

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS FÍSICAS
Sección Departamental de Arquitectura de Computadores y
Automática



TÉCNICAS DE CLASIFICACIÓN AUTOMÁTICA DE USO DE
SUELOS AGRÍCOLAS Y FORESTALES BASADAS EN
IMÁGENES DIGITALES

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

Antonia Macedo Cruz

Bajo la dirección de los doctores

Matilde Santos Peñas
Gonzalo Pajares Martinsanz

Madrid, 2013



UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS

**Departamento: Arquitectura de Computadores y
Automática**

Programa de Doctorado: Ingeniería de Sistemas y Automática

**Técnicas de clasificación automática de uso de
suelos agrícolas y forestales basadas en
imágenes digitales**

Presentada por:

Antonia Macedo Cruz

Bajo la dirección de los doctores

Matilde Santos Peñas

Gonzalo Pajares Martinsanz

Madrid, 2012



UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS

**Departamento: Arquitectura de Computadores y
Automática**

Programa de Doctorado: Ingeniería de Sistemas y Automática

**Técnicas de clasificación automática de uso de
suelos agrícolas y forestales basadas en
imágenes digitales**

Presentada por:

Antonia Macedo Cruz

Bajo la dirección de los doctores

Matilde Santos Peñas

Gonzalo Pajares Martinsanz

Madrid, 2012

ÍNDICE GENERAL

<u>ÍNDICE DE FIGURAS.....</u>	<u>V</u>
<u>ÍNDICE DE TABLAS</u>	<u>VII</u>
<u>AGRADECIMIENTOS</u>	<u>IX</u>
<u>RESUMEN.....</u>	<u>XIII</u>
<u>CAPÍTULO 1 INTRODUCCIÓN.....</u>	<u>3</u>
1.1. MOTIVACIÓN.	3
1.2. OBJETIVOS.....	6
1.3. METODOLOGÍA	7
1.4. APORTACIONES DE LA INVESTIGACIÓN	8
1.5. ORGANIZACIÓN DE LA TESIS	13
<u>CAPÍTULO 2 ESTADO DEL ARTE</u>	<u>17</u>
2.1. VISIÓN GENERAL	17
2.2. ESTADO DEL ARTE DEL RECONOCIMIENTO DE PATRONES.	18
2.3. REPRESENTACIÓN Y PROCESAMIENTO DE IMÁGENES DIGITALES	26
2.3.1. ADQUISICIÓN DE DATOS	26
2.3.2. MEJORAMIENTO DE CONTRASTE	31
2.3.3. SEGMENTACIÓN DE IMÁGENES EN COLOR.....	37
2.3.4. CLASIFICACIÓN NO SUPERVISADA.....	45

2.4. MÉTODOS DE SEGMENTACIÓN POR UMBRALIZACIÓN	50
2.4.1. MÉTODO DE OTSU	55
2.4.2. UMBRALIZACIÓN POR ISODATA	58
2.4.3. ALGORITMO DE UMBRALIZACIÓN FUZZY	59
2.5. MÉTODOS DE AGRUPAMIENTO	60
2.5.1. AGRUPAMIENTO <i>FUZZY C-MEANS</i>	63
2.6. EVALUACIÓN DEL SISTEMA DE CLASIFICACIÓN NO SUPERVISADA DE COBERTURA TERRESTRE.....	67
2.6.1. ÍNDICES DE VALIDACIÓN PARA UN CLASIFICADOR NO SUPERVISADO.....	68
2.6.2. VALIDACIÓN DEL NÚMERO DE AGRUPAMIENTOS (ÍNDICES DE VALIDACIÓN INTERNA).....	69
2.6.3. INFORMACIÓN DE REFERENCIA PARA LA VALIDACIÓN TEMÁTICA.....	75
2.6.4. MATRICES DE ERROR	81
2.7. RESUMEN DEL ANÁLISIS BIBLIOGRÁFICO	89
<u>CAPÍTULO 3 APLICACIÓN DE ESTRATEGIAS DE PROCESAMIENTO Y CLASIFICACIÓN A</u>	
<u>IMÁGENES AÉREAS DE COBERTURA TERRESTRE.....</u>	95
3.1. INTRODUCCIÓN	95
3.2. RECOPIACIÓN Y PROCESAMIENTO DE LA INFORMACIÓN	96
3.2.1. MEJORA DE LA CALIDAD DE LA IMAGEN	99
3.3. ESTRATEGIA DE CLASIFICACIÓN NO SUPERVISADA.....	101
3.3.1. SEGMENTACIÓN Y CODIFICACIÓN.....	102
3.3.2. CLASIFICACIÓN	113
3.4. APLICACIÓN DEL MÉTODO DE CLASIFICACIÓN NO SUPERVISADA PROPUESTO	115
3.4.1. SEGMENTACIÓN DE PÍXELES POR UMBRALIZACIÓN SIMPLE Y CODIFICACIÓN.....	115
3.4.2. CLASIFICACIÓN	121

3.4.3. INTERPRETACIÓN DE RESULTADOS	123
3.5. APLICACIÓN DEL CLASIFICADOR <i>OTSU-CLUSTERING</i> PROPUESTO.....	124
3.6. VALIDACIÓN DEL CLASIFICADOR.....	129
3.6.1. VALIDACIÓN TEMÁTICA MEDIANTE LA MATRIZ DE ERROR	129
3.6.2. UNIDADES DE VALIDACIÓN.....	131
3.6.3. OBTENCIÓN DE LAS MATRICES DE VALIDACIÓN.....	134
3.6.4. VALIDACIÓN DEL NÚMERO DE CLASES.....	139
3.7. CONCLUSIONES DEL CAPÍTULO	143
<u>CAPÍTULO 4 APLICACIÓN DE ESTRATEGIAS DE PROCESAMIENTO Y CLASIFICACIÓN A</u>	
<u>IMÁGENES AGRÍCOLAS CAPTADAS POR CÁMARAS CONVENCIONALES.</u>	<u>147</u>
4.1. INTRODUCCIÓN	147
4.2. RECOPIACIÓN Y PRE-PROCESAMIENTO DE LA INFORMACIÓN	149
4.2.1. MUESTREO DE LA ZONA DE ESTUDIO	151
4.2.2. PRE-PROCESAMIENTO DE LAS IMÁGENES A COLOR	153
4.3. APLICACIÓN DE LA ESTRATEGIA DE CLASIFICACIÓN NO SUPERVISADA	153
4.3.1. SEGMENTACIÓN DE IMÁGENES A COLOR POR COMBINACIÓN DE UMBRALES	154
4.3.2. CODIFICACIÓN DE LA INFORMACIÓN	156
4.3.3. CLASIFICACIÓN DE LA IMAGEN EN COLOR	157
4.4. RESULTADOS Y DISCUSIÓN	159
4.4.1. DESCRIPCIÓN DEL CULTIVO DE AVENA	159
4.4.2. VALIDACIÓN DEL CLASIFICADOR.....	162
4.4.3. RESUMEN Y DISCUSIÓN	172
4.5. CONCLUSIONES DEL CAPÍTULO	174

<u>CAPÍTULO 5 CONCLUSIONES Y TRABAJOS FUTUROS</u>	179
5.1. CONCLUSIONES	179
5.2. LÍNEAS FUTURAS DE TRABAJO.	182
<u>BIBLIOGRAFÍA</u>	185
<u>ANEXO 1. VARIANZAS DENTRO DE CADA CLASE Y ENTRE CADA PAR DE CLASES, DESPUÉS DE APLICAR UMBRALIZACIÓN MÚLTIPLE</u>	213
<u>ANEXO 2. RESULTADOS DE LOS NÚMEROS ÓPTIMOS DE AGRUPAMIENTOS</u>	217
<u>ANEXO 3. TEORÍA DEL COLOR</u>	221
INTRODUCCIÓN A LA TEORÍA DEL COLOR	221
ESPACIO DE COLOR RGB	224
ESPACIOS DE COLORES UNIFORMES EN CUANTO A PERCEPCIÓN	226
ESPACIO DE COLOR L*A*B*	227
ESPACIO DE COLOR L*U*V*	228

ÍNDICE DE FIGURAS

Figura 2.1. Bandas del espectro electromagnético.	29
Figura 2.2. Representación espacial de un píxel en una localización x,y correspondiente a una imagen con tres componentes espectrales.....	31
Figura 2.3. Principios de la técnica de mejoramiento del contraste.	34
Figura 2.4. Esquema comúnmente usado para la segmentación de imágenes a color. ..	42
Figura 2.5. Tipos de muestreo de campo.....	81
Figura 3.1. Esquema del procesamiento y clasificación de las imágenes.	96
Figura 3.2. Localización de la cuenca del río La Sabana, estado de Guerrero, México.	97
Figura 3.3. Imagen e histograma modelo.	99
Figura 3.4. Imágenes que necesitan ecualización de luminancia.	100
Figura 3.5. Igualación del histograma.	100
Figura 3.6. Imagen en color representada por mapas bidimensionales con representación de valores RGB.	105
Figura 3.7. Binarización de una imagen a color.....	107
Figura 3.8. Ejemplo del proceso de codificación de la imagen en color.....	112
Figura 3.9. Fotografía aérea en color RGB.	116
Figura 3.10. Histograma que muestra la distribución de los niveles digitales de la imagen RGB.	117
Figura 3.11. Gráfico de separabilidad por nivel digital.....	118
Figura 3.12. Resultado de la codificación en dos clases por componente espectral	118
Figura 3.13. Imagen resultante de la mezcla de las tres componentes espectrales después de ser clasificado cada uno en dos clases de textura.....	121

Figura 3.14. Imagen clasificada.....	122
Figura 3.15. Imagen original en modelo de color RGB, con más detalles.....	124
Figura 3.16. Clasificación de cobertura terrestre con nivel digital promedio.	128
Figura 3.17. Clasificación de cobertura terrestre por asignación de colores por píxel y clase.....	128
Figura 3.18. Ejemplo de comparación entre imagen de referencia e imagen clasificada a nivel de píxel.	131
Figura 3.19. Ejemplo de comparación entre imagen de referencia e imagen clasificada tomando los polígonos como unidades de validación.	132
Figura 3.20. Comparación visual de la clasificación obtenida por <i>Otsu-Clustering</i> y <i>Fuzzy C-Means</i>	142
Figura 4.1. Diagrama del sistema de clasificación no supervisada de imágenes a color por combinación de umbrales y reagrupamiento de clases.	149
Figura 4.2 Selección y delimitación de las muestras del cultivo de avena para ser fotografiado.	152
Figura 4.3. Muestra de cultivo de avena poco afectado por las heladas.....	161
Figura 4.4 Muestra de cultivo de avena muy afectado por las heladas.....	161
Figura 4.5 Muestra de cultivo de avena (toma superficial).....	162
Figura 4.6. Datos de referencia: unidades de validación.....	164
Figura 4.7. Trazado de polígonos de muestreo.....	165
Figura A3.1. Diagrama de cromaticidad.	222
Figura A3.2 Tetraedro de color RGB. Los puntos a lo largo de la diagonal principal tienen valores de gris, desde el negro en el origen al punto (1 1 1).	226
Figura A3.3. Modelo de color CIELAB.	228
Figura A3.4. Diagrama de cromaticidad de las coordenadas u' y v'	229

ÍNDICE DE TABLAS

Tabla 2.1. Ejemplo matemático de una matriz de error.....84

Tabla 2.2. Ejemplo matemático de una matriz de error *fuzzy*.....87

Tabla 3.1. Momentos estadísticos del histograma por componente espectral y valor de umbral de la imagen de la figura 3.9.119

Tabla 3.2. Nivel digital promedio por clase y componente espectral de la figura 3.13.120

Tabla 3.3. Varianza dentro de cada clase y entre clases de la figura 3.13.....121

Tabla 3.4. Nivel digital promedio de la imagen clasificada (figura 3.14), por clase y componente espectral.122

Tabla 3.5. Reconocimiento de la cobertura terrestre de la figura 3.9 *a*.....123

Tabla 3.6. Varianza dentro de cada clase y entre clases (figura 3.15).....125

Tabla 3.7. Umbrales de cada componente espectral de la figura 3.15.....126

Tabla 3.8. Nivel digital promedio por clase y por componente espectral con dos126

Tabla 3.9. Nivel digital promedio por clase y componente espectral del resultado de la clasificación final.....127

Tabla 3.10 Varianzas dentro de cada clase (diagonal principal) y entre cada par de clases (fila-columna), después de aplicar umbralización múltiple y reagrupamiento de clases solapadas.....127

Tabla 3.11. Distribución del número de muestras (unidad de validación el píxel) por categoría temática.....133

Tabla 3.12. Distribución del número de muestras (unidad de validación el polígono) por categoría temática.....133

Tabla 3.13. Matriz de error fuzzy, las muestras son los píxeles.....136

Tabla 3.14. Matriz de error fuzzy, las muestras son los polígonos.	137
Tabla 3.15. Número óptimo de agrupamientos según 5 índices.....	141
Tabla 4.1. Matriz de error Fuzzy y Determinista.....	167
Tabla 4.2. Precisión del clasificador y errores de comisión	169
Tabla 4.3. Precisión del experto y errores de omisión.....	171
Tabla 4.4. Precisión del clasificador y del experto por umbralización combinada (CT) y simple (IS, OT, FU).....	171
Tabla A3.9. Varianzas dentro de cada clase (diagonal principal) y entre cada par de clases (fila-columna), después de aplicar umbralización múltiple.....	213

AGRADECIMIENTOS

El que esta tesis se haya realizado y terminado ha sido debido al esfuerzo y ánimo de mucha gente para la que debo y quiero tener unas palabras de agradecimiento.

En primer lugar quiero agradecer a mis directores de tesis, Dra. Matilde Santos Peñas por su invaluable apoyo, por haberme acogido desde un primer momento en este hermoso país y por todas sus enseñanzas durante el proceso doctoral tanto como coordinadora del programa de doctorado así como directora de tesis, y profesora del curso de Control Inteligente, además de su preocupación por mi bien estar tanto a nivel profesional como personal a lo largo de todos estos años. Gracias a sus enseñanzas ayuda paciencia y comprensión para corregirme y animarme a seguir adelante, y sobre todo gracias a esa confianza que puso en mi. Gracias por esos momentos que me regalaste para charlar en horas de comida. Quiero agradecerle también la gentileza que tuvo al compartirme su cubículo durante la estancia de investigación que nos financió el Banco Mundial. Su apoyo, ayuda y dedicación han sido encomiables.

A mi otro director Dr. Gonzalo Pajares Martinsanz quiero agradecerle por su continuo apoyo, guía y enseñanzas durante todos estos años, como director y profesor del curso Visión por Computador. Gracias a los dos, ha sido una experiencia estupenda y muy gratificante, espero poder seguir manteniendo el contacto de ambos en el futuro.

A mis profesores Sebastián Dormido Bencomo y Alfonso Urquía Moraleda por sus enseñanzas durante el curso de Simulación de Sistemas Dinámicos; a mis profesores Pérez de Madrid y Pablo, Carolina Mañoso Hierro, por sus enseñanzas durante el curso de Control Predictivo; a todos ellos les quiero agradecer porque me encaminaron hacia el conocimiento científico y tecnológico que soporta esta investigación.

A los miembros del Tribunal designado para la defensa de la Tesis Doctoral, Dr. Jesús Manuel de la Cruz García, Dra. Ángela Ribeiro Seijas, Dr. José Antonio Martín

Hernández, Dr. José María Girón Sierra, por su valioso tiempo y dedicación para revisar este trabajo de tesis.

A todos los profesores y compañeros de la facultad de Ciencias Físicas y de la Facultad de Informática, que me acogieron con cariño y siempre me dieron su apoyo durante mi estancia en España, a Pedro y todos, Gracias.

A la Fundación Carolina (España) y la Secretaría de Relaciones Exteriores-(México), por el co-financiamiento de beca para estudios doctorales en la Universidad Complutense de Madrid, España, y en especial a los coordinadores de los convenios A Dña. Piedi Martín Díez de Baldeón quien como Coordinadora de Becas de la Fundación Carolina-España, me brindó con esmero y amabilidad su valioso apoyo; a Dña. Ma. De Jesús Urtarán, responsable de la embajada de México en España.

A Eva Luz Hernández Durán como subdirectora de intercambio académico para Europa, de la Secretaría de Relaciones Exteriores en México, por las gestiones realizadas para hacer posible el apoyo económico del cofinanciamiento de beca.

Al Banco Mundial por el financiamiento económico para desarrollar una estancia de investigación en Madrid España, bajo el programa de becas “Robert S. McNamara Fellowships Program”, esto después de concluirse el financiamiento de la Fundación Carolina y Secretaria de Relaciones Exteriores, México.

Al Colegio de Postgraduados por el apoyo brindado, a través del Subprograma de Formación de Profesores Investigadores, y en especial a la Dra. Ma. Cristina Guadalupe López Peralta y la Lic. Ma. Lourdes Larque Saavedra, por todas las facilidades brindadas para realizar las gestiones administrativas internas en el Colegio de Postgraduados.

Al Dr. Enrique Mejía Sáenz por todo el apoyo que siempre me brindó como enlace académico del postgrado de Hidrociencias. Al Dr. Leonardo Tijerina, Dr. Exebio, y al Dr. Gavi por sus buenas recomendaciones durante las gestiones de beca, y a todos los profesores y compañeros del postgrado en Hidrociencias que de alguna manera siempre me expresaron sus buenos deseos.

Todo lo que soy o algún día seré se lo debo a mi familia, su constante apoyo me han hecho caminar por la vida con paso firme, ellos son lo más importante para mí, por tanto quiero dedicar esta tesis a lo que más quiero en este mundo, a mis hijos Angélica Yuridia y Christopher Ulianov, por ellos todo merece la pena. Pero no tan solo debo agradecerles, también les debo pedir disculpas por no haberles atendido como merecen. Espero que algún día, cuando vean este trabajo, sepan comprender por qué su madre pasaba tanto tiempo delante del ordenador y de vez en cuando perdía los nervios sin motivo aparente. A mi esposo Isidro Villegas Romero que hizo hasta lo imposible por conseguir su permiso sabático y posdoctoral para poder estar con nosotros durante ésta etapa tan importante de mi vida, el es, sin duda, el que mejor me conoce y por tanto el que más mérito tiene en aguantarme. Te quiero.

Pero nada de todo esto se habría hecho posible sin la enseñanza que con tanto esmero he recibido de mis padres, mis dos grandes referencias en la vida, mi padre Donato Macedo Avellaneda que con tanto cariño me dio los mejores valores, el ha sido siempre para mí un ejemplo de esfuerzo y coraje en la vida. Mi padre quien siempre me animó a luchar y a seguir adelante, y aunque sé que le dolía saber que a su avanzada edad no nos podríamos ver y abrazar por algunos años siempre se guardo esos sentimientos y me dio las más grandes bendiciones. Y a pesar de la diferencia de horarios y la distancia entre México y España siempre estuvo pendiente de la comunicación entre ambos. El 1 de septiembre del 2010 dios lo acogió en su gloria, pero sé que donde está me sigue bendiciendo y está muy orgulloso de ser mi padre, como siempre me lo expresó. Gracias papacito por todo el amor incondicional que siempre me diste.

A mi madre Ma. Félix Cruz, por su cariño y bendiciones que me ha expresado siempre.

A mis hermanas Adabella y Herlinda Macedo Cruz, por todo su cariño y apoyo, amigas y compañeras en las buenas y en las malas, sea por mensaje o por teléfono siempre estamos unidas; las quiero mucho mis dos hermosas hermanas. A mis sobrinos Juan, Kevin, Brandon, Stephanie y Brian.

A mis padres de pila, mamá Cristina Macedo y Wences Cárdenas una pareja ejemplar, gracias por todo el cariño y comprensión, gracias por esas palabras de ánimo que siempre

estuvieron presentes para con nosotros. Mamá Cristina, mi hermana y madrina, dios la acogió en su santa gloria el 20 de enero del 2008, siento mucho no haber estado con ustedes en esos momentos, pero sé que desde el cielo me bendices y que si estuvieras con nosotros físicamente me acompañarías a Madrid a la defensa de tesis. Los tengo presentes en mi corazón siempre y a todos mis hermanos de pila y sobrinos. Los quiero.

A mis hermanos de padre, Rosa, Aurelia y Dámaso Macedo, por ese cariño que siempre me expresaron, el 20 de enero del 2010 mi hermano Dámaso fue acogido por el creador, se que desde ahí me vez y me bendices manito.

A María de Jesús y mi ahijada Kristal que siempre han estado con migo y me han apoyado siempre.

Sé que tengo una familia muy numerosa y no quisiera dejar de mencionar a ninguno en este trabajo tan importante para mí, quiero decirles que todos han contribuido de alguna manera para que esto llegue a su fin, gracias tíos y tías por esas bendiciones y buenos deseos, primos, primas, sobrinos, ahijadas y ahijados, comadres, cuñadas, vecinos y amigos por todos esos momentos que hemos compartido y esos buenos deseos.

Por último quiero dar gracias Dios por darme más de lo que merezco, gracias por cada minuto que me das, gracias por la familia que me elegiste y por todas esas personas maravillosas que has puesto en mi camino. Gracias.

RESUMEN

A través de los años se han desarrollado diversos algoritmos de clasificación de imágenes captadas por sensores aerotransportados, incluso existe software comercial de Sistemas de Información Geográfica (GIS, *Geographic Information Systems*) que incluyen, dentro de sus extensiones, la clasificación de imágenes como una herramienta de utilidad. La mayoría de los métodos que se incorporan en estas aplicaciones son sobre todo modelos de clasificación supervisada, y los no supervisados requieren por lo general el conocimiento de cierta información, por ejemplo, el número de clases o el intervalo de clases, tal como ocurre con los algoritmos de clasificación llamados *c*-medias (*c-mean*), *k*-medias (*k-mean*), o los *c*-medias-difusos (*fuzzy-c-mean*). Uno de los algoritmos no supervisados que ha demostrado una gran eficiencia en clasificación de imágenes en general y aéreas en particular es el conocido como ISODATA; si bien no resulta fácil su aplicación ya que habitualmente el usuario no conoce, a priori, determinados parámetros necesarios para el método tales como el número de grupos espectrales presentes en una imagen, el valor idóneo de dispersión interna de las muestras en los grupos o las distancias entre grupos. Para resolver el problema de configurar los parámetros del algoritmo se establece un proceso interactivo, donde el usuario establece unos valores arbitrarios iniciales que se van modificando a la luz de los resultados obtenidos en las sucesivas pruebas de verificación.

En esta tesis doctoral se trabaja con imágenes aéreas a color e imágenes captadas con cámaras convencionales, el análisis se basa en el estudio de las tres bandas correspondientes al espectro visible, que muestran la cobertura de distintos tipos de vegetación u otros elementos en el suelo, con una aplicación directa en el ámbito de la agricultura. Llámense *imágenes aéreas* a las imágenes digitales captadas por sensores aerotransportados (fotografía aérea en formato digital). Por el contrario las *imágenes captadas por cámaras convencionales*, son aquellas obtenidas a nivel de suelo mediante una cámara digital de uso común. Thorp y Tian (2004) denominan a los del primer grupo sensores remotos y al segundo sensores terrestres. Se han aplicado diferentes

métodos y algoritmos para procesar estas imágenes con el fin de obtener información relevante que permita tomar decisiones y medidas que beneficien la producción agrícola. Puesto que se han analizado dos casos de estudio distintos, la metodología varió un poco para cada caso. Así, en el caso de la clasificación de texturas relativas a la cobertura terrestre captada en las fotografías aéreas, en primer lugar se extrae la información más relevante de cada componente espectral de las imágenes mediante una técnica de segmentación por umbralización (simple o múltiple), aplicando el método de Otsu (modificado) sobre cada componente espectral, de manera que al fusionar las tres componentes umbralizadas se reduce por un lado el tamaño de la imagen resultante, a la vez que se mantienen las características necesarias para establecer una primera clasificación. En segundo lugar, se aplican procesos de reagrupamiento de aquellas clases que presentan solapamiento, de forma que se llegue a un número de clases mínimo pero significativo (óptimo). La etapa de reagrupamiento es conocida como etapa de decisión, ya que es en la cual se decide a qué clase pertenece cada píxel.

En cuanto a la aplicación agrícola derivada de las imágenes captadas por cámaras convencionales, y puesto que la zona es fácilmente accesible, en la etapa de segmentación se hicieron algunos cambios respecto al caso anterior, de manera que se aplicaron tres algoritmos de segmentación por umbralización simple (el método de Otsu (Otsu 1979), el algoritmo de umbralización *Fuzzy* (Huang y Wang 1995) y el algoritmo Isodata (Ridler y Calvard, 1978), así como la combinación de los mismos.

En cuanto a los modelos de color, se trabajó con el modelo de color RGB, y con el modelo de color CIELab (Robertson 1976).

Con el fin de verificar la bondad y calidad de las clasificaciones se han aplicado métodos de validación de las particiones obtenidas con la clasificación, analizando la separación entre las clases, así como dentro de éstas mediante índices internos, tales como el índice de Davies y Bouldin (1979), de Calinski y Harabasz (1974), de Krzanowski y Lai (1985), de Dunn (1974) y de Hartigan (1985). También se ha utilizado lo que se conoce como índice externo para evaluar la precisión temática, tal es el caso de los denominados índices de Kappa, los cuales se obtienen mediante la matriz

de error determinista y difusa. Esta evaluación se ha llevado a cabo elaborando previamente las imágenes que sirven de base para el análisis conocidas técnicamente como *ground truth*, tanto para la zona que comprende las fotografías aéreas como para la zona analizada con imágenes captadas desde tierra. En ambos casos se han obtenido resultados satisfactorios.

Capítulo 1

Introducción



CAPÍTULO 1 INTRODUCCIÓN**1.1. MOTIVACIÓN.**

En México, la principal institución encargada de desarrollar la cartografía temática oficial es el Instituto Nacional de Estadística y Geografía (INEGI). Sin embargo, la Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) en el Programa Estratégico Forestal para México 2025, afirma que los inventarios forestales realizados hasta ahora han sido en general poco útiles para fines de planificación y acciones de producción y restauración agrícolas, destacando que no se cuenta con información suficientemente fiable acerca de la superficie forestal del país tanto por entidad federativa como por cuencas y por especies. Es decir, a pesar de esos estudios, no se dispone de datos periódicamente actualizados que constituyan una base sólida para la planificación forestal.

La cartografía temática desarrollada por INEGI sobre las coberturas nacionales de uso del suelo y vegetación están hechas a escala mediana, 1:250.000. Se denomina serie I y fue elaborada en el periodo de 1979 a 1991. Las técnicas y las herramientas que se utilizaron para generar esta información fueron fotografías aéreas en formato de 23 x 23 cm de ancho y alto, obtenidas a escala 1:50.000, y fotografías aéreas del Sistema Nacional de Fotografía Aérea (SINFA) a escala 1:75.000, todas ellas en blanco y negro. Esta información fue volcada en un mapa a escala 1:250.000, siendo la base que se toma para la mayoría de los estudios geográficos que se desarrollan en México. Algunos de los problemas que lo hacen ineficiente provienen del hecho de que en algunos casos el tamaño promedio de parcelas es tan pequeño que no es visible a esta escala.

La Ley General de Desarrollo Forestal Sostenible (reforma del 2008, Título tercero de la política nacional en materia forestal, Capítulo II, Artículos 44, 45 y 57, y el Plan Estratégico Forestal (PEF) para México 2025, así como el Sistema Nacional de Información Estadística y Geográfica (SNIEG) coordinado por el INEGI, mencionan la necesidad de desarrollar un Sistema Nacional de Información Ambiental y de Recursos Naturales, así como el Inventario Nacional Forestal y de Suelos. Para ello es necesario desarrollar una adecuada metodología de clasificación automática de imágenes que permita procesar eficientemente grandes cantidades de información (puesto que las características de las imágenes aéreas o de satélite así lo requieren), y generar esa base

de información con los diferentes grupos de cobertura terrestre presentes en imágenes digitales, sean aéreas, tomadas desde un satélite o desde tierra.

La necesidad de la clasificación automática o no supervisada surge por el hecho de que con frecuencia no se conoce a priori ni la cobertura vegetal ni los tipos de suelos, debido fundamentalmente a los diferentes factores presentes en las imágenes (aéreas y terrestres), por lo que resulta ciertamente difícil fijar determinados parámetros que requieren los métodos de clasificación no automáticos y por tanto de naturaleza supervisada. Además, de esta manera se asegura que las zonas inaccesibles en las que se desconoce la cobertura vegetal o tipos de suelo puedan también considerarse mediante el sistema de clasificación, sin necesidad de excluir posibles áreas no identificadas por el hecho de limitar el número de clases en dicha clasificación.

Además, México necesita consolidar un diagnóstico de peligros, vulnerabilidad y riesgo que pueda ser actualizado de manera constante dada la geografía y condiciones climatológicas del país. De manera que cuando se trata de reconocer los daños causados por eventos meteorológicos como huracanes, que es el caso de las imágenes aéreas analizadas en el presente trabajo, las zonas se vuelven inaccesibles, por lo que el reconocimiento automático es la única alternativa.

La metodología de clasificación debe poderse aplicar a imágenes en color con al menos las tres bandas espectrales características del espectro visible, puesto que, cuando ocurre algún siniestro y se sobrevuelan las zonas afectadas, estas imágenes son las más fáciles y económicas de obtener.

Otro campo de aplicación pertinente y necesaria se da en el área agrícola, por lo que una de las aplicaciones propuestas se desarrolla en torno a la cuantificación de la densidad de planta afectada y no afectada por bajas temperaturas, por tanto, la cobertura pasa a ser irrelevante cuando los cultivos están muy afectados por las heladas. Además, el suelo suele estar completamente cubierto por las plantas, en el caso que nos ocupa de avena, que es el cultivo de cereal analizado, debido a la extensa cobertura vegetal los métodos basados en el cálculo de índices de vegetación (comúnmente usados) son inviables o inadecuados.

Cuando las imágenes son capturadas bajo condiciones de iluminación natural, la respuesta del sensor es proporcional a la energía de la luz proyectada sobre la superficie.

Esta energía depende de las longitudes de onda correspondientes al espectro visible reflejada por los objetos (plantas y suelo). Cada longitud de onda produce una respuesta diferente que se asigna a un color estándar. Por tanto, para hacer frente al problema de iluminación una de las propuestas existentes en la literatura consiste en usar un espacio de color alternativo al modelo RGB, tal es el caso del CIELAB o $L^*a^*b^*$ (Robertson, 1976).

Con el anterior planteamiento, y para probar la eficacia del sistema de clasificación no supervisada, diseñado y descrito en el presente trabajo, se solicitó al Instituto de Geografía de la UNAM las imágenes aéreas captadas mediante el vuelo realizado después del ciclón tropical “huracán Paulina” que impactó las cuencas hidrográficas de los estados de Guerrero y Oaxaca. Este fenómeno fue originado por un meteoro que causó un alto impacto ecológico debido a que provocó inundaciones, erosión hídrica, movimientos en masa y generó gran cantidad de sedimentos (Villegas y col., 2009). En dichas imágenes el objetivo principal de la clasificación lo constituyen las parcelas inundadas de la cuenca hidrográfica del río La Sabana, Guerrero. Para el estudio del impacto del huracán sobre los suelos afectados, se cuenta con fotografías tomadas por un sensor digital Kodak DCS-420c aerotransportado, que proporciona imágenes espectrales en el modelo de color RGB. Además, dadas las características de la zona de estudio, se sabe que existen algunos lugares inaccesibles físicamente, por lo que se desconoce el número de clases de cobertura posibles a identificar.

Por otro lado, se dispone de fotografías captadas por un sensor digital Canon CCD 1/2.3'', tomadas desde tierra al cultivo de avena afectada por bajas temperaturas (heladas) ocurridas durante el mes de octubre de 2010 en México. En este caso se trata de evaluar el impacto de este otro fenómeno meteorológico sobre el cultivo y de forma indirecta en el suelo. Las fotografías se tomaron 15 días después de la última helada en un ambiente al aire libre, con alta variabilidad de las condiciones de iluminación natural. Como se ha mencionado previamente, para hacer frente al problema de iluminación se aplica el espacio de color CIELAB o $L^*a^*b^*$ (Robertson, 1976), sin descartar el modelo RGB, que también se ha utilizado, como se verá posteriormente. El objetivo de investigación con estas imágenes consiste en clasificar la cobertura terrestre con cultivo de avena, y cuantificar la densidad de plantas afectadas y no afectadas por las heladas.

1.2. OBJETIVOS

Una vez planteada la necesidad de desarrollar un sistema de clasificación automático para analizar la cobertura del suelo, con fines forestales y agrícolas, se deducen los objetivos generales de la presente Tesis Doctoral que se enuncian a continuación:

1. Diseñar un sistema de clasificación inteligente (automático y totalmente no supervisado), que sea capaz de encontrar las diferentes clases de cobertura terrestre presentes en una imagen digital en color. La información generada debe servir como base a un experto para la identificación y clasificación del uso del suelo y evaluación de daños, incluyendo cultivos de avena.
2. Configurar el sistema de forma que sea capaz de detectar automáticamente los grupos existentes en la fotografía aérea digital sin que requiera de otros parámetros de entrada ni otro tipo de información adicional excepto la imagen a procesar. Esto facilitará el procesamiento de grandes cantidades de datos.
3. Configurar el sistema de forma que sea capaz de detectar mediante análisis de color los distintos tipos de suelos y en el caso de las plantas de avena, las partes verdes (no afectadas), verde amarillento (poco afectadas), secas (muy afectadas) por las heladas.
4. Generar medidas cuantitativas y cualitativas tanto para la validación temática como para el número de agrupamientos que resulta de aplicar el clasificador automático.

De estos objetivos generales se desprenden los siguientes objetivos específicos:

- Recopilar y analizar documentación sobre el estado del arte en torno a la clasificación no supervisada de imágenes a color.
- Diseñar la estrategia de clasificación no supervisada (automática) para imágenes en color. Para ello, se probarán distintas estrategias, se modificarán, fusionarán y compararán para analizar su eficiencia para esta aplicación concreta.

- Aplicar los índices de validación de las clasificaciones que permiten determinar el número óptimo de agrupamientos, así como la veracidad de la clasificación temática de cada píxel.
- Generar nuevas figuras para medir la eficiencia del sistema de clasificación.
- Generar los datos de referencia o patrones que permitan diseñar un esquema de validación temática para medir la fiabilidad del clasificador.
- Analizar los resultados de la clasificación cuantitativa y cualitativamente con respecto a la información que proporcionan y el objetivo de este procesamiento.
- Divulgar y difundir los resultados de la investigación mediante publicaciones científicas y presentaciones en congresos y otros foros científicos.

1.3. METODOLOGÍA

La presente investigación reflejada en esta memoria se ha realizado siguiendo las directrices marcadas por el método científico. Inicialmente se ha realizado una búsqueda de información bibliográfica en todas las fuentes disponibles, utilizando principalmente para ello el acceso a través de la red del Consorcio Madroño al que están suscritas las Universidades de la Comunidad de Madrid, y en concreto la Universidad Complutense. Se han estudiado los antecedentes en el área. Posteriormente, se han desarrollado las etapas de trabajo que se indican a continuación, manteniendo actualizadas las referencias bibliográficas mediante una revisión continua de las mismas.

Las etapas de desarrollo seguidas han sido:

- Definición del problema a raíz tanto de las necesidades identificadas como de la motivación establecida previamente.
- Planteamiento de los objetivos de la investigación.
- Diseño y desarrollo de la metodología.
- Obtención de resultados experimentales.
- Clasificación, análisis e interpretación de los datos.
- Diseño y desarrollo de la estrategia de validación del clasificador.
- Conclusiones generales.

- Descripción de líneas futuras de investigación abiertas durante el proceso.

Este proceso abarca desde la adquisición de datos, pre-procesamiento de la información, desarrollo de las diferentes etapas del sistema de clasificación, hasta la obtención de resultados y validación del sistema.

1.4. APORTACIONES DE LA INVESTIGACIÓN

A partir de los objetivos propuestos y teniendo en cuenta los aspectos destacados anteriormente, se ha tratado de dar solución a los problemas planteados y de difundir en la comunidad científica una serie de estrategias que se han mostrado eficaces para el tipo de aplicación propuesto, las cuales poseen el aporte añadido de que pueden extenderse a otros problemas de naturaleza similar. Las contribuciones realizadas en el marco del presente trabajo de investigación se recogen en las siguientes publicaciones:

Artículos publicados en revistas indexadas en el Journal Citation Report (JCR)

Autores: A. Macedo, G. Pajares, M. Santos

Título: “Unsupervised classification with ground cover color images”

Revista: *Agrociencia*; ISSN: 1405-3195

Volumen, pág. (año): 44, 6, 711-722, 2010

Calidad: JCR 2010: 0.294; *Agriculture, multidisciplinary* 33/55 (Q3)

1. En este trabajo se presenta la estrategia de clasificación aplicada a imágenes a color captadas por sensores aéreos, para diferentes tipos de cobertura terrestre presentes en la imagen. El espacio de color usado para la clasificación fue el RGB, y la etapa de segmentación se llevó a cabo mediante el método de Otsu (1979) para la umbralización simple de las tres componentes espectrales; además se extendió el método de Otsu para la multi-umbralización de cada componente espectral. En este artículo se muestra claramente la estrategia que sigue el sistema para decidir la necesidad de aplicar umbralización simple o múltiple, así como la eficiencia del reagrupamiento de clases hasta lograr que la similitud dentro de cada clase de cobertura sea menor que la varianza entre las mismas. Este trabajo es, junto con el anterior, precursor de la investigación desarrollada en relación a los contenidos de la presente memoria (Capítulo 3). En concreto, la incidencia más relevante estriba en el concepto de multi-umbralización y la

estrategia de clasificación, que mediante la evaluación de la similitud dentro de los grupos y separación entre grupos se logra un número óptimo de clases.

Autores: A. Macedo-Cruz, G. Pajares, M. Santos, I. Villegas-Romero
Título: “Digital Image Sensor-Based Assessment of the Status of Oat (Avena sativa L.) Crops after Frost Damage”
Revista: *Sensors*; ISSN: 1424-8220
Volumen, pág. (año): 11, 6, 6015-6036, 2011
Índices de Calidad: JCR 2010: 1.771; Instruments & Instrumentation 14/61 (Q1)

2. En este trabajo se aplica la estrategia de clasificación no supervisada para cuantificar la densidad de plantas de avena afectadas y no afectadas por heladas; las imágenes utilizadas fueron tomadas por una cámara convencional con un sensor digital CCD. El espacio de color usado para la clasificación es $L^*a^*b^*$, y la etapa de segmentación se llevó a cabo mediante la combinación de tres estrategias de umbralización: *a*) el método de Otsu (Otsu, 1979); *b*) el algoritmo Isodata (Ridler y Calvard, 1978) y *c*) la umbralización *Fuzzy* (Huang y Wang 1995). La fusión de estas técnicas de umbralización automática, el diseño de la estrategia de clasificación y la metodología de validación temática, fueron las principales contribuciones del artículo, que generó resultados totalmente satisfactorios. Los planteamientos que se investigaron en este trabajo en relación al análisis de la cobertura vegetal en campos de cultivo dañados por heladas, son exactamente los mismos que se formulan en la presente memoria en relación al análisis de la cobertura vegetal y de suelos que se propone en la memoria (capítulo 4). Se trata pues de un artículo clave para avalar el método y los planteamientos en el trabajo que se presenta.
-

Artículos publicados en memorias de congresos internacionales con ISBN:

Autores: Macedo Cruz, Antonia, Santos Peñas, Matilde, Pajares Martinsanz, Gonzalo, Villegas R. Isidro.

Título: “Intelligent classification of agricultural land covers in the low catchments of “La Sabana” river, Guerrero”.

Fechas, lugar: 27-03/12/2011, México.

Publicación: Conference Proceedings, IEEE SECCIÓN MÉXICO, Vigésima-segunda Reunión Internacional de Otoño, ROC&C’2011.

Volumen, (Núm.) ISBN: COMPUTACION (CP-76)140. ISBN: 978-607-95630-1-1.

3. En este trabajo se aplica la estrategia de clasificación no supervisada para el reconocimiento de cultivos agrícolas de naturaleza arbórea perenne, tales como los cocoteros y huertos de cítricos. Mediante el análisis espectral de fotografías aéreas a color y la densidad de planta se determinaron las áreas de cultivo con necesidades de mejoramiento. Las imágenes usadas para la clasificación fueron obtenidas de Google Earth y procesadas en modelo de color RGB. La metodología de procesamiento fue muy similar a la anterior, sólo que de acuerdo a las características de las imágenes, se aplicó umbralización simple como proceso de segmentación y reagrupamiento de clases para clasificar los objetos que presentaban algún tipo de solapamiento. Los resultados se validaron temáticamente mediante la denominada matriz de error *fuzzy*.

Autores: Macedo Cruz, A., Pajares Martinsanz, G., Santos Peñas, M.

Título: Unsupervised Classification of Images in RGB Color Model and Cluster Validation Techniques

Fechas, lugar: 12-15/07/2010, Las Vegas, Nevada, USA, July 12-15.

Publicación: WORLDCOMP2010 (The 2010 World Congress in Computer Science Computer Engineering and Applied Computing). Proc. 2010 Int. Conf. on Image Processing, Computer Vision, and Pattern Recognition (ICCV’10), Hamid R. Arabnia,

Leonidas Deligiannidis, Gerald Schaefer, Ashu M. G. Solo (Eds.), pp. 526-532, CSREA Press, ISBN 1-60132-154-6.

Índices de Calidad: Paper Acceptance Rate: 28%

4. En este trabajo se ha aplicado la estrategia de clasificación no supervisada para imágenes de cobertura terrestre en el modelo de color RGB. Si bien, en este caso, el objetivo es evaluar la bondad del clasificador no supervisado mediante la aplicación de cinco índices de validación interna, a saber: Davies y Bouldin (1979), el de Calinski y Harabasz (1974), el índice de Krzanowski y Lai (1985), el índice de Dunn (1974), el índice de Hartigan (1985). Los planteamientos que se investigaron en este trabajo, respecto a la validación del número óptimo de agrupamientos, son fundamentales en un modelo de clasificación no supervisada donde se desconoce el número de grupos a clasificar, como es el caso de la clasificación de la cobertura terrestre presente en las fotografías aéreas. Por tanto en el capítulo tres se aplican dichos índices de validación interna.

Autores: Macedo Cruz, A., Pajares Martinsanz, Gonzalo, Santos Peñas, Matilde, Villegas Romero, Isidro.

Título: Unsupervised classification of ground cover based on the Otsu's method

Fechas, lugar: 10-12/nov/2010, Manzanillo, Colima, México.

Publicación: Conference Proceedings, ROPEC'2010 INTERNACIONAL, ISBN: 978-607-95476-1-5.

Pág. (año): 114-121 (2010)

5. En este trabajo se clasifica la cobertura terrestre captada en las fotografías aéreas en formato digital a color mediante el sistema de clasificación no supervisado propuesto, y el bien conocido método de reconocimiento de patrones denominado *Mean Shift* (Comaniciu y Meer, 2002). A diferencia del anterior, para la validación cuantitativa se construye una matriz de error *Fuzzy* (Congalton, y Green, 2009). Parte de los planteamientos de este artículo, relativos a los índices de validación interna, se describen en el capítulo tercero de la presente tesis; mientras que la validación mediante la matriz de error

Fuzzy, se aplica tanto en el capítulo tercero como en el cuarto, siendo su objetivo evaluar la precisión temática. Aunque el planteamiento propuesto en esta publicación ha sido desarrollado con mayor detalle en el capítulo uno.

Autores: A. Macedo Cruz, G. Pajares, M. Santos, E. Mejía Sáenz
Título: Reconocimiento de patrones del paisaje por umbralización en imágenes tele-detectadas
Fechas, lugar: Cuernavaca, Morelos, México, 7-9/10/2009.
Publicación: IEEE Sec. Morelos, ISBN: 978-607-95255-1-4
(año): 2009

6. Esta ponencia, presentada en el congreso del IEEE indicado en la propia referencia, muestra una estrategia de reconocimiento de patrones por umbralización aplicada a paisajes en imágenes tele-detectadas. En este caso, se comparan cualitativamente los resultados obtenidos mediante tres modelos de color (RGB, $L^*a^*b^*$ y $L^*u^*v^*$). Los modelos CIELab y CIELuv se han aplicado para hacer frente a los problemas de iluminación, por lo que los planteamientos investigados en el presente trabajo, han sido retomados para la clasificación de imágenes de cultivos de avena afectados por heladas, y descrito en el capítulo tercero de la presente investigación.

Capítulos en libros

Autores: Macedo Cruz, A. Santos Peñas, M., Pajares Martinsanz, G. Villegas Romero, I.
Título del libro: Deforestation Around the World
Título del capítulo: Unsupervised classification of aerial images based on the Otsu's method
Fechas de publicación: Febrero 2012.
Editorial: InTech OpenAcces (<http://www.intechweb.org/>)
Editor: Dr. Paulo Moutinho
ISBN: 979-953-307-313-1

7. En este trabajo se comparan dos estrategias de clasificación aplicadas a imágenes aéreas a color para el análisis de diferentes tipos de cobertura terrestre presentes en las imágenes. Las estrategias de clasificación son: la conocida como agrupamiento *fuzzy c-means* y la de clasificación no supervisada, que mediante dos procesos (segmentación y clasificación) logra clasificar los diferentes objetos presentes en la cobertura terrestre. Además se desarrolla un proceso de validación del número de agrupamientos mediante índices internos. Parte del tema desarrollado en esta publicación se aborda el capítulo tres.

Presentación de cartel

Autores:	A. Macedo Cruz, G. Pajares, M. Santos.
Título:	La fotografía aérea a color en la clasificación automática.
Fechas, lugar:	León Guanajuato, México 13-15/05/2009.
Publicación:	Participación de la Mujer en la Ciencia, FCMT15:130

8. En este trabajo se presenta la estrategia de segmentación por umbralización múltiple, mediante la adaptación del método de Otsu (1979). Dicho planteamiento ha sido superado y publicado en la contribución referenciada en el punto dos.

1.5. ORGANIZACIÓN DE LA TESIS

La memoria de la tesis se estructura de la siguiente manera:

- Este primer capítulo de introducción se describen la motivación y objetivos que han originado el trabajo de investigación, así como las principales aportaciones del mismo que han sido presentadas en foros científicos internacionales de divulgación.
- En el **capítulo 2** se presenta el estado del arte, donde se revisan diferentes estrategias de segmentación y clasificación existentes en la literatura, que por su relevancia proporcionan la base de la investigación.

- El **capítulo 3** describe la aplicación concreta y la metodología desarrollada para lograr un sistema automático de clasificación de la cobertura terrestre. Partiendo del pre-procesamiento de las imágenes originales, continuando con la clasificación, donde se detallan las diferentes etapas del proceso, para llegar al algoritmo general de clasificación. Una vez concluida la clasificación se agrega un apartado significativo referido al proceso de evaluación. El capítulo concluye con un apartado en el cual se detallan los resultados más relevantes obtenidos, tanto cualitativos como cuantitativos, en base a los diferentes modelos de validación aplicados. Finalmente, se presenta un apartado de análisis y discusión de resultados.
- En el **capítulo 4** se describe la aplicación concreta y la metodología desarrollada para lograr la clasificación y cuantificación de los daños por heladas causados al cultivo de avena. Partiendo del muestreo fotográfico, y continuando con la clasificación, donde se detallan las diferentes etapas del proceso, para llegar al algoritmo general de clasificación. Posteriormente se desarrolla la etapa de evaluación y por último se concluye el capítulo con un apartado en el cual se detallan los resultados más relevantes obtenidos, tanto cuantitativos como cualitativos, cerrando el capítulo con un apartado de análisis y discusiones de resultados.
- En el **capítulo 5** se detallan las conclusiones extraídas y se definen las líneas futuras de investigación.

La memoria se completa con las referencias bibliográficas empleadas a lo largo de todo el trabajo, que sustentan la investigación. Contiene además tres Anexos, los dos primeros con datos derivados de la investigación y el tercero relativo a la descripción de los modelos de color, considerado de interés por su relevancia en el proceso de clasificación, dado el origen de las imágenes.

Capítulo 2

Estado del arte



CAPÍTULO 2 ESTADO DEL ARTE

2.1. VISIÓN GENERAL

La finalidad de este apartado es explicar minuciosamente aquellas técnicas que reviertan mayor interés, por los algoritmos utilizados, en la clasificación de imágenes en color. No se abarcará el amplio espectro que contempla este abundante campo de investigación con el fin de centrarse únicamente en aquellos aspectos más destacables con respecto a los objetivos planteados en la presente investigación.

Para tener un concepto global del estado actual del arte, se anima al lector a profundizar en este campo mediante la publicación de Lu y Weng (2007), donde se realiza un estudio bastante exhaustivo de la taxonomía de los métodos de clasificación de imágenes, aplicados a la clasificación de la cobertura terrestre. A pesar del año de publicación, sus planteamientos siguen estando vigentes hoy día. En Cheng y col. (2001) y Lucchese y Mitra (2001) se desarrolla una amplia revisión bibliográfica sobre los métodos de segmentación y clasificación aplicados a imágenes en color. Ribeiro y col. (2011) analizan específicamente la cobertura del suelo con control de residuos en imágenes agrícolas desde el punto de vista de la Agricultura de Conservación, el análisis se basa en el empleo de técnicas basadas en algoritmos genéticos.

Los aspectos tratados en este capítulo se estructuran de la siguiente manera:

1. En primer lugar se realiza un recorrido por la literatura para identificar los fundamentos teóricos del reconocimiento de patrones, en el cual se abordan de manera genérica los métodos de clasificación, así como la estructura general de los sistemas de clasificación, sección 2.2.
2. A continuación se expone un apartado sobre la representación y procesamiento de la imagen digital, el cual trata sobre las etapas principales de un sistema de clasificación de imágenes, desde la adquisición de datos, el mejoramiento del contraste, continuando con una revisión del estado del arte relativo a segmentación y clasificación de imágenes a color, sección 2.3.

3. En tercer lugar se realiza una revisión de la teoría de los métodos de segmentación por umbralización, desarrollando aquéllos que sirven de base para los capítulos 3 y 4, sección 2.4.
4. En la sección 2.5 se trata sobre métodos de agrupamiento con la finalidad de explicar el algoritmo de agrupamiento *fuzzy c-means*, puesto que éste se utiliza para la comparación entre el algoritmo propuesto y la evaluación del número óptimo de agrupamientos, resultante del algoritmo de clasificación no supervisado propuesto.
5. En la sección 2.6 se realiza una revisión bibliográfica sobre los índices de validación interna, que se aplicaron para la evaluación del clasificador no supervisado, concluyendo con la estrategia de validación temática denominada matriz de error.
6. Por último, se presenta un resumen sobre los aspectos más relevantes en relación al análisis bibliográfico realizado.

2.2. ESTADO DEL ARTE DEL RECONOCIMIENTO DE PATRONES.

De acuerdo con la definición dada por Pajares y Cruz (2007a, b), así como por Theodoridis y Koutroumbas (2009) y Bow (2002), el reconocimiento de patrones es la disciplina científica cuyo objetivo se orienta hacia la clasificación de objetos en un cierto número de categorías o clases.

Cuando hablamos de técnicas de reconocimiento nos estamos refiriendo a reconocimiento de patrones, objetos o formas que podemos percibir. Un patrón no necesariamente se limita a ser un objeto o forma visible, sino que puede ser también un conjunto de datos. Aunque en el presente trabajo de investigación se trata de reconocer los diferentes objetos o patrones visibles en la cobertura terrestre, cabría igualmente aplicar las técnicas desarrolladas a los no visibles, como por ejemplo, los datos correspondientes a la radiación electromagnética captada por algunos sensores remotos.

De acuerdo con Bow (2002), un *patrón* puede ser definido como una descripción estructural o cuantitativa de un objeto o alguna otra entidad de interés. Dependiendo de la aplicación, estos objetos pueden ser imágenes o señales en forma de onda, o cualquier tipo de medida que necesite ser clasificada.

Un sistema automático de Reconocimiento de Patrones puede ser modelado de la misma manera que los procesos perceptuales del ser humano. Por ejemplo, en el reconocimiento de imágenes de cobertura terrestre, en primer lugar, el intérprete identifica el patrón visual asociado a cada cubierta, de acuerdo a una serie de criterios: tono, textura, forma, contexto, disposición u otros tal y como aparecen en la fotografía o imagen digital a analizar. Posteriormente, delimita sobre el resto de las fotografías las zonas que se corresponden con ese modelo previamente definido. En otras palabras, asigna a cada categoría unas determinadas superficies, en función de la semejanza de éstas con el patrón-tipo identificado previamente. Por último, verifica sobre el terreno la interpretación realizada.

Este es también el esquema propio del reconocimiento digital. Basta traducir los criterios analógicos por digitales. De la misma forma, el reconocimiento digital, también referido en la literatura como clasificación digital, se inicia caracterizando los patrones que definen en la imagen las distintas categorías objetivo. Se trata de un reconocimiento basado en los valores numéricos de los píxeles, por lo que esta caracterización también debe ser numérica. Esto es, se trata de obtener el Nivel Digital (*ND*), o mejor aún, el rango de *ND* que identifica a cada categoría, para todas las bandas espectrales contenidas en la imagen y que intervienen en la clasificación.

En la actualidad existe un conjunto de métodos orientados a la clasificación de la cobertura terrestre, en Wilkinson (2005) así como en Lu y Weng (2007) se presentan los principales métodos de clasificación desarrollados entre 1990 y 2007, cuya validez sigue actualmente vigente. Aun así, la clasificación sigue siendo en la actualidad tema abierto de investigación, ya que los resultados de la misma constituyen la base para muchas aplicaciones ambientales y socioeconómicas, entre otros ámbitos. Gracias a los avances de la informática y la tecnología, se continúa en la búsqueda de sistemas fiables, robustos y automáticos con capacidad de adaptación a la gran variedad de situaciones posibles. Además, la clasificación de la cobertura terrestre es un tema complejo, ya que se deben considerar objetos con propiedades heterogéneas, a las cuales se les añade la complejidad de ser capturadas bajo condiciones de iluminación natural (Chen y col., 2003 y 2004; Tseng y col., 2008).

En Theodoridis y Kouroumbas (2009) así como en Bow (2002) y Jain (1999), se encuentran muy desarrollados diversos aspectos básicos del reconocimiento de patrones tanto de naturaleza supervisado como no supervisado. En Gonzalez y Woods (2010) así como en Pajares y Cruz (2007*a, b*), se desarrollan con gran detalle las diferentes técnicas de clasificación de imágenes en color, así como algunas aplicaciones.

De manera que los métodos de clasificación de acuerdo con el proceso de construcción se han dividido en dos grupos: supervisados y no supervisados (Gonzalez y Woods, 1996; Duda y col, 2001; Escalera, 2001; Bow, 2002; Chuvieco, 2008).

a) *Supervisado*: los algoritmos de clasificación supervisada operan usualmente sobre la información suministrada por un conjunto de muestras, patrones, ejemplos o prototipos que se consideran representantes de clases relevantes, además, se asume que poseen una etiqueta de clase correcta, identificada en la literatura como conjunto de entrenamiento.

Así pues, Meyer y col. (2004) bajo el esquema de clasificación supervisada, utilizan el agrupamiento *Fuzzy* para segmentar regiones de interés basado en el verdor y el enrojecimiento de las imágenes. El algoritmo EASA (*Environmentally Adaptive Segmentation Algorithm*) por sus siglas en inglés, propuesto en Tian (1998), se basa en su capacidad de adaptación para la detección de las plantas verdes a través de un proceso de aprendizaje supervisado. Dicho método fue probado en Ruiz-Ruiz y col. (2009) bajo el espacio de color HSI (*Hue-Saturation-Intensity*) para hacer frente a la variabilidad de la iluminación. En la línea de los métodos anteriormente mencionados existe un conjunto de algoritmos de clasificación supervisada aplicados con base a las necesidades y condiciones existentes, como los desarrollados por Bhaskaran y col. (2010), Saeed (2006), Inampudi y col. (2002), entre otros. Civco (1993) y Yang y col. (1999) revelaron que algunos de los métodos supervisados, como por ejemplo las redes neuronales artificiales son ineficientes cuando se trata de un gran número de bandas espectrales. Shawe-Taylor y Cristianini (2004), así como Camps-Valls y col. (2007), aseveran que los clasificadores basados en núcleos “kernels” son capaces de manejar eficientemente entradas de datos de alta dimensión, y hacer frente a las muestras de ruido de una manera robusta. Sin embargo, la dificultad principal con los métodos supervisados es que el proceso de aprendizaje depende en gran medida de la calidad del conjunto de datos de formación, que sólo es útil para imágenes simultáneas, o para

imágenes conteniendo el mismo número de clases y tomadas en las mismas o similares condiciones. Además, el conjunto de muestras con cierta frecuencia no está disponible, o éstas se obtienen en un número muy reducido, dado el alto costo que representa el etiquetado de las citadas muestras.

b) No supervisado: los algoritmos no supervisados se caracterizan por carecer del conocimiento acerca de las etiquetas de los patrones. Para clasificar es necesario aplicar estrategias, aprovechando las características espectrales de la imagen, que no requieran del etiquetado previo de los patrones. Bow (2002) subdivide los algoritmos de clasificación no supervisada en dos grupos: aquellos donde se conoce el número de clases y aquellos otros en los que dicho número es desconocido. El principal objetivo es, pues, descubrir las principales similitudes y diferencias de los patrones para organizarlos en grupos sin necesidad de un guiado específico.

De acuerdo con Duda y *col.* (2001), existen cinco razones principales por las cuales el uso de técnicas de aprendizaje no supervisado resulta de interés para ciertas aplicaciones:

1. La recolección de datos y su posterior etiquetado en un conjunto de datos muy extenso puede suponer un coste muy elevado.
2. También puede ser de interés actuar en sentido contrario: aprender con una gran cantidad de datos sin etiquetar, y sólo usar supervisión para identificar los distintos grupos encontrados.
3. En muchas aplicaciones las características de los patrones cambian lentamente con el tiempo. Si estos cambios pueden rastrearse en un proceso de ejecución sin supervisar, podremos obtener un resultado más idóneo.
4. Podemos usar métodos no supervisados para encontrar características que serán útiles para la categorización.
5. En el comienzo de la investigación puede ser útil tener una visión general de la naturaleza y la estructura de los datos.

La metodología de un sistema de clasificación no supervisada se puede dividir en tres etapas fundamentales (Duda, 1973, 2001).

1. **Adquisición de datos**, en la que se obtiene una representación del objeto como resultado de un conjunto de mediciones.
2. **Extracción de características**, donde se realiza un proceso interpretativo cuyo resultado se considera como una nueva representación del objeto de la que se extrae información relevante sobre el mismo.
3. **Toma de decisiones**, que corresponde a la clasificación propiamente dicha o proceso de identificación. Theodoridis y Koutroumbas (2009) dividen esta última etapa en tres fases:
 - *Medida de similitud o distancia*. Para saber cuánto se asemejan o cuánto difieren dos vectores de características o dos objetos, es necesario definir una métrica de similitud o distancia. Existen distintas medidas de similitud, y escoger la adecuada es una tarea a la que se debe prestar atención ya que esta medida afecta directamente a la formación de los grupos o clases resultantes.
 - *Criterio para el agrupamiento*. Dependiendo del tipo de grupos que se quiera encontrar y del conocimiento o no de los parámetros de entrada, se deberán tener en cuenta unos criterios u otros para agrupar los datos.
 - *Selección del algoritmo*. Una vez definida la medida de similitud y los criterios de agrupamiento podemos elegir un algoritmo que se adapte a nuestros requisitos. Existen muchos tipos de algoritmos que ofrecen distintas soluciones a diversos problemas, por ello las fases anteriores son muy importantes.

Además de las tres etapas antes mencionadas, Theodoridis y Koutroumbas (2009) añaden dos etapas más:

4. **Validación de los resultados**. Una vez obtenidos los resultados es necesario verificar que son correctos y cuantificar los fallos en la clasificación.
5. **Interpretación de los resultados**. El objetivo final de las técnicas de agrupamiento es obtener algún significado o estructura que nos ofrezca algún tipo de información que antes no poseíamos, o nos ayude a obtener un modelo para otros casos. En esta fase se obtiene esa información o modelo.

En cuanto a la *adquisición de datos*, la función del *sensor* consiste en proporcionar una representación concreta de los elementos a ser clasificados, con el objetivo de entender completamente las propiedades que distinguen al objeto.

La *extracción de características* se encarga de recuperar información discriminatoria con el objetivo de eliminar información redundante o irrelevante, reducir la dimensionalidad del problema y maximizar los rasgos discriminatorios. A esta etapa la identificaremos en adelante como la etapa de *segmentación*.

La etapa de *clasificación* tiene como objetivo asignar los objetos de clase desconocida a la categoría apropiada, es decir, etapa en la que se asigna al objeto una etiqueta de clase con base el *algoritmo seleccionado*, en el cual se considera la medida de similitud y los criterios de agrupamiento que se adapten a los requisitos del problema a resolver.

La etapa de *validación de resultados* tiene como objetivo calificar la calidad de los resultados obtenidos; el proceso de agrupamiento va a proporcionar siempre una salida, si bien ésta puede ser correcta o no, por lo que es importante verificar que el resultado obtenido es el esperado.

Los algoritmos de clasificación no supervisada han sido aplicados a una gran variedad de disciplinas científicas, entre otras, algunas tales como segmentación de imágenes y visión por computador (Jain, 1966 y 2000), minería de datos para extraer conocimiento, recuperación de información (Salton, 1980; Zhai, 2004; Cutting, 1992; Dhillon, 2002), análisis de datos heterogéneos (Cadez, 2001), aplicaciones Web (Heer, 2001, Foss, 2001), biología computacional para el análisis de ADN (Xu, 2002), aplicaciones de bases de datos espaciales (Xu, 1998; Ester, 2000) y muchas otras aplicaciones como las descritas posteriormente en el presente trabajo y específicamente orientadas hacia la Agricultura de Precisión y Conservación o con fines forestales, algunas de ellas con ciertos parámetros de entrada y otras totalmente no supervisadas.

En cuanto a teledetección espacial, existen una serie de factores que han favorecido la aplicación de técnicas de clasificación, entre los que se encuentra como prioritario el aumento de la resolución óptica de las imágenes de satélites artificiales, que en la actualidad llegan hasta 20 o 30 cm/píxel, siendo incluso menores en las imágenes procedentes de sensores aerotransportados. Así como la incorporación de sensores de alta resolución en aplicaciones militares y vehículos aéreos no tripulados,

proporcionando la resolución espacial necesaria para obtener información acerca de los objetos, áreas o fenómenos en la superficie terrestre a diferentes escalas. Estos sensores miden la intensidad de la energía emitida o reflejada por los objetos según las diferentes longitudes de onda del espectro electromagnético. De acuerdo con Cortijo y col. (1997) así como con Lu y Weng (2007), una de las aplicaciones más importantes a medio y largo plazo es la generación de mapas temáticos, donde se etiqueta cada píxel de la imagen mediante una regla de clasificación que especifica el tipo de objeto existente en la zona de referencia. Otra aplicación muy importante y cada vez más demandada se encuentra en el área de la agricultura, que constituye uno de los temas de referencia en el presente trabajo.

Sin embargo, los sensores remotos son caros y las imágenes no siempre son fáciles de obtener. Por lo tanto, es necesario considerar también el análisis espacial del campo agrícola mediante el uso de cámaras digitales convencionales, que proporcionan la información necesaria y al mismo tiempo son fáciles de usar y más accesibles. En Burgos-Artizzu y col. (2010) y en Gottschalk y col. (2010), las imágenes se toman mediante una cámara digital colocada en un tractor robotizado con un ángulo cenital, la información proporcionada por las imágenes se ha aplicado a estudios específicos en el ámbito de la Agricultura de Precisión para la identificación y clasificación de texturas, tanto pertenecientes a los cultivos de cereal y maíz analizados como a las malas hierbas existentes en los mismos.

La importancia que ha adquirido el manejo de este tipo de información es tal que existen numerosas aplicaciones que emplean las nuevas tecnologías de tratamiento de imágenes, inteligencia artificial y técnicas de reconocimiento, entre las que podemos citar:

- *Gestión territorial*: para medición y cálculo de parcelas agrícolas y uso del suelo (Ceballos-Silva y López-Blanco, 2003; Tseng, 2008; Jochem y col., 2010).
- *Control de cultivos en agricultura de precisión*: en aplicaciones para determinar el estado de los cultivos, control de riegos agrícolas, aplicación de herbicidas y pulverización de forma selectiva (Guijarro y col., 2011; Burgos-Artizzu, 2007, 2009).

- *Agricultura de conservación*: con fines de análisis de residuos útiles y persistentes en el suelo tras la recolección (Ribeiro y col., 2011).
- *Vigilancia y evaluación de catástrofes*: tales como incendios, inundaciones, avalanchas de nieve o deforestación (Chuvieco y col., 2007; Plaza y Du, 2009; Guijarro y col., 2009a, b; Pajares y col., 2010).
- *Localización de infraestructuras civiles*: tales como carreteras, caminos forestales, cañadas reales, etc. (Chi, 2008; Bhaskaran y col., 2010; Baik y col., 2003; Res, 2008).

Debido al amplio abanico de aplicaciones derivadas del procesamiento de datos teledetectados, diversas empresas y organismos desarrollan o utilizan aplicaciones para abordar la problemática de la clasificación en el ámbito de la agricultura, de las cuales destaca Dimap (2012), empresa española dedicada al procesamiento de imágenes; implantada, en España, Estados Unidos, Chile, Perú, Ecuador, Brasil y México. Proespacio (2012) es una asociación española de empresas del sector espacial donde algunas de ellas se centran en el objeto de esta tesis.

En México podemos mencionar algunas compañías y organismos como SIGSA (2012), empresa Mexicana líder en aplicaciones geográficas; INEGI (2012), Instituto Nacional de Estadística y Geografía; CENAPRED (2011), Centro Nacional de Prevención de Desastres; y el Instituto de Geografía de la Universidad Nacional Autónoma de México (UNAM, 2012).

Para desarrollar las aplicaciones de tratamiento de imágenes, la mayoría de las empresas e instituciones anteriormente citadas utilizan principalmente herramientas comerciales tales como: ERDAS Imagine (2012), ENVI (2012), PCI_Geomatic (2012) u otras tecnologías comerciales como Image Processing Toolbox de MATLAB (2012), en este último caso con fines de análisis previo al desarrollo de las aplicaciones.

La funcionalidad de las aplicaciones mencionadas, aunque en continuo avance, resulta en muchos casos insuficiente para abordar las propuestas de algunos de los proyectos demandados. En general, poseen módulos de clasificación supervisada, estos módulos suelen estar basados en clasificadores clásicos, tales como: clasificadores estadísticos bayesianos, algoritmos competitivos, agrupamiento borroso o *fuzzy* o redes neuronales

entre otros, clasificadores que se presentan en (Pajares y Cruz, 2007a, b). Además, en muchos casos los métodos tradicionales de clasificación no ofrecen los resultados esperados, e incluso aún siendo suficientes, siempre cabe la posibilidad de introducir mejoras o modificaciones precisas según las necesidades del momento.

Sobre esta base de razonamiento, la clasificación de la cobertura terrestre se presenta claramente como una de las aplicaciones susceptibles de mejora, por lo cual surge la necesidad de investigar nuevas alternativas para el procesamiento de grandes cantidades de información en el ámbito del análisis de dicha cobertura mediante imágenes a color, procedentes tanto de sensores remotos, aerotransportados o convencionales en tierra, como es el caso de las imágenes utilizadas en los estudios que se presentan en este trabajo.

2.3. REPRESENTACIÓN Y PROCESAMIENTO DE IMÁGENES DIGITALES

2.3.1. ADQUISICIÓN DE DATOS

La imagen digital se obtiene a partir de la energía electromagnética que reflejan y emiten los objetos del mundo real, y que es recibida por dispositivos de captura, como cámaras digitales fotográficas o de vídeo, sensores electrónicos, escáner u otros, distribuyéndola convenientemente sobre el dispositivo sensible a dicha energía. El resultado es la matriz rectangular constituida por N filas y M columnas ($N \times M$), cuyos elementos son los valores de los niveles digitales que se atribuyen a cada unidad de información o píxel.

Desde un punto de vista físico, una imagen puede considerarse como un objeto plano cuya intensidad luminosa y color puede variar de un punto a otro.

Las imágenes utilizadas en la presente investigación se pueden clasificar de acuerdo al modo en que se ha efectuado la captura de información sobre el campo y los criterios de clasificación de Thorp y Tian (2004). Un primer conjunto lo constituyen las imágenes teledetectadas, que son captadas mediante dispositivos embarcados en satélites artificiales o aeronaves (fotografías aéreas). Un segundo conjunto corresponde a las imágenes capturadas desde tierra, cuya captura se lleva a cabo mediante sensores de imagen convencionales, es decir cámaras fotográficas.

El término *teledetección* se refiere a la detección e identificación de objetos mediante el uso de sensores (cámaras) o de otros dispositivos de detección que se encuentran a distancias relativamente grandes respecto de los elementos objeto de estudio, si bien la información obtenida procede, como siempre, de la energía electromagnética reflejada o emitida por los objetos que se encuentran en la superficie terrestre (Ruiz, 1998; Mather, 2004).

Ya sea que se trate de imágenes teledetectadas o convencionales, cuando la luz incide en un material, una parte de la misma, dependiendo de la longitud de onda, es absorbida por el propio material y otra parte es reflejada. Se podría definir *reflectancia* como la proporción entre la intensidad de luz reflejada por un objeto y la intensidad de luz incidente en él (Meer, 2002). El término *radiancia* se refiere a la densidad del flujo radiante transmitido por unidad de área en la superficie terrestre. Si definimos *emisividad* como la cantidad de radiación por superficie en todas las direcciones del espacio, entonces la *reflectancia* será la relación existente entre la radiación incidente y la *emisividad* de una superficie. Con todo ello podemos deducir que el sensor deberá ser capaz de medir la reflectancia de la superficie, si bien ésta será distinta en función del tipo de cobertura (suelo desnudo, agua, vegetación). Además, los sensores pueden tener distinto grado de actividad en el proceso. Así podremos distinguir entre sensores pasivos, que se limitan exclusivamente a recoger la energía reflejada o emitida por la superficie terrestre, y los activos, que son capaces de emitir energía para posteriormente recibirla, una vez haya sido ésta reflejada por la superficie terrestre. Este es el caso de los sensores del tipo de Radar de Apertura Sintética (SAR, *Synthetic Aperture Radar*), Sanchez-Llado y col. (2011). En la presente investigación sólo se usan imágenes procedentes de sensores pasivos.

Por lo tanto, la imagen digital puede ser considerada como energía en forma de unión entre los campos de fuerza eléctrica y magnética, formada por un amplio rango de longitudes de onda, que componen a su vez el espectro electromagnético. Como sabemos, el espectro electromagnético aunque continuo suele dividirse en franjas entre las cuales se encuentra la región a la que nuestra vista es sensible, que abarca una muy pequeña porción de la totalidad del espectro, denominado espectro visible.

En el campo de la teledetección, cada sensor es diseñado para operar en una o varias bandas del espectro en función de los objetivos perseguidos. De esta forma, se generarán imágenes mono, multi o hiper espectrales en función del número de bandas capaces de ser discriminadas y medidas. La figura 2.1 muestra un esquema de la distribución de las bandas en el espectro electromagnético. Este concepto está directamente ligado con la denominada *resolución espectral*, y se define como “la habilidad para registrar y discriminar información en detalle, lo que depende del efecto combinado de sus distintos componentes” (Sobrino y col., 2000). La resolución del sensor implicará mayor o menor calidad en la imagen que genere.

Las imágenes en color son realmente tri-espectrales, de esta forma un píxel puede considerarse como un vector tridimensional, desde el punto de vista de su representación, cuyas componentes son las intensidades en las tres bandas o canales espectrales: *Rojo*, *Verde* y *Azul*, ya que la información de la imagen a color se desglosa en tres imágenes correspondientes a cada una de las bandas del espectro visible (Pajares y Cruz, 2007a, b).

Resolución espacial: es el tamaño, medido en unidades de terreno, de la mínima unidad de la imagen o píxel. Tiene relación directa con la sección angular observada, que a su vez se mide en radianes. Este tipo de resolución suele ser generalizada como la resolución global de la imagen. Así se suele decir que cuanto menor es el tamaño del píxel, mayor es la resolución de la imagen, cuando realmente se debería decir que mayor es la resolución espacial.

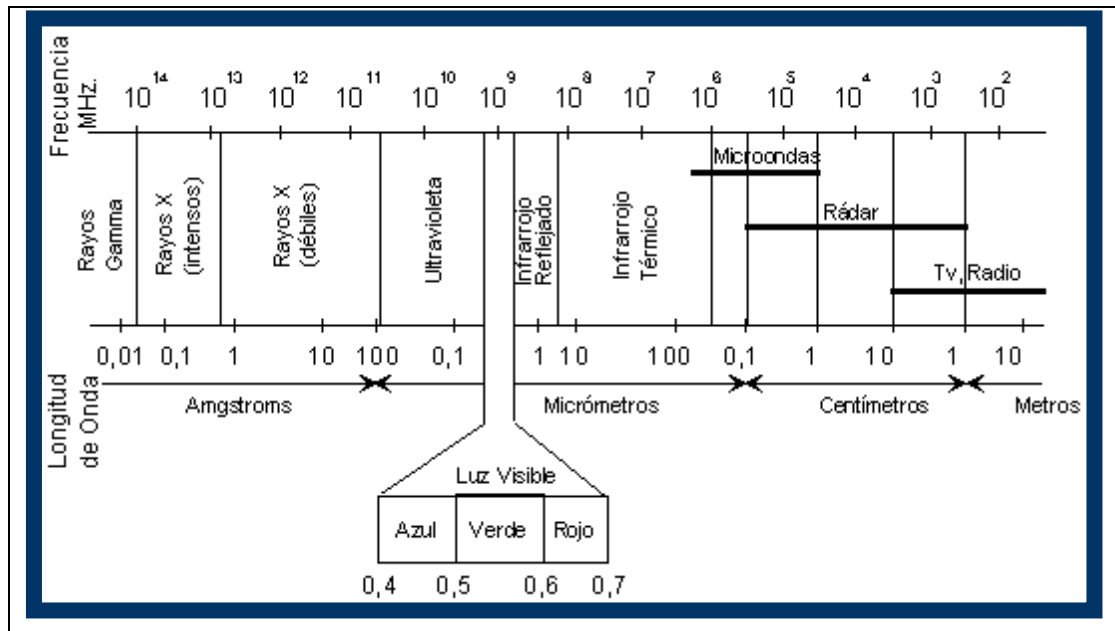


Figura 2.1. Bandas del espectro electromagnético.

Resolución radiométrica (R_{rad}): hace referencia a la sensibilidad de un sensor y expresa su aptitud en una banda para diferenciar señales electromagnéticas de energía diferente (Sobrinó y col., 2000). Esta resolución viene condicionada por el número de Niveles Digitales (ND) de la imagen. Siguiendo la fórmula general, $R_{rad} = 2^n$ bits, para una imagen de 8 bits, el intervalo de ND será 0-255 (puesto que $2^8=256$), y para una imagen de 11 bits el intervalo será de 0-2.047, de manera que a mayor número de ND mayor será el número de intervalos de digitalización de la señal.

Según estos tamaños, una imagen digital con 8 bits y una dimensión de 256×256 píxeles ocupa 64 KB de memoria, con una dimensión de 512×512 píxeles ocupa 256 KB, y si se trata de una imagen de 1.024×1.024 ocupa 1 MB de memoria. En este último caso, al referirnos a una imagen en color de ese tamaño dado, se necesitarán 3 MB de memoria para ser almacenada puesto que son 3 bandas de 1 MB cada una.

Con lo anterior podemos definir una imagen como: “una distribución bidimensional de energía electromagnética que solamente adquiere carta de naturaleza cuando la superficie que la sustenta está iluminada por una superficie radiante” (Pinilla, 1995).

En general, la reflectancia tendrá dependencia espacial, con lo cual en una imagen monocromática vendrá dada por la siguiente expresión: $R = f(x,y)$.

Entonces la imagen digital no es más que una matriz de números, y como toda matriz está compuesta por filas y columnas. Los valores contenidos en cada celda se denominan ND . Dicho valor numérico es función de la resolución radiométrica, $ND = f(x,y)$, siendo en este caso x e y las filas y las columnas respectivamente.

Siguiendo esta definición, una imagen multispectral estará compuesta por tantas matrices como bandas tenga la imagen, donde cada componente de $ND=(ND_1, ND_2, \dots, ND_M)$ representa una banda.

Es decir, una imagen con tres bandas espectrales es conocida como una imagen en color, puesto que el color se representa por la reflectancia de las tres bandas visibles Rojo, Verde, Azul, conocidas universalmente como RGB (derivado de Red, Green, Blue). Por tanto, cada píxel está compuesto por tres valores ND , siendo muy común su representación en 8 bits. Si cada píxel lo denotamos con el símbolo i entonces $i \equiv ND$ representa un determinado píxel en el espacio de color RGB .

$$i = \begin{bmatrix} i_R \\ i_G \\ i_B \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2-1)$$

Con esta representación simplemente estamos diciendo que un píxel i se compone de tres ND : el primero corresponde a la componente R , el segundo a G y el tercero a B . Si queremos considerar también su localización, usamos la notación (x,y) como sigue:

$$i(x, y) = \begin{bmatrix} i_R(x, y) \\ i_G(x, y) \\ i_B(x, y) \end{bmatrix} = \begin{bmatrix} R(x, y) \\ G(x, y) \\ B(x, y) \end{bmatrix} \quad (2-2)$$

De esta forma, si la imagen es de tamaño $N \times M$, x puede variar entre 0 y N , e y entre 0 y M .

De manera que para representar cualquier composición en color, con sus tres componentes espectrales, sólo se debe sustituir la notación RGB por el número de banda usada, o la representación del modelo de color en cuestión. No existe unanimidad a la hora de referirse a los niveles de intensidad en una determinada localización espacial (x,y) , pudiéndose encontrar diversas nomenclaturas, tales como $f(x,y)$, $i(x,y)$, $f(i,j)$, $I(i,j)$,

$g(x,y)$, $g(i,j)$, $X(x,y)$, $X(i,j)$, de manera que i, j o bien x,y se refieren a las coordenadas espaciales, mientras f, i, g, I , o X en cualquier localización (x,y) o (i,j) es proporcional al ND (intensidad) de la imagen en ese punto (Pajares y Cruz, 2007a,b). Sin pérdida de generalidad, en lo que sigue nos referiremos a los valores digitales indistintamente con los términos, ND , intensidad o nivel de gris.

En la figura 2.2, se puede observar una representación de una localización espacial (x,y) de manera que a dicha localización se asignan las tres componentes espectrales de la imagen en color compuesta por sus tres bandas.

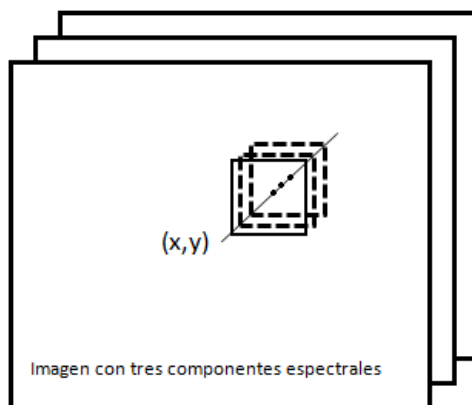


Figura 2.2. Representación espacial de un píxel en una localización x, y correspondiente a una imagen con tres componentes espectrales, donde el píxel se encuentra enmarcado por un cuadrado.

2.3.2. MEJORAMIENTO DE CONTRASTE

En el momento de captura de la imagen, cada nivel digital viene determinado por una serie de elementos que condicionan las características de los datos capturados, como son la energía solar, las particularidades de los elementos de la superficie, los componentes de la atmósfera atravesada por la energía (que provocan absorción y dispersión atmosférica) o las características de los sensores (ganancias y desfases), influyendo en los niveles de intensidad de la imagen. Por lo tanto, los datos obtenidos en ocasiones no se pueden utilizar en su estado original y deberán someterse a un conjunto de tratamientos previos que algunos autores denominan correcciones radiométricas, o mejoramiento de contraste.

El propósito de los métodos de mejoramiento de contraste consiste en aumentar la calidad de una imagen. Por lo general diversos autores como Jafar y Ying (2007),

Coltuc y col. (2006), Subr y col. (2006), Kabir y col. (2009) o Maini y Aggarwal (2010) entre otros, coinciden en la clasificación de los métodos de mejoramiento de contraste (que identificaremos en adelante como MMC), clasificándolos en globales y locales. Los métodos globales se subdividen a su vez en dos grupos: *MMC global por histograma* y *MMC global por expansión de contraste lineal*; y los globales por histograma a la vez se subdividen en *MMC globales por ecualización automática de histograma* y *por igualación de histograma*.

Los métodos de mejoramiento de contraste locales se dividen en dos grupos: *locales por histograma* y *locales por gradiente*.

Métodos de mejoramiento de contraste global. Las técnicas de mejoramiento de contraste global transforman los datos originales usando las estadísticas calculadas para toda la imagen. Por lo general se usan para remediar problemas como condiciones excesivas o pobres de iluminación en el ambiente de origen.

Método de mejoramiento global por ecualización automática de histograma, es un método ampliamente utilizado para el mejoramiento de contraste. En general un histograma con una distribución de los niveles de gris concentrada en una determinada zona presenta un contraste muy bajo. En principio, la ecualización de histograma incrementa el contraste de una imagen mediante la transformación de su histograma en uno uniforme que abarque el rango dinámico completo de niveles de intensidad. Esto se basa en la suposición de que para una máxima transmisión de información, la distribución percibida (histograma) de niveles de intensidad en una imagen o banda espectral debe ser uniforme, es decir, determina automáticamente la función de transformación que permite que la banda resultante de la imagen tenga un histograma uniforme.

A pesar de la simplicidad y la definición implícita de la función de transformación en el método de ecualización de histograma, existen algunos problemas asociados. El primero es que resulta en un cambio significativo en el valor medio del brillo de la imagen. Esto es debido a que la ecualización transforma el histograma original en una distribución uniforme con un valor medio en la mitad del rango de niveles de intensidad, independientemente del valor medio de la imagen original. Segundo, el método de ecualización de histograma puede causar un contraste excesivo o bien una saturación de

los niveles de intensidad, alcanzando valores máximos dentro del rango de representación. Y tercero, la función de transformación de la ecualización de histograma es capaz de mejorar el contraste global de una imagen y podría o no aumentar el contraste local, debido a que está basado en el contenido global de la imagen.

Método de mejoramiento global por igualación de histograma o por especificación de histograma. De acuerdo con Gonzalez y Woods (2008), bajo este esquema de ecualización, el contraste de la imagen original se modifica mediante la especificación del histograma deseado. Es decir, consiste en la adaptación del contraste de una imagen (imagen origen) al de otra (imagen destino) mediante modificación de histograma de la imagen origen.

Método de mejoramiento global por expansión de contraste lineal, es una técnica simple de mejoramiento de contraste, que intenta aumentar el contraste de una imagen incrementando su rango de niveles de gris a un rango deseado de valores, como puede ser el rango completo de valores que el tipo de imagen permita (Gonzalez y Woods, 2008). Este método se diferencia de la ecualización automática de histograma en que sólo puede aplicar una función lineal de escalado a los píxeles de la imagen.

La figura 2.3 muestra esquemáticamente un ejemplo pedagógico sobre el principio usado en las técnicas de mejoramiento del contraste. Se asume un dispositivo de salida capaz de desplegar 256 niveles de intensidad. El histograma muestra los *NDs* en el rango de 58 a 158. Si estos valores de la imagen se desplegaran directamente, sólo una pequeña porción del total disponible se usaría. Los niveles de despliegue de 0 a 57 y de 159 a 255 no se están utilizando.

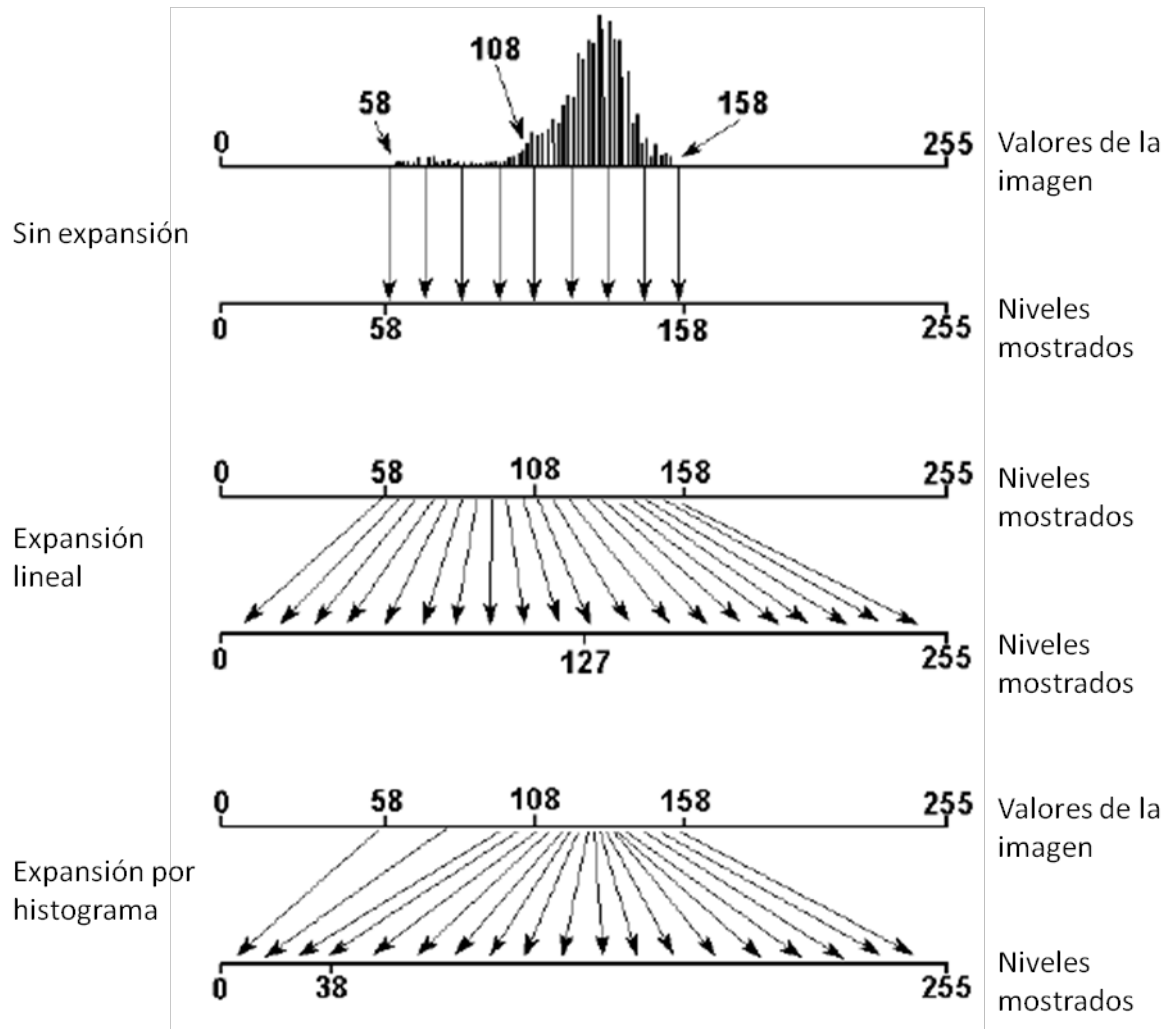


Figura 2.3. Principios de la técnica de mejoramiento del contraste.

Usando el método de *expansión de contraste lineal* el rango de la imagen (58 a 158) se extenderá a la totalidad del rango posible de despliegue, esto es, entre 0 y 255. Puesto que en este método el valor de *ND* bajo el histograma (58) se asigna al extremo negro y un valor alto (158) se asigna al extremo blanco. El resto de los valores de los píxeles se distribuyen linealmente entre estos dos extremos. Una desventaja del “contraste lineal” es que asigna los mismos niveles de visualización tanto a valores de *ND* que ocurren pocas veces como a los que ocurren con mayor frecuencia. Aunque este método aplica el intervalo (m, M) sobre el de $(0, 255)$ en la mayor parte de los casos se produce una imagen con poco contraste. Aunque todas las tonalidades de gris del despliegue son utilizadas, la mayor parte de los píxeles toman valores medios dentro del rango total de

niveles de gris. Esto se debe a la distribución más o menos normal dentro de los valores mínimo y máximo del histograma.

En cambio, la técnica de ecualización por histograma aplica una técnica no-lineal. En este método los valores de los *ND* se distribuyen según su frecuencia de aparición en el histograma. Por ejemplo el rango de valores de la imagen entre 109 y 158 se extiende sobre una gran porción de los niveles de despliegue (39 a 255). Una pequeña porción (0 a 38) se reserva para los valores menos frecuentes de la imagen entre 58 y 108.

Métodos de mejoramiento de contraste local. Los métodos locales consideran sólo las condiciones de pequeñas áreas de la imagen (locales), las cuales pueden variar considerablemente sobre una misma imagen. Los métodos de mejoramiento local suelen encontrarse en la literatura también como mejoramiento espacial, que consiste en modificar el valor de un píxel basándose en los valores de los píxeles de sus vecinos inmediatos. El número de vecinos a considerar depende del tamaño de la vecindad considerada, conocida generalmente como ventana.

Las vecindades más comúnmente utilizadas son matrices de dimensión 3×3 , 5×5 , 7×7 o incluso de mayor orden. La ventana se desplaza a través de la imagen y tiene en cuenta todos los valores de los *ND* que estén dentro de la ventana. Cada valor *ND* del píxel es multiplicado por el coeficiente correspondiente asociado a la ventana. Los valores resultantes se suman y el valor final se divide por la suma de los coeficientes de la ventana. El valor original del píxel central se reemplaza por el nuevo valor así obtenido.

Método de mejoramiento local por histograma. El procedimiento global de ecualización de histograma puede ser fácilmente extendido a mejoramiento local de contraste, llamado ecualización local de histograma. En la ecualización local de histograma se define primero una ventana cuadrada o rectangular, moviéndose el centro de la misma píxel por píxel a través de toda la imagen. Esta extensión de la ecualización de histograma permite a cada píxel adaptarse a su vecindario. Así se puede lograr un mayor contraste para todas las ubicaciones de la imagen. Sin embargo, la ecualización local de histograma produce una modificación no natural de la imagen y bordes en puntos donde la transformación local cambia abruptamente, debido al rápido cambio del histograma local. Esto se debe a que la ecualización local de histograma es sólo la extensión local

de la ecualización automática del histograma, por tanto, hereda sus problemas de saturación.

La especificación de histograma también tiene su contraparte en los enfoques locales de mejoramiento de contraste. ¿Cómo se define el histograma deseado?, en este sentido Jafar y col. (2007), proponen un nuevo método llamado Especificación Local Automática de Histograma (*Automatic Local Histogram Specification - ALHS*), que proporciona automáticamente el óptimo mejoramiento de contraste con una distorsión mínima en la apariencia de la imagen. Básicamente, la especificación local automática del histograma es aplicada localmente como la ecualización local de histograma. Sin embargo, para modificar el píxel en el centro del bloque se especifica el histograma de salida deseado para ese bloque. Así el método de especificación de histograma se utiliza para encontrar el nuevo valor del píxel. La idea central consiste en la determinación del histograma deseado para cada bloque. La especificación local automática del histograma define el histograma de tal manera que éste sea lo más cercano a la distribución uniforme y, al mismo tiempo, su valor medio de intensidad tenga una mínima desviación respecto a su valor original.

MMC locales por aumento de gradiente, en general el aumento de gradiente se encuentra asociado con un aumento de contraste. En este sentido, Subr y col. (2006) proponen un algoritmo que logra el mejoramiento del gradiente mediante la maximización de una función objetivo, obteniendo una imagen resultante con un mayor contraste que se controla con un único parámetro.

De acuerdo con Lu y Weng (2007), cuando se trabaja con información teledetectada durante un único periodo es mejor no realizar correcciones, a menos que sea absolutamente necesario por la mala calidad de la imagen, puesto que dichas correcciones también pueden influir de alguna manera en el proceso de clasificación posterior. Sin embargo, cuando se procesa información de varios años es necesaria igualar la iluminación del conjunto de imágenes, puesto que son tomadas a distintas horas y periodos del año. Con base en lo anterior se decidió no aplicar mejoramiento de contraste a todas las imágenes procesadas, sólo se aplicó a una minoría que consideramos lo necesitaban por las razones expuestas previamente.

Otro aspecto a considerar es que la metodología de mejoramiento de contraste originalmente está diseñada para imágenes en escala de gris, extenderla a imágenes de color no es sencillo, ya que se debe definir una estrecha relación de orden entre píxeles de la imagen en color. La solución que más se ha utilizado al respecto consiste en transferir el procesamiento de imágenes en color a nivel de gris simplemente mediante la representación en un espacio de color donde se pueda separar la intensidad (luminancia) para procesar solamente la componente de luminancia. Para lo cual se selecciona algún espacio de color de entre los que se denominan color de televisión (YUV, YIQ, YC_aC_b) y los de percepción (HSI, CIELAB) (Coltuc y col., 2006), de esta forma se procesa la componente de luminancia y se vuelve posteriormente al modelo de color original (RGB).

2.3.3. SEGMENTACIÓN DE IMÁGENES EN COLOR

Desde hace tiempo se reconoce que el ojo humano puede distinguir miles de tonos de color e intensidades, si bien tan sólo dos docenas de tonos de gris. Razón por la cual a menudo los objetos no se pueden extraer utilizando la escala de grises, pero sí la información de color, ya que el color proporciona información adicional a la intensidad. Por tanto, la información de color es útil y hasta necesaria para el reconocimiento de patrones en tareas relacionadas con la visión por computador. Así, la adquisición y el hardware de procesamiento de imágenes en color han mejorado considerablemente, llegando a ser mucho más accesibles para hacer frente a la complejidad computacional causado por el aumento de espacio y cómputo requeridos por las imágenes de color de alta dimensión. Lo anterior ha hecho que el procesamiento de imágenes en color sea cada vez más práctico, como viene siendo habitual desde hace ya tiempo (Lucchesse y Mitra, 2001).

La segmentación debe verse como un proceso que a partir de una imagen, se obtiene una representación de la misma en la que cada píxel tiene asociada una etiqueta distintiva del objeto al que pertenece. Así, una vez segmentada la imagen dada, se podría formar una lista de objetos, consistentes en las agrupaciones de los píxeles que posean la misma etiqueta.

Duda y col. (1973, 2001) definen la segmentación como “la extracción de un objeto de interés de una imagen arbitraria”. Gonzalez y Woods (2008) la definen como “la subdivisión de una imagen en sus partes constituyentes”.

Para Bow (2002) así como para Theodoridis y Koutroumbas (2009), se trata de la operación de bajo nivel relacionada con la partición de las imágenes, mediante la determinación de regiones disjuntas y homogéneas, es decir, consiste en extraer las características principales que permitan aislar objetos considerados de interés o de importancia de acuerdo al problema que se plantee. Pajares y Cruz (2007 a,b) agregan que durante el proceso de segmentación se extrae de la imagen cierta información subyacente en la misma para su posterior uso.

Puesto que los elementos básicos constituyentes de las imágenes son los píxeles, en este trabajo se han considerado como las características fundamentales de las imágenes, cuya propiedad más relevante es su naturaleza espectral, y utilizando como principal atributo las tres componentes espectrales derivadas del color.

De acuerdo a la complejidad de la escena presente en la imagen a segmentar, ésta puede ser *segmentación completa* o *segmentación parcial*. Si se da una segmentación completa, la segmentación termina cuando los objetos extraídos de la imagen se corresponden unívocamente con las distintas regiones disjuntas a localizar en la misma. En el caso contrario, escenas complejas, el resultado de la segmentación podría ser un conjunto de regiones homogéneas superpuestas, en cuyo caso, la imagen parcialmente segmentada deberá ser sometida a un tratamiento posterior con el fin de conseguir una segmentación completa.

En las últimas dos décadas el tratamiento del color ha recibido una importante atención por su alta calidad como propiedad descriptora (Guijarro y col., 2011; Fonseca y col., 2011; Herrera y col., 2009; Jiji y Ganesan, 2008; Xie, 2008; Rud y col., 2006; Hu y Tao, 2005; García-Alegre, 2000; Comaniciu y Meer, 1997; Chang y col., 1994). En Skarbek y Koschan (1994), se pueden encontrar los avances sobre segmentación de imágenes a color por umbralización de 1971 a 1994; y en Lucchese y Mitra (2001) se encuentran de 1981 al año 1999. En Cheng y col. (2001) o Correa-Tomé (2011), se desarrolla un análisis comparativo con un elevado nivel de profundidad sobre distintas

representaciones del color, tales como los modelos de color RGB, HSI, YIQ, YUV, así como algunas representaciones CIE, incluyendo las características monocromas.

En las técnicas de segmentación y clasificación de imágenes se utilizan con profusión distintos modelos de color como RGB (R, red; G, green y B, blue), HSI (H, hue, S, saturation e I, intensity), YIQ, YUV o $L^*u^*v^*$. En general, destacan como los más eficientes los modelos RGB y HSI (Cheng y col., 2001). El modelo RGB es muy utilizado en técnicas basadas en gráficos por computador y específicamente en aplicaciones donde el computador como dispositivo resulta relevante, ya que posee una gran dependencia de las características del computador donde se visualiza y representa el gráfico o la imagen (Lucchese y Mitra, 2001).

Mendoza y col. (2006) señalan que el modelo HSI está más bien relacionado con el concepto de percepción humana, y bajo este mismo concepto Yang y col., (2008) agrupa las representaciones de color $L^*a^*b^*$ y $L^*u^*v^*$.

En Trias-Sanz (2008) se realiza un estudio sobre diferentes modelos de color para la segmentación de imágenes de satélite correspondientes a entornos rurales, llegando a la conclusión de que en algunos casos resulta más conveniente utilizar el modelo de color RGB junto con ciertos modelos de color transformados.

Cheng y col. (2001) desarrollan un análisis comparativo, con un elevado nivel de profundidad, sobre distintas representaciones del color, incluyendo las características monocromas. En dicho trabajo se concluye con una serie de ventajas e inconvenientes.

Para la clasificación de imágenes capturadas bajo iluminación no controlada, se obtienen mejores resultados aplicando el modelo de color CIE $L^*a^*b^*$, ya que éste es menos dependiente de la iluminación, considerándose aproximadamente uniforme, es decir, la distancia entre dos colores en un espacio de color lineal corresponde a las diferencias percibidas entre ellos. Por lo tanto, proporciona una representación objetiva del color y su uso es esencial para aplicaciones en las que los resultados deben coincidir con los de la percepción humana (Mendoza y col., 2006; Sangwine, 2000; Macedo y col., 2011).

Dentro del análisis del color existen otras aproximaciones relativas a la percepción visual, donde a partir del histograma se identifican pequeñas variaciones del mismo con

el fin de suprimirlas, consiguiendo así un cierto suavizado, y con ello la homogenización de las regiones de texturas (Vázquez y col., 2007).

Los trabajos más ampliamente difundidos sobre segmentación de imágenes a color se basan en enfoques de segmentación a través de niveles de gris con diferentes representaciones de color, es decir se aplica alguna propuesta de segmentación especialmente diseñada para imágenes en escala de gris más algún modelo de color, como se muestra en la figura 2.4. En el Anexo 3 se incluyen los aspectos básicos relacionados con la teoría del color, dada la importancia que ésta posee.

Desde hace ya varias décadas se pueden encontrar trabajos dedicados a la revisión de técnicas de segmentación para imágenes en escala de grises, destacan en esta línea los de Fu y Mui (1981) o Pal y Pal (1993) y más recientemente Devi y Varadarajan (2011).

Como se ha mencionado previamente, una imagen a color se forma por sus tres componentes espectrales, y una segmentación basada en píxeles trata de etiquetar cada píxel como perteneciente a uno de los grupos subyacentes en la imagen bajo consideración. Una alternativa muy común es convertir la imagen de color en una imagen en escala de gris. En la literatura se pueden encontrar un gran número de métodos para llevar a cabo esta transformación (Gonzalez y Woods, 2008; Theodoridis y Koutroumbas, 2009).

Cunha (2003) para realizar un estudio de caracterización de hojas de una planta convierte las imágenes de color a escala de grises, aplica histogramas para ecualizarlas, reduce el ruido, y finalmente binariza la imagen dejando en color negro el objeto de interés, en este caso la hoja, detectando también los bordes de la misma.

Incluso en algunos casos se pasa del modelo de color RGB a algún otro modelo de color que permita separar la componente de intensidad de las componentes de color, de entre ellos destaca por su importancia el modelo HSI (H, *hue*, S, *saturation* y I, *intensity*), donde la banda de intensidad queda disponible para su procesamiento durante la segmentación, tras lo cual se vuelve a convertir de nuevo al modelo de color original RGB, como se expone en Pajares y Cruz (2007a, b). Sin embargo Levachkine y col. (2003) mencionan que bajo esta opción el tono es inestable a baja saturación.

Para hacer frente al tratamiento de las imágenes en color, considerando las tres componentes espectrales, por lo general se toma como referencia una de las formas siguientes:

2. El procesamiento de cada canal individual mediante la aplicación directa de métodos de segmentación basados en niveles de gris (Guijarro y col., 2011; Tellaeché y col., 2008a,b, 2011; Burgos-Artizzu, 2011, 2010, 2007; Sainz-Costa, 2011; Gashnikov y col., 2009; Xie, 2008; Xie y Mirmehdi, 2007; Du y col., 2004; Shih y Cheng, 2005; Levachkine y Torres, 2003; Macedo y col., 2010): los tres canales se suponen independientes entre sí y sólo se tienen en cuenta las interacciones espaciales.
3. La descomposición de la imagen en los canales de luminancia y cromáticos (Shih y Cheng, 2005; Mendoza y col., 2006): transformación del espacio de color para que la segmentación se aplique al canal de luminancia y las características cromáticas a los canales de esta naturaleza, es decir a los cromáticos, cada uno en una manera específica. La selección del espacio de color por lo general depende de la aplicación.
4. La combinación de interacción espacial dentro de cada canal y la interacción entre los canales espectrales (Kurugollu y col., 2001; García-Alegre y col., 2000; Chenaoua y col., 2003; Meyer y Neto, 2008): las técnicas de segmentación en niveles de gris se aplican en cada canal, teniendo en cuenta las interacciones de píxel entre diferentes canales, además, de cada canal individual.

También se han probado técnicas que combinan los tres canales:

5. Aprovechar las tres componentes espectrales para la segmentación de color (Woebbecke y col. 1995, Neto, 2004; Mendoza y col., 2006; Gashnikov y col., 2009; Mezaris y col., 2004; Ribeiro y col., 2005; Polidorio y col., 2003; Tellaeché y col., 2008a,b, 2011; Burgos-Artizzu y col., 2011, 2010, 2007; Sainz-Costa, 2011): la distribución espacial y espectral se maneja simultáneamente. Las principales dificultades que surgen en la representación efectiva, es generalizar y discriminar los datos respecto de las tres componentes espectrales.

En este último caso, los métodos por lo general se enfocan hacia la resolución de un problema específico y con un modelo de color determinado. Como por ejemplo Mendoza y col. 2006, que mediante los modelos de color RGB, $L^*a^*b^*$ y HSV,

implementan una técnica de análisis de imágenes, especialmente diseñada para medir el color del plátano y el pimiento rojo, donde el procesamiento básico para extraer el objeto del fondo se desarrolló mediante estrategias de segmentación de imágenes en escala de gris, tales como filtro paso bajo, binarización por umbralización y operador laplaciana, una vez que se extrae el objeto se usan los modelos de color para medir la coloración del fruto. Mezaris y col. (2004) trabajan con el modelo de color $L^*a^*b^*$, iniciando con un proceso de segmentación para reducir la imagen mediante un algoritmo similar al de las k -medias, introduciendo el número de grupos iniciales e iniciando el agrupamiento por la esquina superior izquierda, posteriormente mediante la implementación de una máscara y el modelo de color $L^*a^*b^*$ se consigue la segmentación. Otros métodos como Neto (2004) o Ribeiro y col. (2005) combinan las tres bandas de color de forma que cada una posee un peso específico en función del objetivo a segmentar, en ambos casos la identificación es sobre plantas en campos de cultivo de cereal o maíz.

Sin embargo, bajo este enfoque se limita a un exclusivo modelo de color, mientras que en imágenes teledetectadas en ocasiones no son precisamente las tres componentes del espectro visible las que se procesan. En Trias-Sanz y col. (2008) se hace un estudio sobre diferentes modelos de color para la segmentación de imágenes de satélite correspondientes a entornos rurales, llegando a la conclusión de que en algunos casos resulta más conveniente utilizar ciertos modelos de color transformados, además del modelo RGB, que también es empleado.



Figura 2.4. Esquema comúnmente usado para la segmentación de imágenes a color.

En cuanto a la descomposición de los canales de luminancia y sus características cromáticas, existen métodos como el propuesto por Shih y Cheng (2005), basados en el crecimiento de regiones por selección automática de semillas. La imagen de entrada se encuentra en el modelo de color RGB y se transforma al espacio de color $YCbCr$, porque bajo dicho modelo las componentes de intensidad y cromáticas pueden ser controladas independientemente. Por tanto, para la segmentación se consideran las tres componentes espectrales, concluyendo con un algoritmo de fusión de regiones, basado en la aplicación de una ventana de dimensión 3×3 , la distancia Euclídea como medida de similitud y un umbral de 0,05, que determina el límite para dicha similitud.

Fonseca y col. (2011), trabajan con imágenes multiespectrales, combinando tres bandas para producir una imagen RGB, posteriormente, las imágenes son transformadas al modelo de color HSI, de forma que es posible aplicar una corrección de contraste sobre la componente I de intensidad, tras lo cual se vuelve al modelo de color RGB para desarrollar el procesamiento de la imagen.

En cuanto a la combinación de las componentes espectrales, un ejemplo se presenta en Kurugollu y col. (2001) quienes desarrollan un trabajo de segmentación por multi-umbralización con imágenes a color. Su técnica consta de dos etapas, multiumbralización y fusión. La multiumbralización se aplica para cada par de componentes espectrales de la imagen RGB (RG, RB, GB), y mediante el histograma seleccionan los picos que aparecen como centros de clase, de manera que el valor de cada pico se usa como semilla para la segmentación de cada par de bandas. La segunda etapa consiste en fusionar los tres pares de componentes en un único par, de forma que la etiqueta final se obtiene al fusionar las clases de los tres pares seleccionando la etiqueta predominante, es decir, por mayoría de coincidencias.

Nimbarte y Mushrif (2010), aplican un método de segmentación por umbralización a imágenes en modelo de color RGB. Para la umbralización se toma cada componente de color de manera independiente para separar en regiones cada componente, lo cual se lleva a cabo mediante dos métodos de umbralización: el de mínima varianza de clase y el método de Otsu. Para mezclar las regiones se calcula un umbral con base a la regla JND “*Just Noticeable Difference*”, de manera que dos regiones vecinas pueden ser mezcladas en una única si ambas son similares, estableciendo la similitud mediante el

cómputo de la distancia Euclídea en función de que ésta sea mayor que un umbral previamente establecido. Desde nuestro punto de vista, esta metodología es adecuada, sólo que los resultados de la clasificación dependen del valor obtenido por JND y por tanto del peso asignado para el cálculo del JND. La separabilidad entre cada grupo dependerá de la experiencia del experto y por tanto del peso asignado, algo que ha de establecerse a través de diferentes pruebas que deberán ejecutarse previamente para encontrar la mejor separación.

Guijarro y col. (2011) desarrollan un sistema de segmentación automática basándose en el cálculo de índices de vegetación obtenidos por la metodología propuesta por Meyer y Camargo (2008), la combinación de dichos índices logra separar la intensidad entre las plantas, el suelo y el cielo. Posteriormente, mediante histogramas de nivel de gris extraen contrastes máximos y mínimos, combinan los índices, y aplican algoritmos de umbralización automática como Otsu o el valor promedio del histograma; además establecen diferencias entre los resultados proporcionados por ambos métodos de umbralización, que son corroboradas y analizadas en Burgos-Artizzu y col. (2011). El objetivo consiste en identificar las partes relativas a la vegetación mediante el concepto de verdor, los rojizos para extraer el suelo, y los índices azules para localizar el cielo. Finalmente, para el proceso de clasificación en la imagen RGB se aplica un algoritmo de agrupamiento difuso, bajo el cual, como se sabe, es necesario proporcionar ciertos parámetros de entrada como es el número de clases a extraer, logrando finalmente una buena segmentación de las imágenes naturales en campos de cultivo de maíz. Sin embargo, desde nuestro punto de vista este algoritmo no permite su generalización para todo tipo de imágenes naturales, particularmente por el hecho de que requiere conocer de antemano las clases a extraer y por tanto el número de éstas.

Ribeiro y col. (2005), basándose en la estrategia de las propiedades del plano de color de pigmentación *cian* en un modelo de color CMYK, construyen una imagen en escala de grises donde las zonas de cubierta vegetal tienen mayor intensidad que el resto. Además incrementan la separación de la capa vegetal del resto mediante una combinación lineal de los planos de color primarios RGB y el uso de tres coeficientes (r , g , b), los cuales se ajustan mediante métodos de optimización para obtener el efecto deseado; en esta misma línea se encuentran los métodos de segmentación utilizados en Tellaeche y col. (2008a, b, 2011), Burgos-Artizzu y col. (2011, 2010, 2007) o Sainz-

Costa (2011). Dichas técnicas se orientan principalmente a la identificación de las partes verdes existentes en las imágenes, algo que difiere sustancialmente de los objetivos planteados en este trabajo.

La alternativa de procesar las tres componentes por separado puede resultar muy conveniente en el caso de los sensores multi-espectrales, principalmente en el ámbito de la teledetección, ya que tanto da que los canales pertenezcan al espectro visible como a otras bandas del espectro con diferentes longitudes de onda e incluso con diferentes sensores que proporcionan imágenes obtenidas también con diferentes longitudes de onda, tal es el caso de los planteamientos formulados por Simone y col. (2002). Sin embargo, cuando se trate de combinar las bandas adquiridas por diferentes sensores multiespectrales es recomendable revisar previamente los planteamientos formulados en Fonseca y col. (2011) en cuanto a los diferentes aspectos a considerar para la selección de bandas, los cuales no se contemplan en los objetivos del presente trabajo de investigación.

Como se mencionó antes, en la segmentación de imágenes complejas, por lo general se desarrolla una segmentación parcial, lo cual implica que después del proceso de segmentación se necesite aplicar alguna metodología de fusión de componentes, tal y como se observa en Kurugollu y col. (2001); Santos y col. (2003) o Nimbarte y Mushrif (2010).

El paso siguiente a la segmentación es el de clasificación propiamente dicha, cuyos aspectos relacionados con el conjunto de imágenes de estudio se abordan a continuación.

2.3.4. CLASIFICACIÓN NO SUPERVISADA

En la actualidad existen diferentes métodos avanzados de clasificación, y por tanto categorizados de diferentes maneras. Dentro de los clasificadores no supervisados existe un conjunto de categorías entre las cuales se pueden enmarcar los métodos de clasificación, por ejemplo, de acuerdo a las características aplicadas se pueden distinguir entre aquellos con características espectrales, los que aplican características espaciales y los que utilizan la combinación de ambas. En cuanto a la información espectral usada se pueden categorizar según la clasificación por píxeles, por objetos o

por campos. En Lu y Weng (2007) se desarrolla una taxonomía de los métodos de clasificación de imágenes así como sus ventajas e inconvenientes.

En la fase de segmentación de imágenes presentada en el apartado anterior, se espera que los segmentos resultantes (conjuntos de píxeles etiquetados) sean homogéneos con respecto a las características representadas en los correspondientes espacios de color. Sin embargo, no hay garantía alguna de que estas regiones muestren también compactación espacial (clases separadas), que es una segunda propiedad deseable en aplicaciones de clasificación junto con la homogeneidad. La compactación espacial o separación significa que las clases o clústeres deben estar convenientemente separados unos de otros.

Para medir cuán espaciadas se encuentran las clases, existen tres métodos comunes: el primero se denomina *vinculación individual*, que mide la distancia entre los miembros más cercanos de los grupos; el segundo, *ligamiento completo*, que mide la distancia entre los miembros más distantes; y el tercero *comparación de los centroides*, que mide la distancia entre los centros de las clases. El más comúnmente usado dentro de la clasificación no supervisada es el tercero, que resulta ser el finalmente elegido para la presente investigación.

Por lo general, la compactación espacial se logra ya sea por subdivisión y fusión, o por crecimiento de regiones de la imagen, mientras la homogeneidad se adopta como criterio para dirigir estos dos procesos; por tanto, es necesario determinar cuándo dos objetos del espacio son “parecidos” y cuándo no. Con este fin se definen las funciones de similitud o de disimilitud, entre las que se encuentran las métricas o distancias (Lucchese y Mitra, 2001; Theodoridis y Koutroumbas 2009; Gonzalez y Woods 2008).

De manera que la etapa de clasificación parte de un conjunto de regiones homogéneas superpuestas, parcialmente segmentadas que deberán ser sometidas a un proceso de compactación espacial con el fin de conseguir una clasificación completa y apropiada.

Bajo esta concepción se han revisado en la literatura algunas aplicaciones, las cuales asumen como premisa el hecho de que para lograr una clasificación completa ha de desarrollarse previamente una segmentación parcial.

Cheng y col. (2001) durante la etapa de segmentación extraen regiones homogéneas, es decir, regiones de similares características mediante la obtención de una serie de

umbrales, lo que definen como homogeneidad fuzzy. Tras la segmentación vuelven a aplicar un criterio de similitud para fusionar las regiones. El problema que se menciona en dicho trabajo es el que surge como consecuencia de trabajar en el modelo de color RGB con sus tres componentes integradas, ya que se sabe que dichas componentes están altamente correlacionadas debido a su dependencia de la iluminación; por tanto, los umbrales de contraste, que guían la segmentación, pueden resultar no fiables. Además, desde nuestro punto de vista, otro de los inconvenientes lo constituye la necesidad de especificar el número de clases con carácter previo, algo que resulta desconocido cuando se trata de determinar la cobertura terrestre en áreas complejas.

Wangenheim y col. (2008) aplican la metodología de segmentación jerárquica basada en crecimiento de regiones, mediante el uso de medidas de similitud para separar las clases que son diferentes y fusionar las que son similares. Las medidas de similitud se establecen en relación a las componentes de color. El problema que plantea esta aproximación estriba en la necesidad de realizar iteraciones sucesivas hasta encontrar los parámetros de entrada más apropiados, lo que exige un conocimiento previo del tipo de análisis a realizar, algo que en nuestro caso no es posible debido al desconocimiento de la cobertura vegetal existente en las zonas bajo estudio.

Baraldi y Parmiggiani (1996), siguiendo la metodología basada en la división y fusión de regiones, aplican medidas de similitud y disimilitud para separar las clases que son diferentes y fusionar las que son similares. Este trabajo se sitúa en la línea de los propuestos en Wangenheim y col. (2008).

Tellaeche y col. (2008a, b, 2011), plantean técnicas de clasificación en imágenes agrícolas con el fin de aplicar tratamientos selectivos sobre las malas hierbas. Tras un proceso de segmentación previa mediante umbralización basado en Ribeiro y col., (2005), segmentan la imagen en lo que llaman celdas de actuación para discernir si una de tales celdas requiere o no tratamiento; por consiguiente, se trata de clasificadores bi-clase basados en técnicas de la teoría de Bayes, la Teoría de la Decisión Multicriterio y Máquinas de Vectores Soporte respectivamente. Dada la naturaleza bi-clase de dichas técnicas, su aplicación al caso multi-clase, que nos ocupa en el presente trabajo, no resulta apropiada.

Kurugollu y col. (2001), una vez segmentada la imagen, bajo la metodología mencionada previamente, se desarrolla la etapa de fusión y etiquetado (clasificación), consistente en la fusión de los tres pares de componentes en un único par, de forma que la etiqueta final se obtiene por selección de la etiqueta predominante, es decir, por mayoría de coincidencias. Tras la fusión de componentes se realiza un filtrado espacial por mayoría.

En Jhansen y col. (2011) después de la segmentación se aplica un proceso de reagrupamiento, si bien no por características espectrales, sino con base a alguna característica especial, tal como la pendiente.

Guijarro y col. (2009a, b) aplican técnicas de combinación de clasificadores simples para análisis de la cobertura terrestre con imágenes en el modelo de color RGB, captadas por sensores aerotransportados; concretamente se basan en los siguientes clasificadores: *Fuzzy clustering*, Bayes, ventana de Parzen, Cuantización Vectorial o Mapas Autoorganizativos. La combinación se lleva a cabo mediante técnicas basadas en la Teoría de la Decisión Multicriterio *Fuzzy* y en el método de Enfriamiento Simulado como estrategia de optimización. En esta última línea se encuentra el método propuesto por Pajares y col. (2010), que utilizan como técnica de optimización la red neuronal de Hopfield. A pesar de que la naturaleza intrínseca de los clasificadores utilizados para la combinación es supervisada, mediante el diseño de una estrategia que asume un determinado número de clases y evalúa la bondad de la clasificación para diferentes particiones, las estrategias planteadas llegan a ser de naturaleza no supervisada. La validación de las particiones se realiza mediante el cómputo de los denominados índices de validación, que se abordan más adelante en este mismo trabajo. La idea de no supervisión, junto con la validación de las particiones se aplican en el trabajo que se presenta, si bien desde perspectivas diferentes. El buen rendimiento de las técnicas anteriormente mencionadas se justifica sobre la base de que las áreas sobre las que se analiza la cobertura terrestre son amplias y abarcan diferentes tipos de vegetación bien diferenciados sobre el terreno, algo que no ocurre en el caso de las imágenes que se estudian en el presente trabajo.

Los anteriores algoritmos de clasificación no supervisada son aplicados a imágenes a color, mediante sus características espectrales a nivel de píxel.

Modelos aplicados a la clasificación de la cobertura terrestre

La mayoría de los métodos de clasificación pueden ser agrupados de la siguiente manera.

- Clasificación por índices espectrales visibles, para la identificación de las plantas verdes, incluidos los cultivos y las malezas, frente al suelo y otras estructuras (piedras, escombros y elementos extraños presentes en el campo), aplicados por Meyer y Camargo Neto (2008); Tellaeché y col. (2008a, b, 2011); Guijarro y col. (2011), Burgos-Artizzu y col. (2009, 2010, 2011) o Sainz-Costa (2011) entre otros.
- Enfoques no supervisados para imágenes hiperespectrales, procedentes de sensores remotos, como los trabajos de Du (2007) o Shah y col. (2004), donde cada píxel es un modelo de combinación lineal de las señales espectrales pertenecientes al espacio hiper-espectral. Continuando con la línea de imágenes teledetectadas, Secord y Zakhor (2007), con imágenes LIDAR, clasifican zonas urbanas para identificar árboles mediante segmentación por crecimiento de regiones y clasificación mediante el algoritmo de Maquinas de Vectores Soporte Ponderadas.
- Criterios específicos de histograma basados en umbrales, incluyendo umbralización dinámica. Por lo general sólo se consideran dos clases (las plantas y el fondo). En Reid y Searcy (1987), una función de decisión se estima bajo el supuesto de que las clases siguen la distribución de Gauss. El método de Otsu (1979), se ha aplicado a las imágenes en escala de gris considerando un problema bi-clase (Ling y Ruzhitsky, 1996; Shrestha y col., 2004). En Gebhardt y col. (2006), se aplicó un enfoque de umbral, donde las imágenes son previamente transformadas de RGB a un modelo de intensidad de la escala de grises. Este método fue mejorado más tarde con la homogeneidad local y las operaciones morfológicas en Gebhardt y Kauhbauch (2007). Kirk y col. (2009) aplican una combinación de verdes y de intensidad de las bandas espectrales rojo y verde, y calculan un umbral automático para un problema bi-clase asumiendo dos funciones *gaussianas* de densidad de probabilidad asociadas al suelo y a la vegetación, respectivamente. En Meyer y Camargo Neto (2008), se aplica el

método de Otsu, con umbral automático por histograma, para la binarización de la imagen una vez que se extrae el verdor; a continuación se obtienen los índices de diferencias normalizadas. Después de diferentes experimentos, los autores concluyen que un umbral de cero es suficiente para la aplicación propuesta, y por lo tanto el método de Otsu finalmente no se aplica en ese trabajo. En esta misma línea de binarización por índices de vegetación se encuentran los trabajos de Tellaeché y col. (2008a, b, 2011); Burgos-Artizzu y col. (2009, 2010, 2011), Guijarro y col. (2011) o Ribeiro y col. (2005).

- Otros algoritmos, como el de *Mean Shift*, fue aplicado en Zheng y col. (2009), bajo la suposición de que la segmentación de la vegetación verde y el fondo se puede considerar como un problema de la segmentación en dos clases; se validó la separabilidad de las clases a través de una red neuronal y el discriminante lineal de Fisher; los espacios de color utilizados fueron RGB, LUV y HSI.

Liu y col. (2007), con el objetivo de reducir la dimensionalidad de imágenes hiper-espectrales, captadas por sensores aéreos sobre los derrames de petróleo, aplican el algoritmo de agrupamiento RPCCL ("*Rival Penalization Controlled Competitive Learning*"), mejorando en este caso la selección de la semilla inicial, dada la importancia de dicha selección. Esto significa que para encontrar la clasificación adecuada es necesario ejecutar el algoritmo varias veces con diferentes semillas.

2.4. MÉTODOS DE SEGMENTACIÓN POR UMBRALIZACIÓN

Este apartado se inicia con una breve revisión de métodos de umbralización existentes en la literatura, continuando con la descripción de tres algoritmos de umbralización clásicos, que constituyen la base para la construcción de otra estrategia de segmentación conocida como combinación de umbrales.

Existen diferentes planteamientos sobre los cuales se aplican técnicas de segmentación por umbralización, especialmente en aplicaciones que requieren procesar cantidades elevadas de datos, distinguiéndose entre umbralización global y local.

La *umbralización global* es aquella en la cual se establece un único umbral para segmentar la imagen, útil cuando solamente existen dos regiones de píxeles. Una de las aplicaciones más comunes es extraer un objeto del fondo de una imagen. La *umbralización local* o multiumbralización es la que se determinan varios umbrales, cada uno de los cuales separa las regiones que identifican los diferentes objetos. Kirby y Rosenfeld (1979) fueron pioneros en la utilización de métodos de umbralización local; dentro de la umbralización local difusa también se encuentran diversas investigaciones al respecto, entre ellas la de Cheng y Chen (1999). Una extensión del método de umbralización local es el de *umbral dinámico*, en el que se define para cada píxel una zona vecina N a la que se asigna un valor de umbral adecuado. Un ejemplo se puede encontrar en Martin y col. (2009), así como en Macedo y col. (2010). En este caso es imprescindible disponer de un método para calcular el valor umbral adecuado de manera automática. Sobre este contexto se han desarrollado diversas investigaciones presentándose varias técnicas de segmentación que, por lo general, se diseñan para resolver un problema en particular. Gonzales-Barron y Butler (2006), compararon los resultados de siete técnicas de umbralización, concluyendo que el método de Otsu es el que mejores resultados proporciona, con un alto grado de uniformidad y logrando además una correlación mejor que el *k-means* para los valores de umbral óptimo. En esta investigación se menciona que no existe un único método que pueda ser considerado bueno para todas las imágenes, por lo que de acuerdo a las necesidades de cada problema puede ser uno más apto que otro.

En Sezgin y Sankur (2004), se realiza una amplia revisión bibliográfica en la cual se pueden encontrar los principales métodos de umbralización por histograma, bajo la cual se analizan, por ejemplo, los picos, valles y curvaturas del histograma suavizado (Kurugollu y col., 2001). También existen los *métodos de umbralización basados en agrupamiento*, donde las muestras de niveles de gris se agrupan en dos partes, objeto y fondo, o alternativamente se modelan como una mezcla de dos gaussianas, como en Otsu (1979), Kittler y Illingworth (1986), o el método de umbral iterativo de Riddler (1978), también conocido como ISODATA, además del método de umbralización por mínimo error (Lloyd, 1985) y el de agrupamientos difusos (Jawahar y col. 1997).

Así pues, determinar el umbral óptimo es una de las etapas más importantes en los procesos de segmentación y clasificación en general y en el caso que nos ocupa en

particular. La determinación de umbral resulta útil además en aplicaciones basadas en compresión y recuperación de imágenes, como por ejemplo Chang y col., (2008) o Tai y col. (1998), que mediante técnicas de BTC (*block truncation coding*) y algoritmos genéticos comprimen y recuperan imágenes en color. Las técnicas BTC requieren determinar el umbral óptimo para cada bloque puesto que pasan por un proceso de binarización por bloque y por componente espectral. Aunque esto último no constituye el objetivo inmediato en el presente trabajo, en el futuro podría ser de utilidad tanto para almacenar como para recuperar las imágenes clasificadas, que se supone formarán un conjunto cada vez más amplio.

La umbralización se aplica incluso a sistemas de control, como por ejemplo, en Burgos-Artizzu y col. (2007), donde para calibrar la dosis de herbicida precisó de un procesamiento previo de imágenes del cultivo que determinara tanto el estado de crecimiento de la mala hierba como el índice de cobertura, para ello se utilizaron técnicas de segmentación por binarización.

Los resultados de Sezgin y Sankur muestran que de los cuarenta métodos de umbralización probados con imágenes en escala de gris, los seis con mejor calidad promedio fueron: “Cluster-Kittler”, “Cluster-Lloyd” y “Cluster-Otsu”, desarrollados por Kittler e Illingworth (1986), Lloyd (1985) y Otsu (1979), respectivamente, así como “Entropy-Kapur”, “Entropy-sahoo” y “Entropy-Yen”, aplicados por Kapur y col., (1985), Sahoo y col., (1997) y Yen y col., (1995).

Chang y col. (2006) realizó un análisis comparativo de doce técnicas de umbralización, entre las que se encuentra el método de Otsu, argumentando que ésta es una de las técnicas que ha proporcionado muy buenos resultados en umbralización por lo que no podía faltar en dicho análisis. Probando imágenes bajo luz controlada y no controlada, en el análisis comparativo se evaluó la uniformidad y la inspección visual. En cuanto a la uniformidad, Otsu se clasifica dentro de los tres mejores, y en cuanto a la inspección visual el método de Otsu consiguió los mejores resultados de entre todos los analizados.

Mientras tanto Stathis y col. (2008), evalúan treinta algoritmos de umbralización con imágenes en escala de gris, clasificándolos en globales y locales e incluyendo el algoritmo de Otsu en ambas clasificaciones.

Puesto que existen numerosos algoritmos de umbralización, éstos se han aplicado a diferentes aéreas de estudio, en algunos casos como apoyo tanto en clasificación supervisada como no supervisada. Por ejemplo, en visión de robots no tripulados el procesamiento en tiempo real exige algoritmos robustos, por lo que lo ideal es simplificar la información mediante binarización. Silveira (2001) diseña una estrategia para guiar un robot aéreo mediante visión en carreteras aplicando una estrategia de umbral adaptativo para binarizar la imagen y encontrar las líneas de carretera. Otro de los campos en que se aplica procesamiento en tiempo real es en la industria, por ejemplo en la avícola para clasificar el huevo defectuoso (Ribeiro y col. 2000; García-Alegre, 2000). Yan y col. (2007) lo aplican para segmentar las líneas eléctricas captadas por sensores aerotransportados. Y continuando con la segmentación de imágenes aerotransportadas, Cao y col. (2005), también clasifican los objetos creados por el hombre apoyándose en algunas técnicas de umbralización. Bicego y col. (2003), clasifican las calles y caminos con imágenes aéreas. Tuo y col. (2004) aplican histograma directo para mejorar las imágenes captadas por diferentes tipos de sensores aerotransportados. Burgos-Artizzu y col. (2010), así como Tellaache y col. (2008a, b, 2011), aplican un proceso de binarización de la imagen a color durante la segmentación con el fin de discernir entre plantas de cultivo, malas hierbas y suelo. Zhang y col. (2008) centran su esfuerzo en la umbralización con fines de procesamiento en tiempo real.

Montiel y col. (2005) proponen utilizar la información proporcionada por las distribuciones de los niveles espectrales a través del análisis de los correspondientes histogramas de las imágenes que contienen las texturas.

Dentro de los trabajos de segmentación en color, Vázquez y col. (2007), así como Tkalcic (2003), identifican a partir del histograma pequeñas variaciones del color con el fin de suprimirlas, consiguiendo así un cierto suavizado y con ello la homogenización de las regiones de texturas.

Roaf y col. (2008) desarrollan una segmentación por umbralización de imágenes RGB, si bien los umbrales de cada banda se obtienen mediante el análisis de los valores, medios, máximos y mínimos que selecciona el experto en la materia.

Sirmacek y col. (2008) y Unsalan y Boyer (2005), procesan imágenes multiespectrales aplicando segmentación mediante *k-means-cluster*, consiguen la binarización mediante técnicas de bordes, logrando identificar edificios y calles.

En cuanto a los trabajos en los cuales se aplica multi-umbralización, Lim y Lee (1998), proponen un método basado en dos etapas. La primera es una multi-umbralización gruesa, mediante un filtro de espacio escalar, seleccionando manualmente el número de clases. En la segunda etapa aplican técnicas de agrupamiento del tipo *Fuzzy C-Means* para refinar la segmentación, ajustando posteriormente el número de clases.

Continuando con la multi-umbralización, Jhansen y col. (2011) aplican clasificación basada en píxeles para el análisis de una zona ribereña con imágenes LIDAR, de manera que clasifican la corriente central de los arroyos determinando los umbrales empíricamente.

En el ámbito de clasificación de imágenes de alta resolución, tales como las proporcionadas por el satélite IKONOS, Cheng y col. (2008) desarrollan un trabajo para clasificar la vegetación existente, en el cual normalizan las diferencias de los índices de vegetación (NDVI) y seleccionan de forma manual, con base a la visualización del histograma de NDVI, un determinado umbral, el cual se aplica para discernir entre dos tipos de vegetación. El trabajo concluye afirmando que hace falta un esquema de soporte teórico sólido que permita determinar el umbral.

Además de la existencia de los diferentes métodos de umbralización simples, también se han desarrollado diversas aplicaciones sobre la combinación de éstos. Dentro de ellas se encuentran las de Guijarro y col. (2011), que aplican un modelo de segmentación automática de texturas en imágenes agrícolas, en el cual como paso previo a la clasificación se desarrolla una metodología de binarización, aplicando los algoritmos de manera independiente y también combinados. Finalmente, después de comparar los resultados, afirman que el modelo de combinación de algoritmos proporcionó mejores resultados que los individuales.

Herrera y col. (2011) también combinan tres métodos de umbralización durante la segmentación de imágenes hemisféricas de los entornos forestales analizados con fines de inventarios forestales y mantenimiento.

El resultado de cada método de segmentación dependerá del algoritmo en concreto, del valor de los parámetros y de la medida de similitud/disimilitud adoptada. Es decir para realizar el agrupamiento de los objetos, es necesario determinar cuándo dos objetos del espacio son parecidos y cuándo no. Con este fin se definen las funciones de similitud o de disimilitud, entre las que se encuentran las métricas o distancias.

A continuación se describen tres de los métodos de umbralización que han proporcionado buenos resultados en imágenes naturales donde se han aplicado, y que se han utilizado en la presente investigación tanto de manera independiente como combinados.

2.4.1. MÉTODO DE OTSU

El método de Otsu es llamado así en honor a *Nobuyuki Otsu*, que lo propuso en 1979.

Petrou (1999) lo describe como un método que no depende del modelado de las funciones de densidad de probabilidad. Es una técnica no paramétrica y no supervisada de umbralización óptima para separar dos clases entre sí. Cada clase viene caracterizada por su centroide y por el área que forma la unión de los puntos asociados a dicho centroide. El algoritmo busca la maximización de la distancia entre las dos clases y la minimización de sus áreas. En concreto, se utiliza la *varianza*, que es una medida de la dispersión de valores en cada clase, en el caso que nos ocupa se trata de la dispersión de los niveles de gris. Dado que son dos las clases involucradas, se calcula el cociente entre ambas variancias y se busca un valor umbral para el que este cociente sea máximo, aplicando un procedimiento que utiliza únicamente los momentos acumulados de orden cero y primer orden del histograma en nivel de gris.

La operación umbral es considerada como la división de los píxeles de una imagen de L niveles de gris en dos grupos, de forma intuitiva por ejemplo objetos y fondo. Es decir, una imagen en escala de gris se puede modelar como una función de intensidad bidimensional de tamaño N ($m \times n = N$), donde m es el número total de filas, y n el número de columnas. Los elementos de la imagen $i(x,y)$ se representan por los L niveles de gris ($i = 0, 1, 2, \dots, L-1$). El número de píxeles en el nivel i se denota por $h(i)$, de esta forma es posible construir el histograma, que incluye todos los niveles de intensidad o de gris. El histograma se normaliza y se considera como una distribución de

probabilidad. La probabilidad de ocurrencia de cada nivel de gris $p(i)$ se representa por la ecuación (2-3) con sus propiedades asociadas,

$$p(i) = \frac{h(i)}{N}, \quad p(i) \geq 0, \quad \sum_{i=0}^{L-1} p(i) = 1 \quad (2-3)$$

La ecuación (2-4) define el momento acumulado de orden cero $w(t)$ o probabilidad acumulada; la ecuación (2-5) representa el momento acumulado de primer orden $\mu(t)$ o media de clases ponderada del histograma hasta el t -ésimo nivel; finalmente el nivel medio total μ_T de la imagen viene dado por la ecuación (2-6).

$$w(t) = \sum_{i=0}^t p(i) \quad (2-4)$$

$$\mu(t) = \sum_{i=0}^t i \cdot p(i) \quad (2-5)$$

$$\mu_T = \sum_{i=0}^{L-1} i \cdot p(i) \quad (2-6)$$

Por ejemplo, supongamos que se busca dividir los píxeles de la imagen en dos clases, C_1 y C_2 (fondo y objeto, o viceversa) por un umbral $T(t) = t$, $0 < t < L-1$, donde C_1 denota todos los píxeles de la imagen con valor de intensidad en el rango $(0, \dots, t)$ y C_2 los píxeles con valores en el rango $(t+1, \dots, L-1)$. Por tanto, la probabilidad acumulada de que un píxel se asigne a la clase C_1 se obtiene mediante w_1 . Dicho de otra manera, w_1 es la probabilidad de que la clase C_1 ocurra. Por ejemplo, si $t = 0$, la probabilidad de que C_1 tenga algún píxel asignado es cero. De igual manera, la probabilidad de la clase C_2 es obtenida por w_2 ; de esta forma el valor de intensidad media ponderada de los píxeles asignados a la clase C_1 se obtiene mediante μ_1 , y la de la clase C_2 mediante μ_2 , ecuación 2-7).

$$w_1 = w(t) = \sum_{i=0}^t p(i); \quad \mu_1 = \frac{\mu(t)}{w(t)} = \frac{\sum_{i=0}^t i \cdot p(i)}{w_1} \quad (2-7)$$

$$w_2 = w(t) = \sum_{i=t+1}^L p(i) = 1 - w_1; \quad \mu_2 = \frac{\mu(t)}{w_2} = \frac{\sum_{i=t+1}^L i \cdot p(i)}{w_2} = \frac{u_T - u(t)}{1 - w(t)}$$

Es fácil verificar la siguiente relación para cualquier elección de t :

$$w_1 + w_2 = 1 \quad \text{y} \quad w_1 u_1 + w_2 u_2 = u_T \tag{2-8}$$

Las varianzas de clases están dadas por la ecuación (2-9), donde se definen los momentos estadísticos acumulativos de segundo orden.

$$\sigma_1^2 = \frac{\sum_{i=0}^t (i - \mu_1)^2 p(i)}{w_1} \quad \sigma_2^2 = \frac{\sum_{i=t+1}^L (i - \mu_2)^2 p(i)}{w_2} \tag{2-9}$$

Con el propósito de evaluar el mejor umbral del nivel t se usa la dimensión métrica normalizada de separabilidad de clases, utilizada en el análisis discriminante, ecuación (2-10). Este punto de vista está motivado por la conjetura de que un umbral que proporcione la mejor separación de clases en los niveles de gris representará el mejor umbral.

$$\eta(t) = \frac{\sigma_B^2(t)}{\sigma_G^2} \quad \text{o} \quad \lambda(t) = \frac{\sigma_B^2(t)}{\sigma_W^2(t)} \tag{2-10}$$

donde σ_G^2 es la *varianza global*.

$$\sigma_G^2 = \sum_{i=0}^{L-1} (i - \mu_T)^2 \cdot p(i) \tag{2-11}$$

Usando el análisis discriminante, la *varianza entre clases* $\{\sigma_B^2(t)\}$ de la imagen umbralizada definida por Otsu (1979) para el bi-nivel es:

$$\sigma_B^2(t) = w_1 (\mu_1 - \mu_T)^2 + w_2 (\mu_2 - \mu_T)^2 = w_1 w_2 (\mu_2 - \mu_1)^2 = \tag{2-12}$$

$$\sigma_B^2(t) = \frac{[\mu_T w(t) - \mu(t)]^2}{w(t)[1 - w(t)]}$$

La varianza dentro de las clases se representa por la ecuación (2-13).

$$\sigma_w^2(t) = w_1 \sigma_1^2 + w_2 \sigma_2^2 \quad (2-13)$$

Por tanto, el problema se reduce a un problema de optimización para buscar el umbral t que maximice la función objetivo, ecuaciones (2-14) y (2-15). Es decir, el umbral óptimo t_{Otsu} es aquél en el cual la *distancia entre las clases* obtiene el valor máximo.

$$\sigma_B^2(t_{Otsu}) = \max_{1 \leq t < L} \sigma_B^2(t) \quad (2-14)$$

$$\eta(t_{Otsu}) = \max_{1 \leq t < L} \eta(t) \quad \lambda(t_{Otsu}) = \max_{1 \leq t < L} \lambda(t) \quad (2-15)$$

2.4.2. UMBRALIZACIÓN POR ISODATA

Isodata (*Iterative Self-Organizing Data Analysis*) es una técnica simple iterativa desarrollada por Ridler y Calvard en 1978, se encuentra descrito en Pajares y Cruz (2007a, b). El objetivo del algoritmo Isodata consiste en dividir el histograma en dos partes que representan a su vez dos sub-regiones en la imagen. Se trata de un método iterativo que obtiene el umbral mediante los siguientes pasos: a) calcula el valor medio de la imagen o primer umbral, el cual determina dos clases formadas por los píxeles cuya intensidad es menor y mayor respectivamente que dicho umbral; b) en la iteración siguiente calcula los valores medios de las clases establecidas en el paso previo, el promedio de esos dos valores medios determina un nuevo umbral; c) el nuevo umbral determina a su vez otras dos clases sobre las que se repite el mismo proceso que en el paso anterior. El proceso se repite hasta que dos valores de umbral, en iteraciones consecutivas, difieren menos de un valor pequeño previamente prefijado.

Los valores de los píxeles del histograma se denotan por $h(0), h(1), \dots, h(L-1)$, donde $h(i)$ especifica el número de píxeles de una imagen cuyo con valor i en escala de gris, y $L-1$ es el valor máximo del píxel en la imagen en dicha escala. La estimación inicial en t_i es el valor medio. Entonces, para valores menores o iguales que éste, $t \leq t_i$, se calcula

el promedio $\mu_1(t)$; para el resto, se obtiene $\mu_2(t)$, ecuación (2-16). Estos valores medios son los que se calculan en los pasos anteriores hasta obtener el umbral final buscado t_f .

$$\mu_1(t) = \sum_{i=0}^t i \cdot h(i) / \sum_{i=0}^t h(i) \quad \mu_2(t) = \sum_{i=t+1}^{L-1} i \cdot h(i) / \sum_{i=t+1}^{L-1} h(i) \quad (2-16)$$

2.4.3. ALGORITMO DE UMBRALIZACIÓN FUZZY

Huang y Wang (1995) aplicaron la teoría de conjuntos difusos para dividir el histograma de una imagen en dos regiones, cuyo objetivo es minimizar una medida de borrosidad, que se define a tal efecto. Esta medida se puede expresar a través de términos diferentes, una de ellas es la entropía. La función de pertenencia, $\mu_F(I(x, y))$, puede ser vista como una función característica que representa la borrosidad de un píxel determinado (x, y) de la imagen I .

$$\mu_F(I(x, y)) = \begin{cases} \frac{1}{1 + |I(x, y) - \mu_1(t)| / D} & \text{if } I(x, y) \leq t \\ \frac{1}{1 + |I(x, y) - \mu_2(t)| / D} & \text{if } I(x, y) > t \end{cases} \quad (2-17)$$

donde los niveles promedios de gris $\mu_1(t)$ y $\mu_2(t)$ son los definidos en la ecuación (2-16), que diferencian dos regiones en la imagen para un valor determinado de umbral t ; D es un valor constante que se establece con el fin de que se cumpla la relación siguiente: $0.5 \leq \mu_F(I(x, y)) \leq 1$. De acuerdo con Gonzalez-Barron (2006), hemos optado por la entropía de la imagen, $E(I)$, obtenida mediante la ecuación (2-18), como medida de borrosidad por medio de la función de Shannon, ecuación (2-19).

$$E(I) = \frac{1}{MN \ln 2} \sum_{i=0}^{L-1} S(\mu_F(i)) h(i) \quad (2-18)$$

$$S(\mu_F(i)) = -\mu_F(i) \ln[\mu_F(i)] - [1 - \mu_F(i)] \ln[1 - \mu_F(i)] \quad (2-19)$$

El umbral óptimo, t_f , se obtiene minimizando la medida de borrosidad $E(I)$ según la ecuación (2-20) como sigue,

$$t_F = \arg \min E(I) \quad (2-20)$$

2.5. MÉTODOS DE AGRUPAMIENTO

En la sección 2.3.4 ya se revisaron algunas aportaciones con respecto a los métodos de agrupamiento, relativos a la clasificación no supervisada. El enfoque proporcionado allí se centró en métodos de clasificación donde previamente se aplica algún proceso de segmentación parcial. Ahora el enfoque que se formula en esta sección se orienta hacia la descripción de métodos de agrupamiento o clasificación sin procesamiento previo, cuyo objetivo es identificar un método de agrupamiento que permita validar los resultados de la estrategia propuesta en el presente trabajo de investigación

Generalmente, los métodos de agrupamiento suelen dividirse de diferentes maneras. Siguiendo la taxonomía establecida por Theodoridis y Kouroumbas (2009) los algoritmos de agrupamiento se dividen en categorías según el procedimiento que utilizan para agrupar los objetos:

- a) *Algoritmos jerárquicos*: como su nombre indica, construyen una jerarquía de agrupamientos, uniendo o dividiendo los grupos de acuerdo a una cierta función de similitud/disimilitud entre los grupos. En otras palabras, construyen un árbol de *clústeres* llamado dendograma. Los métodos de agrupamiento jerárquicos se categorizan en aglomerativos (*bottom-up*) y divisivos (*top-down*). Un agrupamiento aglomerativo, generalmente, comienza con grupos unitarios (*singleton clusters*) y, recursivamente, une dos o más clústeres, siempre de acuerdo a alguna función de similitud/disimilitud. Un agrupamiento divisivo, generalmente, comienza con un único clúster en el que están todos los puntos o datos y, recursivamente, divide el *clúster*, ya sea minimizando o maximizando alguna función que estime el agrupamiento óptimo. El proceso continúa hasta que se alcanza algún criterio de finalización, generalmente el número k de *clústeres*. Entre las ventajas de los algoritmos de agrupamiento jerárquicos se pueden mencionar la flexibilidad con respecto al nivel de granularidad, su fácil manejo y su aplicabilidad en relación a cualquier tipo de atributo. Entre las desventajas se encuentra la no existencia de un criterio de parada claro, salvo el número de clústeres, además de que una vez contruidos éstos, no vuelven a ser

visitados con el fin de mejorarlos (Wangenheim y col., 2008; Gashnikov y col., 2009; Trias-Sanz y col., 2008; HandlyKnowles, 2005).

- b) *Métodos por partición*: emplean diferentes técnicas de reubicación para asignar los puntos o datos a uno de los k *clústeres*; algunos buscan puntos específicos para luego reubicar los restantes, otros tienen un enfoque probabilístico. A diferencia de los métodos jerárquicos tradicionales, en algunos casos los *clústeres* son revisitados después de construidos, de forma que con los puntos reubicados se mejora el agrupamiento. Estos algoritmos muchas veces asumen un conocimiento a priori del número de *clústeres* en que debe ser dividido el conjunto de datos. La idea más usada es hallar los centroides, uno para cada *clúster*, y luego ubicar los restantes puntos en el grupo del centroide más cercano. Este método tiene como desventaja que falla cuando los puntos de un *clúster* se sitúan muy próximos al centroide de otro grupo (Hu y col., 2008; Chu y col., 2002; Hartigan y Wong, 2009; HandlyKnowles, 2005; Pajares y Cruz, 2002; Mezaris y col., 2004).
- c) *Algoritmos basados en densidad*: tratan de identificar *clústeres* en zonas altamente pobladas. Emplean diferentes técnicas para determinar los grupos: por grafos, basadas en histogramas, “kernels”, aplicando la regla K-NN o tratando de descubrir subgrupos denso-conectados (Shu y col., 2003; Márcio y col., 2008).
- d) *Métodos basados en rejillas*: trabajan con los datos indirectamente, construyendo resúmenes de los propios datos sobre el subconjunto del espacio de atributos, realizan una segmentación del espacio y seleccionan segmentos apropiados (Stenberg y col., 2008; Secord y Zakhor, 2007)
- e) *Métodos basados en co-ocurrencia*: consideran que el concepto de similitud por sí solo no es suficiente para agrupar algunos datos, por lo que introducen otros conceptos tales como relaciones entre los vecinos (Li y col., 2008).

Otros autores como Bow (2002), para su división toman en cuenta la existencia o no de una función criterio a optimizar, clasificándolos pues en *directos* e *indirectos*:

- a) *Directos o heurísticos*: son los que no optimizan ninguna función objetivo, es decir, simplemente se basan en alguna observación previa que define la heurística (Corney y col., 2008; Yang y col., 2008).
- b) *Indirectos o por optimización*: son los que optimizan alguna función objetivo. Un ejemplo de optimización es el de Angus (2007), quien propuso una función de optimización Multi-Objetivo a la que denominó PACO por sus siglas en ingles de “*Population-based Ant Colony Optimization*”.

Asimismo, Bow (2002) menciona que suele ser habitual clasificar los algoritmos de agrupamiento según su construcción en *aglomerativos*, *divisivos* y *mixtos* (Chen y col., 2002).

Por otra parte, existe la posibilidad de distinguir el tipo de algoritmos en función del conocimiento que se tenga a priori del número de grupos:

- a) *Si se conoce el número de clústeres*, donde se encuadran los algoritmos siguientes: minimización de la suma de la distancia al cuadrado, ISODATA, técnica de búsqueda dinámica del óptimo agrupamiento y el método de agrupamiento borroso o *fuzzy c-means* (Saeed, 2006, Hu y col., 2008).
- b) *Si no se conoce el número de clústeres*, entre los que se encuentran los siguientes como más relevantes: el método heurístico o también conocido como conjunto de muestras adaptativas, el algoritmo de Batchelor y Wilkin, el algoritmo de agrupamiento jerárquico basado en los k vecinos cercanos (Gómez, 1994; Macedo y col., 2010).

De entre los algoritmos antes mencionados, se seleccionó el de agrupamiento *fuzzy c-means* como método de comparación para validar los resultados frente al método propuesto en este trabajo, las razones que justifican su elección son las siguientes:

- a) Se trata de un método ampliamente utilizado, con resultados satisfactorios, en distintas aplicaciones, incluyendo la clasificación de texturas en imágenes naturales (Guijarro y col., 2009a,b; Pajares y col., 2010; Herrera y col., 2011).
- b) Es un método supervisado que requiere el conocimiento previo del número de clústeres, lo cual es factible en tanto en cuanto el algoritmo propuesto en este trabajo proporciona esta información.

- c) Dado que el método proporciona una partición, resulta factible aplicar índices de validación, sección 2.6.2, con respecto a la partición proporcionada por el método propuesto en este trabajo de investigación. De esta forma, es posible la comparación objetiva de los resultados obtenidos.
- d) Dada la naturaleza borrosa del mismo, es posible la asignación de píxeles a diferentes clases con distintos grados de pertenencia, algo perfectamente asumible y coincidente con la realidad en el caso de las imágenes analizadas en la presente investigación.

2.5.1. AGRUPAMIENTO FUZZY C-MEANS

“*Fuzzy Clustering*” ampliamente conocido como tal en su terminología inglesa, consiste en dividir n objetos, identificados por $x \in X$ y caracterizados por p propiedades, en c *clústeres* o *grupos*. El término *fuzzy*, introducido ya con anterioridad, es sinónimo de los términos difuso y borroso en español, por lo que se usan indistintamente.

Con el fin de formalizar el método, sea el conjunto de datos $X = \{x_1, x_2, \dots, x_n\} \in \mathfrak{R}^p$ un subconjunto del espacio real p -dimensional \mathfrak{R}^p . Cada $x_k = \{x_{k_1}, x_{k_2}, \dots, x_{k_p}\} \in \mathfrak{R}^p$ se denomina *vector de características*, x_{k_j} es la j -ésima característica de la observación x_k .

Como se ha mencionado previamente y centrándonos en las imágenes que se utilizan en este trabajo de investigación, las características son los píxeles y sus propiedades son las componentes espectrales en el modelo de color RGB, por tanto, $p = 3$ y $x_k = \{R, G, B\} \in \mathfrak{R}^3$.

Esta estrategia fue originalmente introducida por Bezdek (1981). Su descripción se puede encontrar en Bezdek (2005), Duda y col. (2001) o Pajares y Cruz (2007a,b), entre otros.

Puesto que los elementos de un clúster deben ser tan similares entre sí como sea posible y a la vez deben ser tan diferentes a los elementos de otros clústeres como también sea posible, el proceso de clasificación se controla por el uso de medidas de similitud basadas en distancias. Así la similitud o la diferencia entre dos puntos x_k y x_l puede interpretarse como la distancia entre esos puntos.

La distribución de las muestras del conjunto X en clases se conoce como *partición*. Si se desea realizar una partición del conjunto X en c clases, se tendrán $S_i \{i = 1, \dots, c\}$ subconjuntos de X , donde cada subconjunto constituirá una clase. Una partición puede enfocarse desde dos perspectivas: *fuzzy* y *no fuzzy*. Una partición *no fuzzy* se conoce en terminología inglesa como *crisp*. Desde el punto de vista *fuzzy* para cada clase S_j se puede definir un conjunto borroso μ_i sobre el universo X , $\mu_i : X \rightarrow [0,1]$ que asigna a cada muestra x_k lo que se conoce como grado de pertenencia de dicha muestra al subconjunto S_j , denominada clase a partir de ese momento.

Los valores tomados corresponden al intervalo continuo (0,1). Esto se puede expresar como $\mu_i(x_k) \equiv \mu_{ik}$. En el caso de conjuntos *no fuzzy*, una muestra x_k pertenece a una clase S_i solamente y no pertenece al resto de las clases. Esto se expresa mediante los valores discretos $\{0, 1\}$, siendo $\mu_{ik} = 1$ para indicar que pertenece y $\mu_{ik} = 0$ para expresar que no pertenece. En el caso de conjuntos borrosos se dice que una muestra puede pertenecer a diferentes clases y así se habla por ejemplo de que x_k pertenece a una clase S_i con μ_{ik} grado de pertenencia y a otra clase S_j con μ_{ij} grado de pertenencia.

Dado el conjunto de muestras $X = \{x_1, x_2, \dots, x_n\}$ y el conjunto V_{cn} de todas las matrices reales de dimensión $c \times n$, con $2 \leq c < n$, se puede obtener una matriz de grado de pertenencia representando una partición de las muestras en clases de la siguiente manera $U = \{\mu_{ik}\} \in V_{cn}$. Tanto en el supuesto "*crisp*" como en el *fuzzy* se deben cumplir las siguientes condiciones:

$$\mu_{ik} \in \{0,1\} \text{ crisp } o \quad \mu_{ik} \in [0,1] \text{ fuzzy} \quad 1 \leq i \leq c; \quad 1 \leq k \leq n$$

$$\sum_{i=1}^c \mu_{ik} = 1 \quad 1 \leq k \leq n \tag{2-21}$$

$$0 < \sum_{k=1}^n \mu_{ik} < n \quad 1 \leq i \leq c$$

La localización de un *clúster* S_i se representa por su centro $v_i = \{v_{i_1}, v_{i_2}, \dots, v_{i_p}\} \in \mathfrak{R}^p$ con $i = 1, \dots, c$, alrededor del cual se concentran los objetos.

La definición básica para llevar a cabo el problema de la partición *fuzzy* para $m > 1$ consiste en minimizar la siguiente función objetivo:

$$\min z_m(U; v) = \sum_{k=1}^n \sum_{i=1}^c m_{ik}^m \|x_k - v_i\|_G^2 \quad (2-22)$$

G es una matriz de dimensión $p \times p$ simétrica y definida positiva. Así se puede definir una norma general del tipo,

$$\|x_k - v_i\|_G^2 = (x_k - v_i)^t G (x_k - v_i) \quad (2-23)$$

Diferenciando la función objetivo para v_i (suponiendo constante U) y μ_{ik} (suponiendo constante v), y aplicando la condición de que $\sum_{i=1}^c \mu_{ik} = 1$, se obtiene,

$$v_i = \frac{1}{\sum_{k=1}^n (\mu_{ik})^m} \sum_{k=1}^n (\mu_{ik})^m x_k \quad i = 1, \dots, c \quad (2-24)$$

$$m_{ik} = \frac{\left(\frac{1}{\|x_k - v_i\|_G^2} \right)^{2/m-1}}{\sum_{j=1}^c \left(\frac{1}{\|x_k - v_j\|_G^2} \right)^{2/m-1}} \quad i = 1, \dots, c; k = 1, \dots, n \quad (2-25)$$

El exponente m se conoce como peso exponencial y reduce la influencia del ruido al obtener los centros de los clústeres, reduciendo la influencia de los valores pequeños de μ_{ik} (puntos lejos de v_i) frente a valores altos de μ_{ik} (puntos cerca de v_i). Cuanto mayor sea $m > 1$, mayor es dicha influencia.

Por tanto, para llevar a cabo el proceso de agrupamiento es necesario definir los siguientes parámetros:

- a) Número de clústeres c : se asigna el resultado obtenido del sistema de clasificación Otsu propuesto en la presente investigación.
- b) El peso exponencial m : para el presente caso se tomó el valor 2, por ser el valor ampliamente utilizado según diversos estudios previos (Bezdek, 2005).
- c) Máximo número de iteraciones: establecido en 100 por haberse constatado suficiente, ya que dicho valor no ha sido alcanzado nunca en los experimentos llevados a cabo en este trabajo.
- d) Valor mínimo que determina el cambio en la función objetivo entre dos iteraciones sucesivas, fijado al valor de $1e-5$.
- e) La matriz G que induce la norma.

El proceso de agrupamiento se detiene cuando el número máximo de iteraciones es alcanzado, o cuando la variación de la función objetivo entre dos iteraciones consecutivas se sitúa por debajo del valor previamente establecido.

Especial mención requiere el estudio de la matriz G que determina la forma del grupo. Si se elige la norma Euclídea, entonces G es la matriz identidad I y la forma de los clústeres se asume que constituyen una hiper-esfera. G también se puede elegir como una matriz diagonal con $G_D = [diag(\sigma_j^2)]^{-1}$ o la norma de Mahalanobis con $G_M = [cov(x)]^{-1}$, donde σ_j^2 denota la varianza de la característica j y cov la covarianza. En el presente caso G es la matriz identidad.

El criterio que sigue el clasificador *fuzzy* para determinar a qué clase pertenece x_k consiste en seleccionar el máximo valor de los grados de pertenencia de entre todos los obtenidos, y asignar x_k a la clase asociada con este valor máximo. Formalmente esto se expresa como sigue:

$$x_k \in w_j \mid \mu_{kj} > \mu_{kl}, \forall l \neq j, j = 1, \dots, c \quad (2-26)$$

2.6. EVALUACIÓN DEL SISTEMA DE CLASIFICACIÓN NO SUPERVISADA DE COBERTURA TERRESTRE

El objetivo de los métodos de clasificación no supervisada es encontrar agrupamientos significativos presentes en un conjunto de datos. Estas agrupaciones se conocen como particiones. Uno de los elementos clave en los métodos de clasificación es determinar la bondad o calidad de las particiones obtenidas. Un criterio para medir dicha calidad consiste en buscar grupos de forma que los elementos dentro de cada grupo sean lo más similares posible entre sí y por consiguiente, se encuentren muy bien agrupados alrededor del centro o representante del clúster. Simultáneamente, la calidad también se mide en función de la separación entre los clústeres, de forma que a mayor separación entre los centros de los clústeres mayor calidad.

En los problemas de clasificación no supervisada la determinación del número óptimo de clústeres constituye un aspecto esencial de la clasificación, ya que a priori resulta desconocido. En este sentido, la validación de la partición resulta ser un paso fundamental en la clasificación no supervisada, que es el caso que nos ocupa en la presente investigación.

Los interrogantes que nos planteamos durante el proceso de validación del clasificador no supervisado propuesto son:

1. ¿El número de agrupamientos generado por el clasificador es el óptimo?
2. ¿Cuál es la correspondencia que existe, temáticamente hablando, entre las clases asignadas a un píxel con nuestro clasificador y las asignadas en la imagen de referencia creada con la colaboración de un experto?

Las formas de evaluar la calidad de los agrupamientos ha ido evolucionando desde la simple observación de resultados, con un "parece bueno", dado que este proceso de análisis de datos es muy subjetivo, hasta la aplicación de diferentes índices de validación utilizados para tal fin, así como por la comparación sobre el terreno de localizaciones concretas identificadas en las imágenes analizadas, hasta llegar al estudio de la matriz de error con datos deterministas o mediante las técnicas *fuzzy*.

Con el objetivo de plasmar el estado del arte en cuanto a los índices de validación que permiten determinar el número óptimo de agrupamientos, así como la veracidad de la

clasificación temática de cada píxel, en esta sección se realiza una descripción de los métodos más comúnmente utilizados en la literatura sobre la validación de un clasificador no supervisado en relación al número de *clústeres*. Específicamente, los indicadores de validación, aplicados a la presente investigación son: el índice de Calinski y Harabasz (1974), el de Davies y Bouldin (1979), el índice de Krzanowski y Lai (1985), el de Dunn (1974) y el de *Hartigan* (1985). Además para evaluar la precisión temática se aplica como método de validación externa, el índice de kappa derivado de la matriz de error.

2.6.1. ÍNDICES DE VALIDACIÓN PARA UN CLASIFICADOR NO SUPERVISADO

Como se ha mencionado previamente, un índice de validación de agrupamientos proporciona una medida cuantitativa y objetiva del resultado de la partición obtenida, y su valor óptimo se usa para determinar el número óptimo de clústeres en el caso de los métodos no supervisados (Halkidi y Vazirgiannis, 2001; 2002a; 2002b; Wang y col., 2009).

Existen diversas técnicas de validación de agrupamientos, diferenciándose principalmente entre internas y externas (Halkidi y col., 2002a; Milligan y Cooper, 1985; Handl y col. 2005; Wang y col., 2009). Estos dos grupos de técnicas difieren radicalmente en sus enfoques, y encuentran aplicación en distintas condiciones experimentales. Las medidas de *validación interna* tienen como objetivo evaluar el resultado de un algoritmo de agrupamiento utilizando sólo las cantidades y las características inherentes al conjunto de datos; es decir, evalúa la partición en base a los datos y las distancias entre ellos, con el objetivo de evaluar si la estructura es intrínsecamente adecuada en función de los datos disponibles (Milligan y Cooper, 1985; Jain y Dubes, 1988; Jain y col., 1999; Theodoridis y Koutroumbas, 2009).

Las medidas de *validación externas* comprenden aquellos métodos que evalúan los resultados de la agrupación asumiendo el conocimiento de la partición correcta de los datos. Es decir, la verdadera partición es conocida, que se conoce como *ground truth* o *standard gold*. Evidentemente, en un contexto real de clasificación no supervisada, la partición correcta no es conocida, pero puede ser que en un contexto experimental de evaluación de algoritmos de agrupamiento sí lo sea. La validación se desarrolla sólo con una muestra representativa de los datos.

En la literatura consultada en relación a esta temática, dichos conceptos y métodos aparecen publicados en diversos trabajos entre los que destacan los siguientes: Rousseeuw (1987), Geva (2000), Levine (2001), Dimitriadou (2002), Lange (2004), Gusarova (2005), Mufti (2005), Bertrand (2006), Bouguessa (2006), Barzily (2008), Theodoridis y Koutroumbas (2009), Congalto (2009).

2.6.2. VALIDACIÓN DEL NÚMERO DE AGRUPAMIENTOS (ÍNDICES DE VALIDACIÓN INTERNA)

Las medidas de validación interna toman un agrupamiento y el conjunto de datos como entrada para evaluar la calidad de la información y usan las mismas categorizaciones que usó el método de agrupamiento. Por tanto se aplican las tres categorías fundamentales necesarias a considerar en una clasificación de calidad:

- *Compacidad.* Este grupo comprende los índices de validación que permiten evaluar la compacidad de los grupos o la homogeneidad interna entre ellos. Es decir, los miembros de cada grupo deben situarse lo más cerca posible entre sí. Una medida común de su diseño compacto es la varianza dentro de cada grupo, que debe ser minimizada para conseguir máxima compacidad.
- *Conectividad.* Permite evaluar qué buena es una partición dada según el concepto de conexión, es decir, hasta qué punto una partición observa la densidad de elementos locales y de los grupos de datos, conjunto con sus vecinos más cercanos en el espacio de datos.
- *Separación.* Se basa en medidas que cuantifican el grado de separación dentro de cada agrupamiento. Es decir, cómo de espaciados están unos clústeres de otros. Existen tres métodos comunes para medir la distancia entre dos clases diferentes: el primero se denomina *vinculación individual*, que mide la distancia entre los miembros más cercanos de los grupos; el segundo *ligamiento completo*, que mide la distancia entre los miembros más distantes; finalmente, el tercero *comparación de los centroides*, que mide la distancia entre los centros de los grupos.
- *Combinaciones.* La literatura proporciona diversos enfoques de mejora que combinan algunas de las medidas antes enumeradas. Técnicas que calculan una

puntuación final tomando como referencia la combinación, lineal o no lineal, de algunas de las anteriores medidas.

Existen diferentes índices para la validación interna (Bezdek y Pal, 1998; Bouguessa y col., 2006; Chou y col., 2004; Günter y Bunke, 2003; Milligan y Cooper, 1985; Pal y Biswas, 1997; Wu y col., 2009). Como ya se ha indicado, se basan en los conceptos de cohesión de los grupos y la separación entre agrupamientos. Tal y como ocurre en los métodos de agrupamiento, en general tampoco existe una base teórica robusta sobre qué método es el más apropiado (Bouguessa y col., 2006), todo depende de la aplicación concreta sobre la que se desee realizar la medida de validación. Por otra parte, a pesar del gran número de índices propuestos no existen comparativas extensas y completas de ellas. Probablemente la comparativa más importante sea la realizada por Milligan y Cooper (1985) hace ya casi tres décadas.

Generalmente, los índices internos se utilizan para la búsqueda del número óptimo de clases (Azuaje, 2002; Bel y col., 2005; Dimitriadou, 2002; Dudoity Fridlyand, 2002; Guidi y col., 2009).

Para evaluar la precisión del clasificador en cuanto al número de agrupamientos generados, nos inclinamos por aplicar una combinación de la primera (compacidad) y la tercera (separación) características, dado que las dos clases de medidas muestran tendencias opuestas en el sentido de que mientras que dentro de un clúster mejora la homogeneidad con un número creciente de grupos, la distancia entre los grupos tiende a deteriorarse. Los siguientes índices se fundamentan en este principio de compacidad y separación: Davies y Bouldin (1979) (DB), Dunn (1974) (Dunn), Calinski y Harabasz (1974) (CH), Krzanowski y Lai (1985) (KL), Hartigan (1985) (Han).

Los índices CH y DB, de acuerdo con el estudio de Milligan y Cooper (1985), se sitúan dentro de los diez mejores índices de validación respecto del número de agrupamientos, quedando en primer lugar CH y en décimo DB. Más recientemente Maulik y Bandyopadhyay (2002) seleccionaron los índices de CH y DB entre otros para evaluar el número óptimo de agrupamientos en cinco conjuntos de datos así como la bondad de un índice de validación recientemente desarrollado, obteniendo en la mayoría de los casos el mismo número óptimo entre CH y el algoritmo en cuestión por lo cual lo declaran confiable. Recientemente Bel y col. (2005) para determinar el número de

grupos estables comparan CH y KL, frente al índice de estabilidad de Bertrand y Bel (2005) para dos conjuntos de datos, uno de datos artificiales y el conocido conjunto de datos Iris (UCI, 2012), concluyendo que para el conjunto de datos Iris, KL resulta ser el mejor.

Los índices de validación que permiten obtener el número óptimo de agrupamientos son utilizados para aquellos clasificadores no supervisados en los cuales uno de sus parámetros de entrada es el número de grupos a extraer. Al no tener conocimiento del número de grupos, estos índices se ejecutan después del algoritmo de agrupamiento varias veces, variando el número de grupos en cada ejecución desde un cierto mínimo a un valor máximo, luego se calcula el valor del índice, y finalmente se obtiene la agrupación que proporciona el mejor valor del índice. Esta tarea requiere mucho tiempo de cómputo y cuando se trata de imágenes grandes (teledetectadas), el proceso resulta muy costoso.

En el presente caso de estudio, el interés se centra en medir la calidad en cuanto al número de agrupamientos que genera el método propuesto. Dado que el método no necesita ningún parámetro, es necesario evaluar si el número de agrupamientos resultantes es el óptimo en comparación con otro modelo que genera resultados similares, pero donde uno de los parámetros de entrada es el número de agrupamientos, como es el caso de *Fuzzy clustering*.

A continuación se describen los índices de validación utilizados para evaluar el método propuesto en este trabajo de investigación.

Davies Bouldin (1979) (DB), ampliamente utilizado en la literatura analizada. Usa como medida de similitud de una clase el promedio de las distancias de sus puntos a su centroide, mientras que como medida de separabilidad entre clases utiliza la distancia entre los grupos, que en nuestro caso sería la distancia Euclídea entre los centros de las clases. Bolshakova y Azuaje (2003) validaron la clasificación mediante los índices de DB y Dunn aplicando las medidas de distancia Euclídea, Manhattan y Chebychev para calcular el número óptimo de agrupamientos, resultando que en la mayoría de los casos estudiados, las tres medidas coinciden en cuanto al número de agrupamientos sugerido. Según lo cual, en nuestro caso se ha elegido la Euclídea por su simplicidad.

El índice DB se define de la siguiente manera:

Sea $U \leftrightarrow X: X_1 \cup \dots \cup X_c$, donde X_i representa el i -ésimo grupo de la partición, y $d(X_i, X_j)$ se define como la distancia entre X_i y X_j (distancia entre dos grupos); $\Delta(X_i)$, $\Delta(X_j)$, representan la distancia dentro de los grupos i y j , es decir, $\Delta(X_i)$ es la medida de dispersión dentro del grupo X_i . Por último, c es el número de clases de la partición U .

$$DB(U) = \frac{1}{c} \sum_{i=1}^c \max_{i \neq j} \left\{ \frac{\Delta(X_i) + \Delta(X_j)}{d(X_i, X_j)} \right\} \quad (2-27)$$

Para analizar la ecuación (2-27) con más detalle, la desglosamos de la siguiente manera

$$R_{ij} = \left\{ \frac{\Delta(X_i) + \Delta(X_j)}{d(X_i, X_j)} \right\} \quad (2-28)$$

$$R_i = \max_{i=1, \dots, c; i \neq j} R_{ij} \quad ; \quad j = 1, \dots, m \quad (2-29)$$

por tanto:

$$DB(U) = \frac{1}{c} \sum_{i=1}^c R_i \quad (2-30)$$

De manera que la *distancia* (d_{ij}) entre los grupos X_i y X_j , en un espacio l -dimensional se define como:

$$d_{ij} = d(X_i, X_j) = \|\bar{x}_i - \bar{x}_j\| = \sqrt{\sum_{k=1}^l |\bar{x}_{ik} - \bar{x}_{jk}|^2} \quad (2-31)$$

La *dispersión* (s_i) de un grupo X_i es definida como,

$$s_i = \Delta(X_i) = \sqrt{\frac{1}{n_i} \sum_{x \in X_i} \|x - \bar{x}_i\|^2} \quad (2-32)$$

donde n_i es el número de elementos en X_i .

El *índice de similitud* R_{ij} entre dos grupos X_i y X_j debe satisfacer las siguientes condiciones (Davies y Bouldin, 1979):

1. $R_{ij} \geq 0$
2. $R_{ij} = R_{ji}$
3. Si $s_i=0$ y $s_j = 0$ entonces $R_{ij} = 0$
4. Si $s_j > s_k$ y $d_{ij} = d_{ik}$ entonces $R_{ij} > R_{ik}$
5. Si $s_j = s_k$ y $d_{ij} < d_{ik}$ entonces $R_{ij} > R_{ik}$.

Estas condiciones establecen que R_{ij} es no negativo y simétrico.

Para escoger el número de agrupamientos adecuado se toma el valor del número de grupos (c) que **minimiza** el *índice de similitud* de *Davies Bouldin*, lo que significa que los grupos son más compactos y están más separados unos de los otros.

El índice DB resulta ser muy utilizado para encontrar el número óptimo de clases; entre los trabajos que lo referencian podemos citar los siguientes: Chou y col. (2004); Gunter y Bunke (2003); Halkidi y col. (2001; 2002b); Kovács y col. (2005); Speer y col. (2005); Saitta y col. (2007).

El índice de Dunn (Dunn, 1974), al igual que DB, este índice estima cuán compactos y separados están los clústeres, se define como:

$$D(U) = \min_{1 \leq i \leq c} \left\{ \min_{\substack{1 \leq j \leq c \\ j \neq i}} \left\{ \frac{d(X_i, X_j)}{\max_{1 \leq k \leq c} \{\Delta(X_k)\}} \right\} \right\} \tag{2-33}$$

donde $\Delta(X_k)$ representa como en el caso anterior la varianza dentro del grupo X_k , y $d(X_i, X_j)$ se define como la distancia entre los grupos X_i y X_j ; c es el número de clases. El objetivo principal de esta medida es aumentar al máximo las distancias entre los grupos, minimizando las distancias dentro de cada grupo.

Para escoger el número de agrupamientos adecuado se toma el valor de c que *maximice* el índice de Dunn, porque eso significa que los grupos son más compactos y están más separados entre sí.

Este índice tiene como inconveniente que considera grupos compactos y bien separados, por lo que al calcular la compacidad según la distancia de los puntos a los centros, no detecta bien la forma de los grupos; si además éstos están solapados, no es posible determinar las fronteras entre los grupos, proporcionando resultados erróneos acerca del número de grupos.

Azuaje (2002) aplicó el índice de Dunn con seis diferentes medidas de separación entre los agrupamientos y tres medidas de compacidad dentro de los clústeres.

El índice de Calinski y Harabasz (1974) (CH) se define como:

$$CH(k) = \frac{trB_k / (k-1)}{trW_k / (n-k)} \quad (2-34)$$

En este caso, para cada número de agrupaciones $k \geq 2$. Donde k denota el número de grupos, n es el tamaño de la partición, B_k es la suma total de las distancias entre clases al cuadrado, W_k es la suma de las distancias entre elementos de un mismo clúster al cuadrado, también conocida como distancia inter-clúster. Como medida de distancia, una de las más utilizadas es la Euclídea; tr denota la traza de una matriz.

Por lo tanto, un número óptimo de agrupaciones se define como un valor de k que *maximiza* $CH(k)$.

El índice CH ha sido utilizado para encontrar el número óptimo de agrupamientos, comparándolo con otros índices, y en la mayoría de los casos se obtienen relativamente buenos resultados. Entre los diferentes trabajos que lo aplican se encuentran entre otros los siguientes: Ben-Hur y col. (2002), Bel y col. (2005), que analizaron satisfactoriamente su estabilidad; Casillas y col. (2003); Guidi y col. (2009); Lam y Yan (2007); Halkidi y col. (2002b), Dudoit y Fridlyand (2002), Dimitriadou y col. (2002) y Milligan y Cooper (1985).

El índice de Krzanowski y Lai (1985) (KL) se define como:

$$KL(k) = \left| \frac{diff_k}{diff_{k+1}} \right| \quad (2-35)$$

$$diff_k = (k-1)^{2/p} W_{k-1} - k^{2/p} W_k \quad (2-36)$$

donde p denota el número de características en el conjunto de datos, W_k es el definido previamente. Por lo tanto, un valor de $k \geq 2$ es óptimo cuando se *maximiza* $KL(k)$.

El índice KL al igual que los anteriores ha sido aplicado para encontrar el número óptimo de agrupamientos, entre los diferentes trabajos que lo aplican se encuentran entre otros el de Dudoit y Fridlyand (2002).

Índice de Hartigan (1985), se define por:

$$Han(k) = \left(\frac{trW_k}{trW_{k+1}} - 1 \right) (n - k - 1) \quad (2-37)$$

donde $k \geq 1$, denota el número de grupos, tr y W_k son las definidas previamente.

Por lo tanto, un valor de k es óptimo cuando éste *minimiza* $Han(k)$.

El índice de Hartigan ha sido aplicado para encontrar el número óptimo de agrupamientos, por diferentes autores tales como Dudoit y Fridlyand (2002), y en algunos casos como en Dimitriadou y col. (2002) y Milligan y Cooper (1985) se aplica el índice Hartigan (1975), definido como $Han(k) = \log(B_k / W_k)$.

2.6.3. INFORMACIÓN DE REFERENCIA PARA LA VALIDACIÓN TEMÁTICA

Tal y como se mencionó previamente, los índices de validación externa se reducen a la comparación de dos particiones: la generada por el clasificador y la denominada *ground truth* o de referencia.

De manera que al tratarse de clasificación de cobertura terrestre, ya sea mediante fotografías aéreas digitales (teledetectadas) o fotografías convencionales tomadas desde tierra, la validación consiste en medir la correspondencia entre la clase temática asignada a un píxel por el propio clasificador y la "*verdadera clase*" a la que pertenece, determinada por el experto sobre el terreno (Congalton, 2009).

United States Geological Survey (USGS, 1990) define la exactitud de los datos espaciales como: "la similitud de los resultados de observaciones, cálculos o estimaciones a los verdaderos valores o a los valores aceptados como verdad".

Pero la información “*ground truth*” no siempre se puede conseguir, por lo que es necesario en ocasiones generarla con la ayuda de algún experto conocedor de la zona de estudio, y siempre considerando los diversos factores esenciales en el diseño del control de calidad de la clasificación, tales como: la *unidad de muestreo*, el *tamaño de la muestra* y el *diseño del muestreo* (Congalton, 2010).

Por tanto, en esta sección se describen los factores aplicados para el diseño de la información de referencia “*ground truth*”, que constituye un elemento esencial en la validación de los resultados obtenidos por el clasificador.

Unidad de muestreo

Se entiende como *unidad de muestreo* los fragmentos de la imagen clasificada que serán seleccionados para evaluar la calidad. Congalton (2009) cita cuatro opciones: el píxel simple, las agrupaciones de píxeles, los polígonos o la agrupación de polígonos.

Existen numerosas recomendaciones sobre qué unidad de muestreo tomar en función del tipo de trabajo, propósito, escala, medios e informaciones auxiliares, pues una unidad puede presentar ventajas e inconvenientes frente a otras.

Aronoff (1989) recomienda el uso del píxel individual, argumentando que si incrementamos el nivel de detalle aumentamos también la posibilidad de ocurrencia de errores y también la fiabilidad. Janssen (1994) recomienda el uso del píxel si esta unidad de muestreo es la utilizada en la clasificación, y aconseja el uso de polígonos sólo para aquellos casos en los que existen problemas de accesibilidad en el terreno. Congalton (1986) prefiere un conjunto de píxeles, indicando que es más fácil reconocer esta superficie en los datos de referencia. Dicks y Lo (1990) aconsejan áreas cuando se evalúen cubiertas de terreno.

De acuerdo con lo anteriormente expresado, no existe unanimidad de criterio, en el caso que nos ocupa la clasificación no supervisada se realiza a nivel de píxel tanto para las imágenes tele-detectadas como para las captadas desde tierra. Parece lógico, por tanto, elegir el píxel como unidad de muestreo. Esto es lo que se hace en el caso de las primeras ya que desde el punto de vista de la fotointerpretación, resulta relativamente fácil el análisis del píxel como unidad. Sin embargo, en el caso de las segundas, la verificación de un píxel en campo resulta prácticamente imposible, ya que se analizan hojas de las plantas de avena (secas y verdes), por lo que desde este punto de vista la

identificación del tamaño y forma de las hojas resulta más evidente que el simple análisis por píxeles. Esta es la razón por la cual, en este segundo caso, se utilizan polígonos como unidades de validación.

Tamaño de la muestra

El tamaño de la muestra se refiere al número de sitios de verificación utilizados para estimar la fiabilidad del clasificador. Cuanto más grande sea el tamaño de la muestra, más precisa será la evaluación. Sin embargo, por razones de coste y tiempo computacional, es conveniente determinar el tamaño de muestra mínimo posible para alcanzar los objetivos de la evaluación. No existe un consenso en cuanto al mínimo, Hay (1979) concreta el número de muestras en 50 para cada clase, de igual manera Congalton (1991) recomienda verificar al menos 50 unidades por clase, y de 75 a 100 si el área de estudio es superior a 400.000 ha o si existen más de 12 categorías.

Belar y col. (2001) señalan que existen dos corrientes estadísticas que permiten calcular el número de unidades de muestreo: el modelo binomial aplicado a poblaciones finitas, y la distribución multinomial. El *modelo binomial* distingue entre muestras correctas e incorrectas (Davis, 1986; Aronoff, 1982). Ginevan (1979) y Aronoff (1985) utilizan el modelo binomial para generar tablas de tamaño de muestra y error aceptable; en este supuesto no es necesario conocer el número de clases existentes en la clasificación.

Por otro lado está la corriente avalada por Tortora (1978) y utilizada por Congalton (1999, 2009) para teledetección, que asocia la clasificación de la imagen a una *distribución multinomial*. En estos casos es necesario conocer previamente el número de clases que tendrá la clasificación, así como la proporción de cada una de ellas. Sobre este tipo de distribución, Medina (1998) recopila los trabajos de distintos autores sobre el tamaño óptimo para distribuciones multinomiales, desde Cochran (1953), Quesenberry y Huts (1964), Goodman (1965) a Tortora (1978).

En nuestro caso se aplica la distribución multinomial para seleccionar el número de muestras de las imágenes de referencia, dado que previamente el clasificador proporcionó el número de clases que se están evaluando. Por tanto, a continuación se describe la distribución multinomial desde el punto de vista de la validación de datos.

Distribución Multinomial

Una distribución multinomial es una distribución de probabilidad discreta. Cuando se realiza un muestreo para estimar parámetros de distribuciones multinomiales, el objetivo se debe centrar en calcular intervalos de confianza para cada una de las categorías de la variable. El procedimiento para construir intervalos de confianza simultáneos para distribuciones multinomiales es basarse en una aproximación a la distribución Chi^2 (χ^2).

El procedimiento propuesto por Tortora (1978) parte de dividir una muestra, imagen en este caso, de tamaño n en k categorías exhaustivas y mutuamente excluyentes. Supongamos que $\Pi_i, i = 1 \dots k$, es la proporción de población para la i -ésima clase ($\sum \Pi_i = 1$) y n_i la frecuencia observada en la i -ésima categoría de una muestra aleatoria de tamaño n , entonces para un valor determinado de α , se quiere determinar el conjunto de intervalos $S_i, i = 1, \dots, k$, tal que:

$$\Pr \left\{ \bigcap_{i=1}^k (\Pi_i \in S_i) \right\} \geq 1 - \alpha \quad (2-38)$$

Es decir, se espera que la probabilidad de que todo intervalo S_i contenga al verdadero valor del parámetro Π_i sea menor a $1 - \alpha$.

Si N es el tamaño total de la población y se incorpora el factor de *corrección por población finita* (*cpf*) además de la varianza de cada Π_i , se tiene una aproximación a los intervalos de confianza (Cochran, 1953).

$$\begin{aligned} \Pi_i^- &= \Pi_i - [B(N - n)\Pi_i(1 - \Pi_i) / (N - 1)n]^{1/2} \\ \Pi_i^+ &= \Pi_i + [B(N - n)\Pi_i(1 - \Pi_i) / (N - 1)n]^{1/2} \end{aligned} \quad (2-39)$$

donde B es el percentil superior correspondiente al $(\alpha / k) \times 100$ de una distribución Chi cuadrado (X^2), con un grado de libertad ($\chi^2(1, \alpha/k)$).

Para la determinación del tamaño de muestra se necesita definir la precisión de cada clase o parámetro de la distribución multinomial. De esta manera, si se desea una precisión absoluta b_i para cada celda; entonces se tiene que:

$$\begin{aligned}\Pi_i - b_i &= \Pi_i - [B\Pi_i(1 - \Pi_i) / n]^{1/2} \\ \Pi_i + b_i &= \Pi_i + [B\Pi_i(1 - \Pi_i) / n]^{1/2}\end{aligned}\tag{2-40}$$

Despejando el valor de b_i (precisión absoluta de la muestra):

$$b_i = [B\Pi_i(1 - \Pi_i) / n]^{1/2}\tag{2-41}$$

de donde se obtiene, para el caso más desfavorable, que el tamaño de muestra necesario para estimar cada celda con una precisión b_i es $n = \text{máx}_i \{B\Pi_i(1 - \Pi_i) / b_i^2\}$

En caso de que se desee incorporar la corrección por finitud (cpf) se convierte en:

$$n = \text{máx} \left\{ BN\Pi_i(1 - \Pi_i) / \left[b_i^2(N - 1) + B\Pi_i(1 - \Pi_i) \right] \right\}\tag{2-42}$$

Como en la práctica se desconoce el valor aproximado de la proporción Π_i que se desea estimar, se supone el caso más desfavorable: $\Pi_i = 1/2$ y $b_i = b$ para $i = 1, \dots, k$.

Por lo tanto, sólo se debe hacer el cálculo de k para determinar el tamaño de la muestra, uno para cada par (b_i, Π_i) , desde $i=1 \dots k$, y seleccionar el máximo n como el tamaño de muestra deseado.

Esquema de muestreo

En el diseño de la distribución de las muestras se contemplan varias propuestas que van desde el básico diseño *aleatorio simple*, muy extendido por su facilidad de aplicación estadística, a establecer puntos de control de forma sistemática, existiendo múltiples opciones entre ambos. Dada la importancia de este aspecto, también existen numerosas recomendaciones, por ejemplo, Janssen (1994) habla del uso del diseño aleatorio simple, siempre que haya transcurrido tiempo entre la toma de datos de la imagen y la clasificación, mientras que Congalton (1999, 2009) recomienda utilizar el aleatorio

simple antes de la clasificación y el aleatorio estratificado una vez realizada la clasificación.

Principales métodos de muestreo

Los tipos de muestreo de campo más ampliamente difundidos en la actualidad son: aleatorio simple, aleatorio estratificado, sistemático, sistemático no alineado y por conglomerados (Congalton, 1988; Rosenfeld, 1982), en la figura 2.6 se muestran ejemplos aclaratorios respecto de los tipos de muestreo indicados.

En el *Aleatorio Simple*, los elementos a verificar se eligen de tal forma que todos cuenten con la misma probabilidad de ser seleccionados, y que la elección de uno no influya en la del siguiente. Gracias a este carácter probabilístico, este muestreo es el más sólido estadísticamente hablando, aunque sus desventajas radican en los elevados costes computacionales debidos al desplazamiento y a la existencia de cierta probabilidad de no recoger adecuadamente la variación espacial presente en la imagen clasificada.

En el *Aleatorio Estratificado*, la muestra se toma dividiendo la población en regiones o estratos, de acuerdo a una variable auxiliar, por ejemplo, grado de afectación en caso de cultivos. Su ventaja fundamental radica en que aporta información sobre subconjuntos de la población y en la posibilidad de reducir el error de muestreo si la variable auxiliar se elige correctamente, mientras que su desventaja consiste en que resulta más difícil de diseñar que el anterior.

En el muestreo *Sistemático* las muestras se distribuyen a intervalos regulares, a partir de un centro de origen señalado al azar. Su ventaja principal radica en garantizar un muestreo sistemático íntegro de la superficie a verificar. Sin embargo, esto acarrea serias limitaciones al permitir estimaciones erróneas si existiera algún tipo de patrón periódico en el área observada, como puede ser la geometría de las parcelas en los campos de cultivo. Por otra parte, no garantiza una estimación probabilística ya que no existe aleatoriedad en la selección de las muestras.

El *Sistemático no Alineado* se puede considerar como una variante del anterior, producto de la variación de una de las dos coordenadas espaciales de manera aleatoria. Esto permite introducir el factor aleatoriedad, a la vez que reduce el error debido a la periodicidad, todo ello manteniendo la revisión completa de todo el terreno.

El muestreo *Por Conglomerados*, como su nombre indica, consiste en seleccionar, en un punto escogido aleatoriamente dentro del terreno, un grupo de observaciones en lugar de muestras independientes, de acuerdo a un esquema diseñado previamente. El hecho de contar con grupos de muestras tiene la ventaja de reducir los gastos de transporte, pero es complejo en su realización y menos preciso que los anteriores.

En la investigación propuesta en este trabajo, la toma de muestras captadas por cámaras convencionales, para la cuantificación de la densidad de planta afectada y no afectada por heladas, se desarrolló mediante un esquema de muestreo aleatorio estratificado (figura 2.5 *b*).

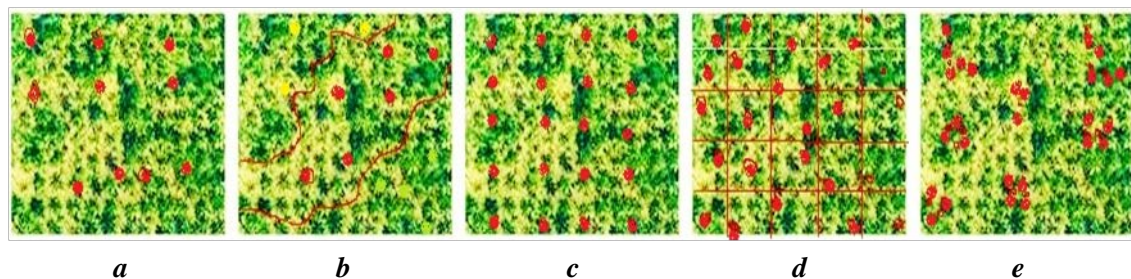


Figura 2.5. Tipos de muestreo de campo: (a) aleatorio, (b) estratificado, (c) sistemático, (d) sistemático no alineado y (e) por conglomerados.

2.6.4. MATRICES DE ERROR

Entre los métodos más comunes para evaluar la precisión de una clasificación, el más extendido en teledetección es el de la *matriz de errores*, que consiste en la representación numérica bidimensional de las muestras utilizadas como evaluación, con tantas filas y columnas como clases tiene la clasificación bajo análisis. Las filas representan las clases asignadas en el proceso de clasificación, y las columnas las clases reales obtenidas a partir de la observación sobre el terreno o bien mediante muestreo de campo y sobre las que se tiene la certeza del tipo de cubierta que representan, cuya información se denomina *datos de referencia* o *ground truth* (Congalton y Geen, 2009). Se asumirá que se han identificado puntos de control en el campo o a través de una fuente de mayor exactitud, y que se comparará el atributo para cada punto con el correspondiente. En la intersección de la fila I con la columna J se contabiliza el número de casos en los que en el clasificador tenían asignada la clase I, y en el campo se

encontró la clase J. La diagonal principal de la matriz indica el número de acuerdos entre los datos de referencia y los arrojados por el clasificador, esto es muestras correctamente clasificadas.

La matriz de error resulta de gran utilidad para la visualización de resultados con imágenes de clasificación y, quizás lo más importante, para medir los resultados estadísticamente. De acuerdo con Congalton (2010), una matriz de error es la única forma de comparar cuantitativamente los resultados de dos clasificadores distintos, pudiéndose construir bajo una concepción rígida o suave. La *matriz de error tradicional* o también denominada *determinista* se construye bajo una concepción rígida, es decir, para su construcción se asume que dada una discrepancia entre los resultados del clasificador y lo que se observa en el campo sólo hay una opción binaria de “correcto” o “incorrecto”, que sería sinónimo de “coincide” y “no coincide”, y por tanto tiene sólo una etiqueta.

La *matriz de error difusa*, también denominada *fuzzy*, se construye bajo una concepción suave, es decir, permite alternativas intermedias para los casos en los cuales se tienen distintos grados de pertenencia, sobre todo para los píxeles mixtos o de frontera (Gopal y Woodcock, 1994; Congalton y Green, 1999; Congalton, 2009). Por tanto, bajo esta perspectiva se debe establecer un conjunto de reglas difusas, particularmente en el caso de que existan discrepancias entre los resultados del clasificador y lo que se observa en el campo.

Matriz de error convencional y su representación matemática

Una matriz de error convencional es una matriz cuadrada de números ordenados en filas y columnas que expresa el número de unidades de muestreo asignados a una categoría particular, resultado del clasificador bajo análisis, en relación con el número de unidades de muestra asignados a esa categoría por el experto (datos de referencia, asumiendo que son datos clasificados con mayor exactitud). Una de las características de la matriz de error convencional es que cada unidad de validación sólo puede ser marcada por una etiqueta. Es decir, si hay discrepancia entre lo que representa el dato y lo que se observa en el campo sólo hay una opción binaria de “correcto” o “incorrecto”, por lo que cada unidad de validación se acumulará en la matriz de error dentro de los correctos (diagonal principal) o en los incorrectos (fuera de la diagonal principal).

A partir de la matriz de error pueden deducirse varios índices. En primer lugar la *fiabilidad global*, que se obtiene dividiendo el total de píxeles correctamente clasificados (suma de los elementos de la diagonal principal) entre el número total de píxeles en la muestra de evaluación.

Además, pueden estudiarse las precisiones individuales de cada clase mediante dos índices complementarios. El primero de ellos se calcula dividiendo el número de píxeles correctos de una clase entre el número total de píxeles de esa clase obtenido a partir de los datos de referencia (el total de la columna correspondiente). Este valor expresa la proporción de píxeles pertenecientes a una clase que han sido correctamente clasificados, dando una idea de los errores por defecto. A este índice se le conoce como *fiabilidad del experto*, también identificado en la literatura como *fiabilidad del productor*, ya que a éste le interesará saber los aciertos en la clasificación de una determinada zona en el terreno.

Por otra parte, dividiendo el número total de píxeles de una clase correctamente clasificados entre el total de píxeles que fueron clasificados dentro de esa clase (la suma de los valores de la fila correspondiente) se obtiene información sobre los errores por exceso. A este índice se le denomina *fiabilidad del clasificador* o *fiabilidad del usuario*, y expresa la probabilidad de que un píxel clasificado dentro de una clase pertenezca realmente a dicha clase. En última instancia, al usuario de los datos le interesará saber la veracidad de la información contenida en la imagen clasificada.

Representación matemática de la matriz de error

Supongamos que n muestras son distribuidas en k^2 celdas, donde cada muestra se asigna a una de las k categorías de la clasificación temática a evaluar (por lo general, las filas), y de forma independiente a una de las mismas k categorías en el conjunto de datos de referencia o *ground truth* (columnas), donde n_{ij} denota el número de muestras clasificadas en la categoría i ($i = 1, 2, \dots, k$) en la imagen clasificada por nuestro clasificador y la categoría j ($j = 1, 2, \dots, k$) en el conjunto de datos de referencia (tabla 2.1).

Tabla 2.1. Ejemplo matemático de una matriz de error

		REFERENCIA (j)			TR
		1	2	k	n _{+j}
CLASIFICACION (i)	1	n ₁₁	n ₁₂	n _{1k}	n ₁₊
	2	n ₂₁	n ₂₂	n _{2k}	n ₂₊
	k	n _{k1}	n _{k2}	n _{kk}	n _{k+}
TC	n _{+j}	n ₊₁	n ₊₂	n _{+K}	n

TC: Total de columnas

TR: Total de renglones

j: Columnas

i : Filas

donde

$$n_{i+} = \sum_{j=1}^k n_{ij} \tag{2-44}$$

representa el número de muestras clasificadas en la categoría *i* por el modelo de clasificación, mientras que el número de muestras clasificadas en la categoría *j* del conjunto de datos de referencia viene dado por la ecuación (2-45).

$$n_{+j} = \sum_{i=1}^k n_{ij} \tag{2-45}$$

Así pues, la fiabilidad global o precisión global (FG), de los resultados del clasificador y los datos de referencia, se puede calcular de la siguiente manera,

$$FG = \frac{\sum_{i=1}^k n_{ii}}{n} \tag{2-46}$$

La fiabilidad del productor o experto se obtiene como sigue:

$$j = \frac{n_{ij}}{n_{+j}} \tag{2-47}$$

La fiabilidad de usuario o del clasificador se obtiene mediante,

$$i = \frac{n_{ii}}{n_{i+}} \tag{2-48}$$

Por último, sea p_{ij} el porcentaje de muestras en la i,j -ésima celda, correspondiente a n_{ij} , esto es $p_{ij} = n_{ij}/n$. Entonces p_{i+} y p_{+j} se definen mediante las dos ecuaciones siguientes,

$$p_{i+} = \sum_{j=1}^k p_{ij} \tag{2-49}$$

$$p_{+j} = \sum_{i=1}^k p_{ij} \tag{2-50}$$

Otro de los índices utilizados para calcular la fiabilidad de una clasificación es el denominado *coeficiente kappa*, \hat{k} que se define a partir de la matriz de error como sigue,

$$\hat{k} = \frac{n \cdot \sum_{i=1}^k p_{ii} - \sum_{j=1}^k p_{i+} \cdot p_{+j}}{n^2 - \sum_{i=1}^k p_{i+} \cdot p_{+j}} \tag{2-51}$$

Donde k representa el número de filas de la matriz, p_{ij} el número de observaciones en la fila i y la columna j , p_{i+} y p_{+i} los valores totales marginales de la fila i y de la columna j , respectivamente, y n es el número total de observaciones. La $\sum p_{ii}$ representa las muestras correctamente clasificadas y la

$$\sum p_{i+} p_{+j} = (n_{+1} \times n_{1+}) + (n_{+2} \times n_{2+}) + (n_{+k} \times n_{k+}).$$

El *coeficiente kappa* expresa la proporción en la reducción del error al aplicar la clasificación respecto al error que se hubiera producido realizando una asignación de clases completamente aleatoria. A diferencia de *FG*, este coeficiente lleva implícita información sobre los elementos marginales, fuera de la diagonal principal, de la matriz. Además recoge en un sólo valor información sobre el proceso y permite la comparación directa de varias matrices.

En resumen, mediante la construcción de la matriz de error antes expuesta, es posible identificar la probabilidad de que un clasificador no supervisado reconozca un objeto presente en la imagen como lo hizo el experto, así como la probabilidad de que el clasificador cometa errores en la clasificación, tanto para cada objeto o clase como de manera general. Sin embargo, este tipo de evaluación puede dar resultados en ocasiones demasiado estrictos. Esto es debido al hecho de que se limita a calificar si el objeto fue reconocido por el clasificador dentro del mismo tema ubicado por el experto, calificándolo como correcto o incorrecto, no existiendo posibilidad de determinar que es posible que se encuentren las dos, si bien con distintos grados de pertenencia.

Matriz de error *fuzzy*

La teoría de conjuntos borrosos resulta muy útil en aquellas situaciones en que los datos y sus relaciones no pueden describirse en términos matemáticos precisos (Santos, 2011). Esta rama de la lógica ha tenido un fuerte desarrollo en los últimos años, fundamentalmente por su capacidad para manejar problemas en los cuales existe una ambigüedad esencial. Como en el caso de la cobertura terrestre en la cual se mezclan varias clases (por ejemplo, arbustos y pasto, agua y arena), donde puede ocurrir que el clasificador identifique todas esas mezclas con distintos grados de pertenencia. Por tanto, es mejor emplear una aproximación que combine la matriz de error y alguna medida de borrosidad, es decir, se incorporan junto a los valores de bien o mal clasificado, alternativas intermedias para dar cabida a los píxeles mixtos o frontera (Gopal y Woodcock, 1994; Congalton y Green 1999, Congalton 2009, 2010), así como para aquellos píxeles en los cuales la variabilidad es a menudo difícil de controlar. El uso de la *matriz de error fuzzy* es pues una herramienta muy potente en el proceso de evaluación de la exactitud. Detalles sobre este enfoque relativo a la matriz de error difuso se pueden encontrar en Green y Congalton (2004).

La matriz de error *fuzzy* es, al igual que la matriz de error tradicional, una matriz cuadrada, con tantas filas y columnas como clases posee la partición derivada de la clasificación bajo análisis. Como en el caso anterior, se asume que se han identificado puntos de control en el campo (o en una fuente de mayor exactitud) y que se comparará el atributo para cada punto con el correspondiente en los datos bajo análisis. En la

intersección de la fila i con la columna j se contabiliza el número de casos en los que en el sistema el punto de control tenía asignada la clase i , y en el campo se encontró la clase j .

La diferencia con la matriz de error tradicional estriba en la manera de contabilizar y representar los resultados en la matriz de acuerdo con la reglas *fuzzy*. Los elementos fuera de la diagonal principal de la matriz contienen dos valores, tabla 2.2. El primer valor de las celdas fuera de la diagonal principal representa las etiquetas que, aunque no son absolutamente correctas, se consideran aceptables según el criterio establecido mediante alguna regla difusa. El segundo valor indica las etiquetas que son inaceptables, es decir, se clasificaron de manera equivocada en dicha clase. La diagonal principal contiene sólo un valor y es aquél que se considera absolutamente correcto.

Por lo tanto, las celdas fuera de la diagonal principal de la matriz contienen dos números, que pueden ser utilizados para distinguir las etiquetas inciertas, que más que erróneas, probablemente se encuentran en los márgenes de una clase y que con base a las reglas *fuzzy* establecidas se pueden contabilizar como aceptables o erróneas.

Representación matemática de la matriz de error fuzzy

Tabla 2.2. Ejemplo matemático de una matriz de error *fuzzy*

		REFERENCIA (j)						TR		
		1		2		...		k	n_{i+}	
CLASIFICACION (i)	1	n_{11}		na_{12}	ne_{12}	na_{1k}	ne_{1k}	n_{1+}
	2	na_{21}	ne_{21}	n_{22}		na_{2k}	ne_{2k}	n_{2+}
	⋮	⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮
	k	na_{k1}	ne_{k1}	na_{k2}	ne_{k2}	n_{kk}		n_{k+}
TC	n_{+j}	n_{+1}		n_{+2}		...		n_{+k}		n

TC: Total de columnas

n : número total de muestras

TR: Total de renglones

na : número de muestras aceptable

j : Columnas; **i :** Filas

ne : número de muestras erróneas

Las ecuaciones (2-52) y (2-53) permiten calcular el número de muestras clasificadas en la categoría i por el modelo de clasificación, consideradas como aceptables o erróneas respectivamente, dentro de la calificación lingüística difusa.

$$na_{i+} = \sum_{j=1}^k (na_{ij}) \quad (2-52)$$

$$ne_{i+} = \sum_{j=1}^k (ne_{ij}) \quad (2-53)$$

La ecuación (2-54) representa el total de muestras clasificadas en la categoría i por el modelo de clasificación.

$$n_{i+} = \sum_{j=1}^k (n_{ii} + na_{ij} + ne_{ij}) \quad (2-54)$$

La ecuación (2-55) representa el número de muestras clasificadas en la categoría j en el conjunto de datos de referencia.

$$n_{+j} = \sum_{i=1}^k (n_{ij} + na_{ij} + ne_{ij}) \quad (2-55)$$

Así pues, la fiabilidad global *fuzzy* o precisión global *fuzzy* (PGF), establecida entre los resultados del clasificador y los datos de referencia, se puede calcular de la siguiente manera,

$$PGF = \frac{\sum_{i=1}^k (n_{ii} + na_{ij})}{n} \quad (2-56)$$

PGF representa, pues, la probabilidad general de que la información esté bien clasificada (Congalton, 2009; Ariza, 1996).

La fiabilidad del experto o precisión del productor, datos *ground truth*, puede obtenerse mediante la siguiente expresión,

$$j = \frac{n_{jj} + na_{+j}}{n_{+j}} \quad (2-57)$$

La fiabilidad del clasificador o precisión del usuario puede ser calculada por,

$$i = \frac{n_{ii} + na_{i+}}{n_{i+}} \quad (2-58)$$

Así pues, la diferencia entre el 100% y la precisión del experto proporciona los errores de exclusión o de omisión, es decir, $100 - j$, y la diferencia entre el 100% y la precisión del clasificador proporciona los errores de inclusión o de comisión, es decir, $ei = 100 - i$.

Por último, como en el caso de la matriz de error no difusa, sea p_{ij} el porcentaje de muestras en la i,j -ésima celda, correspondiente a n_{ij} , esto es $p_{ij} = n_{ij}/n$, entonces p_{i+} y p_{+j} vienen definidos por las expresiones siguientes,

$$p_{i+} = \sum_{j=1}^k p_{ij} \quad (2-59)$$

$$p_{+j} = \sum_{i=1}^k p_{ij} \quad (2-60)$$

En el presente trabajo de investigación se han utilizado ambas matrices de error para evaluar el clasificador propuesto.

2.7. RESUMEN DEL ANÁLISIS BIBLIOGRÁFICO

El objetivo principal de un algoritmo de clasificación, en nuestro caso no supervisado, es conseguir separar los diferentes elementos objeto de la clasificación, presentes en la imagen, de manera que dentro de cada grupo sean lo más homogéneos posible y a la vez las clases estén bien separadas entre sí.

Para conseguirlo, en este trabajo y también de forma general, se plantean tres etapas. La etapa de pre-procesamiento, en la cual se prepara la información necesaria para la clasificación, transformándola al modelo de color necesario, normalizando los datos o realizando algunas correcciones encaminadas a la mejora del contraste de la imagen. La segunda etapa es la de segmentación, cuyo resultado podría ser un conjunto de regiones homogéneas, probablemente superpuestas que se tendrán que evaluar, reagrupar y reconocer durante la tercera etapa que es la clasificación propiamente dicha.

Cuando se trata de clasificar imágenes en color, los métodos de segmentación de nivel de gris se pueden aplicar directamente a cada componente del espacio de color con el que se esté trabajando, tras lo cual, los resultados se combinan convenientemente para obtener un resultado final satisfactorio.

En el método desarrollado en este trabajo, la alternativa de procesar las tres componentes de color por separado, en los modelos de color utilizados, resulta conveniente tanto en el ámbito de la teledetección como de las imágenes captadas en tierra. Es más, los planteamientos formulados para las imágenes en color y sus tres componentes espectrales visibles, pueden aplicarse a cualquier otra aplicación que involucre tres bandas espectrales, no necesariamente pertenecientes al espectro visible, o extenderse a un mayor número de bandas espectrales, proporcionadas por un sensor multi-espectral o incluso por varios sensores capaces de capturar datos en varias longitudes de onda dentro de cualquier banda del espectro.

Lu y Weng (2007) desarrollan un análisis de los diferentes algoritmos de clasificación con la finalidad de encontrar técnicas para mejorar el rendimiento en las clasificaciones, en cuyo trabajo mencionan que un sistema de clasificación está diseñado sobre la base de la necesidad del usuario, la resolución espacial de determinados datos obtenidos, la compatibilidad con el trabajo anterior de procesamiento de imágenes, los algoritmos de clasificación disponibles y las limitaciones de tiempo. Mientras que Hu y Damper (2008) afirman que dependiendo de la aplicación, una misma estrategia puede conducir a resultados diferentes.

Bajo los planteamientos anteriores, para realizar la clasificación de cobertura terrestre mediante fotografías en formato digital con sus tres componentes de color RGB, el coste de etiquetarlas por un experto es muy alto, sobre todo por la necesidad de muestrear grandes cantidades de terreno y especialmente cuando el acceso es complicado, por tanto la alternativa más óptima es la clasificación no supervisada. En este capítulo se han estudiado diversas técnicas de esta índole, así como de naturaleza supervisada. En todos los casos bajo la perspectiva de su aprovechamiento en el diseño de la estrategia también no supervisada propuesta en la presente investigación, cuyo fundamento se basa en la umbralización de las bandas espectrales por separado, razón por la cual se han estudiado métodos de segmentación basados en umbral.

La estrategia de clasificación propuesta se evalúa mediante los denominados índices de validación de las particiones, que permiten tanto determinar intrínsecamente la bondad de la misma como su calidad frente al método *Fuzzy C-Means*, elegida como una de las técnicas más ampliamente usadas.

Capítulo 3

Aplicación de estrategias de procesamiento y clasificación a imágenes aéreas de cobertura terrestre





CAPÍTULO 3 APLICACIÓN DE ESTRATEGIAS DE PROCESAMIENTO Y CLASIFICACIÓN A IMÁGENES AÉREAS DE COBERTURA TERRESTRE.

3.1. INTRODUCCIÓN

En este capítulo se describen tanto los materiales utilizados como los diversos procesos de segmentación y clasificación aplicados a las imágenes aéreas bajo análisis. El objetivo principal consiste en desarrollar una metodología que permita clasificar la cobertura terrestre a partir de imágenes captadas por sensores remotos, de forma totalmente automática, sin necesidad de introducir más parámetros de entrada que las tres componentes espectrales en el modelo de color apropiado. La forma de llevarlo a cabo se sintetiza en la figura 3.1.

En la primera fase, denominada *recolección de la información*, se describen las características principales de las imágenes digitales utilizadas, que como ya hemos indicado se trata de imágenes captadas por sensores aéreos, así como la manera en que éstas fueron obtenidas.

La siguiente fase comprende el *pre-procesado de la imagen original*, en la cual se describe, de manera general, la forma en que se preparan las imágenes que originalmente se encuentran en el modelo de color RGB, para facilitar la segmentación y clasificación de la cobertura terrestre. En ocasiones las imágenes no poseen calidad apropiada para el tratamiento directo y se requiere realizar algunas correcciones tales como: realzar el contraste, normalizar los niveles digitales *ND*, destacar estructuras específicas de la imagen o convertir los datos originales (RGB) a otro modelo de color. Este es el caso del planteamiento formulado en este trabajo, tal y como se describe convenientemente en la correspondiente sección.

Una vez pre-procesada la imagen, se entra en el procedimiento de clasificación propiamente dicho, que consta de tres etapas básicas, a saber: *segmentación*, *codificación* y *decisión*. Estas tres etapas se aplican en forma secuencial en la presente investigación, aunque dependiendo de las características de la imagen pueden aplicarse en paralelo. Es decir, se inicia la segmentación, a continuación se codifica la información extraída de cada componente espectral, posteriormente se fusionan las

componentes similares y, dependiendo del resultado obtenido, se decide si se realiza la codificación final o se regresa al proceso de segmentación, incrementando el número de umbrales hasta superar la condición de clasificación exitosa, que se identifica como que el número de clases alcanza un determinado máximo. De manera general, se puede decir que en realidad se trata de dos procesos básicos: *segmentación* y *decisión*, puesto que la codificación se lleva a cabo en dos fases, primeramente por cada componente espectral de forma separada durante la segmentación y finalmente en la imagen de color con sus tres componentes espectrales.

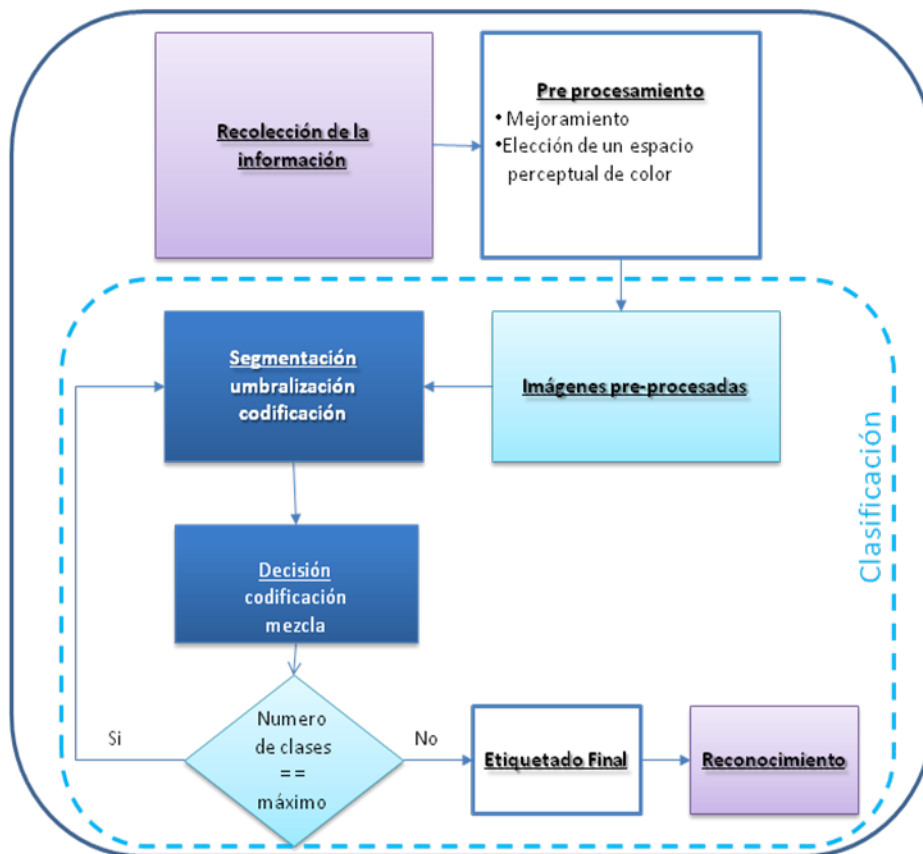


Figura 3.1. Esquema del procesamiento y clasificación de las imágenes.

3.2. RECOPIACIÓN Y PROCESAMIENTO DE LA INFORMACIÓN

Para la presente investigación se usaron 80 fotografías aéreas a color en formato digital, propiedad del Instituto de Geografía de la Universidad Autónoma de México (UNAM), tomadas en octubre de 1997. Las fotografías corresponden a la cuenca hidrográfica del río La Sabana, Guerrero. Dichas imágenes fueron tomadas después de presentarse en 1997 el Huracán de categoría 4, llamado “Paulina”, con el objetivo de diagnosticar las

pérdidas en la agricultura por inundaciones, en concreto en el municipio de Acapulco, Guerrero.

La cuenca hidrográfica del río la Sabana se localiza en la parte Sur del Estado de Guerrero. Pertenece a la Región Hidrológica 19 y al municipio de Acapulco en México. La cuenca tiene una superficie de 46.802,53 ha y un perímetro de 154 km. Sin embargo, la zona correspondiente a las imágenes a clasificar sólo abarca un área situada en la parte baja de la cuenca de aproximadamente 3.973 ha. En la figura 3.2 se muestra la ubicación de la zona analizada sobre el mapa genérico.

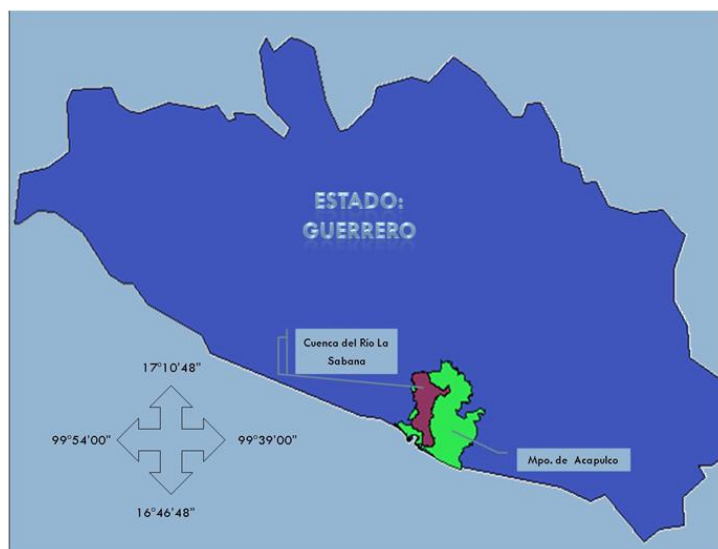


Figura 3.2. Localización de la cuenca del río La Sabana, estado de Guerrero, México.

Por lo tanto, uno de los principales objetivos es la identificación de los cuerpos de agua, así como de las parcelas agrícolas inundadas y el arrastre de los sedimentos sobre las mismas. Otra aplicación que se dio a nivel costero fue la clasificación de cultivos perennes, con la finalidad de clasificar las plantaciones de cocotero, cítricos y mango.

El Instituto de Geografía de la UNAM utilizó un sensor digital Kodak DCS-420c, que proporciona imágenes en color de 1524 x 1012 píxeles, provista de una lente Sigma de 14mm f/2.8, con un campo visual de 35mm, aproximadamente igual al de una lente de 180mm instalada en una cámara aérea con negativo de 23 x 23 cm de ancho y alto. La razón por la que se utiliza una lente de distancia focal más corta es debida a que el tamaño del sensor también es menor, por lo que sólo capta la porción central de la imagen proyectada por la lente. Al disparo de la cámara se asocian, a través de la hora GMT (*Greenwich Mean Time*), los datos fotográficos para cada imagen, tales como

distancia focal, distancia enfocada, velocidad de obturación, diafragma, método de medición de exposición, modo de exposición y factor de compresión como los más relevantes.

Las imágenes se encuentran bajo el sistema de coordenadas UTM (*Universal Transverse Mercator*), zona 14, donde las medidas se establecen en metros.

En una imagen en color en el modelo de color RGB, cada píxel se caracteriza por sus tres componentes del espectro visible. Por tanto, las imágenes cuentan con una *resolución* de tres bandas en dicho espectro visible, dado que la resolución espectral indica el número y anchura de las bandas en las que el sensor puede captar radiación electromagnética. Las componentes espectrales también suelen estar referidas en la literatura como canales o bandas espectrales; por tanto en este trabajo utilizamos los tres términos indistintamente.

La *resolución radiométrica* es de 0 a 255 niveles digitales (*ND*), correspondiéndose con una representación de 8 bits por píxel/canal; la resolución radiométrica se refiere al número de intervalos de intensidad que pueden captarse.

La *resolución espacial* es variable, entre 1:15.000 y 1:20.000. La resolución espacial representa el tamaño del píxel en relación a las dimensiones reales del espacio o escena incluidas en un píxel. En algunos casos se emplea el concepto de IFOV (*Instantaneous Field Of View*) que se define como la sección angular (en radianes) observada en un momento determinado.

Las imágenes utilizadas en nuestros experimentos se han formateado convenientemente según el sistema de coordenadas de la imagen. Para algunos casos se usaron imágenes de tamaño original (1012 × 1524 píxeles), y en otros casos se dividieron en cuatro partes, quedando finalmente, en este último caso, de un tamaño de 506 × 757 píxeles.

Las coordenadas geográficas de las imágenes se han mantenido por su utilidad para la validación del clasificador. En el momento de la clasificación no hay necesidad alguna para su consideración, ya que no intervienen para nada en la clasificación, sin embargo, sí resultan absolutamente necesarias para la validación de los resultados, ya que permiten la asociación y ubicación de las estructuras en la imagen con respecto al terreno.

3.2.1. MEJORA DE LA CALIDAD DE LA IMAGEN

Para el presente estudio se usó principalmente la ecualización por igualación o especificación de histograma para mejorar el contraste de las imágenes, que demandaban esta operación. Para ello se aplicó la metodología planteada por Coltuc y col. (2006).

En la figura 3.3, la cual se ha seleccionado como imagen representativa de las fotografías aéreas que no necesitan ecualización, se observa su histograma (figura 3.3 *b*) que fue tomado como referencia para igualar las imágenes que sí necesitan mejoras, como las mostradas en la figura 3.4, por ejemplo.

De manera que, como primer paso, se transformaron las imágenes a corregir así como la imagen inicial cuyo histograma servirá de muestra, pasando del modelo de color original RGB al modelo de color YIQ (Luminance, In-phase, Quadratic), como se especifica en el capítulo dos. A continuación se extrae la componente *Y* que contiene la luminancia sobre la que se aplica la ecualización de histograma para mejora de la calidad.

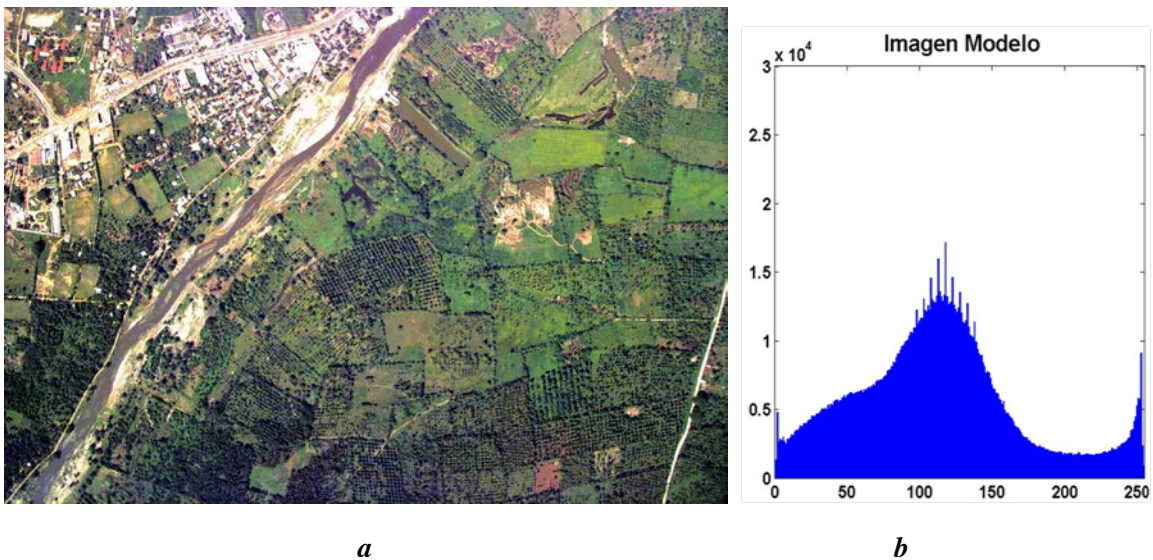


Figura 3.3. Imagen e histograma modelo: **(a)** imagen aérea original, **(b)** histograma funcional de la componente de luminancia.

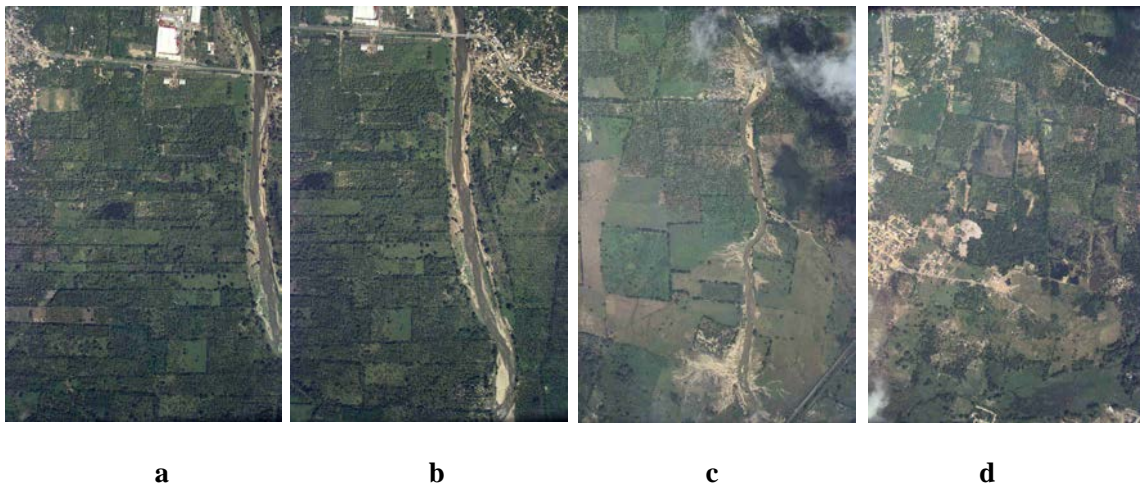


Figura 3.4. Imágenes que necesitan ecualización de luminancia.

Después de obtener la luminancia de cada imagen, los valores son normalizados y se calcula el histograma de la imagen con iluminación funcional.

En la figura 3.5 *b* se puede observar el histograma origen de la componente Y, extraída de la imagen original mostrada en la figura 3.5 *a*. Este histograma será igualado al histograma de la componente Y de la imagen original (figura 3.3 *a*) mostrado en la figura 3.3 *b*, obteniendo como resultado el histograma de la figura 3.5 *d*, correspondiente a la figura 3.5 *c*. Obsérvese en este caso, la mayor similitud de este último histograma en relación al modelo utilizado.

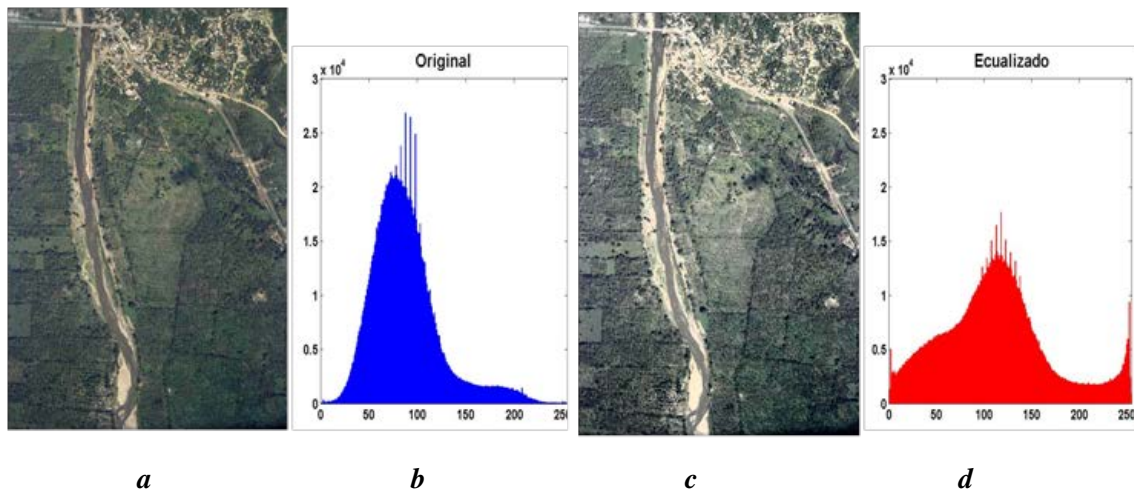


Figura 3.5. Igualación del histograma: (a) imagen origen, (b) histograma origen, (c) imagen destino, (d) histograma ecualizado.

Una vez que las imágenes han sido ecualizadas, se aplica la correspondiente transformada inversa, volviendo de nuevo a su modelo original RGB, para ser procesadas por el algoritmo de clasificación.

3.3. ESTRATEGIA DE CLASIFICACIÓN NO SUPERVISADA

En este apartado se expone la estrategia metodológica del sistema de clasificación de la cobertura terrestre a partir de las imágenes aéreas antes descritas. Al tratarse de imágenes en color, cada píxel se caracteriza por un vector de tres componentes, y el valor de cada componente depende del espacio de color (RGB, CIELAB, YIQ,...) con el que se esté trabajando. Dado que la mayoría de los métodos de segmentación se han diseñado para imágenes en escala de gris, haciendo mención a cada una de las componentes espectrales por separado, se decidió aplicar *segmentación por umbralización* mediante una búsqueda automática de umbrales, considerando las imágenes en color como la composición de tres funciones bidimensionales independientes, es decir, cada componente se procesa por separado.

La estrategia de clasificación aplicada es a nivel de píxel, clasificando cada píxel en la imagen de forma individual teniendo en cuenta sus propiedades espectrales. De forma general consta de cinco pasos:

- *Separar* las tres componentes espectrales.
- *Establecer* un sistema de segmentación de píxeles por umbralización, el cual puede ser simple o múltiple. El objetivo de la umbralización simple es encontrar el *ND* óptimo que permita dividir los píxeles de la imagen en dos grupos (los píxeles con *ND* cercano a cero pertenecerán a un grupo, y los píxeles con nivel alto pertenecerán al segundo). Para la umbralización múltiple se buscan más de dos *ND* óptimos que permitan dividir el conjunto de píxeles en más de dos segmentos.
- *Asociar* a cada píxel de cada componente espectral una etiqueta que permita identificar la posición y el grupo al que pertenece. Así si una componente espectral es segmentada por dos umbrales, los píxeles correspondientes tendrán asignada alguna de las tres etiquetas que identifiquen a qué región o grupo pertenece. Las etiquetas para este caso ejemplificado serán etiquetas numéricas entre cero y dos, dado que el número de regiones es tres.

- *Combinar* las tres componentes espectrales. Es decir, se vuelve a conformar la matriz tridimensional perteneciente a cada imagen en color, la cual trae asociada sus tres vectores de etiquetas que permitirá re-etiquetar las diferentes combinaciones resultantes.
- Una vez segmentada la imagen e integrada con sus tres componentes espectrales, se desarrolla el proceso de decisión, cuya función objetivo consiste en minimizar las desviaciones entre las observaciones que pertenecen al mismo grupo o clase y maximizar las distancias entre los centros de los grupos. Para ello se estudia la similitud de las muestras entre sí dentro de una misma clase y las diferencias con las muestras de clases distintas, re-agrupándolas si es necesario o volviendo al proceso de segmentación (paso 2) para incrementar el número de umbrales. Finalmente, el algoritmo de clasificación concluye cuando todos los píxeles de la imagen han sido asignados a la clase que en teoría les corresponde.

Los cinco pasos anteriormente enumerados se pueden agrupar en dos: *segmentación* y *clasificación*, puesto que la etapa de codificación de píxeles se aplica en ambos procesos.

Además, una vez obtenidos los resultados, es necesario verificar el grado de acierto en la clasificación, para lo que se aplican los índices de validación interna y externa seleccionados al efecto, según los criterios establecidos en el capítulo dos. Finalmente se concluye con la interpretación de los resultados.

3.3.1. SEGMENTACIÓN Y CODIFICACIÓN

La segmentación es la operación de bajo nivel relacionada con la extracción de información relevante de las imágenes. Consiste en identificar las características principales que permitan aislar las estructuras consideradas de interés de acuerdo al problema que se plantea. En el caso que nos ocupa, las características son los píxeles cuya propiedad más relevante es la intensidad de cada componente espectral, es decir, el color.

Una segmentación basada en píxeles trata de etiquetar cada píxel como perteneciente a una de las clases subyacentes en la imagen, cuya naturaleza y número resultan desconocidos a priori. Como ya se ha mencionado previamente, en el caso que nos ocupa se procesa cada componente espectral de manera independiente. Para ello se

utiliza el método de segmentación por umbralización, siendo ésta una de las técnicas más ampliamente utilizadas que permite la agrupación de los píxeles en regiones homogéneas (Gonzalez y Woods, 2008). Procesar las componentes separadamente tiene la ventaja de que puede generalizarse para tratar imágenes multispectrales, teniendo en cuenta que la segmentación de las imágenes en color implica inicialmente la elección de un espacio de representación apropiado.

El método de Otsu es una de las técnicas que ha proporcionado buenos resultados como método de umbralización (Chang y col., 2006; Gonzales-Barron y Butler, 2006), considerándose incluso como algoritmo de umbralización global y local (Stathis y col. 2008), o como agrupamiento binario (Sezgin y Sankur, 2004).

La etapa de segmentación aplicada en esta investigación consiste en realizar subdivisiones sucesivas del histograma a través de la umbralización por el método de Otsu (1979).

La umbralización de una imagen a color implica que en cada una de las tres bandas se realiza una búsqueda exhaustiva de la varianza máxima entre las clases. Este mismo principio se aplica para seleccionar múltiples umbrales (multi-umbralización). El proceso se inicia con una umbralización sencilla: cada componente espectral se divide en dos regiones ($d = 2$), etiquetadas como 0 y 1. Como son tres las componentes, cada píxel posee 3 etiquetas ($z = 3$), por lo que se pueden obtener como máximo 8 grupos o clases ($d^z = 8$). Con el número de región y la etiqueta para cada píxel se genera la matriz de códigos de clase, que almacena la posición y número de clase que le corresponde a un píxel en concreto. Se evalúan los resultados decidiendo si se realiza un nuevo proceso de segmentación o no; en el primer caso se incrementa en uno el número de umbrales por componente espectral, obteniéndose ahora 2 umbrales. Cuando el número de umbrales por componente es mayor que uno se denomina umbralización múltiple. Si por ejemplo $d = 3$ y $z = 3$, habría 9 segmentos en total. Así se puede ir incrementando el valor de d hasta conseguir el número óptimo de agrupamientos. Una vez integradas las tres componentes espectrales de la imagen a color con sus etiquetas asociadas por píxel y por componente, el máximo número de clases generadas es d^z .

Las principales contribuciones de este trabajo al proceso de segmentación están en la manera de integrar los segmentos obtenidos por componente espectral en una única

imagen RGB, y la manera de generar los códigos de clase para cada píxel, así como el algoritmo de umbralización múltiple y el criterio de decisión entre umbralización simple o múltiple.

A continuación se detallan las etapas del proceso de segmentación para las imágenes en color objeto de estudio, iniciando el proceso por la separación de las componentes espectrales y su normalización. Luego se presenta el proceso de segmentación de píxeles por umbralización simple y se expone la extensión para umbralización múltiple, en ambos casos mediante la aplicación del método de Otsu. Por último, se describe el proceso de codificación de píxeles por componente espectral. Se concluye con la integración de la imagen en color con sus correspondientes etiquetas por píxel y por clase, y la generación de los códigos de etiquetas para la imagen bajo procesamiento.

3.3.1.1. SEPARACIÓN DE COMPONENTES ESPECTRALES

Las imágenes a color pertenecientes al espacio RGB se pueden representar por matrices de $m \times n \times 3$, donde cada una de las tres submatrices de dimensión $m \times n$ contiene los valores espectrales de luminancia correspondientes al Rojo, Verde y Azul.

El objetivo de este proceso es extraer cada una de las tres componentes espectrales de manera que se pueda tratar como una imagen en escala de gris. En la figura 3.6 se muestran los $N = m \times n$ píxeles con sus NDs entre 0 y 255, además de la correspondiente localización espacial respecto de la imagen (x,y) , que en la matriz de datos bidimensional se le denomina celda o píxel.

La separación de las componentes espectrales es independiente del modelo de color que se esté utilizando. La única salvedad proviene del hecho de que en las transformaciones, para pasar de un modelo de color a otro, los NDs no se mantienen generalmente en el rango mencionado previamente, pudiendo situarse en cualquier otro dependiendo de la función de transformación aplicada.

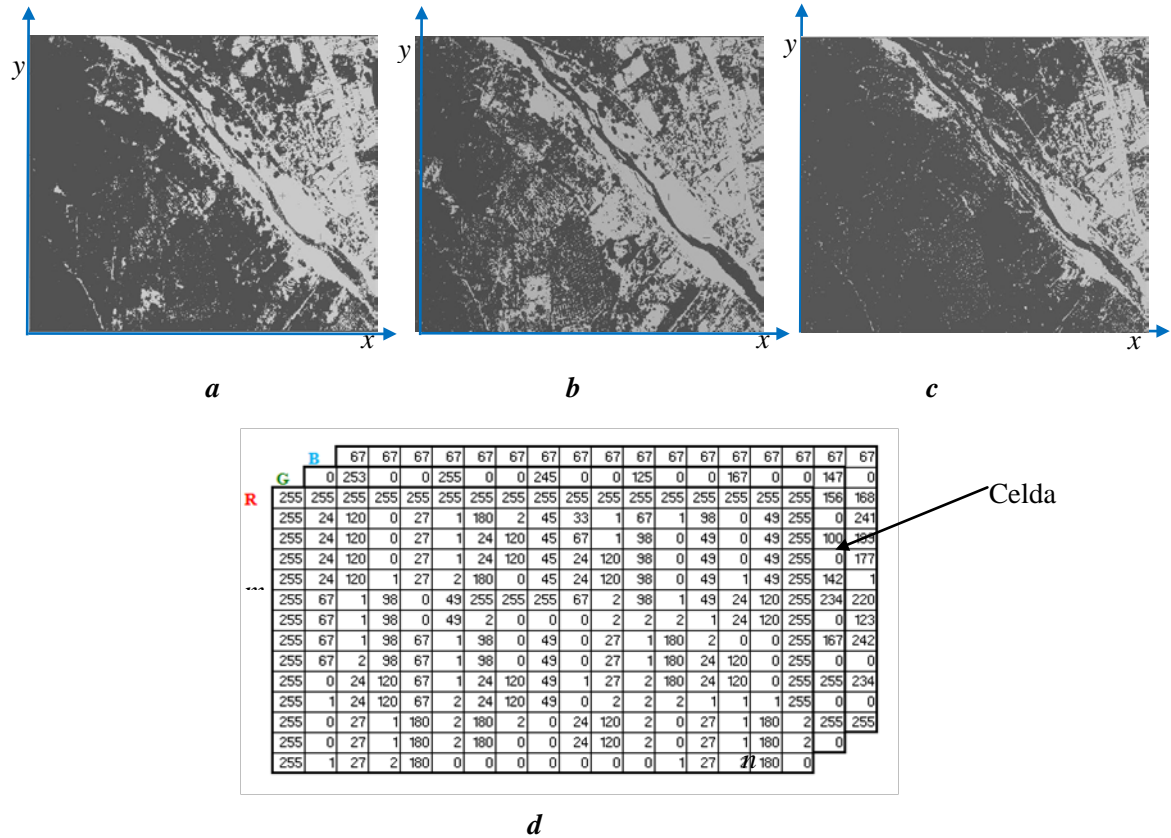


Figura 3.6. Imagen en color representada por mapas bidimensionales: (a), (b) y (c) representación de valores RGB respectivamente, (d) representación de una matriz tridimensional.

3.3.1.2. SEGMENTACIÓN DE IMÁGENES A COLOR POR UMBRALIZACIÓN SIMPLE

La idea básica de la umbralización simple, en imágenes a color, consiste en seleccionar automáticamente un umbral óptimo de nivel de gris por cada componente espectral simple, para poder separar dos partes significativas en cada una de dichas componentes, teniendo como base el propio histograma.

El problema principal a resolver en este caso es la manera de integrar los tres componentes de color, que permita clasificar adecuadamente los objetos presentes en la imagen.

Con el fin de formalizar de forma genérica el método, sea el conjunto de muestras $X = \{d_1, d_2, \dots, d_n\} \in \mathfrak{R}^s$ un subconjunto del espacio real s -dimensional \mathfrak{R}^s . Cada vector de dimensión s $d_k = \{d_{k_1}, d_{k_2}, \dots, d_{k_s}\} \in \mathfrak{R}^s$ se corresponde con una componente espectral dada, siendo d_{k_j} la j -ésima propiedad o atributo de la muestra relativa a la

componente espectral d_k . Como se ha mencionado previamente, y centrándonos en las imágenes que se utilizan en este trabajo de investigación, las características son los píxeles y sus propiedades las componentes espectrales en el modelo de color RGB; por tanto, $s = 3$ y $d_k = \{R, G, B\} \in \mathfrak{R}^3$.

De manera que al aplicar un determinado método para obtener un cierto umbral, a partir de la imagen, $i_s = f(x,y)$ en escala de gris, para cada componente espectral se obtienen los tres umbrales siguientes para el conjunto de las tres imágenes $t_{s1} = [t_{R1}, t_{G1}, t_{B1}]$.

Dada una banda espectral, el objetivo es determinar el valor t_{Otsu} de umbral que maximice la varianza entre clases o modos del histograma en dicha banda. Tomando como referencia las ecuaciones (2-10) a (2-15) que definen el método de Otsu (1979) para la extracción del umbral óptimo, se asume que la varianza global es mayor que cero ($\sigma_G^2 > 0$), ya que únicamente puede ser nula cuando todos los niveles de intensidad de la imagen son exactamente iguales. Es decir, siempre que en la imagen exista una única clase de píxeles $\eta = 0$, ya que la separabilidad de una única clase en relación a sí

misma es cero. La función objetivo $\eta(t) = \frac{\sigma_B^2(t)}{\sigma_G^2}$ está definida por la ecuación (2-10), de

forma que, al igual que $\sigma_B^2(t)$, se evalúa para encontrar t_{Otsu} seleccionándose aquel valor que proporciona el máximo, $\sigma_B^2(t_{Otsu}) = \max_{1 \leq t < L} \sigma_B^2(t)$, en caso de existir, y se obtiene

el promedio de los valores de t que los han generado.

Una vez obtenido el valor t_{Otsu} , que permitirá codificar la información de la imagen de entrada asociada a una de las bandas espectrales, $i_s = f(x,y)$, primeramente es necesario binarizar esa banda $\tilde{i}_s = g(x,y)$ con el valor correspondiente $t_{s1} = t_{Otsu}$, mediante la ecuación (3-1). En la figura 3.7 se muestra la binarización correspondiente a las tres bandas espectrales en escala de gris, imágenes (a, b, c); (d, e, f) indican los valores umbral de R (121), G (113), y B (103), respectivamente. Las imágenes binarizadas por medio de la ecuación (3-1) con los mencionados umbrales se representan en (g, h, i), donde las etiquetas “0” (negro) corresponden a los píxeles más oscuros, con intensidad menor o igual que 121 en (g), que 113 en (h) y que 103 en (i), mientras que con

etiquetas “1” se identifican los píxel más claros, es decir con valores superiores a los mencionados umbrales.

$$\tilde{i}_s = \begin{cases} 0 & \text{si } i_s \leq t_{s1} \\ 1 & \text{si } i_s > t_{s1} \end{cases} \quad (3-1)$$

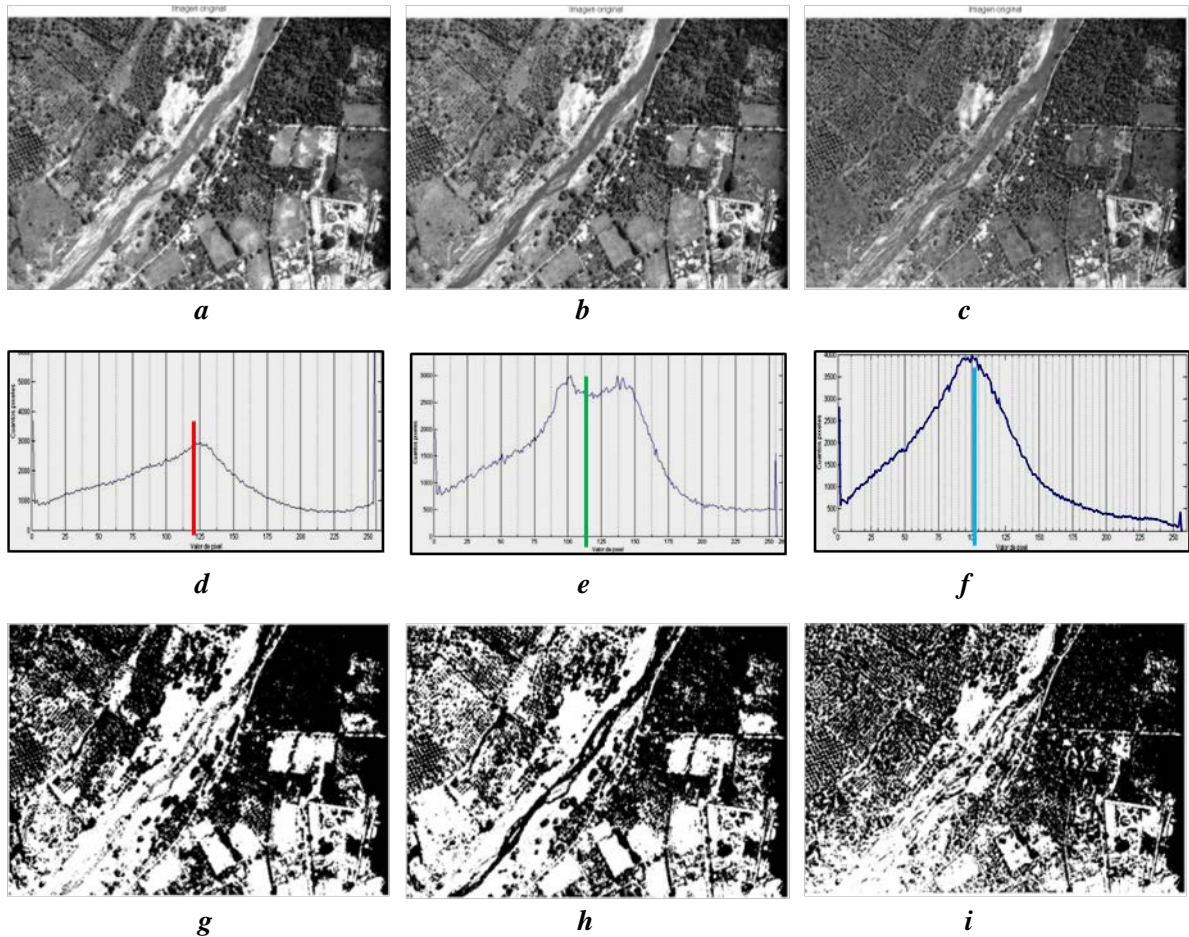


Figura 3.7. Binarización de una imagen a color: (a), (b) y (c) muestran los tres componentes espectrales R, G, B respectivamente; (d), (e) y (f) representan los umbrales de los tres histogramas R, G, B; (g) componente R binarizado, (h) componente G binarizado, (i) componente B binarizado.

3.3.1.3. EXTENSIÓN DEL MÉTODO DE OTSU A UMBRALIZACIÓN MÚLTIPLE DE IMÁGENES A COLOR

El método que se presenta en este apartado es en realidad una adaptación para imágenes en color del método anterior. Multi-umbralizar una imagen a color consiste en separar los píxeles de cada componente espectral de la imagen en M grupos de píxeles o clases, mediante la obtención de los valores del vector de umbrales $t = (t_1, \dots, t_{M-1})$. Si cada

componente espectral contiene N píxeles con $L-1$ niveles de gris, $[0, \dots, L-1]$, y denotamos cada clase con el símbolo C , entonces C_1 contiene todos los píxeles con nivel de gris desde $[1, \dots, t_1]$, C_2 desde $[t_1+1, \dots, t_2]$, ..., C_i desde $[t_{i-1}+1, \dots, t_i]$, ..., y C_M desde $[t_{M-1} +1, \dots, L]$. Los umbrales óptimos $(t_1^{Otsu}, t_2^{Otsu}, \dots, t_{M-1}^{Otsu})$ se obtienen maximizando la ecuación (3-2), tal y como se describe en (2-13).

$$\sigma_B^2(t_j) = \frac{[\mu_T w(t_j) - \mu(t_j)]^2}{w(t_j)[1 - w(t_j)]} = \sum_{j=1}^M w(t_j) [\mu^*(t_j) - \mu_T]^2 \quad (3-2)$$

donde

$$w(t_j) = \sum_{i \in c_j} p(i) \quad \mu(t_j) = \sum_{i \in c_j} i \cdot p(i) \quad (3-3)$$

Siendo $w(t_j)$ la probabilidad acumulada hasta el t -ésimo nivel de la j -ésima clase C_j , y $\mu(t_j)$ la media de clases ponderada de la j -ésima clase (C_j) hasta el t -ésimo nivel. El nivel medio total μ_T de la imagen se obtiene mediante la ecuación (3-4).

$$\mu_T = \sum_{j=1}^M w(t_j) \mu^*(t_j) = \sum_{i=0}^{L-1} i \cdot p(i) = \mu(L) \quad (3-4)$$

A diferencia de la segmentación simple, en la cual se busca un umbral óptimo, ahora se deben elegir $M-1$ umbrales óptimos, maximizando para cada uno de ellos la *varianza entre las clases* $\sigma_B^2(t_j)$.

$$\sigma_B^2(t_1^{Otsu}, t_2^{Otsu}, \dots, t_{M-1}^{Otsu}) = \max_{1 \leq t_1 < \dots < t_{M-1} < L} \{ \sigma_B^2(t_1, t_2, \dots, t_{M-1}) \} \quad (3-5)$$

Por tanto,

$$\sum_{j=1}^M w(t_j) = 1 \quad \text{y} \quad \mu^*(t_j) = \sum_{i \in c_j} (i \cdot p(i)) / w(t_j) \quad (3-6)$$

Al finalizar dicho proceso se tendrá como resultado una matriz de umbrales, con tantas filas como componentes espectrales posea la imagen analizada. En los experimentos

llevados a cabo en este trabajo son tres las imágenes, y tantas columnas como valores umbrales se han obtenido. Dichos valores se utilizarán en el siguiente apartado para codificar la información subyacente en la imagen procesada.

3.3.1.4. CODIFICACIÓN DE PÍXELES POR COMPONENTE ESPECTRAL

El proceso de codificación consiste en asociar una etiqueta numérica a cada píxel, ubicado en su correspondiente localización espacial (x, y) en coordenadas de la imagen, y en relación a cada una de las componentes espectrales que componen dicha imagen, siempre según el grupo al que pertenece.

Se denominan *grupos* o *clases* al conjunto de píxeles identificados por el mismo código numérico asignado.

Como se mencionó anteriormente, los métodos de umbralización simple dividen el histograma en dos regiones; por tanto, para una imagen que contiene tres componentes espectrales se obtienen seis sub-regiones. Sobre cada una de las dos regiones obtenidas en cada banda espectral cabe la posibilidad de aplicar una nueva umbralización, entrando así en lo que podemos denominar multi-umbralización. Con un segundo umbral sobre una cualquiera de las dos partes se obtienen tres particiones del histograma por componente espectral, nueve en total para las tres componentes espectrales. Si se aplica un tercer umbral obtendríamos cuatro particiones por canal, y así sucesivamente. Puesto que cada píxel de cada componente espectral se representa por su nivel de gris i , y su componente espectral s , siendo $s=0=R$; $s=1=G$ y $s=2=B$, entonces $\{i_0(x, y), i_1(x, y), i_2(x, y)\} \equiv \{R, G, B\} \in \mathfrak{R}^3$ (si se usa el modelo de color RGB, o cualquier otro modelo con tres componentes espectrales). Además, el segmento al que pertenece se identifica por una tercera etiqueta k , siendo $k = [1, 2, \dots, M]$.

Por tanto, suponiendo que el número de umbrales por canal es finalmente M , tendremos $\{t_{01}, t_{02}, \dots, t_{0M-1}\}$, para el canal R , y de la misma manera $\{t_{11}, t_{12}, \dots, t_{1M-1}\}$ para la componente G , y $\{t_{21}, t_{22}, \dots, t_{2M-1}\}$ para la componente B , donde como siempre t representa el umbral correspondiente al nivel de gris de la banda en cuestión.

Sobre esta base, cada píxel i puede ser codificado como $\tilde{i}_s(x, y) = f(x, y)$ según sus componentes espectrales a través de la ecuación siguiente.

$$\tilde{i}_s = \begin{cases} 0 & \text{si } i_s \leq t_{s1} \\ 1 & \text{si } t_{s1} < i_s \leq t_{s2} \\ 2 & \text{si } t_{s2} < i_s \leq t_{s3} \\ \vdots & \\ M-1 & \text{si } i_s > t_{sM-1} \end{cases} \quad (3-7)$$

donde s denota la componente espectral correspondiente. Así, si las imágenes de cobertura terrestre a clasificar se encuentran en el modelo de color RGB entonces $s = R, G$ ó B ; si las imágenes se encuentran en modelo de color CIE $L^*a^*b^*$, entonces $s = L, a$ ó b ; y t_{si} son los umbrales consecutivos.

Mediante el uso de la imagen codificada y la imagen original es posible construir el vector de propiedades de cada grupo (valor promedio y varianza respecto de la intensidad dentro de cada grupo).

3.3.1.5. INTEGRACIÓN DE LAS COMPONENTES ESPECTRALES Y CODIFICACIÓN

A continuación se describe el proceso resultante de integrar las tres componentes espectrales de la imagen a color (mezcla) con sus códigos de etiqueta asociados y asignados durante el proceso de segmentación de componentes individuales. Además se expone la metodología seguida para etiquetar las clases resultantes de la imagen.

Puesto que cada componente espectral fue segmentada y codificada de manera independiente, ahora el proceso consiste en la mezcla de las tres componentes espectrales como una matriz de datos tridimensional.

Cada píxel $i \equiv (x, y)$ tiene asociado un código de clase formado por tres valores numéricos $(\tilde{i}_0, \tilde{i}_1, \tilde{i}_2)$, de forma que una misma combinación de códigos identificará a todos los píxeles que pertenecen a la misma clase.

Si la segmentación de los píxeles de cada componente espectral se lleva a cabo mediante un único umbral, las etiquetas numéricas codificadas corresponden al sistema de numeración en base dos (0, 1); si son dos umbrales, las etiquetas numéricas codificadas corresponderán al sistema de numeración en base tres (0, 1, 2), y así sucesivamente para cada componente espectral. Por ejemplo, la figura 3.8 *a*, representa

los códigos de etiquetas integrados en cada píxel para una imagen segmentada y codificada por un código de numeración base dos.

Así como en la codificación fue necesario asociar a cada píxel una etiqueta por componente espectral que identificara cada región segmentada, también resulta necesario identificar cada uno de los grupos generados en la imagen en color. La siguiente matriz de etiquetas se construye con el objetivo de asociar a cada localización espacial de la imagen, (x, y) , la clase a la que pertenece el píxel correspondiente. Por tanto la matriz debe ser del mismo tamaño que la imagen original, de manera que permita guardar en cada celda (localización) el valor de la etiqueta de clase a la que pertenece el píxel correspondiente.

La ecuación (3-8), nos permite generar los códigos de etiquetas bajo un sistema de numeración decimal, figura 3.8 *c*, donde cada etiqueta representa una determinada clase o grupo.

$$\tilde{i} = \tilde{i}_2(b^2) + \tilde{i}_1(b^1) + \tilde{i}_0(b^0) = \sum_{j=2}^0 \tilde{i}_j(b^j) \quad (3-8)$$

- \tilde{i} : es el resultado, en código numérico decimal, que representará la nueva etiqueta de cada píxel en la imagen a color.
- b : es el número de regiones en que se dividen los histogramas de las bandas espectrales. Sus valores son 2, 3 ó 4 según que se extraigan 1, 2, ó 3 umbrales respectivamente. En términos de la teoría fundamental de la numeración, b representa la base numérica (Alcalde y col., 1994).
- \tilde{i}_j : identifica la etiqueta asociada a cada píxel en cada componente espectral, donde el subíndice j determina el número de componente espectral. En el caso de las imágenes en color se tiene que $j = [0, 1, 2]$, representando las bandas R, G y B respectivamente.

El número de combinaciones posible será el máximo valor de $\tilde{i} + 1$, ya que las etiquetas generadas varían de 0 a \tilde{i} . Por ejemplo, el máximo valor generado (en un sistema de numeración decimal) cuando los códigos de etiquetas por componente espectral se representan por un sistema de numeración base 3 (0, 1 y 2) es 27, lo que significa que al

combinar las tres componentes espectrales segmentadas podrán obtenerse como máximo 27 grupos o clases diferentes: $\tilde{i} = 2(3^2) + 2(3^1) + 2(3^0) = 18 + 6 + 2 = 26 + 1 = 27$.

Un ejemplo de este sistema de codificación es el que se describe en la figura 3.8, en **a** se representa la matriz de códigos de etiquetas (\tilde{i}_j , en un sistema de numeración base 2) por componente; en **b** los códigos de color que representa cada etiqueta, y finalmente en **c** los códigos de clase en sistema de numeración base 10, generados a partir de los códigos de cada componente espectral.

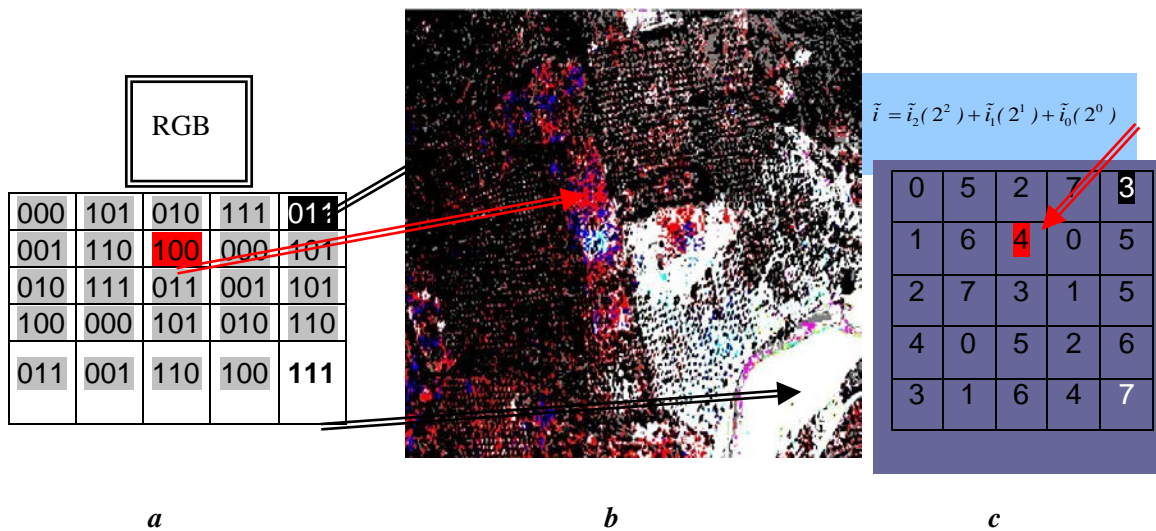


Figura 3.8. Ejemplo del proceso de codificación de la imagen en color: (a) valores de los píxeles con los códigos de etiqueta de cada componente espectral, (b) imagen en color segmentada, (c) código de clase que identifica los segmentos generados en la imagen a color.

Mediante el uso de la matriz de códigos de clases de la imagen a color se facilita el cálculo de los descriptores estadísticos del histograma, ya que una misma localización (x, y) de cada píxel puede ser identificada en cualquiera de las imágenes (original, segmentada o etiquetada).

Hasta este momento se han mostrado los resultados de la segmentación por umbralización, así como el resultado de la combinación de los tres vectores de etiquetas, considerando la localización espacial de cada píxel. El siguiente paso consiste en determinar la clasificación propiamente dicha.

3.3.2. CLASIFICACIÓN

Durante la fase previa de segmentación es de esperar que las regiones resultantes del agrupamiento de píxeles sean *homogéneas* con respecto a las características del color. Sin embargo, no existe garantía alguna de que estas regiones muestren también *separación espacial*, que resulta ser una segunda propiedad deseable en aplicaciones de clasificación junto con la de homogeneidad.

Es decir, una buena clasificación debe ser tal que los elementos de la clase deben agruparse mutuamente lo máximo posible a la vez que las clases se diferencian o separan al máximo. Estos son los criterios que miden los índices de validación de las particiones descritos en el capítulo dos. Pues bien, con el objetivo de garantizar el mejor cumplimiento posible de los criterios de homogeneidad y separación, se desarrolla un algoritmo de clasificación cuyo fundamento es el principio que expone Otsu (1979), tratando de minimizar las desviaciones entre las observaciones que pertenecen al mismo grupo, y de maximizar las distancias entre los centros de los grupos.

Puesto que la homogeneidad se traduce en el hecho de que los píxeles de cada grupo deben estar lo más cerca posible unos de otros, el criterio apropiado es el de minimizar la varianza dentro de cada grupo. Los píxeles de las regiones fragmentadas pueden tener que ser reacomodados (fusionados en una clase) para conseguirlo.

Respecto del criterio de separación espacial lo que se busca es exactamente maximizar las distancias entre los centros de los grupos.

Ambos criterios son los mismos que se aplican en la segmentación respecto de la extracción de umbrales por Otsu, si bien ahora sobre las clases de la partición que se está estableciendo. Desde esta perspectiva el clasificador propuesto puede denominarse Otsu-Clustering para imágenes en color. Esto representa una importante contribución desde el punto de vista de métodos de clasificación de naturaleza no supervisada.

3.3.2.1. MEDIDAS DE HOMOGENEIDAD Y SEPARACIÓN EN LA CLASIFICACIÓN

Sabemos que los elementos de las clases serán píxeles con tres valores, obtenidos a partir del proceso de umbralización previo. Pues bien, para cada clase se calcula el promedio de los valores de sus píxeles mediante la ecuación (3-9).

$$\boldsymbol{\mu}_k \equiv (\mu_r^k, \mu_g^k, \mu_b^k) = \frac{1}{N_k} \sum_{i_k \in C_k} \mathbf{i}_k \quad (3-9)$$

Se calcula también la varianza dentro de cada clase y entre las clases, denotadas por σ_k y σ_{hk} , respectivamente, lo cual se lleva a cabo mediante las ecuaciones (3-10) y (3-11). Evidentemente la varianza dentro de cada clase σ_k está relacionada únicamente con la clase k (C_k), y la varianza entre dos clases C_k y C_h , denotada por σ_{kh} , involucra ambas clases k y h , con $k \neq h$.

$$\sigma_k = \frac{1}{d \cdot N_k} \sum_{i_k \in C_k} \left[(i_r^k - \mu_r^k)^2 + (i_g^k - \mu_g^k)^2 + (i_b^k - \mu_b^k)^2 \right]^{1/2} \quad (3-10)$$

$$\sigma_{kh} = \frac{1}{d} \left[(\mu_r^k - \mu_r^h)^2 + (\mu_g^k - \mu_g^h)^2 + (\mu_b^k - \mu_b^h)^2 \right] \quad (3-11)$$

donde d es la dimensión del vector de componentes, tres en nuestro caso.

Basándose en la varianza, podemos fusionar aquellas clases que poseen un alto grado de similitud en sus componentes espectrales. La similitud es un concepto definido de la siguiente manera. Dada las clases C_k y C_h , $k \neq h$, ambas se fusionan en una única clase si $\sigma_k \geq \sigma_{hk}$ ó $\sigma_h \geq \sigma_{kh}$, es decir, la varianza dentro de cada clase debe ser menor que la varianza entre las clases. Si la varianza de alguna clase es mayor que la existente entre ellas, eso significa que ésta se solapa con otra clase, por tanto la mejor opción es su fusión. Este proceso de re-agrupamiento se repite con todos los píxeles de todas las clases hasta que todas las *varianzas entre clases* son mayores que sus correspondientes *varianzas dentro de las clases*. Sin pérdida de generalidad, si se fusionan los píxeles de dos clases, la clase resultante será re-etiquetada con la clase con valor de varianza más pequeña, mientras que la etiqueta que queda libre se elimina definitivamente. Este hecho no afecta en absoluto al proceso de clasificación.

La idea que subyace en el proceso anterior se fundamenta en la hipótesis de que si se ha logrado una buena partición, las clases obtenidas deben estar debidamente separadas, sin superposiciones importantes, con lo cual no sería necesaria la fusión. Por el

contrario, si las clases llegan a superponerse significa que la varianza espectral entre dos clases σ_{kh} es menor que la varianza individual dentro de las clases σ_k y σ_h .

Por tanto, durante el proceso anteriormente expuesto, pueden darse dos situaciones:

1. que una o más clases se hayan fusionado.
2. que no haya sido fusionada ninguna clase.

En el primer caso se dice que la clasificación fue exitosa y se generan los correspondientes vectores de propiedades de cada clase, a la vez que el etiquetado final con el correspondiente reconocimiento de las clases.

El segundo caso podría ser claro indicio de que por las características espectrales de la imagen (varianzas espectrales bien definidas), es posible encontrar más agrupaciones. Por lo tanto, para conseguirlo, el número de umbrales por componente espectral debe incrementarse en uno, lo cual conduce a repetir el ciclo de clasificación (segmentación, codificación, mezcla y decisión).

3.4. APLICACIÓN DEL MÉTODO DE CLASIFICACIÓN NO SUPERVISADA PROPUESTO

3.4.1. SEGMENTACIÓN DE PÍXELES POR UMBRALIZACIÓN SIMPLE Y CODIFICACIÓN

Para ilustrar el desarrollo de este apartado se utiliza una imagen a color representativa del total de las imágenes analizadas (figura 3.9 *a*), sobre la cual se aplica el proceso de umbralización simple y clasificación.

La imagen digital fue captada por sensores aéreos con las siguientes características: tamaño 1524×1012 , lo que hace un total de 1.542.288 valores espectrales digitales por componente, correspondientes a una resolución radiométrica de 256 niveles. En las figuras 3.9 *b*, *c* y *d*, se muestran las tres componentes espectrales que la integran.

Como se ha explicado previamente, el objetivo de la primera etapa de segmentación consiste en encontrar cuál de los 256 *ND* es el que permite discriminar entre dos grupos de cobertura terrestre dentro de una componente espectral. Siguiendo la filosofía del método de Otsu, si el umbral seleccionado es adecuado, la varianza existente dentro de cada grupo de píxeles generado por el umbral debe ser lo más pequeña posible, y al mismo tiempo, la varianza entre los dos grupos debe ser lo más alta posible, con un mínimo promedio de error incluido en la asignación de píxeles. Es decir, se calcula el

cociente entre ambas variancias y se busca un valor umbral para el que este cociente sea máximo.

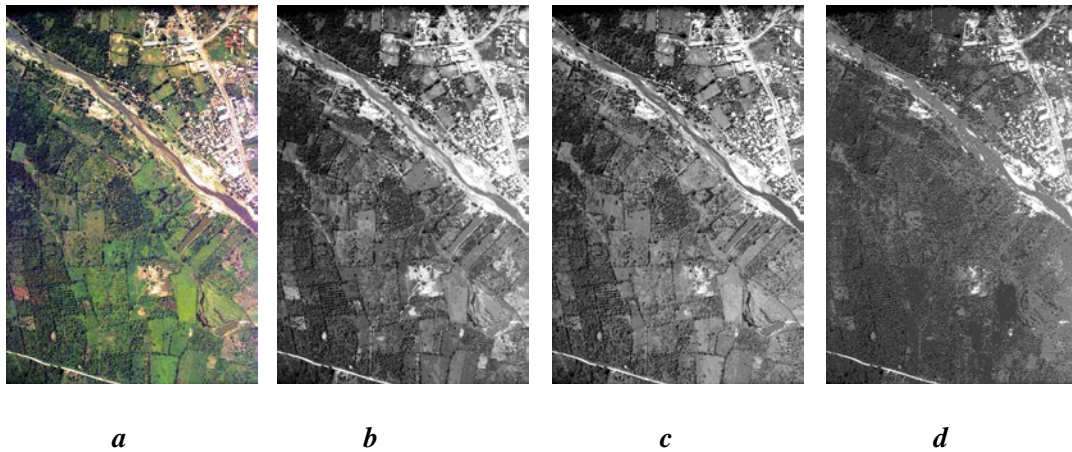


Figura 3.9. (a) Fotografía aérea en modelo de color RGB, (b) componente R (rojo), (c) componente G (verde), (d) componente B (azul).

Obsérvese sobre las bandas espectrales (figura 3.9 **b**, **c**, **d**) que los píxeles representados con tonalidades claras (cercanos al blanco) corresponden a ND altos en dicha banda, mientras que los píxeles de color oscuro, a los ND cercanos al cero.

El primer paso consiste en calcular el histograma de frecuencias de los niveles de gris, el cual acumula toda la información estadística sobre una imagen o cualquier fragmento de ella. Para cada nivel de gris se puede determinar primeramente el histograma de la imagen o, lo que es lo mismo, el conjunto del número de píxeles para cada uno de los valores digitales. La figura 3.10 muestra los correspondientes histogramas para las tres bandas espectrales consideradas en la imagen de la figura 3.9. Dividiendo el valor del histograma para un determinado ND , i , por el número total de píxeles existentes en la imagen se obtiene la probabilidad de ocurrencia de dicho ND , $p(i)$, (sección 2.4.1). A partir del histograma es posible calcular el nivel digital promedio de la imagen μ_T (media total) y la variación de los niveles digitales alrededor de la media σ_G^2 (varianza total).

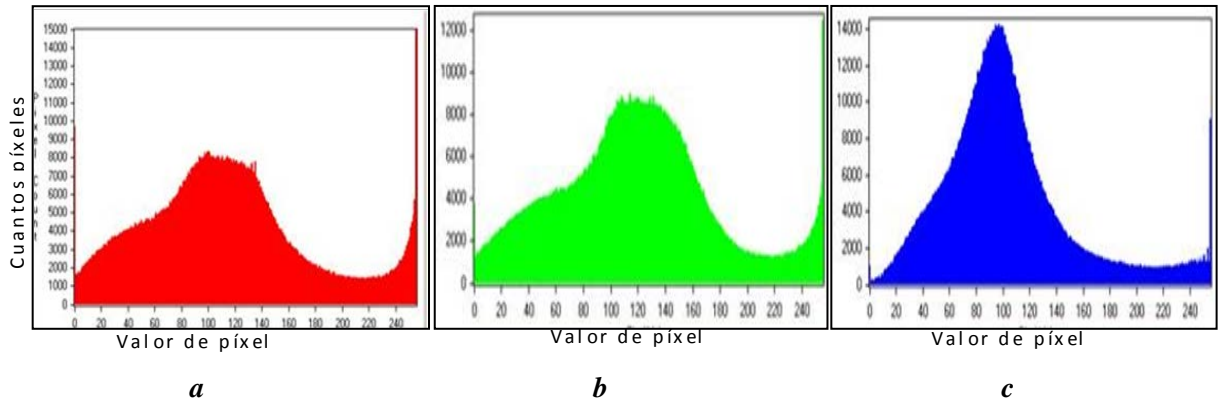


Figura 3.10. Histograma que muestra la distribución de los niveles digitales de la imagen RGB de la figura 3.9: **(a)** componente R, **(b)** componente G y **(c)** componente B.

El objetivo es encontrar el mejor valor ND o umbral t que establezca la separación entre dos partes del histograma. Para ello se aplica el proceso descrito en la sección 2.4.1, obteniendo la probabilidad acumulada $w(t)$, el momento acumulado de primer orden $\mu(t)$ o media de clases ponderada, así como la varianza $\sigma^2(t)$ o momento acumulado de segundo orden. Nótese que las variables anteriores dependen de t . El problema se reduce a un problema de optimización, donde el umbral óptimo t_{Otsu} es el valor del $ND(t)$ que obtiene la máxima variación en la función objetivo $\eta(t)$, tal y como se establece en la sección antes mencionada. La figura 3.11 muestra la variación de la función objetivo con respecto a los distintos umbrales para la banda R, observándose que la máxima variación se produce en $t = 132$, que representa el umbral óptimo. De igual manera para las componentes G y B se determinan los correspondientes umbrales con valores 114 y 127 respectivamente.

Una vez establecidos los correspondientes umbrales, los píxeles se etiquetan de acuerdo con ellos. Así con respecto a la banda R, a todos los píxeles con valores en dicha banda inferiores a 132 se les asigna la etiqueta 0 mientras que al resto se les asigna la etiqueta 1. Del mismo modo se procede para las bandas G y B, obteniéndose los resultados mostrados en la figura 3.12. Por ejemplo, en la banda R las etiquetas con valor cero (negro) representan áreas con mayor cubierta vegetal, mientras que las de valor uno (blanco) se asignan al suelo con poca cubierta vegetal.

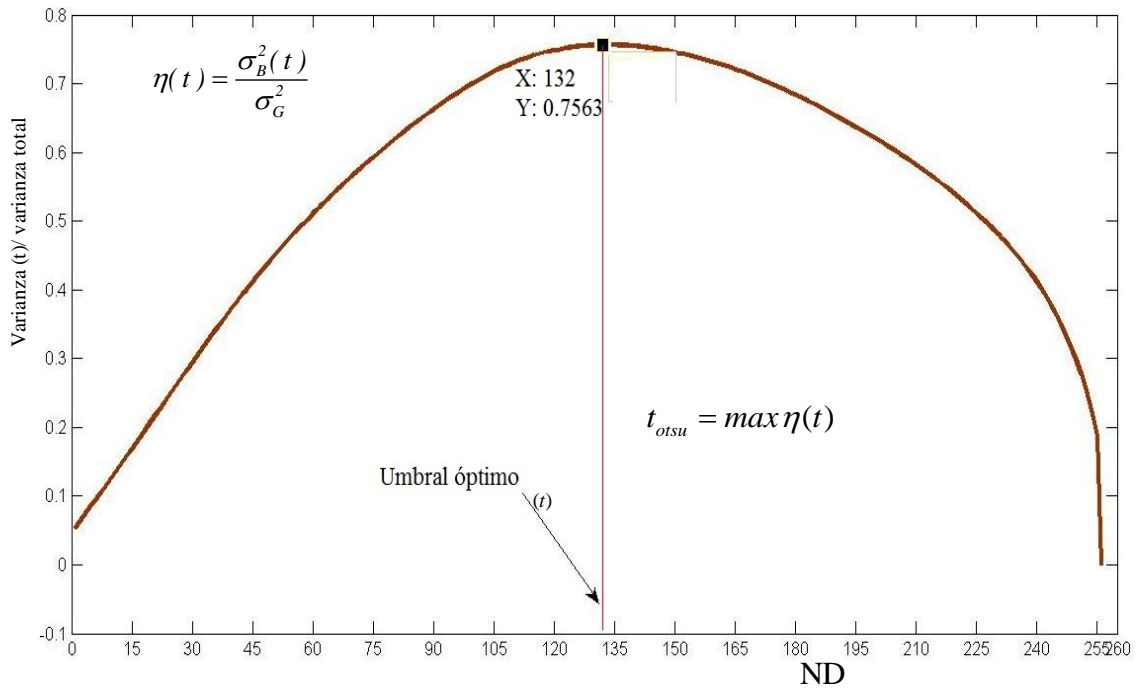


Figura 3.11. Gráfico de separabilidad por nivel digital (umbral óptimo para la componente R).

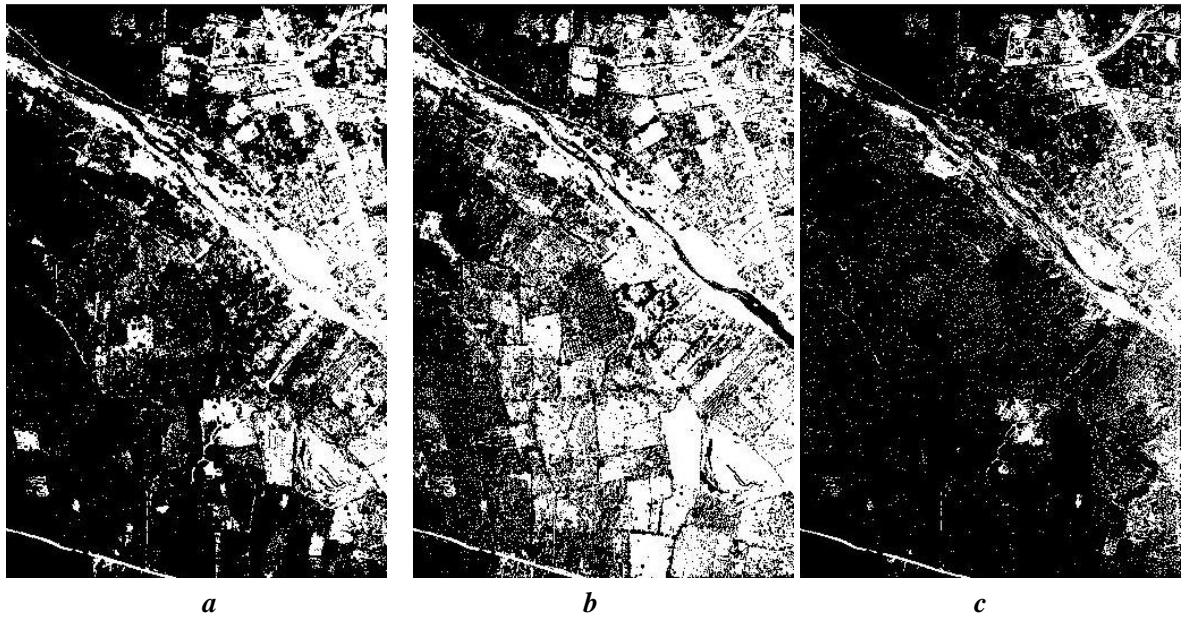


Figura 3.12. Resultado de la codificación en dos clases por componente espectral de la figura 3.9 **a**: (a) banda R, (b) banda G y (c) banda B.

Tras la codificación se obtienen tres matrices de etiquetas (una por componente espectral), con códigos binarios que identifican para cada posición (x, y) , el segmento al que pertenecen con respecto a cada histograma. Además, para cada uno de los

segmentos establecidos por los umbrales se calculan los principales momentos estadísticos del histograma, que se muestran en la tabla 3.1.

Tabla 3.1. Momentos estadísticos del histograma por componente espectral y valor de umbral de la imagen de la figura 3.9.

Medidas		Rojo	Verde	Azul
Segmento 1	Probabilidad de ocurrencia	0,69	0,50	0,81
	Nivel digital promedio	76,02	68,27	79,13
Segmento 2	Probabilidad de ocurrencia	0,31	0,50	0,19
	Nivel digital promedio	187,52	159,41	174,15
Promedio Total		110,46	114,03	97,36
Varianza Total		3509,30	3197,80	2218,20
Varianza Máxima Normalizada		0,76	0,65	0,63
Umbral Óptimo		132,00	114,00	127,00

3.4.1.1. MEZCLA DE COMPONENTES ESPECTRALES Y REPRESENTACIÓN

Dado que se han utilizado tres umbrales, uno por banda, se han generado ocho clases, que se muestran en la figura 3.13. A modo de ejemplo, los píxeles pertenecientes a la clase 2 fueron etiquetados con sus tres valores a cero, que en este caso representa la masa arbórea, mientras que los píxeles etiquetados con sus tres valores a uno corresponden a la clase 8, representando principalmente suelos o áreas carentes de vegetación y por tanto con un alto contenido de arcillas.

Para la representación mostrada en la figura 3.13 se han calculado los valores promedio de los *ND* en cada banda, una vez que los píxeles han sido etiquetados convenientemente. En la tabla 3.2 se muestran dichos valores para las diferentes clases. La combinación de estos valores para las tres bandas R, G, B generan los colores mostrados en la mencionada figura.

Tabla 3.2. Nivel digital promedio por clase y componente espectral de la figura 3.13

Clases	Rojo	Verde	Azul
1	61,30	67,62	66,12
2	112,29	99,09	136,76
3	111,02	133,03	94,85
4	121,56	132,42	135,98
5	147,11	104,18	102,27
6	143,27	102,89	140,40
7	156,49	151,35	106,11
8	215,07	199,51	179,48

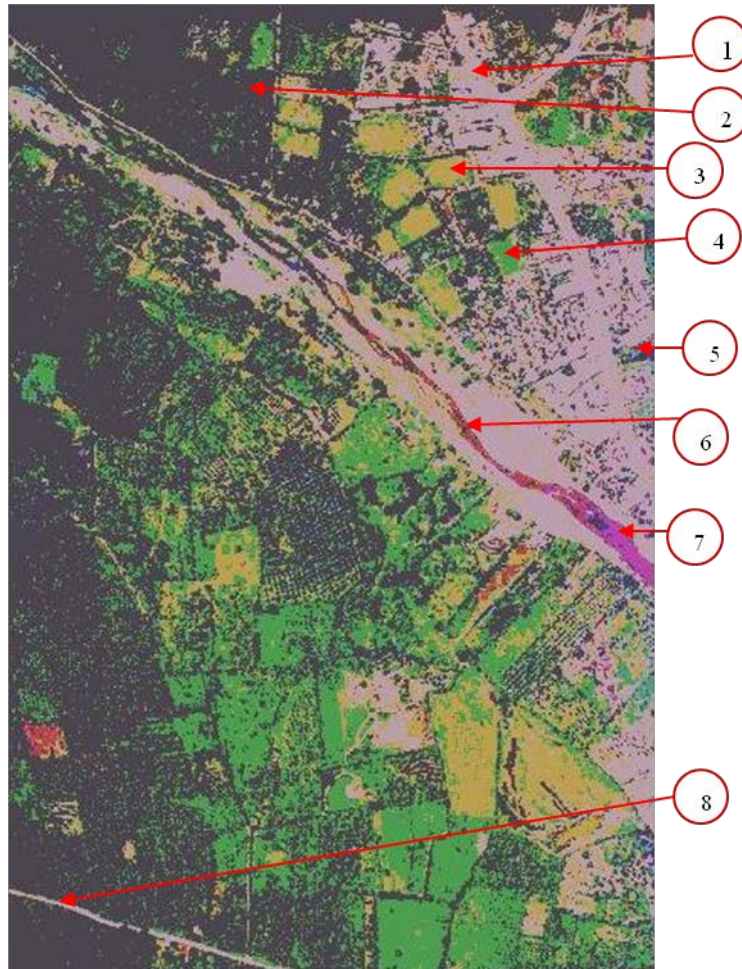


Figura 3.13. Imagen resultante de la mezcla de las tres componentes espectrales después de ser segmentado cada uno en dos clases de textura.

3.4.2. CLASIFICACIÓN

En este proceso se decide la clase a la que se asignan los píxeles una vez evaluadas sus características tras el proceso de segmentación. Los valores de la diagonal principal de la tabla 3.3 corresponden a las varianzas dentro de cada clase. Los valores fuera de la diagonal corresponden a las varianzas entre cada par de clases.

Tabla 3.3. Varianza dentro de cada clase y entre clases de la figura 3.13

Clases	1	2	3	4	5	6	7	8
1	42,56	1904,74	1670,09	2463,94	2532,55	2927,66	3855,50	10681,35
2	1904,74	1001,02	700,33	371,11	566,25	320,06	1162,10	4905,68
3	1670,09	700,33	133,08	565,21	515,37	830,90	699,27	4570,73
4	2463,94	371,11	565,21	775,67	511,47	330,42	517,71	3338,42
5	2532,55	566,25	515,37	511,47	662,19	484,69	741,93	3936,06
6	2927,66	320,06	830,90	330,42	484,69	1124,60	877,14	3591,14
7	3855,50	1162,10	699,27	517,71	741,93	877,14	220,21	2263,89
8	10681,35	4905,68	4570,73	3338,42	3936,06	3591,14	2263,89	255,29

A modo de ejemplo, para ilustrar mejor el proceso, se puede observar que los valores de las varianzas dentro de las clases 2, 4, 5 y 6, son mayores que algunas de las varianzas entre clases. Por ejemplo, la varianza dentro de la clase 2 es 1001,02; la de la clase 3 es 133,08; y la varianza entre ambas clases es de 700,33. Como la varianza entre ambas clases es menor que la de la clase 2, ésta se debe reagrupar con alguna otra. Sin embargo la clase 3 no necesita fusión ya que se encuentra bien separada de acuerdo con el criterio aplicado, su varianza es menor que las varianzas entre todas las clases (de hecho su varianza es el menor valor de toda la fila y columna).

Siguiendo con la comparación entre la clase 2 y la clase 4, la varianza dentro de la clase 4 es de 775,67 y la varianza entre ambas resulta ser de 371,11. En este caso ambas clases cuentan con varianza dentro de la clase mayor que la varianza entre ambas, por tanto deben fusionarse en una única clase.

Como resultado del reagrupamiento, de las ocho clases iniciales cuatro de ellas fueron fusionadas en una única clase. Por tanto el número de clases resultante para esta imagen resulta ser finalmente cinco. En la tabla 3.4 se muestran los promedios y el número de píxeles por clase tras el proceso de fusión. En la figura 3.14 se muestran los mismos resultados de forma grafica.

Tabla 3.4. Nivel digital promedio de la imagen clasificada (figura 3.14), por clase y componente espectral.

Clases	Rojo	Verde	Azul
1	61,30	67,62	66,12
2	132,91	112,22	124,28
3	111,02	133,03	94,85
4	156,49	151,35	106,11
5	215,07	199,51	179,48

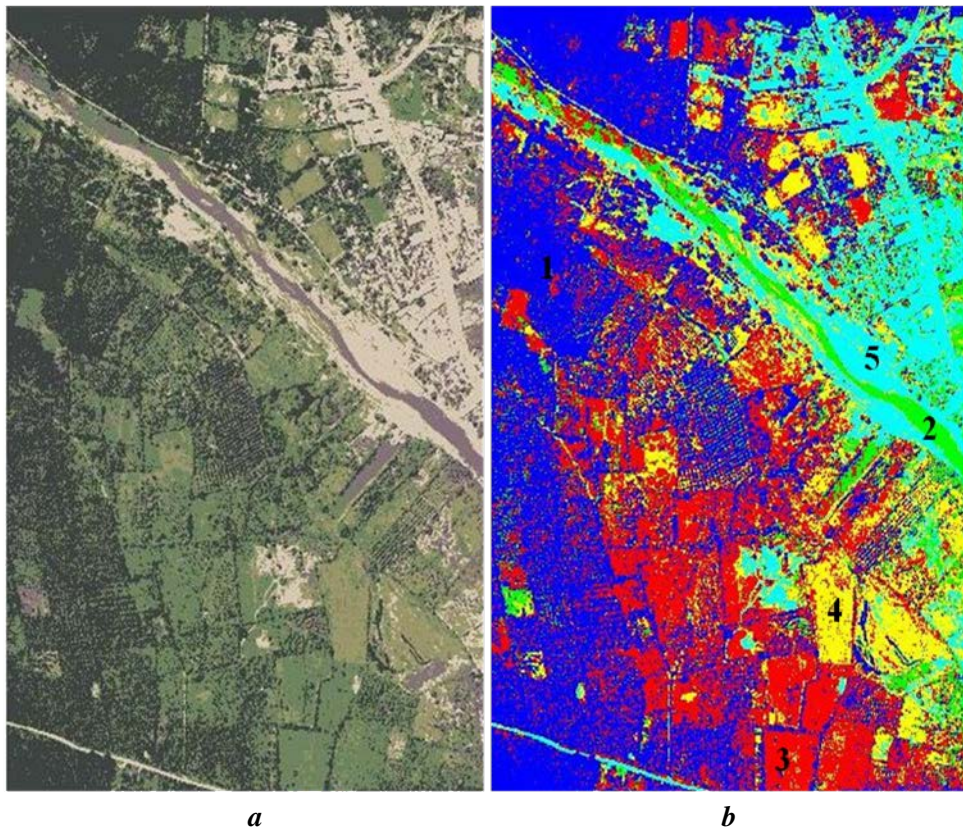


Figura 3.14. Imagen clasificada: (a) promedios de color RGB, (b) asignación de colores por clase a cada píxel.

3.4.3. INTERPRETACIÓN DE RESULTADOS

Una vez que todos los píxeles de la imagen de cobertura terrestre han sido clasificados por el método no supervisado propuesto, es necesario analizar los resultados.

La interpretación más relevante debe ser dada por un experto reconocedor de la zona de estudio, o con habilidad para foto-interpretar imágenes geográficas, cuya finalidad última consiste en la asignación de nombres reales a las etiquetas de las clases. Es decir, en base a los resultados de la clasificación no supervisada y la experiencia del experto, verificar el uso que se le está dando a cada una de las clases encontradas y realizar el etiquetado definitivo de las clases, como se muestra en la tabla 3.5.

Por tanto, como resultado de este proceso de análisis e interpretación, el clasificador agrupó de manera automática los diferentes objetos o texturas del paisaje en cinco clases, generando la imagen clasificada mostrada en la figura 3.14. El 36,82 % de la superficie de la imagen contiene vegetación arbórea, representada en la imagen clasificada por el color azul. El 8,68 % corresponde a los cuerpos de agua, identificados por el color verde, y el 27,65% contiene pastos cultivados de altura media, representados por el color rojo. El 14,18 % contiene pastos de crecimiento corto y con poca densidad, color amarillo, y el 12,67 % son zonas sin cobertura vegetal, color verde-azul. Dentro de esta última se consideran los sedimentos acumulados sobre la ribera del río por efecto del huracán Paulina, así como las calles y los edificios. La cantidad de píxeles agrupados en cada clase se muestra también en la tabla 3.5.

Tabla 3.5. Reconocimiento de la cobertura terrestre de la figura 3.9 *a*.

C_k	Uso que tiene el suelo de acuerdo con la cobertura terrestre	Área (píxeles)	Cobertura (%)	Color
1	Vegetación arbórea (agrícola perene)	567.933	36,82	Azul
2	Cuerpos de agua	133.887	8,68	Verde
3	Pasto cultivado de altura media (agrícola de ciclo corto)	426.380	27,65	Rojo
4	Pastos corto poco denso (agrícola de ciclo corto)	218.634	14,18	Amarillo
5	Zonas sin cobertura vegetal	195.454	12,67	Verde-Azul

C_k : Número de clase de cobertura terrestre. **Área**: Área en píxeles, por cada imagen.

De acuerdo con las características de la imagen presentada en la figura 3.9 *a*, clasificada con el método propuesto en este trabajo, el proceso ha concluido exitosamente en una única iteración, es decir, no hay necesidad de un segundo proceso de segmentación.

Similares consideraciones pueden extraerse del conjunto total de imágenes analizadas, si bien obteniendo diferente número de clases finales según el proceso de fusión realizado, que en este caso es siempre igual o inferior a ocho como corresponde al proceso de umbralización simple.

3.5. APLICACIÓN DEL CLASIFICADOR *OTSU-CLUSTERING* PROPUESTO

Con esta clasificación se pretende demostrar la alternativa y eficiencia del método de clasificación propuesto cuando se presentan más de ocho clases en una imagen a color. Los resultados se muestran a partir de la etapa de clasificación, puesto que las etapas de segmentación y codificación son exactamente igual que en el caso de umbralización simple, si bien ahora aplicando umbralización múltiple.

Para ilustrar los resultados obtenidos se seleccionó una imagen representativa del total de imágenes de cobertura terrestre que por sus características (mayores detalles visuales por la mayor resolución espacial), puedan presentar un mayor número de clases, figura 3.15.



Figura 3.15. Imagen original en modelo de color RGB, con mayor acercamiento.

En la tabla 3.6 se muestran las varianzas dentro y entre las clases obtenidas. Las varianzas dentro de cada clase (diagonal principal) son menores en los 8 casos que sus varianzas entre cada par de clase (los valores de la misma fila y columna). Esto indica que las clases se encuentran bien definidas con la simple mezcla de sus componentes espectrales binarizadas. Sin embargo, puede ocurrir que algunas clases se hayan sobre-agrupado, dado que al tener binarizada la imagen (dos clases por componente) y siendo tres componentes, al combinarlas se limita a $2^3 = 8$ clases.

Con base a este razonamiento se decide recurrir nuevamente al proceso de segmentación, si bien incrementando en uno el número de umbrales, aplicando la umbralización múltiple. Puede ocurrir que después del proceso de decisión nuevamente obtengamos como resultado ocho clases bien definidas, si bien cabe la posibilidad de encontrar un número mayor. En cualquier caso, este número se obtiene de forma automática mediante el método propuesto.

Tabla 3.6. Varianza dentro de cada clase y entre clases (figura 3.15).

C_i	1	2	3	4	5	6	7	8
1	1,65	1256,92	1465,21	1902,38	1705,04	2216,85	2808,48	6791,79
2	1256,92	73,32	338,09	265,92	345,69	318,04	746,80	2840,90
3	1465,21	338,09	38,88	256,81	234,32	444,61	341,92	2295,53
4	1902,38	265,92	256,81	83,63	322,04	232,30	312,41	1796,17
5	1705,04	345,69	234,32	322,04	142,16	337,28	436,29	2182,49
6	2216,85	318,04	444,61	232,30	337,28	144,96	502,14	1818,26
7	2808,48	746,80	341,92	312,41	436,29	502,14	31,72	1225,94
8	6791,79	2840,90	2295,53	1796,17	2182,49	1818,26	1225,94	7,00

Una vez procesados los datos de la imagen mostrada en la figura 3.15 por el método de umbralización múltiple se obtienen los umbrales mostrados en la tabla 3.7. Recordemos que con dos umbrales por componente espectral los píxeles de cada componente pueden ser asignados a tres regiones distintas. Así, para la componente roja, si el nivel de gris es menor que 121 se asigna a cada píxel una etiqueta numérica para indicar que pertenecen a una región dada; si el valor del píxel está entre 121 y 144 pertenece a otra región, y con nivel de gris mayor que 144 es del tercer grupo. De igual manera se determinan las regiones para las componentes G y B.

Tabla 3.7. Umbrales de cada componente espectral de la figura 3.15.

Núm.	Rojo (R)	Verde (G)	Azul (B)
1	121	113	103
2	144	136	106

Núm.: número de umbrales

El número posible de clases con dos umbrales y por tanto tres regiones es 27 ($3^3 = 27$). En este caso los valores promedios obtenidos para las 27 clases generadas se muestran en la tabla 3.8, y sus correspondientes varianzas dentro de cada clase y entre las mismas en la tabla A3.9, mostrada en el Anexo 1 por su tamaño. Se puede observar que las varianzas dentro de cada clase (diagonal principal), no siempre son menores que las varianzas entre cada par de clases, hecho que muestra la necesidad de reagrupar aquellas clases que presentan solapamiento hasta conseguir la mayor compacidad posible dentro de las clases (varianzas mínimas) y bien separadas (maximizando las varianzas entre las clases).

Tabla 3.8. Nivel digital promedio por clase y por componente espectral con dos umbrales.

Clases	Rojo	Verde	Azul
1	60,61	63,58	56,66
2	101,32	97,09	112,70
3	102,14	97,30	115,25
4	107,32	121,39	85,84
5	108,85	121,73	95,50
6	112,12	122,50	115,89
7	113,99	142,82	89,06
8	114,36	142,59	99,10
9	115,02	142,28	115,92
10	128,77	100,79	103,78
11	127,34	101,79	104,01
12	129,93	100,72	120,75
13	130,38	125,80	105,03
14	130,18	125,57	103,98
15	130,83	126,19	120,61
16	133,39	146,45	109,92
17	132,56	146,50	104,03
18	134,48	146,73	120,57
19	151,70	104,38	88,70
20	152,44	106,00	103,98
21	158,23	125,37	125,95

Clases	Rojo	Verde	Azul
22	153,90	126,94	92,23
23	154,96	127,14	104,06
24	158,42	127,25	125,99
25	160,62	150,52	95,94
26	161,80	151,88	104,04
27	199,28	184,43	157,29

El resultado final tras la fusión de 18 grupos fue de 9 clases, cuyos promedios y varianzas se presentan en las tablas 3.9 y 3.10.

Las figuras 3.16 y 3.17 muestran respectivamente sendas imágenes en las que aparecen los códigos numéricos asignados a cada píxel en la ubicación (x, y), así como la clase a la que dichos píxeles pertenecen.

Tabla 3.9. Nivel digital promedio por clase y componente espectral del resultado de la clasificación final.

Clases	Rojo	Verde	Azul
1	81,98	94,48	78,83
2	113,34	110,97	125,98
3	123,74	136,16	123,67
4	127,31	151,42	94,40
5	125,27	152,58	116,51
6	127,44	151,87	128,20
7	141,59	116,63	99,88
8	143,23	115,12	116,48
9	144,62	110,11	134,86

Tabla 3.10 Varianzas dentro de cada clase (diagonal principal) y entre cada par de clases (fila-columna), después de aplicar umbralización múltiple y reagrupamiento de clases solapadas.

C _i	1	2	3	4	5	6	7	8	9
1	52,93	1241,20	1913,77	1777,01	2628,19	2841,37	3777,68	3716,53	8727,24
2	1241,20	64,70	460,53	265,92	637,78	318,04	746,80	2840,90	4139,33
3	1913,77	460,53	184,00	256,81	225,92	647,11	341,92	2295,53	3281,00
4	1777,01	265,92	256,81	78,57	451,73	232,30	312,41	1796,17	3015,44
5	2628,19	637,78	225,92	451,73	210,99	425,96	296,85	2182,49	2155,11
6	2841,37	318,04	647,11	232,30	425,96	141,70	502,14	1818,26	2754,97
7	3777,68	746,80	341,92	312,41	296,85	502,14	33,19	1225,94	1331,75
8	3716,53	2840,90	2295,53	1796,17	279,97	1818,26	1225,94	9,34	1916,80
9	8727,24	4139,33	3281,00	3015,44	2155,11	2754,97	1331,75	1916,80	205,95



Figura 3.16. Clasificación de cobertura terrestre con nivel digital promedio.

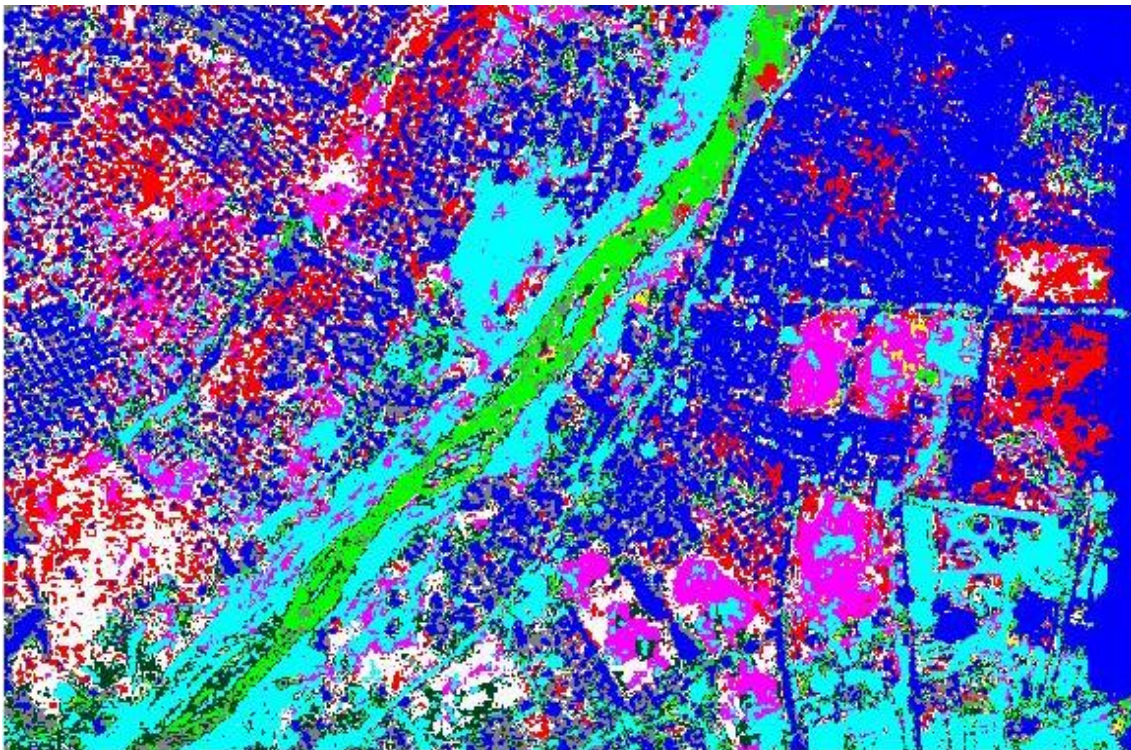


Figura 3.17. Clasificación de cobertura terrestre por asignación de colores por píxel y clase.

Similares consideraciones y resultados se han obtenido para todas las imágenes analizadas de la cuenca inundada. Dependiendo de las fotografías el número de clases varía al aplicar el método multi-umbral, que en cualquier caso se establecen de forma automática.

3.6. VALIDACIÓN DEL CLASIFICADOR

El objetivo de la validación del proceso de clasificación es medir de forma objetiva y cuantitativa la bondad o calidad del resultado obtenido. En particular interesa evaluar el grado de concordancia entre las clases asignadas a cada píxel de la imagen clasificada y un conjunto de datos de referencia (*ground truth*). Es decir, si la clasificación está de acuerdo o no con el etiquetado de la clasificación temática determinada por un experto conocedor de la zona de estudio. Para esta validación se ha aplicado uno de los métodos más extendido en cuanto a imágenes tele-detectadas, el de la matriz de errores (Congalton y Green, 2009; Lunetta y Lyon, 2004; Lewis 2004, Más y col., 2003; Couturier y col., 2008, Buendía-Rodríguez y col. 2008; Niclós y col., 2009).

Otro aspecto que también interesa evaluar es el número de agrupamientos que automáticamente genera el clasificador.

Generalmente los algoritmos de agrupamiento de naturaleza supervisada requieren como parámetro de entrada el número de clases existentes. Sin embargo, cuando se trata de clasificación de cobertura terrestre con imágenes tele-detectadas, resulta difícil conocer con precisión y antelación dicho valor. Para resolver este problema mediante algoritmos supervisados, la decisión más adecuada consiste en aplicar el algoritmo considerando un número diferente de clases cada vez. Posteriormente se evalúa la calidad de las diferentes clasificaciones obtenidas, eligiendo de entre todas ellas la mejor. La diferencia con respecto a nuestro planteamiento es que ese número se determina de forma automática. En ambos casos la evaluación se realiza mediante los denominados índices de validación, descritos en la sección 2.6.2.

3.6.1. VALIDACIÓN TEMÁTICA MEDIANTE LA MATRIZ DE ERROR

Una matriz de error expresa mediante sus filas el número de unidades (es decir, píxeles o grupos de píxeles) asignados a una categoría determinada por un proceso de clasificación, frente a las unidades asignadas a esa misma categoría mediante otro

proceso de clasificación distinto, que es considerado como datos de referencia o *ground truth*, en este caso expresado en las columnas. La comparación entre ambos, si se analiza de manera tradicional, produce el número de muestras bien o mal clasificadas. A este tipo de matriz de error, en la cual se tienen dos opciones de calificación (bien y mal), la vamos a identificar a partir de ahora como matriz de error tradicional (MET).

Congalton y Green (1993 y 2009) recomiendan utilizar la matriz de error como punto de partida para la identificación de fuentes de error, y no sólo para calcular el valor del error en la clasificación. Por ejemplo, la variación de la interpretación humana puede tener un impacto significativo en lo que se considera correcto y lo que no lo es. Si la interpretación fotográfica se utiliza como referencia en una evaluación y esa interpretación no es completamente correcta, entonces los resultados van a ser engañosos. Lo mismo ocurre si se comparan las observaciones en tierra frente a las medidas del suelo.

Con el fin de tratar esa ambigüedad, Gopal y Woodcock (1994), así como Congalton y Green (2010), propusieron el uso de los conjuntos borrosos para “permitir el reconocimiento explícito de la posibilidad de que la ambigüedad puede existir en relación con la etiqueta de la clasificación apropiada”. En este enfoque además de correcto (acuerdo) e incorrecto (desacuerdo), pueden incluirse una variedad de respuestas tales como aceptable, comprensible pero equivocada, etc. Congalton y Green (1999, 2004, 2009) presentan una metodología de evaluación difusa que no sólo se ocupa de la ambigüedad sino que también permite que los resultados de la evaluación se presenten en una matriz de error más completa, a la que denominaremos “Matriz de Error *Fuzzy*” (MEF).

La MEF, al igual que la MET, es una matriz cuadrada de números en la cual las filas representan las categorías determinadas por el clasificador, y las columnas los datos de referencia. En la MEF se incorporan, junto a los valores de bien o mal clasificado, alternativas intermedias para dar cabida a los píxeles mixtos o frontera.

La MEF proporciona una información más completa para validar las etiquetas temáticas, a partir de la cual se puede derivar tanto la evaluación determinista como la difusa.

Para aplicar dicho método de validación del clasificador se definieron los datos de referencia con la colaboración directa de un experto conocedor de la zona de estudio (Villegas, 2006).

3.6.2. UNIDADES DE VALIDACIÓN

La mayoría de las revisiones bibliográficas se inclinan por seleccionar sólo una unidad (píxel) para la validación. Sin embargo en el presente trabajo se han utilizado dos tipos de unidades de validación o muestreo de distinto tamaño, el píxel y el polígono, para facilitar la comprobación.

Se seleccionó el píxel como unidad de validación ya que el clasificador utiliza éste como característica principal. Además esta unidad de medida es muy adecuada ya que al superponer dos imágenes geográficamente ubicadas (Georreferenciadas mediante el sistema de coordenadas UTM), visualmente es fácil identificar si un píxel que en la imagen de referencia (original) pertenece a una clase temática coincide con la misma asignada en la clasificación. Obsérvese un claro ejemplo en los cuerpos de agua (píxeles 50-54) de la figura 3.18.

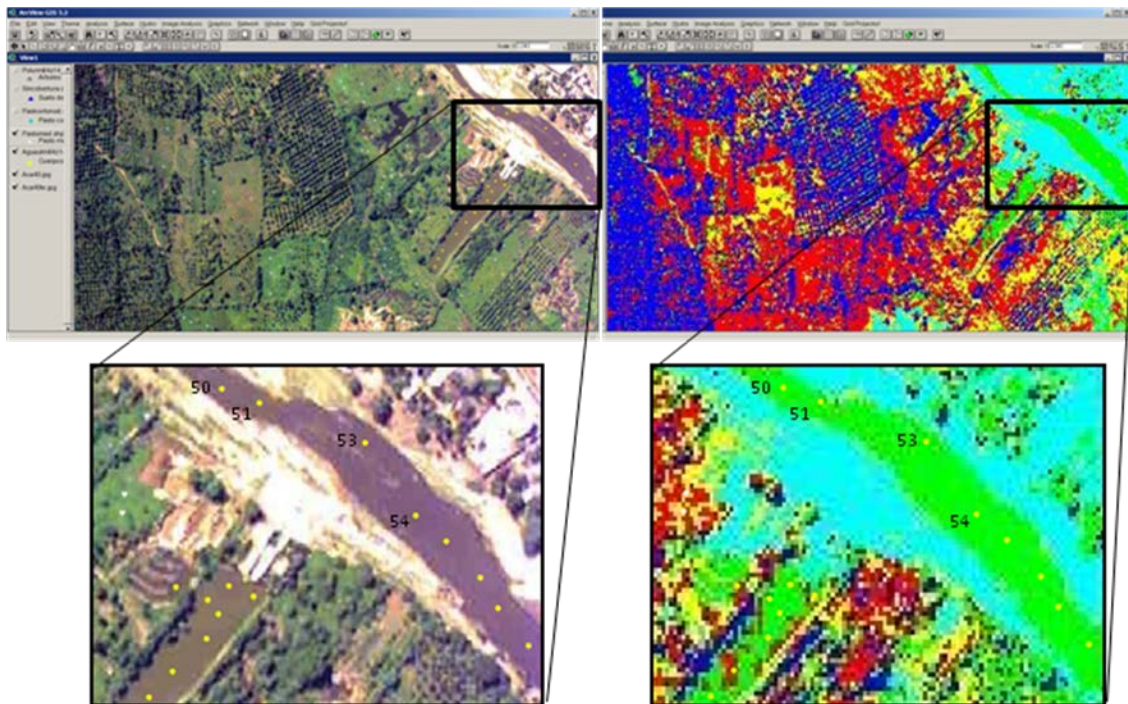


Figura 3.18. Ejemplo de comparación entre la imagen de referencia y la imagen clasificada a nivel de píxel.

Por otro lado, en algunos casos es necesario realizar visitas de campo para asegurarse que efectivamente la clase temática que se asigna a un objeto en la fotografía aérea es correcta. Pero es muy difícil conocer con exactitud dónde se sitúa ese píxel en la imagen original. Por lo tanto, utilizar un único píxel como unidad de muestreo puede ocasionar que muchos de los errores representados en la matriz de error sean errores de posición en lugar de errores temáticos. Para solventar esto se identifica el objeto en el cual está situado el píxel, trazándose a continuación un polígono con diferentes formas y tamaños que se han utilizado también grupos de píxeles como unidades de validación. Así existen menos posibilidades de cometer un error posición. En la figura 3.19 se muestra un ejemplo de esta estrategia junto con las diferentes ubicaciones de los polígonos considerados.

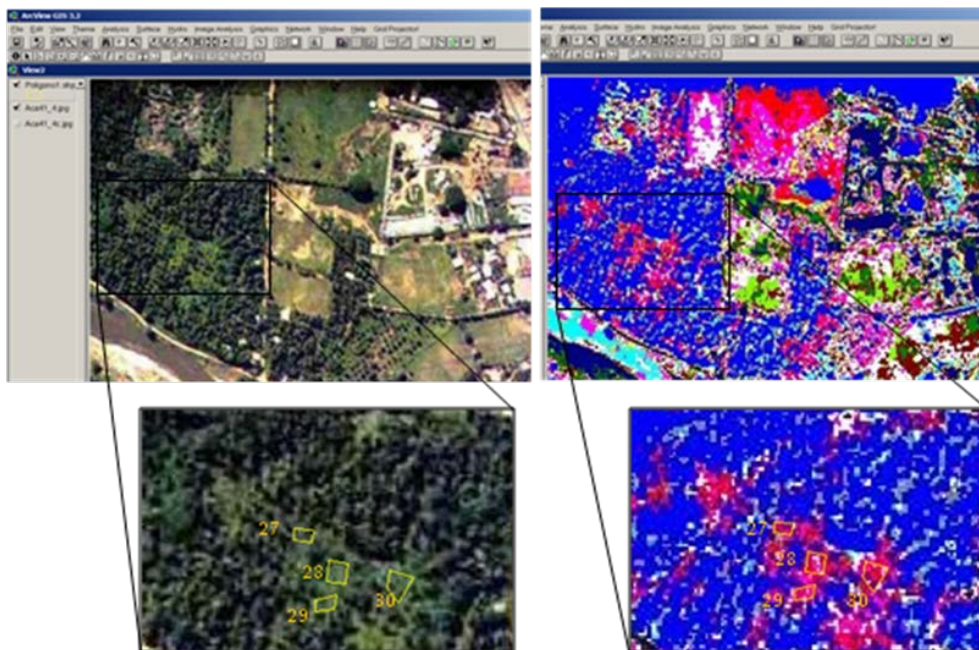


Figura 3.19. Ejemplo de comparación entre imagen de referencia e imagen clasificada tomando los polígonos como unidades de validación.

Otra cuestión importante, además del tamaño de la muestra, es determinar cuántas unidades de validación son necesarias. Existen diferentes trabajos de investigación sobre este tema, como el de Medina (1998). Algunos autores sostienen que el mínimo debe ser 50 muestras, otros que dependen del espacio físico o geográfico que contiene la imagen a clasificar, proponiéndose diferentes metodologías de cálculo para este número.

En el presente estudio decidimos adoptar la propuesta de Congalton y Green (2009), aplicando la distribución multinomial detallada en las secciones 2.6.3 y 4.4.2.1.

Siguiendo la referencia anterior, se obtuvieron 6180 unidades de muestreo para validar el clasificador cuando la unidad es el píxel (tabla 3.11) y 819 muestras cuando se usan polígonos (tabla 3.12).

Tabla 3.11. Distribución del número de muestras (unidad de validación el píxel) por categoría temática.

Clase (uso del suelo)	Pertenencia por clase %	No. muestras
Vegetación arbórea (agricultura perene)	37	2270
Cuerpos de agua	9	540
Pasto de altura media (agricultura de ciclo corto)	28	1710
Pastos corto poco denso (agricultura de ciclo corto)	14	880
Zonas sin cobertura vegetal	13	780
Total	100	6180

Tabla 3.12. Distribución del número de muestras (unidad de validación el polígono) por categoría temática.

Clase (uso del suelo)	Pertenencia por clase	No. muestras
Árboles (Agricultura perene)	32	260
Pasto seco (agricultura de ciclo corto)	4	39
Pasto de altura media (agricultura de ciclo corto)	3	33
Cuerpos de agua	2	24
Pastos de alto (agricultura de ciclo corto)	7	72
Pastos cortos y densos (agricultura de ciclo corto)	37	243
Pasto corto poco denso (agricultura de ciclo corto)	6	63
Montículos de arena húmeda (arrastre de sedimento)	1	10
Zonas sin cobertura vegetal	8	75
Total	100	819

La obtención de las muestras para la validación fue la denominada *aleatoria estratificada*, es decir, para cada una de las categorías temáticas (clases), dado un número de muestras a extraer, éstas se seleccionan de manera aleatoria.

Se usó el software diseñado para sistemas de información geográficos ArcView (2012), versión 3.2, de manera que al tener geo-posicionadas todas las imágenes originales y las que resultan de la clasificación, al mostrarlas en pantalla quedarán totalmente solapadas entre sí.

Con este software al realizar el zoom (acercamiento de la imagen) se conoce exactamente la unidad de medida en tierra que se selecciona en cada píxel o polígono.

3.6.3. OBTENCIÓN DE LAS MATRICES DE VALIDACIÓN

Para el análisis de los resultados derivados de este trabajo de investigación se decidió usar la MEF porque, además de lo expuesto previamente (sección 3.6.1), permite realizar el análisis global y particular para cada una de las clases obtenidas. Además las características de esta matriz permiten desarrollar el análisis tanto desde el punto de vista *fuzzy* como determinista, dado que se cuenta con toda la información necesaria.

Cada celda de la MEF que no forma parte de la diagonal principal contiene dos valores separados por punto y coma [;]. No se aplica la separación con comas tradicional para evitar confusiones con la representación numérica decimal. El primer valor representa las etiquetas que, aunque sin ser absolutamente correctas, se consideran aceptables según el criterio de flexibilidad difuso establecido mediante alguna regla de naturaleza difusa. El segundo valor indica las etiquetas que son inaceptables (incorrectas). La diagonal principal contiene sólo un valor, el que se considera absolutamente correcto.

Para la evaluación de cada unidad de validación se estableció un conjunto de reglas difusas respecto a las etiquetas asignadas, que se sintetizan como sigue:

- **Absolutamente correctas:** se anota en la diagonal principal, cuando las unidades de validación seleccionadas en la imagen de referencia de una clase particular se superponen en un 100% con la misma clase temática en la imagen clasificada.
- **Aceptable:** primer valor en las celdas no diagonales, cuando existe un 50% o más de solapamiento entre las unidades marcadas en la imagen de referencia para una clase en particular, y la misma clase temática en la imagen clasificada. Se considera aceptable ya sea porque dichos píxeles se ubican entre dos clases o porque el tipo de textura está mezclada y la complejidad en la imagen de referencia no permite establecer los límites exactos.
- **Error** (error inaceptable): se anota a la derecha del punto y coma en las celdas no diagonales (segundo valor de la celda), cuando las muestras seleccionadas en la imagen de referencia para una clase determinada no se solapan con esa clase en la imagen clasificada, es decir, se ubican en otra clase temática o el solapamiento entre esas

unidades de validación en ambas imágenes es menor del 50%. Por lo tanto la clase de la muestra de la imagen clasificada no coincide con la de referencia.

Aplicando estas reglas se elaboró la MEF mostrada en la tabla 3.13 (píxel como unidad de validación) y tabla 3.14 (polígono como unidad de validación). Se obtuvieron trabajando con 10 de las 64 imágenes clasificadas en el primer caso, y con 3 de las 16 imágenes en el segundo.

Los detalles sobre este tipo de matrices y su construcción están detallados en la sección 2.6.4.

Tabla 3.13. Matriz de error *fuzzy* (MEF), las muestras de validación son los píxeles clasificados bajo umbralización simple.

REFERENCIA							Precisión del usuario				
							DETERMINISTA		FUZZY		
CLASIFICACION							Totales Bien		Totales Bien		
	Arb	Cag	Pam	Pcp	Zsc	RDC	muest.	(%)	muest.	(%)	
	Arb	2270	0;20	20;50	0;0	0;0	2360	2270.00	96%	2290.00	97%
	Cag	0;0	430	20;30	70,70	0;0	620	430.00	69%	520.00	84%
	Pam	0;0	20;60	1570	30;20	0;0	1700	1570.00	92%	1620.00	95%
	Pcp	0;0	0;0	10;10	690	0;0	710	690.00	97%	700.00	99%
	Zsc	0;0	10;0	0;0	0;0	780	790	780.00	99%	790.00	100%
TDR	2270	540	1710	880	780	6180	5740.00		5920.00		

Precisión del productor

	Total bien	Arb	Cag	Pam	Pcp	Zsc
		iDetermin.	2270 100%	430 80%	1570 92%	690 78%
Fuzzy		2270 100%	460 85%	1620 95%	790 90%	780 100%

PRECISIÓN GENERAL			
DETERMINISTA		FUZZY	
5740 / 6180	93%	5920 / 6180	96%
Kappa	90%		94%

Arb: Árboles
 Cag: Cuerpos de agua.
 Pam: Pasto de altura media.
 Pcp: Pasto corto poco denso .

Zsc: Zonas sin cobertura vegetal.
 TDR: Total de datos de referencia.
 TDC: Total de datos clasificados.

Tabla 3.14. MEF con las muestras como polígonos y clasificado bajo umbralización múltiple.

		REFERENCIA										Precisión del clasificador			
		Arb	Pse	Pal	Cag	Pme	Pcd	Pcp	Mah	Zsc	TPC	DETERMINISTA		FUZZY	
											Total	Porcien	Total	Porcien	
CLASIFICACIÓN	Arb	260	0;0	3;3	3;0	0;0	0;0	0;0	0;0	0;0	269	260	97%	266.00	99%
	Pse	0;0	33	0;0	0;0	9;3	3;3	0;0	0;0	0;0	51	33	65%	45.00	88%
	Pal	0;0	0;0	15	0;0	0;0	0;0	0;0	0;0	0;0	15	15	100%	15.00	100%
	Cag	0;0	0;0	0;0	21	0;0	0;0	0;3	1;0	0;0	25	21	84%	22.00	88%
	Pme	0;0	0;0	0;0	0;0	45	15;6	3;0	0;0	0;0	69	45	65%	63.00	91%
	Pcd	0;0	0;0	9;3	0;0	15;0	210	15;12	0;0	0;0	264	210	80%	249.00	94%
	Pcp	0;0	0;0	0;0	0;0	0;0	3;0	30	0;0	0;0	33	30	91%	33.00	100%
	Mah	0;0	0;0	0;0	0;0	0;0	0;3	0;0	9	0;0	12	9	75%	9.00	75%
	Zsc	0;0	6;0	0;0	0;0	0;0	0;0	0;0	0;0	75	81	75	93%	81.00	100%
	TPR	260	39	33	24	72	243	63	10	75	819	698		783	

		Precisión del productor (experto)													
Determ.	Total bien	260.00	33.00	15.00	21.00	45.00	210.00	30.00	9.00	75.00					
		100%	85%	45%	88%	63%	86%	48%	90%	100%					
Fuzzy	Total bien	260.00	39.00	27.00	24.00	69.00	231.00	48.00	10.00	75.00					
		100%	100%	82%	100%	96%	95%	76%	100%	100%					

PRECISIÓN GENERAL			
DETERMINISTA		FUZZY	
698/819	85%	783/819	96%
Kappa	81%		94%

Arb: Árboles
 Pse: Pasto seco
 Pal: Pasto alto
 Arb: Árboles

Cag: Cuerpos de agua
 Pme: Pastos mediano
 Pcd: Pastos cortos y densos
 Pcp: Pasto corto poco denso

Mah: Montículos de arena húmeda
 Zsc: Zonas sin cobertura vegetal
 TPL: Total de polígonos clasificados
 TPR: Total de polígonos de referencia

3.6.3.1. ANÁLISIS DE LAS MATRICES DE ERROR

Las imágenes de referencia seleccionadas para la evaluación presentan cinco clases de uso de cobertura terrestre, tabla 3.13; es decir, corresponden a la clasificación de imágenes con menos de 8 clases obtenidas con umbralización simple. A la vista de los resultados mostrados en las tablas se puede concluir lo siguiente:

- La clase que obtiene la mejor precisión del clasificador es la de zonas sin cobertura vegetal puesto que la determinista es del 99% y la *fuzzy* del 100%. Dentro de esta clase se consideran los sedimentos acumulados sobre la ribera del río por efecto del huracán Paulina, así como las calles y los edificios.
- El segundo lugar, en cuanto a precisión del clasificador, se encuentra el pasto corto poco denso, obteniendo una precisión determinista del 97% y *fuzzy* del 99%.
- La cobertura arbórea también presenta una buena precisión, del 96% y 97% para los casos determinista y difuso respectivamente.
- Con buena precisión se clasifican los pastos de altura mediana (92% y 95% respectivamente para el caso tradicional y el difuso).
- Dentro de las clases de cobertura terrestre la que obtiene la precisión de clasificación más baja es la de cuerpos de agua, del 69% en el caso determinista y 84% en el difuso. En la tradicional resulta ser muy baja debido probablemente a que se trata de escurrimientos causados por el huracán, lo que significa que son aguas poco profundas en las cuales, en parte, aparecen montículos de arena y en ocasiones hasta un poco de pasto. Por lo tanto, con esos datos de referencia es difícil definir claramente lo que sería realmente agua debido a la similitud de los niveles de intensidad. La clasificación de los cuerpos de agua no resulta ser un problema exclusivo de este algoritmo. Efectivamente, Zhu y col. (2000) señalan ya numerosas dificultades para diferenciar agua y bosque usando imágenes aéreas. Polidório y col. (2003) y Rowe y Grewe (2001) reportan el mismo problema al separarlos de las sombras; del mismo modo que Saeed (2006) lo indica al clasificar cuerpos de agua poco profundos.
- En resumen, podemos afirmar que el clasificador propuesto tiene una precisión global muy buena dado que la *precisión general* del modelo de clasificación

para imágenes con menos de 8 clases es del 93% desde la perspectiva del análisis determinista y del 96% para el difuso.

Los resultados presentados en la tabla 3.14 corresponden a la clasificación de imágenes con más de 8 clases de cobertura, realizada por umbralización múltiple. Las conclusiones de su análisis son las siguientes:

- Las clases que obtienen una mejor precisión del clasificador son los pastos altos.
- En segundo lugar se encuentra la cobertura arbórea con el 97% de precisión determinista y el 99% *fuzzy*.
- Las zonas sin cobertura vegetal y los pastos cortos poco densos también obtienen un mínimo error, dado que alcanzan una precisión tradicional del 93% y 91% respectivamente, y una precisión *fuzzy* del 100%.
- Los cuerpos de agua y los pastos cortos densos obtienen un error aceptable, con una precisión mayor que el 80%.
- Sin embargo los pastos secos, pasto mediano y los montículos de arena húmeda presentan una precisión determinista menor que el 80% ya que las componentes espectrales de los pastos y la humedad son muy parecidos. Si consideramos aquellos píxeles que se encuentran en los límites de una manera más suave (*fuzzy*), la precisión ya es aceptable para los pastos, mayor que el 88%.
- En resumen, para las imágenes clasificadas con más de 8 clases se obtiene una precisión general determinista del 85% y *fuzzy* del 96%.

De acuerdo con Anderson y col. (1976), un sistema de clasificación eficiente debe tener una precisión general de al menos el 85%. Esta norma es una de las diez que fueron escritas para asegurar la validez de los sistemas de clasificación de cobertura terrestre, resultando ser una de las más aceptadas universalmente (Martín-H y col., 2010). Esto significa que el clasificador aquí propuesto cumple dicha norma, por lo que en definitiva puede considerarse un buen clasificador para análisis de cobertura terrestre a través de imágenes aéreas.

3.6.4. VALIDACIÓN DEL NÚMERO DE CLASES

Como se ha indicado reiteradamente, el clasificador propuesto es de naturaleza no supervisada y no requiere ningún parámetro de entrada, de manera que automáticamente realiza las agrupaciones que considera adecuadas, validándolas durante el mismo

proceso de clasificación mediante la maximización de la varianza entre clases a la vez que minimiza la varianza dentro de clases.

Los índices de validación interna, al igual que los algoritmos de agrupamiento, generan diferentes resultados según la información que procesen. Por tanto, un índice de validación puede ser el mejor para un algoritmo de clasificación y un conjunto de datos determinado, pero no es posible generalizarlo para todos los algoritmos y aplicaciones. De manera que para evaluar el clasificador propuesto, identificado como *Otsu-clustering*, hemos seleccionado cinco índices de validación (DB, CH, Dunn, KL, Han), cuyos resultados se muestran en la tabla 3.15.

Los índices de validación para determinar el número óptimo de clases son generalmente utilizados para clasificadores no supervisados en los cuales uno de sus parámetros de entrada es el número de grupos a extraer, como se ha indicado en la sección 2.6.2.

Dado que nuestro clasificador no requiere ese parámetro de entrada, tampoco permitirá ir probando con un mínimo y un máximo número de clases para medir cuál es el óptimo. Para poderlo comparar ha sido necesario seleccionar un clasificador no supervisado que proporcione resultados similares al propuesto, donde sí haya que especificar el número de clases. Se ha optado por el método *Fuzzy C-Means* (FCM) (Bezdek, 1981; Dunn, 1993) por su probada eficiencia en numerosos ámbitos de aplicación, además de su naturaleza difusa que encaja bien con los planteamientos formulados en este trabajo respecto de la MEF.

Estos índices se calculan aplicando varias veces el algoritmo de agrupamiento, variando el número de clases en cada ejecución de un cierto valor mínimo a un máximo. Así, de las 80 imágenes clasificadas se seleccionaron un 11% de manera aleatoria, clasificando cada una de ellas mediante el FCM, y partiendo desde 2 hasta un máximo de 8 clases para compararlas con aquellas en las que nuestro clasificador daba menos de 8, y desde 2 hasta 18 para casos con más de 8 clases.

La tabla 3.15 del Anexo 2 muestra los resultados obtenidos para los agrupamientos realizados sobre nueve muestras, según los cinco índices de validación internos (DB, CH, Dunn, KL, Han), para las veintisiete clases que corresponden a esta situación. Las muestras M1 a M7 son imágenes con un número de clases menor que ocho, y sobre ellas se ha aplicado el método de umbralización simple. Por el contrario, las muestras

M8 y M9 corresponden a imágenes sobre las que se ha aplicado el método multi-umbral con dos umbrales por banda espectral.

La tabla 3.16 muestra una síntesis de los resultados mostrados en la tabla 3.15, en la que aparece el número de clases óptimo obtenido mediante el cálculo de los cinco índices de validación interna. La última columna presenta el número de clases generadas por el clasificador Otsu-Clustering propuesto. Cuando alguno de los números óptimos de clases de algún método coincide con el número generado por nuestro clasificador, se marca en negrita y se agrega un asterisco a su derecha.

Tabla 3.16. Número óptimo de agrupamientos según los cinco índices de validación.

Muestras	<i>Fuzzy-Clustering</i>					Otsu-Clustering
	DB	CH	Dunn	KL	Han	
M1	4	5*	3	2	6	5
M2	3*	6	3*	2	6	3
M3	5*	6	2	2	6	5
M4	6*	5	2	2	6*	6
M5	5	6*	2	2	6*	6
M6	3*	6	2	2	6	3
M7	4*	6	2	2	6	4
M8	10	13	11	8	16*	16
M9	3	6	3	3	13*	13

DB: Índice de Davies Bouldin (1979)

KL: Índice de Krzanowski y Lai (1985)

CH: Índice de Calinski y Harabasz (1974)

Han: Índice de Hartigan (1985).

Dunn: Índice de Dunn (1974).

M1...M9: Muestra 1... Muestra 9

Observamos que siempre existe correspondencia al menos con un índice, siendo para el DB con el que se obtiene la máxima concordancia en el caso de umbral simple, y con Han para multi-umbral. Por tanto, en base a los resultados obtenidos por los cinco índices de validación, se puede concluir que el número de agrupamientos generado por nuestro clasificador puede considerarse apropiado. Una observación final resulta del hecho de que cada muestra evaluada se respalda con al menos uno o dos de los índices internos calculados para tal fin y con el método Otsu-Clustering.

En la figura 3.20 se muestra, a modo de ejemplo, una comparación visual sobre los resultados de la clasificación realizada mediante el algoritmo FCM (*a, d*) y el Otsu-Clustering (*c, f*) para las imágenes originales mostradas en (*b, e*).

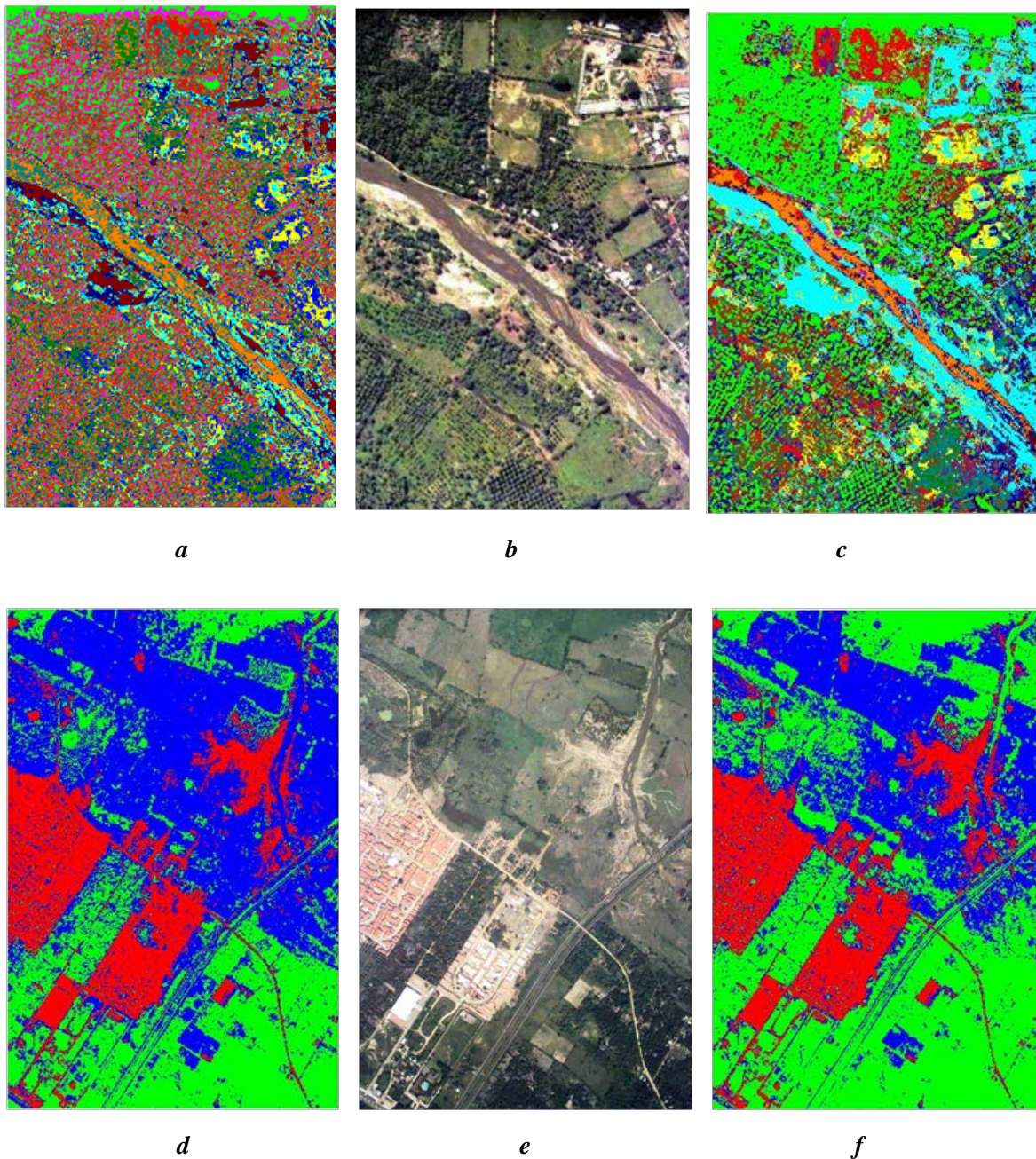


Figura 3.20. Comparación visual de la clasificación obtenida por *Otsu-Clustering* y *Fuzzy C-Means*: **(b, e)** imágenes originales; **(a)** clasificación de **b** mediante *Fuzzy C-Means* para nueve clases, **(c)** clasificación de **b** mediante *Otsu-Clustering* para nueve clases; **(d)** clasificación de **e** mediante *Fuzzy C-Means* para tres clases, **(f)** clasificación de **e** mediante *Otsu-Clustering* para tres clases.

En el caso de la clasificación de la imagen original de la figura 3.20 **b**, el método de *Otsu-Clustering* obtiene 9 clases, las mismas que se determinan para la clasificación con el algoritmo *FCM*. En este caso se observa cómo el método *Otsu-Clustering* realiza una clasificación más nítida, de forma que las estructuras que cubren el suelo se aprecian con mucha más precisión, identificándose claramente tanto en la imagen clasificada

como en la original, algo que no ocurre con los resultados mostrados en la figura 3.20 *a* obtenidos mediante *Fuzzy C-Means*. Para la imagen mostrada en la figura 3.20 *e*, el método de Otsu-Clustering obtiene 3 clases de cobertura, valor que se asigna como parámetro de entrada para la clasificación por el algoritmo FCM; en este caso se aprecia un alto grado de similitud respecto de los resultados de ambos clasificadores.

3.7. CONCLUSIONES DEL CAPÍTULO

Tras el procesamiento y la clasificación realizados sobre imágenes en color de cobertura terrestre se pueden extraer las siguientes conclusiones:

- El método de Otsu (1979) mediante su técnica de umbralización adaptativa resulta ser altamente eficiente para separar las dos partes de que se compone cada banda espectral. Al aplicarlo a imágenes a color, con tres componentes espectrales, se generarán hasta ocho clases.
- Mediante el mismo criterio de Otsu es posible separar más de dos áreas entre sí en cada banda espectral. Aplicando multi-umbralización se pueden obtener hasta un total de n^3 grupos diferenciados para las n bandas espectrales, siendo n igual a tres en el caso de las imágenes en color analizadas.
- La eficiencia del método de clasificación propuesto, denominado Otsu-Clustering, se verifica mediante una serie de índices de validación y de un reconocido modelo de clasificación de imágenes a color *Fuzzy C-means*, que permiten comprobar si el número de clases obtenido de forma no supervisada es el correcto.
- Por otro lado, también es posible validar la eficacia del clasificador mediante la Matriz de Error *Fuzzy*, con la que se obtienen resultados muy satisfactorios tanto para umbralización simple como múltiple.
- Un inconveniente de cualquier método de clasificación no supervisada, y por tanto del propuesto en esta investigación, deriva del hecho de que siempre se requiere la concurrencia de un experto para etiquetar temáticamente las clases generadas por el clasificador, de tal forma que en algún momento podría inducir algún tipo de error. Este inconveniente se puede ver compensado por el hecho de la automatización, que permite identificar clases sin un conocimiento previo sobre las mismas.

UNIVERSIDAD COMPLUTENSE DE MADRID, ESPAÑA

Capítulo 4

Aplicación de estrategias de procesamiento y clasificación a imágenes agrícolas captadas por cámaras convencionales





CAPÍTULO 4 APLICACIÓN DE ESTRATEGIAS DE PROCESAMIENTO Y CLASIFICACIÓN A IMÁGENES AGRÍCOLAS CAPTADAS POR CÁMARAS CONVENCIONALES.

4.1. INTRODUCCIÓN

El objetivo de este capítulo consiste en aplicar los métodos y procedimientos propuestos en este trabajo de investigación para clasificar la cobertura terrestre de imágenes digitales obtenidas por cámaras convencionales, directamente a nivel del suelo, sin necesidad de recurrir a plataformas aéreas o de satélites. En concreto, se analizan imágenes de campos de cultivo de avena que han sido afectados por fuertes heladas. El tratamiento permite cuantificar la densidad tanto del cultivo afectado por las bajas temperaturas como de la parte no afectada.

La estimación de los daños es un reto importante debido a que permitirá una predicción más precisa de la producción de cultivos, y por lo tanto posibilita la planificación en base al rendimiento estimado. En todo caso, la cuantificación de los daños es muy útil ya que permite el registro de datos e información para futuros pronósticos.

El procedimiento general se describe en el esquema mostrado en figura 4.1, cuyos detalles se describen a lo largo del capítulo.

El fundamento consiste en aplicar una estrategia de clasificación no supervisada ya que las imágenes pueden presentar componentes espectrales muy distintos, debido a los diferentes factores que pueden ocasionar daños en los cultivos de avena. Además, los enfoques supervisados no pueden ser entrenados con los patrones de textura apropiados ya que no se conocen de antemano.

El sistema de cuantificación de la densidad de plantas afectadas por heladas se desarrolla de la siguiente manera:

- *Recopilación de la información*, la población total del cultivo de avena corresponde a una superficie total de 20 hectáreas. En esta etapa se describen las características principales de las imágenes digitales utilizadas y la manera en que éstas fueron obtenidas, así como el método de muestreo utilizado.
- *Pre-procesado de la imagen original*, proceso en el cual se describe de manera general cómo se preparan las imágenes que originalmente se encuentran en el

modelo de color RGB, para su segmentación y clasificación. En concreto, la transformación de los datos originales (RGB) al modelo de color ($L^*a^*b^*$), y la normalización de los valores de intensidad.

- *Segmentación*, consiste en la combinación de tres algoritmos de umbralización, los cuales se aplican de manera sucesiva a cada una de las bandas espectrales del espacio de color, L^* , a^* y b^* , produciendo la partición de los correspondientes histogramas en varias regiones.
- *Codificación*, las regiones obtenidas de la combinación de umbrales son codificadas por componente espectral, y a partir de los códigos individuales se genera el código global de clases de la imagen en color.
- *Algoritmo de clasificación*, se aplica el mismo algoritmo de decisión expuesto en el capítulo tres para la clasificación de las fotografías aéreas.
- *Validación*, que permite la verificación de los resultados de clasificación mediante la comparación con los criterios establecidos a tal fin.

Por tanto, en el presente capítulo se desarrollan cada una de las etapas mencionadas. Posteriormente se presentan los resultados de la clasificación, su validación y la discusión relativa a dichos resultados junto con las conclusiones pertinentes extraídas.

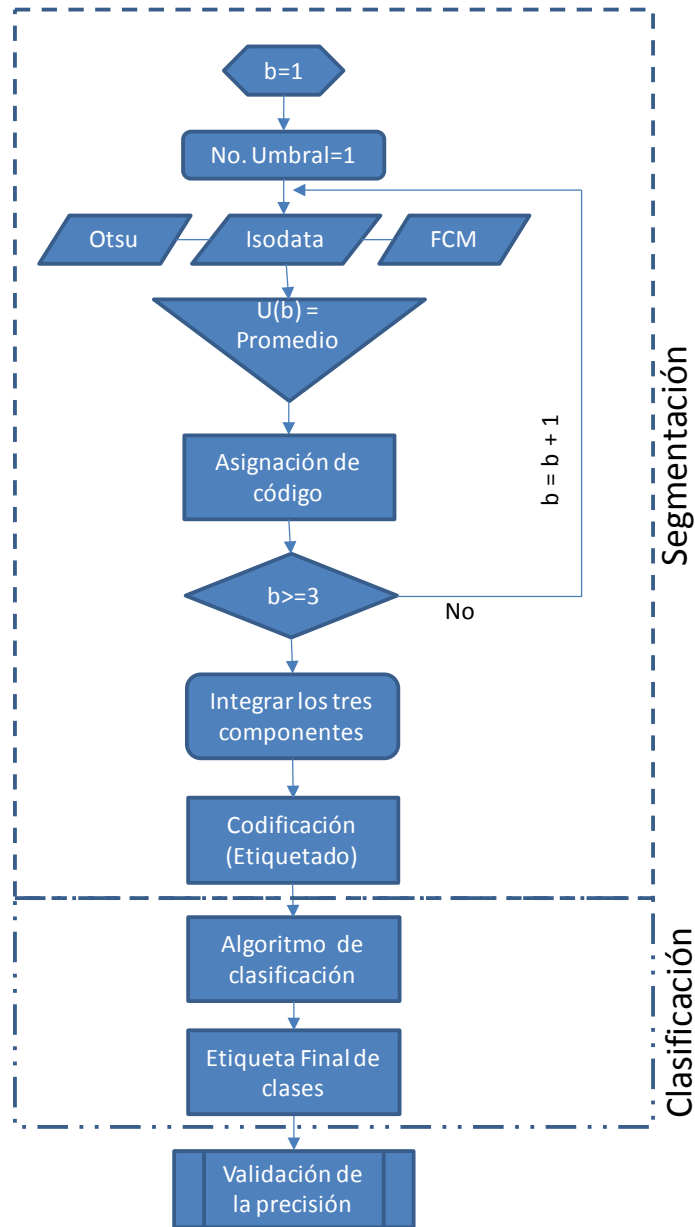


Figura 4.1. Diagrama del sistema de clasificación no supervisada de imágenes a color por combinación de umbrales y reagrupamiento de clases.

4.2. RECOPIACIÓN Y PRE-PROCESAMIENTO DE LA INFORMACIÓN

Las imágenes utilizadas son fotografías de un cultivo de avena afectado por las bajas temperaturas (heladas) que tuvieron lugar entre los meses de octubre y noviembre de 2010, en México. Las fotografías digitales tomadas con cámaras convencionales fueron captados 15 días después de la última helada, durante dos días consecutivos, tratando de evitar avances significativos en el daño y por tanto considerando exclusivamente el daño producido por dicho agente meteorológico. Las condiciones climáticas de estos

dos días eran diferentes, uno se correspondía con un día soleado y el otro nublado. Además las muestras fueron adquiridas a diferentes horas del día (por la mañana, al mediodía y por la tarde), con alta variabilidad de las condiciones de iluminación natural. Por tanto, estas situaciones permiten la evaluación del método bajo condiciones ambientales diferentes, principalmente en la iluminación recibida por las cámaras durante la captura de las imágenes.

De acuerdo con Webb y col. (1994), el grado de tolerancia a las heladas mostrado por una planta depende en gran medida de la etapa de desarrollo en la que se encuentre. Antes del inicio de la floración, por lo general de 8 a 10 semanas después de la germinación, las plantas de cereales son capaces de soportar el frío extremo. Sin embargo, la etapa de desarrollo más susceptible a daños por heladas es el período anterior a la floración, después de la emergencia de la espiga. En resumen, los efectos de las heladas sobre los cereales son más relevantes durante y después de la emergencia de la espiga, la planta se vuelve muy sensible, y las heladas intensas a menudo causan severos daños tanto en el tallo como en las hojas, llegando a secar la planta.

En México, en el año 2010 las heladas se presentaron entre octubre y diciembre (de 10 a 60 días), con temperaturas mínimas tan bajas como $-13\text{ }^{\circ}\text{C}$, dañando considerablemente los cultivos de avena. Los daños por congelación de las hojas se suelen producir en el rango de $-2\text{ }^{\circ}\text{C}$ a $-4\text{ }^{\circ}\text{C}$.

Uno de los objetivos a conseguir como resultado del análisis de dichas imágenes es la cuantificación de la densidad de plantas afectadas por las bajas temperaturas, tanto para determinar el daño producido en los cultivos como para tomar medidas preventivas en situaciones futuras similares.

De forma más específica, tal y como se ha comentado previamente, además con este tipo de imágenes también se pretende verificar la robustez del método de clasificación propuesto frente a la variabilidad de la iluminación.

Las fotografías fueron captadas con un sensor de imagen convencional (Canon digital IXUS 85 IS, sensor: CCD 1/2.3", longitud focal: 35 mm). La resolución de las imágenes es de 3.648×2.736 píxeles. El sensor ha sido calibrado con anterioridad para la estimación de los parámetros intrínsecos (distancia focal y distorsión radial), y extrínsecos (matriz de desplazamientos y ángulos de rotación) (Tsai, 1987). Esto es

necesario para determinar la superficie real fotografiada con el fin de poder estimar la superficie total del campo de cultivo.

Una de las ventajas de este tipo de imágenes es el relativamente bajo coste de la cámara comparado con los dispositivos aerotransportados. Además, con este tipo de fotografías el estudio puede realizarse en distintos ámbitos tales como agricultura de precisión, selección de cosecha, clasificación de frutas, etc.; otra gran ventaja es que la resolución de estos sensores es alta, lo cual permite efectuar mediciones con cierta precisión.

Sin embargo se tiene la desventaja de que con dichas imágenes la cobertura del terreno es muy baja y, por lo tanto, para el caso agrícola debe realizarse un muestreo adecuado para que los resultados puedan generalizarse. Otra desventaja es que para zonas de difícil acceso las muestras no se pueden tomar con la facilidad deseada, o simplemente no se pueden tomar.

4.2.1. MUESTREO DE LA ZONA DE ESTUDIO

Con el fin de obtener un conjunto de fotografías representativas de la extensión de la zona de estudio, de manera que abarquen la población total, se realizó un muestreo fotográfico sobre el campo del cultivo de avena. Como se ha comentado en el capítulo dos existen diversos métodos de muestreo en la literatura. En la investigación descrita en este trabajo se aplicó el denominado “aleatorio estratificado” de la siguiente manera: la población total del cultivo de avena se dividió en dos regiones, según el grado de afectación. La primera región, muy afectada por las heladas (prácticamente seca), con la típica tonalidad amarillenta; la segunda, menos afectada, con tonalidades situadas entre el verde y el amarillo, con predominio de la tendencia hacia los verdes. Para cada una de estas regiones se aplicó un muestreo aleatorio que consistió en tomar 100 muestras/ha (50 imágenes de las zonas más afectadas y 50 del área poco afectada), de una superficie total de 20 hectáreas; es decir, se tomaron 2000 muestras.

Para delimitar las muestras se construyó un dispositivo formado por cuatro postes de madera situados formando un cuadrado, que delimita una superficie de 1 m², y que se sitúan sobre el terreno para delimitar el espacio a fotografiar. Cada muestra se midió y se registró convenientemente, como se puede observar en la figura 4.2 *a*. Una vez que se ha seleccionado la muestra del campo, se fotografía con la cámara que se coloca en posición cenital, a una altura de 1.5 m. En la figura 4.2 *b* se observa un área de avena

sin daños. La figura 4.2 *c* muestra la geometría de exploración, sobre la superficie del suelo, mediante la proyección de perspectiva proporcionada por el sistema óptico asociado a la cámara.

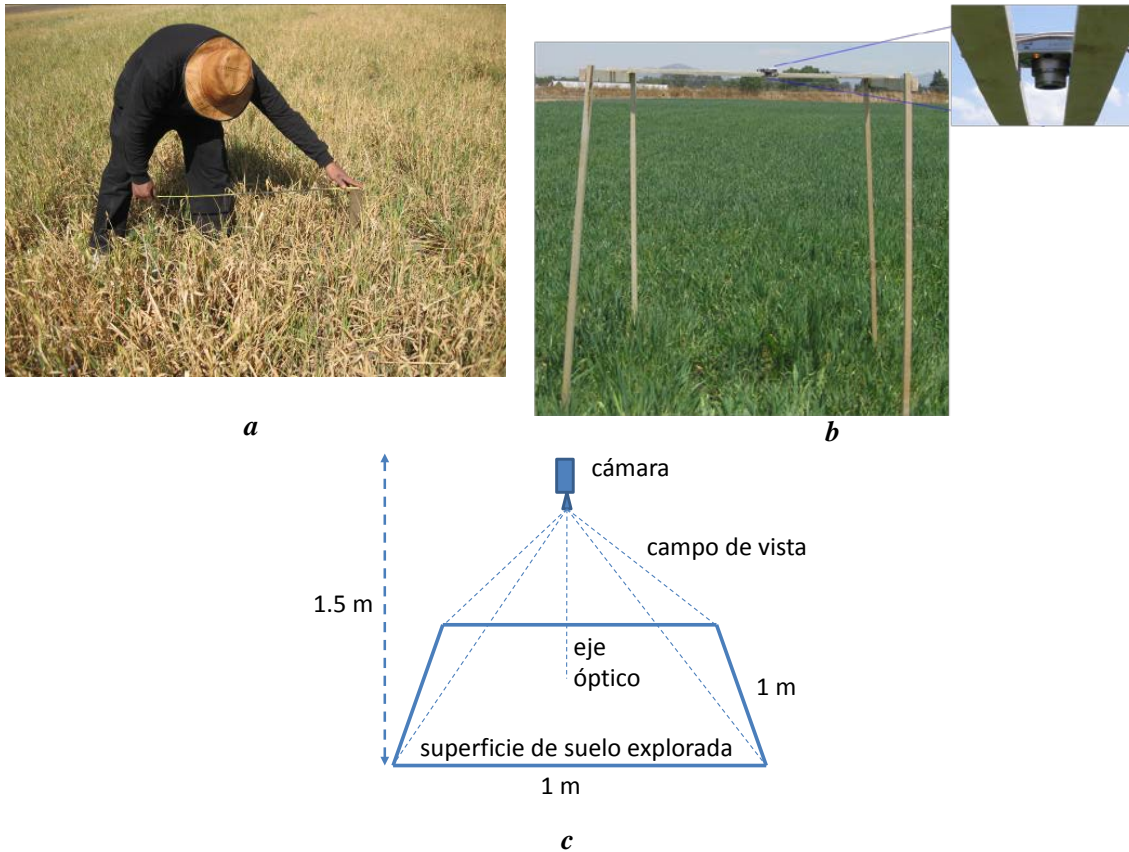


Figura 4.2 (a), (b) Selección y delimitación de las muestras del cultivo de avena para ser fotografiado, (c) representación del sistema geométrico con el eje óptico perpendicular al suelo.

Sobre cada muestra fotografiada se establece un punto de control, de manera que durante el proceso de validación de la clasificación sea posible ubicar los puntos necesarios. Estos puntos de control fueron geo-codificados, es decir, se asignaron coordenadas geográficas, latitud-longitud a cada uno de ellos (León y col., 2007), cuya finalidad es su posicionamiento y localización.

Desde el punto de vista de la clasificación de la cobertura terrestre, bajo el cual se llevó a cabo este trabajo, no es necesario tener en cuenta las coordenadas geográficas de las imágenes. Sin embargo para la validación del clasificador sí son imprescindibles, dado que esto facilita enormemente la comparación entre las clases de la imagen original y la clasificada.

4.2.2. PRE-PROCESAMIENTO DE LAS IMÁGENES A COLOR

Las imágenes de los campos de cultivo son capturadas en un entorno al aire libre, es decir, con alta variabilidad de las condiciones de iluminación natural. La respuesta del sensor es proporcional a la energía de la luz proyectada sobre su superficie, y depende de las longitudes de onda, dentro del espectro visible, reflejadas por los objetos. En el caso que nos ocupa son las plantas y el suelo. Cada longitud de onda produce una respuesta diferente a la que se asigna un color estándar.

Para hacer frente al problema de variación en la iluminación se propone el uso del espacio de color CIELAB (1976), también denominado $L^*a^*b^*$ (Robertson, 1976), ya que según los resultados experimentales realizados con él, resulta ser menos dependiente de la iluminación que otros modelos de color, como se expuso en el capítulo dos. CIE es la abreviatura en francés de la Comisión Internacional de la Luz (Commission Internationale d'Eclairage), donde la componente L^* representa la luminosidad de color, y las otras dos, a^* y b^* , son las componentes de cromaticidad (dimensión del color oponente), representadas mediante un conjunto de ejes de color contrastantes (rojo-verde y amarillo-azul). De acuerdo con Mendoza y col. (2006), este modelo de color se considera aproximadamente uniforme, es decir, la distancia entre dos colores en un espacio de color lineal corresponde a las diferencias percibidas entre ellos. Sangwine (2000) afirma que este modelo proporciona una representación objetiva del color y su uso es esencial para aplicaciones en las que los resultados deben coincidir con los de la percepción humana, como en este caso. Así pues, las imágenes se transforman del modelo de color RGB al CIE $L^*a^*b^*$ mediante las transformaciones matemáticas que se describen en el Anexo 3. Una vez que se tienen las imágenes en el espacio de color deseado, se separan las tres componentes espectrales, procediéndose a la normalización de los niveles digitales contenidos en ellas; a continuación se aplica el correspondiente proceso de clasificación que se describe a continuación.

4.3. APLICACIÓN DE LA ESTRATEGIA DE CLASIFICACIÓN NO SUPERVISADA

La primera fase de la clasificación es la *segmentación* de las imágenes, la cual se lleva a cabo mediante la combinación de tres métodos de umbralización, con el objeto de combinarlos y posteriormente separar en dos grupos cada una de las componentes espectrales de la imagen ya procesada.

La imagen en el espacio de color CIELab posee tres componentes: L^* , a^* y b^* . Para cada uno se debe calcular de forma automática un umbral basándose en la combinación de los tres métodos de umbralización seleccionados, a saber: Isodata, Otsu y *Fuzzy*. Aplicando esta umbralización sucesivamente a cada una de las bandas espectrales del espacio de color, L^* , a^* y b^* , se produce la partición de los correspondientes histogramas en varias regiones, dependiendo del número de umbrales establecido. La combinación de estas regiones determina las diferentes clases donde cada píxel puede ser clasificado. Dependiendo de la naturaleza del problema el número de regiones es variable, si bien debe ser suficiente para hacer frente al problema de clasificación.

A continuación se desarrolla un proceso de codificación, el cual se divide en dos etapas: asignación y etiquetado. La primera consiste en asignar una de las clases posibles a cada píxel para cada una de las componentes espectrales. El etiquetado es el proceso de codificación de la imagen, con sus tres componentes espectrales integradas, que consiste en la asignación de un código numérico a cada píxel con localización de coordenadas de imagen (x,y) , que permita identificar la posición y la clase a la que pertenece. Por último, el algoritmo de decisión o clasificación se encarga de evaluar cada uno de los grupos establecidos procedentes de la clasificación mediante los correspondientes índices de validación descritos en el capítulo dos. Durante el proceso de clasificación se realiza una reasignación de píxeles hasta conseguir que los valores proporcionados por los índices de validación sean los que expresen la mejor calidad de la partición.

Gracias a su flexibilidad, la estrategia de clasificación propuesta llega a ser no supervisada con un alto grado de similitud respecto de la expuesta en el capítulo previo. Se admite un número variable de clases que se ajusta a los requisitos del problema. La sinergia aprovechada en las estrategias de segmentación y clasificación resulta ser una de las principales aportaciones de este trabajo. La fusión de los diferentes enfoques mejora la segmentación de las imágenes agrícolas utilizadas en estos experimentos, detectando los daños en los cultivos de avena.

4.3.1. SEGMENTACIÓN DE IMÁGENES A COLOR POR COMBINACIÓN DE UMBRALES

Como ya se ha expuesto previamente, tanto en la sección 2.4 como posteriormente en la sección 3.3.2.1, existen diferentes métodos de umbralización por división del histograma que han sido ampliamente comentados en la literatura, reiterando que no

existe un criterio unánime sobre cuál es mejor. El tema es que cada método suele producir resultados diferentes debido a que formula distintas hipótesis sobre el contenido de la imagen. En Gonzales-Barron y Butler (2006), se evalúan siete algoritmos de umbralización en imágenes con alta variabilidad de textura, como las utilizadas en nuestro problema de clasificación. Los métodos analizados son: 1) Ridler y Calvard (1978), denominado Isodata; 2) algoritmo de Otsu (1979); 3) umbral por mínimo error, de Kittler y Illingworth (1986); 4) algoritmo *K-means* (Hartigan, 1975); 5) histograma por entropía (Pun, 1980); 6) el método Preservación del Momento (*Moment preserving*), de Tsai (1985); y 7) umbral difuso (*fuzzy*) de Huang y Wang (1995). De acuerdo con Gonzales-Barron y Butler (2006), los mejores resultados se obtienen con Isodata, Otsu, umbral difuso y el método preservación del momento, por este orden. Isodata y *Moment preserving* obtuvieron resultados muy similares. En base a este estudio hemos seleccionado los tres primeros con el fin de diseñar nuestra estrategia basada en combinación de umbrales, cuyos detalles se proporcionan a continuación.

Para cada componente espectral en el espacio de color CIELab obtenemos tres umbrales, t_I , t_O y t_F , que corresponden a la aplicación de los tres algoritmos de umbralización seleccionados, Isodata, Otsu y *Fuzzy* respectivamente, descritos en la sección 2.4. El problema consiste en encontrar una combinación de ellos que proporcione un valor de umbral único y a la vez genere los mejores resultados para la clasificación. Según Kuncheva (2004), en clasificación existen diferentes métodos y estrategias de combinación. La más simple es la fusión mediante el cómputo de la media aritmética, que en el caso de los tres umbrales resulta ser $t = (t_I + t_O + t_F)/3$. Se ha podido constatar que el uso de la media como método de fusión produce resultados satisfactorios en la clasificación final en esta aplicación. Una de las ventajas de la media es que los valores máximo y mínimo del umbral no se aplican estrictamente, ya que se obtiene un valor intermedio. El promedio del umbral combinado se aplica sobre cada componente espectral, L^* , a^* y b^* , y se denota como t_L , t_a , y t_b , respectivamente. Por tanto, cada componente espectral se divide en dos regiones mediante su correspondiente umbral.

4.3.2. CODIFICACIÓN DE LA INFORMACIÓN

La codificación es el proceso mediante el cual se asocia a cada píxel un código numérico que identifica su posición (x,y) en la imagen y proporciona información sobre la clase a la que pertenece. El algoritmo de clasificación no supervisada que se describe a continuación se fundamenta en dos procesos secuenciales bien diferenciados relativos a la codificación. El primero, al que denominamos *proceso de asignación*, consiste en otorgar un código numérico a cada píxel en su localización (x,y) y para cada componente espectral; es decir, al tratarse de una imagen en color se dispone de tres componentes espectrales que se codifican de manera independiente. El segundo proceso de codificación se aplica tras la integración de las tres componentes espectrales. En este proceso se hace uso de los códigos de etiqueta generados para cada componente espectral individual, para posteriormente generar los códigos finales de clasificación de los píxeles en la imagen. Este segundo proceso se denomina *etiquetado de los grupos*.

4.3.2.1. PROCESO DE ASIGNACIÓN

Dado un píxel i , perteneciente a la imagen original RGB y localizado en (x,y) , sobre él se aplica la correspondiente transformación desde el espacio de color RGB al CIE $L^*a^*b^*$, donde sus tres componentes espectrales se denotan como $L^*(x,y) = i_L$, $a^*(x,y) = i_a$, y $b^*(x,y) = i_b$. Como se mencionó anteriormente, los métodos de umbralización dividen el histograma en dos regiones. Puesto que existen tres componentes espectrales, obtendremos seis sub-regiones. Si es necesario podemos aplicar umbrales sucesivos sobre cada componente espectral, de manera que con un segundo umbral se obtiene tres particiones del histograma por componente espectral. Si se aplica un tercer umbral obtendríamos cuatro particiones, y así sucesivamente. Por lo tanto, suponiendo que el número de umbrales por componente es finalmente M , tendremos $t_{L1}, t_{L2}, \dots, t_{LM}$, para el canal L^* , y de la misma manera, $t_{a1}, t_{a2}, \dots, t_{aM}$ para el canal a^* , y $t_{b1}, t_{b2}, \dots, t_{bM}$, para el b^* . Sobre esta base, cada píxel i puede ser codificado como \tilde{i}_s según sus componentes espectrales a través de la ecuación (4-1).

$$\tilde{i}_s = \begin{cases} 0 & \text{if } i_s \leq t_{s1} \\ 1 & \text{if } t_{s1} < i_s \leq t_{s2} \\ 2 & \text{if } t_{s2} < i_s \leq t_{s3} \\ \vdots & \\ M & \text{if } i_s > t_{sM} \end{cases} \quad (4-1)$$

donde s denota la componente espectral, es decir, $s = L, a^*$ ó b^* , y t_{si} son los umbrales correspondientes.

Por ejemplo, sabemos que en el espacio de color CIE $L^*a^*b^*$ los valores de L^* se sitúan en el rango $[0, 100]$, mientras que a^* y b^* varían en el rango $[-110, 110]$. Por lo tanto, aplicando dos umbrales $t_{a1} = -20$ y $t_{a2} = 60$ para la componente espectral a^* , un píxel se codifica como 0, 1 ó 2 según su valor espectral a^* sea inferior a -20, esté entre -20 y 60, o sea mayor que 60, respectivamente.

4.3.2.2. ETIQUETADO DE LOS GRUPOS GENERADOS EN LA IMAGEN $L^*a^*b^*$

Una vez que la imagen ha sido codificada según sus componentes espectrales individuales, el siguiente paso es el etiquetado de las clases existentes, es decir, asignar a cada uno de los píxeles el código de clase al que pertenece. Dado que se han obtenido M umbrales y, por consiguiente, se tienen n particiones del histograma por canal, siendo $n = M+1$, el número de posibles combinaciones es n^d , donde d es el número de componentes espectrales. Al tratarse del espacio de color CIE $L^*a^*b^*$, $d = 3$. El número de combinaciones representa el número de clases. Sólo es necesario identificar cada grupo por su etiqueta y asignar a cada píxel la etiqueta correspondiente del grupo acuerdo a la ecuación (4-2). Así pues, dado el píxel $i \equiv (x, y)$ con los códigos \tilde{i}_L, \tilde{i}_a , y \tilde{i}_b obtenidos según la ecuación (4-1), estas etiquetas pueden generarse mediante \tilde{p}_i como sigue,

$$\tilde{p}_i = n^2 \tilde{i}_L + n \tilde{i}_a + \tilde{i}_b \quad (4-2)$$

4.3.3. CLASIFICACIÓN DE LA IMAGEN EN COLOR

Sea C_k el número de clases posibles, donde k identifica una clase entre 1 y n^d . Cada clase contiene N_k píxeles de la imagen original, y cada píxel posee tres valores

correspondientes a los tres niveles de las tres bandas espectrales con $d=3$. Es un espacio tridimensional, siendo sus elementos vectores con las componentes espectrales de los píxeles en el espacio CIELab, es decir, $\mathbf{i}_k \equiv (i_L^k, i_a^k, i_b^k)$ para el píxel $i \equiv (x, y)$, donde el índice k indica que el píxel y sus componentes espectrales pertenecen a la clase C_k .

Para cada clase se calcula el valor medio de los miembros pertenecientes a esa clase,

$$\boldsymbol{\mu}_k \equiv (\mu_L^k, \mu_a^k, \mu_b^k) = \frac{1}{N_k} \sum_{i_k \in C_k} \mathbf{i}_k \quad (4-3)$$

Basándose en las potencialidades del método de Otsu, es posible calcular la varianza espectral *dentro de cada clase y entre las clases*, denotadas por σ_k y σ_{kh} respectivamente, ecuaciones (4-4) y (4-5). Evidentemente σ_k está relacionada únicamente con la clase C_k , mientras σ_{kh} involucra dos clases, C_k y C_h , con $k \neq h$.

$$\sigma_k = \frac{1}{d \cdot N_k} \sum_{i_k \in C_k} \left[(i_L^k - \mu_L^k)^2 + (i_a^k - \mu_a^k)^2 + (i_b^k - \mu_b^k)^2 \right]^{1/2} \quad (4-4)$$

$$\sigma_{kh} = \frac{1}{d} \left[(\mu_L^k - \mu_L^h)^2 + (\mu_a^k - \mu_a^h)^2 + (\mu_b^k - \mu_b^h)^2 \right] \quad (4-5)$$

Con el valor de estas varianzas podemos fusionar determinadas clases en función de sus similitudes espectrales. La similitud es un concepto definido de la siguiente manera. Dada las clases C_k y C_h , $k \neq h$, ambas se fusionan en una única clase si $\sigma_k \geq \sigma_{kh}$ ó $\sigma_h \geq \sigma_{kh}$. La hipótesis que sustenta esto se basa en el hecho de que si se ha logrado una buena partición, las clases obtenidas deben estar debidamente separadas, sin superposición entre ellas, en cuyo caso no es necesaria su fusión. Por el contrario, si dos clases se solapan, significa que la varianza espectral entre ellas σ_{kh} es menor que la varianza individual dentro de las clases σ_k y σ_h , requiriéndose, en este caso, un nuevo agrupamiento. Este proceso de re-agrupamiento se repite hasta que todas las varianzas entre clases son mayores que sus correspondientes varianzas dentro de las clases. Sin pérdida de generalidad, si se mezclan los píxeles de dos clases, los píxeles de la clase resultante fusionada serán re-etiquetados con la etiqueta de la clase con valor de

varianza más pequeña, mientras que la otra etiqueta se elimina definitivamente. Esto no afecta al proceso de clasificación ya que sólo se modifican las etiquetas.

En el caso de terrenos de cultivo de avena el número de clases es siempre menor que ocho porque no existen más estados posibles para la plantación.

4.4. RESULTADOS Y DISCUSIÓN

Con el propósito de mostrar el funcionamiento de la clasificación no supervisada propuesta en este capítulo se describen a continuación las imágenes objeto de estudio y cómo se ha aplicado esta estrategia sobre las mismas.

4.4.1. DESCRIPCIÓN DEL CULTIVO DE AVENA

Se ha seleccionado una superficie del cultivo de avena de 20 hectáreas. De esta superficie se tomaron al azar 2000 muestras de 1 m², que fueron fotografiadas con el sensor CCD descrito en la sección 4.1. Las fotografías reflejan zonas del cultivo con diferente grado de daño, hecho que se puede observar en las imágenes seleccionadas como representativas del conjunto total (figuras 4.3 *a* y 4.4 *a*). El grado de afectación fue evaluado por un experto en agricultura y gestión de pérdidas.

- 1) **Descripción del área de cultivo de la fotografía 4.3 *a*:** la densidad de planta verde en promedio es de 53,15%, por lo que se considera que muestra un bajo grado de daños por heladas. De esta zona de cultivo se seleccionaron 1.000 muestras al azar, el 10% de las cuales se usaron como datos de referencia para la evaluación del método de clasificación propuesto. El resto, es decir 900, fueron clasificadas y comparadas con las imágenes de prueba.
- 2) **Descripción del área de cultivo de la fotografía 4.4 *a*:** en esta zona el promedio de densidad de planta verde estimado es del 26,6%; el resto se considera planta seca y suelo. Por tanto, esta zona es un ejemplo de alto grado de damnificación por heladas. El número de muestras usadas para clasificación y de referencia es similar a las señaladas en el punto anterior.

El objetivo del proceso de clasificación es identificar las cuatro clases que se han definido según el grado de afectación del cultivo debido a las muy bajas temperaturas. La primera clase es la de planta poco afectada, la cual muestra en su mayoría las componentes espectrales correspondientes a verdes. La segunda clase es la avena

altamente afectada por las heladas, encontrándose en un estado muy seco, es decir, las componentes de color son principalmente amarillentas. El tercero representa un estado intermedio, donde las plantas presentan tonalidades de tipo verde-amarillento, esto es, sin ser totalmente verdes tampoco están secas. La cuarta clase representa el suelo oscuro o con sombras. Desde el punto de vista agrícola, las plantas del primer grupo van a seguir creciendo y desarrollándose hacia la madurez y, a menos que se vieran afectadas por una nueva helada u otros agentes, es posible estimar la futura cosecha que se obtendrá a partir de ellas. El segundo grupo debe desecharse desde el punto de vista de expectativas de cosecha. Del mismo modo, el cuarto tampoco se considera a estos efectos. Por último, el tercer grupo se estima, siempre según el experto, que contribuye con alrededor del 40-60% para calcular los rendimientos de cosecha. Obviamente, esta cuantificación se calcula en el momento de la detección, de forma que los daños posteriores no son considerados.

Por tanto, existen tres diferentes clases de textura de plantas de avena que tienen que ser reconocidas con el fin de identificar el estado del cultivo, y una cuarta clase que corresponde a sombras y suelo. Esto significa que en una determinada imagen pueden identificarse las cuatro clases o sólo algunas de ellas. Por lo tanto, de acuerdo con el procedimiento descrito, un único umbral es suficiente ya que no es necesario identificar más de ocho clases.

Dado que al haber tres componentes en el espacio de color esto da lugar a ocho clases, de las cuales sólo cuatro corresponden a la realidad, será necesaria la reagrupación de los píxeles para pasar de ocho a cuatro categorías. Esto se puede lograr mediante la relajación de los criterios de fusión con el fin de forzar la unión de las cuatro clases con el mayor grado de solapamiento, lo que se determina por la identificación de las mayores diferencias entre σ_k ó σ_h y σ_{kh} , ecuaciones (4-4) y (4-5).

Las figuras 4.3 **b** y 4.4 **b** muestran los resultados de la clasificación de las imágenes originales presentadas en las figuras 4.3 **a** y 4.4 **a**, respectivamente. Las etiquetas de las cuatro clases están representadas por colores: 1) verde, para la clase que representa la planta de la avena que no se ve afectada por las bajas temperaturas (color real); 2) de color amarillo los cultivos de avena en estado seco debido a las heladas; 3) de color rojo, corresponde a las plantas en una fase intermedia entre verde y seca; 4) azul,

corresponde a la categoría de suelo y sombra. Las clases se identifican como sigue: GO (avena verde), DO (avena seca), HD (avena entre verde y amarilla), y SG (suelo oscuro o con sombras). Como se ha indicado anteriormente, en algunas imágenes no necesariamente se presentan las cuatro categorías, como ocurre en la figura 4.5 *b*, en la que no aparece suelo ya que está totalmente cubierto con plantas de avena en sus tres posibles estados.

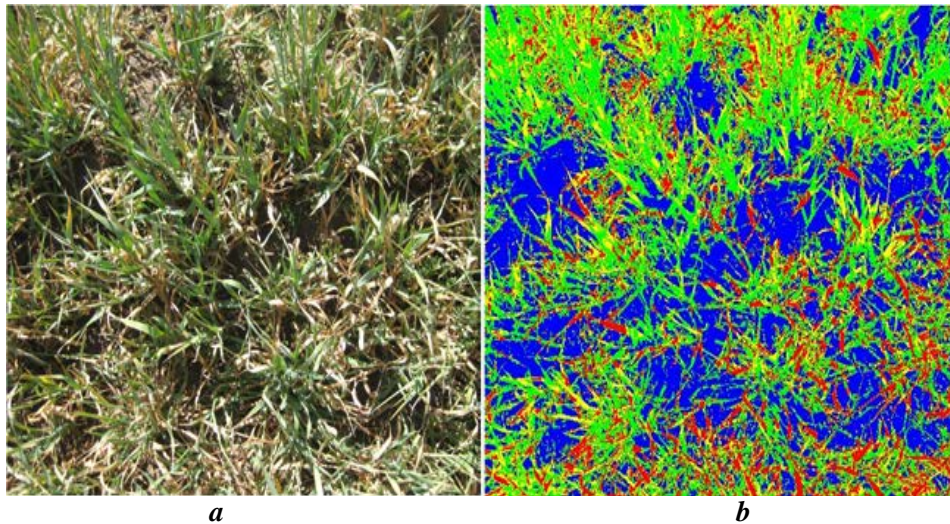


Figura 4.3. Muestra de cultivo de avena poco afectado por las heladas: (a) imagen original, (b) resultados de la clasificación (cuatro re-agrupamientos).

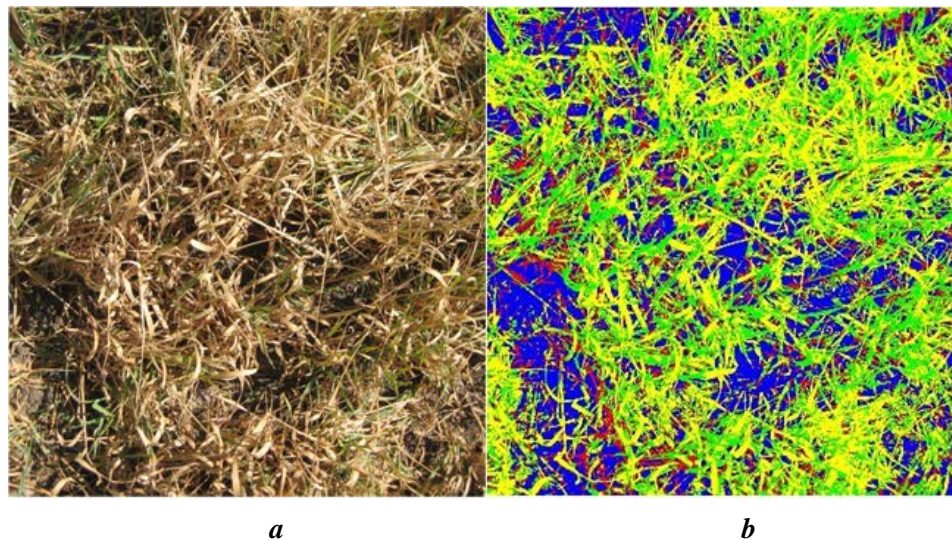


Figura 4.4. Muestra de cultivo de avena muy afectado por las heladas: (a) imagen original, (b) resultados obtenidos de la clasificación (cuatro re-agrupamientos).

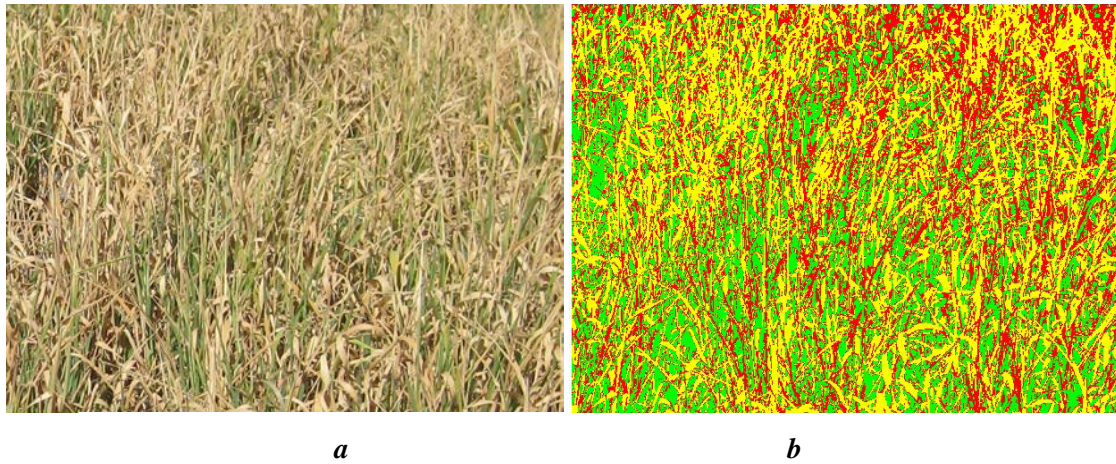


Figura 4.5. Muestra de cultivo de avena (toma superficial): **(a)** imagen original, **(b)** resultados obtenidos de la clasificación.

4.4.2. VALIDACIÓN DEL CLASIFICADOR

Como es bien sabido, la validación de un proceso de clasificación se refiere al grado de concordancia entre las clases asignadas a cada píxel de la imagen clasificada y un conjunto de datos de referencia determinados por un experto. Para obtener una estimación cuantitativa se calculó la matriz de error difusa en base a la metodología propuesta por Congalton y Green (2009) y Green y Congalton (2004), detallada en la sección 2.6.

Uno de los requisitos necesarios para realizar la validación del clasificador es la selección del tipo de unidades de muestra, también llamadas de validación o referencia, así como su tamaño y la técnica de muestreo.

4.4.2.1. SELECCIÓN DE LAS UNIDADES DE VALIDACIÓN

En este caso se van a utilizar como unidades de validación agrupaciones de píxeles en forma de polígonos con diferentes formas y tamaños. De este modo se pueden analizar los errores temáticos sin llegar a confundirlos con los errores de posición.

El número de unidades de validación depende de las dimensiones de la imagen y el número de clases. Siguiendo las pautas establecidas en el capítulo previo, aquí también hemos aplicado una distribución multinomial (Congalton y Green, 2009) para su selección. Se determina a partir de una distribución chi-cuadrado (χ^2), utilizando el nivel de confianza deseado y el porcentaje de la imagen cubierto por cada clase, de acuerdo con la ecuación (4-6).

$$n_s = B\Pi_k (1 - \Pi_k) / b_k^2 \quad (4-6)$$

Donde Π_k es la fracción de la superficie de la imagen que pertenece a la clase k ; B es un coeficiente que se obtiene a partir de la distribución chi-cuadrado con un grado de libertad, y el parámetro $(1 - b/k)$; b es la precisión (en este caso, $b = 0,05$ ya que el grado de confianza se ha establecido en el 95% y b representa el 5%); k es el número de clases o categorías. Por ejemplo, si hay cuatro categorías en nuestro esquema de clasificación ($k = 4$), y el nivel de confianza deseado es del 95% (con precisión del 5%), seleccionando una clase particular que cubra el 16% de los píxeles de la imagen, esto es, $\Pi_k = 16\%$, el valor de B debe ser determinado a partir de la tabla de chi-cuadrado (χ^2) con un grado de libertad y $(1 - 0,05/4) = 0,9875$. En este caso el valor apropiado de B es $\chi^2(1, 0,9875) = 6,36640$. Por lo tanto, de acuerdo con los datos y aproximaciones anteriores, el resultado del número de unidades de muestreo es el siguiente:

$$n_s = 6,36640(0,16)(1 - 0,16) / 0,05^2 = 342$$

Si los píxeles de la imagen corresponden a las cuatro clases, y cada una de ellas cubre el 35%, 16%, 16% y 33% respecto de la imagen total, entonces los 342 polígonos se distribuirán en cuatro categorías, en concreto habrá 120, 55, 55 y 112 para cada una de ellas, respectivamente.

4.4.2.2. TÉCNICA DE MUESTREO

Una vez determinado el número de polígonos de cada clase (ecuación 4-6), la imagen original es analizada por un experto que asigna las clases correspondientes a estos polígonos, en base a los siguientes criterios:

- (a) Cada polígono debe contener sólo píxeles de una clase, por lo tanto el tamaño y forma de cada polígono se debe ajustar a los objetos de la imagen. Por ejemplo, los polígonos relacionados con la clase de las hojas de avena (mayoría de las imágenes analizadas) tienen formas irregulares, como se muestra en la figura 4.6 *a*.
- (b) Los polígonos representados en la imagen original (figura 4.6 *a*) son conocidos como datos de referencia o datos de campo, los cuales deben ser validados por un experto.

(c) Para visualizar la correspondencia de los polígonos dibujados sobre la fotografía original con las clases presentes en las imágenes clasificadas (figuras 4.6 *a* y *b*), se usaron herramientas propias de Sistemas de Información Geográfica (SIG), concretamente ArcView (2012), versión 3.2, considerando que ambas imágenes (original y clasificada) poseen idénticas coordenadas UTM (figura 4.7 *a* y *c*). De esta forma los polígonos correspondientes se pueden superponer tanto en las imágenes clasificadas (figura 4.7 *c*) como en las originales (figura 4.7 *a*), lo que permite cuantificar la correspondencia de las etiquetas.

La figura 4.6 muestra la distribución de estos polígonos en ambas imágenes para las siguientes categorías: suelo sombreado (línea blanca), avena verde (línea roja), avena seca (línea en negro), y avena en un estado intermedio (línea rosa). La imagen de referencia ha sido etiquetada por los expertos y la clasificada por la estrategia propuesta.

Para cualquier polígono en la imagen clasificada (figura 4.6 *b* o figura 4.7 *b*), cuya ubicación y etiqueta son conocidas, es posible calcular el número de píxeles que han sido clasificados correcta o incorrectamente. Esta correspondencia constituye la base para el cálculo de las matrices de error, según se describe en la siguiente sección.

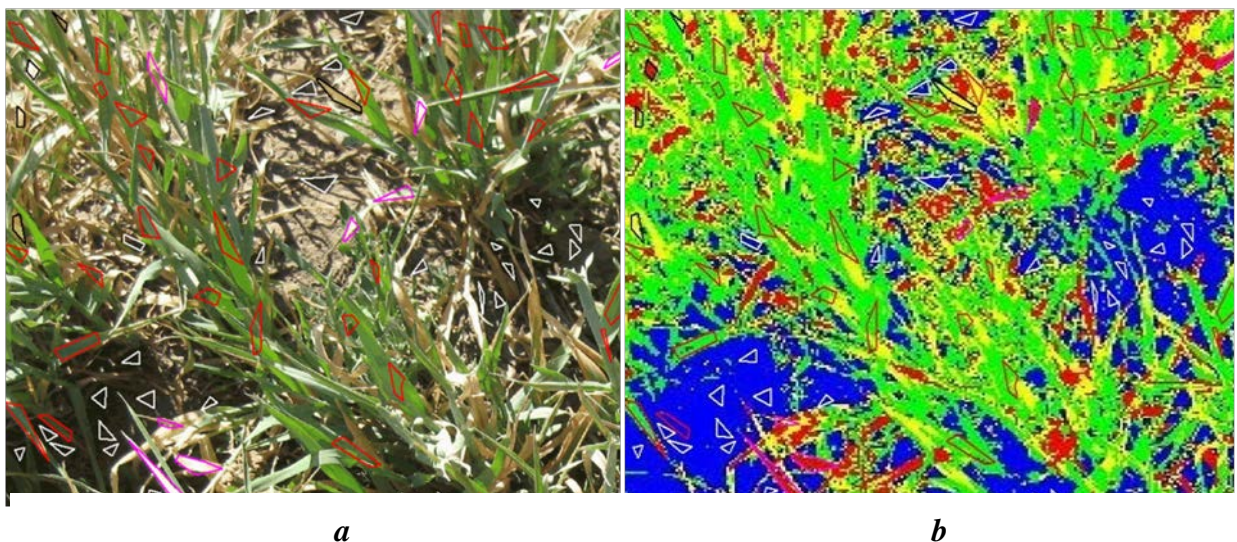


Figura 4.6. Datos de referencia: (a) unidades de validación (polígonos) en la imagen original, (b) unidades de validación (polígonos) en la imagen clasificada.

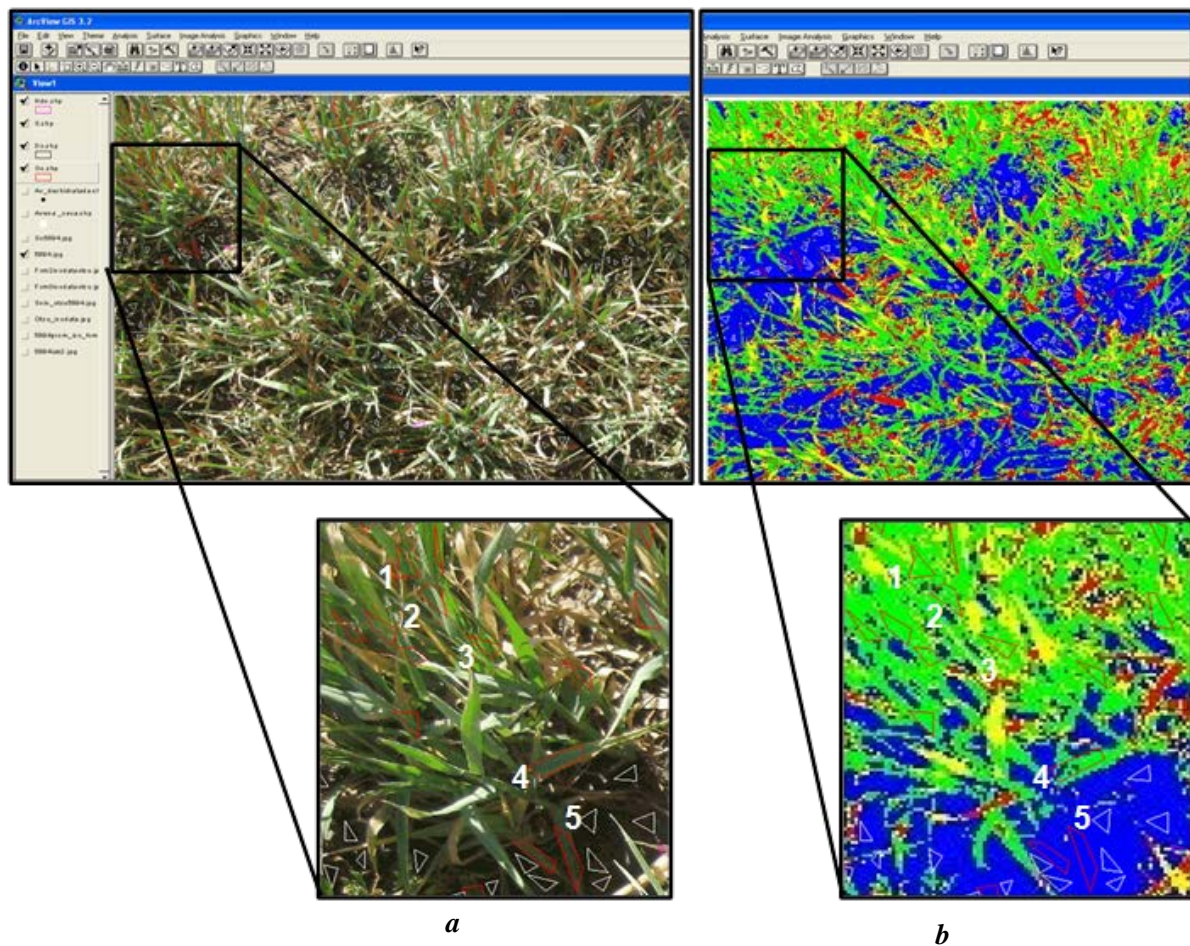


Figura 4.7. Trazado de polígonos de muestreo: (a) imagen original geo-referenciada, (b) superposición de polígonos en la imagen clasificada.

4.4.2.3. CÁLCULO DE LA MATRIZ DE ERROR DETERMINISTA Y DIFUSA

Las matrices de error son una forma muy efectiva de obtener y representar la eficiencia y precisión del clasificador, así como las de cada clase individual. Además permiten identificar los errores de inclusión (errores de comisión) y los errores de exclusión (errores de omisión). Un error de comisión se comete cuando una unidad de validación se considera como perteneciente a una categoría cuando realmente no pertenece a ella. Un error de omisión se produce cuando una unidad de validación se excluye de una categoría a la cual pertenece. Todos y cada uno de los errores constituyen una omisión en la categoría correcta y una inclusión en una categoría equivocada.

Con el fin de obtener los valores de la matriz de error se han comparado las etiquetas asignadas a cada polígono en las imágenes clasificadas y las clases asignadas por el

experto en las imágenes originales (datos de referencia). La correspondencia espacial entre los dos conjuntos de datos proporciona una medida cuantitativa de la exactitud de nuestro procedimiento de clasificación (Congalton y Green, 2009). Hemos comparado 200 polígonos de las 2000 muestras disponibles. Los resultados de cada clase se muestran en la tabla 4.1, la cual representa al mismo tiempo la matriz de error determinista (MET) y la matriz de error *fuzzy* o difusa (MEF) bajo los mismos criterios y planteamientos realizados en los capítulos precedentes.

Los valores de la diagonal principal muestran el número de unidades de validación correctamente clasificadas. Las celdas fuera de la diagonal principal contienen un par de valores, separados por punto y coma; el primer valor representa el número de muestras que, aunque no son absolutamente correctas, son consideradas como aceptables en cuanto a su clasificación, de acuerdo con las reglas difusas previamente establecidas. El segundo valor indica aquellas unidades de validación cuya clasificación es considerada inaceptable, incluso para la aproximación *fuzzy*, es decir, se identifican como mal clasificadas. Para la matriz de error determinista o clásica estos dos valores se suman y se obtiene un único valor que determina el error de la clasificación.

Por ejemplo en la tabla 4.1 los dos valores de la primera fila, tercera columna (1,3) son 400 y 600. Desde el punto de vista determinista, hay 1000 unidades de validación o polígonos (400+600) clasificados erróneamente en la categoría GO por el método propuesto, ya que de acuerdo a los datos de referencia deberían ser HD (columna). Esto significa que 1000 polígonos han sido excluidos de la categoría correcta, HD, e incluidos en la categoría incorrecta, GO. Sin embargo, bajo el criterio difuso la clasificación de 400 de esos polígonos fue considerada aceptable, mientras que las restantes 600 muestras sí están mal clasificadas.

Tabla 4.1. Matriz de error *Fuzzy* y Determinista.

		DATOS DE REFERENCIA				Total Clasificado
		GO	DO	HD	SG	
DATOS DE LA CLASIFICACION	GO	22600	600;600	400;600	0;0	24800
	DO	0;0	8800	600;400	0;0	9800
	HD	0;0	600;200	9200	0;0	10000
	SG	600;800	0;0	0;0	22800	24200
Total U. de Muestréo		24000	10800	11200	22800	68800

GO: avena verde
 DO: avena seca
 HD: avena deshidratada (entre verde y amarilla)
 SG: suelo oscuro o con sombras.

El conjunto de reglas difusas establecidas para determinar la precisión de la clasificación son las siguientes:

- **Absolutamente correctas:** una muestra se considera clasificada de forma absolutamente correcta cuando se superpone el 100% de la superficie de esa unidad en la imagen de referencia de una categoría particular a la misma clase y ubicación en la imagen clasificada. Se anotan en la diagonal principal de la matriz de error.
- **Aceptable:** una unidad de validación se considera clasificada de forma aceptable cuando existe al menos el 50% de solapamiento entre la superficie del polígono de referencia para una categoría particular y la misma clase y superficie en la imagen clasificada. Por ejemplo, un polígono en la imagen original (de referencia) que pertenece a la clase GO se proyecta sobre dos clases en la imagen clasificada, GO y SG. Si por lo menos el 50% de esa muestra coincide con la clase GO se considera aceptable. Esto añadirá 1 al contador izquierdo en la celda que corresponde a la columna GO (no-diagonal principal), fila SG.

- **Error (inaceptable):** se considera un error de clasificación de este tipo cuando un polígono de referencia de una categoría determinada se superpone en más del 50% de su área con una clase totalmente equivocada. Por ejemplo, un polígono que en la imagen de referencia pertenece a la clase DO está en la clase GO de la imagen clasificada. Esto añade 1 al valor a la derecha de la celda, colocado en la fila GO, columna DO.

Además de que esta matriz muestra claramente los errores de omisión y comisión, la matriz de error *fuzzy* y la determinista pueden ser usadas para calcular otras medidas de validación del clasificador como la precisión global, la precisión del experto, y la precisión del clasificador, que son de interés para la realización de inventarios en el ámbito de la agricultura.

4.4.2.4. PRECISIÓN DEL CLASIFICADOR Y ERRORES DE COMISIÓN

De acuerdo con Anderson (1976), un sistema de clasificación eficiente debe tener una precisión general de al menos el 85%. A continuación se diferencian dos conceptos relativos a la precisión.

La **precisión global determinista** (es decir, tradicional) resulta ser simplemente la suma de los valores de la diagonal principal (es decir, las muestras correctamente clasificadas), dividida por el número total de unidades de validación en toda la matriz de error, resultando para la matriz de error de la tabla 4.1: $63400/68800 = 92\%$.

La **precisión global de la evaluación *fuzzy*** se calcula como la suma de los valores de la diagonal principal (muestras correctamente clasificadas) más las que se consideran aceptables (el primer valor de las celdas que no están en la diagonal principal), dividida por el número total de unidades de validación de toda la matriz de error difusa, que para la tabla 4.1 es: $66200/68800 = 96\%$.

La precisión global, tanto determinista como *fuzzy*, expresa lo bien que el clasificador identifica el tipo de cobertura terrestre. Es decir, indica cuántas veces un tipo de cobertura obtenido como resultado de la clasificación concuerda con la realidad.

Esto mismo se puede obtener para cada clase. Entonces la precisión determinista de cada categoría se calcula dividiendo cada celda de la diagonal principal (número de unidades de validación correctamente clasificados para esta clase) por el total de muestras clasificadas en dicha categoría (total de la fila). La evaluación de la precisión

difusa por categoría se calcula mediante la suma del valor de cada celda en la diagonal principal y los valores que se consideran aceptables (primer valor de cada celda de la fila, fuera de la diagonal principal) dividida por el total de polígonos clasificados (total de la fila). La tabla 4.2 muestra las precisiones del clasificador junto con los errores de comisión en la versión determinista y *fuzzy*, cuyos datos para el cálculo se obtienen de la tabla 4.1.

Para facilitar el análisis se diseñaron tres columnas en cada caso: la denominada “Total”, que contiene el número total de muestras bien clasificadas; “precisión del clasificador”, que se calculó dividiendo Total clasificados/Total. La columna “error de comisión” se calcula mediante la diferencia con el 100% de la “precisión del clasificador” para cada clase.

Tabla 4.2. Precisión del clasificador y errores de comisión.

Clases	Determinista			Fuzzy		
	Total	precisión del clasificador	error de comisión	Total	precisión del clasificador	error de comisión
GO	22600	91%	9%	23600	95%	5%
DO	8800	90%	10%	9400	96%	4%
HD	9200	92%	8%	9800	98%	2%
SG	22800	94%	6%	23400	97%	3%

Por ejemplo, la precisión del clasificador para la categoría DO (avena seca) se calcula dividiendo el número total de muestras correctamente clasificadas en esa categoría (8800 para el caso determinista y 9400 para el caso *fuzzy*) entre el número total de unidades de validación contabilizadas como de avena seca en esa fila (8800+600+400=9800, ver tabla 4.1). Por tanto, existe un 90% de probabilidad de que un píxel clasificado como avena seca por el clasificador propuesto pertenezca en realidad categoría esa clase, bajo la perspectiva determinista, y habría un 96% de probabilidad si se calcula con el criterio difuso (tabla 4.2, fila DO).

4.4.2.5. PRECISIÓN DEL EXPERTO Y LOS ERRORES DE OMISIÓN

La precisión del experto se calcula para cada clase, y describe la habilidad de clasificar una categoría en particular en base a la realidad del suelo. Este cálculo se realiza dividiendo el número total de unidades de validación correctamente clasificadas en una categoría determinada (valor de la diagonal principal) entre el número total de unidades de muestreo de esa clase (Total en columna, tabla 4.1). Por ejemplo, $22600/24000 = 0,94$ para la clase GO. Es decir, en este caso se toman en cuenta los errores de exclusión (errores de omisión).

La tabla 4.3 muestra los valores obtenidos al calcular esta precisión bajo el enfoque determinista y *fuzzy*, utilizando los valores correspondientes de la tabla 4.1.

Por ejemplo, si estamos interesados en la capacidad para clasificar la categoría DO, esta precisión se obtiene dividiendo el número total de unidades correctamente clasificadas de esta clase (8800 para el caso determinista y 10000 en el difuso) entre el número total de las unidades de validación (polígonos) de avena seca (DO), según lo indicado por los datos de referencia (10800, total en columna DO). Esta división resulta en una precisión del 81% (determinista) y 93% (*fuzzy*), que son bastante buenas.

La precisión y errores del experto para cada clase se muestran en la tabla 4.3, tanto para el enfoque determinista como para el *fuzzy*. De nuevo se han estructurado los datos en tres columnas para facilitar el análisis. La denominada “Total” contiene el número total de muestras bien clasificadas. La columna denominada “precisión del experto” se calculó dividiendo Total de unidades muestreadas por categoría (tabla 4.1)/Total. La llamada “error de comisión” se calcula mediante la diferencia entre el 100% y “precisión del experto” para cada categoría.

Tabla 4.3. Precisión del experto y errores de omisión.

Clases	Determinista			Fuzzy		
	Total	precisión del experto	error de omisión	Total	precisión del experto	error de omisión
GO	22600	94%	6%	23200	97%	3%
DO	8800	81%	19%	10000	93%	7%
HD	9200	82%	18%	10200	91%	9%
SG	22800	100%	0%	22800	100%	0%

4.4.2.6. PRECISIÓN DE LA CLASIFICACIÓN PARA UMBRALIZACIÓN SIMPLE Y COMBINADA (ERRORES DE OMISIÓN Y COMISIÓN)

La tabla 4.4 resume la precisión del clasificador y del experto, así como los errores de comisión y de omisión. Se comparan los resultados obtenidos tras la aplicación de cuatro estrategias de umbralización, tanto desde el enfoque determinista como difuso. Estos métodos son: combinado de umbralización (CT), estrategia de umbral simple por Isodata (IS), Otsu (OT) y *Fuzzy* (FU). Los valores que se muestran en esta tabla son el resultado del promedio de los resultados de las cuatro categorías (GO, DO, HD, SG). Como puede observarse, el mejor rendimiento se obtiene con la estrategia propuesta, CT, en términos de precisión y error.

Tabla 4.4. Precisión del clasificador y del experto por umbralización combinada (CT) y simple (IS, OT, FU).

	Determinista				Fuzzy			
	Precisión (%)		Error (%)		Precisión (%)		Error (%)	
	Clasific	Expert	Comisión	Omisión	Clasific	Expert	Comisión	Omisión
CT	91.78	89.45	8.22	10.55	96.44	95.08	3.56	4.92
IS	87.71	87.70	12.29	12.30	94.83	94.01	5.17	5.99
OT	85.67	83.10	14.33	16.90	90.99	88.20	9.01	11.80
FU	88.44	89.03	11.56	10.97	94.97	94.89	5.03	5.11

4.4.3. RESUMEN Y DISCUSIÓN

Para identificar el nivel de daños en los cultivos de avena, fundamentalmente debido a las heladas, se ha desarrollado una estrategia de clasificación no supervisada que incluye tres fases: a) aplicación de un método de umbralización de naturaleza automática; b) procedimiento de clasificación mediante combinación de umbrales y fusión de clases similares, y c) aplicación de medidas de validación y precisión del clasificador mediante el cálculo de la matriz de error.

Para la umbralización automática se combinaron los tres mejores métodos de los siete analizados por Gonzales-Barron y Butler (2006): Otsu, Isodata y *Fuzzy*. Los dos primeros algoritmos producen resultados similares, siendo su diferencia con respecto al tercero que éste genera los valores de umbral más bajos. En Valdovinos y Sánchez (2007) y Kuncheva (2004) se afirma que la combinación de diferentes clasificadores ha demostrado su utilidad para mejorar los resultados de la clasificación. Basándonos en esta idea, en nuestra propuesta hemos aplicado la combinación de umbrales, tomando como valor umbral definitivo el promedio de estos tres umbrales para cada banda espectral del espacio de color CIELab. Se ha comprobado a través de los experimentos realizados que la fusión proporciona mejores resultados que cuando se consideran los umbrales por separado, lo cual permite afirmar que la ventaja radica en la elección de un valor intermedio en relación a los tres proporcionados por los métodos individuales.

La precisión de un proceso de clasificación se refiere al grado de concordancia entre la imagen clasificada y las clases observadas en las parcelas, conocidas éstas en la literatura como “*ground truth*” o datos de referencia. Para mostrar cuantitativamente esta precisión se calculó la matriz de error que nos permite identificar algunas fuentes de error y no simplemente el valor "error". La matriz de error *fuzzy* se utilizó para ampliar los resultados a fin de considerar la incertidumbre en el etiquetado.

Como se muestra en la tabla 4.3, las clases que tienen más errores de omisión son HD (*fuzzy*) y DO (determinista). La precisión y errores de omisión para la categoría HD puede ser analizada teniendo en cuenta los valores de la tabla 4.1, tercera columna (HD), última fila (Total de muestras). De 11200 unidades de muestreo clasificadas como HD por el experto, 9200 se asignan correctamente, 600 fueron clasificadas erróneamente como GO, y 400 como DO. En cuanto a la estrategia difusa, 400 se

encontraban entre GO y HD, y 600 entre DO y HD, clasificaciones que se consideran aceptables. Por lo tanto, la precisión de los expertos (tabla 4.3) para la categoría de HD es del 82% (9200/11200), con un error de omisión del 18% en el caso determinista. La precisión *fuzzy* del experto es del 91% (10200/11200), con un error de omisión difuso del 9%. El mismo análisis puede realizarse para cada categoría.

La precisión del clasificador para la misma categoría HD puede obtenerse a partir de la tabla 4.1. Los valores de la tercera columna (HD), última fila (Total de muestras), 11200, representa el número de unidades de validación que fueron clasificadas como HD por el experto. Nuestro clasificador asigna esta categoría a 10000 polígonos (tercera fila, última columna). En la matriz de error difusa es posible observar que, de estos 10000 polígonos, 9200 fueron clasificados correctamente (valor de diagonal). Además, existen 200 unidades de muestra que fueron clasificadas como HD cuando los datos de referencia muestran que en realidad pertenecen a la clase DO. Las otras 600 unidades de muestreo fueron clasificadas como categorías intermedias entre DO y HD y, por tanto, desde este punto de vista difuso su clasificación se considera aceptable (tercera fila, segunda columna, “600; 200”). Con un número total de 10.000 unidades de muestreo clasificadas de esta clase, la precisión determinista resulta ser del 92% (9.200/10.000), mientras que la difusa es del 98% (9.800/10.000) (tabla 4.2). Es decir, esta categoría es la que tiene el error de comisión *fuzzy* más bajo (2%). Dicho de otra manera, el clasificador posee un 98% de probabilidad de clasificar correctamente las plantas de avena deshidratada (HD).

Estos resultados permiten la cuantificación de los daños causados a los cultivos debidos al efecto de temperaturas muy bajas. De las 20 hectáreas analizadas, solamente una superficie de 7,6 hectáreas no se ha visto afectada por las heladas. Sin embargo, las bajas temperaturas han dañado parcialmente una superficie de 2,7 hectáreas, causando una pérdida prácticamente total de 9,7 ha. Basándose en el protocolo sugerido por el SIAP (2011) y el uso de los datos históricos, es posible predecir que la cosecha estimada será de 126,92 t/ha, lo cual se deduce a partir de las 7,6 hectáreas no afectadas por las heladas.

4.5. CONCLUSIONES DEL CAPÍTULO

En este capítulo, al igual que en el anterior, se presenta el diseño de un nuevo clasificador no supervisado que permite la cuantificación de los daños causados por bajas temperaturas en un campo de avena. La diferencia con respecto al clasificador expuesto en el capítulo anterior, estriba en el modelo de color y en el número de umbrales utilizados para la división de las bandas espectrales en regiones. Las imágenes se obtienen mediante una cámara digital de sensor CCD, es decir, con un dispositivo de un costo relativamente bajo. Hemos utilizado el modelo de color CIE $L^*a^*b^*$, ya que se trata de un modelo de color próximo a la percepción humana y por tanto más adecuado en aplicaciones agrícolas con gran variabilidad en las condiciones de iluminación ambiental.

La estrategia de clasificación se basa en la fusión de tres técnicas de umbralización, generando de manera automática tantas clases como sean necesarias para la aplicación. De acuerdo con Martin-H y col. (2009), esta estrategia podría ser catalogada como agrupación dinámica. Esta flexibilidad en el número de clases constituye una de las principales ventajas de este método. El clasificador es capaz de identificar correctamente las plantas de avena afectadas por bajas temperaturas (heladas), distinguiendo claramente las plantas secas, las plantas verde, y las zonas sin vegetación (tierra).

Para los productores a gran escala esta metodología resulta de gran utilidad por el hecho de facilitar la monitorización de los cultivos, de tal manera que se puede llegar a estimar la pérdida de plantas y por tanto la reducción de la producción. También resulta útil para otros propósitos tales como los seguros agrícolas, ya que facilita la evaluación de la superficie dañada y no dañada a la vez que se estima la futura cosecha. Extendiendo el ámbito de aplicación, el mismo procedimiento puede ser adecuado para cuantificar los efectos de las plagas, enfermedades, sequías u otros fenómenos que pueden causar daños en las plantas.

Mediante estos experimentos se llega a la conclusión de que es posible utilizar este tipo de sensores basados en imagen para evaluar el estado fenológico de los cultivos de ciclo corto, como por ejemplo, la avena, el trigo y la cebada, después de haber sido afectados

por agentes climatológicos, algo que hasta donde hemos podido investigar no se ha planteado con anterioridad.

Además, actualmente los cambios bruscos de temperatura son muy frecuentes por lo que el riesgo de afectación en los cultivos es mayor. Esto exige contar con un sistema dinámico de clasificación de imágenes que cuantifique el nivel de afectación de la planta oportunamente para facilitar tomas de decisiones relativas a la estimación de la cosecha.

Otra contribución importante de esta propuesta, con respecto a las observaciones visuales de los expertos, es la estimación cuantitativa de los daños. Las decisiones tomadas por los expertos y productores, que a veces son influenciados por la fatiga u otros aspectos subjetivos, se basan principalmente en aspectos cualitativos. Esta ventaja, junto con las anteriormente mencionadas, justifica el uso del sistema de visión para la clasificación automática en cultivos de avena y por extensión en otros cultivos.

La implementación de esta estrategia podría ser muy útil para diversos agentes de la producción, como productores, comerciantes y aseguradoras agrícolas, ya que permitirá hacer predicciones más precisas sobre la producción.

La aplicación de la estrategia propuesta para el análisis de las superficies de cultivo se ha realizado tras la aparición de daños por fenómenos meteorológicos, si bien es posible su uso en cualquier otra situación donde la clasificación de las texturas o elementos presentes en la imagen constituye el elemento de interés.

UNIVERSIDAD COMPLUTENSE DE MADRID, ESPAÑA

Capítulo 5

Conclusiones y trabajos futuros



CAPÍTULO 5 CONCLUSIONES Y TRABAJOS FUTUROS

El presente capítulo tiene como finalidad resumir las contribuciones más relevantes del trabajo de investigación desarrollado, plasmadas en la presente memoria, así como resaltar los resultados más importantes a modo de conclusiones globales, identificando las líneas de trabajo futuras que puedan dar lugar a la ampliación de la presente investigación.

5.1. CONCLUSIONES

La investigación desarrollada se ha centrado en la identificación, de forma automática, de las diferentes texturas existentes en las imágenes analizadas. Estas imágenes se han obtenido tras la ocurrencia de sendos fenómenos meteorológicos naturales, cuales son huracán y heladas.

En el primer caso se trata de identificar los efectos del huracán sobre la cobertura vegetal, así como en las zonas carentes de la misma. De forma genérica, el objetivo subyacente se centra en el análisis del suelo con fines de evaluación de daños y su posterior uso, principalmente con fines agrícolas.

En el segundo caso, se trata de evaluar los daños producidos por heladas sobre campos de cultivo de avena, cuyo fin último consiste en estimar la disminución en la producción y por consiguiente el rendimiento final para este caso especial de uso del suelo en agricultura.

Con tal propósito y para ambos casos, se ha desarrollado un clasificador de naturaleza no supervisada, cuya finalidad última estriba en su capacidad para el análisis de gran cantidad de datos de imágenes con fines agrícolas y de forma automática. En cada fase del clasificador se han introducido modificaciones o propuestas novedosas que lo hacen suficientemente eficiente para las aplicaciones propuestas.

La unidad base para el análisis de las imágenes ha sido el píxel, cuyas propiedades son las que proporcionan sus correspondientes componentes espectrales, según el modelo de color utilizado en cada caso.

Las principales conclusiones extraídas de este trabajo se exponen a continuación en base a las etapas principales del proceso de clasificación propuesto, agrupándose en: segmentación, clasificación y validación.

Segmentación: para las imágenes en los modelos de color RGB y CIELab se aplicó un proceso de segmentación de píxeles por umbralización, dado su buen comportamiento en imágenes procedentes de entornos naturales. Para la búsqueda del umbral óptimo se eligió el método de Otsu en el caso de las imágenes aéreas (capítulo tres) y tres métodos (Isodata, Otsu y *Fuzzy*) combinados para las imágenes de superficie (capítulo cuatro). Los umbrales se obtienen a partir de cada una de las componentes espectrales en los respectivos modelos de color utilizados, obteniendo así dos o más regiones por componente espectral. En cualquier caso se trata de técnicas de umbralización adaptativa que obtienen el umbral óptimo mediante la maximización de la distancia entre las dos clases en que se divide el histograma y la minimización de sus áreas. El uso habitual de estas técnicas es su aplicación en binarización de imágenes; si bien, en el presente trabajo se extendieron y adaptaron para el caso de la multi-umbralización, donde se contemplan más de dos regiones, que constituyen la base del clasificador.

Dependiendo del número de umbrales y por tanto de regiones consideradas, el número de clases varía. Así, cuando se utilizan dos umbrales por banda espectral se obtienen ocho posibles clases, para los modelos de color RGB y CIELab, mientras que si se utilizan tres umbrales el número de clases posibles resulta ser de 27. Dependiendo del tipo de aplicación se determina la necesidad de uno o más umbrales. Por ejemplo, en el caso de las imágenes de cobertura terrestre se determina que el número de clases nunca supera el valor de ocho, por lo que la umbralización simple resulta ser suficiente.

Clasificación: el objetivo de la clasificación de la cobertura terrestre consiste en identificar los diferentes tipos de texturas en base a las características del color de los píxeles. Para lograr este objetivo, se parte del método de Otsu (1979), estableciendo dos o más grupos o regiones en el histograma de las bandas espectrales utilizadas. El número de grupos depende del número de umbrales obtenido, en cualquier caso se mide tanto la similitud dentro de cada región como entre las regiones resultantes. Si las regiones presentan un determinado grado de solapamiento se procede a la fusión de las mismas, hasta que todas las varianzas dentro de cada grupo sean menores que las varianzas entre los grupos. Por tanto, la idea principal ha consistido en la realización de un proceso de reagrupamiento de clases hasta encontrar el número óptimo de agrupamientos mediante la maximización de la distancia entre las clases y la minimización de sus áreas de solapamiento.

En conclusión, el método de Otsu (1979) ha resultado ser un método eficiente tanto para la binarización por umbralización como para la multiumbralización, proporcionando los mecanismos necesarios para llevar a cabo la fusión de clases y en consecuencia facilitar el diseño de un procedimiento de clasificación de naturaleza no supervisada, identificado como Otsu-Clustering.

Validación: este proceso se ha orientado en dos sentidos. En primer lugar se verificó si el número de agrupamientos generado por el modelo propuesto es óptimo en comparación con determinados *índices de validación interna* definidos con tal fin. Dada la coincidencia con los resultados arrojados por el índice de validación de Davies Bouldin, podemos concluir que el número de agrupamientos generado por nuestro clasificador para imágenes con menos de ocho clases resulta ser óptimo. Igualmente se puede afirmar que funciona satisfactoriamente para más de ocho clases, validado por el índice de Hartigan.

Se llevó a cabo por último un proceso de validación temática mediante la denominada matriz de error. Se han aplicado dos tipos de análisis, el tradicional o determinista, y el *fuzzy* o difuso. Previamente se determinan las imágenes de referencia o “*ground truth*”. Como conclusiones relevantes derivadas de la evaluación temática se concluye que para imágenes con menos de ocho clases se consigue una precisión global determinista del 93% y difusa del 96%. Para imágenes que presentan más de ocho clases la precisión de la clasificación disminuye al 85% en el caso tradicional, manteniéndose en el 96% en la difusa.

Mediante la metodología de clasificación propuesta se concluye que es posible discriminar las zonas inundadas por efecto del huracán, mediante fotografías aéreas digitales, con una precisión *fuzzy* de entre el 85% y el 96%.

Por tanto como colofón, puede afirmarse que el clasificador propuesto resulta ser suficientemente fiable para las aplicaciones propuestas. De este modo puede usarse eficientemente tanto para imágenes con menos de ocho clases (obtenidas mediante umbralización simple) como para imágenes con más de ocho clases (obtenidas con multi-umbralización).

5.2. LÍNEAS FUTURAS DE TRABAJO.

La estrategia de clasificación no supervisada propuesta se ha comportado de forma robusta en los experimentos realizados, por lo que su aplicación es posible en otras áreas agrícolas o incluso en ámbitos de otra naturaleza. En este sentido caben algunas mejoras que planteamos como desarrollos futuros.

- El método ha mostrado su eficiencia en imágenes a color en dos modelos concretos, RGB y CIELab. Cabe plantear su aplicación a otros modelos de color que se adecuen mejor a los diferentes tipos de fotografías
- Del mismo modo, es posible su utilización en imágenes hiper-espectrales cubriendo un mayor amplio rango del espectro y por tanto con un número de bandas superior a las tres utilizadas en los modelos de color antes mencionados.
- En los métodos propuestos solamente se han considerado los valores espectrales de los píxeles de forma individual. Se podría proponer una ampliación considerando la información contextual contenida en los píxeles vecinos respecto del que se está evaluando en un momento determinado.
- Cabría también la posibilidad de su aplicación para conjuntos de imágenes procedentes de diferentes sensores, donde cada imagen se trataría del mismo modo que las bandas espectrales analizadas en este trabajo.
- Los procesos experimentales realizados no necesitan requisitos específicos de tiempo real, puesto que el estudio se realiza en momentos posteriores a la ocurrencia de los fenómenos meteorológicos. No obstante, cabe la posibilidad de que determinadas aplicaciones requieran procesamiento en tiempo real, por lo que se podría abordar esta posibilidad, máxime teniendo en cuenta que los procedimientos propuestos son altamente paralelizables desde el punto de vista computacional.

BIBLIOGRAFÍA

1. Alcalde, E.; García, M.; Peñuelas, S. (1994). *Informática básica, serie Informática de gestión*. McGraw-Hill, España, 393 p.
2. Anderson, J.R.; Hardy, E.E.; Roach, J.T.; Witmer, R.E. (1976). A land use and land cover classification system for use with remote sensor data. In: *Geological Survey; Government Printing Office: Washington, DC, USA*.
3. Arc View Gis. (2012). Environmental Systems Research Institute Inc. Redlands. USA. <http://www.esri.com/software/arcgis/arcview/index.html>, (accessed March 22, 2012)
4. Ariza, F.J.; Pinilla, C.; Borque, M.J. (1996). Control de Calidad del Proceso de Clasificación de Imágenes de Satélite. *Mapping*. 34: 74-86
5. Aronoff, S. (1985). The Minimum Accuracy Value as an Index of Classification Accuracy. *Photogrammetric Engineering & Remote Sensing*. 51(1):99-111
6. Asmus, V. V.; Buchnev, A. A.; Pyatkin, V. P. (2009). Classification system in the software of the Earth remote sensing data processing. *Pattern Recognition and Image Analysis*, 19: 69-74.
7. Azuaje, F. (2002). A cluster validity framework for genome expression data. *Bioinformatics*. 18:319-320
8. Baik, S.W.; Ahn, S.M.; and Lee, J.W. (2003). Adaptive Segmentation of Remote-Sensing Images for Aerial Surveillance. *Computer Analysis of Images and Patterns*, 2756:549-554.
9. Baraldi, Andrea. (1996). Single Linkage Region Growing Algorithms Based on the Vector Degree of Match. *Geoscience and Remote Sensing*, 34(1): 137-148.
10. Barzily, Z.; Volkovich, Z., Akteke-Örtürk, B. and Weber, G-W. (2008). Cluster Stability Using Minimal Spanning Trees. *International Conference 20th EURO Mini Conference Continuous Optimization and Knowledge-Based Technologies (EurOPT-2008)*, pp. 248-252.
11. Bel Mufti, G.; Bertrand, P.; and El Moubarki, L. (2005). Determining the number of groups from measures of cluster stability, In: *Proceedings of International Symposium on Applied Stochastic Models and Data Analysis*, pp. 404-412.

12. Belart, P.; Pinilla, C.; Ariza, F.J. (2001). Control de calidad de clasificaciones de imágenes por el método de los itinerarios. *Mapping*. Revista internacional de ciencias de la tierra. 73: 34-48.
13. Ben-Hur, A.; Elisseeff, A.; and Guyon, I. A. (2002). Stability based method for discovering structure in clustered data. *Pac Symp Biocomputing*, 7: 6-17.
14. Bertrand, P.; and Mufti, G. B. (2006). Loevinger's measures of rule quality for assessing cluster stability. *Computational Statistics and Data Analysis*, 50(4): 992-1015.
15. Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York.
16. Bezdek, J.C.; and Pal, N.R. (1998). Some new indexes of cluster validity. *IEEE Transactions on Systems, Man and Cybernetics*, 28(B):301-315
17. Bezdek, James C.; Keller, James; Krisnapuram, Raghu; Pal, Nikhil. (2005). *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing (The Handbooks of Fuzzy Sets)*. Springer, 776 p.
18. Bhaskaran, Sunil; Paramananda, Shanka; and Ramnarayan, Maria. (2010). Per-pixel and Object-oriented Classification Methods for Mapping Urban Features Using Ikonos Satellite Data. *Applied Geography*, 30(4): 650-665.
19. Bicego, M.; Dalfini, S.; Vernazza, G.; Murino, V. (2003). Automatic Road Extraction from Aerial Images by Probabilistic Contour Tracking. *Image Processing*, vol. 3. (ICIP 14-17 sep 2003). pp. III- 585-588.
20. Bolshakova, N.; and Azuaje, F. (2003). Cluster validation techniques for genome expression data. *Signal Processing*, 83(4): 825-833.
21. Bouguessa, Mohamed ; Wang, Shengrui; y Sun, Haojun. (2006). An objective approach to cluster validation. *Pattern Recognition Letters*, 27(13):1419-1430.
22. Bow, S. T. (2002). *Pattern Recognition and Image Preprocessing*. 2nd ed. Marcel Dekker. New York, NY, USA. 698 p.
23. Broek, E.L.; Rikxoort, E.M. (2004). Evaluation of color representation for texture analysis, in: *Proceedings of the 16th Belgium Netherlands Artificial Intelligence Conference*, R. Verbrugge, N. Taatgen, and L. R. B. Schomaker, (Eds.), Groningen The Netherlands, 35-42.

24. Buendía-Rodríguez, Enrique; Islas-Gutiérrez, Fabián y Guerra-De la Cruz, Vidal. (2008). Soil Erosion Identification with Landsat ETM+ Images, in Tlaxco and Terrenate, Tlaxcala, Mexico. *Terra Latinoamericana*. 26(1):1-9.
25. Burgos-Artizzu, X.; Ribeiro, A.; Guijarro, M.; Pajares, G. (2011). Real-time image processing for crop/weed discrimination in maize fields. *Computers and Electronics in Agriculture*, 75:337-346.
26. Burgos-Artizzu, X.P.; Ribeiro, A.; Santos, M. (2007). Controlador Borroso Multivariable para el Ajuste de Tratamiento en Agricultura de Precisión. *RIAI*, 4(2): 64-71.
27. Burgos-Artizzu, X.P.; Ribeiro, A.; Tellaeché, A.; Pajares, G.; Fernández-Quintanilla, C. (2010). Analysis of natural images processing for the extraction of agricultural elements. *Image Vision Computing*, 28:138-149.
28. Burgos-Artizzu, X.P.; and Ribeiro, A. (2009). Controlador Borroso Multivariable para el ajuste de tratamientos en Agricultura de Precision. *RIAI*, 4: 64-71.
29. Cadez, I.; Smyth, P.; and Mannila, H. (2001). Probabilistic modelling of transactional data with applications to Profiling, Visualization and Prediction. In *Proceedings of the 7th ACM SIGKDD*, 37-46, San Francisco, CA.
30. Calinski, R.; and Harabasz, J. (1974). A dendrite method for cluster analysis. *Commun Statistics*, 3:1-27.
31. Camps-Valls, G.; Bandos, T. V.; and Zhou, D. (2007). Semi-supervised graph-based hyperspectral image classification. *IEEE transactions on Geoscience and Remote Sensing*, 45: 3044-3054.
32. Cao, Guo; Yang, Xin; Mao, Zhihong. (2005). A two-stage level set evolution scheme for man-made objects detection in aerial images. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. CVPR, 1: 474- 479.
33. Carneggie, D. M.; y Lauer, D. T. (1966). Uses of multiband remote sensing in forest and range inventory. *Photogrammetria* Vol. 21: 115-141.
34. Casillas, A. Gonzalez, R. Martínez. (2003). Algoritmo de Clustering On-Line Utilizando Metaheurísticas y Técnicas de Muestreo. *Procesamiento del Lenguaje Natural*, pp. 57-64.

35. Ceballos-Silva, A. (2003). Evaluating Biophysical Variables to Identify Suitable Areas for Oat in Central Mexico: a Multi-criteria and GIS Approach. *Agriculture, ecosystems & environment*, 95: 371-377.
36. CENAPRED (2012). Centro Nacional de Prevención de Desastres. <http://www.cenapred.gob.mx/es/> (accessed March 22, 2012)
37. Chang, C.C., Lin, C.Y., and Fan, Y.H. (2008). Lossless Data Hiding for Color Images Based on Block Truncation Coding. *Pattern Recognition*, 41: 2347-2357.
38. Chang, C.I.; Du, Y.; Wang, J.; and Guo, S.M. (2006). Survey and Comparative Analysis of Entropy and Relative Entropy Thresholding Techniques. *Vision, Image and, 153(6)*: 837-850.
39. Chang, M. M.; Sezan, M. I.; Tekalp, A. M. (1994). Adaptive Bayesian segmentation of color images. *Journal Electronic Imaging*, 3: 404-414.
40. Chen, D.; Stow, D.A.; and Gong, P. (2004). Examining the effect of spatial resolution and texture window size on classification accuracy: an urban environment case. *International J. of Remote Sensing*, 25:2177-2192.
41. Chen, G.; Jaradat, S. A.; Banerjee, N.; Tanaka, T. S.; Ko, M. S. H.; & Zhang, M. Q. (2002). Evaluation and Comparison of Clustering Algorithms in Analyzing ES Cell Gene Expression Data. *Statistica Sinica*, 12: 241-262.
42. Chen, Q.; Luo, J.; and Zhou, C. (2003). A geostatistic based segmentation approach for remotely sensed images. *24th Asian Conference on Remote Sensing, Busan, South Korea*. Pp:2-4.
43. Chen, W. and Fang, K. (2008). Multilevel thresholding algorithm based on particle swarm optimization for image segmentation. In: *Control Conference, 27th Chinese, Kunming, Yunnan, China*, pp:348-351.
44. Chenaoua, K. S.; Bouridane, A., and F. Kurugollu, A. (2003). Unsupervised histogram based color image segmentation,” in: *10th IEEE Int. Conf. Electronics, Circuits and Systems*, 1: 240–243.
45. Cheng, H. D. and Chen, Y. H. (1999). Fuzzy partition of two-dimensional histogram and its application to thresholding. *Pattern Recogn.* 32: 825-84.
46. Cheng, H.D.; Jiang, X.H.; Sun, Y.; Wang, J. (2001). Color image segmentation: advance and prospects. *Pattern Recognition*, 34:2259-2281.

47. Cheng, Wen-Chun; Chang, Jyh-Chian; Chang, Chien-Ping; Su, Yu and Tu, Te-Ming. (2008). A Fixed-Threshold Approach to Generate High-Resolution Vegetation Maps for IKONOS Imagery. *Sensors*. 8: 4308-4317.
48. Chi, C. and Liu, Z. (2008). Interpretation of Landslide from SPOT-5 Imageries in the Three Gorges Reservoir Area. In: *Earth Observation and Remote Sensing Applications EORSA*. Pp 1-5.
49. Chou, H. C.; Su, M. C.; and Lai, E. (2004). A New Cluster Validity Measure and Its Application to Image Compression. *Pattern Analysis and Applications*, 7: 205-220, (SCI, ED).
50. Chuvieco, E. (2008). *Teledetección Ambiental*. 3a ed. Ariel. Barcelona, España. 592 p.
51. Chuvieco, E.; and Santis, A. (2007), Reflexiones sobre el empleo de modelos empíricos y de simulación en la estimación de variables ambientales: aplicaciones en gestión de incendios forestales. *XII Congreso de la Asociación Española de Teledetección. Hacia un mejor entendimiento de la dinámica global y regional*. (sep. 19-21) Mar del Plata, Argentina, pp. 161-169.
52. Civco, D. L. (1993). Artificial neural networks for land-cover classification and mapping. *International Journal of Geophysical Information Systems*, 7(2):173-186.
53. Cochran, W. G., (1953). *Sampling Technics*. John Wiley & Sons, Inc.
54. Coltuc, D.; Bolon, P. and Chassery, J-M. (2006). Exact histogram specification. *IEEE Transactions on Image Processing* 15(5):1143-1152.
55. Comaniciu, D.; and Meer, P. (2002). Mean shift: a robust approach toward feature space analysis. *IEEE PAMI.*, Vol. 24(5):603-619.
56. Comaniciu, D.; Meer, P. (1997). Robust analysis of feature spaces: color image segmentation. In: *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, San Juan, PR, pp. 750-755.
57. Congalton, R. (2010). How to Assess the Accuracy of Maps Generated from Remotely Sensed Data. *Manual of Geospatial Science and Technology*, 2nd Edition. John Bossler. (Editor). Taylor & Francis, Boca Raton, FL pp. 403-421

58. Congalton, R. and Green, K. (1993). A practical look at the sources of confusion in error matrix generation. *Photogrammetric Engineering and Remote Sensing*, 59: 641–644.
59. Congalton, R. G.; and Green, K. (1999). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Lewis Publishers, Boca Raton, FL. 160 p.
60. Congalton, R.G. (2004) Putting the map back in map accuracy assessment. *Remote Sensing and GIS Accuracy Assessment*; Lunetta, R.S., Lyon, J.G., Eds.; Lewis: Boca Raton, FL, USA; 292 p.
61. Congalton, R.G.; Green, K. (2009). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. 2nd ed.; CRC/Taylor & Francis: Boca Raton, FL, USA; p. 178.
62. Corney, David; Byrne, Emma; Buxton, Bernard and Jones, David. (2008). A Logical Framework for Template Creation and Information Extraction. *Data Mining: Foundations and Practice*, pp. 79-108.
63. Correa-Tome, F.E., Sanchez-Yanez, R.E. and Ayala-Ramirez, V. (2011). Comparison of perceptual color spaces for natural image segmentation tasks. *Optical Engineering*, 50(11):117-203.
64. Cortijo, F.J.; Pérez de la Blanca, N.; Abad, J.; Damas, S. (1997). A comparison of multispectral image classifiers using high dimensional simulated data sets. In: *Proceedings of VII National Symposium on Pattern Recognition and Image Analysis*. Barcelona, Spain; pp. 365-370.
65. Couturier, S; Vega, A; Mas, J-F; Tapia, V.; López-Granados, E. (2008). Evaluación de confiabilidad del mapa del Inventario Forestal Nacional 2000: diseños de muestreo y caracterización difusa de paisajes. *Investigaciones Geográficas. Boletín del Instituto de Geografía, UNAM*, (67), pp. 20-38. ISSN 0188-4611
66. Cutting, D., Karger, D., Pedersen, J. and Tukey, J. (1992). Scatter/gather: a cluster-based approach to browsing large document collection. In: *Proceeding of the 15th ACM SIGIR Conference*, 318-329, Copenhagen, Denmark.
67. Davies, D.L.; and Bouldin, D.W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 1(2): 224-227.

68. Davis, J.C., (1986). *Statistics and Data Analysis in Geology*. Wiley & Sons, New York. 656 p.
69. Devi, P.L., Varadarajan, S. (2011). Image Segmentation and Techniques: a review. *International Journal of Advanced Research in Technology*, 1(2):118-127.
70. Dhillon, I.; Mallela, S.; and Kumar, R. (2002). Enhanced Word Clustering for Hierarchical Text Classification. *Proceeding of the 8th ACM SIGKDD*, 191-200, Edmonton, Canada
71. Dicks. S.E.; and T. H. C. Lo., (1990). Evaluation of Thematic Map Accuracy in a Land-Use and Land-Cover Mapping Program. *Photogrammetric Engineering & Remote Sensing*. 55(4): 475-478.
72. Dimap (2012). Digital Image Processing, S.L. <http://www.dimap.es/>
73. Dimitriadou, E.; Dolnicar, S.; and Weingessel, A. (2002) An examination of indexes for determining the Number of Cluster in binary data sets. *Psychometrika*. 67(1): 137-160.
74. Du, Y.; Chang, C. I.; and Thouin, P.D. (2004). Unsupervised approach to color video thresholding. *Optical Eng.* 43(2):282-289.
75. Du. Q. (2007). Unsupervised real-time constrained linear discriminant analysis to hyperspectral image classification. *Patt. Recog.* 40:1510-1519.
76. Duda, R. O. and Hard, P.E. (1973). *Pattern Classification and scene analysis*. Wiley. New York.
77. Duda, R. O.; Hard, P.E.; and Stork, D.G. (2001). *Pattern Classification*. 2nd ed. John Wiley. New York, NY, USA. 654 p.
78. Dudoit, S.; and Fridlyand, J. (2002). A prediction-based resampling method for estimating the number of clusters in a dataset. *Genome Biology*, 3(7): 0036.1-21.
79. Dunn. J. (1974). Well separated clusters and optimal fuzzy partitions. *J. Cybernetics*, 4: 95-104.
80. Envi (2012). <http://www.exelisvis.com/language/en-us/products/services/envi.aspx>. (accessed March 20, 2012)
81. Erdas Imagine (2012). <http://www.erdas.com/products/ERDASIMAGINE/ERDASIMAGINE/Details.aspx>, (accessed March 20, 2012)

82. Escalera H.; De la; A. (2001). *Visión por Computador, Fundamentos y Métodos*. Pearson Educación. Madrid, España. 304 p.
83. Ester, M.; Frommelt, A.; and Kriegel, H-P. and Sander, J. (2000). Spatial Data Mining: Database Primitives, Algorithms and Efficient DBMS Support, 193-216, in: Springer, *Data Mining and Knowledge Discovery. Computer Science*.
84. Fonseca, L.; Namikawa, L.; Castejon, E.; Carvalho, L.; Pinho, C. and Pagamisse, A. (2011). Image Fusion for Remote Sensing Applications. *Image Fusion and Its Applications*. ISBN 978-953-307-182-4.
85. Foss, A.; Wang, W.; and Zaane, O. (2001). A non-parametric approach to Web log analysis. *1st SIAM ICDM, Workshop on Web Mining*. 41-50, Chicago, IL.
86. Fu K.S. and Mui, J.K. (1981). A Survey on Image Segmentation. *Pattern Recognition*. 13:3-16,
87. García-Alegre, M.C; Ribeiro, A; Guinea, D.; y Cristobal, G. (2000). Eggshell defects detection based on color processing. *SPIE, Electronic Imaging Conf.*, CA.
88. Gashnikov, M. V.; Chernov, a. V.; and Chupshev, N. V. (2009). Color Correction of Vehicle Images During the Sequential Registration of Color Channels. *Pattern Recognition and Image Analysis*, 19(1): 106-108.
89. Gebhardt, S.; Kaühbauch, W.A. (2007). A new algorithm for automatic Rumex obtusifolius detection in digital image using colour and texture features and the influence of image resolution. *Precis. Agric.*, 8:1-13.
90. Gebhardt, S.; Schellberg, J.; Lock, R.; Kaühbauch, W.A. (2006). Identification of broad-leaved dock (Rumex obtusifolius L.) on grass land by means of digital image processing. *Precis. Agric.*, 7:165-178.
91. Ginevan, M. E., (1979). Testing Land-Use Map Accuracy: Another look. *Photogrammetric Engineering & Remote Sensing*. 45(10):1371-1377.
92. Gómez, A. Darío M. (1994). *Reconocimiento de Formas y Visión Artificial*. Assison-Wesley Iberoamericana.
93. Gonçalves, M L.; Márcio L. A.; Netto and Costa, J. A. F. (2008), Cluster Analysis of Land-Cover Images Using Automatically Segmented SOMs with Textural Information. *Intelligent Data Engineering and and Automated Learning. Computer Science*, 5326:483-490.

94. Gonzales-Barron, U.; Butler, F. (2006). A comparison of seven thresholding techniques with the k-means clustering algorithm for measurement of bread-crumbs features by digital image analysis. *J. Food Eng.* 74:268-278.
95. Gonzalez, R. and Woods, R. E. (1996). Tratamiento Digital de Imágenes. Addison Wesley, Wilmington, EE.UU. 774 p.
96. Gonzalez, R. and Woods, R. E. (2008). Digital Image processing, third ed. Prentice Hall, New Jersey, EE.UU. 954 p.
97. Goodman, A. Leo. (1965). Simultaneous Confidence Intervals for Multinomial Proportions. *Technometrics*, 7(2):247-254.
98. Gottschalk, R.; Burgos-Artizzu, X.P.; Ribeiro, A.; Pajares, G.; Miralles, A.S. (2010). Real-time image processing for the guidance of a small agricultural field inspection vehicle. *Int. J. Intell. Syst. Tech. Appl.* 8:434-443.
99. Guidi, L., Ibanez, F., Calcagno, V., Beaugrand, G. (2009). A new procedure to optimize the selection of groups in a classification tree: Applications for ecological data. *Ecological Modelling*, 220:451-461
100. Guijarro, M. Pajares, G. (2009a). On combining classifiers through a fuzzy multicriteria decision making approach: Applied to natural textured images. *Expert Systems with Applications*, 36:7262-7269.
101. Guijarro, M.; Pajares, G. and Herrera, P. J. (2009b). Image-Based Airborne Sensors: A Combined Approach for Spectral Signatures Classification through Deterministic Simulated Annealing. *Sensors*, 9:7132-7149.
102. Guijarro, M.; Pajares, G.; Riomoros, I.; Herrera, P.J.; Burgos-Artizzu, X.; Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture*, 75:75-83.
103. Gunter, S.; Bunke H. (2003). Validation Indices for Graph Clustering. *Pattern Recognition Letters*. 24:1107-1113.
104. Halkidi, M.; Batistakis Y.; and Vazirgiannis, M. (2002a). Cluster validity methods: part I. *SIGMOD Rec.*, 31(2):40-45.
105. Halkidi, M.; Batistakis, Y.; and Vazirgiannis, M. (2001). On Clustering Validation Techniques. *Intelligent Information Systems Journal*, 17(2-3): 107-145.

106. Halkidi, M.; Batistakis, Y.; and Vazirgiannis, M. (2002b). Cluster Validity Methods: Part II. *SIGMOD Record*, 31(3):19-27.
107. Handl, J.; Knowles, J.; and Kell, D. B. (2005). Computational cluster validation in post-genomic data analysis. *Bioinformatics (Oxford, England)*, 21(15): 3201-3212.
108. Hartigan, J. A. and Wong, M. A. (2009). A K-Means Clustering. *Society*, 28(1): 100-108.
109. Hartigan, J.A. (1985). Statistical theory in clustering. *J. Classification*. 2:63-76.
110. Hartigan, J. A. (1975). *Clustering algorithms*. New York: Wiley.
111. Hay, A.M., (1979). Sampling Designs to Test Land-Use Map Accuracy. *Photogrammetric Engineering & Remote Sensing*. 45(4):529-533.
112. Heer, J. and Chi, E. (2001). Identification of Web user traffic composition using multimodal clustering and information scent. *1st SIAM ICDM, Workshop on Web Mining*, 51-58, Chicago, IL.
113. Herrera, P. Javier; Pajares, Gonzalo; Guijarro, María. (2011). A segmentation method using Otsu and fuzzy k-Means for stereovision matching in hemispherical images from forest environments. *Applied Soft Computing*, 11: 4738-4747.
114. Herrera, P.J.; Pajares, Gonzalo; Guijarro, M.; and Ruz, J.J. (2009a). Segmentation and Stereoscopic Correspondence in Images Obtained with Omnidirectional Projection for Forest Environments. *Advances in Stereo Vision, Intechopen.com*. pp. 41-56.
115. Herrera, Pedro Javier; Pajares, Gonzalo; Guijarro, M.; Ruz, José J.; Cruz, Jesús M.; and Montes, Fernando. (2009b). A Featured-based Strategy for Stereovision Matching in Sensors with Fish-eye Lenses for Forest Environments. *Sensors*, 9(12):9468-92.
116. Hu, Chunchun; Meng, Lingkui and Shi, Wenzhong. (2008). Fuzzy Clustering Validity for Spatial Data. *Geo-spatial Information Science*, 11(3):191-196.
117. Hu, R. and Damper, R.I. (2008). A 'Non Panacea Theorem' for classifier combination. *Pattern Recognition*, 41(8):2665-2673.

118. Hu, X, Tao, CV and Prenzel, B. (2005). Automatic Segmentation of High-resolution Satellite Imagery by Integrating Texture, Intensity, and Color Features. *Photogrammetric engineering and remote*, 71(12): 1399-1406.
119. Huang, L.K.; Wang, M.J. (1995). Image thresholding by minimising the measures of fuzziness. *Patt. Recog.* 21:41-51.
120. Inampudi, R.B.; Guntupalli, S.P.; and Rao, A.A. (2002), Knowledge Based Approach for Automated Digital Image Processing. In: Geoscience and Remote Sensing Symposium, *IGARSS'02. IEEE International*. Pp. 1340-1342.
121. INEGI (2012). Instituto Nacional de Estadística y Geografía. <http://www.inegi.org.mx/> (accessed March 20, 2012)
122. Jafar, I. and Ying, H. (2007). A new method for image contrast enhancement based on automatic specification of local histograms. *Int. Journal of Computer Science and Network Security*, 7(7):1-10.
123. Jain, A. K.; and Dubes, R. C. (1988). *Algorithms for Clustering Data. Prentice-Hall advanced reference series*. Prentice-Hall, Inc., Upper Saddle River, NJ. 320 p.
124. Jain, A. K.; Duin, R.; and Mao, J. (2000). Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 22(1):4-37.
125. Jain, A. K.; Murty, M. N.; and Flynn, P. J. (1999). Data clustering: a review. *ACM Computing Surveys*, Vol. 31 (3):264-323.
126. Janssen Lucas L. F.; Van der Wel F. J. M. (1994). Accuracy Assessment of Satellite Derived Land-Cover Data: A Review. *Photogrammetric Engineering & Remote Sensing*. Vol. 60(4):419-426.
127. Jawahar, C. V.; Biswas, P. K. and Ray, A. K. (1997). Investigations on fuzzy thresholding based on fuzzy clustering. *Pattern Recog.* 30(10):1605-1613.
128. Jiji, G.W.; Ganesan, L. (2008). A new approach for unsupervised segmentation. *Applied Soft Computing Journal*.
129. Jochem, Andreas; Hollaus, Markus; Rutzinger, Martin and Höfle, Bernhard. (2010). Estimation of Aboveground Biomass in Alpine Forests: A Semi-Empirical Approach Considering Canopy Transparency Derived from Airborne LiDAR Data. *Sensors*, 11(1): 278-295.

130. Johansen, Kasper, Tiede, Dirk, Blaschke, Thomas, Arroyo, Lara a. and Phinn, Stuart (2011), Automatic Geographic Object Based Mapping of Streambed and Riparian Zone Extent from LiDAR Data in a Temperate Rural Urban Environment, Australia. *Remote Sensing*, 3(6): 1139-1156.
131. Kabir, Hasanul; and Al-Wadud, A. (2009). Brightness Preserving Image Contrast Enhancement Using Weighted Mixture of Global and Local Transformation Functions. *The International Arab Journal of Information Technology*, 7(4): 403-411.
132. Kapur, J. N.; Sahoo, P. K.; and Wong, A. K. C. (1985). A new method for gray-level picture thresholding using the entropy of the histogram. *Graph. Models Image Process.* 29:273-285.
133. Kirby, R. L. and Rosenfeld, A. (1979). A note on the use of gray level, local average (gray level) space as an aid in threshold selection. *IEEE Trans. Syst. Man Cybern.* SMC-9:860-864.
134. Kirk, K.; Andersen, H.J.; Thomsen, A.G.; Jørgensen, J.R. (2009). Estimation of leaf area index in cereal crops using red- green images. *Biosyst. Eng.* 104:308-317.
135. Kittler, J.; Illingworth, J. (1986). Minimum error thresholding. *Patt. Recog.* 19: 41-47.
136. Kovács, F.; Legány, C.; and Babos, A. (2005). Cluster validity measurement techniques. Proceedings of the 6th International Symposium of Hungarian Researchers on *Computational Intelligence, Budapest.*
137. Krzanowski, W.; and Lai, Y. (1985). A criterion for determining the number of groups in a dataset using sum of squares clustering. *Biometrics.* 44:23-34.
138. Kumar, S.; Ghosh, J.; Crawford, M.M. (2001). Best-bases feature extraction for pairwise classification of hyperspectral data. *IEEE Trans. Geoscience and Remote Sensing*, 39(7):1368-1379.
139. Kuncheva, L.I. (2004). *Combining Pattern Classifiers: Methods and Algorithms.* Wiley: New York, NY, USA; p. 376.
140. Kurugollu F; Sankur, B.; Harmanci, A.E. (2001). Color image segmentation using histogram multithresholding and fusion. *Image and Visión Computing* 19:915-928

141. Lam, B. S. Y.; Yan, Hong. (2007). Assessment of Microarray Data Clustering Results Based on a New Geometrical Index for Cluster Validity. *Soft Comput.*, 341-348
142. Lange, Tilman; Roth, Volker; Braun, M. L y Buhmann J.M. (2004). Stability-based validation of clustering solutions. *Neural Computation*, 16:1299-1323.
143. León, Emilio A.; Duffy, Sergio J.; Stevenson, Mark A.; Lockhart, Caryl; Späth, Ernesto J.A. (2007). *Sistema AVE de información Geográfica para la asistencia en la vigilancia epidemiológica de la influenza aviar, basado en el riesgo*. FAO e INTA.
144. Levachkine, S. ; M. Torres & M. Moreno. (2003). Intelligent Segmentation of Color Geo-Images. IEEE International Geoscience and Remote Sensing Symposium, IGARSS', Toulouse, France, IEEE Catalog Number: 03CH37477C, Library of Congress Number: 2003104493, ISBN: 0-7803-7930-6.
145. Levine, E; y Domany, E. (2001). Resampling method for unsupervised estimation of cluster validity. *Neural Computation*, 13:2573-2593.
146. Li, Hongfen; Hu, Guangdao; Li, Jiang-feng. (2008). Multi-Classifer Systems (MCSs) of Remote Sensing Imagery Classification Based on Texture Analysis. *ISICA*, pp. 740-749
147. Lim, Y.W; Lee, S.U. (1990). On the color image segmentation algorithm based on the thresholding and fuzzy c-means techniques. *Pattern Recognition*, 23:935-952.
148. Ling, P.P.; Ruzhitsky, V.N. (1996). Machine vision techniques for measuring the canopy of tomato seedling. *J. Agric. Eng. Res.* 65:85-95.
149. Liu, Xuefeng ; Ji, Guangrong ; Zhao, Wencang ; Cheng; Junna. (2007). Improved Approach of Seed Point Selection in RPCCL Algorithm for Aerial Remote Sensing Hyper-spectral Data Clustering with Data Dimensionality Reduction. *Software Engineering, Artificial Intelligence*. 8th ACIS International Conference on Networking, and Parallel/Distributed Computing, (SNPD 2007), 1:860-865.
150. Lloyd, D.E. (1985). *Automatic target classification using moment invariant of image shapes*. Technical Report, RAE IDN AW126, Farnborough, UK.

151. Lu, D.; and Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sensing*, 28(5):823-870.
152. Lucchese, L. and Mitra, S. K. (2001). Color image segmentation: a state-of-the-art survey. *Proc. Indian Nat. Sci. Acad. (INSA-A)* 67(2):207-221.
153. Lunetta R. S.; and Lyon, J. G. (2004). *Remote sensing and GIS accuracy assessment*. Lewis Publishers, Boca Raton, FL, USA. 292p.
154. Macedo, A.; Pajares, G.; Santos, M. (2010). Unsupervised classification with ground color cover images. *Agrociencia*, 44(6):711-722. ISBN: 1405-3195
155. Macedo-Cruz, A.; Pajares, G.; Santos, M.; Villegas-Romero, I. (2011). Digital image sensor-based assessment of the status of Oat (*Avena sativa* L.) Crops after frost damage. *Sensors*, 11:6015-6036. ISSN 1424-8220.
156. Macedo Cruz, A.; Santos, M.; Pajares, G.; Villegas-Romero, I. (2012). Unsupervised classification of aerial images based on the Otsu's method. In: *Deforestation Around the World*. Paulo Moutinho (ed), InTech, pp.171-186. ISBN 978-953-51-0417-9 <http://www.intechopen.com/books/deforestation-around-the-world>
157. Macedo Cruz, A.; Santos, M.; Pajares, G; Villegas Romero, I. (2011). Intelligent classification of agricultural land covers in the low catchments of "La Sabana" river, Guerrero". In: *IEEE Sección México, Vigésima segunda Reunión Internacional de Otoño, ROC&C'2011*, Acapulco, Guerrero, México (nov 27 - dic 3) COMPUTACION (CP-76)140. ISBN: 978-607-95630-1-1.
158. Macedo-Cruz, Antonia; Pajares-Martinsanz, Gonzalo; Santos-Peñas, Matilde. (2010). Unsupervised classification of images in RGB color model and cluster validation techniques. In: *WORLDCOMP2010, Proceedings of the International Conference on Image Processing, Computer Vision, & Pattern Recognition, (IPCV 2010)*. Hamid R. Arabnia, Leonidas Deligiannidis, Gerald Schaefer, Ashu M. G. Solo (eds.), Las Vegas, Nevada, USA, July 12-15. CSREA Press 2010 (Eds) 2:526-532. ISBN 1-60132-154-6.
159. Macedo Cruz, A; Pajares Martinsanz, G; Santos Peñas, Matilde; Villegas Romero, I. (2010). Unsupervised classification of ground cover based on the Otsu's method. In: *Memoria digital del ROPEC'2010 INTERNACIONAL, XII*

- Reunión de Otoño de Potencia, Electrónica y Computación*, nov 10-12, Manzanillo, Colima, México. 2010:114-121. ISBN: 978-607-95476-1-5.
160. Macedo Cruz, Antonia; Pajares Martinsanz, Gonzalo; Santos Peñas, Matilde. (2009). Reconocimiento de patrones del paisaje por umbralización en imágenes teledetectadas. In: *VII Congreso Internacional Sobre Innovación y Desarrollo Tecnológico*. Memoria del (CIINDET-2009), ISBN: 978-607-95255-1-4.
161. Maini, Raman; and Himanshu, Aggarwal. (2010). A Comprehensive Review of Image Enhancement Techniques. *Arxiv preprint arXiv:1003.4053*, 2(3): 8-13.
162. Martín-H, J.A.; Santos, M.; Farias, G.; Duro, N.; Sanchez, J.; Dormido, R.; Dormido-Canto, S.; Vega, J.; Vargas, H. (2009). Dynamic clustering and modeling approaches for fusion plasma signals. *IEEE Trans. Instrum. Meas.* 58: 2969-2978.
163. Martín-H, J.A.; Santos, M.; de Lope, J. (2010). Orthogonal variant moment features in image analysis. *Inform. Sci.* 180:846-860.
164. Mas, J.F; Reyes, J; Pérez, A. (2003). Evaluación de la confiabilidad temática de mapas o de imágenes clasificadas: una revisión. *Investigaciones Geográficas*, 51. Boletín del Instituto de Geografía, UNAM, pp. 53-72.
165. Mather, P. M. (2004). *Computer processing of remotely-sensed images: An introduction*. Third edition, Willey & sons.
166. Maulik, U.; Bandyopadhyay, S. (2002). Performance evaluation of some clustering algorithms and validity indices. *IEEE Trans Pattern Anal Mach Intell*, 24(12): 1650-1654
167. Medina, F. (1998). Tamaño Óptimo de Muestra en Encuestas de Propósitos Múltiples. CEPAL. *Memoria del Taller Regional sobre Planificación de Encuestas en Hogares*, Santiago de Chile. (agosto98).
168. Meer, F. (2002). Physical principles of optical remote sensing. *Spatial statistics for remote sensing*. 1:27-40.
169. Mendoza, F.; Dejmek, P.; Aguilera, J.M. (2006). Calibrated color measurements of agricultural foods using image analysis. *Postharvest Biol. Technol.* 41(3): 285-295.

170. Meyer, G.E.; Camargo Neto, J. (2008). Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* 63(2):282-293.
171. Meyer, G.E.; Camargo Neto, J.; Jones, D.D.; Hindman, T.W. (2004). Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images. *Comput. Electron. Agric.* 42:161-180.
172. Mezaris V.; Kompatsiaris I.; Srintzis MG. (2004). Still image segmentation tools for object-based multimedia applications. *International Journal of pattern recognition and artificial intelligence*, 18(4):701-726.
173. Milligan, G.W.; and Cooper, M.C. (1985). An examination of procedures for determining the number of cluster in a data set. *Psychometrika*, 50:159-179.
174. Montiel, E.; Aguado, A.S.; and Nixon, M.S. (2005). Texture classification via conditional histograms. *Pattern Recognition Letters*, 26:1740-1751.
175. Mufti, G. B.; Bertrand, P.; and Moubarki, L. E. (2005). Determining the number of groups from measures of clusters validity. In: *Proceedings of ASMDA2005*, pp. 404-414.
176. Neto, J.C. (2004). A Combined Statistical-Soft computing approach for classification and mapping weed species in minimum tillage systems. *EDT collection for University of Nebraska - Lincoln*. Paper AAI3147135.
177. Niclós, R; Estrela, M. J.; Valiente, J. A. y Barberá, M. J. (2009). Generación de Mapas de usos de suelo periódicos a escala regional con imágenes MODIS: aplicación a la Comunidad Valenciana. *Teledetección: Agua y desarrollo sostenible. XIII Congreso de la Asociación Española de Teledetección*. Calatayud, 23-26 de septiembre 2009, pp. 213-216.
178. Nimbarte, Nita M.; and Mushrif, Milind M. (2010). Multi-level Thresholding Algorithm for Color Image Segmentation. *Second International Conference on Computer Engineering and Applications*,(19-31mar) 2:231-233.
179. Otsu, N. (1979). A threshold selection method from gray level histogram. *IEEE Trans. Systems Man and Cybernetics*. 9(1):62-66.
180. Pajares M, G.; de la Cruz, M. (2001). *Visión por Computador, Imágenes Digitales y Aplicaciones*. Rama, Madrid España. 764 p.

181. Pajares, G. and Cruz, J. M. (2002). Clasificación de Texturas Naturales mediante K-Means. *Revista Electrónica de Visión por Computador* <http://www.revc.uab.es/revista/06/> (ISSN 1575-5258), 6:1-18, Ed. Centre de Visió per Computador, Universidad Autònoma, Barcelona.
182. Pajares, G. and Cruz, J.M. (2007b). *Visión por Computador: Imágenes Digitales y Aplicaciones*, 2ª ed. RA-MA, Madrid. 768 p.
183. Pajares, G. and Cruz, J.M. (2007a). *Ejercicios resueltos de Visión por Computador*, RA-MA, Madrid. 338 p.
184. Pajares, G., Guijarro, M., Ribeiro, A. (2010). A Hopfield Neural Network for combining classifiers applied to textured images. *J. Neural Networks*, 23(1):144-153.
185. Pal, N. R. and Biswas, J. (1997). Cluster Validation Using Graph Theoretic Concepts. *Pattern Recognition*, 30(6):847-857
186. Pal, N.R. and Pal, S.K. (1993). A Review on Image Segmentation Techniques. *Pattern Recognition*, 26(9):1277-1294.
187. Pinilla, C. (1995). *Elementos de Teledetección*. RA-MA. Madrid.
188. Plataniotis K.N.; Venetsanopoulos A.N, (2000). *Color Image Processing and Applications*. Springer, Berlín, Alemania. 375 p.
189. Plaza, A.; Du, Q.; Chang, Y.L. (2009). High Performance Computing for Hyperspectral Image Analysis: Perspective and State-of-the-art. *Geoscience and Remote Sensing Symposium, 2009 IEEE International, (IGARSS 2009)*, 5:72-75.
190. Polidório, A M.; Flores, F. C. ; Imai, N. N.; Tommaselli, A. M. G.; Franco, C. (2003). Automatic Shadow Segmentation in Aerial Color Images. XVI Brazilian Symposium on *Computer Graphics and Image Processing (SIBGRAPI'03)*, pp.270-277.
191. Proespacio (2012). Agrupación de empresas. <http://www.proespacio.org/proespacio/index.php> (accessed on 2 April 2012).
192. Pun, T. (1980). A new method for grey-level picture thresholding using the entropy of the histogram. *Signal Process.* 2:223-237.
193. Quesenberry, C.P. and Hurst, D. C. (1984). Large Sample Simultaneous Confidence Intervals for Multinomial proportions. *Technometrics*, 6(2):191-195.

194. Raof, A.; Zaleha, S.; Sahidan, S.I.; Mashor, N. Noor, S.S, Idris, M and Hasan, H (2008), Color Thresholding Method for Image Segmentation Algorithm of Ziehl-Neelsen Sputum Slide Images. *Library*, (Cce): 212-217.
195. Reid, J.F.; Searcy, S.W. (1987). Vision-based guidance of an agricultural tractor. *IEEE Control Syst.* 7:39-43.
196. Ribeiro, A., M. C. García-Alegre, D. Guinea, and G. Cristobal. (2000). Automatic rules generation by G.A. for eggshell defect classification. *Eur. Congr. Comput. Methods Appl. Sci. Eng.*, Barcelona, Spain."
197. Ribeiro, A.; Fernández-Quintanilla, C.; Barroso, J.; García-Alegre, M.C. (2005). Development of an image analysis system for estimation of weed. In: *Proceedings of the 5th European Conf. On Precision Agriculture (5ECPA)*, pp. 169-174.
198. Ribeiro, A., Ranz, J., Burgos-Artizzu, X.P., Pajares, G., Sánchez del Arco, M.J. and Navarrete, L. (2011). An Image Segmentation Based on a Genetic Algorithm for Determining Soil Coverage by Crop Residues. *Sensors*, 11:6480-6492.
199. Ridler, T. W. and Calvard, S. (1978). Picture thresholding using an iterative selection method. *IEEE Transactions on Systems, Man and Cybernetics*, 8(8): 630-632.
200. Robertson, A.L. (1976). The CIE 1976 color difference formulae. *Color Res. Appl.* 2:7-11.
201. Rousseeuw, P. J. (1987). Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Journal of Computational and Applied Mathematics.* 20:53-65.
202. Rowe N.C, L.L. Grewe. (2001). Change detection for linear features in aerial photographs using edge-finding. *IEEE Transactions on Geoscience and Remote Sensing* 39:1608-1612.
203. Rud, R.; Shoshany, M.; Alchanatis, V.; and Cohen, Y. (2006). Application of spectral features' ratios for improving classification in partially calibrated hyperspectral imagery: a case study of separating Mediterranean vegetation species. *Journal Real-Time Image Processing*, 1:143-152.

204. Ruiz, L. A. (1998). Introducción al tratamiento digital de imágenes en teledetección. *Publicaciones de la Universidad politécnica de Valencia*.
205. Ruiz-Ruiz, G.; Gómez-Gil, J.; Navas-Gracia, L.M. (2009). Testing different color spaces based on hue for the environmentally adaptive segmentation algorithm (EASA). *Comput. Electron. Agric.* 68:88-96.
206. Saeed, F. (2006). Digital Image Processing of High Resolution Aerial Photograph of Shallow Marine Sanctuary. Victoria, Australia. In: *Space Technologies, international Conference*. pp.179-184.
207. Sahoo, P., Wilkins, C.; and Yeager, J. (1997). Threshold selection using Renyi's entropy. *Pattern Recognition*, 30:71-84.
208. Sainz-Costa, N., Ribeiro, A., Burgos-Artizzu, X.P., Guijarro, M., Pajares, G. (2011). Mapping Wide Row Crops with Video Sequences Acquired from a Tractor Moving at Treatment Speed. *Sensors*, 11:7095-7109.
209. Saitta, S.; Raphael, B. and Smith, I.F.C. (2007). A bounded index for cluster validity. In: *Computer Science (United States)*, 4571:174-187.
210. Salton, G. (1980). Automatic Information Retrieval. *Computer Publication*, 13(9): 41-56.
211. Sangwine, S.J. (2000). Colour in image processing. *Electronics Communication Engineering Journal*, 12:211-219.
212. Sanchez-Llado, F.J., Pajares, G., López-Martínez, C. (2011). Improving the Wishart Synthetic Aperture Radar image classifications through Deterministic Simulated Annealing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66:845-857.
213. Santos, M. (2011). Un enfoque aplicado del control inteligente. *RIAI*, 8(4):283-296.
214. Santos, M.; Pajares, G.; Portela, M.; Cruz, de la J.M. (2003). A new wavelets image fusion strategy. Lecture Notes in Computer Science. *Springer-Verlag, Berlin Heidelberg*, 2652:919-926.
215. Secord, J.; Zakhor, A. (2007). Tree Detection in Urban Regions Using Aerial Lidar and Image Data. *IEEE Geoscience and Remote Sensing Letters*, 4(2): 96-200.

216. Sezgin, Mehmet; and Sankur, B. (2004). Survey over Image Thresholding Techniques and Quantitative Performance Evaluation. *Journal of Electronic Imaging*, 13(1):146-165.
217. Shah, C.A; Arora, M.K.; Varshney, P.K. (2004). Unsupervised classification of hyperspectral data: An ICA mixture model based approach. *Int. J. Remote Sens.* 25:481-487.
218. Shawe-Taylor, J.; and Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*. Cambridge University Press.
219. Shih F, Cheng S. (2005). Automatic seeded region growing for color image segmentation. *Image and Vision Computing*. 23(10):877-886.
220. Shrestha, D.S.; Steward, B.L.; Birrell, S.J. (2004). Video processing for early stage maize plant detection. *Biosystems Engineering*, 89:119-129.
221. Shu, G.; Zeng, B.; Chen, Y. P. and Smith, O. H. (2003) Performance assessment of kernel density clustering for gene expression profile data. *Comparative and Functional Genomics*, 4(3):287-299.
222. SIAP. (2012). Anuario Estadístico de la Producción Agrícola; Available online: <http://www.siap.gob.mx> (accessed on 11 April 2012).
223. SIGSA (2012). Sistemas de Información Geográficos S.A. de C.V. <http://www.sigsa.info/> (accessed on 11 April 2012).
224. Silveira, G.F.; Carvalho, J.R.H.; Madrid, M.K.; Rives, P.; Bueno, S.S. (2001). A fast vision-based road following strategy applied to the control of aerial robots. Proceedings of XIV Brazilian Symposium, *Computer Graphics and Image Processing*, pp. 226-231.
225. Simone, G.; Farina, A.; Morabito, F.C.; Serpico, S.B.; & Bruzzone, L. (2002). Image fusión techniques for remote sensing applications. *Information Fusion*, 3:3-15.
226. Sirmacek, B.; Unsalan, C. (2008). Building Detection from Aerial Images Using Invariant Color Features and Shadow Information. In 23rd International Symposium, *Computer and Information Sciences*, Istanbul, Turkey (Oct. 27-29), pp. 1-5.
227. Skarbek, W. and Koschan, A. (1994). Colour Image Segmentation - A Survey. *Technical Report University Berlin*, pp. 94-32.

228. Sobrino, J. A.; Raissouni, N.; Kerr, Y.; Oliosio, A; López-García, M.J.; Belaid, A.; (2000). *Teledetección*. Universidad de Valencia, Valencia (España), ISBN 84-370-4220-8.
229. Speer, N; Spieth, C., and Zell, A. (2005). Biological Cluster Validity Indices Based on the Gene Ontology. *Advances in Intelligent Data Analysis*, 6th International conference on Intelligent Data Analysis (IDA 2005), Madrid, Spain, LNCS 3646:429-439.
230. Stathis P., Kavallieratou E., Papamarkos .N. (2008). An Evaluation Technique for Binarization Algorithms. *Journal of Universal Computer Science*, 14(18): 3011-3030.
231. Subr, K. ; Majumder, A.; and Irani, S. (2006). Contrast enhancement of images using human contrast sensitivity. In: *Proceedings of the 3ed Symposium on Applied Perception in Graphics and Visualization*. 4(3):69-76.
232. Tai, S.C.; Chen, W.J.; Chen, P.J. (1998). Genetic algorithm for single bitmap absolute moment block truncation coding of color images. *Opt. Eng.* 37(9):2483-2490.
233. Tellaeche, A.; Burgos-Artizzu, X.P.; Pajares, G.; Ribeiro, A. (2008a). A vision-based method for weeds identification through the Bayesian precision theory. *Pattern Recognition*, 41:521-530.
234. Tellaeche, A; Burguos-Artizzu, X.; Pajares, G; Ribeiro, A; Fernandez-Quintanilla, C. (2008b). A new vision-based approach to differential spraying in precision agriculture. *Computers and Electronics in Agriculture*, 60:144-155.
235. Tellaeche, A., Pajares, G., Burgos-Artizzu, X.P., Ribeiro, A. (2011). A computer vision approach for weeds identification through Support Vector Machines. *Applied Soft Computing*, 11:908-915.
236. Theodoridis, S. and Koutroumbas, K. (2009). *Pattern Recognition*. Fourth ed. Academic Press, ELSEVIER, California, USA.
237. Thorp, K.R and Tian L.F (2004). Review of remote sensing of weeds in agriculture. *Prec. Agric.* 5:477-508.
238. Tian, S. (1998). Environmentally adaptive segmentation algorithm for out-door image segmentation. *Comput. Electron. Agric.* 21:153-168.

239. Tkalcic, J.T.M. (2003). Colour spaces-perceptual, historical and applicational background. *IEEE region 8 EUROCOM*, pp. 304-308.
240. Tortora, R. D. (1978). A Note on Sample Size estimation for Multinomial Populations. *The American Statistician*, 32(3):100-102.
241. Tortora, R. D., (1978). A Note on Sample Size estimation for Multinomial Populations. *The American Statistician*, 32(3):100-102.
242. Trias-Sanz, R.; Stamon, G.;Louchet, J. (2008). Using colour, texture, and hierarchical segmentation for high-resolution remote sensing. *ISPRS Journal of Photogrammetry & Remote Sensing*, 63:156-168.
243. Tsai, C. M. and Lee, H. H. (2002). Binarization of color document images via luminance and saturation color features. *IEEE Trans. Image Process.* IP-11: 434-451.
244. Tsai, R.Y. (1987). Metrology using off-the-shelf TV cameras and lenses. *IEEE J. Robot. Autom.* 3:323-344.
245. Tsai, W.H. (1985). Moment-preserving thresholding: A new approach. *Comput. Vis. Graph. Image Process.* 29:377-393.
246. Tseng, M. H., S. J. Chen, G. H. Hwang, and M. Y. Shen. (2008). A genetic algorithm rule-based approach for land-cover classification. *J. Photogrammetry and Remote Sensing*, 63(2):202-212.
247. Tuo, Hongya and Zhang, L; and Liu, Y. (2004), Multisensor Aerial Image Registration Using Direct Histogram Specification. *Networking, Sensing and Control.* 2:807-812
248. UCI (2012). UCI Machine Learning Repository. <http://archive.ics.uci.edu/ml/>, (accessed March 19, 2012).
249. UNAM (2012). Universidad Nacional Autónoma de México. <http://www.igeograf.unam.mx/web/sigg/inicio/index.html>, (accessed March 19, 2012).
250. Unsalan, C.; Boyer, K.L. (2005). A system to detect houses and residential street networks in multispectral satellite images. *Comput. Vis. Image. Und.* 98:423-461.
251. Valdovinos, R.M.; Sánchez, J.S. (2007). Performance analysis of classifier ensembles: Neural networks versus nearest neighbor rule. In: Proceedings of the

- 3rd Iberian Conference, *Pattern Recognition and Image Analysis (IbPRIA)*, Girona, Spain, pp. 105-112.
252. Vazquez, E.; Baldrich, R.; Vazquez, J.; and Vanrell, M. (2007). Topological Histogram Reduction towards colour segmentation. (J. Martí, J.M. Benedí. A.M. Mendonça and J. Serrat (Eds.)). *Pattern Recognition and Image Analysis*. Lecture Notes in Computer Science, Springer-Verlag, Berlin, 4477, Part I, 55-62.
253. Villegas R, I. (2006). *Degradación del suelo por impacto del huracan Paulina, en la cuenca del río la sabana, Guerrero*. (Tesis doctoral, Colegio de Postgraduados), pp. 71-96.
254. Villegas-Romero, Isidro; Oropeza-Mota, J. Luis; Martínez-Ménes, Mario; Mejía-Sáenz, Enrique. (2009). Path and relation rain-runoff caused by hurricane Pauline in the Sabana river, Guerrero, México. *Agrociencia*, 43(4): 345-356.
255. Wang, K.; Wang, B.; and Peng L. (2009). CVAP: Validation for cluster analyses. *Data Science Journal*. 8(20):88-93
256. Wangenheim, A., Bertoldi, R.F., Abdala, D.D. and Richter, M.M. (2007). Color image segmentation guided by a color gradient network. *Pattern Recognition Letters*, 28:1795-1803.
257. Wangenheim, A.; Bertoldi, R.F.; Abdala, D.D.; Richter, M.M.; Priese, L.; and Schmitt, F. (2008). Fast Two-step Segmentation of Natural Color Scenes Using Hierarchical Region-growing and a Color-gradient Network. *Journal of the Brazilian*, 14(4): 29-40.
258. Woebbecke, D.M., Meyer, G.E., von Bargen, K., Mortensen, D.A., (1995). Shape features for identifying young weeds using image analysis. *Transactions on American Society of Agricultural Engineering*, 38(1):271-281.
259. Webb, M.S.; Uemura, M.; Steponkus, P.A. (1994). Comparison of freezing injury in oat and rye: Two cereals at the extremes of freezing tolerance. *Plant Physiol*. 104:467-478.
260. Wilkinson, G. G. (2005). Results and Implications of a Study of Fifteen Years of Satellite Image Classification Experiments. *IEEE Trans. Geoscience Remote Sensing*, 45(3):433-440.

261. Woebbecke, D.M.; Meyer, G.E.; Bargen, K.V.; Mortensen, D.A. (1995). Shape features for identifying young weeds using image analysis. *Trans. ASAE*, 38:271-281.
262. Wu, Kuo-Lung; Yang, Miin-Shen y Hsieh, June-Nan. (2009). Robust cluster validity indexes. *Pattern Recognition*, 42(11):2541-2550.
263. Wyszecky G.; Stiles, W.S. (1982). *Color Science. Concepts and Methods, Quantitative Data and Formulae*. 2ª Ed., Wiley, Nueva York (EE.UU.).
264. Xie, X. (2008). A Review of Recent Advances in Surface Defect Detection using Texture Analysis Techniques. *Electronic Letters on Computer Vision and Image Analysis*, Special Issue on Quality Control by Artificial Vision, 7(3):1-25.
265. Xie, X.; and Mirmehdi, M. (2007). Colour Image Segmentation using Texems. *Annals of the British Machine Vision Association*, (6):1-10.
266. Xu, X., Ester, M. and Kriegel, H-P. and Sander, J. (1998). A distribution-based clustering algorithm for mining large spatial datasets. In: *Proceedings of the 14th ICDE*, Orlando, FL., pp. 324-331.
267. Xu, Y., Olman, V. and Xu, Do. (2002). Clustering gene expression data using a graph-theoretic approach: an application of minimum spanning trees. *Bioinformatics*, 18(4):536-545.
268. Yan, Guangjian; Li, Chaoyang; Zhou, Guoqing; Zhang, Wuming; Li, Xiaowen. (2007). Automatic Extraction of Power Lines From Aerial Images. *IEEE Geoscience and Remote Sensing Letters*, 4(3):387-391.
269. Yang, A.Y., Wright, J., Ma, Y., Sastry, S.S. (2008). Unsupervised segmentation of natural images via lossy data compression. *Computer Vision and Image Understanding*, 110:212-225.
270. Yang, H.; Meer, F. Van Der; Bakker, W. and Tan, Z. J. (1999). A back-propagation neural network for mineralogical mapping from AVIRIS data. *International Journal of Remote Sensing*, 20(1):97-110.
271. Yatskiv, I. y Gusarova, L. (2005). Methods of Cluster Analysis Results Validation. Proceedings of International Conference RelStat' 2004. *Transport and Communication*. 6(1):775-80.
272. Yen, J. C., Chang, F. J.; and Chang, S. (1995). A new criterion for automatic multilevel thresholding. *IEEE Trans. Image Process*. IP-4:370-378.

-
273. Yu, H., Li, M., Zhang, H.J., Feng, J. (2002). Color texture moments for content-based image retrieval. *International Conference on Image Processing*, 3:24-28.
 274. Zhai, C.; and Lafferty, J. (2004). A study of smoothing methods for language models applied to information retrieval. *ACM Transactions on Information Systems (TOIS)*, 22(2):179-214.
 275. Zhang, Jianqing; Ke, Tao; and Sun, Mingwei. (2008). The Parallel Computing Based on Cluster Computer in the Processing of Mass Aerial Digital Images. *Information Processing*, pp. 389-393.
 276. Zheng, L.; Zhang, J.; Wang, Q. (2009). Mean-shift-based color segmentation of images containing green vegetation. *Comput. Electron. Agric.* 65:93-98.
 277. Zhu, Z.L.; Yang, L.M.; Stehman, S.V.; Czaplewski, R.L. (2000). Accuracy assessment for the U.S. geological survey regional land-cover mapping program: New York and New Jersey region. *Photogramm. Eng. Remote Sens.* 9(1):163-167.

ANEXOS



ANEXO 1. VARIANZAS DENTRO DE CADA CLASE Y ENTRE CADA PAR DE CLASES, DESPUÉS DE APLICAR UMBRALIZACIÓN MÚLTIPLE.

Tabla A3.9. Varianzas dentro de cada clase (diagonal principal) y entre cada par de clases (fila-columna), después de aplicar umbralización múltiple.

C _i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	43,41	956,28	1335,33	1360,19	1579,64	1867,55	2324,58	2386,28	2570,00	1607,65	1731,65	2156,37	2078,37	2190,63
2	956,28	703,10	42,71	236,94	220,28	240,36	727,13	699,88	709,78	300,27	276,77	341,62	438,90	438,19
3	1335,33	42,71	397,98	347,37	206,03	214,31	729,10	670,33	676,74	337,53	216,06	257,64	432,44	376,10
4	1360,19	236,94	347,37	271,42	109,61	301,11	153,82	180,33	335,46	199,99	215,32	462,87	170,87	205,83
5	1579,64	220,28	206,03	109,61	952,85	47,38	164,61	136,26	147,58	203,49	167,34	220,04	151,45	136,00
6	1867,55	240,36	214,31	301,11	47,38	461,16	276,62	135,68	130,45	321,42	169,18	190,35	259,46	118,60
7	2324,58	727,13	729,10	153,82	164,61	276,62	895,92	74,32	240,51	594,00	569,13	684,19	131,85	151,57
8	2386,28	699,88	670,33	180,33	136,26	135,68	74,32	2942,03	47,44	578,97	542,74	582,15	142,34	122,86
9	2570,00	709,78	676,74	335,46	147,58	130,45	240,51	47,44	1268,46	637,24	550,83	580,48	266,11	129,60
10	1607,65	300,27	337,53	199,99	203,49	321,42	594,00	578,97	637,24	510,11	90,17	367,03	203,37	224,43
11	1731,65	276,77	216,06	215,32	167,34	169,18	569,13	542,74	550,83	90,17	1439,13	93,39	199,95	188,47
12	2156,37	341,62	257,64	462,87	220,04	190,35	684,19	582,15	580,48	367,03	93,39	576,60	388,46	226,11
13	2078,37	438,90	432,44	170,87	151,45	259,46	131,85	142,34	266,11	203,37	199,95	388,46	396,79	72,42
14	2190,63	438,19	376,10	205,83	136,00	118,60	151,57	122,86	129,60	224,43	188,47	226,11	72,42	1151,98
15	2503,40	466,04	390,99	443,27	171,91	117,06	357,20	153,71	120,18	423,46	218,77	216,26	328,16	92,26
16	2849,84	898,15	858,90	281,92	253,25	290,04	103,93	109,37	207,48	681,84	651,76	733,33	140,48	145,79
17	2964,68	922,13	865,01	318,50	264,71	241,85	137,22	110,46	113,10	704,24	666,39	704,75	165,04	146,08
18	3235,84	952,94	886,06	517,49	301,91	257,15	359,28	163,02	126,61	793,47	679,57	705,62	360,47	175,32
19	2829,96	944,05	859,65	660,99	594,99	602,46	685,81	672,05	717,01	188,02	219,96	416,22	217,01	237,78
20	2969,65	969,12	845,05	691,95	606,76	551,68	672,79	648,57	642,47	220,86	210,08	193,46	222,47	208,81
21	3488,08	1122,42	977,96	959,37	735,73	657,03	869,58	744,03	717,82	551,17	319,01	231,75	512,03	303,07

C_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
22	3207,22	1063,91	957,03	718,12	641,16	617,06	533,31	525,67	550,91	319,54	319,46	427,76	188,49	195,13
23	3343,44	1107,14	977,31	764,61	674,45	613,75	570,52	554,04	539,41	349,70	332,47	326,21	220,91	204,79
24	3815,89	1247,31	1098,12	1022,91	798,62	715,50	803,92	670,40	633,34	627,94	419,82	358,28	525,09	311,13
25	4172,25	1595,06	1469,85	974,86	885,40	830,80	714,67	703,77	703,20	885,99	859,61	903,80	368,01	364,75
26	4354,53	1683,51	1547,91	1042,50	944,30	872,95	766,13	750,38	731,75	957,30	925,24	940,45	417,12	405,46
27	8727,24	4361,24	4079,74	3548,76	3109,36	2893,65	2935,95	2651,12	2505,39	3312,09	3008,55	2867,59	2508,88	2182,72

Continuación de tabla A3.9 (columnas 15 a 27)

C_i	15	16	17	18	19	20	21	22	23	24	25	26	27
1	2503,40	2849,84	2964,68	3235,84	2829,96	2969,65	3488,08	3207,22	3343,44	3815,89	4172,25	4354,53	8727,24
2	466,04	898,15	922,13	952,94	944,05	969,12	1122,42	1063,91	1107,14	1247,31	1595,06	1683,51	4361,24
3	390,99	858,90	865,01	886,06	859,65	845,05	977,96	957,03	977,31	1098,12	1469,85	1547,91	4079,74
4	443,27	281,92	318,50	517,49	660,99	691,95	959,37	718,12	764,61	1022,91	974,86	1042,50	3548,76
5	171,91	253,25	264,71	301,91	594,99	606,76	735,73	641,16	674,45	798,62	885,40	944,30	3109,36
6	117,06	290,04	241,85	257,15	602,46	551,68	657,03	617,06	613,75	715,50	830,80	872,95	2893,65
7	357,20	103,93	137,22	359,28	685,81	672,79	869,58	533,31	570,52	803,92	714,67	766,13	2935,95
8	153,71	109,37	110,46	163,02	672,05	648,57	744,03	525,67	554,04	670,40	703,77	750,38	2651,12
9	120,18	207,48	113,10	126,61	717,01	642,47	717,82	550,91	539,41	633,34	703,20	731,75	2505,39
10	423,46	681,84	704,24	793,47	188,02	220,86	551,17	319,54	349,70	627,94	885,99	957,30	3312,09
11	218,77	651,76	666,39	679,57	219,96	210,08	319,01	319,46	332,47	419,82	859,61	925,24	3008,55
12	216,26	733,33	704,75	705,62	416,22	193,46	231,75	427,76	326,21	358,28	903,80	940,45	2867,59
13	328,16	140,48	165,04	360,47	217,01	222,47	512,03	188,49	220,91	525,09	368,01	417,12	2508,88
14	92,26	145,79	146,08	175,32	237,78	208,81	303,07	195,13	204,79	311,13	364,75	405,46	2182,72
15	425,01	297,14	165,23	140,66	439,58	226,31	255,13	352,03	214,57	253,90	423,25	398,73	1979,28

C_i	15	16	17	18	19	20	21	22	23	24	25	26	27
16	297,14	479,92	45,82	266,20	597,04	553,15	677,83	202,72	221,18	462,44	273,61	306,94	2133,10
17	165,23	45,82	1029,72	91,13	617,52	562,37	598,03	206,34	208,74	301,27	258,71	285,09	1823,33
18	140,66	266,20	91,13	381,77	720,90	570,71	575,07	353,76	210,20	229,35	328,47	265,14	1544,54
19	439,58	597,04	617,52	720,90	1902,30	99,19	502,14	172,52	202,38	543,84	704,84	766,11	2817,26
20	226,31	553,15	562,37	570,71	99,19	5549,04	155,03	158,45	148,98	221,08	650,50	702,23	2373,74
21	255,13	677,83	598,03	575,07	502,14	155,03	1871,42	438,06	214,15	150,44	732,58	718,24	2166,71
22	352,03	202,72	206,34	353,76	172,52	158,45	438,06	800,16	62,70	423,80	180,29	218,05	1981,24
23	214,57	221,18	208,74	210,20	202,38	148,98	214,15	62,70	2072,63	160,49	180,15	204,60	1586,57
24	253,90	462,44	301,27	229,35	543,84	221,08	150,44	423,80	160,49	609,42	390,56	258,28	1266,33
25	423,25	273,61	258,71	328,47	704,84	650,50	732,58	180,29	180,15	390,56	602,21	36,23	1495,43
26	398,73	306,94	285,09	265,14	766,11	702,23	718,24	218,05	204,60	258,28	36,23	1161,28	1112,20
27	1979,28	2133,10	1823,33	1544,54	2817,26	2373,74	2166,71	1981,24	1586,57	1266,33	1495,43	1112,20	184,63

ANEXO 2. RESULTADOS DE LOS NÚMEROS ÓPTIMOS DE AGRUPAMIENTOS

Tabla 3.15. Resultados de los números óptimos de agrupamientos según los cinco índices de validación internos (DB, CH, Dunn, KL, Han) para las muestras 1 a 9 y las 27 clases.

	Valor del índice para el número de clases k				
Índice / k	2	3	4	5	6
<i>Muestra 1</i>					
DB *	0,796	0,594	0,593	0,660	0,690
CH**	1 938 800,0	3 086 200,0	3 311 200,0	3 356 300,0	3 341 400,0
Dunn**	2,225	2,236	1,75	1,79	1,45
KL **	10 752 000 000,0	6 357 400 000,0	5 177 300 000,0	4 606 300 000,0	4 266 200 000,0
Han *	6 773 300 000,0	3 056 300 000,0	2 054 600 000,0	1 575 300 000,0	1 292 000 000,0
<i>Muestra 2</i>					
DB *	0,729	0,458	0,504	0,621	0,674
CH **	2 077 300,0	3 103 100,0	3 804 700,0	4 141 500,0	4 216 500,0
Dunn **	2,581	2,581	2,013	1,729	1,496
KL **	7 121 400 000,0	4 359 100 000,0	3 158 100 000,0	2 622 000 000,0	2 369 800 000,0
Han *	4 486 200 000,0	2 095 700 000,0	1 253 300 000,0	896 720 000,0	717 710 000,0
<i>Muestra 3</i>					
DB *	0,634	0,518	0,636	0,511	0,634
CH **	2 315 600,0	3 056 700,0	3 323 700,0	3 569 400,0	3 770 900,0
Dunn **	2,476	1,403	1,227	1,135	1,068
KL **	5 769 200 000,0	3 809 600 000,0	3 068 600 000,0	2 591 500 000,0	2 269 800 000,0
Han *	3 634 400 000,0	1 831 400 000,0	1 217 800 000,0	886 290 000,0	687 410 000,0

<i>Muestra 4</i>					
DB *	0,658	0,637	0,657	0,666	0,620
CH **	2 291 600,0	2 309 700,0	2 435 000,0	2 737 700,0	2 730 600,0
Dunn **	2,831	1,968	1,240	1,145	0,953
KL **	5 675 500 000,0	4 627 400 000,0	3 904 100 000,0	3 208 300 000,0	2 978 600 000,0
Han *	3 575 400 000,0	2 224 600 000,0	1 549 300 000,0	1 097 200 000,0	902 080 000,0
<i>Muestra 5</i>					
DB *	0,564	0,607	0,625	0,478	0,579
CH **	3 393 700,0	3 850 500,0	4 290 400,0	4 490 700,0	4 609 800,0
Dunn **	3,011	2,001	1,720	1,647	1,429
KL **	5 827 100 000,0	4 077 500 000,0	3 167 800 000,0	2 716 300 000,0	2 448 200 000,0
Han *	3 670 800 000,0	1 960 300 000,0	1 257 100 000,0	928 950 000,0	741 450 000,0
<i>Muestra 6</i>					
DB *	0,700	0,548	0,581	0,555	0,564
CH **	2 231 700,0	3 485 900,0	4 075 200,0	4 541 500,0	4 776 900,0
Dunn **	2,349	2,064	1,920	1,783	1,696
KL **	6 653 700 000,0	3 864 700 000,0	2 895 200 000,0	2 347 000 000,0	2 054 200 000,0
Han *	4 191 600 000,0	1 858 000 000,0	1 149 000 000,0	802 650 000,0	622 130 000,0
<i>Muestra 7</i>					
DB *	0,641	0,475	0,474	0,518	0,634
CH **	2 641 000,0	3 860 900,0	4 349 500,0	4 457 700,0	4 631 900,0
Dunn **	2,858	2,613	2,067	1,785	1,972
KL **	7 662 200 000,0	4 533 800 000,0	3 487 200 000,0	3 047 600 000,0	2 699 100 000,0
Han *	4 826 900 000,0	2 179 600 000,0	1 383 900 000,0	1 042 300 000,0	817 430 000,0

<i>Muestra 8</i>	Valor del índice para el número de clases k										
	7	8	9	10	11	12	13	14	15	16	18
DB *	0,728	0,733	0,641	0,534	0,611	0,736	0,675	0,642	0,667	0,643	0,627
CH**	111 850	103 070	111 510	102 890	119 380	108 280	124 530	116 680	115 220	117 000	101 520
Dunn**	1,226	1,200	1,287	1,189	1,337	0,766	1,175	1,099	0,798	0,454	0,445
KL**	405 940	415 260	369 060	382 330	320 690	340 580	289 350	299 660	295 760	284 680	312 930
Han*	110 930	103 820	85 297	82 371	64 837	64 977	52 335	51 587	48 627	44 834	45 561

<i>Muestra 9</i>	Valor del índice para el número de clases k										
	3	4	5	6	7	8	9	12	11	12	13
DB *	0,466	0,549	0,603	0,661	0,741	0,707	0,715	0,806	0,841	0,885	0,885
CH**	81 009	92 707	97 091	98 928	97 496	95 589	92 151	88 734	84 874	81 663	78 199
Dunn**	2,591	2,035	1,926	1,704	1,627	1,449	1,307	1,226	1,111	1,114	1,0394
KL**	145 160	111 380	95 871	86 683	82 147	79 118	78 070	77 608	78 026	78 313	79 214
Han*	69 788	44 201	32 787	26 252	22 449	19 779	18 044	16 720	15 775	14 941	14 328

*El óptimo número de agrupamientos es el valor que *minimiza* la función (es decir, valor mínimo de la fila).

** El óptimo número de agrupamientos es el valor que *maximiza* la función (es decir, valor máximo de la fila).

ANEXO 3. TEORÍA DEL COLOR

INTRODUCCIÓN A LA TEORÍA DEL COLOR

El color representa una sensación creada en respuesta a la excitación del sistema visual humano por radiaciones electromagnéticas conocidas como luz.

La caracterización de la luz es la materia central de la ciencia del color. Si la luz es acromática o sin color su único atributo es la intensidad. Así, el término nivel de gris indica la medida escalar de intensidad que toma valores desde negro a blanco, pasando por toda la gama de grises.

La luz cromática posee un espectro de energía electromagnética entre 400 y 700 nm, como puede observarse en el diagrama de cromaticidad de la figura A3.1, las tres características básicas que se utilizan para describir una fuente cromática de luz son la radiación, luminancia y brillo.

En el tratamiento del color existen unos parámetros básicos cuyas definiciones se presentan a continuación (Platanitos, 2000).

a) Intensidad (I). Es una medida, sobre una parte del espectro electromagnético, del flujo de potencia que es radiado desde una superficie, o incidente en ella, y se expresa en unidades de vatios por metro cuadrado.

b) Brillo (Br). Es un atributo relacionado con la sensación visual relativo al hecho de que un área aparente mayor o menor emisión de luz.

c) Luminancia (Y). Es la potencia radiada ponderada por una función de sensibilidad espectral característica de la visión humana. Fue definida por la CIE (siglas francesas de la Comisión Internacional de Iluminación, Commission Internationale de L'Eclairage).

d) Luminosidad (L^*). Es la respuesta perceptiva no lineal a la luminancia que posee la visión humana. La no linealidad es prácticamente logarítmica.

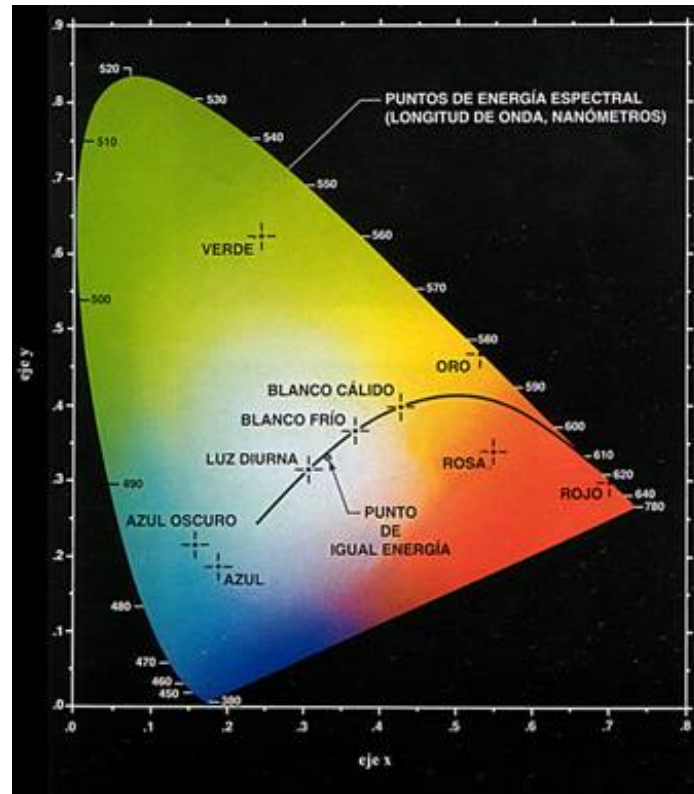


Figura A3.1. Diagrama de cromaticidad

e) Tono (H). Es una propiedad del color asociada con la longitud de onda dominante en una mezcla de ondas de luz, por tanto, representa el color dominante tal y como es percibido por un observador. Así, cuando se dice que un objeto es rojo, verde, amarillo o de otro color, es en realidad su tono lo que se especifica.

f) Saturación (S). Se refiere a la pureza relativa o a la cantidad de luz blanca mezclada con un tono. Los colores puros del espectro están completamente saturados y no contienen luz blanca. El grado de saturación disminuye a medida que se añade más luz blanca.

El *tono* y la *saturación* describen la *chrominancia*. Los humanos interpretan el color basándose en su *luminosidad* (L^*), *tono* (H) y *saturación* (S) (Gonzalez, 1996).

Para poder realizar un tratamiento de imágenes en color es necesario establecer un modelo de representación de dicho color. Dependiendo de la aplicación concreta que se vaya a implementar, se elegirá una representación u otra. El proceso de elección depende de cómo se generen las señales de color y de qué información se necesite extraer de ellas.

Se pueden distinguir cuatro modelos básicos de color (Plataniotis, 2000):

a) *Modelos colorimétricos*. Basados en medidas físicas de la reflectancia espectral. Suelen trabajar con tres filtros de colores primarios y un fotómetro, dando lugar a un *diagrama de cromaticidad CIE*.

b) *Modelos de color psicofísicos*. Fundamentados en la percepción humana del color. Estos modelos se basan en criterios subjetivos de observación y referencias comparativas (por ejemplo, el modelo Munsell de color), o bien se construyen a partir de experimentos que cumplan con la percepción humana del color (por ejemplo, modelos de Tono, Saturación y Luminosidad).

c) *Modelos de color fisiológicos*. Basados en los tres colores primarios (los tres tipos de conos de la retina humana). El espacio de color rojo-verde-azul RGB (*Red-Green-Blue*) utilizado en el *hardware* de las computadoras es el más conocido.

d) *Modelos de colores contrapuestos*. Hacen uso de parejas de colores primarios opuestos entre sí, como el amarillo-azul o el rojo-verde.

En 1931 la CIE planteó unas curvas tipo para un hipotético *Observador Estándar*. Estas curvas especifican cómo puede ser transformada una determinada distribución espectral de potencia de un estímulo externo (luz radiante visible incidente en el ojo) en un conjunto de tres números que especifican un color.

La intensidad, definida como el flujo de potencia que se radia desde un objeto, es una medida lineal de la luz. El valor lineal R es proporcional a la intensidad del flujo de potencia radiado desde un objeto alrededor de la banda de los 700 nm del espectro visible. El valor lineal G corresponde a la banda de los 546.1 nm y el valor B a la de los 435.8 nm; éstos constituyen el sistema de referencia R_{CIE} G_{CIE} B_{CIE} , y que generalmente se denota con las siglas RGB simplemente (Plataniotis, 2000).

El sistema CIE RGB primario no puede representar todos los posibles colores reproducibles. De ahí que CIE propusiera el sistema artificial primario XYZ, que da lugar a curvas espectrales siempre positivas, y por tanto, aplicable a un gran número de situaciones prácticas. Las coordenadas XYZ están linealmente relacionadas con las RGB mediante la matriz. Dada por (a3-1).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.490 & 0.310 & 0.200 \\ 0.177 & 0.813 & 0.011 \\ 0.000 & 0.010 & 0.990 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (\text{a3-1})$$

Las coordenadas de crominancia vienen dadas por,

$$x = \frac{X}{X+Y+Z}; \quad y = \frac{Y}{X+Y+Z} \quad (\text{a3-2})$$

El blanco de referencia se representa por $X = Y = Z = 1$. La componente Y es equivalente a la luminancia. En este trabajo de investigación se utiliza precisamente esta matriz de conversión.

El sistema CIE se basa en la descripción del color a través de la componente de luminancia Y junto con las dos componentes adicionales X y Z (Wyszecki, 1982). Los valores tri-estímulos XYZ pueden usarse para describir cualquier color, dando lugar al espacio de color CIE XYZ. Su uso juega un papel muy importante en la teoría del color en imágenes ya que otros espacios de color se deducen a partir del XYZ mediante transformaciones matemáticas. Por ejemplo, el espacio lineal RGB puede transformarse a partir de éste por una simple transformación matricial. De igual forma, pueden obtenerse otros espacios de color a partir del XYZ como son el YIQ (Luminance, In-phase, Quadrature-phase chrominances), usado en el estándar de televisión NTSC (National Television Standard Committee), espacio de tono, saturación e intensidad HSI (Hue, Saturation, Intensity), etc. La CIE también estandarizó otros dos espacios de color llamados $L^*u^*v^*$ y $L^*a^*b^*$, que son uniformes de acuerdo con la percepción humana, por ello tienen gran aplicación en la práctica [Wyszecki, 1982]. La componente L^* representa la luminosidad, mientras que las otras dos (u^*v^* o a^*b^*) las componentes de crominancia.

ESPACIO DE COLOR RGB

El sistema de representación RGB lineal es independiente del dispositivo utilizado para adquirir la imagen, y se suele emplear para asegurar la consistencia en el color a través de distintos dispositivos (Plataniotis, 2000).

El espacio de color RGB es el espacio de color más extendido y el que utilizan la gran mayoría de las cámaras de video y fotografías para construir una imagen en color, este modelo está basado en el sistema de coordenadas cartesianas y el sub-espacio de color de interés es el tetraedro mostrado en la figura A3.2. Según esta representación los valores RGB se sitúan en tres vértices y el cian, magenta y amarillo se ubican en los otros tres vértices opuestos. El negro corresponde al origen y el blanco al vértice más alejado del origen. En este modelo, la escala de grises se extiende desde el negro al blanco a lo largo de la diagonal principal que une esos dos puntos, y los colores se representan como puntos dentro del tetraedro definidos por vectores desde el origen, por conveniencia se asume que los valores han sido normalizados de modo que el tetraedro de la figura A3.2 es unario, es decir todos los valores de R, G y B se sitúan en el rango [0,1]. Las imágenes en este modelo se forman por la combinación en diferentes proporciones de cada uno de los colores primarios RGB (Pajares y Cruz, 2007a, b).

Una de las ventajas de trabajar sobre el mismo espacio de color con que fue capturada la imagen es que esto permite evitar la alteración de las propiedades del color durante el proceso de segmentación, propia de los errores de conversión y transformación, y por otro lado conseguir una mayor velocidad de segmentación al ahorrar tiempo de procesamiento en conversión y redondeo.

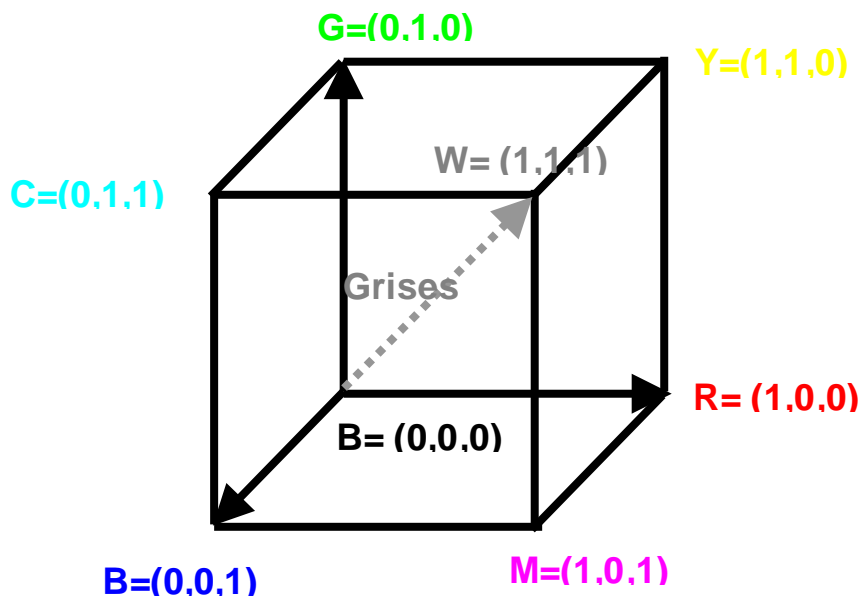


Figura A3.2 Tetraedro de color RGB. Los puntos a lo largo de la diagonal principal tienen valores de gris, desde el negro en el origen al punto (1 1 1).

ESPACIOS DE COLORES UNIFORMES EN CUANTO A PERCEPCIÓN

La sensibilidad visual a pequeñas diferencias entre colores tiene una gran importancia en la percepción del color (Plataniotis, 2000). Un sistema que se use para especificar un color tiene que ser capaz de representar cualquier color con una alta precisión. Todos los sistemas actualmente disponibles para ello se basan en el modelo de color CIE XYZ.

Se busca un sistema de representación uniforme en cuanto a percepción del color, de manera que diferencias entre colores, en dicho sistema de representación, impliquen diferencias de colores tal y como las percibe el ojo humano.

La CIE tardó más de una década en encontrar una transformación de XYZ a un espacio de color uniforme. Al final no se encontró un único sistema y en 1976 la CIE estandarizó dos espacios, $L^*u^*v^*$ y $L^*a^*b^*$, como perceptivamente uniformes (Wyszecki, 1982). Ambos espacios son igual de buenos en cuanto a uniformidad de percepción y proporcionan muy buenas estimaciones de diferencias de color (distancias) entre dos vectores de color. Los dos sistemas se basan en la luminosidad percibida L^* y en un conjunto de ejes de colores contrapuestos (rojo-verde frente a amarillo-azul).

ESPACIO DE COLOR L*a*b*

Para mejorar la representación del color, la CIE desarrolló en 1976 el modelo de color L*a*b* (figura A3.3), la luminosidad percibida por un observador estándar se supone que sigue la luminancia física (L*) (una cantidad proporcional a la intensidad) de acuerdo a una ley de raíz cúbica, es decir, viene definida por (Plataniotis, 2000).

En el nuevo espacio X, Y, Z, la luminosidad es calculada por la ecuación (a3-3),

$$L^* = \begin{cases} 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16 & \text{if } \frac{Y}{Y_n} > 0.008856 \\ 903.3 \left(\frac{Y}{Y_n} \right)^{1/3} & \text{si no} \end{cases} \quad (a3-3)$$

Las componentes a* y b* son obtenidos por,

$$\begin{aligned} a^* &= 500 \left[f \left(\frac{X}{X_n} \right) - f \left(\frac{Y}{Y_n} \right) \right] \\ b^* &= 200 \left[f \left(\frac{Y}{Y_n} \right) - f \left(\frac{Z}{Z_n} \right) \right] \end{aligned} \quad (a3-4)$$

Donde $f(t) = t^{1/3}$ para $t > 0.008856452$, y $7.787 t + 16/116$ para otro valor. Normalizados X_n, Y_n, Z_n , son los valores tri-estímulo CIE XYZ del punto blanco de referencia (los tres colores primarios en un modelo de color de componente aditivo). La división de la función $f(t)$ en dos dominios se hace para prevenir una pendiente infinita en $t = 0$. Se asume que la función $f(t)$ sea lineal para $t = t_0$ y que coincide con la raíz cúbica de t para $t = t_0$, tanto en valor ($t_0^{1/3} = at_0 + b$) como en pendiente ($1/(3t_0^{2/3}) = a$). Los valores de b fueron ajustados a 16/116.

El modelo de color L*a*b* es probablemente el modelo de color más completo (figura A3.3.) Se usa habitualmente para describir todos los colores que puede ver el ojo humano (Wyszecki, 1982). Las diferencias de color que se perciben como iguales en este espacio de color tridimensional, tienen distancias muy próximas entre ellas.

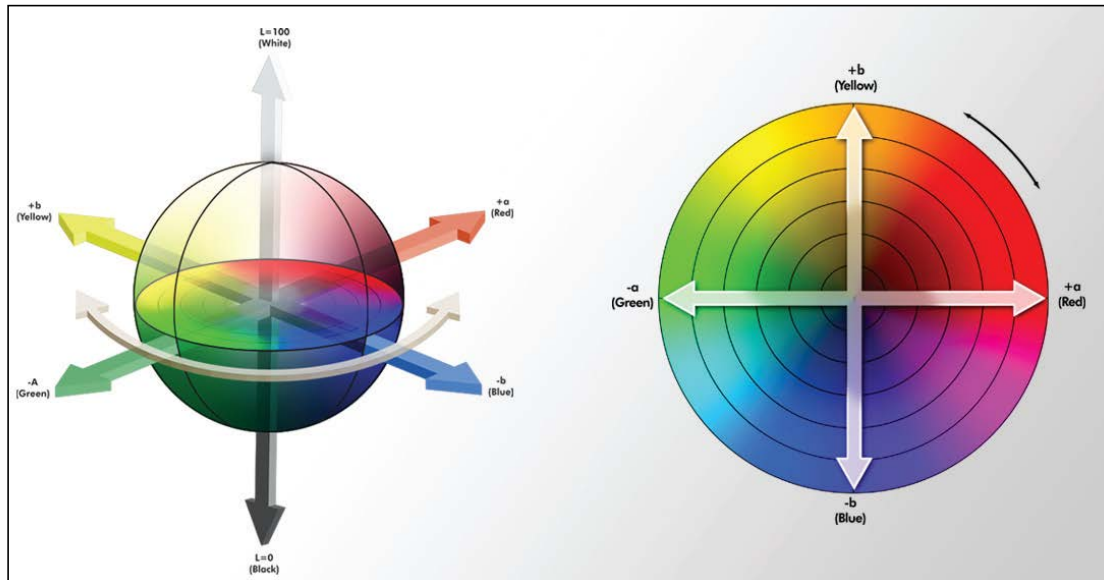


Figura A3.3. Modelo de color CIELAB

ESPACIO DE COLOR $L^*u^*v^*$

Partiendo del sistema XYZ definido en (a3-1) y de las coordenadas de cromaticidad x e y definidas en (a3-2), se hallan las coordenadas u' y v' de la siguiente forma,

$$u' = \frac{4X}{X + 15Y + 3Z} = \frac{4x}{-2x + 12y + 3} \tag{a3-5}$$

$$v' = \frac{9Y}{X + 15Y + 3Z} = \frac{9y}{-2x + 12y + 3}$$

La luminosidad se calcula como en (a3-3) y las otras dos componentes vienen dadas por

$$\begin{aligned} u^* &= 13L^*(u' - u'_0) \\ v^* &= 13L^*(v' - v'_0) \end{aligned} \tag{a3-6}$$

Las coordenadas Y_0, u'_0, v'_0 corresponden al blanco de referencia que se escoja. Para el estándar D_{65} , $Y_0 = 1, u'_0 = 0.1978$ y $v'_0 = 0.4683$. En la figura A3.4 se muestra el diagrama de cromaticidad para las coordenadas u' y v' .

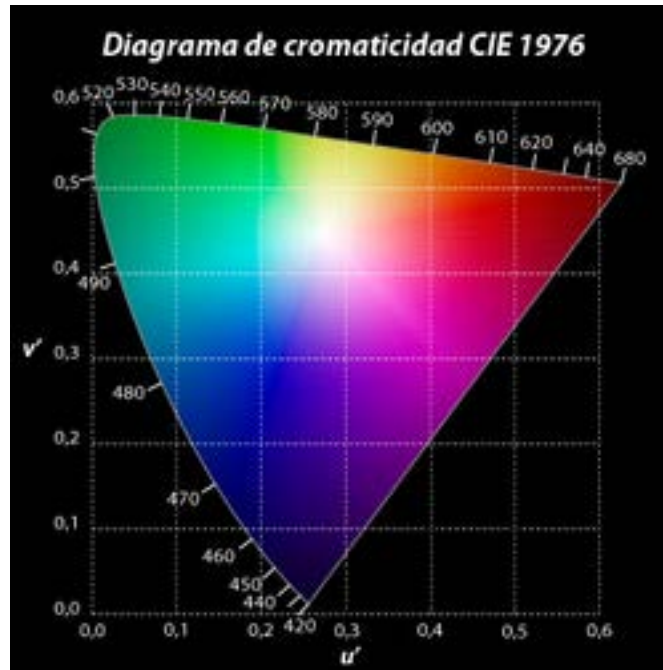


Figura A3.4. Diagrama de cromaticidad de las coordenadas u' y v' .

La fórmula para calcular la diferencia de color entre dos puntos para el plano $L^*u^*v^*$ según la distancia Euclídea, se define en Plataniotis (2000) como sigue.

$$\Delta C = \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2} \tag{a3-7}$$

La cromaticidad, el tono y la saturación vienen dados por

$$Cr_{uv}^* = \sqrt{(u^*)^2 + (v^*)^2}$$

$$H_{uv} = \arctan\left(\frac{v^*}{u^*}\right) \tag{a3-8}$$

$$S_{uv}^* = \frac{Cr_{uv}^*}{L^*}$$

Se observa que las componente a^* y b^* no dependen de la luminancia, en el sentido que pueden tomar valores distintos de cero para luminancia nula. En el sistema $L^*u^*v^*$ esto no ocurre (Schmid, 1999).

Los sistemas de representación del color llamados uniformes tienen la propiedad de que distancias Euclídeas calculadas en dichos espacios, corresponden a diferencias en color tal y como las percibe el ojo humano. Para este trabajo se ha escogido el sistema de color uniforme $L^*u^*v^*$ y $L^*a^*b^*$, que es un sistema de coordenadas cartesiano ortogonal que cumple con dicha propiedad (Wyszecky, 1982).