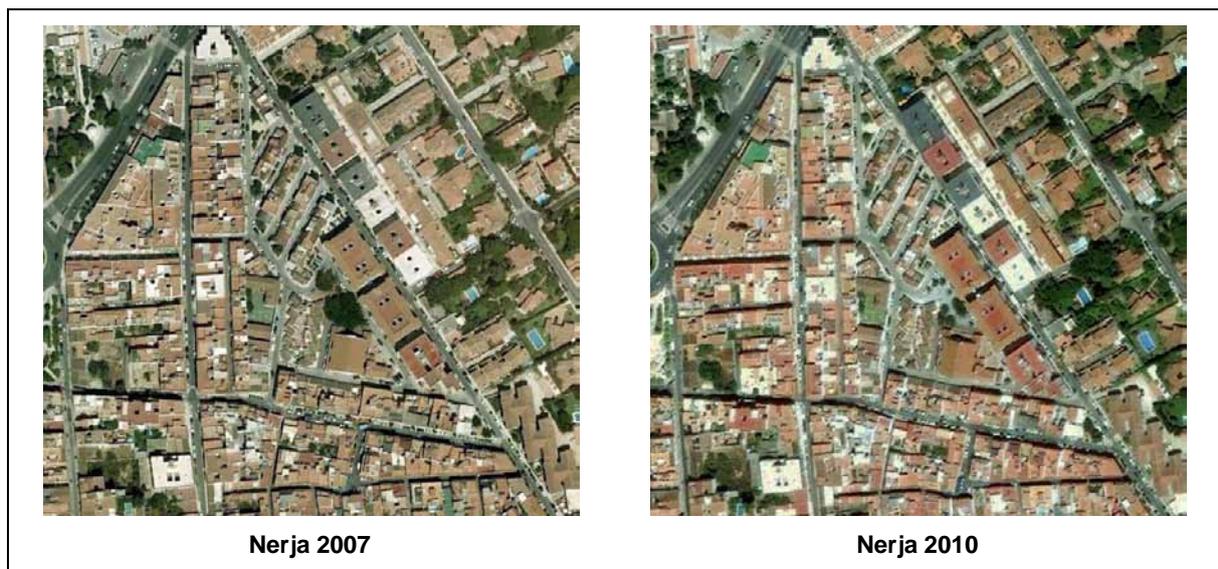


Figura 4-24. Imágenes de zona de Nerja.



R. Victoria 2004	R. Victoria 2006
------------------	------------------

Figura 4-25. Imágenes de zona de Rincón de la Victoria.



Figura 4-26. Imágenes de zona de Huelva.

Usaremos tres pares de imágenes para las pruebas en las cuales se anotará su tiempo aproximado de ejecución en segundos, el número de puntos característicos obtenidos en las imágenes maestra y esclava, el número de emparejamientos, el porcentaje de emparejamientos respecto al número de puntos característicos candidatos y el número de *outliers*, ya sean los descartados por RANSAC, por encontrarse repetidos en distintas escalas cuando se usa la técnica SIFT o por la aplicación de la técnica estadística de los cuartiles. Los parámetros empleados en las pruebas se muestran en la Tabla 4-1 y las imágenes en las figuras 4-25, 4-26 y 4-27.

Experimento 1: Imágenes tomadas de Google Earth de una zona de Nerja adquiridas el 15/08/2007 y el 08/06/2010 con un porcentaje de sombras similar y aproximadamente en la misma dirección a las que se ha aplicado una corrección radiométrica. Tamaño 680 x 620 píxeles. Se muestran en la Figura 4-24.

Experimento 2: Imágenes de Rincón de la Victoria. Tamaño 1250 x 700 y resolución espacial de 0.6 metros/píxel. Se muestran en la Figura 4-25 y tienen las características siguientes:

	07/07/2004	31/07/2006
Hora de adquisición	11:04	11:23
Elevación del Sol	68.2°	67.4°
Acimut del Sol	124.6°	140.1°
Ángulo desde el nadir	12.8°	13.6°

Tabla 4-2. Datos adquisición imágenes de Rincón de la Victoria.

Experimento 3: Imágenes de Huelva. Tamaño 1100 x 700 píxeles y resolución espacial de 0.6 metros/píxel. Los datos de adquisición se muestran en la tabla siguiente y las imágenes en la Figura 4-26.

	11/12/2005	25/05/2007
--	------------	------------

Hora de adquisición	11:34	11:41
Elevación del Sol	28.9°	71.4°
Acimut del Sol	167.8°	147.0°
Ángulo desde el nadir	6.9°	5.0°

Tabla 4-3. Datos adquisición imágenes de Huelva.

Todas las técnicas propuestas buscan correspondencias en una imagen en niveles de gris. Por tanto, cuando realizamos la corrección geométrica de dos imágenes RGB, primero se realiza de forma automática una conversión a gris. En caso de querer realizar detección de cambios en otro tipo de imágenes, podemos usar una de las bandas de las que dispone o transformar la imagen, con cualquiera de los métodos propuestos y comentados posteriormente (Tasseled-Cap y PCA) o a otro espacio de color que pueda resultar de nuestro interés.

En el caso ideal de que el registro de imágenes fuera perfecto, la imagen esclava quedaría posicionada exactamente sobre la maestra y, por tanto, si usamos el método de diferencias (apartado 6.1.1) en la detección de cambios posterior, los cambios quedarían perfectamente identificados con un valor de diferencia distinto de cero y los píxeles sin cambio con un valor de cero. Podemos entonces usar la diferencia de los niveles de gris de los píxeles como una estimación del error cometido en el registro. Cuanto mejor sea el ajuste, menor será la diferencia en los niveles de gris de los píxeles.

De nuevo usaremos la expresión del error cuadrático medio:

$$E = \sqrt{\frac{1}{n * m} \sum_{i=1}^n \sum_{j=1}^m (f_m(x_i, y_j) - f_s(x_i, y_j))^2} \quad (4.5)$$

donde $f_m(x_i, y_j)$ y $f_s(x_i, y_j)$ son niveles de gris del pixel (i, j) en la imagen maestra y esclava respectivamente, y los valores n y m , el número de columnas y filas de las imágenes.

Además de la expresión de error anterior, estudiaremos el error cuadrático medio de la localización de los píxeles para transformaciones polinómicas hasta orden 4, el que hemos llamado error RMS Total en (4.4).

En las técnicas de correlación cruzada e información mutua (ECC), el número de puntos candidatos seleccionados como correspondencias dependerá en gran medida del valor mínimo de correlación cruzada y correlación de entropía respectivamente. Un valor bajo hará que aumente el número de correspondencias establecidas, pero también que aumente el número de *outliers*. Un valor demasiado alto, dará lugar a un número reducido de correspondencias, más precisas, pero a veces insuficientes para una buena transformación.

En sus versiones con ventana de búsqueda también influye el espacio de búsqueda seleccionado. Si este es insuficiente, no se establecerán las correspondencias adecuadamente. Si lo aumentamos, el tiempo de ejecución del algoritmo se incrementará exponencialmente.

En cuanto a la técnica de tracking, no hay ningún parámetro que influya excesivamente en el número y calidad de las correspondencias establecidas. Se obtienen algunas diferencias cuando variamos el número de niveles de las pirámides empleadas en el algoritmo, aunque las diferencias obtenidas en la medida de error que estamos usando en las pruebas son mínimas. Para asegurar la corrección de los puntos establecidos, cuanto mayor sea el desplazamiento sufrido por los píxeles de una imagen a otra, mayor deberían ser el tamaño de la ventana y el número de niveles.

Por otro lado, si recordamos de un apartado anterior, los puntos SIFT y SURF se emparejan con sus correspondientes puntos en la imagen esclava comparando sus descriptores. Para añadir robustez al algoritmo y evitar correspondencias erróneas, ambos se aseguran de que el vecino más próximo lo sea significativamente más que el segundo más próximo. Con este fin, el ratio de las distancias n_1 y n_2 a los dos vecinos más próximos debe cumplir que sea inferior a un valor determinado. En (Lowe, 2004) descartan todos aquellos pares cuyo ratio supera el valor 0.8, y en la implementación de (Hess, 2010) sólo son seleccionados aquellos que no superen el valor constante de 0.49. Este valor es muy importante pues influye en el número de puntos característicos emparejados y asegura la corrección evitando seleccionar pares posiblemente erróneos. En otras palabras, con un valor alto de este parámetro se obtendría mayor número de pares, pero por otro lado, también más pares erróneos. Debido a la gran influencia de este valor en los resultados, se ha eliminado la constante del código fuente y se ha añadido un parámetro a la función de extracción para poder modificarlo.

Podríamos pensar, ya que un mayor número de correspondencias nos proporcionaría un ajuste más preciso, que lo adecuado sería relajar los parámetros anteriores al máximo para permitir el establecimiento del mayor número de ellas. Pero eso también conlleva, como hemos comentado, a emparejar características erróneamente. Por este motivo, en todas las pruebas realizadas se aplica RANSAC para eliminar pares que no se ajusten a la geometría epipolar (como vimos en el apartado 4.2) y, posteriormente, se aplica el método de eliminación de *outliers* por cuartiles. Este último, generalmente, reduce el error E (4.5) aunque en algunos casos puntuales lo incrementa levemente al eliminar correspondencias válidas. En el caso del algoritmo de tracking no suele ser necesario el uso de este filtro, ya que el desplazamiento sufrido por los píxeles es más o menos homogéneo y, al aplicar la técnica, aumenta el error por eliminar algunos pares correctos. Sí que ayuda en el resto de métodos, sobre todo en los detectores SIFT y SURF donde se generan muchos *outliers* cuando se relaja el criterio de las distancias a los vecinos más próximos.

Con todo lo comentado hasta ahora, vemos que estudiar cual de las técnicas de *matching* presentadas ofrece los mejores resultados no es una tarea simple ni fácil, pero vamos a intentar dar respuesta a esta y otras preguntas. En primer lugar, estudiaremos individualmente cada una de ellas modificando el parámetro más decisivo en el número y calidad de las correspondencias. Para NCC y NCCS se obtiene el error E para los valores de correlación mínima en el rango [0.0, 0.7] con un paso de 0.1 (con valores de este parámetro superiores a 0.7 no se obtiene ninguna correspondencia en ninguno de los tres pares de imágenes). Para MI y MIS se varía el valor de mínima correlación de entropía en el rango [0.04, 0.1] con pasos de 0.015 (de nuevo, valores superiores a 0.1, no ofrecen suficientes correspondencias para la corrección). La técnica de tracking se prueba variando el número de niveles de la pirámide de 1 a 5, y las técnicas SIFT y SURF modificando el valor del ratio de las distancias a los vecinos más próximos en el rango [0.1, 1.0] con pasos de 0.1.

Esto no es nada práctico en un estudio de detección de cambios pues puede llevar días de trabajo para unas pocas imágenes, pero con ello intentaremos comprobar si alguna de las técnicas destaca frente a las otras.

El cálculo de la expresión de error E depende de la transformación geométrica que se realice. De las ofrecidas por ENVI, la triangulación, aunque más precisa por realizar ajustes locales, no es adecuada porque marca como píxeles de fondo aquellos que quedan fuera del polígono que encierra a todas las correspondencias. Así que usaremos una transformación polinomial y compararemos los resultados hasta una de orden 4.

De todos los datos obtenidos en las pruebas, en la Tabla 4-4 se presentan aquellos con menor error para cada técnica estudiada y para transformaciones polinómicas hasta orden 4. En cursiva destacan el mínimo error obtenido para cada transformación. En negrita, el valor de error mínimo para cada técnica. Así, por ejemplo, en el experimento 1 la técnica de correlación cruzada normalizada (NCC) obtiene su mejor resultado usando un valor de 0.5 para el parámetro de mínima correlación y aplicando una transformación polinomial de segundo grado a las 129 correspondencias obtenidas en el *matching*. Aunque podemos ver que la técnica que ofrece un mejor resultado para una transformación de orden 2 es la de tracking con puntos de Harris (indicado en cursiva).

Las columnas "Valor del parámetro crítico" nos muestran el valor mínimo de correlación cruzada normalizada para NCC y NCCS, el valor mínimo de correlación de entropía para MI y MIS, el número de niveles de las pirámides para HA-LK y ME-LK y el ratio máximo de las distancias a los vecinos más próximos para SIFT y SURF.

En la Tabla 4-5 se muestran el tiempo de ejecución, número de puntos candidatos iniciales, porcentaje de puntos candidatos seleccionado como emparejamientos, número de *outliers* eliminados por RANSAC y técnica de los cuartiles, porcentaje de *outliers* en el conjunto de emparejamientos, el error E, y por

último el error RMS de ajuste para una transformación hasta orden 4 de los mejores resultados de la Tabla 4-4 (los marcados en negrita). Continuando el ejemplo anterior, en el experimento 1, la técnica NCC obtenía su mejor resultado para una transformación de orden 2 con un error E de 39.0787. Para eso, como podemos ver en la tabla, de 825 puntos candidatos se han establecido 129 correspondencias después de descartar 43 marcadas como *outliers*. El tiempo de ejecución en este caso es alrededor de los 61 segundos.

Debido a la diversidad de técnicas y a que sus resultados dependen de los parámetros que se usen, para compararlas se hará uso en la medida de lo posible de medidas relativas. No podemos concluir que una técnica es mejor que otra por presentar un mayor número de pares de características que otra, pero sí será una buena medida observar cuántos de los puntos característicos candidatos en la imagen maestra terminan siendo seleccionados como una correspondencia.

Como se puede apreciar por los datos en la Tabla 4-5 y la Figura 4-27, las técnicas SIFT y SURF son las que peor resultados obtienen cuando se usan en imágenes de satélite de alta resolución, puesto que de los miles de puntos candidatos sólo logra emparejar en torno al 4% en el mejor de los tres casos y en torno a un 1% en los otros dos donde los cambios producidos son más importantes. Podríamos pensar que esto se debe a que los detectores SIFT y SURF seleccionan como puntos candidatos muchísimos más puntos que Harris o los de mínimos autovalores, pero si comparamos en términos de valor absoluto, tampoco destacan sobre el resto de técnicas. Es más, si observamos del total de pares seleccionados qué porcentaje ha sido marcado como outlier, SIFT y SURF vuelven a destacar frente el resto de técnicas en los tres casos. Estos malos resultados se deben a la sensibilidad del detector SIFT y SURF a los cambios en los elementos de la superficie terrestre, en los ángulos de adquisición, en la iluminación, y en la orientación y superficie de las sombras. Además, en zonas urbanas suelen existir patrones que se repiten, lo que genera puntos (o blobs) característicos muy similares y, por tanto, no seleccionables al no diferenciarse suficientemente entre sí y no superar la restricción del ratio de las distancias a los vecinos más próximos.

Son los algoritmos más rápidos pero generalmente ofrecerán menos pares de puntos, a veces, insuficientes para una buena corrección geométrica. Estas dos técnicas serían útiles solo en el caso de dos imágenes tomadas con características de adquisición muy similares o en aquellos casos en los que se necesite invariancia a rotación y escala.

Los métodos de correlación cruzada e información mutua ofrecen mayores porcentajes de correspondencias respecto a puntos candidatos pero son lentos y aún lo son más los que emplean una ventana de búsqueda. A pesar de que proporcionan más emparejamientos que los métodos SIFT y SURF, el tiempo necesario para calcularlos es excesivo comparados con el obtenido por la técnica de *tracking* (o seguimiento) de características LKT, especialmente, las versiones con ventana de

búsqueda donde el tiempo de ejecución crece exponencialmente con el tamaño de la ventana empleada (véase el incremento de tiempo exagerado en el experimento 2 al usar un tamaño de ventana de 20 píxeles a cada lado, en vez de los 10 píxeles de los otros dos casos). También podemos observar, como cabría esperar, que el porcentaje de emparejamientos en NCCS y MIS es mucho mayor que los de NCC y MI pues aumenta la precisión en la localización de las correspondencias, pero esto también conlleva que sean las técnicas más lentas de todas las presentadas y que su uso con una ventana de búsqueda amplia y un alto número de puntos candidatos sea totalmente impracticable. Los tiempos de ejecución de NCC y MI dependen también del valor del parámetro “Desplazamiento máximo” que determina el radio máximo en el que localizar las correspondencias y permite evitar la comparación de cada punto de la imagen maestra con todos los de la imagen esclava.

El valor de los distintos parámetros críticos en las técnicas NCC, NCCS, MI y MIS refleja como los cambios en las imágenes complican la detección de correspondencias. En el primer experimento (imágenes de Nerja), con una elevación y acimut del sol muy similares como puede observarse visualmente, los valores de correlación cruzada y correlación de entropía que presentan menor error E son superiores que en los otros dos, en los que es necesario relajar este criterio para obtener las correspondencias suficientes. Una variación en el acimut del Sol durante la adquisición dará lugar a sombras en distintas direcciones. De igual forma, las diferentes elevaciones del mismo harán que las sombras arrojadas sean de mayor o menor tamaño.

Estas variaciones pueden observarse en las imágenes de Rincón de la Victoria y Huelva (Figura 4-29) y producen cambios que dificultan la tarea de buscar correspondencias siendo, por tanto, necesario disminuir los valores de correlación mínima. En el caso de que hubiéramos usado los mismos valores de correlación para los tres pares de imágenes, el número de correspondencias detectadas para el primer par de imágenes sería muy superior respecto a los otros dos al haberse producido menos cambios. Es destacable también, los valores tan bajos de correlación de entropía (ECC) usados en el caso de MI y MIS. Mientras que la ECC tiene un rango $[0..1]$, en nuestros experimentos usar valores mayores a 0.1 dificulta demasiado la obtención de correspondencias. Esto nos muestra lo sensible que es este detector a los cambios en la textura, algo que es muy habitual en las imágenes de satélite de muy alta resolución al ser prácticamente imposible obtener imágenes desde un mismo ángulo y con las mismas condiciones de elevación y acimut del Sol.

Hemos probado el tracking con dos tipos de puntos característicos, de Harris y mínimos autovalores de la matriz Hessiana. En ambos casos, la tasa de correspondencias respecto a tiempo de ejecución es similar, al igual que ocurre con el porcentaje de correspondencias respecto al número de puntos candidatos. Por tanto, para usar la técnica de tracking nos es indiferente usar unos u otros pues ambos son estables para el seguimiento. Es suficiente con usar los puntos obtenidos por el

detector Harris, aunque se puede usar ME cuando sea necesario obtener un mayor número de correspondencias.

Respecto al número de correspondencias obtenidas y el tiempo de ejecución, podemos determinar que, de todas las técnicas estudiadas, Lucas-Kanade Tracking usando puntos de Harris o ME es la que mejor resultado ofrece en los experimentos. Se puede comprobar en las gráficas de la Figura 4-27 y la Figura 4-28, cómo las técnicas de tracking y las de correlación cruzada e información mutua con ventanas de búsqueda (NCCS y MIS) son las que encuentran un mayor porcentaje de correspondencias y, por otro lado, las que menos *outliers* eliminan. Esto nos indica que son más precisas en la detección de correspondencias. Las técnicas NCC y MI ofrecen peores resultados debido a la baja repetibilidad de los puntos de Harris en las imágenes de satélite multitemporales y, por tanto, un porcentaje alto es posteriormente filtrado como outlier. SIFT y SURF, además de seleccionar pocas correspondencias, son las que mayor número de *outliers* eliminan en el filtrado.

Además de un conjunto suficiente de pares de puntos, es conveniente que estos estén localizados de la forma más homogénea posible por toda la imagen. Si los puntos que usamos para la transformación no cumplen esto, podríamos obtener imágenes deformadas en aquellas zonas que no cuenten con los emparejamientos necesarios, siendo estas deformaciones más notables cuanto mayor sea el grado del polinomio que se usa para la transformación. Si se usa un ajuste por triangulación, toda la información de la imagen que quede fuera del polígono que engloba todos los pares se perderá, además, será peor el ajuste en las zonas donde estas correspondencias escaseen.

Hemos determinado que LKT es el mejor en cuanto al porcentaje de correspondencias respecto al número de puntos candidatos iniciales y por su tiempo de ejecución. Nos queda, por tanto, estudiar cuál es el mejor ajuste de los ofrecidos por las distintas técnicas. Para ello, haremos uso del Error cuadrático medio en los niveles de gris (4.5) que se lista en la Tabla 4-5 y nos permitirá evaluar el resultado del registro de las imágenes; y el Error RMS Total (4.4) obtenido para un ajuste polinomial hasta orden 4, el cual nos dará una indicación de cuán bueno es el conjunto de pares seleccionados para dicho ajuste. Se debe tener cuidado de no confundir ambas expresiones de error cuadrático medio.

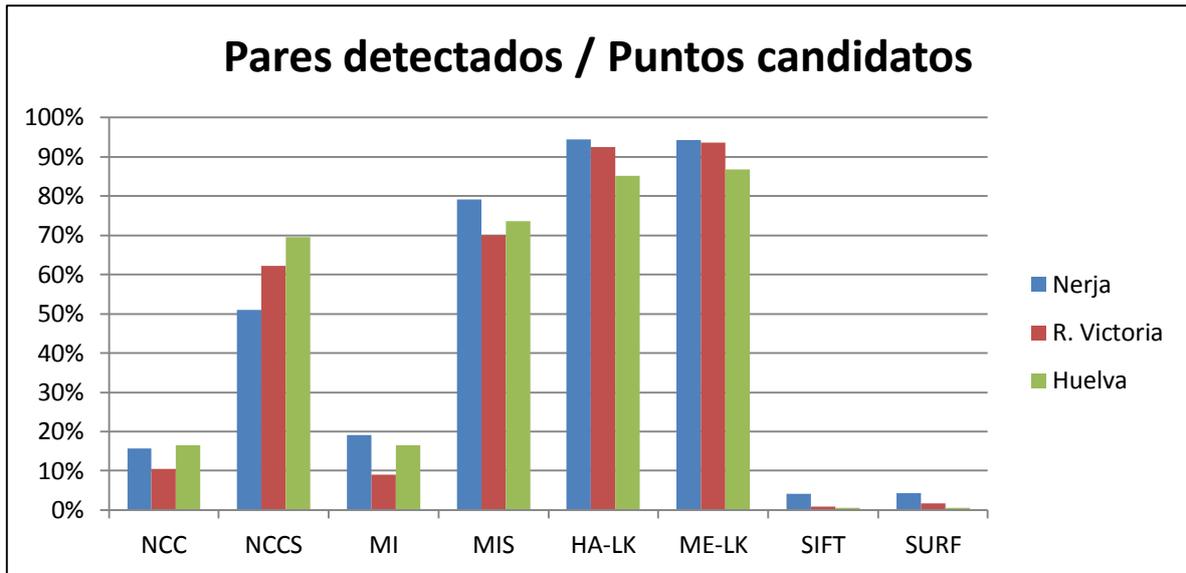


Figura 4-27. Porc. de puntos candidatos seleccionados como emparejamientos tras eliminar *outliers*.

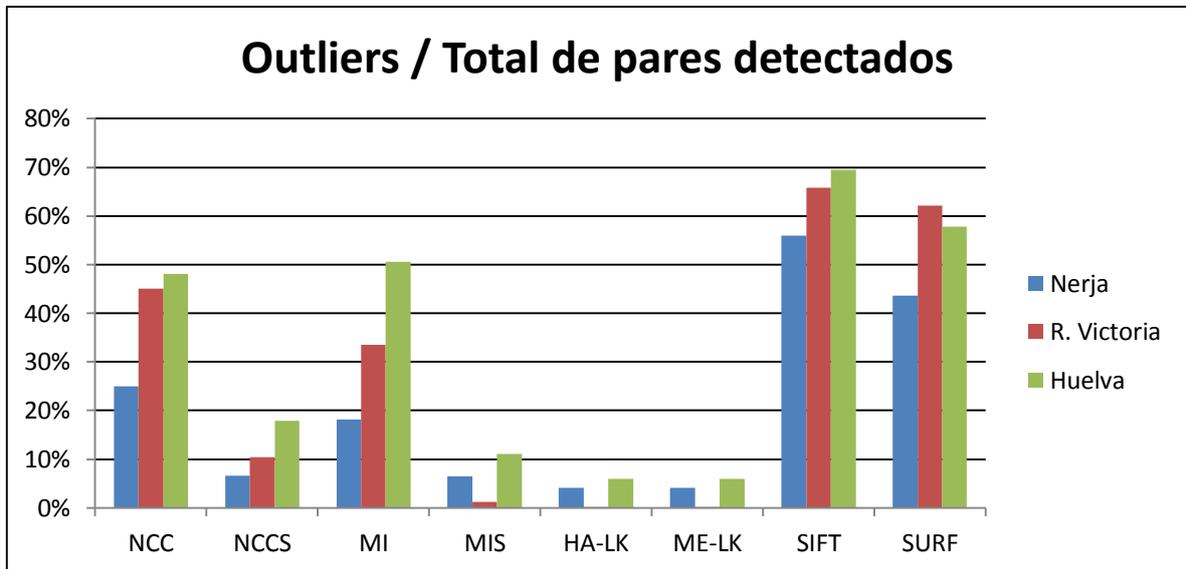


Figura 4-28. Porcentaje de *outliers* respecto al total de correspondencias.

La Figura 4-30 muestra gráficamente el error cuadrático medio en los niveles de gris. Recordemos que, en todos los casos, éste se ha calculado usando los valores para los parámetros más críticos que menor error presentaban tras el registro. Esto no es práctico a la hora de realizar un estudio de detección de cambios, pues la única forma de obtener el mejor valor para estos parámetros es por prueba y error.

	Valor del parámetro crítico				Correspondencias obtenidas				Error E			
	Orden 1	Orden 2	Orden 3	Orden 4	Orden 1	Orden 2	Orden 3	Orden 4	Orden 1	Orden 2	Orden 3	Orden 4
Experimento 1. Nerja												
NCC	0,4	0,5	0,4	0,3	230	129	232	274	39,1145	39,0787	39,5902	40,0617
NCCS	0,6	0,6	0,6	0,5	57	56	54	90	38,9725	38,7985	39,1102	39,1650
MI	0,085	0,07	0,075	0,075	158	216	201	201	39,1111	39,2378	39,7797	40,2522
MIS	0,06	0,085	0,085	0,095	97	87	87	80	39,3537	39,1630	39,0825	39,1115
HA-LK	1	1	1	1	774	788	781	779	38,8544	38,6958	38,6465	38,5532
ME-LK	1	2	1	1	1604	1621	1629	1616	38,8507	38,7044	38,6577	38,5573
SIFT	1,0	0,6	1,0	1,0	243	93	238	228	38,7607	39,2688	39,4518	39,5537
SURF	0,8	0,9	0,9	1,0	148	226	230	362	39,1049	39,0955	39,2131	39,5364
Experimento 2. Rincón de la Victoria												
NCC	0,3	0,3	0,3	0,3	149	148	139	144	46,7296	45,8615	45,4834	45,2906
NCCS	0,3	0,3	0,3	0,3	153	150	148	137	46,7709	45,4381	44,9082	44,1909
MI	0,06	0,07	0,075	0,065	201	138	123	181	46,7633	45,9325	45,4138	45,9007
MIS	0	0,07	0,08	0,07	174	153	117	154	46,8194	45,5550	45,0668	44,4016
HA-LK	5	3	3	5	1275	1273	1243	1272	46,8447	45,8196	45,3516	44,7092
ME-LK	5	4	5	3	2814	2811	2782	2808	46,8551	45,7865	45,2652	44,6851
SIFT	0,8	0,9	0,8	0,8	80	86	74	71	46,7544	46,0175	46,3207	46,6877
SURF	1,0	1,0	1,0	1,0	152	145	141	134	46,8821	47,3645	47,3066	47,9137
Experimento 3. Huelva												
NCC	0,3	0,3	0,3	0,3	151	147	137	150	49,7502	50,2143	50,1697	50,4193
NCCS	0,3	0,1	0,1	0,1	110	139	143	137	49,0924	49,1192	49,1072	48,8091
MI	0,075	0,09	0,07	0,06	152	95	159	191	49,3748	49,4925	50,1384	51,4782
MIS	0,085	0,09	0,08	0,06	117	109	140	145	49,2408	49,2260	49,2189	49,1088
HA-LK	3	3	3	3	809	802	810	783	49,1198	48,9870	48,9724	48,8373
ME-LK	5	3	5	4	2152	2140	2154	2092	49,1259	48,9885	48,9693	48,8607
SIFT	0,9	0,9	0,8	1,0	44	40	37	49	49,8093	51,4618	52,1653	58,1656
SURF	0,9	1,0	0,9	1,0	38	54	41	62	49,8355	50,6622	52,9633	56,3439

Tabla 4-4. Datos de error en los experimentos de corrección geométrica.

	Tiempo de ejecución	Puntos candidatos	Grado Polinomio	Matches	Outliers	%Matches	%Outliers	Error E	Error RMS				
									Orden 1	Orden 2	Orden 3	Orden 4	
Experimento 1. Nerja									52,4692				
NCC	61	825	2	129	43	15,64%	25,00%	39,0787	1,95	1,92	1,82	1,70	
NCCS	114	110	2	56	4	50,91%	6,67%	38,7985	1,49	1,24	1,01	0,70	
MI	62	825	1	158	35	19,15%	18,13%	39,1111	2,77	2,72	2,65	2,52	
MIS	114	110	3	87	6	79,09%	6,45%	39,0825	1,78	1,49	1,29	1,20	
HA-LK	7	825	4	779	33	94,42%	4,06%	38,5532	0,40	0,33	0,26	0,18	

ME-LK	14	1715	4	1616	69	94,23%	4,09%	38,5573	0,40	0,34	0,26	0,19
SIFT	9	5864	1	243	309	4,14%	55,98%	38,7607	2,74	2,71	2,68	2,56
SURF	5	5249	2	226	175	4,31%	43,64%	39,0955	2,93	2,87	2,78	2,62
Experimento 2. Rincón de la Victoria								57,1175				
NCC	62	1376	4	144	118	10,47%	45,04%	45,2906	5,92	4,92	4,23	3,72
NCCS	884	220	4	137	16	62,27%	10,46%	44,1909	5,93	3,83	2,99	1,95
MI	60	1376	3	123	62	8,94%	33,51%	45,4138	5,79	5,07	4,45	4,07
MIS	849	220	4	154	2	70,00%	1,28%	44,4016	5,86	4,36	3,48	2,95
HA-LK	26	1376	4	1272	3	92,44%	0,24%	44,7092	4,22	2,64	1,99	1,41
ME-LK	58	3000	4	2808	6	93,60%	0,21%	44,6851	4,39	2,72	2,09	1,46
SIFT	15	9566	2	86	166	0,90%	65,87%	46,0175	5,28	3,92	3,29	2,58
SURF	17	9346	1	152	250	1,63%	62,19%	46,8821	5,93	5,17	4,32	3,96
Experimento 3. Huelva								62,4385				
NCC	53	919	1	151	140	16,43%	48,11%	49,7502	4,36	4,18	4,09	3,83
NCCS	212	197	4	137	30	69,54%	17,96%	48,8091	2,39	1,94	1,73	1,35
MI	47	919	1	152	156	16,54%	50,65%	49,3748	3,96	3,92	3,82	3,54
MIS	190	197	4	145	18	73,60%	11,04%	49,1088	2,01	1,59	1,48	1,03
HA-LK	15	919	4	783	49	85,20%	5,89%	48,8373	1,19	0,96	0,87	0,53
ME-LK	38	2410	4	2092	133	86,80%	5,98%	48,8607	1,28	1,04	0,92	0,59
SIFT	13	8298	1	44	100	0,53%	69,44%	49,8093	3,57	2,90	2,15	1,52
SURF	10	6323	1	38	52	0,60%	57,78%	49,8355	3,98	3,81	3,32	15,00

Tabla 4-5. Datos comparativos de las técnicas de detección geométrica.

En el primer caso, en el que usamos las imágenes de Nerja, la técnica que mejor resultado presenta es la de tracking. Le siguen SIFT y la búsqueda por correlación cruzada normalizada (NCCS). En las imágenes de Rincón de la Victoria, las mejores técnicas son NCCS y MIS, siguiéndole el tracking. En Huelva, de nuevo el mejor resultado es obtenido por NCCS y tracking, en este caso, el primero ligeramente mejor. Podemos afirmar que en función del error, las técnicas NCC, MI, SIFT y SURF son las menos precisas en los tres casos, aunque SIFT en el caso del par de Nerja (las que menos cambios ha sufrido), queda entre las cuatro mejores. Las más precisas son NCCS, MIS y tracking.

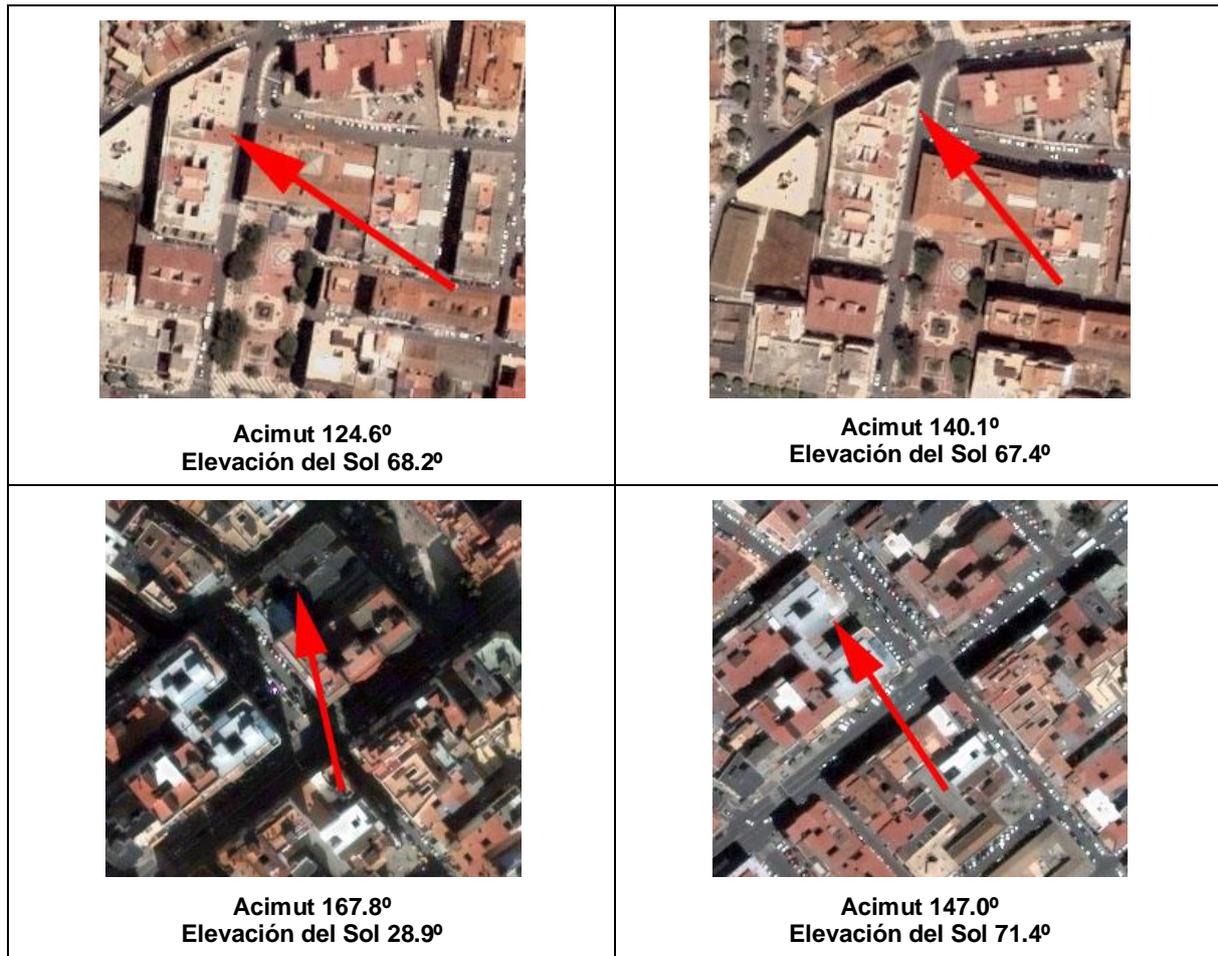


Figura 4-29. Diferencias en la posición del Sol en R. Victoria (arriba) y Huelva (abajo).

A diferencia de lo esperado, NCCS da mejor resultado que MIS para la medida de error que estamos empleando en los tres experimentos. Con esto no podemos concluir que la correlación cruzada sea mejor que la correlación de entropía (ECC) en cualquier caso. El peor resultado se debe a cambios por diferencias en el ángulo de adquisición en las tres imágenes. Al ser la información mutua independiente del color, a la hora de registrar, se intentan alinear correctamente estos cambios, mientras que medidas como la correlación cruzada no los tiene en cuenta. Se observa también que usar una ventana de búsqueda para la correlación cruzada e información mutua reduce el error de registro al aumentar la precisión en la localización de las

correspondencias. Debido a que el buen resultado de NCCS en los tres casos se debe a que hemos buscado el valor de correlación mínima que obtenía el mínimo error y esto no es viable realizarlo en cada estudio de detección de cambios, podemos concluir que la técnica de tracking es la que mejor resultados ofrece generalmente.

Si estudiamos el Error RMS Total (4.4), podemos observar que en los tres casos la técnica de tracking proporciona ajustes polinomiales con un error inferior al resto. Esto es así de forma natural, pues el tracking no consiste en emparejar características, sino en intentar obtener el desplazamiento sufrido por cada punto seleccionado en una imagen (véase más atrás en apartado 3.3) incluso cuando las correspondencias establecidas no están bien localizadas. Como todos los puntos en una región tienden a experimentar un desplazamiento similar, el resultado suele ser adecuado para una transformación con un bajo error. Por eso siempre es conveniente comprobar además del error RMS, si la localización de los emparejamientos es adecuada. LKT funciona mejor cuanto menor es el desplazamiento entre los píxeles de las imágenes maestra y esclava. En algunas ocasiones, para que funcione correctamente será necesario un ajuste RST (rotación, escalado y traslación) inicial. En zonas montañosas suele ser también más difícil que el tracking funcione correctamente, pues además de las deformaciones debidas a los distintos puntos de vista en la adquisición, encontramos las deformaciones debidas al terreno. Cuando las distancias entre los pares de correspondencias son elevadas, es recomendable usar más niveles en la pirámide empleada (se puede establecer como parámetro) y un mayor tamaño de la ventana de seguimiento.

En conclusión, siempre que sea posible se aconseja el uso de la técnica de tracking, por proporcionar un abundante número de pares y por rapidez en su ejecución, aunque en casos como los comentados (gran desplazamiento entre las imágenes o zonas muy accidentadas) podrá ser necesario un ajuste inicial o recurrir a otras técnicas. En los casos en el que el tracking falle, habría que recurrir a las técnicas de correlación e información mutua para buscar correspondencias basadas en área. En cada caso el analista deberá llegar a un equilibrio entre la precisión y velocidad a alcanzar en la corrección.

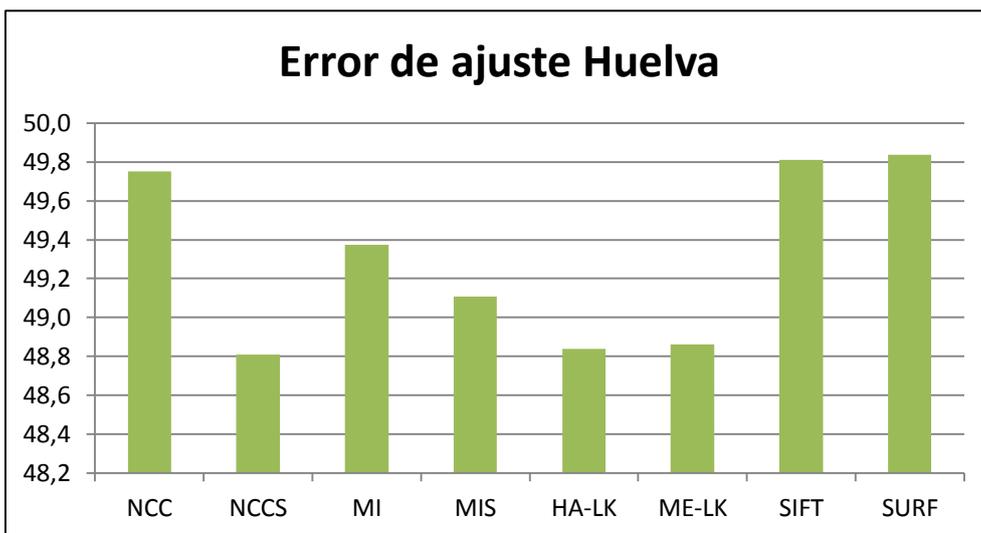
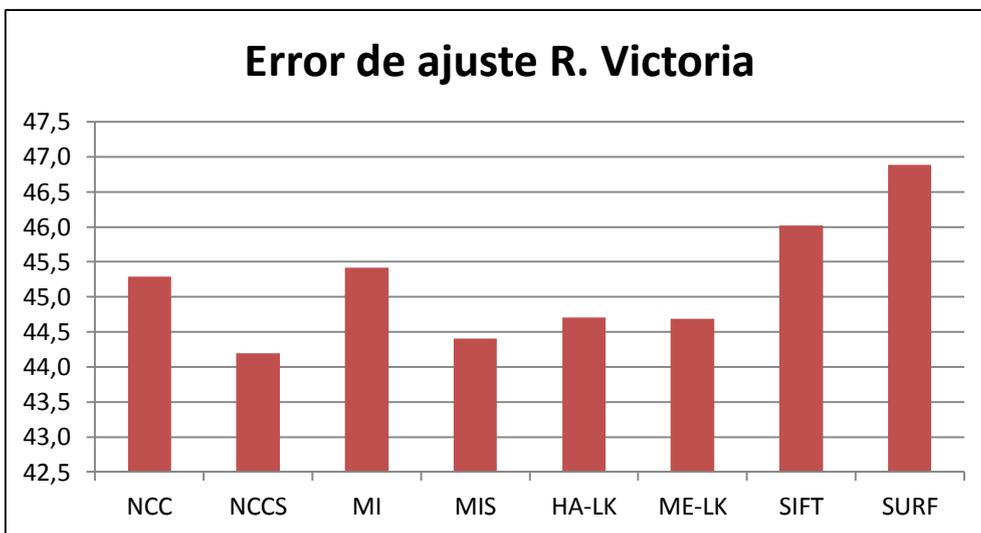
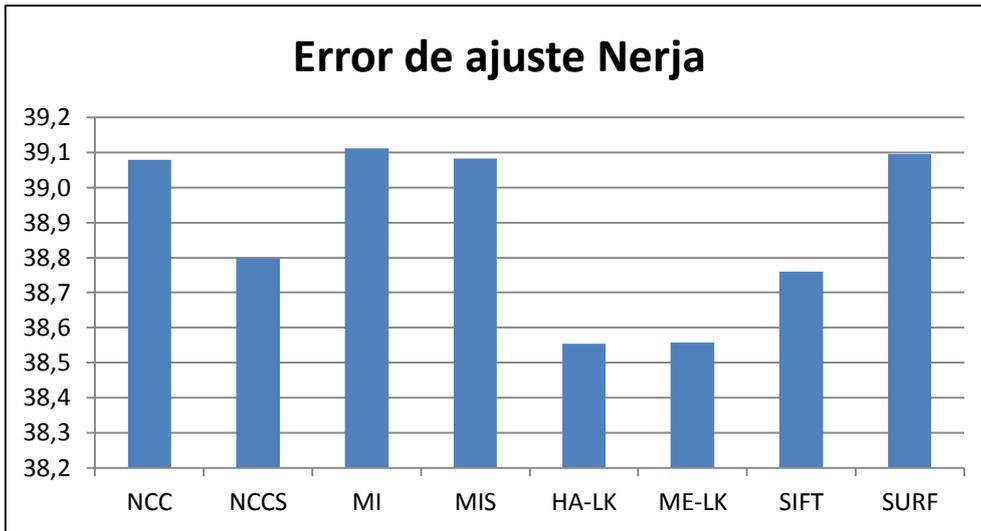


Figura 4-30. Error cuadrático medio de los niveles de gris tras la corrección geométrica.

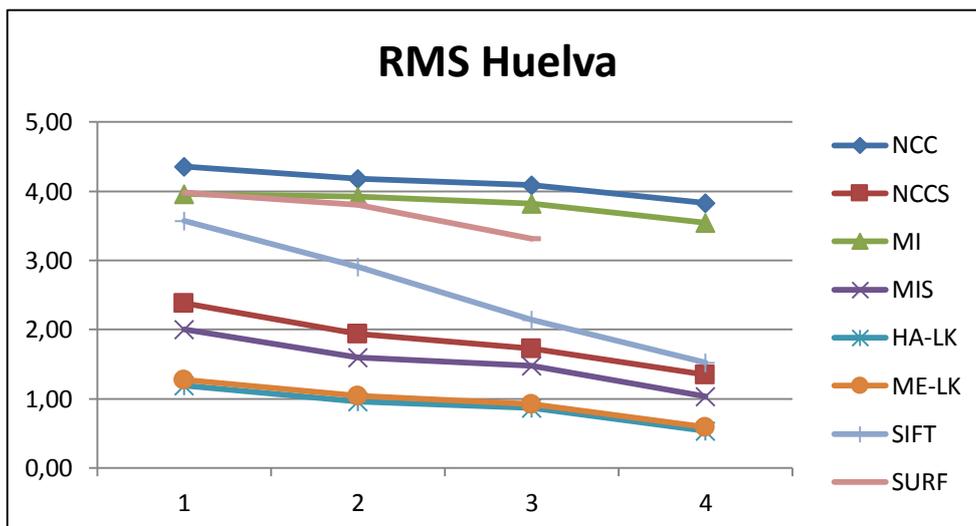
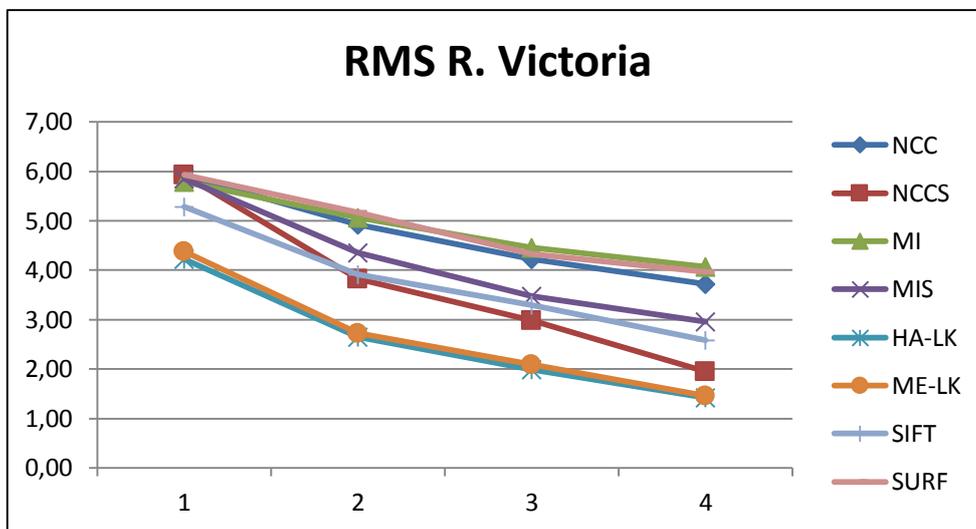
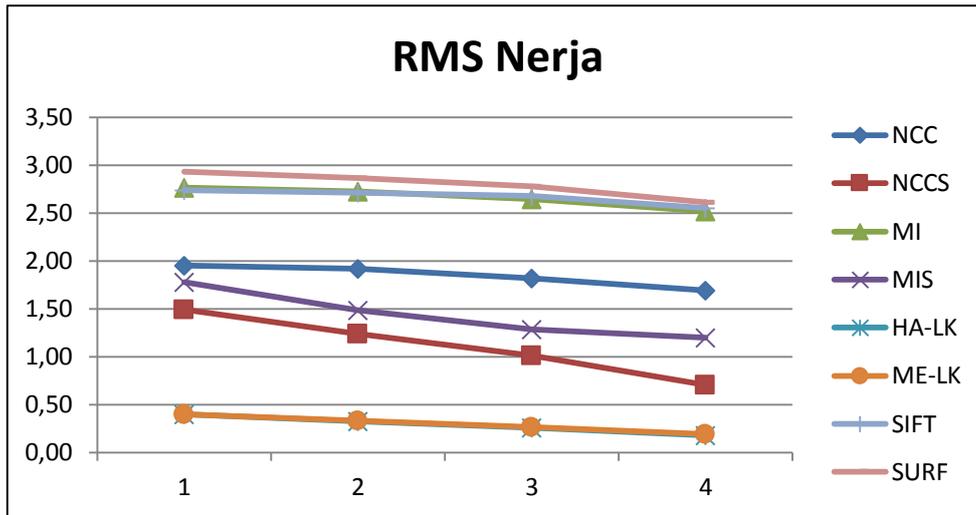


Figura 4-31. Error RMS ajuste polinomial hasta orden 4.

Técnica	Ventajas	Inconvenientes
Detección		
Harris	<ul style="list-style-type: none"> • Muy rápido en la detección. • Permite obtener un conjunto de puntos repartidos homogéneamente. 	<ul style="list-style-type: none"> • Baja repetibilidad debida a los frecuentes cambios que afectan a las imágenes de satélite.
Autovalores mínimos	<ul style="list-style-type: none"> • Muy rápido en la detección. • Permite obtener un conjunto de puntos repartidos homogéneamente. 	<ul style="list-style-type: none"> • Baja repetibilidad debida a los frecuentes cambios que afectan a las imágenes de satélite.
SIFT	<ul style="list-style-type: none"> • Invariantes a escala y rotación. 	<ul style="list-style-type: none"> • Más lentos de obtener que Harris y ME.
SURF	<ul style="list-style-type: none"> • Invariantes a escala y rotación. 	<ul style="list-style-type: none"> • Más lentos de obtener que Harris y ME.
Matching		
Correlación cruzada normalizada	<ul style="list-style-type: none"> • Más precisión en la localización de las correspondencias. • Adecuado cuando las otras técnicas no ofrecen buenos resultados. 	<ul style="list-style-type: none"> • Muy lento de obtener si se usan ventanas de búsqueda. • Requiere corrección radiométrica. • Depende de la repetibilidad de los puntos usados en la versión sin búsqueda.
Mutual Information	<ul style="list-style-type: none"> • Más precisión en la localización de las correspondencias. • Invariancia al color. • Adecuado cuando las otras técnicas no ofrecen buenos resultados. 	<ul style="list-style-type: none"> • Muy lento de obtener si se usan ventanas de búsqueda. • Depende de la repetibilidad de los puntos usados en la versión sin búsqueda. • Sensible a cambios por diferencias en el ángulo de adquisición.
Lucas-Kanade Tracking	<ul style="list-style-type: none"> • Es muy rápido en su ejecución y es el que mejor resultados ofrece en condiciones normales. 	<ul style="list-style-type: none"> • Puede dar lugar a <i>matches</i> mal localizados y, por tanto, no válidos para el registro.
SIFT	<ul style="list-style-type: none"> • Permite el registro de imágenes de distinta resolución espacial y rotación. 	<ul style="list-style-type: none"> • Muy sensible a los cambios, generalmente pocos <i>matches</i> son detectados.
SURF	<ul style="list-style-type: none"> • Permite el registro de imágenes de distinta resolución espacial y rotación. 	<ul style="list-style-type: none"> • Muy sensible a los cambios, generalmente pocos <i>matches</i> son detectados.

Tabla 4-6. Ventajas e inconvenientes de técnicas de detección y *matching* de puntos de control.

Se desaconseja recurrir al uso de características SIFT o SURF para el registro de imágenes de satélite, a no ser que éstas hayan sido adquiridas en mismas condiciones (estación del año, iluminación, ángulo de adquisición...) o sean necesarias para corregistrar imágenes con distinta orientación o escala. Generalmente, ofrecen pocas correspondencias y a veces erróneas, aunque esto puede reducirse con el uso de RANSAC y técnicas de filtrado. En caso de disponer

de imágenes aptas para estas técnicas, son las más rápidas de todas las estudiadas y pueden proporcionarnos un ajuste bueno y suficiente para la detección.

Siempre que el error en el registro no sea demasiado alto, cualquier técnica sería adecuada para una corrección geométrica global siempre que ofrezca suficientes correspondencias y éstas estén distribuidas homogéneamente por toda la imagen, dejando para la fase de detección de cambios precisar la localización de los mismos con un ajuste geométrico local como veremos en el apartado 6.4.

4.6 Futuras mejoras

Existe una variante del algoritmo SURF que reduce el tiempo de cómputo de los puntos característicos, eliminando el cálculo para la invariancia a rotación que, generalmente, no es necesaria cuando se realiza detección de cambios, pues las imágenes de satélite suelen encontrarse orientadas en dirección al norte. Este algoritmo se conoce por U-SURF y es robusto a rotaciones de $\pm 15^\circ$.

Se podrían añadir otros algoritmos de detección de características como ORB (Rublee, Rabaud, Konolige, & Bradski, 2011), STAR (Agrawal, Konolige, & Blas, 2008) o FAST (Rosten & Drummond, 2006) ya incluidos en las últimas versiones de OpenCV, e incluso probar variantes de SIFT como PCA-SIFT (Ke & Sukthankar, 2004) y GLOH (Mikolajczyk & Schmid, 2005) para evaluar si se obtienen mejoras.

Técnicas como la correlación o información mutua con ventana de búsqueda son las que más precisión pueden ofrecer a la hora de localizar las correspondencias en la imagen esclava, incluso con distintas orientaciones. No obstante, esta búsqueda es excesivamente lenta. Como mejora se podrían aplicar técnicas de optimización para acelerarla.

A fin de mejorar los resultados de NCC y MI, se podrían aplicar en la imagen esclava valores menos restrictivos para la mínima correlación cruzada y correlación de entropía respectivamente, con el propósito de aumentar las probabilidades de que todo punto candidato en la imagen maestra puede encontrar su correspondencia.

En las técnicas NCCS y MIS, para reducir el tiempo de ejecución se puede pasar como parámetro un desplazamiento inicial. En futuras versiones se prepararán las funciones para pasarles un fichero de correspondencias iniciales a partir de la cual se pueda calcular esta transformación inicial.

Capítulo 5

Extracción de sombras y otras capas de información

El suelo terrestre se encuentra en constante cambio y éstos pueden ser de muy distinta naturaleza. Antes de afrontar un trabajo de detección de cambios, es necesario definir bien cuáles de ellos nos serán de utilidad para el estudio que estemos realizando. Si tomamos dos imágenes de una misma zona y nos basamos tan sólo en un cambio de color para determinar que un elemento de la superficie terrestre ha cambiado, prácticamente tendríamos que concluir que un porcentaje muy alto de los píxeles han cambiado de un instante a otro incluso después de haberlas corregido radiométrica y geométricamente.

Por este motivo, antes del proceso de detección automática de cambios, es necesario obtener información útil de las imágenes que permita, por un lado, una detección de cambios más rápida y precisa, y, por otro, disponer de más información para la posterior clasificación de los cambios.

Hay ciertos fenómenos que una herramienta automática de detección podría determinar como cambios pero que, en general, no son de interés para el usuario final de las imágenes. Un ejemplo de esto son las sombras y, habitualmente, todas las imágenes aéreas se ven afectadas por ellas. La aparición de sombras en una imagen depende de ángulo de incidencia de los rayos solares y de los elementos de la superficie terrestre iluminados, tales como edificios, árboles, montañas, etc. Una sombra, no supone un cambio en el terreno, de ahí que sea interesante su localización para excluirlas de la detección de cambios. Incluso hay trabajos donde se propone restaurar el brillo de las mismas (Sarabandi, Yamazaki, Matsuoka, & Kiremidjian, 2004) para identificar posibles elementos ocultos.

Otro aspecto que dificulta la detección de cambios es la aparición de nubes. Las nubes ocultan áreas del terreno en las que no se podría detectar cambios, ya que la radiación visible (que es la captada para obtener las imágenes usadas en este proyecto) es reflejada completamente por las nubes impidiendo ver el suelo terrestre.

Una previa localización de nubes en las imágenes permitiría eliminar las zonas afectadas de la posterior detección de cambios.

Además de eliminar sombras y nubes, cualquier otra información extraíble de las imágenes puede ser útil. Todo dependerá del tipo de cambios que se está interesado en obtener. Por ejemplo, para una detección de cambios urbanos o incluso de construcciones en zonas no urbanizables, nos sería muy útil identificar zonas de vegetación para descartarlas de la detección.

Uno de los métodos clásicos de detección de cambios consiste en clasificar las imágenes según la naturaleza del suelo e identificar como cambio aquellos píxeles de la imagen que cambien de categoría. Este método se llama *Post-clasificación* y para ello es necesario obtener previamente toda la información sobre el suelo posible.

5.1 Detección de sombras

En este apartado, vamos a describir un procedimiento para detección de sombras en imágenes que nos permitirá crear máscaras que nos ayudarán a simplificar la posterior detección y clasificación de los cambios.

5.1.1 ¿Qué son las sombras?

Las sombras surgen cuando un objeto se interpone entre otro objeto de la escena y la fuente de iluminación. Los trabajos realizados hasta ahora clasifican las sombras en dos tipos: “self” y “cast”. Las primeras, se dan en la parte de los objetos de la escena que no son iluminadas directamente por la fuente de iluminación. Las segundas, se corresponden con el área proyectada por el objeto en la dirección de la luz. La siguiente Figura explica esta clasificación.

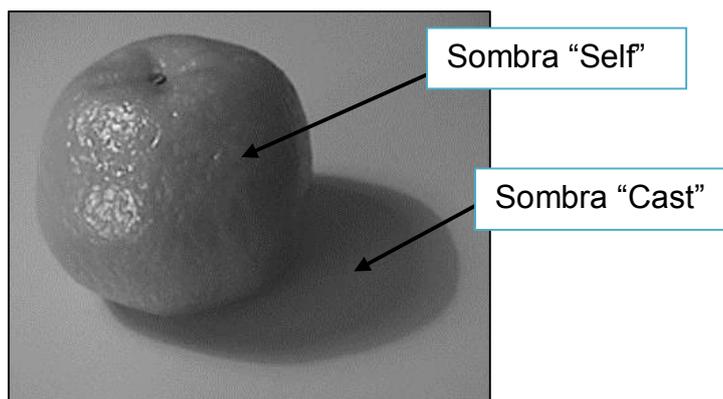


Figura 5-1. Tipos de sombra.

La mayoría de métodos propuestos en la literatura científica, diferencian entre estos dos tipos de sombras y las tienen en cuenta en el proceso de detección. Cuando trabajamos con imágenes aéreas, la diferencia entre estos dos tipos de sombras es inapreciable. Por este motivo, de aquí en adelante se utilizará el término “sombras” para referirse indistintamente a uno u otro tipo.

Las técnicas de detección de sombras se clasifican en: basadas en “modelos” o basadas en “propiedades”. Las primeras, usan modelos que representan un conocimiento a priori de la geometría de la escena, los objetos y la iluminación. Las técnicas basadas en “propiedades”, identifican las sombras usando características de la imagen tales como la geometría, brillo o el color de las mismas. De este segundo tipo, son las técnicas que se han implementado en este proyecto, una de ellas totalmente automática y otras más simples e interactivas que permiten ajustar las máscaras de sombras antes de crearlas.

5.1.2 Espacios de color con características invariantes

En trabajos anteriores, se pone de manifiesto la utilidad de la invariancia de algunos espacios de color ante ciertos fenómenos. Estos espacios de color fueron creados principalmente para el reconocimiento de objetos de color y la recuperación de imágenes desde una base de datos, y al usarlos, se pretende que el reconocimiento o la recuperación no se vean afectados por: cambios en el punto de vista, diferencias en la posición y orientación del objeto, la presencia de sombras, o por cambios en la dirección y color de la luz.

La invariancia a sombras hace que algunos de estos espacios de color sean de utilidad para la identificación de las mismas. Algunos de estos espacios de color son el rgb (RGB normalizado), el tono H y saturación S del espacio HSV, y el espacio de color $c_1c_2c_3$.

En el método que detallaremos más adelante, se usa el espacio $c_1c_2c_3$ para obtener las sombras candidatas, concretamente la componente c_3 debido a su alta respuesta a las sombras como se puede apreciar en la Figura 5-2.

El espacio de color $c_1c_2c_3$ fue creado por (Gevers & Smeulders, 1999) y se trata de una transformación no lineal del espacio RGB. Las expresiones para su cálculo son:

$$c_1 = \arctan\left(\frac{R}{\max\{G, B\}}\right)$$

$$c_2 = \arctan\left(\frac{G}{\max\{R, B\}}\right)$$

$$c_3 = \arctan\left(\frac{B}{\max\{R, G\}}\right)$$

Además de presentar una alta respuesta a las sombras, la componente c_3 conlleva algunos problemas que se han de tener en cuenta: 1) es una componente afectada por ruido que influye por tanto en la clasificación de algunos píxeles y dificulta la obtención precisa de los bordes de las áreas de sombras; 2) su ecuación se vuelve inestable para valores bajos de saturación; y 3) obtiene un alta respuesta para valores altos de azul, como las piscinas en las imágenes de satélite (véase Figura 5.2).



Figura 5-2. Componente C3 de una imagen.

Para reducir el efecto del ruido, la técnica que veremos emplea una versión suavizada de la componente c_3 y la imagen gradiente para determinar los bordes de las áreas de sombras. Con el fin de evitar problemas de clasificación debidos a baja saturación, se hace uso de las componentes S y V del espacio de color HSV.

5.1.3 Método de detección de sombras

En este apartado trataremos el método de detección de sombras propuesto en (Arévalo, González, & Ambrosio, 2008) que detecta sombras en una imagen RGB mediante una técnica de crecimiento de regiones sobre la componente c_3 suavizada. El crecimiento de regiones consiste en localizar un conjunto de píxeles o semillas a las que, posteriormente y de forma recursiva, se van añadiendo aquellos píxeles vecinos que cumplan un conjunto de restricciones. Para que el crecimiento de regiones sea adecuado es necesario un algoritmo fiable que localice al menos una semilla en cada sombra y ninguna fuera de las áreas de sombras.

El primer paso de este algoritmo consiste en obtener las imágenes necesarias para todo el procedimiento. Por un lado, se debe obtener la componente c_3 de la imagen a la que posteriormente se aplica un filtro promedio con una ventana 3x3 para suavizarla. Por otro, se obtiene la imagen HSV de la que extraemos las componentes S y V. A la imagen V además le aplicamos un filtro de Sobel con una ventana 3x3 para obtener la imagen gradiente, donde son fácilmente identificables los bordes.

Una vez obtenidas las imágenes necesarias se procede a la localización de las semillas. En este método, una semilla es una ventana 5x5 píxeles y es situada en aquellos píxeles de la imagen c_3 suavizada que cumple las siguientes restricciones:

- El centro de la ventana es un máximo local en la imagen c_3 suavizada y todos los píxeles de la ventana son mayores que la media de todos los píxeles de la imagen c_3 .

- La media de los píxeles de la semilla en la componente V debe ser menor que un umbral T_V (cumpliendo la hipótesis de que las sombras son oscuras y tiene poca intensidad). Por otro lado, la media en la componente de saturación S debe ser mayor que un umbral T_S para evitar la inestabilidad de c_3 para colores de baja saturación.
- Ninguno de los píxeles de la semilla pertenecen ya a otra semilla previa.

El tamaño de la semilla (5x5) determina el tamaño mínimo de sombra que puede ser detectado por este método, que es de unos 3x3 m² en el caso de imágenes Quickbird (con una resolución de 0.6 m por píxel), o 5x5 m² para una imagen Ikonos (con una resolución de 1 m por píxel). Los autores proponen un valor para los umbrales $T_V = 0.35$ y $T_S = 0.02$, normalizados en el rango [0..1], tras el estudio de un conjunto de imágenes Quickbird. En este proceso, no supone un problema localizar varias semillas en una misma sombra ya que en el crecimiento de regiones terminarán uniéndose en una sola. La siguiente Figura muestra la localización de las semillas en la imagen RGB de la Figura 5-2:



Figura 5-3. Semillas detectadas en sombras.

Podemos ver como en este caso hay una sombra donde no se ha colocado una semilla. Subiendo el umbral T_V muy levemente, conseguimos situar al menos una semilla en la sombra mencionada, pero si se eleva demasiado se corre el peligro de situar semillas en zonas que no corresponden a sombras, dando lugar a falsos positivos.

Una vez localizadas las semillas, se inicia el proceso de crecimiento de regiones. Es un proceso recursivo que se ejecuta para cada uno de los ocho vecinos de cada píxel del borde de una región en crecimiento. Empezando por las semillas, un píxel vecino es clasificado como sombra y añadida a la región si satisface todas las siguientes condiciones:

- No ha sido añadido ya a alguna región.

- La magnitud del gradiente de V está por debajo de un umbral $T_E = 0.30$, o en otras palabras, no es un píxel del borde.
- Está a una distancia de Mahalanobis menor que d_0 de la media \bar{c}_3 de la región, es decir:

$$\frac{|c_3 - \bar{c}_3|}{\sigma} < d_0 \quad (5.1)$$

Valores altos de σ indican una alta dispersión de c_3 dentro de la región, y por tanto se permite una mayor desviación desde la media \bar{c}_3 . Los autores del método proponen un valor de $d_0 = 3$.

- Y además satisface las condiciones impuestas para las semillas sobre el valor de S (que debe superar el valor T_S) y V (que debe ser inferior a T_V).

El proceso finaliza cuando ninguno de los píxeles vecinos ha sido añadido a la región. Después del crecimiento de regiones, pueden quedar pequeños huecos en las sombras debido a píxeles mal clasificados. Para corregir esto, se aplica un cierre morfológico sobre la imagen binaria de sombras. Esta operación consiste en una dilatación de la imagen seguida de una operación de erosión y rellenará los huecos de hasta dos píxeles conservando los bordes originales. Como elemento estructurante para estas operaciones se usa el kernel siguiente:

$$k = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

El resultado de la detección de sombras se muestra en la Figura 5-4 y el diagrama de todo el proceso en la Figura 5-5.

El método propuesto tiene el inconveniente de que no podemos estar seguros de que los parámetros seleccionados han sido los adecuados para una imagen hasta el final del procedimiento. Y aunque los parámetros por defecto propuesto por (Arévalo, González, & Ambrosio, 2008) son los que mejores resultados ofrecen generalmente (al menos para imágenes Quickbird), a veces, dependiendo de la naturaleza de las imágenes, pueden quedar sombras que no han sido seleccionadas al no situar algunas semillas en ellas.



Figura 5-4. Sombras detectadas en una imagen RGB.

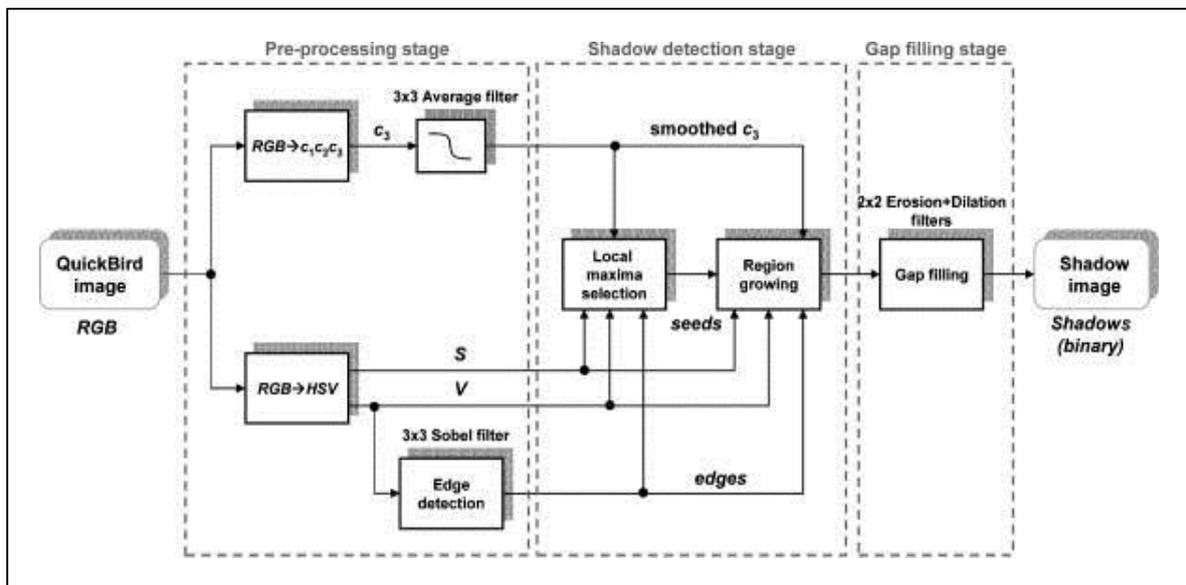


Figura 5-5. Proceso de detección de sombras.

Cuando esto ocurre y por la dificultad de afinar los parámetros del algoritmo, podemos recurrir a algunos de los dos métodos más sencillos e interactivos que comentaremos a continuación. Aunque son menos precisos, pueden ayudarnos a simplificar el proceso de detección de sombras y a obtenerlas de una forma mucho más rápida.

En ambos, el valor de los parámetros se selecciona con controles deslizantes y se observa el resultado en una parte de la imagen. Cuando los valores seleccionados son adecuados, se procede a la generación de la máscara de sombras.

En el primer caso, se deben establecer dos parámetros: un umbral para la componente c_3 y otro para el brillo. Se seleccionará como sombra todo píxel que supere el valor seleccionado para c_3 . Esto último también seleccionará píxeles de color

azul intenso (como las piscinas), por este motivo, también debe cumplir que su brillo no supere el segundo de los umbrales seleccionados.

En la segunda opción, se sigue la técnica propuesta por (Tsai V. J., 2006) donde se hace uso del espacio de color HSI o el espacio YIQ, comentados en la sección 5.5. Como en el primer caso, se debe seleccionar un valor umbral que se aplica a la imagen sintética obtenida a partir de la expresión:

$$\frac{H + 1}{I + 1} \quad (5.2)$$

para la imagen en el espacio HSI o a partir de la imagen obtenida por:

$$\frac{Q + 1}{Y + 1} \quad (5.3)$$

para la imagen en el espacio de color YIQ. Ambas imágenes presentan una alta respuesta a las sombras como se puede apreciar en la Figura 5-6.



Figura 5-6. Sombras usando el espacio de color HSI (izquierda) e YIQ (derecha).

5.2 Restauración de sombras

Con este procedimiento se pretende restaurar el brillo de las zonas afectadas por sombras, eliminando falsos positivos (los provocados por la propia sombra), así como facilitando la detección de posibles cambios ocultos por las mismas.



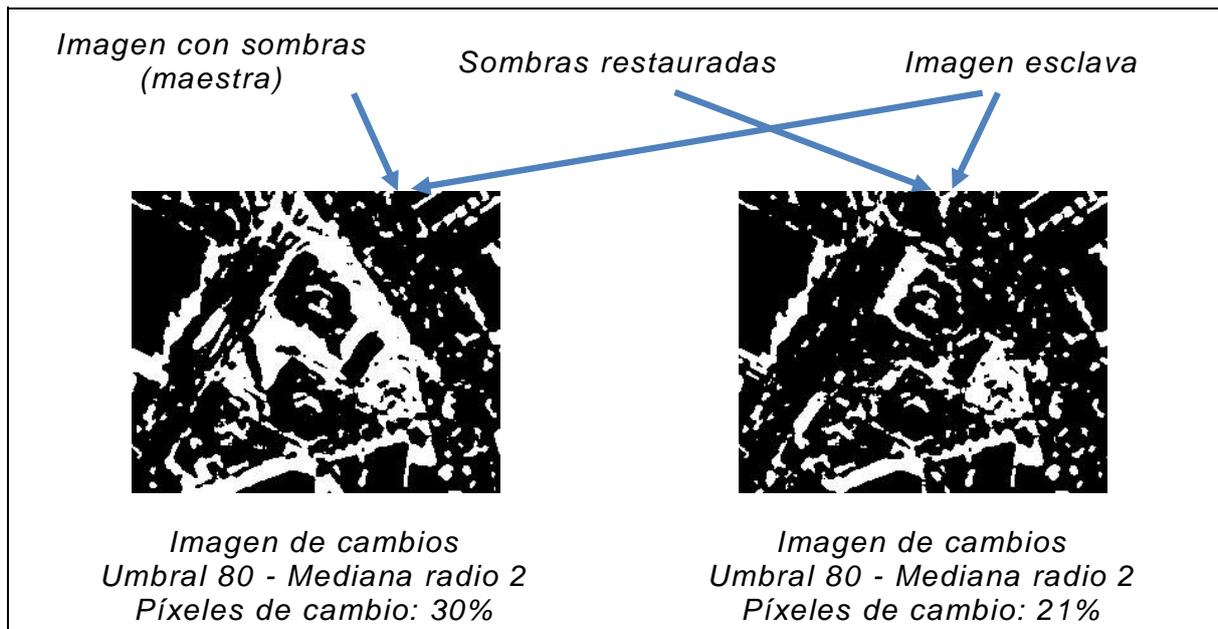


Figura 5-7. Uso de restauración de sombras I.

El procedimiento de restauración de sombras implementado en este proyecto se basa en el método propuesto en (Sarabandi, Yamazaki, Matsuoka, & Kiremidjian, 2004). Para que el brillo de la zona afectada por una sombra pueda ser restaurado, debe existir otra imagen en la que dicha zona carezca de sombras con el fin de extraer su histograma y su información de brillo.

Sean las imágenes IM1 (imagen con sombras), e IM2 (imagen de la misma zona sin sombras), corregistradas y normalizadas radiométricamente, el procedimiento es el siguiente:

1. Comprobar que las dos imágenes tienen mismas dimensiones y tienen misma resolución radiométrica.
2. Crear una máscara que contenga los píxeles de sombra en la imagen IM1 que corresponden con píxeles sin sombra en la imagen IM2.
3. Obtener histograma de la imagen IM2 usando la máscara generada en el paso 2.
4. Generar una imagen IM3 a partir de IM1 (enmascarada) mediante especificación de histograma, usando como referencia el histograma calculado en el paso anterior.
5. Se fusionan las imágenes IM3 e IM1 para obtener la imagen con sombras restauradas. Para ello, se seleccionan de IM3 todos los píxeles enmascarados y de IM1 el resto de píxeles.

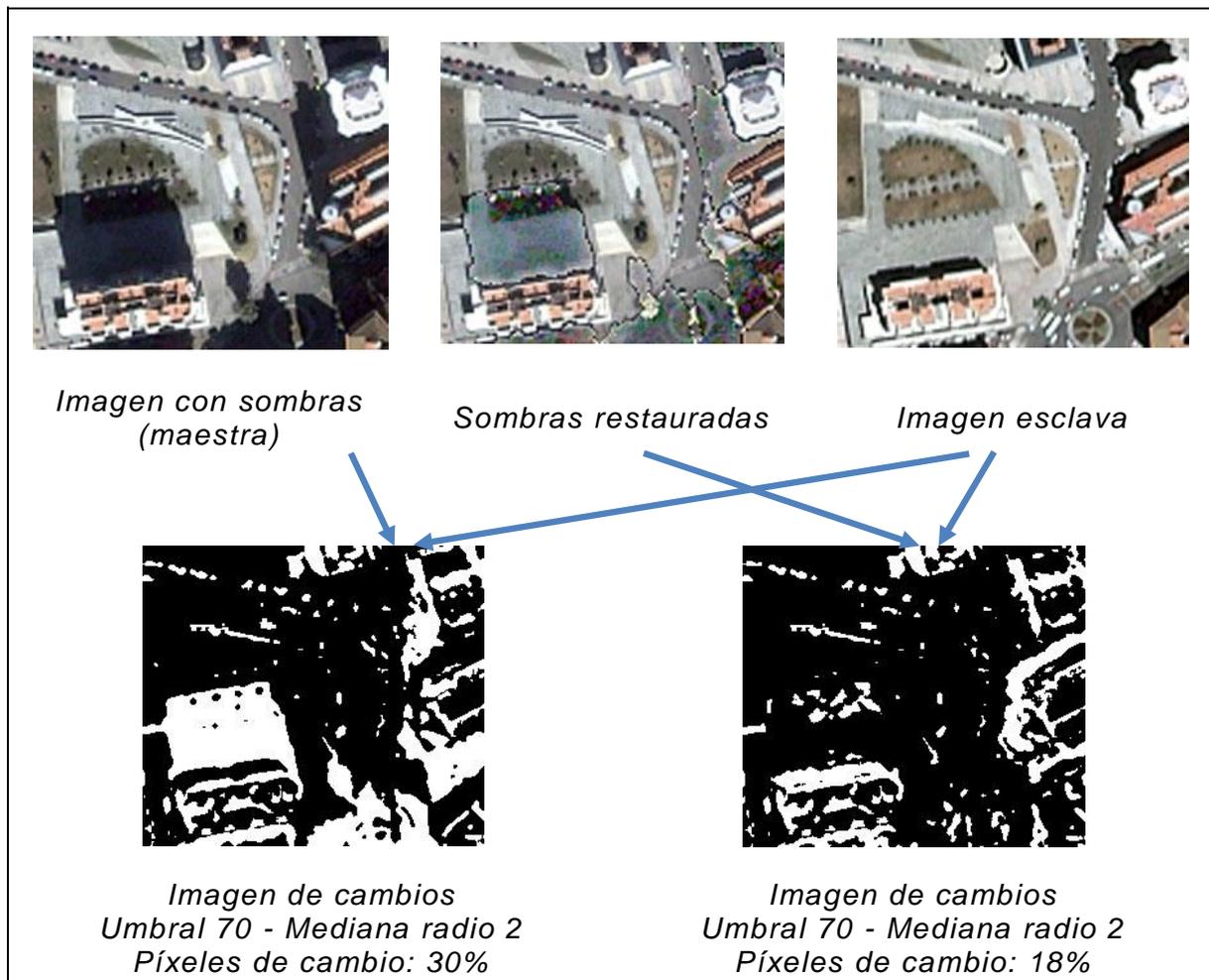


Figura 5-8. Uso de restauración de sombras II.

En las ilustraciones 5-7 y 5-8 se pueden apreciar los cambios generados antes y después de usar el procedimiento de restauración de sombras en dos pares de fragmentos de imagen. En el ejemplo de la Figura 5-7, al restaurar el brillo en las sombras de la imagen maestra (con un porcentaje mayor de sombras que la imagen esclava), el número de píxeles detectado como cambio (tras realizar la diferencia de niveles de gris, aplicar un umbral y un filtro mediana para eliminar cambios muy pequeños), se reduce en un 9%. En el ejemplo de la Figura 5-8, el porcentaje de píxeles de cambios se reduce en un 12%.

5.3 Detección de nubes

La presencia de nubes en las imágenes es otra de las causas de cambios que, generalmente, no son de interés. Aunque las nubes no es uno de los problemas habituales que podemos encontrar en imágenes de satélite, puesto que la mayor parte de los productos ofrecidos por las operadoras de satélite siguen unos criterios de calidad, veremos aquí algunos métodos para su detección para hacer más completa esta memoria.

Como estudiamos en el apartado 2.5, donde se explican la dispersión y absorción de la radiación, el vapor de agua de las nubes dispersa toda la radiación visible (lo que da su color blanco) y tiene una alta absorción en el infrarrojo térmico y parte del infrarrojo medio. Basándonos en estas propiedades, la detección de nubes se convierte en algo simple si estudiamos estas bandas del espectro electromagnético, tal como hacen los satélites meteorológicos. Pero las imágenes más usadas en la detección de cambios no disponen de esta información. Suelen usarse imágenes con tres bandas en el espectro visible (RGB) y, en algunos casos, otra banda en el infrarrojo cercano. Como ya hemos visto, toda esta radiación es reflejada por las nubes y, por tanto, son capturadas en las imágenes con altos valores en los niveles digitales de todas las bandas. Podríamos pensar que para detectarlas y eliminarlas del proceso de detección, basta con aplicar un umbral alto y seleccionar aquellos píxeles cuyo nivel digital supere ese valor umbral. Pero en ese caso, también estaríamos seleccionando otros elementos de la superficie terrestre con una alta reflectividad.

Generalmente, la presencia de nubes lleva consigo la aparición de las sombras que generan al interceptar la radiación solar. En (Wang, Huang, & Rosenfeld, 1991) se propone un método de detección de nubes que tiene en cuenta esta característica y se basa en los siguientes aspectos:

- 1) Las sombras son proyecciones paralelas de las nubes sobre el terreno.
- 2) Las nubes son objetos tridimensionales y aunque su forma generalmente no es la misma que la de su sombra, si la nube es compacta suelen ser muy aproximadas y en cualquier caso, sus centroides coinciden (en la línea que sigue la radiación solar).
- 3) Podemos suponer que en pequeñas áreas todas las nubes tienen la misma altura. Por tanto, la línea que une los centroides de dos nubes es paralela a la línea que une los centroides de sus sombras, aunque de distinta longitud ya que se encuentran a distinta altura.

El algoritmo consiste en unir cada par de nubes candidatas (seleccionadas aplicando un valor umbral por ejemplo) y cada par de sombras con una línea. Posteriormente encontrar pares de estas líneas que sean aproximadamente paralelas:

$$\frac{|S_1 - S_2|}{|S_1| + |S_2|} \leq 0.065$$

donde S_1 y S_2 son las pendientes de las líneas. El cociente de sus longitudes debe estar comprendido en un intervalo razonable:

$$0.75 \leq \frac{L_S}{L_C} \leq 1.05$$

donde L_s y L_C son las longitudes de las líneas que unen las sombras y las nubes respectivamente. Y por último el cociente de sus áreas debe también estar comprendido en un intervalo:

$$0.4 \leq \frac{s_1 + s_2}{c_1 + c_2} \leq 40$$

donde s_1 , s_2 , c_1 y c_2 son la áreas de las sombras y nubes candidatas.

El siguiente paso es filtrar los pares (nube, sombra) seleccionando las nubes candidatas de tres en tres. Si las líneas que las unen tienen un punto en común, también debe tenerlo las líneas de las sombras correspondientes. De todas las líneas, nos quedamos con aquellas en las que esto se cumple al menos para el 25% de todas las combinaciones de tres posibles. Para cada conjunto de tres, basta con mantener dos de los segmentos que las unen y se puede descartar el tercero.

A continuación, se trata de encontrar la función lineal que asigna a cada nube su sombra. Por cada par de nubes y sus correspondientes sombras, se pueden encontrar los parámetros para esta función lineal, que consiste en la línea que une el segmento entre los centroides de las nubes y el segmento entre los centroides de las sombras. Para determinar qué función lineal es la más adecuada de todas las posibles se sigue el siguiente procedimiento:

- 1) Para cada nube candidata c , se construye la sombra $s(c)$ predicha por la función lineal seleccionada.
- 2) Si $s(c)$ solapa cualquier otra nube candidata, se resta ésta de la sombra, llamando al resultado $s'(c)$.
- 3) Se interseca $s'(c)$ con la unión de todas las sombras candidatas en la imagen y se llama al resultado $s''(c)$.
- 4) Se define la puntuación de c como $A(s''(c))/A(s'(c))$ donde $A(x)$ es la función que devuelve el área de x .
- 5) Se define la puntuación de la función lineal escogida como la media de las puntuaciones asignadas a todas las nubes candidatas.

La mejor función será aquella que tenga la mayor puntuación asignada. Serán seleccionadas como nubes, las nubes candidatas c que tengan una puntuación mayor de 0.5 para la función seleccionada. La sombra predicha $s(c)$ para cada nube detectada, debe intersecar al menos una sombra candidata. Se seleccionará como sombra probable para la nube c aquella cuya intersección sea mayor.

Ya que, como hemos comentado, el problema de las nubes no es muy común en las imágenes de satélite, se propone la implementación de este detector para futuras versiones, así como el estudio de los resultados del método para imágenes de muy alta resolución espacial como la de los satélites Worldview, Quickbird e Ikonos.

5.4 Índices de vegetación

La banda del infrarrojo cercano, disponible en varios sensores multiespectrales y aunque no necesaria para una detección de cambios fiable, puede ser muy útil a la hora de detectarlos y de su posterior clasificación. Esta radiación es altamente reflejada por la vegetación viva, que también absorbe la mayor parte de la radiación de la luz roja, lo que le da el color verde.

Esta característica hace que las diferencias espectrales sean muy notables entre vegetación y otros elementos de la superficie terrestre y ha hecho que se desarrollen lo que se conoce como *índices de vegetación*, que permiten detectar y comparar de forma muy eficaz áreas de vegetación. A continuación veremos los índices de vegetación más comunes.

El más simple es el cociente de las bandas del rojo e infrarrojo cercano:

$$RVI = \frac{NIR}{RED}$$

donde *NIR* es el valor digital del infrarrojo cercano y *RED* el de la banda del rojo. Este índice no es lineal, su distribución no es normal y es susceptible a errores de división por cero. El rango habitual para vegetación verde es [2..8].

El más utilizado es el índice de vegetación de diferencia normalizada *NDVI*:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

que reduce el problema de división por cero y con una respuesta lineal en el intervalo [-1..1]. El rango habitual para vegetación verde va de 0.2 a 0.8.

Los dos índices anteriores se ven afectados por señales procedentes del suelo o por influencia de la atmósfera. En un intento de reducir esto surge el índice de vegetación *EVI*:

$$EVI = 2.5 \frac{NIR - RED}{NIR + 6 * RED - 7.5 * BLUE + 1}$$

donde recordemos, *NIR* era el infrarrojo cercano, *RED* la banda del rojo, y *BLUE* es la banda del azul. Su valor oscila entre -1 y 1 y el valor común para la vegetación verde va de 0.2 a 0.8.

El último índice que veremos intenta mejorar el *NDVI* para hacerlo relativamente resistente a factores atmosféricos y se conoce como *ARVI*:

$$ARVI = \frac{NIR - (2 * RED - BLUE)}{NIR + (2 * RED - BLUE)}$$

Su valor se encuentra en el rango $[-1..1]$ y el valor habitual para vegetación verde de nuevo oscila entre 0.2 y 0.8.

En la Figura 5-9 podemos ver el índice NDVI de una imagen. Es fácil apreciar, como las zonas más brillantes corresponden a las zonas más verdes, y como las construcciones humanas como edificios y carreteras quedan muy oscuras. Esto permite diferenciar muy fácilmente zonas urbanas de las zonas con vegetación. Si obtenemos el NDVI de las dos imágenes en las que queremos detectar cambios, podemos aplicar la técnica de diferencia de imágenes que veremos en el siguiente capítulo para obtener cambios entre estos tipos de uso del suelo.

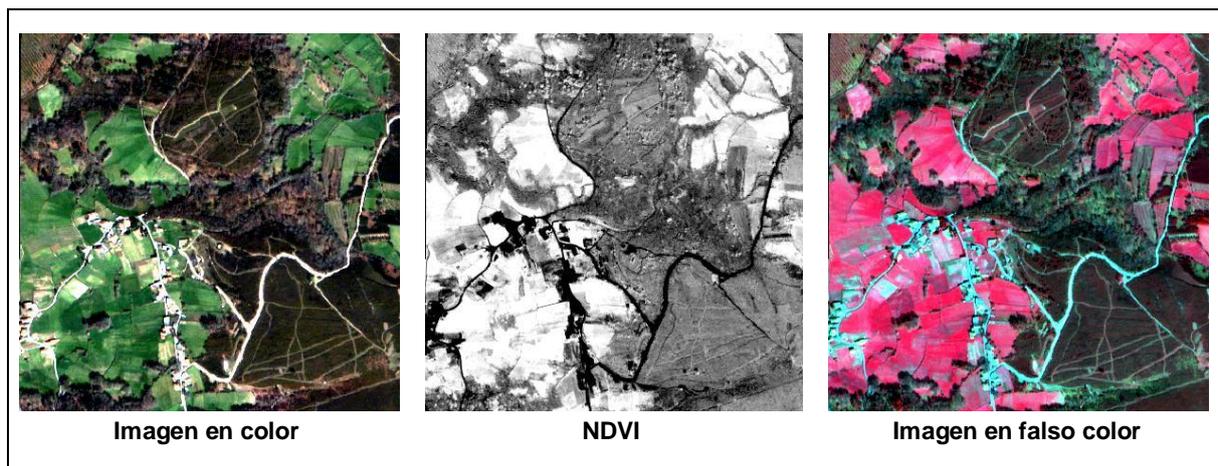


Figura 5-9. Imagen en color verdadero, en falso color y su índice NDVI.

La imagen en falso color es una forma de representar la información del infrarrojo cercano asignándolo a la banda del rojo, la información del rojo en la banda verde y la del verde en el azul. De esta forma obtenemos una imagen en que las zonas de vegetación viva quedan resaltadas con un color rojo muy vivo y que suele ser útil en procesos de detección de cambios por análisis visual.

Otra forma de obtener información sobre zonas de vegetación es usar transformaciones de las bandas como la Tasseled-Cap que estudiaremos en el siguiente capítulo.

5.5 Otras capas de información

No es raro encontrar, cuando trabajamos con imágenes, píxeles sin información. Como las imágenes digitales son cuadradas no queda más remedio que asignarle un valor constante a esos píxeles sin información, que suele coincidir con el valor digital mínimo o máximo posible. Antes de aplicar técnicas de detección habría que enmascarar estos píxeles para que no sean tenidos en cuenta, generando una máscara de fondo.

Para ello sería tan fácil como generar una imagen binaria que seleccione, aplicando un valor umbral, todos los píxeles con ese valor constante. ENVI dispone de una herramienta de generación de máscaras que nos permitirá seleccionar todos

los píxeles en un rango determinado. Por ejemplo, para un fondo negro, podemos crear una máscara que seleccione aquellos píxeles que tengan un valor en el rango [1..255] en cualquiera de las bandas para imágenes de tipo byte. Sólo quedarán fuera de la selección aquellos píxeles con un valor 0 en todas las bandas. Debemos tener en cuenta que los píxeles de fondo pueden ver modificado su valor por la corrección radiométrica y que si creamos las máscaras de píxeles de fondo antes de la corrección geométrica, éstas deben transformarse también.

Se pueden obtener otros tipos de información aplicando métodos de clasificación con el fin de extraer información sobre el uso que se hace del suelo. Puede ser de forma no supervisada, agrupando los píxeles de forma automática por similitud entre ellos, o de forma supervisada pasando un conjunto de píxeles de entrenamiento para cada una de las clases a obtener. Esta información nos será útil para aplicar el método de post-clasificación que selecciona como cambio aquellos píxeles que cambian de una clase a otra y que comentaremos en el capítulo siguiente.

Algunas técnicas de detección de cambios requieren aplicar transformaciones a las bandas de las imágenes como PCA o Tasseled-Cap. Las estudiaremos también en el capítulo dedicado a las técnicas de detección de cambios.

La mayoría de las técnicas de detección se basan en obtener una imagen de cambios en niveles de gris, a la que se aplica un valor umbral global o adaptativo para clasificar los píxeles como cambios o no-cambios. Generalmente, las imágenes base empleadas son en niveles de gris de los sensores pancromáticos o imágenes RGB (algunas con una cuarta banda del infrarrojo cercano) obtenidas de los sensores multiespectrales. En la mayoría de los casos, la resolución espacial de las imágenes multiespectrales ha sido incrementada por un proceso conocido como *pansharpening* que combina la imagen pancromática con la información del sensor multiespectral.

Puesto que es conocido que el espacio de color RGB no representa adecuadamente el sistema de visión humano, podría ser interesante estudiar otros espacios de color como HSV o CIELab para la detección de cambios. Comentaremos aquí los distintos espacios de color y en el apartado 6.3 del capítulo de detección de cambios, estudiaremos con un ejemplo la influencia de los espacios de color en los resultados de la detección.

El espacio de color RGB es el más ampliamente usado y representa los colores en cantidades de los tres colores primarios, rojo, verde y azul. La combinación de los tres colores da lugar al blanco. Si representamos cada componente con un byte, un color RGB vendrá dado por una tupla (R, G, B) donde R, G, y B, son valores comprendidos en el rango [0, 255]. Así, por ejemplo, el negro corresponde a (0, 0, 0), el blanco a (255, 255, 255) y un rojo puro a (255, 0, 0). El inconveniente de este espacio de color es que existe una alta correlación entre sus componentes, no es intuitivo psicológicamente y no es uniforme, pues la distancia entre colores RGB no corresponde con la distancia percibida por el ojo humano.

El espacio HSV surge para hacer la representación de los colores más intuitiva psicológicamente y lo hace en función de tres medidas, el tono o matiz, la saturación y el valor. El matiz corresponde al color, la saturación a la pureza del color y el valor al brillo, tres aspectos que son más intuitivos que las cantidades de colores primarios del espacio RGB. Se representa en forma de cono donde el matiz H es un ángulo con valor en $[0, 360)$ grados, la saturación S es la distancia al eje de brillo negro-blanco con valor en $[0, 1]$ y el valor V representa la altura en el eje negro-blanco con valor en $[0, 1]$. Se obtiene de espacio RGB con las ecuaciones (5.4) donde R, G y B son los valores normalizados en el rango $[0, 1]$ de rojo, verde y azul del espacio RGB.

$$\begin{aligned}
 V &= \max(R, G, B) \\
 S &= \begin{cases} \frac{V - \min(R, G, B)}{V} & \text{si } V \neq 0 \\ 0 & \text{si } V = 0 \end{cases} \\
 H &= \begin{cases} \frac{60(G - B)}{V - \min(R, G, B)} & \text{si } V = R \\ 120 + \frac{60(B - R)}{V - \min(R, G, B)} & \text{si } V = G \\ 240 + \frac{60(R - G)}{V - \min(R, G, B)} & \text{si } V = B \end{cases} \quad (5.4)
 \end{aligned}$$

Otros espacios de color usados en este proyecto son el HSI, YIQ y CIELab. El espacio de color HSI usado en uno de los procedimientos de detección de sombras es muy similar al espacio HSV y se obtiene del espacio de color RGB normalizado con las siguiente expresiones:

$$\begin{aligned}
 r &= \frac{R}{R + G + B}, g = \frac{G}{R + G + B}, b = \frac{B}{R + G + B} \\
 H &= \begin{cases} \cos^{-1} \left(\frac{0.5 \cdot [(r - g) + (r - b)]}{[(r - g)^2 + (r - b)(g - b)]^{\frac{1}{2}}} \right), & \text{si } b \leq g \\ 2\pi - \cos^{-1} \left(\frac{0.5 \cdot [(r - g) + (r - b)]}{[(r - g)^2 + (r - b)(g - b)]^{\frac{1}{2}}} \right), & \text{si } b > g \end{cases} \\
 S &= 1 - 3 \cdot \min(r, g, b) \\
 I &= \frac{R + G + B}{3}
 \end{aligned}$$

donde R, G, B son los valores RGB normalizados en el rango $[0, 1]$, la componente H toma valores en el rango $[0, 2\pi)$ radianes, y las componentes S e I toman valores en el rango $[0, 1]$.

También se ha probado la utilidad del espacio de color YIQ en la detección de sombras. Este se obtiene de las siguientes ecuaciones:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.595716 & -0.274453 & -0.321263 \\ 0.211456 & -0.522591 & 0.311135 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

El espacio CIELAB o Lab es otro espacio de color que estudiaremos para la detección de cambios y que fue creado para que sea perceptivamente más lineal, es decir, para que un cambio entre dos colores en este espacio sea percibido aproximadamente con la misma importancia por el ojo humano. Deriva del espacio conocido como CIEXYZ. Consta también de tres componentes: L es la luminosidad con valor en el rango $[0, 100]$, a y b definen el color, con valores en el rango $[-127, 127]$, el primero entre rojo y verde, y el segundo entre azul y amarillo. Se obtiene con las expresiones siguientes:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$L = 116f\left(\frac{Y}{Y_n}\right) - 16$$

$$a = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right]$$

$$b = 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right]$$

donde X_n , Y_n , y Z_n corresponden al valor CIEXYZ del blanco de referencia, generalmente el conocido como D65 correspondiente a la luz del Sol de mediodía en Europa occidental, con valores $(0.950456, 1.0, 1.088754)$ y

$$f(t) = \begin{cases} t^{\frac{1}{3}} & \text{para } t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29} & \text{para otro valor} \end{cases}$$

Nos queda por estudiar cómo obtener el valor del cambio entre dos colores en cada espacio. Cada uno de los modelos vistos, requerirá de una ecuación de diferencia de color que nos determine la magnitud del cambio producido. Volveremos a hablar de los espacios de color en el capítulo de detección de cambios.

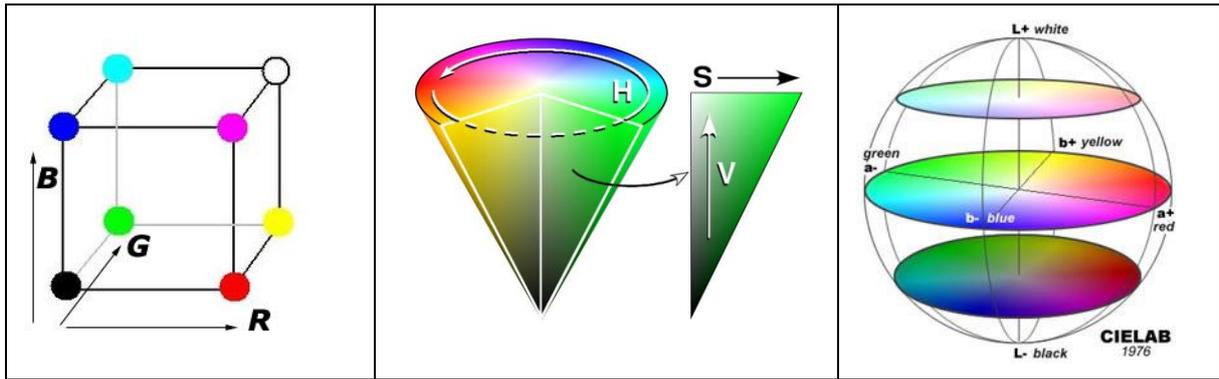


Figura 5-10. Representación de los espacios RGB, HSV y LAB.

Capítulo 6

Detección de cambios

Se conoce como detección de cambios al proceso de identificar diferencias en el estado de un objeto o fenómeno que ha sido observado en diferentes instantes de tiempo. La detección de cambios es útil en aplicaciones tan diversas como análisis de cambios de usos del suelo, evaluación de la deforestación, estudio de los cambios estacionales en la producción de pastos, evaluación de daños, detección de estrés hídrico de cultivos, estudio de cambios ambientales, análisis de cambios urbanos, etc.

Un cambio en la superficie terrestre supone un cambio en los valores de radiancia mayor que los provocados por otros factores como diferencias en las condiciones atmosféricas, en el ángulo del Sol y en la humedad del suelo. Estas variaciones son captadas fácilmente por las imágenes de satélite y aéreas, que son las usadas principalmente en la detección de cambios.

Son numerosas las técnicas ideadas para la detección de cambios y, hoy en día, es tema de numerosas investigaciones debido a la importancia que tiene la observación de la Tierra y los fenómenos que ocurren en ella. Existen trabajos que recopilan y resumen muchas de las técnicas existentes como (Singh, 1989) o (Serpico & Bruzzone, 1999). También, podemos encontrar mucha información sobre detección de cambios y perfectamente clasificada en la web, en una librería para la visión por computador (VisionBib.Com Computer Vision Information Pages).

Una buena detección de cambios debe proporcionar la siguiente información: área y porcentaje de cambios, su localización y distribución espacial, y, a ser posible, información sobre el cambio de tipo de uso del suelo. Sus resultados dependerán de muchos factores, entre otros los siguientes:

- Precisión en la corrección geométrica.
- Correcta normalización radiométrica de las imágenes.
- Disponibilidad de información previa sobre el suelo.
- Complejidad de la zona en estudio.

- Algoritmos o técnicas de detección usadas.
- Habilidad y experiencia del analista.
- Conocimiento o familiaridad con el área de estudio.
- Restricciones de coste y tiempo.

Es difícil determinar cuál de las técnicas de detección de cambios propuestas en los distintos trabajos de investigación es la mejor. Cada problema concreto necesitará de la técnica y estudio adecuado. Por esta razón, en este proyecto haremos un recorrido por las técnicas más habituales y estudiaremos las más usadas, entre ellas, la diferencia de imágenes.

6.1 Técnicas de detección de cambios

Los métodos de detección de cambios pueden ser agrupados en siete categorías: métodos algebraicos, transformaciones, clasificación, modelos avanzados, uso de GIS, análisis visual y otras técnicas que no encajan en ninguno de los grupos anteriores. En este capítulo veremos algunas de las técnicas más conocidas, sencillas e importantes y haremos una revisión de los distintos trabajos en detección de cambios.

6.1.1 Métodos algebraicos

En el grupo de métodos algebraicos tenemos la diferencia de imágenes, regresión de imágenes, división de imágenes, diferencia de índices de vegetación y análisis del vector de cambio (CVA). Todos estos algoritmos tienen en común que es necesario seleccionar un valor umbral para determinar las áreas de cambio. Excluyendo CVA, estos métodos son relativamente simples y fáciles de implementar e interpretar, pero no pueden proporcionar información sobre los cambios, tan sólo clasificar los distintos elementos como cambio o no. La principal desventaja es la dificultad de seleccionar un umbral apropiado para identificar las áreas de cambio. Otro aspecto crítico para las técnicas de esta categoría, es seleccionar las bandas o índices de vegetación más adecuados. Diferentes trabajos han ido apareciendo donde se pone de manifiesto cuales son las mejores bandas o índices a emplear en función del tipo de cambio que se desea detectar y de las zonas de estudio. Veamos a continuación una descripción de estos métodos, sus ventajas e inconvenientes, y algunos artículos donde se han empleado.

Diferencias de imágenes: Consiste en restar píxel a píxel las dos imágenes (corregistradas previamente) para producir una imagen de diferencias que represente los cambios producidos entre las fechas de adquisición de las mismas. Matemáticamente:

$$Dx_{ij}^k = x_{ij}^k(t_2) - x_{ij}^k(t_1) + C$$

donde x_{ij}^k es el nivel de gris del píxel en la banda k , con coordenadas i y j , t_1 es la fecha de adquisición de la primera imagen, t_2 la de la segunda, y C una constante para producir números positivos. Una alternativa consiste en obtener el valor absoluto de la diferencia y eliminar la constante, de esta forma obtenemos una imagen de cambios con el mismo rango de valores que las imágenes de entrada. Las zonas de mayor cambio corresponden a valores mayores en la imagen de cambios.

Es simple de implementar y muy fácil la interpretación de los resultados. Como inconveniente, no puede proporcionar una matriz detallada de cambios y requiere la selección de un umbral que delimite las áreas de cambio. Ejemplos de uso en (Muchoney & Haack, 1994), (Sohl, 1999), (Manavalan, Kesavasamy, & Adiga, 1995).

Regresión de imágenes: En el método de regresión se supone que los píxeles de la imagen tomada en el instante t_2 debe ser una función lineal de los píxeles de la imagen tomada en el instante t_1 . Esta función lineal se puede estimar por regresión de mínimos cuadrados. Sea $\hat{x}_{ij}^k(t_2)$ la predicción del valor obtenido por la línea de regresión, la imagen de diferencias o cambios se puede obtener con la expresión:

$$Dx_{ij}^k = \hat{x}_{ij}^k(t_2) - x_{ij}^k(t_1)$$

De nuevo aplicando un umbral, se obtienen las áreas de cambio. Esta técnica reduce los efectos producidos por diferencias en las condiciones atmosféricas y en el ángulo del Sol al tener en cuenta la media y varianza de los valores de los píxeles. Su cálculo es más complejo que la diferencia de imágenes y es necesario seleccionar las bandas más adecuadas para su cálculo. Como ejemplos de uso encontramos (Singh, 1986) y (Jha & Unni, 1994).

Ratio de imágenes: Es considerado un método rápido de identificación de áreas de cambio. En este método la imagen de cambios se obtiene dividiendo las bandas de las dos imágenes entre sí y aplicando un valor umbral. Se efectúa una comparación píxel a píxel como en los métodos anteriores. La imagen de cambios queda determinada por la expresión:

$$Rx_{ij}^k = \frac{x_{ij}^k(t_1)}{x_{ij}^k(t_2)}$$

donde $x_{ij}^k(t_2)$ representa el valor del píxel en la banda k , en la fila i y en la columna j en el instante t_2 . Si la intensidad o energía reflejada es similar en cada imagen tenemos que $Rx_{ij}^k = 1$, indicando que no existe cambio. En áreas de cambio el valor del ratio será significativamente mayor que 1 o menor que 1 dependiendo de la naturaleza de los cambios entre las dos fechas. La dificultad en esta metodología radica en seleccionar un valor umbral superior e inferior en la distribución no normal de los valores del cociente para determinar los píxeles de cambio. La práctica habitual es seleccionar arbitrariamente valores umbral y probarlos empíricamente para determinar si la detección de cambios es adecuada. Este método reduce el impacto

del ángulo del Sol, sombras y topografía. Como ejemplo de uso tenemos (Prakash & Gupta, 1998).

Diferencia de índices de vegetación: Existen diferencias espectrales notables entre la vegetación y otros componentes de la superficie terrestre. Esas diferencias en respuesta espectral aumentan a medida que la vegetación se hace más densa o más productiva. Se pueden tener en cuenta estas diferencias espectrales para obtener distintos índices que ayuden a la detección de cambios. Entre ellos los estudiados en el apartado 5.4.

La diferencia de estos índices nos puede dar información de aquellas zonas donde la vegetación ha cambiado. Como en los métodos anteriores, es necesario seleccionar un valor umbral para determinar las áreas de cambio. Reduce los efectos de la topografía e iluminación. Como inconveniente, esta técnica puede aumentar el ruido en la imagen resultante. Como ejemplos de uso de estos índices tenemos a (Townshend & Justice, 1995), (Guerra, Puig, & Chaume, 1998), (Lyon, Yuan, Lunetta, & Elvidge, 1998) y (Nelson, 1983).

Análisis del vector de cambio (CVA): Produce dos tipos de información de cambio, la dirección del vector de cambio y su magnitud. Sean dos imágenes de dos bandas, $x = (x_1, x_2)$ un píxel de la imagen tomada en el instante de tiempo t_1 donde x_i es el nivel de gris del píxel x en la banda i , e $y = (y_1, y_2)$ un píxel en la misma posición de la imagen tomada en el instante t_2 , el vector de cambio para cada píxel vendrá dado por la expresión:

$$\vec{D} = (y_1 - x_1, y_2 - x_2)$$

De este vector se puede obtener la magnitud que nos determinará si el píxel pertenece a una zona de cambio o no.

$$|\vec{D}| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2}$$

Aplicando un umbral a la imagen de magnitudes se obtiene un mapa binario de cambios. Además de la magnitud del vector de cambios, se puede hacer uso de la dirección del mismo para obtener una clasificación en función de su naturaleza. Volveremos a ver esto en el apartado 6.5.

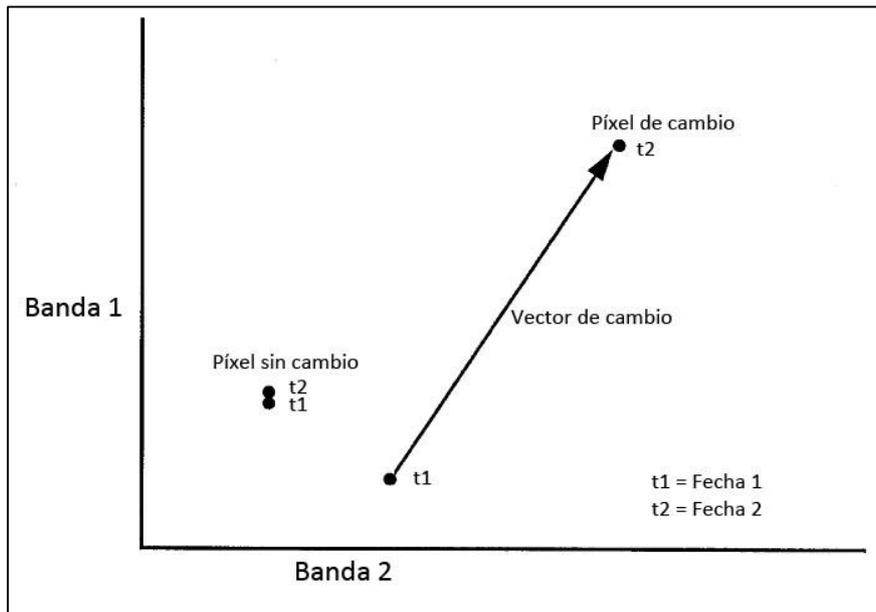


Figura 6-1. Vector de cambios en técnica CVA para imágenes de dos bandas.

La técnica CVA no sólo puede emplearse con imágenes de varias bandas sino también a varias transformaciones de las bandas, como PCA o Tasseled Cap, transformaciones que veremos a continuación.

6.1.2 Transformaciones

Con este conjunto de técnicas se pretende transformar las imágenes para obtener una mejor extracción de cambios. Una de las ventajas de estos métodos es reducir la redundancia entre las bandas de cada imagen y derivar otras componentes que representen mejor la variabilidad de los datos. Como otros métodos ya vistos, tampoco pueden proporcionar una matriz detallada de tipos de cambios y es necesario seleccionar un umbral para determinar las áreas de cambio.

Análisis de componente principales: Es usado para reducir un alto número de bandas de una imagen multispectral en unas pocas componentes principales teniendo en cuenta para ellos la máxima varianza de los datos en las distintas bandas. Con esta transformación se pueden detectar cambios de dos formas diferentes: calculamos las componentes principales para todas las bandas de cada una de las imágenes y, posteriormente, aplicando el método de diferencias, regresión o el módulo del vector del método CVA, se obtienen las áreas de cambios; o crear una única imagen compuesta por todas las bandas de ambas imágenes y obtener sus componente principales, que tras su análisis posterior facilitarán la detección de las áreas de cambios en las menores componentes principales.

El cálculo de las componentes principales consiste en lo siguiente: por cada banda se obtiene la media de todos los niveles de gris y se resta al valor de cada píxel dando lugar a un conjunto de datos con media cero. Con estos datos se obtiene una matriz de covarianza de tamaño $N \times N$ donde N es el número de bandas de la imagen y de ella se calculan los autovalores y autovectores (también conocidos como valores

y vectores propios). Estos autovectores determinan los ejes o direcciones en que los datos presentan una mayor varianza. El autovector correspondiente al mayor autovalor es la componente principal del conjunto de los datos, de esta forma, podríamos desechar aquellas componentes con autovalores menos significativos, y así reducir las dimensiones del conjunto de datos inicial. Una vez ordenados los autovalores y sus correspondientes autovectores de mayor a menor, se procede a transformar los datos iniciales multiplicándolos por cada uno de los autovectores para obtener los datos finales de cada una de las componentes principales. Se puede ver un ejemplo muy detallado con un conjunto de datos de dos dimensiones en (Smith, 2002). Como inconvenientes en esta técnica podemos citar que es dependiente de la escena, los resultados de la detección de cambios a menudo es difícil de interpretar, tampoco proporciona una matriz de cambios de clase y requiere la determinación de umbrales para la identificación de áreas de cambio. Como trabajos de ejemplo tenemos a (Byrne, Crapper, & Mayo, 1980), (Ingebritsen & Lyon, 1985), (Parra, Mouchot, & Roux, 1996), (Kwarteng & Chavez, 1998), (Li & Yeh, 1998), (Jha & Unni, 1994), (Collins & Woodcock, 1996), (Muchoney & Haack, 1994).

Transformación Tasseled Cap: Esta transformación es similar al análisis de componentes principales. La diferencia es que PCA depende de las imágenes y Tasseled Cap es independiente. La también conocida como transformación KT por los autores del mismo (Kauth & Thomas, 1976) se basa en PCA para obtener una transformación de las bandas de las imágenes originales. Se comprobó que cada una de estas componentes principales se correspondían con distintas características físicas de las imágenes Landsat: brillo, verdor y humedad. Trabajos más recientes nos han ofrecidos transformaciones para las imágenes de los satélites de alta resolución Quickbird e Ikonos. Para ello en (Yarbrough, Easson, & Kuzmaul, 2005) y (Horne H. J., 2003) se aplica una análisis de componentes principales a una base de datos de 11 escenas y 195 imágenes multispectrales diferentes respectivamente, obteniendo la transformación que mejor se ajusta a los resultados por el método de mínimos cuadrados.

La transformación para las imágenes Quickbird es la siguiente:

$$\begin{aligned}u_1 &= 0.319 x_{azul} + 0.542 x_{verde} + 0.490 x_{rojo} + 0.604 x_{nir} \\u_2 &= -0.121 x_{azul} - 0.331 x_{verde} - 0.517 x_{rojo} + 0.780 x_{nir} \\u_3 &= 0.652 x_{azul} + 0.375 x_{verde} - 0.639 x_{rojo} - 0.163 x_{nir} \\u_4 &= 0.677 x_{azul} - 0.675 x_{verde} + 0.292 x_{rojo} + 0.011 x_{nir}\end{aligned}$$

Y la transformación para las imágenes Ikonos:

$$\begin{aligned}u_1 &= 0.326 x_{azul} + 0.509 x_{verde} + 0.560 x_{rojo} + 0.567 x_{nir} \\u_2 &= -0.311 x_{azul} - 0.356 x_{verde} - 0.325 x_{rojo} + 0.819 x_{nir} \\u_3 &= -0.612 x_{azul} - 0.312 x_{verde} + 0.722 x_{rojo} - 0.081 x_{nir} \\u_4 &= -0.650 x_{azul} + 0.719 x_{verde} - 0.243 x_{rojo} - 0.031 x_{nir}\end{aligned}$$

donde x_{azul} , x_{verde} , x_{rojo} , y x_{nir} son los valores de los píxeles en las bandas del azul, verde, rojo e infrarrojo cercano respectivamente.

Para detectar cambios usando estas transformaciones se procedería al igual que con la transformación PCA, se transforman las imágenes y se aplica el método de diferencias, regresión o el módulo del vector del método CVA, para obtener las áreas de cambio tras aplicar el umbral adecuado. Al igual que PCA la información obtenida es difícil de interpretar y no proporciona una matriz completa de cambios de clase, además de requerir de valores umbrales correctos para la determinación de las áreas de cambio. Precisan de una calibración atmosférica precisa de cada imagen puesto que esta transformación se basa en las distintas respuestas espectrales de los sensores que adquieren las imágenes. Esta transformación se usa en (Collins & Woodcock, 1996), (Coppin, Nackaerts, Queen, & Brewer, 2001) y (Seto, Woodcock, Song, Huang, Lu, & Kaufmann, 2002).

Otras transformaciones han sido utilizadas en trabajos sobre detección de cambios como la transformación Gramm-Schmidt en (Collins & Woodcock, 1994) y la transformación Chi-Square en (Ridd & Liu, 1998).

6.1.3 Métodos por clasificación

Este conjunto de métodos se basa en imágenes clasificadas en las cuales la calidad y cantidad de muestras de entrenamiento son cruciales para producir buenos resultados de clasificación. La mayor ventaja de estos métodos es la capacidad de proporcionar una matriz de cambios de clase que dé información sobre la naturaleza de los cambios, además de reducir el impacto de las diferencias atmosféricas y ambientales entre las imágenes multi-temporales. No obstante, seleccionar un conjunto adecuado de píxeles de entrenamiento es a menudo difícil, haciendo lento y complejo el proceso de detección de cambios.

Comparación post-clasificación: Consiste en clasificar las imágenes con cualquiera de los métodos supervisados o no-supervisados existentes en la literatura científica con el fin de construir una matriz que refleje los cambios de clase que se han producido. Un píxel será de cambio si su clase es distinta en las dos imágenes empleadas para la detección, pero además, sabremos qué tipo de cambio se ha producido y podremos centrarnos sólo en aquellos que sean de interés, por ejemplo, nuevo suelo urbano o zonas deforestadas recientemente. Tenemos un ejemplo en la Figura 6-2.

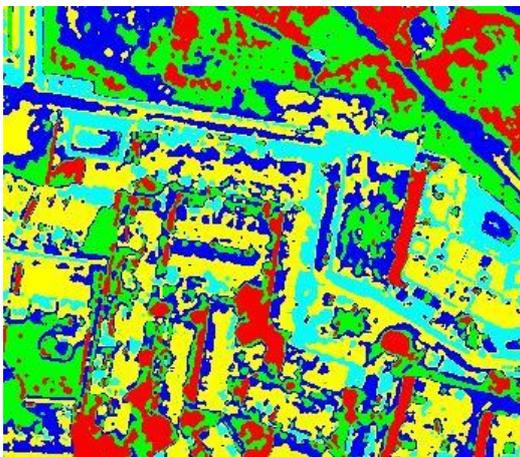
Otros ejemplos de uso en (Brondizio, Moran, Mausel, & Wu, 1994), (Dimiyati, Mizuno, Kobayashi, & Kitamura, 1996), (Mas, 1997), (Mas, 1999), (Castelli, Elvidge, Li, & Turek, 1998), (Miller, Bryant, & Birnie, 1998), (Foody, 2001), (Jensen, Ramsay, Mackey, Christensen, & Sharitz, 1987), (Jensen, Rutchey, Koch, & Narumalani, 1995), (Munyati, 2000) y (Ward, Phinn, & Murray, 2000).



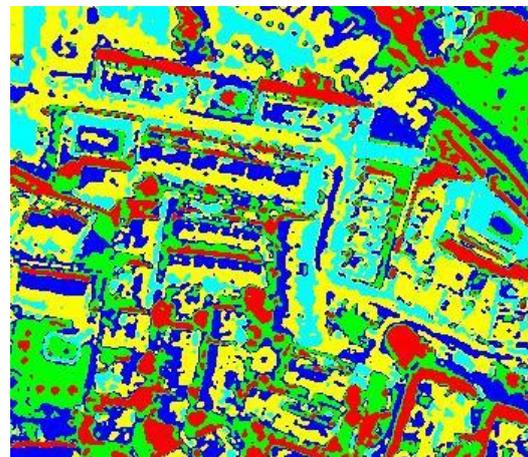
Imagen 1



Imagen 2



Clasificación imagen 1



Clasificación imagen 2

Pixel Count		Reference						
		Initial State						
		Class 1	Class 2	Class 3	Class 4	Class 5	Row Total	Class Total
Final State	Unclassified	0	0	0	0	0	0	0
	Class 1	4892	3195	2889	2268	576	13820	13820
	Class 2	3788	8563	4850	3155	1087	21443	21443
	Class 3	2467	5661	7425	7694	1901	25148	25148
	Class 4	2034	5063	6877	11599	4605	30178	30178
	Class 5	754	1911	2617	4407	4906	14595	14595
	Class Total	13935	24393	24658	29123	13075		
	Class Changes	9043	15830	17233	17524	8169		
Image Difference	-115	-2950	490	1055	1520			

Matriz de cambios de clase

Figura 6-2. Proceso de detección de cambios por post-clasificación.

Existen otros métodos basados en la clasificación de las imágenes, algunos de ellos usando redes neuronales artificiales como en (Gopal & Woodcock, 1996), (Abuelgasim, Ross, Gopal, & Woodcock, 1999), (Dai & Khorram, 1999), (Woodcock, Macomber, Pax-Lenney, & Cohen, 2001) y (Liu & Lathrop, 2002).

6.1.4 Modelos avanzados

En este grupo de técnicas de detección los valores de reflectancia de las imágenes son convertidos a medidas físicas. Los parámetros obtenidos son más intuitivos a la hora de interpretarlos y mejores para extraer información sobre la vegetación que las firmas espectrales. Las desventajas de estos métodos son su alto coste computacional y la dificultad de elaborar procesos adecuados de transformación de valores de reflectancia a parámetros biofísicos.

Algunos de los métodos en esta categoría son LSMA (Linear Spectral Mixture Analysis) usado en (Adams, y otros, 1995), (Roberts, Batista, Pereira, Waller, & Nelson, 1999), (Ustin, Roberts, & Hart, 1998), (Rogan, Franklin, & Roberts, 2002), (Radeloff, Mladenoff, & Boyce, 1999) y (Kressler & Steinnocher, 1996) entre otros, y el modelo de reflectancia Li-Strahler (Li & Strahler, 1985), (Li & Strahler, 1992).

No se han implementado ninguno de estos modelos apostando por la simplicidad y algoritmos más genéricos que permitan la detección de cambios de todo tipo, sin necesidad de imágenes complejas de alto número de bandas o de otro tipo de información asociada a los sensores, como modelos de calibración radiométrica o la respuesta espectral de los mismos.

6.1.5 Modelos GIS

En esta categoría se engloban aquellos métodos que hacen uso de información GIS (Sistemas de Información Geográfica) en el proceso de detección de cambios. Permite la incorporación de capas de información que ayuden a interpretar los distintos elementos de las imágenes y los cambios ocurridos, y la posterior incorporación del resultado al sistema de información geográfica para su análisis. Como desventaja podemos apuntar que el resultado de la detección de cambios depende de la calidad, precisión y escala de las distintas fuentes de información que se empleen.

Algunos trabajos de detección con uso de GIS son (Price, Pike, & Mendes, 1992), (Westmoreland & Stow, 1992), (Mouat & Lancaster, 1996), (Slater & Brown, 2000), (Petit & Lambin, 2001), (Chen, 2002), (Weng, 2002), (Yeh & Li, 2001), (Prol-Ledesma, Uribe-Alcantara, & Diaz-Molina, 2002), (Lo & Shipman, 1990) y (Taylor, Brewer, & Bird, 2000).

6.1.6 Análisis visual

Consiste en la interpretación visual de una composición de imágenes multitemporales. Permite el uso de toda la experiencia y conocimiento del analista responsable de la detección de cambios. La textura, forma, tamaño y patrones de las imágenes son elementos claves para la identificación de cambios por interpretación visual. Estos elementos no son usados generalmente en la detección digital de cambios por la dificultad que supone su extracción de las imágenes. No obstante, un analista experimentando puede hacer uso de todos ellos para ayudarle en la

detección. La desventaja de este método es que requiere mucho tiempo para detección de cambios en grandes áreas.

Una forma simple de análisis visual consiste en asignar una de las imágenes multitemporales a la banda del rojo, la siguiente al verde y una tercera si existe al azul. La composición de color resultante juntos a los distintos elementos de la escena dará información al analista para detectar los cambios acontecidos en ella. A continuación se muestra una de estas composiciones usando la banda roja de las dos imágenes de la Figura 6-2, la primera asignada al rojo, y la segunda al verde y azul.

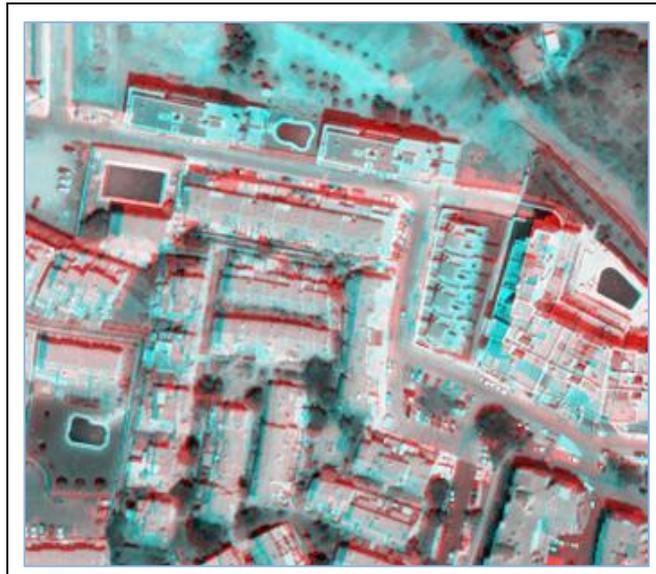


Figura 6-3. Asignación de bandas para análisis visual de cambios.

En ella podemos ver como las áreas de cambios quedan representadas en rojo o en cian dependiendo de la dirección del cambio. En gris quedan aquellas zonas que no han cambiado entre las fechas de adquisición de las imágenes. Por un lado, se puede observar cómo las piscinas, carreteras y jardines quedan en gris (sin cambios), por otro, cómo destaca el movimiento de tierras en cian en la parte superior de la imagen, las zonas en rojo debidas a sombras, y el color cian y rojo producido por los edificios al verse desde distintos ángulos en una y otra imagen.

Trabajos que hacen uso del análisis visual son (Sunar, 1998), (Ulbricht & Heckendorff, 1998), (Sader & Winne, 1992), (Stone & Lefebvre, 1998), (Asner, Keller, Pereira, & Zweede, 2002), (Slater & Brown, 2000).

6.1.7 Otras técnicas

Aquí se engloban trabajos como (Henebry, 1993) que hace uso de medidas de dependencia espacial con imágenes del sensor TM, (Wang F. , 1993) usa sistemas de visión basados en el conocimiento, y (Lambin & Strahler, 1994) combinando varios indicadores obtenidos de imágenes de sensor AVHRR. Es tan grande la variedad de sensores y de aplicaciones de la detección de cambios disponibles, que

constantemente aparecen nuevos trabajos desarrollando nuevas técnicas especializadas en una tarea y tipos de cambios e imágenes concretos.

6.2 Umbralización

Muchos algoritmos de detección de cambios, como los vistos en las categorías de métodos algebraicos y transformaciones, requieren la selección de un valor umbral para diferenciar los píxeles de cambio de aquellos que no lo son. *Umbralizar* una imagen no es más que clasificar los píxeles de una imagen en dos grupos, uno que incluye todos los píxeles por encima del valor umbral y otro con valor menor o igual a éste. A esta clasificación se le conoce como umbralización de dos niveles.

Muchos de los algoritmos existentes pueden generalizarse a una umbralización multinivel donde varios valores umbrales son seleccionados y los píxeles de la imagen clasificados en varias regiones. Generalmente, la umbralización binaria es la más adecuada para la detección de cambios, aunque se podría usar una umbralización multinivel para ofrecer una clasificación menos exacta usando lógica difusa y en vez de asignar a un píxel una clase (cambio o no cambio) asignarle una probabilidad de que el píxel sea de cambio.

La mayoría de las técnicas de umbralización se basan en la forma del histograma de la imagen para seleccionar el valor umbral. Una imagen de cambios ideal, donde los píxeles de cambios presentaran altos niveles de gris frente aquellos sin cambio, presentaría un histograma bimodal con dos picos bien diferenciados y un valle que los separe. En este caso, el valor umbral correspondería con el nivel de gris en el valle, separando los píxeles de ambas regiones. Pero en aplicaciones reales esta separación no está tan clara, incluso el histograma pasa a tener un único pico, dificultando la selección de umbral óptimo.

Muchos algoritmos han sido propuestos para la selección de umbrales en imágenes. En este proyecto estudiaremos la umbralización de Otsu, referencia para muchos trabajos posteriores, pero que supone que el histograma de la imagen es bimodal con dos picos diferenciados; selección de umbral usando ISODATA para una clasificación no supervisada de los píxeles de cambio, que trata de minimizar la distancia al nivel de gris medio de cada una de las dos regiones finales; un algoritmo para selección de umbrales en imágenes con un histograma unimodal, que son los habituales en imágenes de cambios por diferencia o división de imágenes; y, por último, un método que propone conservar los tres primeros momentos de la imagen, para obtener una imagen binaria con dos regiones diferencias y lo más parecida posible a la imagen original. Además presentaremos una técnica de umbralización multinivel.

Ya que determinar un correcto valor umbral es bastante subjetivo, depende de la escena y depende de la habilidad del analista o técnico que realiza la detección, además de los métodos no supervisados anteriores, se propone un método interactivo

que permite seleccionar un umbral de forma que el analista puede comprobar el resultado antes de aplicarlo. Igualmente, permite aplicar umbrales con histéresis, técnica que comentaremos en un apartado posterior.

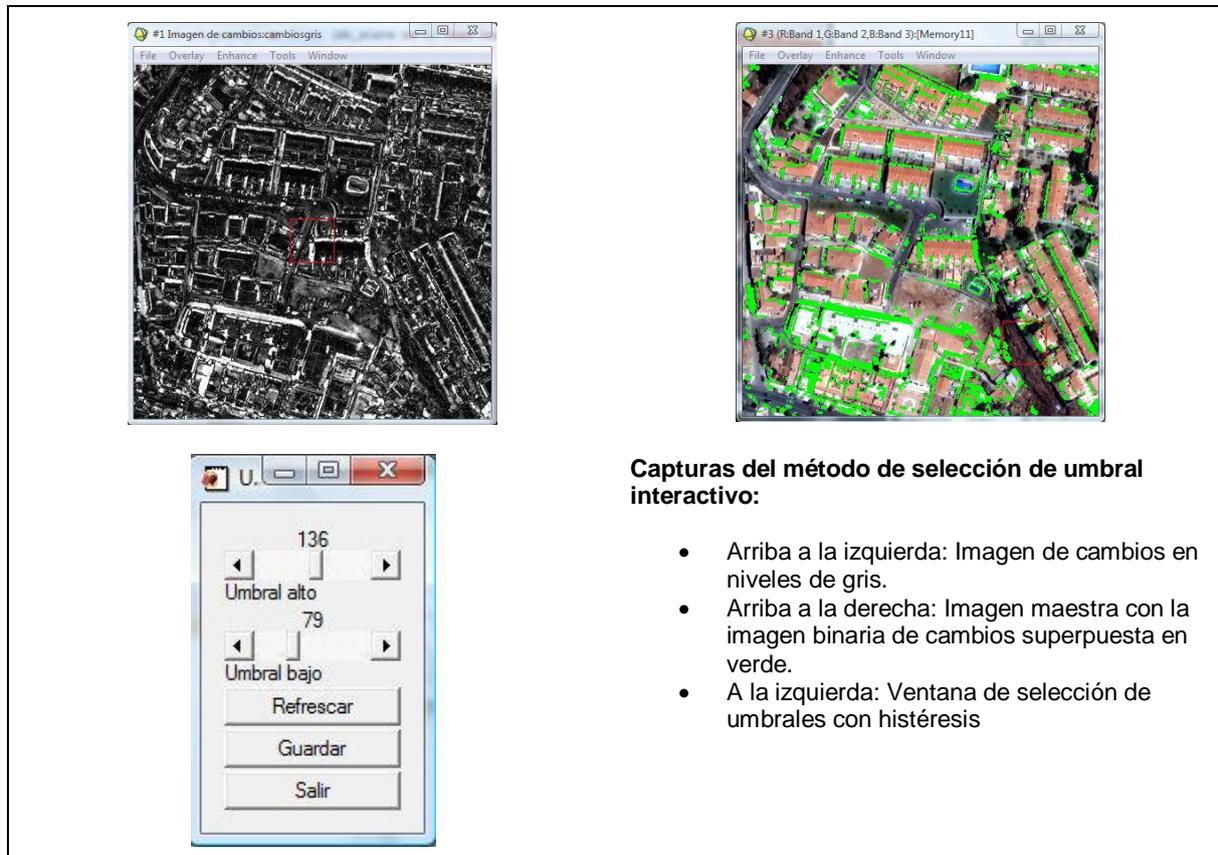


Figura 6-4. Ventanas de la función interactiva de selección de valores umbrales.

6.2.1 Umbralización de Otsu

La umbralización de (Otsu, 1979) es uno de los métodos de referencia en la mayoría de trabajos en los que se desea segmentar una imagen. Utiliza la varianza en los niveles de gris de los píxeles para obtener el valor umbral, de tal forma que la varianza entre los píxeles sea la mínima y que la varianza entre las dos clases generadas sea máxima.

Sea una imagen con L niveles de gris $[1,2,\dots,L]$, n_i el número de píxeles con nivel de gris i , N el número total de píxeles de la imagen y su histograma normalizado al transformarlo en una distribución de probabilidades por las siguientes ecuaciones:

$$p_i = \frac{n_i}{N}, \quad p_i \geq 0, \quad \sum_{i=1}^L p_i = 1$$

Si separamos los píxeles en dos clases C_0 y C_1 por un umbral de valor k , las probabilidades de que un píxel pertenezca a una clase vendrán dadas por:

$$\omega_0 = \Pr(C_0) = \sum_{i=1}^k p_i = \omega(k)$$

$$\omega_1 = \Pr(C_1) = \sum_{i=k+1}^L p_i = 1 - \omega(k)$$

Y las medias de los niveles de gris de los píxeles de cada clase por:

$$\mu_0 = \sum_{i=1}^k i \Pr(i|C_0) = \sum_{i=1}^k \frac{ip_i}{\omega_0} = \frac{\mu(k)}{\omega(k)}$$

$$\mu_1 = \sum_{i=k+1}^L i \Pr(i|C_1) = \sum_{i=k+1}^L \frac{ip_i}{\omega_1} = \frac{\mu_T - \mu(k)}{1 - \omega(k)}$$

donde

$$\omega(k) = \sum_{i=1}^k p_i, \quad \mu(k) = \sum_{i=1}^k ip_i$$

son los momentos de orden cero y orden uno del histograma hasta el nivel k -ésimo y

$$\mu_T = \mu(L) = \sum_{i=1}^L ip_i$$

es la media de todos los niveles de gris de los píxeles de la imagen original. Con estas definiciones se cumple que:

$$\omega_0 \mu_0 + \omega_1 \mu_1 = \mu_T, \quad \omega_0 + \omega_1 = 1 \quad (6.1)$$

Las varianzas de las clases vendrán dadas por las expresiones siguientes:

$$\sigma_0^2 = \sum_{i=1}^k (i - \mu_0)^2 \Pr(i|C_0) = \sum_{i=1}^k (i - \mu_0)^2 \frac{p_i}{\omega_0}$$

$$\sigma_1^2 = \sum_{i=k+1}^L (i - \mu_1)^2 \Pr(i|C_1) = \sum_{i=k+1}^L (i - \mu_1)^2 \frac{p_i}{\omega_1}$$

Como comentamos al principio de la sección se trata de maximizar la varianza entre clases y minimizar la varianza de los píxeles de cada clase. El valor umbral deseado será por tanto el que maximiza el cociente $\lambda = \sigma_B^2 / \sigma_W^2$ donde σ_B^2 es la varianza entre las clases y σ_W^2 la varianza dentro de las clases expresadas por:

$$\begin{aligned}\sigma_W^2 &= \omega_0\sigma_0^2 + \omega_1\sigma_1^2 \\ \sigma_B^2 &= \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 = \omega_0\omega_1(\mu_1 - \mu_0)^2\end{aligned}$$

donde σ_B^2 se ha simplificado usando (6.1). Maximizar λ es equivalente a maximizar la ecuación $\eta = \sigma_B^2/\sigma_T^2$, ya que se cumple que $\eta = \lambda/(\lambda + 1)$ donde:

$$\sigma_T^2 = \sum_{i=1}^L (i - \mu_T)^2 p_i$$

Dado que σ_T^2 es independiente de k , maximizar η es equivalente a maximizar σ_B^2 , el valor umbral óptimo t vendrá dado por:

$$\sigma_B^2(t) = \max_{1 \leq k < L} \sigma_B^2(k)$$

El método de Otsu, se basa en la forma del histograma y supone un histograma al menos bimodal con un valle diferenciado. Las imágenes de cambios que se suelen obtener en la detección de cambios en imágenes de satélite generan histogramas sin valles claramente definidos. Esto hace que este método de umbralización no sea en principio el más adecuado para la umbralización de imágenes de cambios.

6.2.2 Umbralización de histograma unimodal

Habitualmente, cuando realizamos la diferencia o ratio de imágenes para obtener cambios, obtenemos una imagen en niveles de gris con un histograma unimodal. Como hemos comentado, la mayoría de algoritmos de umbralización están ideados para obtener el umbral óptimo de una imagen con histograma al menos bimodal, segmentando la imagen en dos regiones, el objeto principal y el fondo. (Rosin, 2001) propone un método de selección de valor umbral para imágenes con histograma unimodal.

El algoritmo propuesto supone que existe un pico principal en el histograma debido a los píxeles de una de las dos clases de píxeles en la imagen (generalmente, los que no suponen cambios) y que la segunda clase puede no producir un pico fácil de identificar pero que es necesario separar del primero para evitar que los píxeles se clasifiquen erróneamente. Para encontrar el valor umbral, se traza en el histograma una línea recta desde el pico principal hasta la primera entrada vacía del histograma. El valor umbral seleccionado corresponde con el nivel de gris del histograma que se encuentre a la máxima distancia de la línea trazada como representa la siguiente imagen.

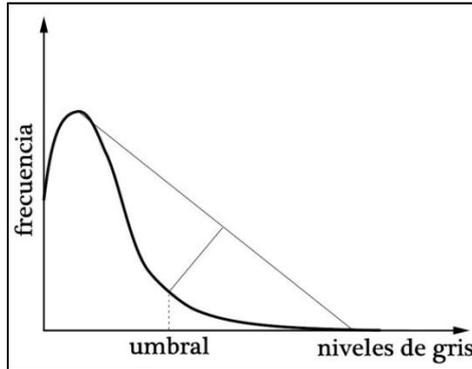


Figura 6-5. Selección de nivel umbral para histograma unimodal.

6.2.3 Umbralización con conservación de los momentos

(Tsai W. , 1985) propone un método de selección de valor umbral basado en el principio de conservación de los momentos de la imagen. Antes de umbralizar la imagen, se calculan los momentos de la imagen de entrada. Entonces, se obtiene el valor umbral de tal forma que los momentos de la imagen umbralizada se mantengan sin cambios. Esta operación es similar a una transformación que recupera una imagen ideal a partir de una versión suavizada de la misma.

Dada una imagen f con n píxeles, y siendo $f(x, y)$ el nivel de gris del píxel (x, y) . Se define el momento i -ésimo m_i de f como sigue:

$$m_i = \frac{1}{n} \sum_x \sum_y f^i(x, y), \quad i = 1, 2, 3 \dots$$

que se pueden calcular a partir del histograma de f de la siguiente forma:

$$m_i = \frac{1}{n \sum_j n_j (z_j)^i} = \sum_j p_j (z_j)^i \quad (6.2)$$

donde n_j es el total de píxeles en la imagen f con el nivel de gris z_j y $p_j = n_j / n$. Además se define $m_0 = 1$. Como comentamos antes, se puede considerar la imagen f como una versión suavizada de una imagen ideal g con píxeles con solo dos niveles de gris z_0 y z_1 , donde $z_0 < z_1$. Se trata entonces de seleccionar un valor umbral que conserve los tres primeros momentos de f , si todos los píxeles con valor menor que el valor umbral son reemplazados por z_0 y todos aquellos por encima son reemplazados por z_1 .

Sean p_0 y p_1 los porcentajes de píxeles por debajo y encima del valor umbral respectivamente, los tres primeros momentos de g corresponden con las siguientes ecuaciones:

$$m'_i = \sum_{j=0}^1 p_j (z_j)^i, \quad i = 1,2,3.$$

Y conservar los tres primeros momentos en g requiere que se cumplan las siguientes igualdades:

$$m'_i = m_i, \quad i = 1,2,3.$$

Sabiendo que $p_0 + p_1 = 1$, las igualdades descritas son equivalentes a:

$$\begin{aligned} p_0 z_0^0 + p_1 z_1^0 &= m_0 \\ p_0 z_0^1 + p_1 z_1^1 &= m_1 \\ p_0 z_0^2 + p_1 z_1^2 &= m_2 \\ p_0 z_0^3 + p_1 z_1^3 &= m_3 \end{aligned} \quad (6.3)$$

donde $m_0 = 1$, y m_i con $i = 1,2,3$ son obtenidos por (6.2). Para encontrar el valor umbral t , primero se resuelven las ecuaciones en (6.3) con el propósito de obtener p_0 y p_1 . Entonces, se obtiene el valor t tal que:

$$p_0 = \frac{1}{n} \sum_{z_j \leq t} n_j$$

Es decir, se trata de encontrar el primer nivel de gris del histograma que supere el porcentaje de píxeles indicado por p_0 .

Estas expresiones pueden generalizarse para una umbralización multinivel, pero para este trabajo no es necesario, pues el propósito es encontrar dos clases de píxeles y segmentar la imagen en dos regiones, la de píxeles de cambio y la de píxeles que no han cambiado.

Las ecuaciones en (6.3) pueden resolverse con las siguientes:

$$\begin{aligned} c_d &= \begin{vmatrix} m_0 & m_1 \\ m_1 & m_2 \end{vmatrix}; \\ c_0 &= \frac{1}{c_d} \begin{vmatrix} -m_2 & m_1 \\ -m_3 & m_2 \end{vmatrix}; \\ c_1 &= \frac{1}{c_d} \begin{vmatrix} m_0 & -m_2 \\ m_1 & -m_3 \end{vmatrix}; \\ z_0 &= \frac{1}{2} \left[-c_1 - (c_1^2 - 4c_0)^{\frac{1}{2}} \right]; \\ z_1 &= \frac{1}{2} \left[-c_1 + (c_1^2 - 4c_0)^{\frac{1}{2}} \right]; \\ p_d &= \begin{vmatrix} 1 & 1 \\ z_0 & z_1 \end{vmatrix}; \end{aligned}$$

$$p_0 = \frac{1}{p_d} \left| \begin{array}{cc} 1 & 1 \\ m_1 & z_1 \end{array} \right|;$$

$$p_1 = 1 - p_0.$$

6.2.4 Selección automática de valor umbral por técnica ISODATA

Esta técnica iterativa fue desarrollada por (Ridler & Calvard, 1978). Y consta de los siguientes pasos:

1. Seleccionar un valor umbral inicial T_0 cuyo valor es la mitad del rango dinámico máximo, es decir, el valor medio entre el mínimo y máximo nivel de gris en los píxeles de la imagen de cambios.
2. Segmentar el histograma en dos partes usando el valor T_0 generando dos regiones, la de cambios (C) y la de píxeles sin cambio (NC).
3. Calcular la media de los niveles de gris de los píxeles en ambas regiones ($m_{N,0}$ y $m_{NC,0}$).
4. Obtener un nuevo valor umbral T_1 como el promedio de las medias obtenidas anteriormente.
5. Volver a segmentar el histograma en dos partes, usando el nuevo valor umbral.
6. Repetir el proceso hasta que los valores umbrales converjan, es decir:

$$T_k = \frac{m_{N,k-1} + m_{NC,k-1}}{2} \text{ hasta } T_k = T_{k-1}$$

Usando esta técnica podemos obtener un valor umbral automáticamente que en algunos casos será suficiente para la detección de cambios, y en otros ayudará al analista a determinar un valor correcto.

6.2.5 Umbralización con histéresis

Un umbral con histéresis consiste en seleccionar dos valores umbral t_0 y t_1 tal que $t_0 \leq t_1$, de forma que todo píxel con valor superior a t_1 quede seleccionado como cambio y todo aquel con valor inferior o igual a t_0 se marque como no-cambio. Los píxeles con valores digitales comprendidos entre estos dos valores se clasificarán como cambio, sólo si están conectados a píxeles ya seleccionados como cambio por el valor umbral t_1 . El algoritmo funciona de la siguiente manera: 1) se segmenta la imagen de cambios en niveles de gris, aplicando los dos valores umbrales, para obtener dos imágenes binarias de cambio: una con los píxeles de cambios debidos al umbral superior t_1 , y la otra con aquellos cuyo nivel de gris superan el valor umbral t_0 ; 2) se dilata la imagen binaria obtenida por el umbral t_1 y se realiza una operación "and" con la obtenida de aplicar el umbral t_0 , de esta forma, incorporamos a la imagen de cambios, los píxeles comprendidos entre los valores t_0 y t_1 que están conectados a los cambios ya seleccionados como definitivos; 3) se realiza esta operación hasta que ningún píxel es añadido a la imagen binaria de cambios definitiva.

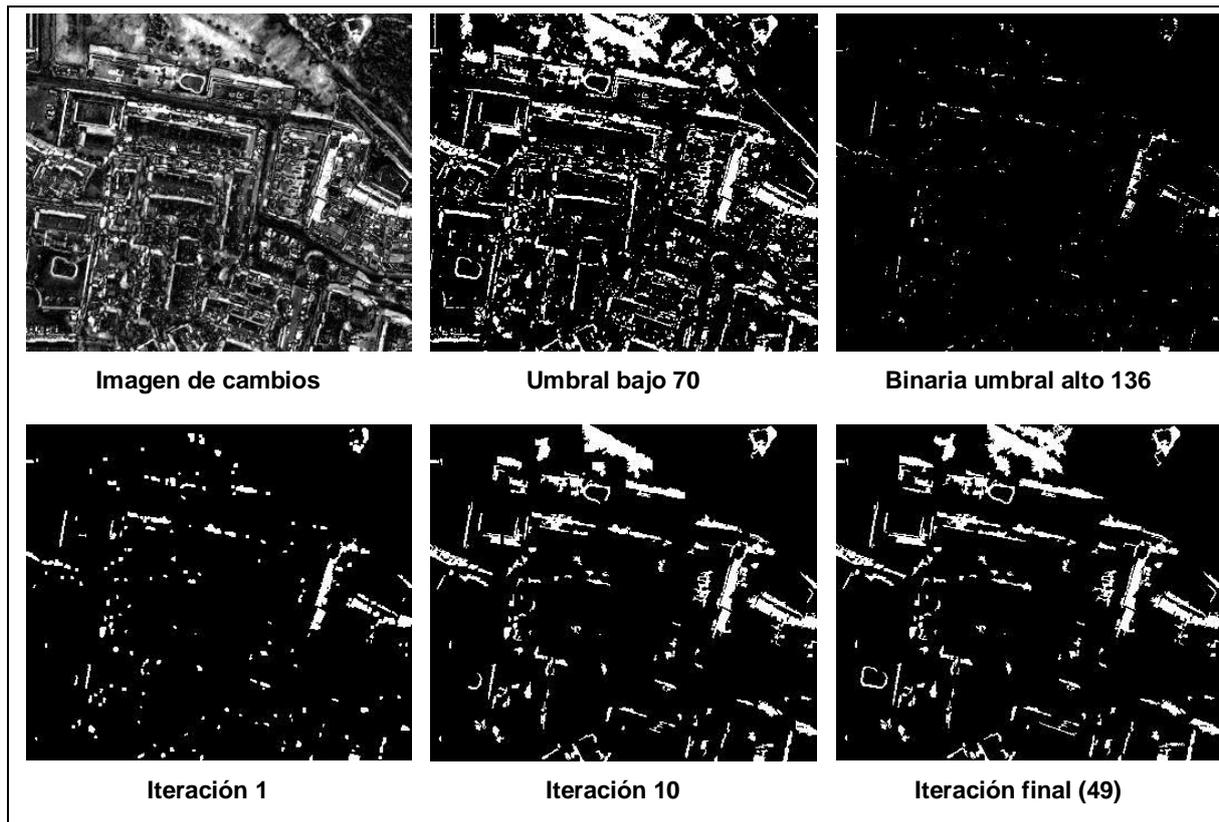


Figura 6-6. Iteraciones en la umbralización con histéresis.

Si usamos un único valor umbral muy alto, muchos píxeles serán clasificados erróneamente como no-cambio, de la misma forma si es muy bajo, muchos píxeles serán clasificados como cambios cuando no lo son. También ocurre que una vez seleccionado un valor umbral adecuado, hay áreas de cambio que no quedan localizadas completamente porque sus valores digitales están en torno al valor umbral. La histéresis soluciona esto, permitiendo seleccionar un valor umbral adecuado y otro inferior para no perder aquellos píxeles con menor nivel de gris pero que pertenecen a un cambio ya identificado y al que están conectados.

6.2.6 Umbralización multinivel

Todos los métodos de umbralización vistos comparten un mismo problema, la dificultad de determinar un umbral óptimo que clasifique los píxeles de la imagen como cambio o no cambio, sin incluir falsos negativos ni positivos. Es prácticamente imposible seleccionar un valor umbral que clasifique perfectamente todos los píxeles.

Esto lleva a usar técnicas como los conjuntos y el razonamiento difuso. Usando estas técnicas los píxeles serán clasificados como cambio o no-cambio pero con un grado mayor o menor que otro píxel dependiendo de lo que se conoce como función de pertenencia.

Sea X el conjunto de todos los píxeles, se puede generar dos clases o conjuntos difusos C y NC para los cambios y los no-cambios, de la forma $C = \{x, \mu_C(x)\}$; $x \in X$ donde $\mu_C(x)$ es la que hemos llamado función de pertenencia. Así un píxel x formará

parte de la clase C en el grado que establezca su función de pertenencia, que asocia a cada píxel un valor en el intervalo $[0,1]$. De tal forma la función de pertenencia de la clase de no-cambios NC queda definida por:

$$\mu_{NC}(x) = 1 - \mu_C(x) \quad (6.4)$$

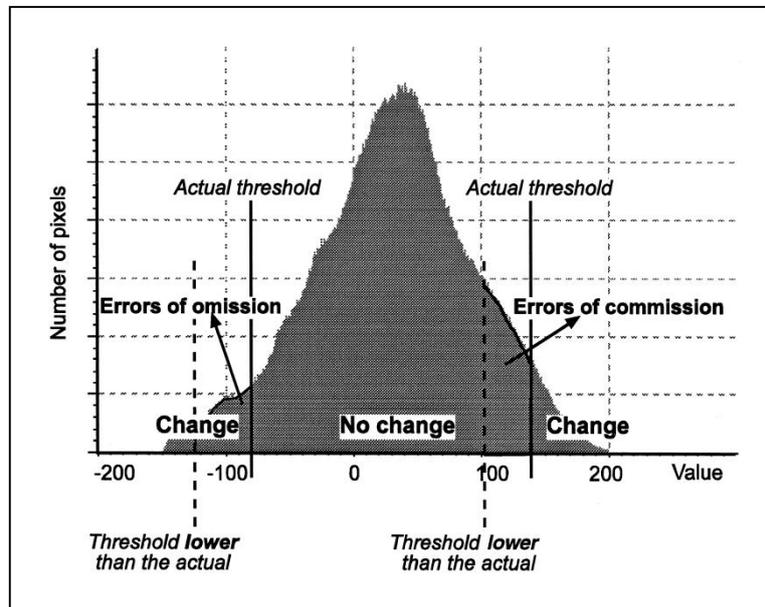


Figura 6-7. Histograma de imagen de cambios por método de diferencias.

Como, generalmente, las imágenes de diferencias presentan un histograma unimodal, con un pico provocado por los píxeles de no-cambio y con forma de campana (Figura 6-7), en (Metternicht, 1999) se propone una función de pertenencia en forma de s dadas por las expresiones:

$$\mu_{NC}(x) = \frac{[(1 - \nu)^{\lambda-1}(x - a)^\lambda]}{[(1 - \nu)^{\lambda-1}(x - a)^\lambda + \nu^{\lambda-1}(b - x)^\lambda]}; x \in [a, b]$$

$$\mu_{NC}(x) = \frac{[(1 - \nu)^{\lambda-1}(c - x)^\lambda]}{[(1 - \nu)^{\lambda-1}(c - x)^\lambda + \nu^{\lambda-1}(x - b)^\lambda]}; x \in [b, c]$$

donde λ es un indicador de cuán rápido se incrementa la pertenencia a un conjunto y ν es el punto de inflexión de la función; a y c los puntos donde la función de pertenencia es 0 y b el punto donde la pertenencia es máxima (igual a 1).

Ya que el pico del histograma de la imagen de diferencias corresponde con píxeles sin cambio, se puede tomar como valor b , el nivel de gris con mayor número de píxeles en el histograma, y como a y c los niveles de gris más bajo y alto respectivamente con valor distinto de 0 en el histograma.

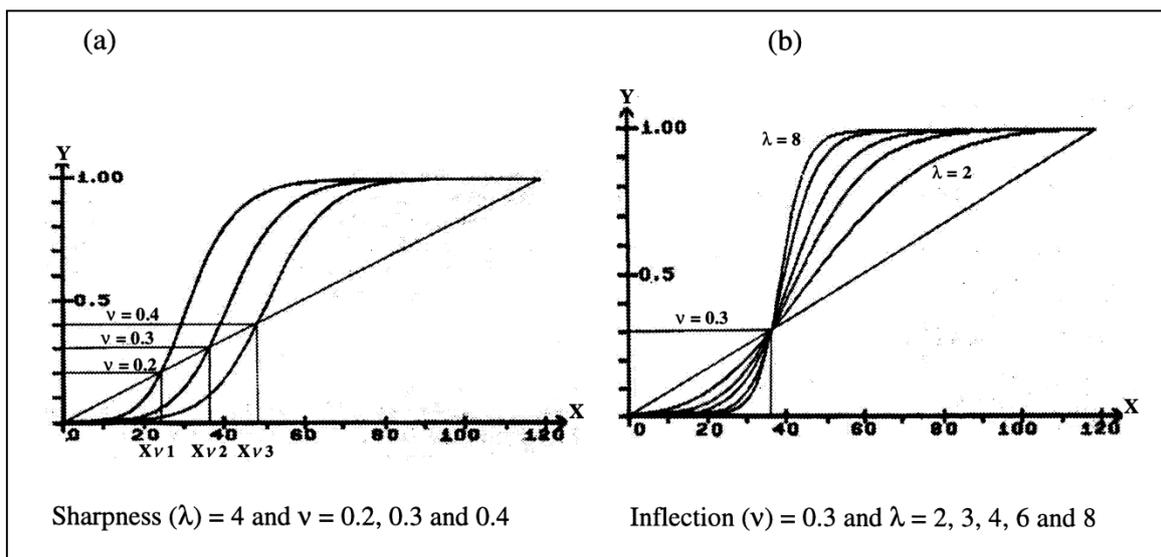


Figura 6-8. Ejemplos de cambio en λ y ν . Tomado de (Metternicht, 1999).

Pero aún establecidos el valor de los parámetros a , b , y c queda por determinar los valores de λ y ν para que la función de pertenencia sea lo más próxima al histograma. Esto requiere de varios ajustes para obtener los valores adecuados. Con el propósito de hacer este procedimiento de la forma más automática y rápida posible, y suponiendo que el histograma carece de cambios bruscos entre los distintos niveles (en otras palabras, tiene una forma similar al de la Figura 6-7), se propone en este proyecto usar el histograma normalizado como función de pertenencia, asignando así un valor de 1 a aquellos píxeles que corresponden al máximo del histograma y 0 a los de los extremos. Esta última sería la función de pertenencia para la clase de píxeles sin cambio, para obtener la de la clase de cambios basta con aplicar (6.4) para obtener su complementario.

En la Figura 6-9 se muestra un ejemplo de umbralización multinivel usando este método. Este tipo de imágenes puede servirnos, posteriormente a la detección de cambios, para analizar posibles falsos negativos y positivos mediante inspección visual.

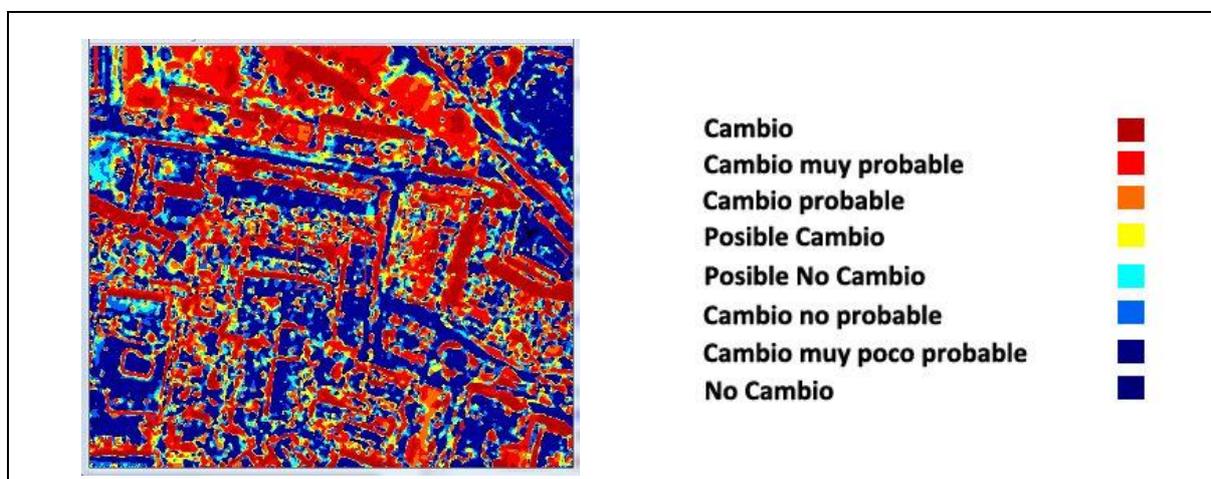


Figura 6-9. Umbralización multinivel por probabilidad de cambio.

6.2.7 Umbralización adaptativa

Las técnicas de Otsu, Histograma Unimodal, Conservación de momentos y umbral ISODATA, obtienen un valor umbral global para toda la imagen de cambios. Pero como hemos visto en otras situaciones, podemos idear un método adaptativo que nos ofrezca un valor umbral para el vecindario de cada píxel.

Para realizar una umbralización adaptativa se realiza un procedimiento similar al que realizábamos en la normalización radiométrica adaptativa bilineal. Se divide la imagen a procesar, en este caso la imagen de cambios en niveles de gris, en ventanas conectadas por puntos de malla tal como vimos en la Figura 3-6. A continuación, se obtiene el umbral óptimo en cada punto de malla por cualquiera de los procedimientos que hemos estudiado y al resto de píxeles se les aplica un umbral obtenido por interpolación bilineal de los 4 puntos de malla vecinos. En el siguiente apartado compararemos los distintos métodos en su versión global y adaptativa.

6.2.8 Estudio comparativo de métodos de umbralización

En este apartado vamos a estudiar con un par de imágenes de ejemplo, el desempeño de cada uno de los métodos vistos de umbralización binaria en sus versiones global y adaptativa. Para ello usaremos las imágenes del experimento 2 del capítulo de correcciones geométricas que son las que más cambios presentan entre las fechas de adquisición (Figura 4-25).

Primero se crean manualmente dos máscaras, una con cambios y otra con zonas que no han sufrido cambio. El mejor método de umbralización será aquel que ofrezca un mayor porcentaje de píxeles de cambio en las zonas seleccionadas por la máscara de cambios (verdaderos positivos) al que llamaremos P_{VP} , y un menor porcentaje en las zonas seleccionadas por la máscara de no-cambios (falsos positivos) al que denominaremos P_{FP} .

	Otsu		Conservación de momentos		Histograma unimodal		ISODATA	
	P_{VP}	P_{FP}	P_{VP}	P_{FP}	P_{VP}	P_{FP}	P_{VP}	P_{FP}
Global	50,29%	4,21%	41,59%	3,19%	50,59%	4,21%	51,85%	4,36%
Adap. 64	54,95%	4,70%	47,42%	4,15%	55,56%	7,89%	55,33%	4,71%
Adap. 128	55,74%	4,75%	47,77%	3,95%	53,14%	6,23%	57,42%	4,87%
Adap. 256	55,36%	4,42%	47,89%	3,53%	40,38%	5,04%	57,67%	4,58%

Tabla 6-1. Comparativa de técnicas de umbralización automática.

En nuestro ejemplo, la técnica que consigue un mayor porcentaje de verdaderos positivos (P_{VP}) es la ISODATA. Generalmente, usar umbralización

adaptativa incrementa el P_{VP} , aunque vemos que en la técnica de histograma unimodal, cae para una ventana de 256 píxeles y el porcentaje de falsos positivos (P_{FP}) es superior que el resto de técnicas. Esta técnica depende de la forma del histograma, por eso se recomienda su uso sólo en su forma global, donde el histograma tiene mucha más información.

La que peor resultados da en este experimento es la de conservación de momentos, que aunque tienen un menor P_{FP} , su P_{VP} también es inferior al del resto.

Una porcentaje de verdaderos positivos en torno al 50% es muy bajo en cualquier proceso de detección de cambios. Siempre es conveniente tener un mayor número de verdaderos positivos aunque eso implique un incremento en el de falsos positivos. Es más fácil eliminar posteriormente falsos positivos que estudiar las imágenes en busca de cambios no detectados. Los valores tan bajos obtenidos en este ejemplo, se deben a haber seleccionado como muestra de cambios, algunos con color muy similar en las dos imágenes, lo que suele ser problemático en la detección de cambios basada en píxeles y el color.

El umbral devuelto por cualquiera de las técnicas, mejor las que ofrecen un mayor P_{VP} , puede servir como punto de partida para un umbralización interactiva, donde el analista aplicando histéresis y pequeños ajustes pueda afinar los umbrales hasta obtener el resultado deseado.

6.3 Espacios de color en la detección de cambios

En la detección de cambios basados en el color, como en las técnicas algebraicas y transformaciones que hemos visto, los resultados obtenidos dependen como cabe esperar del espacio de color empleado en las imágenes.

Cuando realizamos una detección de cambios en imágenes de varias bandas, vimos que podíamos aplicar el método CVA para obtener el módulo del vector de cambio. De esta forma, podemos obtener una imagen de cambios usando cualquier tipo de imagen y espacio de color. Habitualmente, al igual que para este proyecto, tendremos como fuente de información imágenes de tres bandas en el espacio de color RGB. Pero el espacio RGB, tiene el inconveniente de que existe una alta correlación entre sus componentes y no es perceptivamente lineal, es decir, dos vectores de cambio distintos con módulo idéntico, no siempre implica un cambio idéntico en su percepción visual.

Por este motivo, podemos pensar que usar otros espacios de color más adecuados a la percepción visual del ser humano, nos ofrecerá unos mejores resultados en la detección, pues somos nosotros, los analistas y usuarios de los diferentes algoritmos y técnicas, quienes en última instancia decidimos si se ha producido un cambio o no. En este momento, surgen dos preguntas. Por un lado, si es importante la información del color obtenida por los sensores multiespectrales, o

basta con usar las imágenes en niveles de gris de los sensores pancromáticos. Y por otro lado, si usar espacios de color perceptiblemente más lineales que el espacio RGB mejora los resultados de la detección de cambios.

Vamos a estudiar la influencia del espacio de color en la detección de cambios para las imágenes del experimento 2 del capítulo de correcciones geométrica. Como cada espacio de color requeriría de un valor umbral diferente y no es posible usar un único valor umbral para comparar los resultados de la detección para cada uno de ellos, usaremos lo que se conoce como curvas ROC (Receiver Operating Characteristic).

Una curva ROC es una representación del porcentaje de verdaderos positivos frente al porcentaje de falsos positivos para un sistema clasificador binario según varía el valor umbral seleccionado. Para su cálculo, se usan las máscaras de cambios y no-cambios que se usaron en el apartado 6.2.8, y se procede a calcular ambos porcentajes para varios valores umbrales diferentes abarcando todo el rango dinámico de la imagen de cambios.

En la Figura 6-10 vemos tres curvas de ejemplo. Para cada curva se obtiene una lista de porcentajes de verdaderos positivos P_{VP} y de porcentajes de falsos positivos P_{FP} aplicando diferentes valores umbrales. En el eje X se representa los porcentajes P_{FP} y en el eje Y los porcentajes P_{VP} . En el caso de un detector de cambios ideal, obtendríamos la Curva 1, es decir el 100 por 100 de aciertos o verdaderos positivos para un porcentaje 0 de falsos positivos. Por tanto, una técnica o imágenes serán más adecuadas cuanto más se acerque su curva a la Curva 1. Con esto, podemos afirmar que la técnica o las imágenes empleadas para obtener la Curva 2, son más adecuadas que las usadas para la Curva 3.

Usando este sistema pasaremos a intentar dar respuesta a las preguntas que hicimos anteriormente. En la Figura 6-11 podemos comparar las curvas ROC al usar diferencia de gris y el módulo del vector de cambios usando imágenes en RGB, HSV y CIELab. La gráfica muestra los resultados para un porcentaje de verdaderos positivos superior al 50% donde los cambios entre espacios de color se hacen notables. Como podemos apreciar claramente, y tal como podríamos esperar, usando la información del color obtenemos un mejor resultado que usando imágenes en niveles de gris, donde parte de la información del cambio se ha perdido.

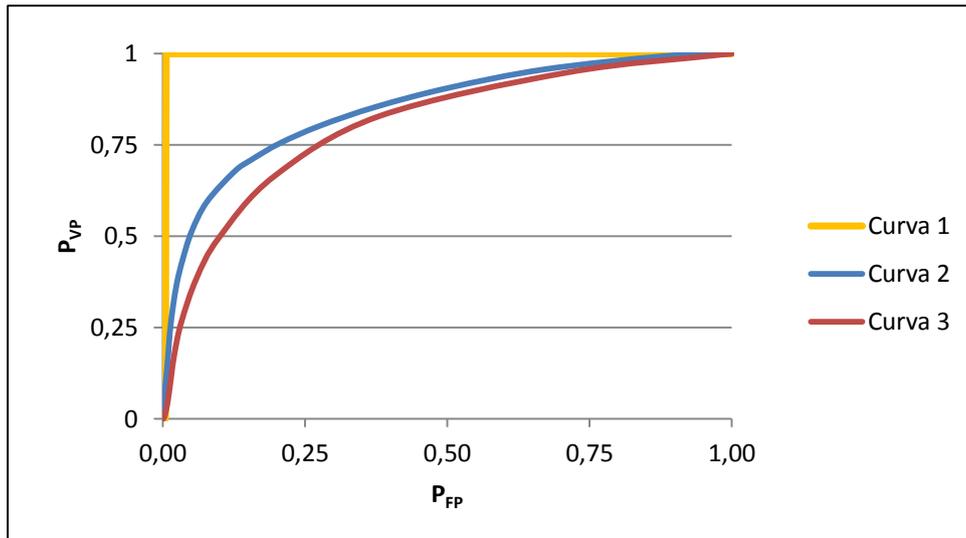


Figura 6-10. Curvas ROC.

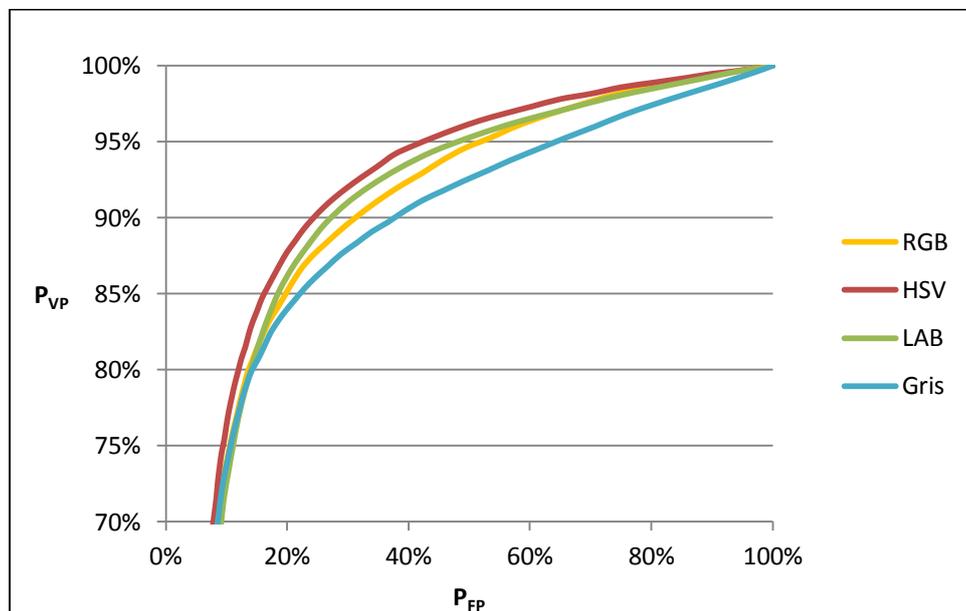


Figura 6-11. Curvas ROC usando imágenes RGB, HSV, LAB e imágenes en niveles de gris.

Para obtener la imagen de cambios usando HSV y LAB debemos usar las siguientes expresiones de diferencia de color. Para dos píxeles con vectores de color en el espacio HSV (h, s, v) y (h', s', v') , la magnitud del cambio vendrá dada por:

$$\Delta E = \sqrt{(vs \cos(2\pi h) - v's' \cos(2\pi h'))^2 + (vs \sin(2\pi h) - v's' \sin(2\pi h'))^2 + (v - v')^2}$$

Para dos píxeles del espacio LAB (L, a, b) y (L', a', b') la magnitud del cambio será:

$$\Delta E_{Lab} = \sqrt{(L - L')^2 + (a - a')^2 + (b - b')^2}$$

Comparando las distintas curvas se puede observar que, para las imágenes empleadas en la prueba, cualquiera de los espacios de color alternativos al RGB ofrece un mejor resultado en la detección de cambios. Podemos concluir, aunque

sería necesario un estudio más exhaustivo con muchos más pares de imágenes, que el color aporta información útil en la detección de cambios y que usar espacios o expresiones de diferencias de color perceptivamente más lineales mejora los resultados.

6.4 Filtrado de cambios

De todos los cambios identificados por cualquiera de las técnicas presentadas, serán de interés sólo un porcentaje de ellos dependiendo del estudio que se esté realizando y del fenómeno que se pretenda observar. Muchos de los píxeles de cambio, se deberán a cambios en las condiciones de adquisición de las imágenes, a sombras, a la presencia de nubes, etc., cambios que generalmente no son de interés. Un filtrado de estos cambios nos facilitará la identificación de aquellos que sí son de interés para el usuario final. En este apartado, estudiaremos distintos falsos positivos que pueden surgir y cómo eliminarlos automáticamente.

El primer filtrado ocurre en la fase de detección. Cuando se aplica un umbral a la imagen de cambios en niveles de gris, también se aplica un filtro mediana para eliminar los cambios debidos a ruido en las imágenes. Generalmente, cuando se realiza un análisis de cambios, sólo son de interés aquellos que tienen un área mayor a un valor determinado. Eliminando los cambios con un área inferior obtendremos un mapa de cambios más claro y más fácil de interpretar. También podemos aplicar un valor máximo, eliminando cambios que superen un área determinada. Esto nos puede ayudar a eliminar cambios que no sean del interés del analista, por ejemplo, grandes extensiones de campo cuyo nivel de gris ha cambiado por la vegetación o por la humedad del suelo.

En el ejemplo siguiente, la mancha verde oscuro representa a un árbol que ha sido talado. Este pequeño cambio puede no ser de interés para el analista y puede ser eliminado automáticamente aplicando un filtrado por área. El ruido en las imágenes, de tipo sal y pimienta, también será eliminado tanto al aplicar el filtro mediana como el filtrado por área.

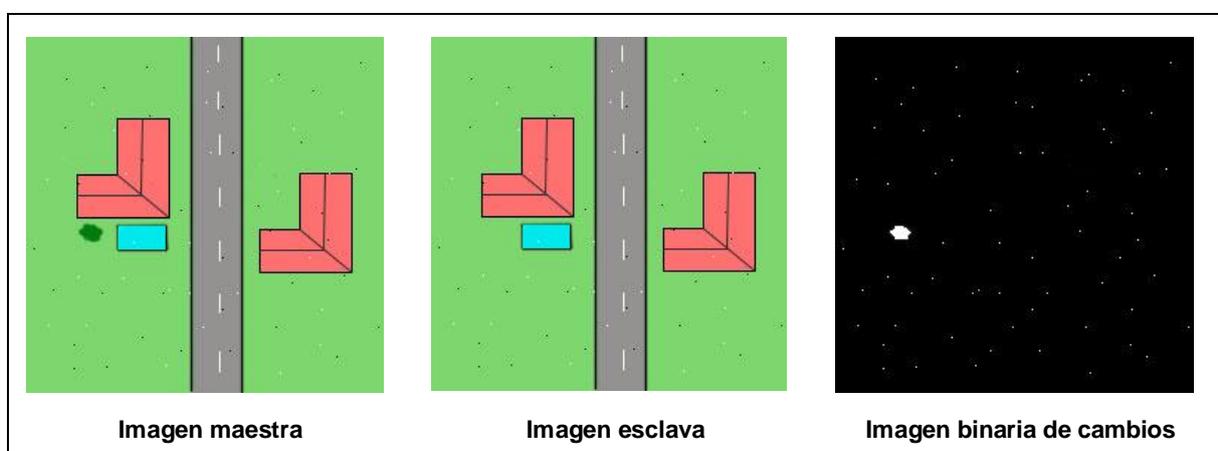


Figura 6-12. Cambios pequeños y ruido.

Otros cambios no deseados aparecen por los errores en el proceso de registro de las imágenes. Por la naturaleza de las imágenes de satélite, el registro no es exacto y siempre existen errores que producen cambios por el mal alineamiento de los distintos elementos en la escena. Estos cambios pueden ser eliminados si se realiza una corrección geométrica local. Con este fin, se hace uso de la correlación cruzada normalizada o la información mutua que vimos en el capítulo de correcciones geométricas, para mover una ventana alrededor del cambio hasta encontrar el desplazamiento que corrige el error de registro y así proceder a detectar de nuevo el cambio utilizando el cambio original como máscara.

En la Figura 6-13 se muestra el procedimiento seguido en una corrección geométrica local para el cambio marcado en rojo. En primer lugar, se crea una ventana en la imagen esclava alrededor del cambio. A continuación, se va desplazando por esta ventana el fragmento de la imagen maestra que contiene el cambio, a la vez que se calcula la función de similitud que se haya seleccionado (NCC o ECC). Por último, se detecta el cambio de nuevo usando el cambio original como máscara y el desplazamiento que haya devuelto el valor máximo para la función de similitud. Al usar el cambio como máscara evitamos la aparición de nuevos cambios por introducir nuevos errores de registro, esta vez en el vecindario del cambio. En el ejemplo que nos ocupa, tras la corrección local el cambio seguirá existiendo y será marcado como cambio definitivo.

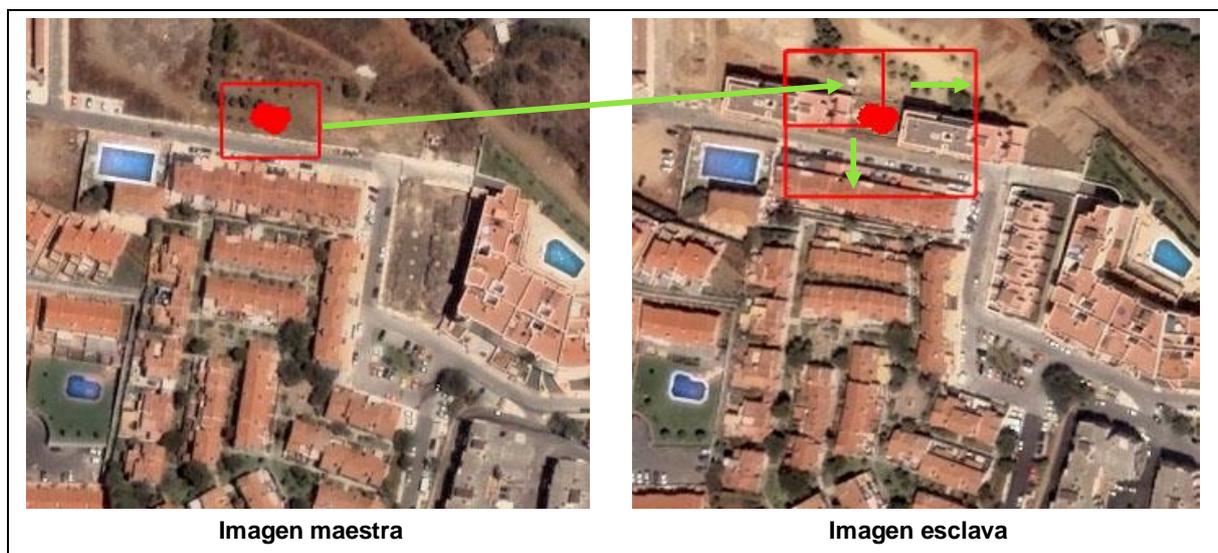


Figura 6-13. Venta para la corrección geométrica local.

En zonas homogéneas o donde los píxeles contienen poca información, tanto la correlación cruzada como la información mutua pueden proporcionar un desplazamiento incorrecto. Para evitar este tipo de correcciones locales erróneas, se aplica de nuevo el algoritmo RANSAC tal como hacíamos en el registro de imágenes. Recordemos que el algoritmo RANSAC toma dos muestras de valores y un modelo matemático, y descarta aquellas muestras que no se ajustan al modelo. El primer conjunto de muestras está formado por los pares (x_c, y_c) donde x_c e y_c son las coordenadas x e y de los centroides del cambio en estudio y sus vecinos hasta un

radio especificado como parámetro. El segundo conjunto lo constituyen los centroides con el desplazamiento calculado para cada cambio $(x_c + \Delta x, y_c + \Delta y)$. De esta forma se puede observar si la corrección local calculada está en consenso con sus vecinos, evitando corregir aquellos cambios de los que se ha obtenido un desplazamiento no fiable.

En la Figura 6-14 se puede observar un cambio, sus vecinos (contenidos en un círculo de radio 150 píxeles) y las direcciones de los desplazamientos calculados para su corrección local. En este ejemplo, el cambio (en el centro del círculo) y sus vecinos requieren un desplazamiento similar por lo que RANSAC determinaría que existe un consenso y se procedería a su corrección. Otros cambios, como el indicado por la flecha, serían marcados como *outliers* por RANSAC, por tanto, no se procedería a corregirlos localmente.

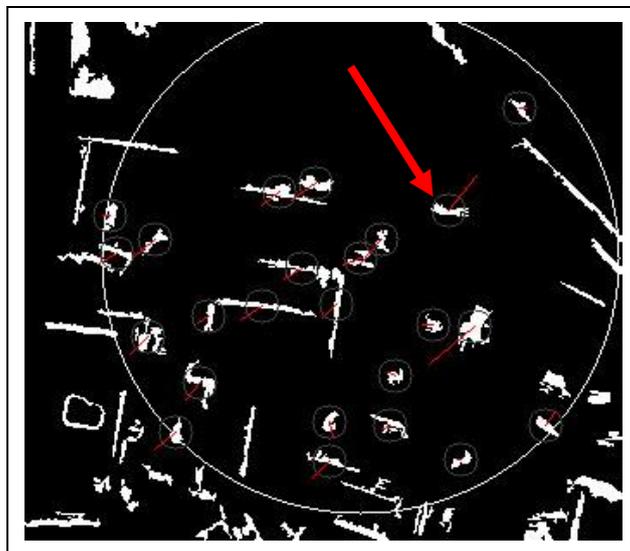


Figura 6-14. Consenso con vecinos en corrección geométrica local.

A continuación veremos distintos falsos positivos que pueden ser eliminados por este proceso de ajuste geométrico local. La Figura 6-15 representa una carretera que no ha sido alineada correctamente y es marcada dos veces como cambio al aplicar cualquiera de los métodos vistos. Tras la corrección local, la carretera quedará alineada y al realizar la detección de nuevo ambos cambios desaparecerán.

En el ejemplo de la Figura 6-16 ocurre algo similar, los edificios no están bien alineados lo que produce cambios en las zonas donde no se solapan. Estos cambios son muy habituales, algunos de ellos serán eliminados en el filtrado por área por resultar demasiado pequeños y otros podrán ser corregidos con el ajuste geométrico local. Al alinear los edificios estos cambios desaparecerán.

Otros cambios no deseados que podemos encontrar frecuentemente son los provocados por sombras. Las imágenes de satélite no siempre son adquiridas con la misma elevación del Sol, esto provoca sombras de distintos tamaños y en distintas direcciones. Las sombras como vimos en un capítulo anterior pueden ser detectadas,

incluso algunas de ellas restauradas. De esta forma eliminamos este tipo de cambios pues pueden ser filtrados antes de la detección. En el caso de que el ajuste geométrico no sea el adecuado la información de sombras no estaría bien localizada y debemos esperar a realizar el ajuste geométrico local para eliminarlas. En ese caso una vez obtenidos los cambios, se eliminan aquellos que son marcados como sombras.

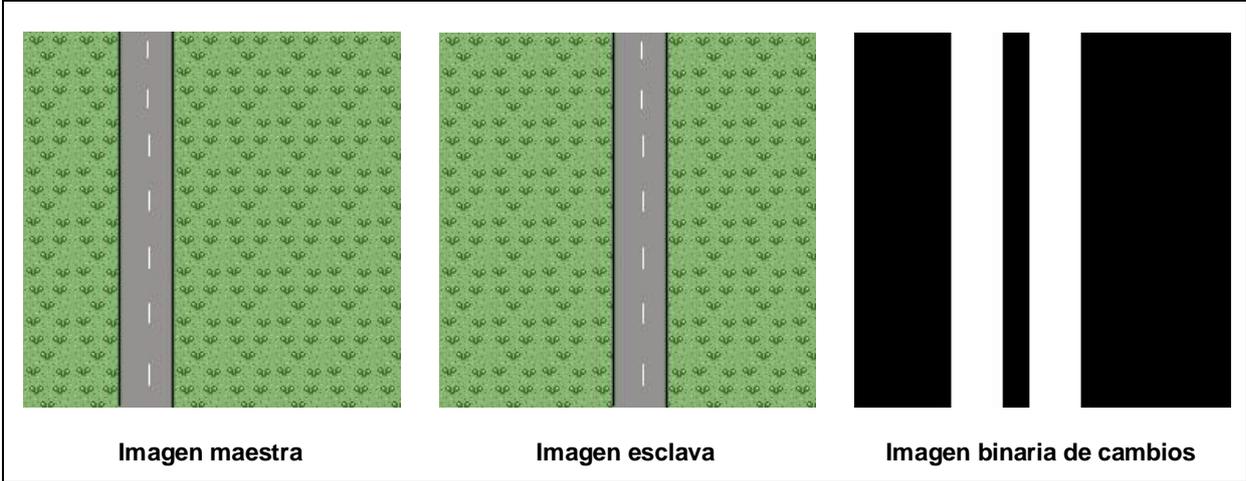


Figura 6-15. Cambios debidos a un elemento mal alineado.

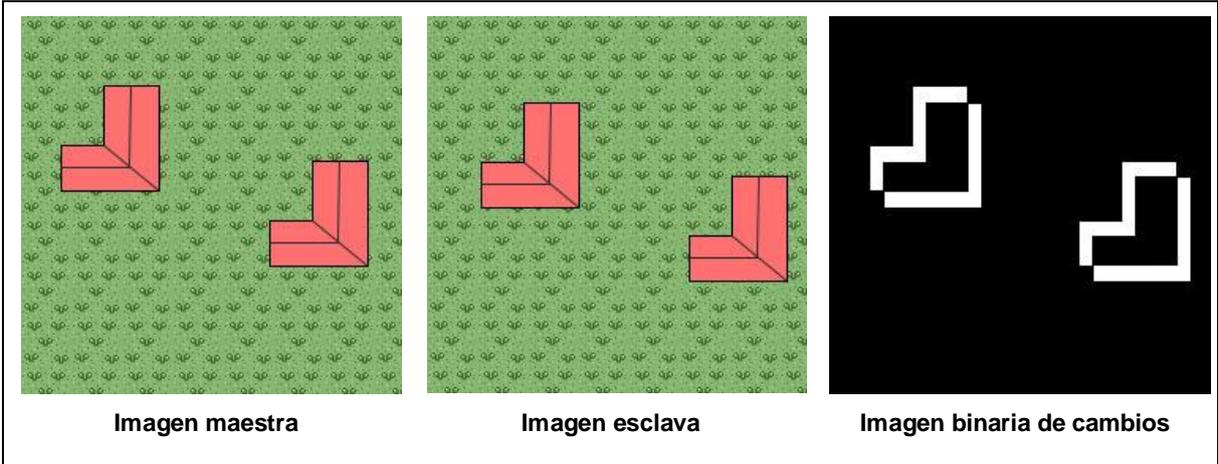


Figura 6-16. Cambios debidos a elementos ligeramente mal alineados.

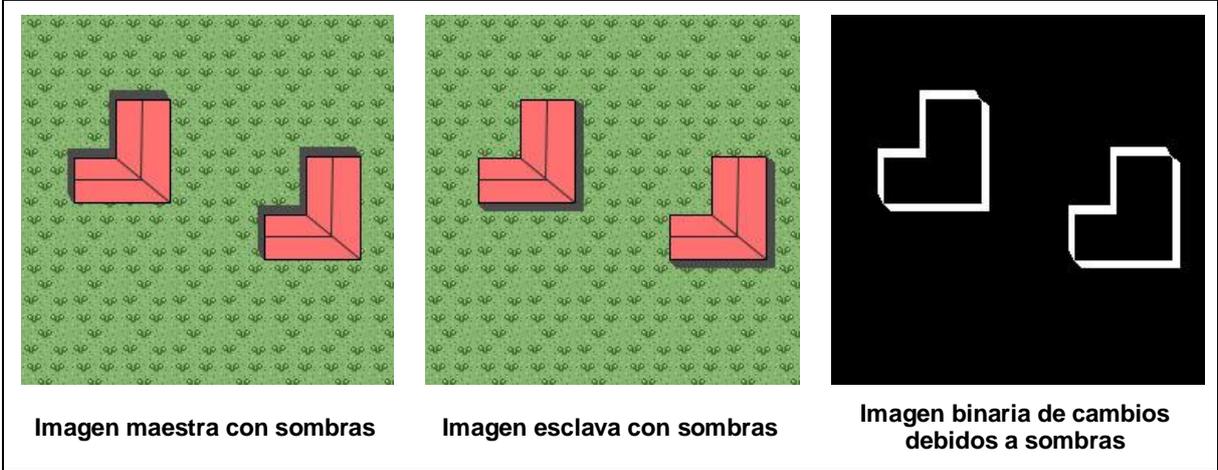


Figura 6-17. Cambios debido a sombras.

Independientemente de la precisión del registro de las imágenes, del ajuste radiométrico, del método de detección elegido, de la selección de un valor umbral adecuado y el filtrado de cambios posterior, hay ciertos cambios que podrían considerarse falsos positivos y que no pueden ser eliminados automáticamente y, por tanto, deberá ser el analista quien los estudie y clasifique.

Si trazamos una línea recta que una el sensor que toma las imágenes con el centro de la Tierra, llamamos nadir al punto de la línea que interseca con la superficie terrestre. Lo ideal para la detección de cambios es que las imágenes de la zona a estudiar sean adquiridas con el ángulo menor posible desde el nadir. Las imágenes adquiridas en vuelos solucionan este inconveniente tomando varias imágenes y seleccionando la parte central de cada una de ellas, asegurando que el ángulo desde el nadir no supere un cierto valor. En el caso de los satélites, esto no es posible y cuando se adquieren imágenes de una zona de la Tierra, se toman desde diferentes ángulos en función de la órbita más cercana a la zona de estudio.

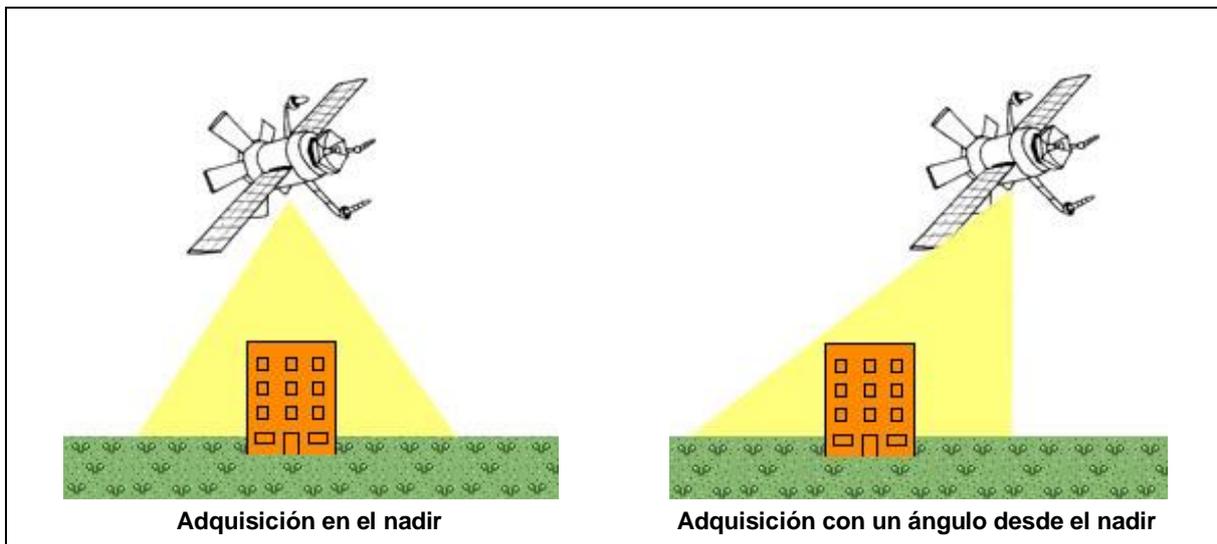


Figura 6-18. Ángulo de adquisición de las imágenes.

Esta diferencia en el ángulo de adquisición puede introducir nuevos elementos en la escena que sólo son visibles desde ciertos puntos de vista. Como podemos ver en la imagen anterior, en el primer caso, el edificio se encuentra en el nadir y en la imagen obtenida por el sensor podrá observarse la terraza del edificio y el suelo de alrededor; en el segundo, la imagen es capturada desde un ángulo diferente y en ella aparecerá además del techo y suelo, parte de la fachada del edificio, quedando oculto parte del suelo alrededor.

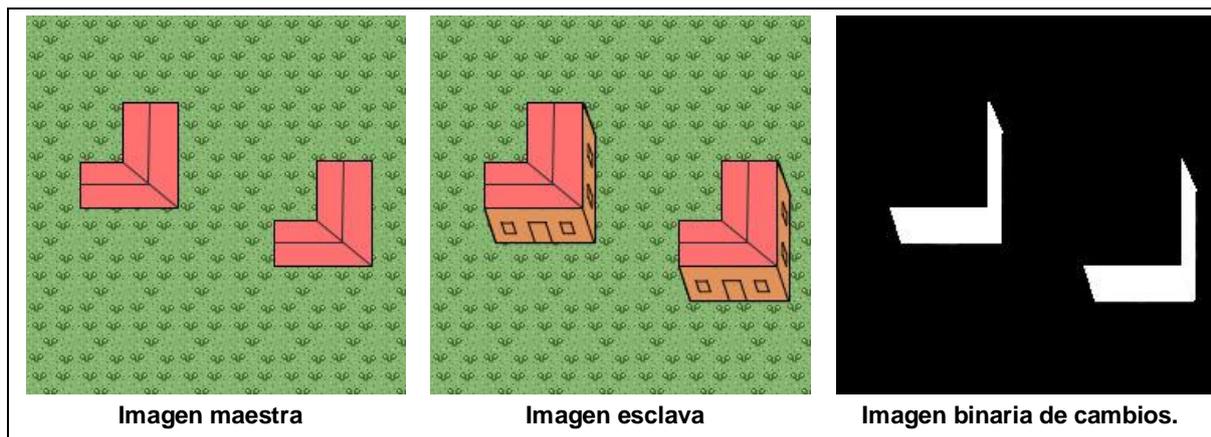


Figura 6-19. Efectos del ángulo de adquisición en los cambios.

Se puede ver en la Figura anterior como diferentes puntos de vista pueden añadir nuevos elementos a la escena, cambios que en la mayoría de los casos no son de interés, ya que no representan un cambio en el uso del suelo o en la estructura de los diferentes elementos, pero que no pueden ser eliminados automáticamente y requieren de la experiencia del analista para su clasificación.

Los falsos positivos no son el único problema a la hora de detectar cambios, también los falsos negativos que requerirán también de la pericia del analista para ser localizados. Los algoritmos de detección más utilizados se basan en la información de color de los distintos elementos de la escena. A veces ocurre que un elemento cambia pero el color es similar en ambas escenas y éste por tanto no es identificado como cambio al no superar el valor umbral que se ha seleccionado.

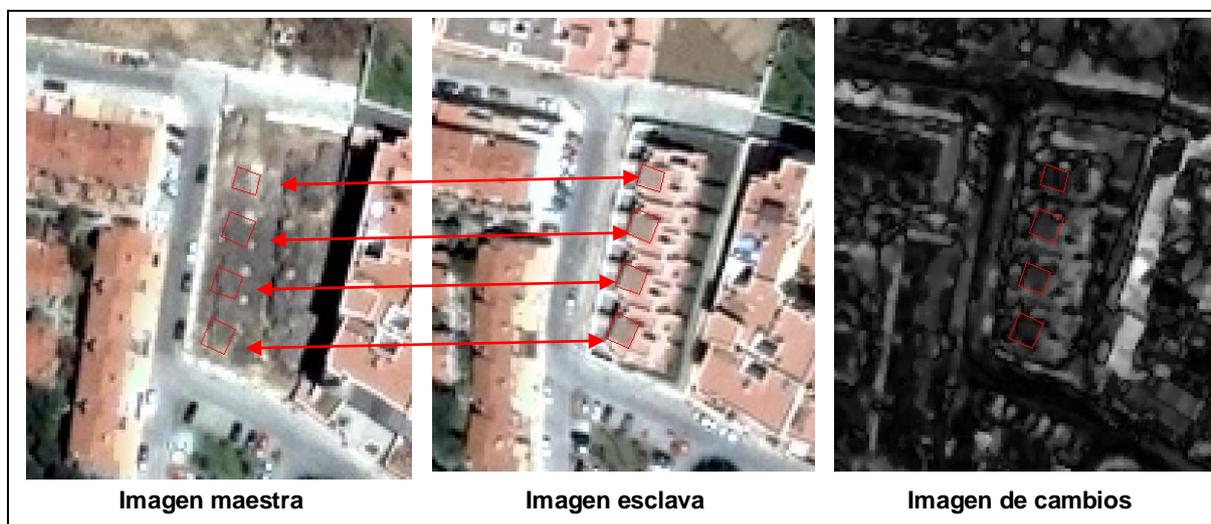


Figura 6-20. Falsos negativos por píxeles con color muy similar.

En la Figura anterior vemos como parte del techo de los nuevos edificios, marcados en rojo, tienen un color muy similar al suelo antes de la construcción. Esto provoca que la imagen de diferencias presente una respuesta baja en estas zonas y al umbralizar para obtener la imagen binaria de cambios, estos cambios reales no sean tenidos en cuenta y desaparezcan de los cambios finales. Usar umbralización con histéresis puede ayudar a prevenir estos falsos negativos.

No a todos los cambios se les puede aplicar una corrección geométrica local ya que algunos de ellos se encuentran cerca del borde de la imagen y, por tanto, no se puede crear una ventana alrededor para encontrar el mejor ajuste. Además, si aplicamos todas las técnicas de filtrado comentadas en el apartado anterior, no podremos diferenciar que cambios han sido corregidos localmente o no. Por este motivo, durante el filtrado se realiza la vectorización de los cambios y se guarda un valor numérico en uno de los campos que indique el tipo de filtrado aplicado:

1. No filtrados: Si no se ha efectuado corrección geométrica local.
2. Borde: No se ha podido corregir por situarse en el borde de la imagen.
3. Indeterminado: No hay suficientes vecinos para llegar a un consenso.
4. Corregido: Se ha efectuado una corrección geométrica local.
5. No corregido: No hay consenso con los vecinos y no se ha corregido.

6.5 Clasificación de cambios

Ya hemos comentado que los cambios que se producen entre dos imágenes aéreas pueden ser de naturaleza muy distinta. En la fase de filtrado, pueden ser descartados muchos de ellos que no sean de interés para el estudio. Los cambios que pasen este filtrado, también se deberán a distintos fenómenos y un proceso posterior de clasificación será necesario. Esta clasificación consistirá en agrupar los cambios en grupos de cambios con características similares.

La clasificación puede ser de forma no supervisada, si es el mismo sistema quien decide a qué clase pertenece cada cambio, o de una forma supervisada, donde es el analista de los cambios quien toma las decisiones o selecciona patrones de entrenamiento para ayudar al sistema. Veremos dos técnicas de clasificación no supervisada, una basada en CVA (Change Vector Analysis) y otra en el algoritmo K-Means; y presentaremos una herramienta para la clasificación manual de los cambios.

6.5.1 Clasificación por el análisis del vector de cambios

Ya estudiamos la técnica CVA en el apartado 6.1.1. Con ella identificábamos los cambios aplicando un valor umbral al módulo del vector de cambios. Pero además de su magnitud, de cada vector podemos obtener también su dirección que nos será de utilidad para la clasificación.

Sean D_1 y D_2 las dos fechas de adquisición de las imágenes, cada píxel vendrá dado como una tupla de n valores para cada fecha $\{(banda_1, banda_2, \dots, banda_n)_{D_1}, (banda_1, banda_2, \dots, banda_n)_{D_2}\}$, donde n es el número de bandas. En la Figura 6-1, veíamos la representación de un vector de cambios para una imagen de tan sólo dos bandas. Ya que los imágenes están normalizadas radiométricamente, podemos transformar estos vectores al *espacio de cambio radiométrico*, donde los píxeles sin cambio quedan dentro de un círculo centrado en el origen y los vectores de cambio pueden representarse como en la Figura 6-21. En esta representación, todos los vectores tienen como origen el origen de coordenadas, los ejes representan

diferencias en el valor de las bandas, y la comparación de magnitudes y direcciones de los vectores se puede realizar fácilmente.

Una forma simple de clasificar estos vectores consiste en dividir el espacio en sectores separados por los ejes de coordenadas (cuatro sectores en el caso de dos bandas). Cada sector agrupará los vectores de cambio en función de si la diferencia de los píxeles en cada banda ha sido positiva o negativa. Así en el ejemplo presentado, los píxeles con vectores de cambios A y C, formarían parte de la misma clase con una diferencia positiva en la banda 1 y una negativa en la banda 2, mientras que el píxel con vector B, formaría parte de la clase en la que la diferencia ha sido positiva para ambas bandas.

Esto mismo puede extrapolarse a un mayor número de bandas n , y será este número el que determine el número de clases de cambios que se obtendrá en el procedimiento:

$$N^{\circ} \text{ clases} = 2^n$$

En la práctica, algunas de las clases pueden quedar desiertas tras la clasificación y el número de clases total se ve reducido. Para imágenes con las tres bandas RGB tendríamos una clasificación de máximo $2^3 = 8$ clases, lo que puede resultar a veces excesivo. Para evitar esto siempre podremos recurrir a algunas de las transformaciones que hemos visto, como PCA, para obtener las dos bandas más significativas (reduciendo a su vez la correlación entre las bandas) y obtener clasificaciones de un menor número de clases.

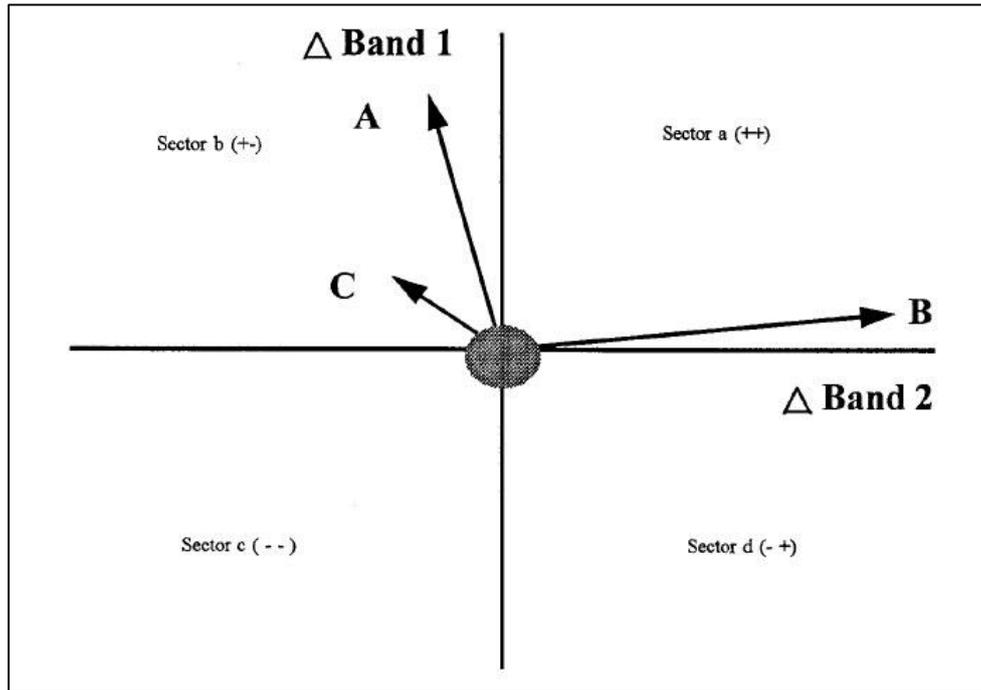


Figura 6-21. Representación del vector de cambios en el espacio de cambio radiométrico.

Para que todos los píxeles de un cambio sean agrupados en la misma clase, debemos buscar un valor que los represente, como pueden ser las medidas estadísticas de la media, moda o mediana, siendo la media la elegida para nuestra implementación. Por tanto, para cada cambio, se procede calculando la media de los píxeles en cada banda y para cada una de las dos imágenes empleadas. El vector de cambios vendrá dado por la diferencia de los valores medios de cada banda. Es decir, para cada cambio C_i , se calcula dos tuplas con las medias de sus píxeles $\bar{x}_i^{D_j}$ donde i indica el número de banda y D_j es la fecha de adquisición $\{(\bar{x}_1^{D_1}, \bar{x}_2^{D_1}, \dots, \bar{x}_n^{D_1}), (\bar{x}_1^{D_2}, \bar{x}_2^{D_2}, \dots, \bar{x}_n^{D_2})\}$. El vector de cambios vendrá dado entonces por la expresión:

$$\vec{C}_i = (\bar{x}_1^{D_2} - \bar{x}_1^{D_1}, \bar{x}_2^{D_2} - \bar{x}_2^{D_1}, \dots, \bar{x}_n^{D_2} - \bar{x}_n^{D_1})$$

En la Figura 6-22 se puede observar el resultado de aplicar esta técnica a un conjunto de cambios delimitados manualmente. En este ejemplo, primero se obtienen las dos componentes principales de las imágenes del experimento 2 de correcciones geométricas. Estas imágenes se clasifican según el método propuesto y nos ofrece la clasificación de la Figura 6-22. Recordemos que para 2 bandas obtenemos como máximo $2^2 = 4$ clases de cambios, cada una representada por un color. En este ejemplo, la clasificación obtenida agrupa en el color azul, cambios debidos a nuevas construcciones; con el color amarillo, nuevas zonas asfaltadas; con el color rojo, movimientos de tierras y nuevos edificios con techos de color rojizo; y con el color verde, zonas donde se ha producido un oscurecimiento en las imágenes, como la recuperación de vegetación seca en zonas no construidas.

Curiosamente, la realización del experimento directamente con las imágenes RGB no ofrece un resultado de interés pues de las $2^3 = 8$ clases posibles, todos los cambios son agrupados en dos de ellas, quedando las otras seis vacías. Esto se debe a la correlación existente entre las bandas RGB que es eliminada con el análisis de componentes principales.



Figura 6-22. Ejemplo de clasificación por dirección del vector de cambios.

6.5.2 Clasificación no supervisada por algoritmo K-Means

Uno de los algoritmos más conocidos para la clasificación no supervisada es el de K-Means. Dados un conjunto de muestras y un conjunto de k centroides, se trata de agrupar las muestras en k conjuntos de tal forma que las distancias a sus respectivos centroides sean las mínimas. Matemáticamente y en el caso que nos ocupa, sean (C_1, C_2, \dots, C_m) el conjunto de vectores de cambio de los m cambios a clasificar en k conjuntos, se trata de encontrar la partición $S = \{S_1, S_2, \dots, S_k\}$ tal que:

$$\arg \min_S \sum_{i=1}^k \sum_{C_j \in S_i} \|C_j - \mu_i\|^2$$

donde μ_i es el vector media de las muestras en S_i .

El algoritmo más común usado consiste en seleccionar un conjunto de semillas o centroides iniciales (uno por cada clase) y realizar varias iteraciones hasta conseguir la partición óptima. Primero se asigna cada vector de cambios a la clase cuya distancia al centroide es menor. A continuación, se calcula los nuevos centroides, como la

media de los vectores de cada grupo. Estas operaciones se realizan hasta que ningún vector de cambio cambia de clase. En caso de no pasar las semillas como parámetro, estas son seleccionadas aleatoriamente de la muestra. La selección de las semillas es un factor decisivo en la clasificación, por eso en algunas ocasiones será importante seleccionarlas manualmente ya que el resultado obtenido depende de las mismas.

K-Means permite clasificar vectores de distinta naturaleza. En nuestra implementación se proponen tres tipos diferentes:

- 1) Diferencia de las medias de los niveles de gris de cada cambio.

Sean $(\bar{r}_1, \bar{g}_1, \bar{b}_1)$, y $(\bar{r}_2, \bar{g}_2, \bar{b}_2)$ las medias de los niveles de gris de un cambio en la imagen menos reciente y la más actual respectivamente. El vector de características a clasificar vendrá dado por:

$$F_c = (\bar{r}_2 - \bar{r}_1, \bar{g}_2 - \bar{g}_1, \bar{b}_2 - \bar{b}_1)$$

- 2) Vector de medias de los niveles de gris de cada banda.

Dadas las definiciones anteriores, el vector de características en este caso será dado por la siguiente expresión:

$$F_c = (\bar{r}_1, \bar{g}_1, \bar{b}_1, \bar{r}_2, \bar{g}_2, \bar{b}_2)$$

- 3) Vector de los momentos de los cambios.

En este caso, es aplicable sólo a imágenes en niveles de gris y el vector de características usado para la clasificación estará constituido por los momentos del cambio en las dos imágenes:

$$F_c = (m_1, v_1, s_1, u_1, m_2, v_2, s_2, u_2)$$

donde m_i es la media de los niveles de gris, v_i es la varianza, s_i el sesgo, y u_i es la medida de la forma del histograma conocida como curtosis. Siendo x_j el valor digital del píxel j de un cambio, los valores comentados se obtienen de las expresiones:

$$m_i = \frac{1}{N} \sum_{j=0}^{N-1} x_j$$

$$v_i = \frac{1}{N-1} \sum_{j=0}^{N-1} (x_j - m_i)^2$$

$$s_i = \frac{1}{N} \sum_{j=0}^{N-1} \left(\frac{x_j - m_i}{\sqrt{v_i}} \right)^3$$

$$c_i = \frac{1}{N} \sum_{j=0}^{N-1} \left(\frac{x_j - m_i}{\sqrt{v_i}} \right)^4 - 3$$

6.5.3 Clasificación manual de los cambios

La clasificación no supervisada, si no se dispone de bandas suficientes (como para discriminar cambios por su naturaleza), o las bandas están correlacionadas como hemos visto, difícilmente nos ofrezca una clasificación perfecta y generalmente será necesaria una revisión posterior por el analista encargado de la detección de cambios. Para ello se ha implementado una herramienta con la que podemos desplazarnos por los cambios sobre una imagen uno a uno y modificar aquellos cambios mal clasificados o incluso asignarles una clase si no la tenían todavía.

6.6 Futuras mejoras

Los métodos algebraicos propuestos buscan obtener cambios radiométricos, es decir, en el nivel de gris o color de los píxeles. En muchos casos, se producen cambios claramente identificables por el ojo humano pero que no han supuesto una diferencia radiométrica suficientemente grande para que el cambio sea detectado. Un ejemplo de esto lo vimos en el apartado de filtrado de cambios. Por este motivo, se propone como futura mejora desarrollar nuevas técnicas no basadas tan sólo en la radiometría de los píxeles, sino que añada información de textura o incluso detección de cambios orientada a objetos. El método de parcelas comentado posteriormente puede ser un buen comienzo, seleccionando como vector de características medidas de la textura como la varianza, sesgo y curtosis, medidas que hemos empleado en la clasificación no supervisada.

Las técnicas de detección estudiadas están basadas en píxeles, es decir, se determina si un píxel es de cambio o no de forma individual e independientemente del resto de píxeles de la imagen. Pero en las imágenes de muy alta resolución, debido a la cantidad de detalles y objetos que se representan en ellas, un píxel es tan sólo parte de un conjunto de píxeles de características similares. Por este motivo, y para tener en cuenta esta peculiaridad, se han propuesto métodos basados en *parcelas* como en (Bruzzone & Prieto, 2000).

Brevemente, la idea de la detección basada en parcelas consiste en lo siguiente: 1) segmentar las imágenes individualmente, 2) obtener las parcelas intersecando las dos imágenes segmentadas (véase Figura 6-23), 3) obtener un vector de características de cada parcela para cada imagen, por ejemplo media y entropía u otra medida de la distribución espacial de los píxeles, 4) obtener imagen de cambios como la magnitud del vector diferencia de los vectores de características, 5) umbralizar la imagen de cambios en niveles de gris para obtener la imagen binaria de cambios.

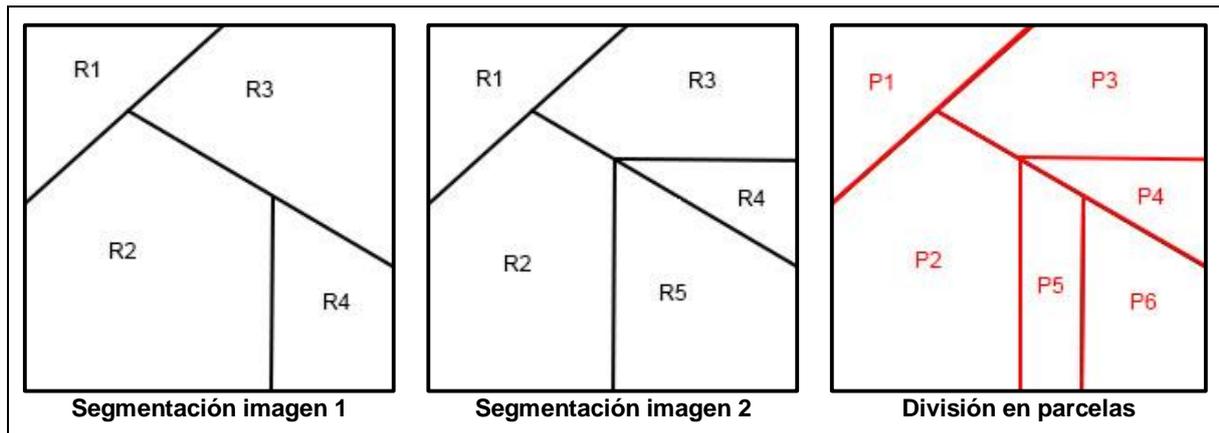


Figura 6-23. Ejemplo de división en parcelas para la detección de cambios.

La detección basada en parcelas puede realizarse en varios niveles (Bovolo, 2009), en la que se modifica el parámetro que controla la segmentación y obteniendo un vector de características mayor. La versión multinivel presenta una mayor precisión en la detección que la de un único nivel o la basada en píxeles.

Las técnicas de umbralización que hemos visto fueron ideadas para cualquier tipo de imagen y tratan de encontrar el umbral óptimo para su binarización. Se propone como futura mejora el estudio de nuevos algoritmos de búsqueda del valor umbral que tengan en cuenta las peculiaridades de las imágenes de cambio.

Se podrían emplear imágenes gradiente de la imagen de cambios con el propósito de detectar fuertes transiciones en los niveles de gris que pueda indicarnos la presencia de un cambio. Y para delimitarlos correctamente, usar la imagen de bordes y crecimiento de regiones.

Unos de los principales inconvenientes del filtrado por corrección geométrica local es que es excesivamente lento al tener que ir desplazando la ventana alrededor de cada cambio para alcanzar el máximo de la medida de similitud que se emplee (MI o NCC). Se propone como futura mejora emplear técnicas de optimización para acelerar esta búsqueda del valor máximo. Al hacerlo, se deberá tener en cuenta la posible aparición de artefactos en la función de similitud cuando la corrección local es a nivel sub-píxel. En (Briechle & Hanebeck, 2001) nos proponen una optimización para la búsqueda por correlación cruzada normalizada.

Capítulo 7

Conclusiones y resumen

La detección de cambios es un área de la teledetección en constante investigación debido a los cambios que se producen en las tecnologías de adquisición de imágenes, la mayor resolución de las mismas, las nuevas necesidades de observación de la Tierra que van surgiendo y la dificultad de crear métodos de detección de cambios universales que sean válidos para cualquier tipo de cambio.

La gran cantidad de información que nos ofrecen los satélites de hoy en día hace necesario el uso de herramientas automáticas que localicen los cambios sin necesidad de expertos analistas ni de información previa del terreno. Los métodos supervisados son menos interesantes por el aumento de costes, en términos de tiempo y esfuerzo, que supone necesitar información obtenida por fotointerpretación o trabajo de campo. Sin embargo, proporcionan mayor información sobre la naturaleza y tipo de cambio, detalles que no se pueden obtener con los métodos no supervisados.

Al igual que un observador podría identificar como cambio lo que para otros no lo es, dependiendo de sus conocimientos o su percepción visual, de manera similar podría ocurrir en la detección automática de cambios. Las herramientas desarrolladas para trabajar con un tipo de imágenes concreto o para detectar cambios de ciertas características (como la construcción de nuevos edificios), puede no ser útil para otro tipo de imágenes o cambios. De ahí que surja la necesidad de seguir un marco de trabajo, como el comentado en el apartado 2.9, que guíe el proceso de diseño del sistema adecuado de detección de cambios para cada estudio que se realice.

En las imágenes de muy alta resolución que ofrecen los satélites más recientes, aparece una gran cantidad de detalles que no eran perceptibles en las de menor resolución. Esto hace más compleja la detección de cambios y que sea necesaria en cada estudio una definición clara del concepto de cambio y los tipos de cambio de interés.

La constante evolución de las distintas técnicas así como el gran número de procesos auxiliares que le suelen acompañar, hacen que un primer acercamiento a la

detección de cambios sea una tarea difícil y lleve a muchos nuevos usuarios de estas tecnologías a perderse entre artículos, algoritmos y técnicas centradas en una única tarea y tipo de imágenes concreto.

En ese sentido, a lo largo del proyecto han surgido dificultades que han llevado a perder mucho tiempo en estudiar e implementar técnicas muy específicas fuera de los objetivos del mismo, o incluso, a pretender elaborar herramientas válidas para todo tipo de cambios e imágenes.

El tiempo empleado en el desarrollo de este proyecto y la experiencia adquirida elaborando mapas de cambios en imágenes de alta resolución Quickbird, han sentado las bases y conocimientos básicos necesarios para llevarlo a su etapa final, dando lugar a un conjunto de herramientas básicas para cualquier detección de cambios, pensando en su reutilización en estudios posteriores y más específicos, y a esta memoria que servirá como manual de referencia para iniciarse en el estudio de este área de la teledetección.

La mayoría de algoritmos implementados para este proyecto se han escrito en forma de funciones pensando en su reutilización. En aquellos donde ha sido posible se devuelve el descriptor del fichero resultado de la operación. Esto permitirá crear nuevos flujos de trabajo encadenando las funciones y la fácil creación de nuevos *scripts* en IDL. Para usuarios no programadores, las funciones también están disponibles desde el menú creado para el proyecto en la barra de menús de ENVI.

En los próximos párrafos, haremos una revisión a modo de resumen de todas las fases de la detección de cambios estudiadas en este proyecto, a la vez que comentaremos las conclusiones obtenidas en cada una de ellas con el fin de dar una visión global de todo el procedimiento.

Todo estudio de detección de cambios debe comenzar con una definición clara de cuáles son los cambios de interés para el usuario final y qué tipo de imágenes se van a emplear para poder determinar qué técnicas o algoritmos son los más adecuados. Generalmente, antes de proceder a identificar los cambios, se realizará un procesado previo que prepare las imágenes y ayude a precisar e identificar mejor los cambios. Este procesamiento previo incluye: una corrección radiométrica para normalizar el color de las imágenes en un mismo sistema métrico y hacerlas comparables; y una corrección geométrica en la que las imágenes son alineadas para que los píxeles con mismas coordenadas representen la misma zona de la superficie terrestre.

Para la corrección radiométrica (Capítulo 3) se han estudiado algunas transformaciones lineales, la especificación de histograma y un método adaptativo con interpolación bilineal. Vimos que la relación entre los píxeles de dos imágenes no suele ser lineal y que, por tanto, los métodos de transformación lineales no suelen ser los más adecuados para una corrección radiométrica. De los dos métodos no lineales

estudiados, el método adaptativo bilineal es el que mejor resultado ha obtenido en las pruebas, aunque una especificación de histograma puede ser suficiente en muchos casos. Todos los métodos se ven afectados, como es de esperar, por los cambios ocurridos en las imágenes. Enmascarar los píxeles de cambio que sean reconocibles a priori, como las sombras, mejorará los resultados de la corrección.

En las pruebas realizadas, se hizo uso del error cuadrático medio en los niveles de gris para comparar el desempeño de las distintas técnicas. El método de regresión simple SR y el adaptativo bilineal son los que menor error presentaron. Debemos observar que aunque el método lineal SR nos lleva a un error bajo, en el caso de imágenes afectadas por muchos cambios, esta transformación puede llevar a una pérdida de contraste y, por tanto, de información radiométrica, debido a la no linealidad de la relación entre los píxeles de las imágenes, lo que supone un efecto indeseable para una adecuada detección de cambios. Este método también tiene el inconveniente de requerir una corrección geométrica precisa previa, algo que generalmente es difícil de alcanzar. Por otro lado, comentar que no se debe caer en la idea de reducir demasiado la ventana del método adaptativo pues, aunque reduce la medida de error, provoca grandes diferencias entre las distintas ventanas y una iluminación no uniforme, otro efecto indeseado teniendo en cuenta que las condiciones de iluminación en una imagen de satélite suelen ser las mismas para todos sus píxeles.

En el Capítulo 4 estudiamos las correcciones geométricas. Estas correcciones constan de cuatro pasos: detección de características, emparejamiento o *matching* de características, estimación del modelo de transformación, y remuestreo de la imagen y transformación. Los dos primeros pasos son los más decisivos de la corrección y son los que hemos estudiado más detenidamente. El sistema se ha diseñado de tal forma que se pueda seleccionar una técnica de *matching* independientemente del algoritmo de detección de características seleccionado. Para los dos últimos pasos, hemos empleado herramientas de transformación y remuestreo ya disponibles en ENVI.

En la detección de características, se han implementado los algoritmos automáticos de extracción de características siguientes: esquinas de Harris, autovalores mínimos de la matriz Hessiana, características SIFT y características SURF (véase apartado 4.1 para conocer detalles de los diferentes algoritmos).

Respecto al *matching* de características, se han implementado las siguientes opciones: *matching* por correlación cruzada normalizada (NCC), búsqueda por NCC, *matching* por información mutua (MI), búsqueda por MI, algoritmo de tracking Lucas-Kanade, *matching* de características SIFT y *matching* de características SURF (todos ellos estudiados de forma detallada en el apartado 4.2).

Hemos puesto a prueba los algoritmos por NCC y MI con puntos de Harris, el tracking con puntos de Harris y mínimos autovalores de la Hessiana, y *matching* por características SIFT y SURF. Llegar a una conclusión de cuál de las técnicas ofrece

un mejor resultado es una tarea muy compleja. Para compararlos hemos hecho uso del error cuadrático medio de los niveles de gris de los píxeles. Sin embargo, no basta con ejecutar cada uno de los algoritmos y calcular el error, pues el desempeño de cada uno de ellos depende de los parámetros que se hayan seleccionado. Estudiando los algoritmos de matching propuestos por separado, hemos determinado en cada caso cuál de los parámetros influye más en su ejecución. Una vez determinado el que hemos llamado parámetro crítico del algoritmo, se ha procedido a obtener el error para varias ejecuciones del mismo en los que se ha ido modificando el valor del parámetro, seleccionando en cada caso el valor que menor error presentaba. Tras obtener el valor de los parámetros y la mejor transformación polinomial posible, hemos comparado todos los algoritmos para intentar extraer algunas conclusiones.

Los métodos SIFT y SURF son los más rápidos, pero han mostrado malos resultados en cuanto al porcentaje de emparejamientos obtenidos respecto al número de puntos candidatos. También, son los que obtienen un mayor porcentaje de *outliers*. Esto último, se debe a la poca robustez de estos métodos ante cambios en las condiciones de adquisición. Los métodos basados en correlación e información mutua (especialmente los que usan ventana de búsqueda) mejoran estos porcentajes pero son de lenta (o muy lenta) ejecución. El tracking Lucas-Kanade obtiene muy buenos resultados en tiempo de ejecución, un alto porcentaje de emparejamientos y un porcentaje bajo de *outliers*, destacando frente al resto de alternativas.

Centrándonos en el error cuadrático medio de los niveles de gris de los píxeles y comparando las técnicas con el valor para los parámetros críticos que más reducía el error, pudimos ver que la técnica de correlación con búsqueda obtenía muy buen resultado en las tres pruebas realizadas. Le sigue la técnica de tracking, que en todas las pruebas obtuvo un error más bajo que SIFT, SURF, y las técnicas NCC y MI sin ventana de búsqueda. Teniendo en cuenta que el tiempo de ejecución de los algoritmos con ventana de búsqueda son excesivamente altos y que obtener el mejor valor para los parámetros críticos es una tarea ardua y costosa, podemos recomendar el uso del tracking en la mayoría de los estudios de detección de cambios, por su rapidez, buena precisión en la búsqueda de emparejamientos y por no depender excesivamente del valor de alguno de sus parámetros (véase apartado 4.5 para un análisis más detallado de los datos obtenidos en las diferentes pruebas).

Todas las técnicas han sido implementadas para que trabajen con imágenes de cualquier tamaño. Pensando en posibles limitaciones de memoria, se puede indicar al sistema que divida las imágenes en fragmentos. Éste se encargará también de procesar los diferentes fragmentos e ir guardando los emparejamientos obtenidos en un mismo fichero de una forma transparente para el usuario.

El Capítulo 5 lo hemos dedicado a estudiar conceptos y técnicas sobre distintos tipos de información que pueden ser de utilidad en una detección de cambios. Dependiendo de los tipos de cambios que hayamos definido como cambios de interés, necesitaremos extraer diferentes capas de información de las imágenes, ya sea para

localizar directamente los cambios de un tipo concreto (por ejemplo cambios en la vegetación), o para cancelar cambios que no son de interés (como pueden ser los provocados por sombras).

En primer lugar, vimos técnicas y espacios de color aptos para la detección de sombras y se han desarrollado métodos automáticos e interactivos para extraerlas. Asimismo, se ha presentado un método de restauración de sombras. Por las pruebas realizadas, aunque visualmente el resultado puede no ser adecuado, comprobamos que la restauración de sombras reduce el porcentaje de píxeles de cambio. La restauración también es útil para hacer visibles cambios reales sobre el terreno que han quedado ocultos tras sombras.

Otro tipo de cambios revisado es el de la presencia de nubes. Hemos comentado un algoritmo de detección nubes, aunque su implementación se deja para futuras mejoras puesto que no suponen normalmente un problema teniendo en cuenta que los satélites comerciales siguen unos criterios de calidad en los que no se puede superar un porcentaje de nubes determinado.

Hemos revisado el concepto de índices de vegetación y algunos de los que dispone ENVI que pueden ser de gran ayuda en aquellos estudios de detección de cambios en los que se vea implicada la presencia de vegetación. En estos casos, es recomendable usar imágenes multiespectrales que dispongan de la banda del infrarrojo cercano.

Finalmente, se ha analizado el uso de máscaras para eliminar información de las imágenes no útil, como pueden ser los píxeles de fondo, y se han explicado los espacios de color que se han usado para los distintos algoritmos de detección de sombras y cambios.

Tras recorrer las etapas de pre-procesamiento de las imágenes, hemos estudiado en el Capítulo 6 técnicas de detección de cambios, la umbralización de imágenes de cambio en niveles de gris, técnicas de filtrado de cambios y algoritmos de clasificación no supervisada.

En un primer apartado, se han revisado la mayoría de técnicas propuestas en artículos de revistas científicas, centrándonos en los métodos algebraicos y transformaciones más habituales, las cuales suelen ser válidas y suficientes para la mayoría de estudios de detección de cambios.

Las técnicas algebraicas basadas en píxeles obtienen una imagen de cambios en niveles de gris que es necesario umbralizar para clasificar todos los píxeles como cambio o no-cambio. En el apartado 6.2 se muestran los cuatro algoritmos automáticos para determinar el valor umbral óptimo que se han implementado para este proyecto. Los hemos puesto a prueba con un par de imágenes y, de los cuatro, la umbralización ISODATA fue la que obtuvo un mayor porcentaje de verdaderos positivos, con un porcentaje de falsos positivos no muy alto y similar al del resto de

algoritmos. El peor resultado en cuanto a verdaderos positivos lo obtiene el método de umbralización por conservación de momentos. Excepto en el método para histograma unimodal, aplicar los algoritmos de forma adaptativa incrementa la tasa de verdaderos positivos con un ligero aumento de la tasa de falsos positivos.

Cuando realizamos una detección de cambios basada en píxeles es conveniente un umbral que proporcione un alto porcentaje de verdaderos positivos. Es preferible que en la detección se localice el mayor número de cambios y eliminar aquellos que no lo son posteriormente, a tener muy pocos falsos positivos y perder información de cambios verdaderos. Por eso a veces, el umbral óptimo seleccionado por alguno de los algoritmos propuestos puede proporcionarnos tasas de verdaderos positivos más bajas de lo deseado. Teniendo en cuenta esto, se ha desarrollado un sistema interactivo de selección del valor umbral en el que el usuario pueda elegir este valor mediante una barra de desplazamiento y observar el resultado en tiempo real antes de aplicarlo. Además, incorpora histéresis, lo que supone una gran mejora a la hora de obtener los cambios, pues permite seleccionar como cambio aquellos píxeles con nivel de gris inferior al umbral que están conectados a cambios ya identificados. A veces, un cambio queda dividido en varios fragmentos por un valor umbral alto. La histéresis también soluciona este inconveniente manteniendo conectados todos los píxeles del cambio.

Por último, dentro del apartado sobre umbralización, hemos presentado la idea de una umbralización multinivel donde el resultado no es una imagen binaria de cambios, sino una clasificación de los píxeles en función de la probabilidad de que sean de cambio o no. Este tipo de imágenes puede resultar útil para determinar rápida y visualmente zonas de riesgo por cambios.

En un tercer apartado del capítulo de detección de cambios, nos hemos preguntado si el espacio de color RGB es el más adecuado para la obtención del mapa de cambios. Usando diferentes espacios y ecuaciones de diferencia de color, y con la ayuda de curvas ROC en las que representamos el porcentaje de verdaderos positivos frente al de falsos positivos, hemos intentado dar respuesta a esta pregunta. En la prueba realizada a un par de imágenes, se ha observado claramente que la detección de cambios es mejor cuando se usan imágenes a color frente a imágenes en niveles de gris. También, pudimos apreciar como otros espacios de color perceptivamente más lineales, como HSV y CIELab, mejoraron los resultados aumentando el porcentaje de verdaderos positivos (véase apartado 6.3).

Hemos comentado la importancia de definir en cada estudio los tipos de cambios de interés y que una forma de afrontar el diseño de un sistema de detección consistía en obtener el máximo de cambios radiométricos posibles y, posteriormente, cancelar aquellos no considerados de interés. Además, en el apartado de umbralización, llegamos a la conclusión de que era preferible obtener un alto porcentaje de verdaderos positivos aunque eso implique obtener demasiados falsos positivos. Por estos motivos, surge la necesidad de una etapa de filtrado posterior a

la detección de los cambios que hemos abordado en el apartado 6.4, y en el cual hemos estudiado diferentes falsos positivos que pueden aparecer y algunas técnicas para eliminarlos.

Se han propuesto tres filtros: un filtrado por área, que elimina cambios pequeños debidos a ruido y permite centrar la detección en cambios de un área determinada; un segundo filtro para intentar corregir cambios debido a ruido en el registro de imágenes, intentando realizar una corrección geométrica local para cada cambio; y, por último, un filtro que permite eliminar los cambios que se deben a sombras cuando éstos no son de interés.

Finalmente, se han propuesto algunos algoritmos no supervisados para la clasificación de cambios. Uno basado en la dirección del vector cambios y otro en el uso de K-means para la clasificación a partir de diferentes vectores de características (véase apartado 6.5). Para que una clasificación no supervisada sea de utilidad, es necesario disponer de las suficientes bandas de información radiométrica, que no exista demasiada correlación entre ellas (como ocurre en el espacio RGB), y que se usen medidas de la textura o distribución de los píxeles de los cambios y no sólo la información del color.

Apéndice A

Detalles de implementación e instalación del software

En este apéndice comentaremos las distintas funciones implementadas e integradas en ENVI para este proyecto. Para conocer los argumentos que recibe cada función se puede consultar la ayuda en HTML que acompaña a esta memoria. Iremos revisando las funciones por fases de la detección de cambios y, por último, el asistente que guía por las distintas fases. Asimismo, comentaremos otras funciones auxiliares que se han implementado y las librerías implementadas en C con ayuda de la librería OpenCV.

A.1 Funciones para la corrección radiométrica.

Con el fin de poder corregir/normalizar imágenes de 11 bits (2048 niveles) de resolución radiométrica con los procedimientos elaborados en IDL/ENVI, es importante cambiar uno de los parámetros de configuración que determina el número máximo de niveles (llamados bins) que serán contados en el cálculo de los histogramas. Por defecto este valor es 256, válido para imágenes de 8 bits. Cuando usamos imágenes de 16 bits (como las imágenes de Quickbird e Ikonos con una resolución radiométrica de 11 bits) los niveles cercanos son agrupados por defecto en 256 grupos o bins.

Los procedimientos para los métodos HZ, MM, MS, SR y EH, están ideados de tal forma que las imágenes puedan estar enmascaradas, y así, eliminar del cálculo de las estadísticas los píxeles que pueden ser identificados como cambio a priori, como las sombras, saturaciones y grandes zonas de cambio reconocibles visualmente. De esta forma, se evita que distorsionen los resultados y, por tanto, el ajuste sea mejor. El método por pseudoinvariantes (PI) y el de conjunto de píxeles oscuros y brillantes (DB) no necesitan máscaras ya que los procesos eliminan los píxeles innecesarios para la normalización.

Otra característica de estos procedimientos es que están preparados para normalizar imágenes de 8 y 16 bits (excepto los procedimientos de los métodos DB y PI). Algunos satélites como Quickbird e Ikonos distribuyen sus imágenes en 8 y 16 bits (en el segundo caso con una resolución radiométrica de 11 bits). El resultado de la normalización dependerá del tipo de imágenes pasadas como entrada. En la versión actual de los procedimientos no son soportadas imágenes de tipo flotante (píxeles en el rango [0,1]). Los métodos DB y PI sólo están preparados para imágenes de 8 bits porque usan algunos umbrales determinados. Para prepararlos para 16 bits se podría escalar los umbrales en el rango [0..2047] (para imágenes con resolución radiométrica de 11 bits).

Las funciones creadas para la corrección radiométrica por los métodos HZ, MM, MS, SR, EH, DB y PI son *haze_correction*, *maximum_minimum_correction*, *ms_normalization*, *sr_normalization*, *histogram_specification*, *dark_bright_normalization* y *PI_normalization*. Todas ellas devuelven el descriptor de fichero de la imagen normalizada (corregida radiométricamente).



Figura A-1. Menú de opciones de correcciones radiométricas.

Todos los procedimientos tras su instalación estarán disponibles en un submenú común llamado “Corrección radiométrica” (Figura A-1). Junto a los ficheros fuente, se encuentra la documentación de cada procedimiento donde se puede comprobar los distintos parámetros que pueden recibir cuando se llaman desde IDL.

A.2 Funciones para la corrección geométrica

Para las correcciones geométricas se han implementado en IDL/ENVI dos funciones y un conjunto de funciones auxiliares. La extracción y *matching* de características puede tener un alto coste computacional dependiendo del número de ellas y el tamaño de las imágenes. Por esto y para poder utilizar bibliotecas de funciones ya existentes, se ha creado una librería dinámica DLL que contiene los cálculos y llama a las funciones necesarias usando el lenguaje C. Esta DLL se llama

GeomCorr.dll y en ella se hace uso de la implementación del método SIFT de Rob Hess (Hess, 2010), y la implementación del método SURF (Bay, Ess, Tuytelaars, & Gool, 2008) y la extracción de esquinas (Shi & Tomasi, 1994) de la librería OpenCV 2.1.0.

La primera función *extract_features* nos permite obtener los distintos puntos característicos o blobs de una imagen mediante Harris, Minimum eigenvalues, SIFT o SURF llamando a la función *extract_harris*, *extract_lk*, *extract_SIFT*, y *extract_SURF*. Devuelve una matriz con las coordenadas de los distintos puntos, incluso los vectores de características en los métodos SIFT y SURF si se solicita. Esta función también está preparada para guardar los puntos detectados en un fichero vectorial de ENVI (.evf). A estos algoritmos se puede acceder también desde el menú de detección de cambios creado para este proyecto (Figura A-2).

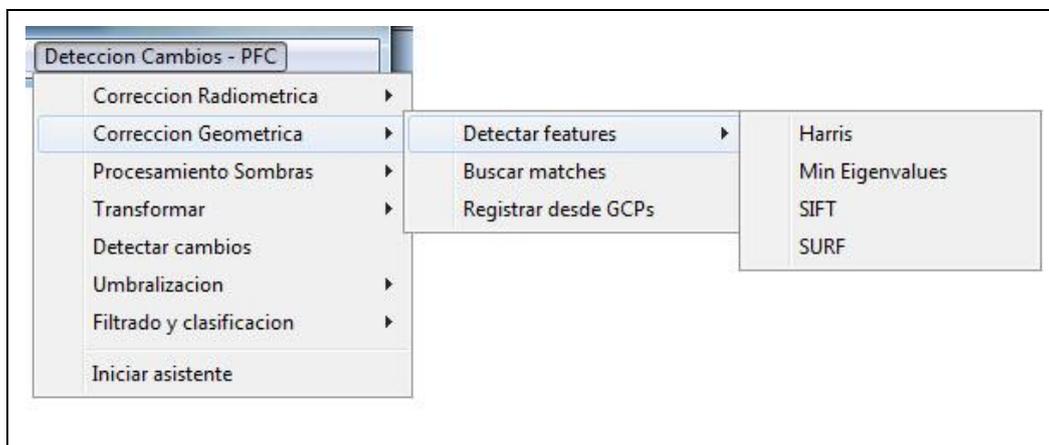


Figura A-2. Menú de opciones de correcciones geométricas.

Además de los parámetros específicos de cada técnica ya estudiados en el capítulo de correcciones geométricas, durante su ejecución y dependiendo del método seleccionado, el sistema solicita otros parámetros que explicaremos a continuación.

El parámetro *Máximo de esquinas* indica cuántas esquinas o puntos característicos se quieren seleccionar como máximo al usar los algoritmos de Harris y mínimos autovalores. Igualmente podemos indicarle, para que se repartan más o menos homogéneamente, una *Distancia mínima* entre los puntos seleccionados. Con el propósito de asegurar aún más la distribución homogénea de los puntos por toda la imagen, podemos hacer uso de los parámetros *Ancho tile* y *Alto tile*. Si estos parámetros tienen valor distinto de cero, la imagen se dividirá en *tilas* de tamaño *Ancho tile* x *Alto tile* y se buscarán tantos puntos como se indique en el parámetro *Máximo de esquinas*. Por ejemplo, si queremos seleccionar un punto característico en cada trozo de 64x64 píxeles de la imagen, debemos establecer el parámetro *Máximo de esquinas* a 1, y los parámetros *Ancho tile* y *Alto tile* a 64. Si estos parámetros tienen valor 0, se buscarán *Máximo de esquinas* en toda la imagen.



Figura A-3. Puntos de control distribuidos homogéneamente.

La segunda función *extract_matches* es la encargada de seleccionar pares de puntos para el registro de las imágenes. Se puede seleccionar las siguientes técnicas de emparejamiento: Correlación cruzada normalizada (NCC), Búsqueda por NCC, Mutual Information, Búsqueda por MI, Lucas-Kanade Tracking, Características SIFT y Características SURF. La diferencia entre las técnicas NCC y MI, y sus versiones de búsqueda, es que en las primeras se intentan emparejar puntos detectados en las imágenes maestra y esclava, y en las segundas se intenta buscar la correspondencia desplazando una ventana por la imagen esclava. Esta función también está disponible desde el menú, que hará que se muestre un cuadro de diálogo que nos permite seleccionar el método de detección de características y el método que emplearemos para su emparejamiento. Para facilitar el trabajo con imágenes grandes y evitar los problemas de memoria que conlleva, se puede indicar al sistema que divida las imágenes de entrada en fragmentos antes de procesarlos. El resultado seguirá siendo un fichero con las coordenadas de los pares de puntos encontrados para la imagen completa, liberando al usuario del procesamiento necesario en estos casos. Además, se podrá indicar si se desea usar el algoritmo RANSAC para filtrar los emparejamientos calculados antes de guardarlos en el fichero de salida.

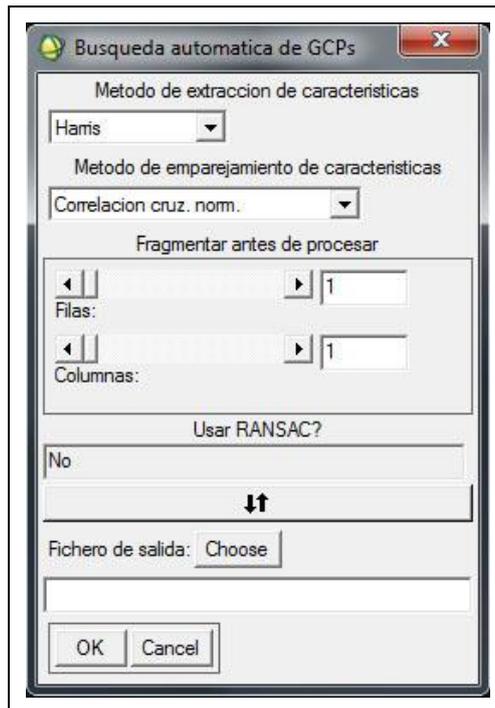


Figura A-4. Selección de métodos de extracción y matching de características.

Una vez establecidos estos parámetros iniciales, el sistema nos solicitará los parámetros para cada una de las técnicas seleccionadas (Figura A-5), con el fin de extraer los emparejamientos, guardarlos en un fichero de texto válido para la transformación posterior y mostrar un mensaje con el número de correspondencias extraídas después de filtrar por RANSAC (en caso de estar activa esta opción). El fichero de texto con los emparejamientos encontrados (GCPs) se debe usar a continuación para registrar las imágenes con la opción correspondiente del menú, o la opción *Warp from GCPs: Image to Image* de ENVI.

Además de estas dos funciones importantes, se han implementando otras funciones auxiliares para la gestión de los parámetros, para la gestión de los distintos puntos y emparejamientos, y para cálculo de medidas de error. La función *pointsToEVF* permite almacenar en un fichero vectorial de ENVI (.evf) una lista de puntos pasada como parámetro. Las funciones *pointsToFile* y *matchesToFile* son usadas para almacenar en un fichero de texto, una lista de puntos y de correspondencias respectivamente. Las coordenadas son corregidas para adaptarlas al sistema de coordenadas de ENVI cuyo origen de coordenadas es (1, 1) en vez del habitual en otros sistemas y librerías como la OpenCV (0, 0).

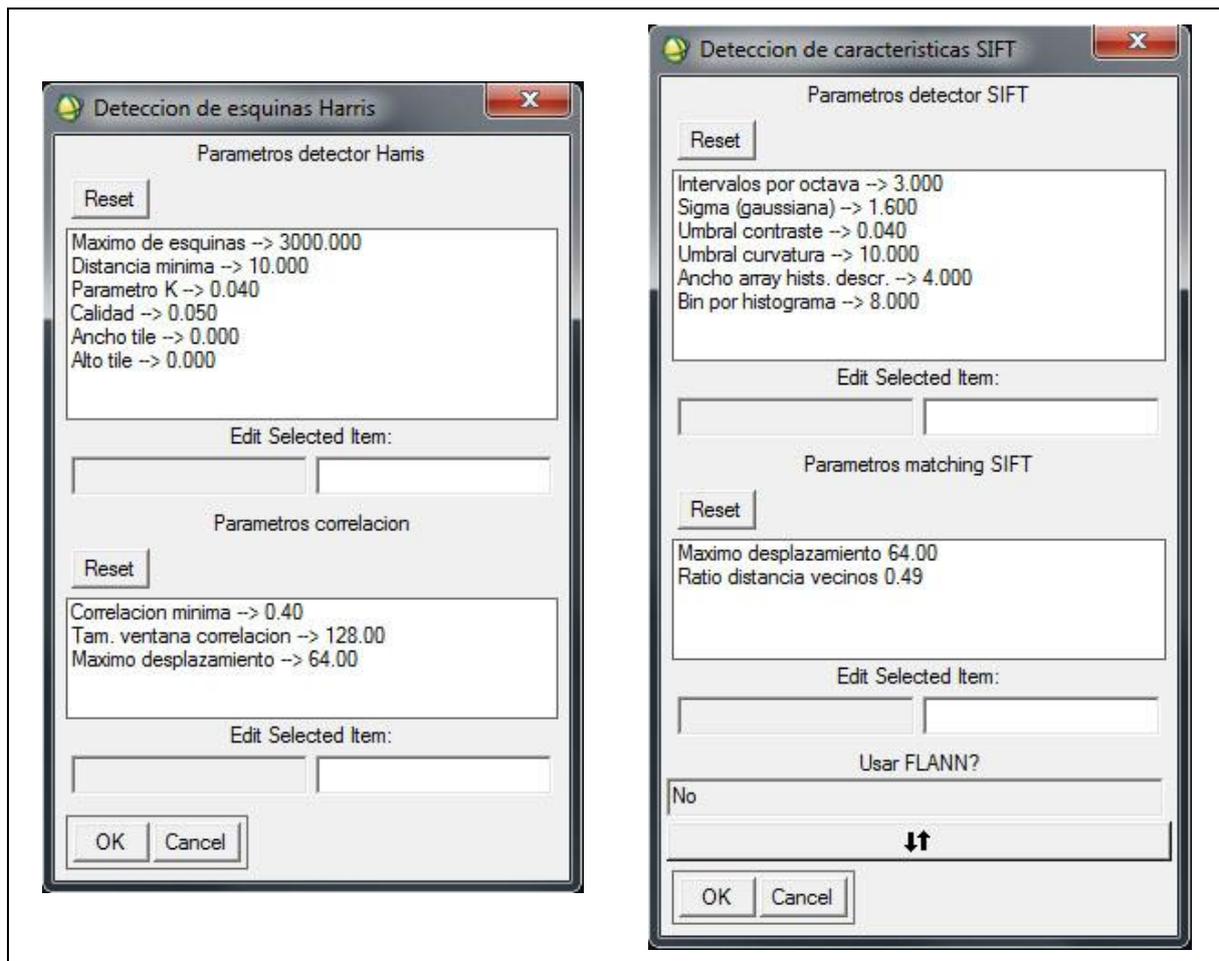


Figura A-5. Ventana de parámetros de los métodos de corrección geométrica.

Cuando la corrección geométrica se realiza después de una fragmentación previa en *tiles*, se hace uso de las funciones *getTileDims* para obtener las coordenadas y dimensiones de los distintos fragmentos a procesar.

Con el propósito de gestionar los parámetros necesarios para los algoritmos de detección y emparejamiento de características, se hace uso de la función *getParameters* que muestra un cuadro de diálogo con los parámetros necesarios para los algoritmos que se hayan seleccionado.

Por último, se han añadido una función para calcular los cuartiles en la distribución de datos, *quartiles*; una para obtener el error RMS para un ajuste polinomial de una distribución de correspondencias y grado pasado como parámetro, *get_polynomial_rms*; y una función *remove_outliers* que elimina aquellas correspondencias con un error mayor al correspondiente al tercer cuartil más 1.5 veces el rango intercuartil (véase explicación en el apartado 4.2).

A.3 Funciones para detección de sombras

Para la detección de sombras se han elaborado dos procedimientos: uno más simple e interactivo usando diferentes espacios de color y otro más complejo que sigue

todos los pasos del método propuesto en esta memoria. Con el primer método se trata de obtenerlas aplicando un valor umbral a imágenes en los espacios de color $c_1c_2c_3$, HSI e YIQ, que presentan una alta respuesta a los píxeles de sombra en alguna componente o combinación de ellas.

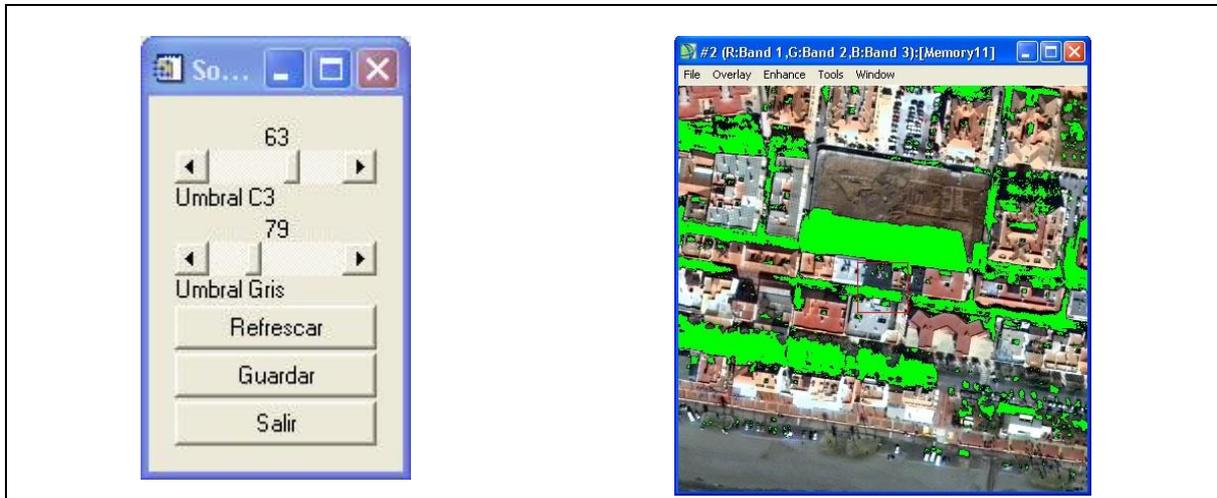


Figura A-6. Ventanas del detector simple de sombras para el espacio de color $c_1c_2c_3$.

En el caso del espacio $c_1c_2c_3$ (Figura A-6), para determinar si un píxel es de sombra, tan sólo se basa en que la componente c_3 de la imagen supere un umbral T_{C_3} y que el valor de la intensidad (componente V del espacio HSV) sea inferior a un umbral T_V . El valor de los umbrales es controlado interactivamente mediante dos controles deslizantes. El resultado se va viendo en una ventana, sobre un fragmento de la imagen completa seleccionado por el usuario. Una vez se está satisfecho con los umbrales indicados, se detectan sombras en toda la imagen y se guarda la máscara binaria con las sombras detectadas. Este método se ha implementado en la función *simple_shadows_detection* que devuelve el descriptor de fichero de la máscara de sombras que se haya generado, o el valor -1 en caso de cancelar la operación. Los detalles de los parámetros pueden consultarse en la ayuda en HTML que acompaña a los archivos del código fuente.

Para el caso de los espacios de color HSI e YIQ, las funciones *simple_shadows_detection_hsi* y *simple_shadows_detection_yiq*, después de convertir las imágenes RGB en los espacios de color correspondientes, calculan las imágenes resultantes de aplicar (5.2) y (5.3). Una vez obtenidas, se les aplica un valor umbral seleccionándolo de forma interactiva mediante una barra de desplazamiento. El resultado puede ir viéndose sobre un fragmento de la imagen antes de aplicarlo a toda la imagen y generar la máscara de cambios resultante. De nuevo, estas funciones devuelven el descriptor de fichero de la máscara generada, un valor de -1 en caso de que la operación sea cancelada.

El segundo método sigue todos los criterios propuestos por (Arévalo, González, & Ambrosio, 2008) y explicados en este proyecto. Es un procedimiento más lento por el número de operaciones y comprobaciones que se realizan. El programa solicita el valor de los parámetros necesarios para obtener la máscara de sombras. No es

interactivo como el primer procedimiento pero es más exacto, como pudieron comprobar los autores del método en su trabajo.

Los parámetros solicitados en el programa son:

- T_E = Umbral máximo imagen de ejes.
- T_S = Umbral mínimo para saturación (componente S) normalizado.
- T_V = Umbral máximo de intensidad (componente V) normalizado.
- N = Tamaño de la semilla ($N \times N$).
- D_{min} = Distancia mínima entre semillas.
- Filt. C3 = Tamaño del filtro promedio para c_3 .
- Filt. Sat = Tamaño del filtro promedio para saturación S.
- d_0 = Multiplicador de la desviación típica en el crecimiento de regiones (5.1).

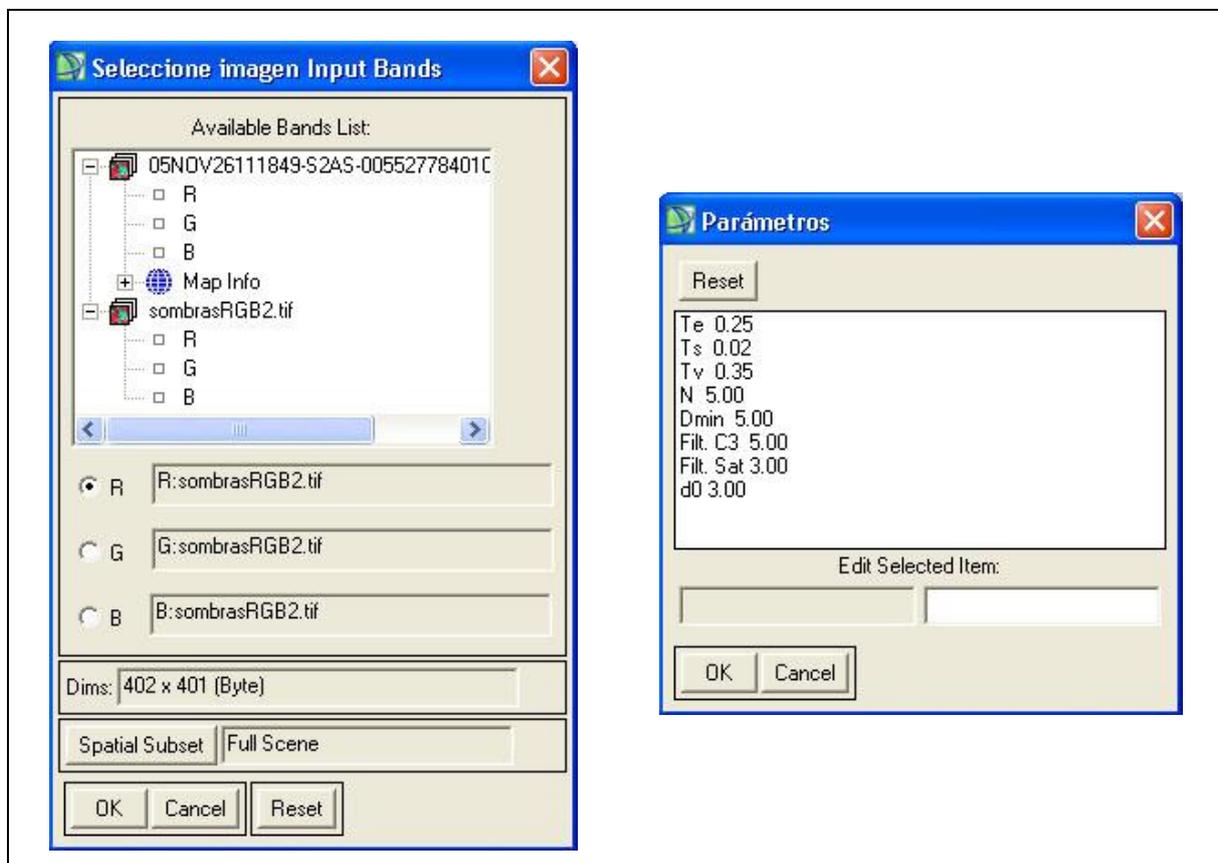


Figura A-7. Captura de ventanas del detector de sombras.

Todas las soluciones procesan las imágenes por “tiles” (fragmentos de la imagen con todas las columnas o “samples” y un conjunto consecutivo de líneas). Esto hace que para imágenes muy grandes, el método automático pueda sufrir resultados no deseados durante el crecimiento de regiones, debido a que los “tiles” se cargan en memoria y procesan uno a uno (con un solape), y puede que algunas sombras queden desconectadas, sobre todo cuando el número de líneas de cada “tile” es

demasiado pequeño como para asegurar que una sombra tiene una semilla en cada uno de ellos (o en el solape). Para reducir este problema se puede incrementar el tamaño de cada “tile” en la configuración de ENVI siempre que la máquina donde se trabaje lo permita. O como alternativa, en estos casos usar el primer procedimiento.

A.4 Funciones para la restauración de sombras

Para implementar este método en ENVI, se han generado un procedimiento y una función con misma denominación: “*shadows_restore*”.

Las imágenes a emplear deben de tener las mismas dimensiones y estar corregistradas. Deben emplearse imágenes de tipo byte si el número de bins en las preferencias de Envi es 256. Se pueden emplear imágenes de tipo Uint si éste es superior a 256.

La función “*shadows_restore*” tiene 5 argumentos, 4 de ellos obligatorios. Recibe las dos imágenes a emplear: *im1*, imagen en la que se desea restaurar las sombras, e *im2*, imagen empleada para obtener la información para la generación de los histogramas; además recibe dos máscaras de sombras, una para cada imagen. Un quinto parámetro opcional permite introducir un nombre para la imagen de salida. Esta función sólo se puede llamar desde IDL. Para llamar al método desde menú, se hace uso del procedimiento de mismo nombre, que se encarga de solicitar las dos imágenes y máscaras de sombras.

A.5 Funciones para la detección de cambios

La principal función implementada para la obtención de la imagen binaria de cambios es *get_changes_image*. El método empleado para obtener la imagen de cambios se selecciona dando un valor numérico al parámetro *method* de entre los siguientes:

- 0 – Módulo del vector de diferencias.
- 1 – Diferencia de imágenes (sin aplicar valor absoluto).
- 2 – Ratio de imágenes.
- 3 – Regresión de imágenes.



Figura A-8. Selección de método de detección de cambios.

Si no se da valor a este parámetro, el sistema solicitará el método mediante una lista desplegable (Figura A-8). En el caso de imágenes de más de una banda sólo es seleccionable la primera de las opciones. El resultado de la función será el descriptor de fichero de la imagen de cambios en niveles de gris. Será necesario aplicar posteriormente una binarización de la imagen aplicando un valor umbral por cualquiera de los métodos implementados que comentaremos más adelante.

La función *get_changes_image* llama a los procedimientos *image_differencing*, *image_ratioing*, e *image_regression* en función del método que se haya seleccionado. Estos procedimientos trabajan directamente con los datos y devuelven el resultado en un parámetro de salida a la función que se encarga de escribirlo en un fichero con nombre especificado por el usuario.

También se han implementado algoritmos de transformación de las imágenes. La función *principal_component_analysis* realiza un análisis de componentes principales de una imagen y devuelve el descriptor de fichero de la imagen resultante. También disponemos de la función *tasseled_cap* que permite la conocida transformación Tasseled-Cap para imágenes de los satélites Quickbird e Ikonos.

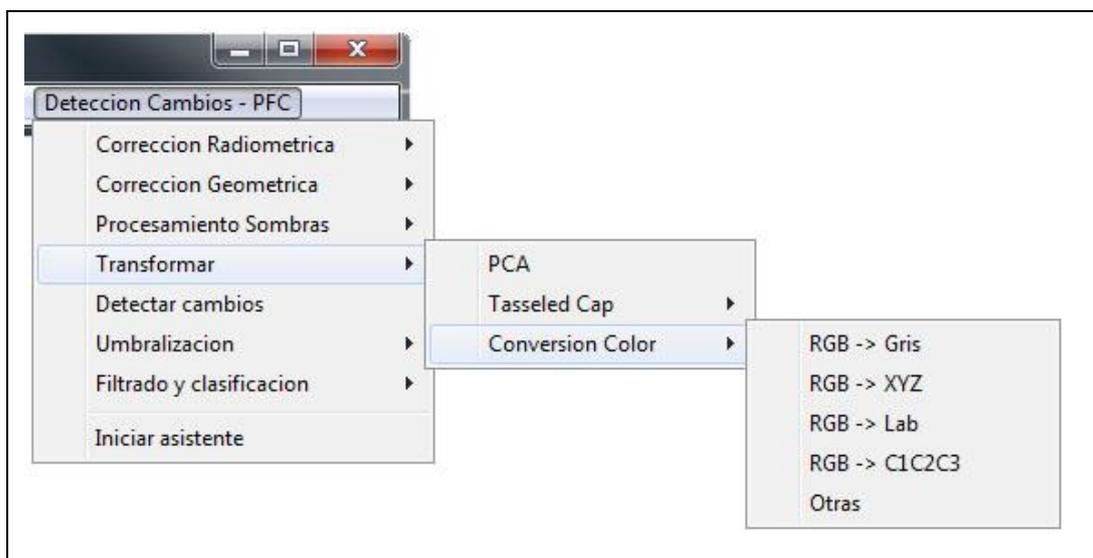


Figura A-9. Menú de opciones de transformación.

Para la binarización de la imagen de cambios en niveles de gris (Figura A-10) disponemos de las funciones *changes_thresholding*, *adaptive_thresholding*, y *interactive_thresholding*. La primera permite aplicar un valor umbral automático y global a toda la imagen, y la segunda es similar pero obtiene varios umbrales y los aplica de forma adaptativa. Ambas llaman a las funciones *isodata_threshold*, *moment_preserving_threshold*, *otsu_threshold*, y *unimodal_threshold* para obtener el valor umbral por los cuatro métodos presentados en la memoria. La función interactiva presenta una herramienta para la selección manual del valor umbral a través de un control deslizante (Figura A-11). Las tres alternativas devuelven el descriptor de fichero de la imagen binaria de cambios resultante de aplicar el valor umbral.

Una última función llamada *fuzzy_thresholding* nos permite obtener una umbralización multinivel, asignando a cada píxel una probabilidad de cambio en vez de asignarle la clase de “Cambio” o “No-Cambio”, como hacen las anteriores.



Figura A-10. Menú de opciones de umbralización.

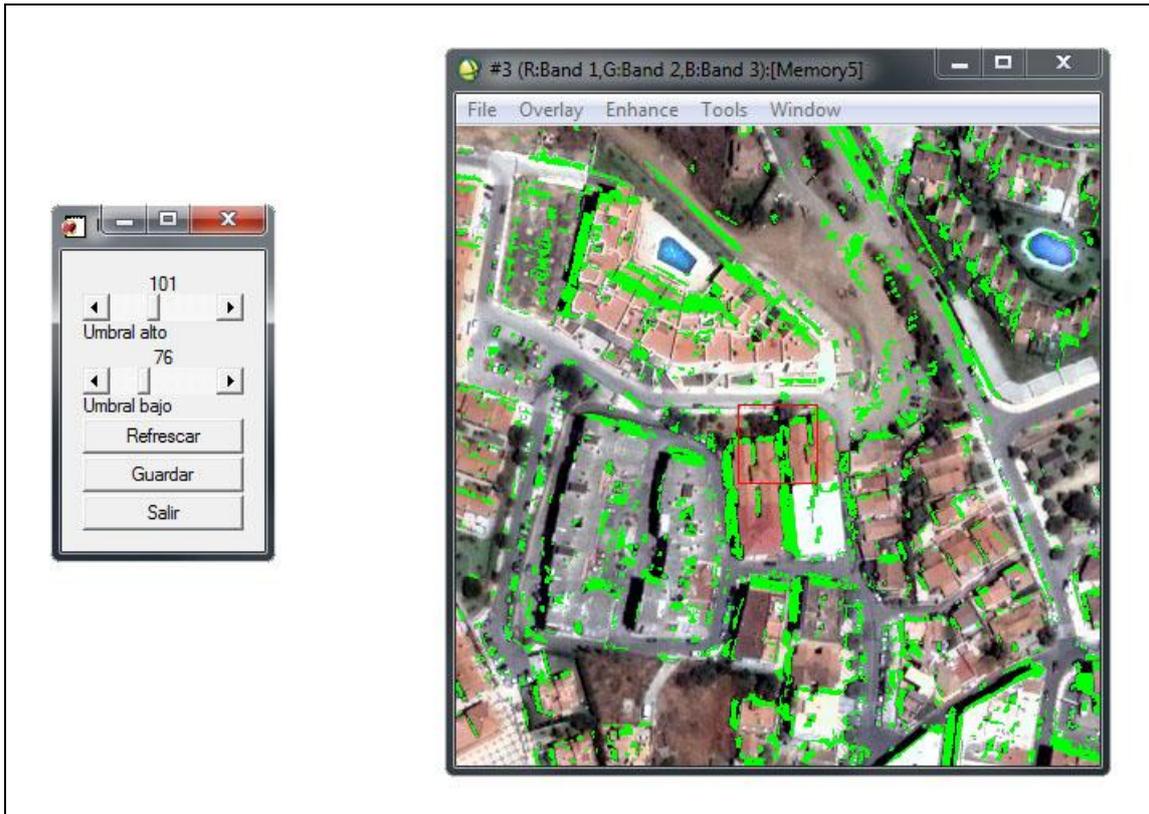


Figura A-11. Selección de umbral con histéresis.

A.6 Funciones para el filtrado y clasificación de cambios

Las funciones de filtrado se han implementado, principalmente, en una librería dinámica externa en lenguaje C que recibe el nombre de *ChangesFilter.dll* con el propósito de poder usar las funciones de la librería *OpenCV* para las operaciones más costosas en recursos. Estas operaciones de filtrado se llaman desde la función *filter_changes_wrapper* que devolverá el valor 0 si todas las operaciones solicitadas se han realizado con éxito. Consultar los posibles códigos de error en la ayuda HTML.

Los tipos de filtrados posibles son un filtrado de cambios por área, un ajuste geométrico local y la eliminación de sombras. Esta función se encargará también de vectorizar los cambios y guardar en su base de datos el tipo de filtrado aplicado a cada uno. La función solicitará al usuario las operaciones de filtrado que se desean realizar y, a continuación, solicitará las imágenes necesarias para su realización y el nombre del fichero *Shapefile* de salida. Si además se desea usar valores distintos a los usados por defecto para los parámetros necesarios, se debe llamar a la función con el argumento */params*.

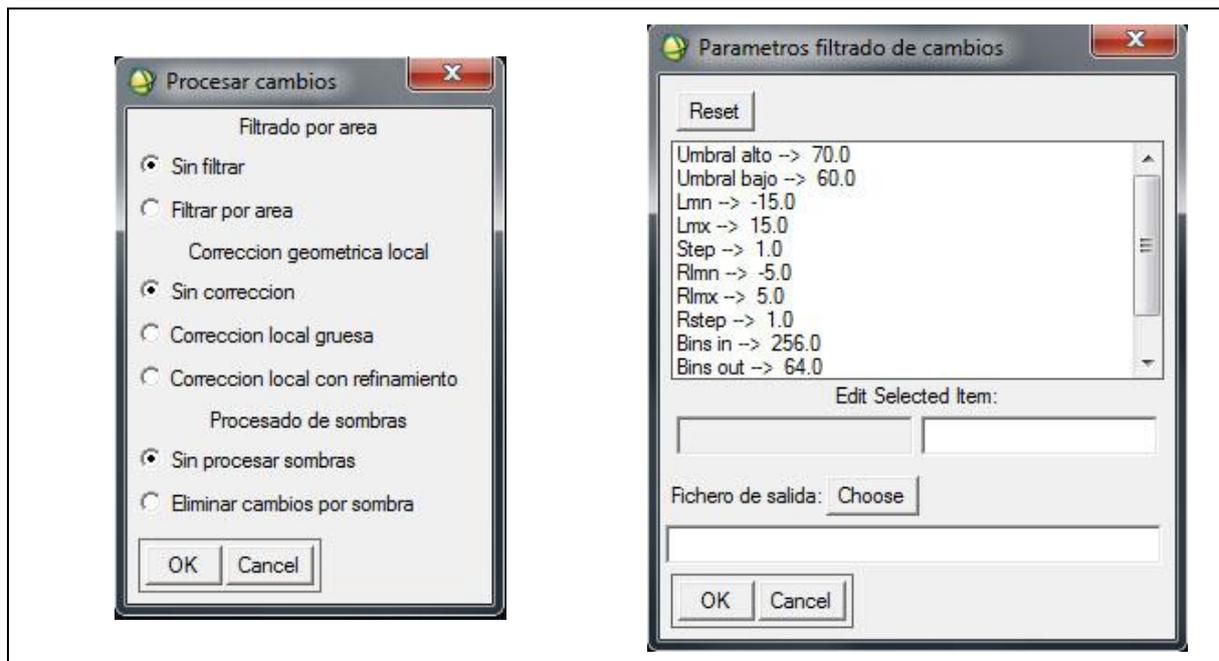


Figura A-12. Opciones de filtrado y parámetros para corrección geométrica local.

Los parámetros solicitados por la función dependerán de las opciones de filtrado seleccionadas (Figura A-12). Para el filtrado por área se pide un área mínima y un área máxima para los cambios. Al parámetro área máxima se le puede asignar el valor -1 si no queremos aplicar un límite superior al tamaño del cambio. Así, por ejemplo, en el caso de querer vectorizar los cambios pero no querer aplicar ningún tipo de filtrado, podremos llamar a la función usando un filtrado por área, pero asignando el valor 0 al área mínima y un valor de -1 al área máxima.

La corrección geométrica local realiza una detección de cambios usando el método de diferencias. Por este motivo, es conveniente recordar los valores umbrales usados para obtener la imagen binaria de cambios que se desea filtrar. Los parámetros *Lmn*, *Lmx*, *Step*, *Rlmn*, *Rlmx* y *Rstep* permiten controlar el desplazamiento de la ventana de búsqueda para la corrección geométrica local. También, se puede controlar el número de *bins* usados para el cálculo del histograma conjunto empleado en el cálculo de la información mutua (es ignorado en caso de usar correlación). Generalmente, estos parámetros no hay que modificarlos. El parámetro *Radio* determina el área circular alrededor de cada cambio que se tiene en cuenta a la hora de decidir si un desplazamiento propuesto es consistente con el de los cambios vecinos. Y, por último, el parámetro *Método alineado* puede tomar dos valores distintos. Si el valor es 0 se empleará la información mutua para la corrección, y si es 1 se aplicará búsqueda por correlación cruzada normalizada.

En un tercer paso, el filtrado de sombras elimina de la imagen binaria de cambios los píxeles de cambio provocados por sombras. Este filtrado es útil en el caso de que una restauración de sombras no sea factible. Esto puede hacer también que cambios ocultos por sombras no sean identificados como tales y es importante usar máscaras de sombras lo más precisas posible.

Durante el filtrado se guarda, en un archivo de base de datos DBF que acompaña al fichero Shapefile de salida, un valor numérico que indica el tipo de filtrado que se le ha realizado. Este valor se puede encontrar en el atributo *Clase_Filtrado* y puede tomar los siguientes valores:

- 1 – No se ha aplicado ninguna corrección local.
- 2 – Corrección no posible por situarse el cambio en el borde de la imagen.
- 3 – No hay suficientes vecinos para comprobar la consistencia de la corrección.
- 4 – Se ha aplicado una corrección geométrica local.
- 5 – No se ha aplicado una corrección geométrica local por no ser consistente con sus vecinos.

El proceso de filtrado tiene en cuenta la proyección de las imágenes a la hora de vectorizar. Por tanto, sus coordenadas pueden estar basadas en píxeles o en la proyección de las imágenes empleadas. Para mostrar los cambios en forma vectorial, ENVI solicitará transformar el fichero en su propio formato EVF (Envi Vector File). Si el fichero de la imagen menos reciente tenía asignada una proyección, será necesario indicar la proyección en el momento de la conversión para que los cambios se localicen correctamente.

Para la clasificación no supervisada de los cambios, disponemos de la función *change_classification*. Dependiendo de sus parámetros, se usará la técnica de análisis de la dirección del vector de cambio (CVA) o el algoritmo K-Means para clasificar los cambios.

En caso de necesitar una clasificación totalmente supervisada se debe llamar a la herramienta proporcionada por el procedimiento *interactive_classification* con el que se puede asignar una clase (incluso definir nuevas clases) para cada cambio. Es importante destacar que estos procedimientos sólo funcionan con ficheros vectoriales que cumplan algunas condiciones como los generados por la herramienta de filtrado. El archivo de base de datos debe contar con un campo *Clase* donde almacenará el valor numérico de la clase asignada a cada cambio. En el caso de la herramienta supervisada, es necesario tener la imagen en la que se revisarán los cambios mostrada previamente en algún display de ENVI.

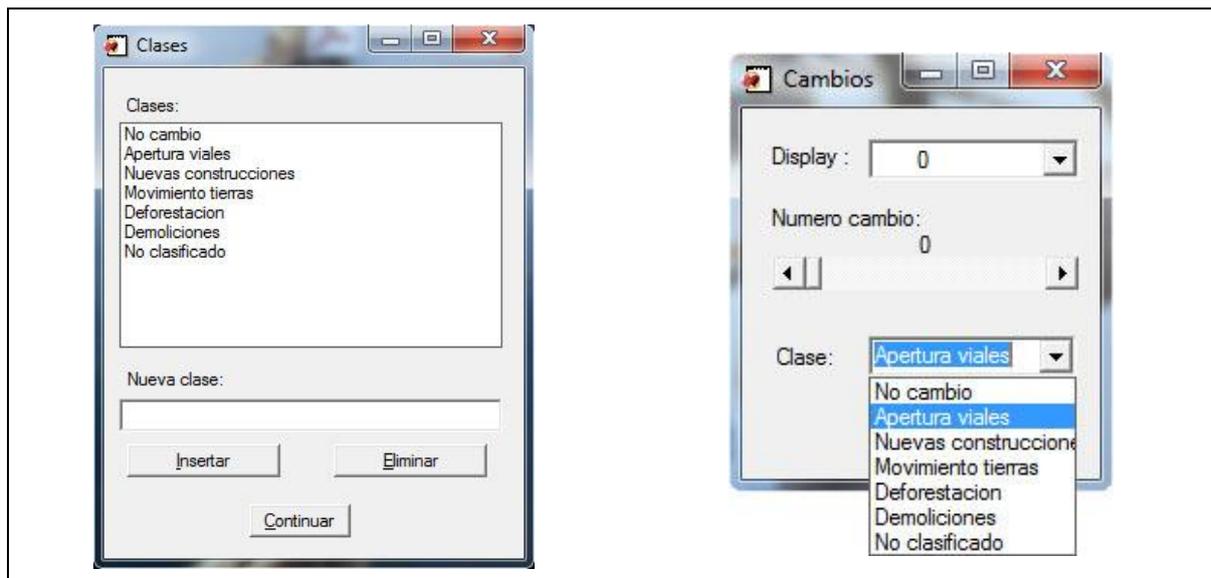


Figura A-13. Herramientas para la definición (izq.) y clasificación manual (der.) de cambios.

A.7 Asistente para la detección de cambios

Con el propósito de facilitar y simplificar todo el proceso de detección de cambios, se ha implementado un asistente que guía al usuario por las distintas fases permitiéndole realizar las operaciones más comunes en cualquier detección de cambios, a la vez que gestiona todos los ficheros que se van generando durante el proceso. Las distintas fases son accesibles mediante botones y las distintas pestañas de la ventana del asistente, como se puede apreciar en la figura a continuación. Todas las operaciones que se van realizando quedan registradas en un fichero de log si se ha especificado en la primera de las fases.

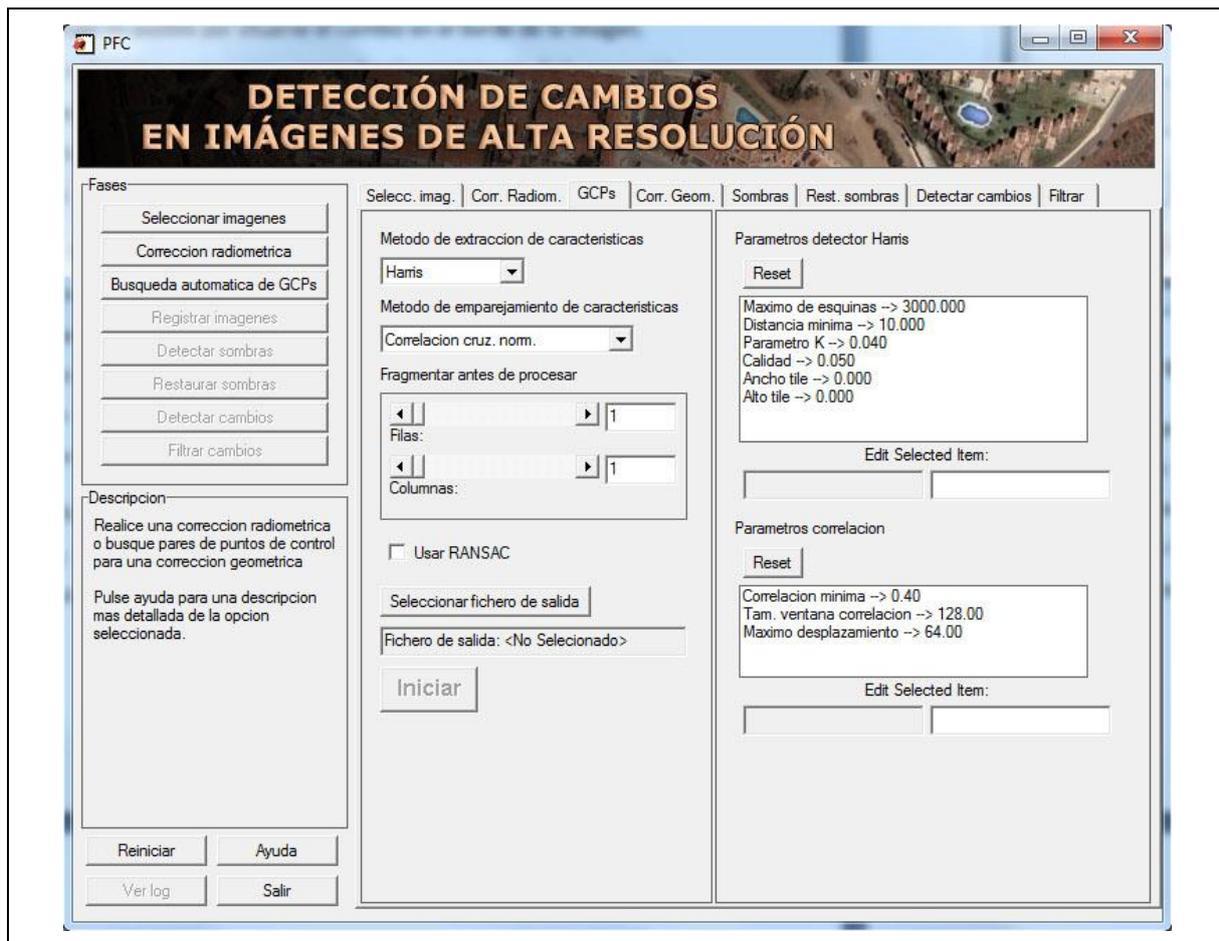


Figura A-14. Asistente para la detección de cambios.

Se puede llamar al asistente desde la opción correspondiente en el menú creado para el proyecto o llamando a la función *wizard*.

A.8 Instalación del software

Todas las funciones se clasifican en carpetas según su categoría. Para instalarlas tan sólo es necesario copiarlas a la carpeta *save_add* de ENVI para que sean compiladas cada vez que se inicia el programa y estén disponibles para su uso.

A las funciones les acompaña una ayuda en HTML que explica la funcionalidad, los requisitos y parámetros de cada una de ellas. Se puede encontrar en el CD que acompaña a la memoria junto al código fuente.

Apéndice B

Listado de funciones.

- adaptive_bilinear_normalization

Aplica el método de normalización radiométrica adaptativo bilineal.

- adaptive_thresholding

Función para umbralización adaptativa de imágenes de cambios.

- changes_classification

Función para la clasificación no supervisada de los cambios.

- changes_thresholding

Función para umbralización de imágenes de cambios.

- classify_cva

Función para la clasificación de cambios por dirección del vector de cambios.

- classify_kmeans

Función para la clasificación de cambios por el algoritmo K-Means.

- color_conversion

Realiza la transformación de una imagen de un espacio de color a otro.

- compare_histograms

Función para obtener el Error Cuadrático Medio entre los histogramas de dos imágenes.

- covariance_matrix

Obtiene la matriz de covarianzas de un array con tantas filas como variables se quiere analizar.

- dark_bright_normalization

Función de normalización por el proceso de selección de los conjuntos oscuros y claros.

extract_features

Procedimiento para llamar a los diferentes métodos de extracción de características.

extract_harris

Función de extracción de esquinas de Harris.

extract_lk

Función de extracción de esquinas propuesto en (Shi & Tomasi, 1994).

extract_matches

Procedimiento de extracción de emparejamientos para el registro de imágenes.

extract_SIFT

Función de extracción de características SIFT.

extract_SURF

Función de extracción de características SURF.

filter_changes

Llamada a la función de la librería ChangesFilter.dll para filtrado de cambios.

filter_changes_wrapper

Función para filtrado de los cambios.

find_features

Llamada a la función de la librería GeomCorr.dll para la extracción de características.

find_matches

Llamada a la función de la librería GeomCorr.dll para la extracción de emparejamientos.

fuzzy_thresholding

Función para umbralización difusa

get_change_info

Función auxiliar para extraer los datos de un cambio para la clasificación no supervisada.

get_changes_image

Función para calcular la imagen de cambios en niveles de gris.

get_data_nbands

Función para leer una región de una imagen desde disco.

get_polynomial_rms

Función para el cálculo del error cuadrático medio de la posición de los puntos de control para una transformación polinómica.

get_shadows_mask

Genera una máscara de sombras usando componente C3 del espacio de color C1C2C3.

get_shadows_mask_hsi

Genera una máscara de sombras usando el espacio de color HSI.

get_shadows_mask_yiq

Genera una máscara de sombras usando espacio de color YIQ.

getTileDims

Función que devuelve las dimensiones de un tile cuando en la corrección geométrica se realiza una división de las imágenes.

haze_correction

Esta función realiza una corrección radiométrica lineal de la forma $x_i = y_i + (y_{min} - x_{min})$, donde x_{min} e y_{min} son los valores digitales que dejan por debajo el 0.1% de los píxeles más oscuros en la imagen a ajustar y de referencia respectivamente.

histogram_specification

Función para corregir radiométricamente una imagen con el método de especificación de histograma.

histogram_specification_arrays

Función para corregir radiométricamente una imagen con el método de especificación de histograma.

hysteresis_thresholding

Aplica un umbral con histéresis a un array.

image_differencing

Obtiene el valor absoluto de la diferencia entre dos arrays.

image_ratioing

Obtiene el ratio entre dos arrays de dos dimensiones (imágenes de una banda).

image_regression

Obtiene la imagen de cambios entre dos arrays de dos dimensiones (imágenes de una banda) usando el método de detección de cambios por regresión de imágenes.

interactive_classification

Procedimiento para la clasificación manual de los cambios.

interactive_thresholding

Función que permite aplicar un umbral con histéresis a una imagen de niveles de gris.

isodata_threshold

Selecciona un valor umbral usando la clasificación no supervisada ISODATA (Ridler and Calvard 1978)

lab2rgb

Convierte una imagen del espacio de color CIELAB al espacio RGB.

lab2xyz

Convierte una imagen del espacio de color CIELAB a CIEXYZ.

linspace

Return a real vector of length N, containing equidistant values between x1 and x2 inclusively.

logical_xor

Operación XOR de los parámetros que recibe.

matchesToFile

Procedimiento para almacenar una lista de emparejamientos en un fichero de texto. Usado para escribir a fichero de texto las coordenadas devueltas por las funciones en OpenCV.

maximum_minimum_correction

Función que normaliza la imagen de tal forma que tenga el mismo DN mínimo y máximo que la imagen de referencia en todas las bandas.

mean_square_error

Función que devuelve el Error Cuadrático Medio entre dos imágenes para cada banda en un array de tantos elementos como bandas tengan las imágenes.

mean_square_error_pro

Procedimiento que llama a la función de cálculo del Error Cuadrático Medio y muestra los resultados en un cuadro de diálogo.

moment_preserving_threshold

Selección de umbral óptimo por conservación de momentos (Tsai (1985) Moment-preserving thresholding)

ms_normalization

Esta función normaliza la imagen a partir de medidas estadísticas.

otsu_threshold

Selección de umbral óptimo por el método de Otsu (1979)

PI_normalization

Normalización por el método de pseudoinvariantes

PI_set

Función de selección del conjunto invariante Propuesto por G, Hong y Y, Zhang en "A Comparative Study on Radiometric Normalization Using High Resolution Satellite Images"

pointsToEvf

Procedimiento para almacenar una lista de puntos en un fichero vectorial de ENVI.

pointsToFile

Procedimiento para almacenar una lista de puntos en un fichero de texto. Usado para escribir a fichero de texto las coordenadas devueltas por las funciones en OpenCV.

principal_component_analysis

Realiza la transformación Principal Component Analysis (PCA) de una imagen.

principal_component_analysis_arrays

Realiza la transformación Principal Component Analysis (PCA) de una imagen contenida en un array.

ProgramRootDir

Encuentra el directorio raíz de un programa.

quartiles

Función para el cálculo de los cuartiles de un conjunto de datos.

remove_outliers

Función de eliminación de outliers en un fichero de puntos de control.

rgb2ccc

Convierte una imagen del espacio de color RGB al espacio C1C2C3 propuesto en Color Based Object Recognition (Gevers y Smeulders, 1999)

rgb2gray

Convierte una imagen del espacio de color RGB en una imagen en niveles de gris.

rgb2hsi

Convierte una imagen del espacio de color RGB al espacio HSI.

rgb2lab

Convierte una imagen del espacio de color RGB al espacio CIELAB.

rgb2l1l2l3

Convierte una imagen del espacio de color RGB al espacio L1L2L3 propuesto en Color Based Object Recognition (Gevers y Smeulders, 1999)

rgb2xyy

Convierte una imagen del espacio de color RGB al espacio xyY

rgb2xyz

Convierte una imagen del espacio de color RGB al espacio CIEXYZ.

rgb2yiq

Convierte una imagen del espacio de color RGB al espacio YIQ.

shadows_detection

Genera una máscara de sombras por el método de crecimiento de regiones (Arévalo et al, 2008).

shadows_restorer

Función que recupera la información de las sombras de una imagen con información de la zona en otra imagen.

shadows_restorer

Procedimiento que recupera la información de las sombras de una imagen, con información de la zona en otra imagen.

simple_shadows_detection

Genera una máscara de sombras usando componente C3 del espacio de color C1C2C3.

simple_shadows_detection_hsi

Genera una máscara de sombras usando el espacio de color HSI.

simple_shadows_detection_yiq

Genera una máscara de sombras usando el espacio de color YIQ.

sr_normalization

Esta función normaliza la imagen a partir de medidas estadísticas.

tasseled_cap

Función para la obtención de la transformación Tasseled Cap de una imagen Quickbird o Ikonos.

tasseled_cap_ikonos

Recibe una imagen Ikonos en un array y obtiene la transformada Tasseled Cap.

tasseled_cap_quickbird

Recibe una imagen Quickbird en un array y obtiene la transformada Tasseled
Cap.

unimodal_threshold

Selección de umbral óptimo para imágenes con histograma unimodal (Rosin
2001)

wizard

Inicia el asistente de detección de cambios.

xyy2xyz

Convierte una imagen del espacio de color xyY al espacio CIEXYZ

xyz2lab

Convierte una imagen del espacio de color CIEXYZ al espacio CIELAB.

xyz2rgb

Convierte una imagen del espacio de color CIEXYZ al espacio RGB.

xyz2xyy

Convierte una imagen del espacio de color CIEXYZ al espacio xyY.

Referencias

- Abuelgasim, A. A., Ross, W. D., Gopal, S., & Woodcock, C. E. (1999). Change detection using adaptive fuzzy neural networks: environmental damage assessment after the Gulf War. *Remote Sensing of Environment*, 70, 208-223.
- Adams, J. B., Sabol, D., Kapos, V., Filho, R. A., Roberts, D. A., Smith, M. O., & Gillespie, A. R. (1995). Classification of multispectral images based on fractions of endmembers: application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, 52, 137-154.
- Agrawal, M., Konolige, K., & Blas, M. R. (2008). CenSurE: Center Surround Extremas for Realtime Feature Detection and Matching. *European Conference on Computer Vision*, 102-115.
- Arévalo, V., González, J., & Ambrosio, G. (2008). Shadow detection in colour high-resolution satellite images. *International Journal of Remote Sensing*, 29(7), 1945-1963.
- Arévalo, V., & González, J. (2008). An experimental evaluation of non-rigid registration techniques on Quickbird satellite imagery. *International Journal of Remote Sensing*, 29(2), 513-527.
- Asner, G. P., Keller, M., Pereira, R., & Zweede, J. C. (2002). Remote sensing of selective logging in Amazonia - assessing limitations based on detailed field observations, Landsat ETM+, and textural analysis. *Remote Sensing of Environment*, 80, 483-496.
- Bay, H., Ess, A., Tuytelaars, T., & Gool, L. V. (2008). Speeded-Up Robust Features. *Computer Vision and Image Understanding (CVIU)*, 110(3), 346-359.
- Beis, J. S., & Lowe, D. G. (2003). Shape indexing using approximate nearest-neighbor search in high-dimensional spaces. *Conference on Computer Vision and Pattern Recognition (CVPR)*, (págs. 1000-1006).
- Bouguet, J. Y. (1999). Pyramidal Implementation of the Lucas Kanade Feature Tracker. *Intel Coporation, Microprocessor Research Labs*.

- Bovolo, F. (2009). A multilevel parcel-based approach to change detection in very high resolution multitemporal images. *IEEE Geoscience and Remote Sensing Letters*, 6(1), 33-37.
- Briechle, K., & Hanebeck, U. D. (2001). Template matching using fast normalized cross correlation. *Aerospace/Defense Sensing, Simulation, and Controls* (págs. 95-102). International Society for Optics and Photonics.
- Brondizio, E. S., Moran, E. F., Mausel, P., & Wu, Y. (1994). Land use change in the Amazon Estuary: patterns of Caboclo settlement and landscape management. *Human Ecology*, 22, 249-278.
- Bruzzone, L., & Bovolo, F. (2013). A novel framework for the design of change-detection systems for very-high-resolution remote sensing images. *Proceedings of the IEEE*, 101(3), 609-630.
- Bruzzone, L., & Prieto, D. F. (2000). An adaptive parcel-based technique for unsupervised change detection. *International Journal of Remote Sensing*, 21(4), 817-822.
- Byrne, G. F., Crapper, P. F., & Mayo, K. K. (1980). Monitoring land cover change by principal component analysis of multitemporal Landsat data. *Remote Sensing of Environment*, 10, 175-184.
- Castelli, V., Elvidge, C. D., Li, C. S., & Turek, J. J. (1998). Classification-based change detection: theory and application to the NALC data sets. En R. S. Lunetta, & C. D. Elvidge, *In Remote Sensing Change Detection: Environmental Monitoring Methods and Applications* (págs. 53-73). Chelsea, MI: Ann Arbor Press.
- Chen, X. (2002). Using remote sensing and GIS to analyze land cover change and its impacts on regional sustainable development. *International Journal of Remote Sensing*, 23, 107-124.
- Collins, J. B., & Woodcock, C. E. (1994). Change detection using the Gramm-Schmidt transformation applied to mapping forest mortality. *Remote Sensing of Environment*, 50, 267-279.
- Collins, J. B., & Woodcock, C. E. (1996). An assessment of several linear change detection techniques for mapping forest mortality using multiemporal Landsat TM data. *Remote Sensing of Environment*, 56, 66-77.
- Coppin, P., Nackaerts, K., Queen, L., & Brewer, K. (2001). Operational monitoring of green biomass change for forest management. *Photogrammetric Engineering and Remote Sensing*, 67, 603-611.

- Dai, X. L., & Khorram, S. (1999). Remotely sensed change detection based on artificial neural networks. *Photogrammetric Engineering and Remote Sensing*, 65, 1187-1194.
- Dimiyati, M., Mizuno, K., Kobayashi, S., & Kitamura, T. (1996). An analysis of land use/cover change using the combination of MSS Landsat and land use map - a case study in Yogyakarta, Indonesia. *International Journal of Remote Sensing*, 17, 931-944.
- Fischler, M. A., & Bolles, R. C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *24(6)*, 381-395.
- Foody, G. M. (2001). Monitoring the magnitude of land-cover change around the southern limits of the Sahara. *Photogrammetric Engineering and Remote Sensing*, 67, 841-847.
- Gevers, T., & Smeulders, A. W. (1999). Colour-based object recognition. *Pattern Recognition*, 32, 453-464.
- González, J. (2000). *Visión por computador*. Ed. Paraninfo.
- J González, C Galindo, V Arevalo, G Ambrosio (2007). Applying image analysis and probabilistic techniques for counting olive trees in high-resolution satellite images. *Advanced Concepts for Intelligent Vision Systems*.
- Gonzalez, R. C., & Woods, R. E. (2002). *Digital Image Processing*. Ed. Prentice Hall.
- Gopal, S., & Woodcock, C. E. (1996). Remote sensing of forest change using artificial neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 34, 398-403.
- Guerra, F., Puig, H., & Chaume, R. (1998). The forest-savanna dynamics from multitemporal Landsat-TM data in Sierra Parima, Venezuela. *International Journal of Remote Sensing*, 19, 2061-2075.
- Hall, F. G., Strebel, D. E., Nickeson, J. E., & Goetz, S. J. (1991). Radiometric rectification: Toward a common radiometric response among multitemporal, multisensor images. *Remote Sensing of Environment*, 35(1), 11-27.
- Harris, C. G., & Stephens, M. (1988). A Combined Corner and Edge Detector. *In 4th Alvey Vision Conference*, (págs. 174-151).
- Henebry, G. M. (1993). Detecting change in grasslands using measures of spatial dependence with Landsat TM data. *Remote Sensing of Environment*, 46, 223-234.

- Hess, R. (2010). *SIFT Library*. Obtenido de <http://blogs.oregonstate.edu/hess/code/sift/>
- Hong, G., & Zhang, Y. (2008). A comparative study on radiometric normalization using high resolution satellite images. *International Journal of Remote Sensing*, 29(2), 1-14.
- Horne, H. J. (2003). A tasseled cap transformation for IKONOS image. *ASPRS 2003 Annual conference proceedings*. Anchorage, Alaska.
- Ingebritsen, S. E., & Lyon, R. J. (1985). Principal component analysis of multitemporal image pairs. *International Journal of Remote Sensing*, 6, 687-696.
- Jensen, J. R., Cowen, D., Narumalani, S., & Halls, J. (1997). Principles of change detection using digital remote sensor data. En J. L. Star, J. E. Estes, & K. C. McGwire (Edits.), *Integration of Geographic Information Systems and Remote Sensing* (págs. 37-54). Cambridge: Cambridge University Press.
- Jensen, J. R., Ramsay, E. W., Mackey, H. E., Christensen, E. J., & Sharitz, R. P. (1987). Inland wetland change detection using aircraft MSS data. *Photogrammetric Engineering and Remote Sensing*, 53, 521-529.
- Jensen, J. R., Rutchey, K., Koch, M. S., & Narumalani, S. (1995). Inland wetland change detection in the everglades water conversation area: using a time series of normalized remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 61, 199-209.
- Jha, C. S., & Unni, N. V. (1994). Digital change detection of forest conversion of dry tropical forest region. *International Journal of Remote Sensing*, 15, 2543-2552.
- Kauth, R. J., & Thomas, G. S. (1976). The Tasseled Cap - a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Proceedings the Symposium Machine Processing of Remoetly Sensed Data*, (págs. 4B-41-4B-51). Purdue University, West Lafayette, Indiana.
- Ke, Y., & Sukthankar, R. (2004). PCA-SIFT: A more distinctive representation for local image descriptors. *Proceedings of the Conference on Computer Vision and Pattern Recognition*, (págs. 511-517). Washington, USA.
- Koenderink, J. J. (1984). The Structure of Images. *Biological Cybernetics*, 50(5), 363-370.
- Kressler, F., & Steinnocher, K. (1996). Change detection in urban areas using satellite data and spectral mixture analysis. *International Archives of Photogrammetry and Remote Sensing*, 31, 379-383.

- Kwarteng, A. Y., & Chavez, P. S. (1998). Change detection study of Kuwait city and environs using multitemporal Landsat Thematic Mapper data. *International Journal of Remote Sensing*, 19, 1651-1661.
- Lambin, E. F., & Strahler, A. H. (1994). Indicators of land-cover change for change vector analysis in multitemporal space at coarse spatial scales. *International Journal of Remote Sensing*, 15, 2099-2119.
- Li, X., & Strahler, A. H. (1985). Geometric-optical modeling of a conifer forest canopy. *IEEE Transactions on Geoscience and Remote Sensing*, 23, 705-721.
- Li, X., & Strahler, A. H. (1992). Geometric-optical bidirectional reflectance modeling of discrete crown vegetation canopy: Effect of crown shape and mutual shadowing. *IEEE Transactions on Geoscience and Remote Sensing*, 30, 276-292.
- Li, X., & Yeh, A. G. (1998). Principal component analysis of stacked multitemporal images for the monitoring of rapid urban expansion in the Pearl River Delta. *International Journal of Remote Sensing*, 19, 1501-1518.
- Lindeberg, T. (1994). Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 21(2), 225-270.
- Liu, X., & Lathrop, R. G. (2002). Urban change detection based on an artificial neural network. *International Journal of Remote Sensing*, 23, 2513-2518.
- Lo, C. P., & Shipman, R. L. (1990). A GIS approach to land-use change dynamics detection. *Photogrammetric Engineering and Remote Sensing*, 56, 1483-1491.
- Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *International Conference of Computer Vision*, (págs. 1150-1157). Corfu, Greece.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91-110.
- Lucas, B. D., & Kanade, T. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision. *Seventh International Joint Conference on Artificial Intelligence (IJCAI-81)*, (págs. 674-679).
- Lyon, J. G., Yuan, D., Lunetta, R. S., & Elvidge, C. D. (1998). A change detection experiment using vegetation indices. *Photogrammetric Engineering and Remote Sensing*, 64, 143-150.
- Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., & Suetens, P. (1997). Multimodality Image Registration by Maximization of Mutual Information. *Transactions on Medical Imaging*, 16(2), 187-198.

- Manavalan, P., Kesavasamy, K., & Adiga, S. (1995). Irrigated crops monitoring through seasons using digital change detection analysis or IRD-LISS 2 data. *International Journal of Remote Sensing*, 16, 633-640.
- Mas, J. F. (1997). Monitoring land-cover changes in the Terminos Lagoon Region, Mexico: a comparison of change detection techniques. *Proceedings of the IV International Conference on Remote Sensing for Marine and Coastal Environments, Orlando, FL, USA (Amsterdam: National Aerospace Laboratory)*, 1, 159-167.
- Mas, J. F. (1999). Monitoring land-cover changes: a comparison of change detection techniques. *International Journal of Remote Sensing*, 20, 139-152.
- Mateu, A., & Ruiz, L. A. (1999). Comparación y evaluación de métodos de normalización radiométrica relativa. *Revista de Teledetección*, 12.
- Metternicht, G. (1999). Change Detection Assesment using Fuzzy Sets and Remotely Sensed Data: An Application of Topographic Map Revision. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 221-233.
- Mikolajczyk, K. (2002). Detection of local features invariant to affine transformations, Ph.D. thesis. Institut National Polytechnique de Grenoble, France.
- Mikolajczyk, K., & Schmid, C. (2005). A performance evaluation of local descriptors. *IEEE Transaction of Pattern Analysis & Machine Intelligence*, 27(10), 1615-1630.
- Miller, A. B., Bryant, E. S., & Birnie, R. W. (1998). An analysis of land cover changes in the northern forest of New England using multitemporal Landsat MSS data. *International Journal of Remote Sensing*, 19, 245-265.
- Moravec, H. P. (1977). Towards Automatic Visual Obstacle Avoidance. *Proc. 5th International Joint Conference on Artificial Intelligence*, (pág. 584).
- Mouat, D. A., & Lancaster, J. (1996). Use of remote sensing and GIS to identify vegetation change in the upper San Pedro river watershed, Arizona. *Geocarto International*, 11, 55-67.
- Muchoney, D. M., & Haack, B. N. (1994). Change detection for monitoring forest defoliation. *Photogrammetric Engineering and Remote Sensing*, 60, 1243-1251.
- Muja, M. (2011). *Fast Library for Approximate Nearest Neighbors*. Obtenido de <http://www.cs.ubc.ca/~mariusm/index.php/FLANN/FLANN>
- Muja, M., & Lowe, D. G. (2009). Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. *International Conference on Computer Vision Theory and Applications (VISAPP'09)*, (págs. 331-340).

- Munyati, C. (2000). Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset. *International Journal of Remote Sensing*, 21, 1787-1806.
- Nelson, R. F. (1983). Detecting forest canopy change due to insect activity using Landsat MSS. *Photogrammetric Engineering and Remote Sensing*, 49, 1303-1314.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems Man Cybernet*, 9(1), 62-66.
- Parra, G. A., Mouchot, M. C., & Roux, C. (1996). A multitemporal land-cover change analysis tool using change vector and principal components analysis. *Proceedings of IGARSS Symposium*, 1, págs. 1753-1755. Lincoln, Nebraska, USA.
- Petit, C. C., & Lambin, E. F. (2001). Integration of multi-source remote sensing data for land cover change detection. *International Journal of Geographical Information Science*, 15, 785-803.
- Prakash, A., & Gupta, R. P. (1998). Land-use mapping and change detection in a coal mining area - a case study in the Jharia coalfield, India. *International Journal of Remote Sensing*, 19, 391-410.
- Pratt, W. K. (1991). *Digital image processing*. Ed. John Wiley & Sons.
- Price, K. P., Pike, D. A., & Mendes, L. (1992). Shrub dieback in a semiarid ecosystem: the integration of remote sensing and GIS for detecting vegetation change. *Photogrammetric Engineering and Remote Sensing*, 58, 455-463.
- Prol-Ledesma, R. M., Uribe-Alcantara, E. M., & Diaz-Molina, O. (2002). Use of cartographic data and Landsat TM images to determine land use change in the vicinity of Mexico city. *International Journal of Remote Sensing*, 23, 1927-1933.
- Radeloff, V. C., Mladenoff, D. J., & Boyce, M. S. (1999). Detecting jack pine budworm defoliation using spectral mixture analysis: separating effects from determinants. *Remote Sensing of Environment*, 69, 156-169.
- Ridd, M. K., & Liu, J. (1998). A comparison of four algorithms for change detection in an urban environment. *Remote Sensing of Environment*, 63, 95-100.
- Ridler, T. W., & Calvard, S. (1978). Picture thresholding using an iterative selection method. *IEEE Transaction on Systems, Man and Cybernetics*, 8, 630-632.
- Roberts, D. A., Batista, G. T., Pereira, J. L., Waller, E. K., & Nelson, B. W. (1999). Change identification using multitemporal spectral mixture analysis: application in eastern Amazonia. En R. S. Lunetta, & C. D. Elvidge (Edits.), *Remote*

Sensing Change Detection: Environmental Monitoring Methods and Applications (págs. 137-161). Ann Arbor Press.

Rogan, J., Franklin, J., & Roberts, D. A. (2002). A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery. *Remote Sensing of Environment*, 80, 143-156.

Rosin, P. (2001). Unimodal thresholding. *Pattern Recognition*, 34(11), 2083-2096.

Rosten, E., & Drummond, T. (2006). Machine learning for high-speed corner detection. *European Conference on Computer Vision*, 1, págs. 430-443.

Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: An efficient alternative to SIFT or SURF. *International Conference on Computer Vision*, 2564 - 2571 .

Sader, S. A., & Winne, J. C. (1992). RGB-NDVI color composites for visualizing forest change dynamics. *International Journal of Remote Sensing*, 13, 3055-3067.

Sarabandi, P., Yamazaki, F., Matsuoka, M., & Kiremidjian, A. (2004). Shadow detection and radiometric restoration in satellite high resolution images. *Geoscience and Remote Sensing Symposium (IGARSS)*, 6, 3744-3747.

Serpico, S. B., & Bruzzone, L. (1999). Change detection. En C. H. Chen (Ed.), *Information Processing for Remote Sensing* (págs. 319-336). Singapore: World Scientific Publishing.

Seto, K. C., Woodcock, C. E., Song, C., Huang, X., Lu, J., & Kaufmann, R. K. (2002). Monitoring land-use change in the Pearl River Delta using Landsat TM. *International Journal of Remote Sensing*, 23, 1985-2004.

Shi, J., & Tomasi, C. (1994). Good Features to Track. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (págs. 593-600).

Silpa-Anan, C., & Hartley, R. (2008). Optimised KD-trees for fast image descriptor matching. *Computer Vision and Pattern Recognition*, (págs. 1-8).

Singh, A. (1986). Change detection in the tropical forest environment of northeastern India using Landsat. En M. J. Eden, & J. T. Parry (Edits.), *Remote Sensing and Tropical Land Management* (págs. 237-254). New York: J. Wiley.

Singh, A. (1989). Digital change detection techniques using remotely sensed data. *International Journal of Remote Sensing*, 10, 989-1003.

Slater, J., & Brown, R. (2000). Changing landscapes: monitoring environmentally sensitive areas using satellite imagery. *International Journal of Remote Sensing*, 21, 2753-2767.

- Smith, L. I. (2002). *A tutorial on Principal Components Analysis*. Obtenido de http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf
- Sohl, T. (1999). Change analysis in the United Arab Emirates: an investigation of techniques. *Photogrammetric Engineering and Remote Sensing*, 65, 475-484.
- Stone, T. A., & Lefebvre, P. (1998). Using multitemporal satellite data to evaluate selective logging in Para, Brazil. *International Journal of Remote Sensing*, 19, 2517-2526.
- Sunar, F. (1998). An analysis of changes in a multi-date data set: a case study in the Ikitelli area, Istanbul, Turkey. *International Journal of Remote Sensing*, 19, 225-235.
- Taylor, J. C., Brewer, T. R., & Bird, A. C. (2000). Monitoring landscape change in the national parks of England and Wales using aerial photo interpretation and GIS. *International Journal of Remote Sensing*, 21, 2737-2752.
- Townshend, J. R., & Justice, C. O. (1995). Spatial variability of images and the monitoring of changes in the normalized difference vegetation index. *International Journal of Remote Sensing*, 16, 2187-2195.
- Tsai, V. J. (Junio de 2006). A comparative study on shadow compensation of color aerial images in invariant color models. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6), 1661-1671.
- Tsai, W. (1985). Moment-preserving thresholding: a new approach. *Computer Vision, Graphics and Image Processing*, 29, 337-393.
- Ulbricht, K. A., & Heckendorff, W. D. (1998). Satellite images for recognition of landscape and land use changes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53, 235-243.
- Ustin, S. L., Roberts, D. A., & Hart, Q. J. (1998). Seasonal vegetation patterns in a California coastal savanna derived from Advanced Visible/Infrared Imaging Spectrometer (AVIRIS) data. En R. S. Lunetta, & C. D. Elvidge (Edits.), *Remote Sensing Change Detection: Environmental Monitoring Methods and Applications* (págs. 163-180). Chelsea: Ann Arbor Press.
- VisionBib.Com *Computer Vision Information Pages*. (s.f.). Obtenido de <http://www.visionbib.com>
- Wang, C., Huang, L., & Rosenfeld, A. (1991). Detecting clouds and cloud shadows on aerial photographs. *Pattern Recognition Letters*, 12(1), 55-64.
- Wang, F. (1993). A knowledge-based vision system for detecting land change at urban fringes. *IEEE Transactions on Geoscience and Remote Sensing*, 31, págs. 136-145.

- Ward, D., Phinn, S. R., & Murray, A. T. (2000). Monitoring growth in rapidly urbanizing areas using remotely sensed data. *Professional Geographer*, 52, 371-386.
- Weng, Q. (2002). Land use change analysis in the Zhujian Delta of China using satellite remote sensing, GIS and stochastic modeling. *Journal of Environmental Management*, 64, 273-284.
- Westmoreland, S., & Stow, D. A. (1992). Category identification of changed land-use polygons in an integrated image processing/geographic information system. *Photogrammetric Engineering and Remote Sensing*, 58, 1593-1599.
- Woodcock, C. E., Macomber, S. A., Pax-Lenney, M., & Cohen, W. B. (2001). Monitoring large areas for forest change using Landsat: generalization across space, time and Landsat sensors. *Remote Sensing of Environment*, 78, 194-203.
- Yarbrough, L. D., Easson, G., & Kuzmaul, J. S. (2005). Quickbird 2 Tasseled Cap Transform Coefficients: A comparison of derivation methods. *Pecora 16: Global priorities in land remote sensing*. Sioux Falls, South Dakota.
- Yeh, A. G., & Li, X. (2001). Measurement and monitoring of urban sprawl in a rapidly growing region using entropy. *Photogrammetric Engineering and Remote Sensing*, 67, 83-90.