

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE INFORMÁTICA
Departamento de Ingeniería del Software e Inteligencia Artificial



TESIS DOCTORAL

Segmentación automática de texturas en imágenes agrícolas

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR

M. Isabel Riomoros Callejo

Directores

Gonzalo Pajares Martinsanz
María Guijarro Mata-García
Ángela Ribeiro Seijas

Madrid, 2016

SEGMENTACIÓN AUTOMÁTICA DE TEXTURAS EN IMÁGENES AGRÍCOLAS

Tesis doctoral

M. Isabel Riomoros Callejo



UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Informática

Departamento de Ingeniería del Software e Inteligencia Artificial

Universidad Complutense de Madrid
Facultad de Informática

Departamento de Ingeniería del Software e Inteligencia Artificial



**SEGMENTACIÓN AUTOMÁTICA DE
TEXTURAS EN IMÁGENES AGRÍCOLAS**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR PRESENTADA POR

M. Isabel Riomoros Callejo

Bajo la dirección de los doctores:

Gonzalo Pajares Martinsanz

María Guijarro Mata-García

Ángela Ribeiro Seijas

Madrid, 2015

Agradecimientos

Del campo y de sus gentes he sacado a lo largo de mi vida la fuerza necesaria para desarrollarme como persona y como profesional y espero, para eso me encuentro en este lugar del camino, dar fruto.

En mí se ha dado la paradoja de admirar aquello con lo que el hombre prehistórico comenzó a sustentarse, la tierra, y a la vez disfrutar con la tecnología, aquello que el hombre del siglo XXI deberá gestionar de manera eficiente y de lo que se beneficiará de manera ilimitada. Pese a que la tecnología es necesaria, no nos podemos olvidar de las personas, tan importantes en los campos, que siembran con su saber, con paciencia vigilan y ayudan en su crecimiento y recogen con satisfacción lo sembrado. Y es a las personas, a las que me voy a dirigir en estos momentos, comenzando por mis directores de tesis.

Gracias a Gonzalo Pajares, por ofrecerme la posibilidad de trabajar junto a él, me ha enseñado mucho, siempre comprometido con mi proyecto, siempre he recibido su ayuda, gracias por su paciencia y buena predisposición. Gracias a María, por su aliento, capacidad de trabajo, dedicación, confianza y por contagiarme de ilusión cuando he dudado, ayudándome a resolver problemas y, en muchas ocasiones, a poner orden en el caos. Gracias a Ángela por su dedicación, apoyo y ayuda.

Gracias a las personas más importantes de mi vida, Valentín y Ángela, ellos, que me han hecho crecer y convertirme, con su ejemplo, en una mujer preparada para “segmentar de manera automática” la vida y eliminar “las malas hierbas” que a veces han querido crecer a mi alrededor. Mi madre, cada día más debilitada me inspira, me llena de esperanza y me demuestra que vale la pena vivir y luchar por lo que quieres. Gracias Ángel Luis, querido hermano, por tantos días compartiendo nuestras historias mientras madurábamos. Gracias Yolanda por tu ayuda y disponibilidad.

Gracias a Juan Antonio, Isabel y Clara, la mejor cosecha que he tenido, han colmado con creces mis expectativas. Siempre me he sentido tan querida y habéis sido tan generosos en las épocas de sequía en mi investigación, épocas en las que me he sentido abatida, en esos momentos siempre he tenido vuestra sonrisa y vuestra ternura.

Gracias a esos amigos y compañeros incondicionales, siempre dispuestos a echar una mano, siempre os he sentido cerca cuando os he necesitado, sin vosotros tampoco hubiera sido posible llegar hasta aquí.

Por último, gracias a los proyectos: Desarrollo de Métodos de Identificación de Texturas Naturales y su Localización Espacial mediante métodos de Visión Artificial (AGL2008-04670-C03-02); Robot Fleets for Highly Effective Agriculture and Forestry Management" Theme 4 NMP-Nanosciences, Nanotechnologies, Materials and New Production Technologies (CP-IP 245986-2 (RHEA) FP7-NMP-2009-LARGE-3; Integración de Información Multisensorial y Aprendizaje Automático para la Detección, Caracterización y Reconocimiento Preciso de Estructuras Naturales en Campos de Cultivo (AGL2014-52465-C4-3-R), que han hecho posible esta investigación.

A mis padres Valentín y Ángela.

Resumen

El control de malas hierbas en grandes extensiones de terreno resulta costoso y a veces contaminante desde el punto de vista medioambiental. El avance en los últimos años de los sistemas de control y automatización en agricultura ha favorecido la aparición del concepto de Agricultura de Precisión (AP), para intervenir o actuar sobre el cultivo, en el momento adecuado y en el lugar preciso.

La presente tesis se ha desarrollado en el marco de dos proyectos de investigación a nivel nacional y uno europeo. Este último dentro del VII programa Marco de la Unión Europea, cuyo principal objetivo era implementar técnicas de AP de forma más eficiente mediante flotas de robots. En todos ellos, que han involucrado robots, éstos están equipados con sus correspondientes sistemas de Visión Artificial con el fin de identificar la existencia de determinadas texturas de interés con fines de tratamientos específicos o navegación segura de los vehículos, evitando obstáculos u otros elementos existentes en el campo de trabajo.

Para la consecución de los objetivos mencionados resulta necesario el análisis de las imágenes procedentes de los campos de cultivo. Estas imágenes están constituidas por diferentes elementos que en su conjunto contienen diferentes tipos de estructuras, tales como: a) plantas verdes alineadas en forma de surcos o dispersas según se trate de cultivo o malas hierbas; b) suelo que constituye el sustrato; c) cielo u otros componentes añadidos. Todos ellos constituyen lo que desde el punto de vista del tratamiento de imágenes se conoce como texturas.

Por tanto, y de forma más específica, la investigación se ha centrado en el análisis de las texturas existentes en las imágenes captadas por los sistemas de visión en campos de maíz, lo cual se ha realizado desde dos puntos de vista: cromático y espacial.

Desde el punto de vista cromático se consideran las propiedades espectrales de las plantas verdes (cultivo de maíz y malas hierbas), así como del suelo y las partes del cielo existentes. Las imágenes se obtienen en el modelo de color RGB (Red, Green, Blue) que constituyen los tres canales espectrales considerados para el análisis. Corresponden a la información procedente del rango de longitudes de onda pertenecientes al espectro visible. Se trata de localizar tales texturas considerando tanto imágenes panorámicas como estudiando exclusivamente las partes localizadas del suelo, que constituyen la región específica de interés dentro del campo de cultivo. En este enfoque se aplica un

procedimiento de clasificación basado en lo que se conoce como agrupamiento borroso para discernir entre tres tipos de plantas y dos de suelo. En el primer caso, se distinguen las plantas de cultivo de las malas hierbas, lo que constituye un aporte relevante a la comunidad científica. La técnica de agrupamiento borroso es por su naturaleza supervisada, si bien se ha aplicado un procedimiento que permite automatizar el proceso, evitando en cierta medida la supervisión, ya que sólo es necesario proporcionar el número de clases que se quieren identificar.

El enfoque espacial tiene en cuenta, no sólo las propiedades cromáticas, sino que añade una componente de tal naturaleza extraída de las plantas verdes para enfatizar dichas texturas mediante su realzado, aplicando la transformada de wavelets. Esta transformada es capaz de descomponer las imágenes que contienen las partes verdes en componentes de aproximación y detalle, procedentes precisamente de las plantas, lo que unido a las propiedades cromáticas permite un mayor realzado de dichas plantas. La transformada de wavelets está basada en una combinación de filtros paso bajo y paso alto por filas y columnas en las imágenes, que convenientemente aplicados permiten extraer la información detallada en tres direcciones espaciales (horizontal, vertical y diagonal). Más específicamente, son los filtros paso alto los que obtienen la información de detalle de las texturas en las plantas verdes, que por la propia naturaleza de éstas se presentan de forma aleatoria en todas las direcciones espaciales. La combinación de propiedades cromáticas con las espaciales proporciona un planteamiento de interés en este tipo de procesos para la identificación de las plantas verdes en campos de cultivo.

En ambos casos, el objetivo es la identificación de las plantas verdes para eliminar las malas hierbas mediante tratamientos específicos o navegación autónoma. Los resultados obtenidos han sido suficientemente satisfactorios justificando ambos planteamientos.

Palabras clave

Agricultura de precisión, automatización de tareas agrícolas, visión por computador, tratamiento de imágenes, identificación de texturas, análisis espectral de imágenes, transformada de wavelets, tratamientos selectivos de eliminación de malas hierbas.

Abstract

The control of weeds in large land areas results expensive and most times pollutant from the environmental point of view. In the last years, important advances through the application of control systems and automation in agriculture have been made possible the appearance of Precision Agriculture (PA), which means to carry out the right intervention, in the right place at the right time.

The research has been supported by two national research projects and also with funding from a research European project in the FP7 framework, where the main goal was the efficient application of PA techniques through a fleet of robots. The three projects have involved robots, equipped with machine vision systems with the aim of identifying the most relevant textures for site-specific treatments or safe navigation for the involved vehicles, avoiding obstacles or existing elements in the agricultural field.

To achieve the above objectives it is required the analysis of images coming from the crop fields. These images contain different elements that on the whole contain different structures, namely: a) green plants aligned in furrows or randomly distributed depending on if they are crops or weeds; b) soil as the basis; c) sky and other added components. They form what is known as textures from the point of view of image processing.

Therefore, and more specifically, the research has been focused on the texture analysis existing in the images acquired by the machine vision systems in maize fields, which has been carried out from two points of view: chromatic and spatial.

From the chromatic point of view, the spectral properties displayed by the green plants (maize crops and weeds), as well as soil and the existing sky parts. The images are captured in the RGB (Red, Green, Blue) color space that form the three spectral channels considered for analysis. They contain the information obtained in the visible wavelength range. The goal is to locate such textures considering panoramic images and also the specific parts in the soil that form the region of interest in the crop field. In this approach a classification procedure, based on fuzzy clustering, is applied, which allows the identification of three types of plants and two of soil. In the first case, crops and weeds plants are identified, making an important contribution to the scientific community. The fuzzy clustering is a supervised approach. Nevertheless an automatic method has been

applied in order to avoid the supervision, where only the number of classes to be identified is required.

The spatial approach considers not only the chromatic properties but also adds a spatial component extracted from the green plants to enhance such textures by applying the wavelets transform. This transform provides useful approximation and detail coefficients from the green plants, which together with the chromatic properties achieve a greater enhancement of these plants. The wavelet transform is based on a combination of low and high pass filters by rows and columns in the images, that conveniently applied allow extracting the required detailed information in three spatial directions (horizontal, vertical and diagonal). More specifically, the high pass filters capture the detailed information from the existing textures in the green plants that, because of the nature of these textures, appear randomly in all spatial directions. The combination of chromatic and spatial properties provides an interesting improvement for plants discrimination in maize fields.

In both approaches the goal is the identification of green plants for weeds removal with site-specific treatments or autonomous navigation. The performance achieved justifies both approaches for the intended purpose.

Keywords

Precision Agriculture, agricultural automation tasks, computer vision, image processing, texture identification, spectral imaging analysis, wavelets transform, site-specific treatments for weeds removal.

Índice

Resumen

Abstract

Índice de figuras

Índice de tablas

Capítulo 1.- Introducción

| | |
|--|----|
| 1.1. Antecedentes | 1 |
| 1.2. Identificación del problema | 6 |
| 1.3. Motivación y objetivos | 8 |
| 1.4. Metodología | 9 |
| 1.5. Aportes a la investigación | 10 |
| 1.6. Organización de la memoria | 14 |

Capítulo 2.- Estado del arte

| | |
|--|----|
| 2.1. Visión general | 18 |
| 2.2. Visión artificial | 20 |
| 2.3. Métodos de segmentación por la respuesta espectral | 24 |
| 2.3.1. Índices de vegetación (índices cromáticos) | 25 |
| 2.3.2. Umbralización | 28 |
| 2.3.3. Segmentación basada en aprendizaje | 30 |
| 2.4. Métodos de segmentación por la respuesta espacial mediante wavelets | 32 |
| 2.5. Conclusión general | 36 |

Capítulo 3.- Fundamentos teóricos

| | |
|--|----|
| 3.1. Imágenes | 41 |
| 3.1.1. Modelos de color | 41 |
| 3.2. Índices de vegetación: detección del verdor | 43 |
| 3.3. Umbralización | 47 |
| 3.3.1. Umbralización manual | 49 |
| 3.3.2. Método Otsu | 50 |
| 3.4. Clasificación: agrupamiento borroso | 52 |
| 3.5. Transformada de wavelets: combinación | 55 |
| 3.5.1. Wavelets Daubechies | 65 |
| 3.5.2. Wavelets Coiflet | 65 |
| 3.5.3. Wavelets Symlets | 65 |
| 3.5.4. Wavelets Biortogonales | 66 |
| 3.5.5. Wavelets Meyer | 66 |

Capítulo 4.- Sistemas automáticos: identificación de texturas

| | |
|--|----|
| 4.1. Método de segmentación automático basado en la combinación de índices cromáticos y agrupamiento borroso | 70 |
| 4.1.1. Segmentación automática | 74 |
| 4.1.2. Clasificación mediante agrupamiento borroso | 83 |
| 4.2. Método para la extracción de verdes mediante un índice cromático e información espacial | 86 |
| 4.2.1. Segmentación de verdes | 89 |
| 4.2.2. Descomposición en wavelets | 92 |
| 4.2.3. Descriptores de textura | 94 |
| 4.2.4. Combinación de información de verdor y textura | 97 |
| 4.2.5. Binarización | 98 |

Capítulo 5.- Análisis de resultados

| | |
|--|-----|
| 5.1. Resultados de la segmentación automática combinando índices cromáticos y agrupamiento borroso | 101 |
| 5.1.1. Imágenes reales de campos de cultivo | |
| 5.1.2. Extracción del verdor | 103 |
| 5.1.3. Identificación de los diferentes tipos de verdes y rojos | 105 |
| 5.2. Resultados de la extracción de verdes mediante un índice cromático e información espacial | 107 |
| 5.2.1. Imágenes de campos de cultivo | 109 |
| 5.2.2. Familias de wavelets | 113 |
| 5.2.3. Análisis ExG | 114 |

Capítulo 6.- Conclusiones y trabajo futuro

| | |
|---|-----|
| 6.1. Conclusiones | 119 |
| 6.1.1. Segmentación automática combinando índices cromáticos y agrupamiento borroso | 119 |
| 6.1.2. Extracción de verdes mediante un índice cromático e información espacial | 120 |
| 6.2. Trabajo futuro | 121 |

| | |
|---------------------------|-----|
| Bibliografía | 123 |
|---------------------------|-----|

Índice de figuras

| | | |
|--------------------|---|----|
| Figura 2.1. | <i>Fases de un sistema de visión artificial por computador.</i> | 21 |
| Figura 3.1. | <i>Histograma de intensidad como suma de dos funciones de densidad de probabilidad.</i> | 48 |
| Figura 3.2. | <i>Cálculo del primer coeficiente.</i> | 57 |
| Figura 3.3. | <i>Desplazamiento de la wavelet.</i> | 58 |
| Figura 3.4. | <i>Escalado (expansión) y comienzo de nuevo desde el principio.</i> | 58 |
| Figura 3.5. | <i>Localización de las bandas de frecuencia en una TDW con cuatro bandas. La convención es fila/columna.</i> | 61 |
| Figura 3.6. | <i>(a) Esquema del algoritmo de filtrado por filas y columnas para la descomposición de la imagen original en cuatro subimágenes. (b) Simbología utilizada.</i> | 61 |
| Figura 3.7. | <i>(a) Esquema del procedimiento de reconstrucción. (b) Simbología utilizada.</i> | 63 |
| Figura 3.8. | <i>Representación gráfica de la Wavelet de Haar.</i> | 64 |
| Figura 3.9. | <i>Wavelets de Daubechies de distintos órdenes.</i> | 65 |
| Figura 4.1. | <i>Esquema de los pasos seguidos para las dos aproximaciones propuestas</i> | 70 |
| Figura 4.2. | <i>Esquema de los pasos seguidos en la estrategia basada en índices cromáticos.</i> | 73 |
| Figura 4.3. | <i>Imagen original mostrando sólo el suelo.</i> | 75 |
| Figura 4.4. | <i>Imagen original mostrando el suelo y el cielo con mucha iluminación.</i> | 75 |
| Figura 4.5. | <i>Imagen original mostrando el suelo y el cielo con poca iluminación.</i> | 75 |
| Figura 4.6. | <i>Diferentes grados de verdes (a) ExG; (b) ExGR.</i> | 77 |

| | | |
|---------------------|---|-----|
| Figura 4.7. | <i>(a)-(b) Imágenes binarias obtenidas mediante la aplicación del histograma de la media y Otsu, (c) las diferencias entre (a) y (b) donde los píxeles blancos corresponden a los verdes ocluidos por Otsu.</i> | 80 |
| Figura 4.8. | <i>(a) ExB imagen obtenida con la ecuación (3.11); (b) (c) imágenes binarias de la imagen de (a) obtenidas mediante la aplicación de la media y Otsu respectivamente.</i> | 81 |
| Figura 4.9. | <i>(a) ExR imagen obtenida de la ecuación (3.10); (b) imagen binaria obtenida aplicando el criterio de la media; (c) diferencias entre las imágenes binarias dadas en la figura 4.6. (b) y 4.7. (b)</i> | 82 |
| Figura 4.10. | <i>Imágenes binarias obtenidas desde la imagen original de la figura 4.3 (a) ExB; (b) ExR; (c) diferencia entre la binarización de imágenes ExB y ExR.</i> | 84 |
| Figura 4.11. | <i>Proceso completo de identificación de la capa vegetal.</i> | 88 |
| Figura 4.12. | <i>Imagen original capturada con el sensor BASLER.</i> | 89 |
| Figura 4.13. | <i>Imagen en RGB.</i> | 90 |
| Figura 4.14. | <i>Aplicamos los índices cromáticos (ExG, ExGR, CIVE y VEG) a la imagen original 4.13.</i> | 90 |
| Figura 4.15. | <i>Índice ExG aplicado a la imagen original, imagen 4.12.</i> | 92 |
| Figura 4.16. | <i>Descomposición multiescala de una imagen mediante la transformada wavelet discreta.</i> | 93 |
| Figura 4.17. | <i>Coefficientes de aproximación y detalle AExG, HExG, VExG y DExG.</i> | 94 |
| Figura 4.18. | <i>Imagen con los verdes obtenida a partir de la nueva AExG con la TDIW.</i> | 98 |
| Figura 4.19. | <i>(a) Imagen original; (b) Imagen binarizada aplicando el umbral de Otsu; (c) Imagen resultante de la binarización con Otsu, de la imagen ExG mejorada.</i> | 99 |
| Figura 5.1. | <i>Imagen original capturada con un sensor mostrando sólo el suelo.</i> | 102 |

| | | |
|---------------------|--|-----|
| Figura 5.2.- | <i>Imagen original mostrando el suelo y el cielo con mucha iluminación.</i> | 102 |
| Figura 5.3. | <i>Imagen original mostrando el suelo y el cielo con poca iluminación.</i> | 103 |
| Figura 5.4. | <i>"Ground truth".</i> | 104 |
| Figura 5.5. | <i>Clasificación de imagen para la figura 5.2.: (a) Tres tipos diferentes de clases de plantas verdes; (b) 2 clases diferentes de suelo.</i> | 106 |
| Figura 5.6. | <i>Resultados de la clasificación de la imagen de la figura 5.1: (a) 2 clases distintas de plantas verdes; (b) 2 clases diferentes de suelo.</i> | 106 |
| Figura 5.7. | <i>Binarización de la imagen (a) mediante el umbral Otsu (a) Imagen original (b) Información de los verdes utilizando ExG; (c) ExG con verdes e información de la textura.</i> | 108 |
| Figura 5.8. | <i>Binarización de la imagen (a) mediante el umbral de Otsu (a) Imagen original (b) Información de los verdes utilizando ExG (c) ExG con verdes e información de la textura.</i> | 108 |
| Figura 5.9. | <i>El "ground-truth" de la ROI con clasificación errónea (puntos negros) solamente verdor y con verdor más textura espacial.</i> | 112 |
| Figura 5.10. | <i>Curva ROC: Clasificación por Otsu; el azul utiliza la información del verdor de ExG, y el naranja utiliza la información de verdor ExG con la información de la textura espacial.</i> | 117 |

Índice de Tablas

| | | |
|-------------------|--|-----|
| Tabla 3.1. | <i>Varios índices utilizando el espectro infrarrojo</i> | 45 |
| Tabla 3.2. | <i>Índices de vegetación basados en el espectro visible</i> | 47 |
| Tabla 5.1. | <i>Porcentaje de error medio comparándolo con el conjunto de imágenes “ground truth”</i> | 104 |
| Tabla 5.2. | <i>Porcentaje medio de error obtenido para las diferentes familias de wavelets</i> | 114 |
| Tabla 5.3. | <i>Porcentaje de error medio y desviación estándar sobre los ROI's en el conjunto de imágenes “ground truth”</i> | 115 |
| Tabla 5.4. | <i>Valores de “precisión”, “recall” y “accuracy”</i> | 117 |

1

Introducción

1.1. Antecedentes

Las prácticas agrícolas tradicionales han estado orientadas hacia la explotación de los recursos agrícolas provocando en muchos casos situaciones no deseables respecto de la contaminación atmosférica, del suelo o del subsuelo. Tradicionalmente el tratamiento con agroquímicos de grandes extensiones de cultivo se realiza con una dosis fija, la más adecuada a la peor situación observada, y de forma uniforme. Recientemente se ha tomado conciencia a distintos niveles considerando que tales actuaciones conllevan una pérdida importante de recursos, además del impacto ecológico que ello supone, no sólo por la contaminación de suelos sino también porque puede dar lugar a la aparición de resistencias difíciles de erradicar. No es lo mismo aplicar un tratamiento a una zona con más alta densidad de vegetación, que a otra con baja densidad e incluso en alguna en la que no existe vegetación.

Centrándonos en el tratamiento de las malas hierbas en campos de cultivo con fines de su eliminación, es cierto, que se debería aplicar un tratamiento más agresivo en aquellas zonas con más alta densidad que en las zonas con menor densidad, incluso no realizar ningún tipo de tratamiento si la densidad de malas hierbas es muy baja o incluso nula.

También resulta de interés considerar el estado de crecimiento del cultivo y su capacidad de competir con la mala hierba, a la hora de ajustar la dosis.

La utilización de productos fitosanitarios de forma generalizada trae como consecuencia la contaminación del suelo, la disminución de la biodiversidad, la contaminación del agua, e incluso peligros para la salud. Además de lo anterior, estas situaciones generan el aumento de los costes de producción. Esto, unido al aumento de población, a la disponibilidad de menor extensión de terreno y agua para el cultivo, hace que las prácticas agrícolas sean cada vez más costosas tanto desde un punto de vista social, como económico o medioambiental.

En el siglo XXI la Política Agraria está otorgando mayor protagonismo a los aspectos medioambientales relacionados directamente con factores que afectan a todo el planeta, como son la generación de gases efecto invernadero, la contaminación de la atmósfera y el suelo por el uso excesivo de plaguicidas, riesgos en la seguridad alimentaria, la degradación del suelo, desaparición de especies y ecosistemas y la gestión inadecuada de los recursos hídricos.

El aumento de la población y la necesidad de producir el alimento suficiente, han hecho que los Comisarios europeos de agricultura aboguen por una agricultura sostenible y las políticas europeas vayan encaminadas a conseguir este objetivo. La Política Agraria Común (PAC) ha experimentado grandes cambios desde su inicio, su objetivo actual es ayudar a la agricultura a cumplir los requisitos del desarrollo sostenible. En estos términos lo especifica la Comisión Europea (CE 2014), cuando emite las directrices aplicables a las ayudas estatales en los sectores agrícola y forestal y en las zonas rurales de 2014 a 2020, en el punto 9 de las disposiciones comunes dice:

“La agricultura ha de adaptarse a las nuevas realidades, enfrentarse a los retos de la seguridad alimentaria, el medio ambiente y el cambio climático y mantener viva la economía agrícola. Como parte integrante de la PAC, la política de desarrollo rural de 2014-2020 contribuirá a lograr los siguientes objetivos:

- 1) Fomentar la competitividad de la agricultura.*
- 2) Garantizar la gestión sostenible de los recursos naturales y la acción por el clima.*

3) *Lograr un desarrollo territorial equilibrado de las economías y comunidades rurales incluyendo la creación y conservación del empleo.*”

Para conseguir estos objetivos la Comisión aconseja tener en cuenta unas prioridades en materia de desarrollo, entre las cuales destacan las siguientes:

- Fomentar la transferencia de conocimientos e innovación.
- Mejorar la viabilidad de las explotaciones agrícolas y la competitividad de los diferentes tipos de agricultura en la diversidad de las regiones, y promover las tecnologías agrícolas innovadoras.
- Restaurar, preservar y mejorar los ecosistemas relacionados con la agricultura.
- Promover la eficiencia de los recursos y fomentar el paso a una economía baja en carbono y capaz de adaptarse al cambio climático.

Estos consejos de la Comisión Europea unidos a que la sociedad se va transformando como consecuencia de los avances tecnológicos hacen que, en agricultura, cada vez sea más importante la inversión en ciencia y tecnología, desde el desarrollo de mejores productos fitosanitarios hasta su aplicación de forma controlada; del mismo modo, cada vez es más frecuente hacer uso de lo que las nuevas tecnologías de la información y la comunicación (TIC) nos ofrecen. Algunas de las tecnologías que se utilizan en la agricultura con tal finalidad son:

- Los Sistemas de Posicionamiento Global por Satélite (GPS), permiten determinar en cada momento la posición, con la precisión que los sistemas GPS determinen, de los equipos y sistemas utilizados en los campos agrícolas durante el trabajo.
- Dispositivos de recolección de datos (sensores y emisores), que permiten detectar, tratar y almacenar la información sobre la variabilidad del cultivo.
- Los Sistemas de Información Geográfica (SIG), que constituyen una integración de *hardware, software y datos geográficos* diseñados para capturar, almacenar, manipular, analizar y desplegar en todas sus formas la información geográficamente referenciada.
- Las tecnologías que son capaces de controlar de forma automática los equipos agrícolas, de suerte que se puedan regular los mismos en tiempo real.

Los computadores en general y los sistemas de visión artificial en particular, ambos en continuo desarrollo, han permitido la automatización de múltiples procesos de una forma bastante eficaz, así como realizar cambios en continuo, sin tener necesidad de detener la tarea que se está realizando. Todo esto ha favorecido la aparición de la denominada Agricultura de Precisión (AP), que engloba un conjunto de técnicas de cultivo dirigidas a realizar operaciones selectivas y más específicamente a ajustar de forma controlada el uso de agroquímicos, considerando la diversidad tanto del medio físico como biológico (Kropff y col. 1997).

Con técnicas propias de AP es posible adaptar los insumos a las necesidades de cada cultivo, dado que las características de los suelos no son uniformes en su distribución espacial, y los fertilizantes y los productos fitosanitarios se pueden aplicar en dosis concretas y en el lugar preciso. Esto no implica que los tratamientos no sean efectivos al reducir los insumos utilizados mediante una aplicación selectiva, tampoco que la cantidad y la calidad de los productos finales sea menor (Earl y col. 1996; Krishna 2013). Con una correcta actuación en el momento adecuado y en el lugar preciso, se puede conseguir una mayor calidad de las producciones.

Para realizar tareas y tratamientos específicos es necesario identificar las variaciones existentes en el campo de cultivo, para ello las imágenes de alta resolución constituyen una buena fuente de información que permite estudiar las condiciones del cultivo y del suelo, así como determinar las zonas donde se requieren tratamientos específicos. Sin embargo, la disponibilidad y los costes a menudo prohibitivos de tales imágenes podrían sugerir un producto alternativo para esta aplicación particular en la AP. En concreto, las imágenes tomadas por sensores remotos, o dispositivos a bordo de vehículos aéreos no tripulados (Unmanned Aerial Vehicles, UAV), son una alternativa potencial dado su bajo coste de operación para la monitorización ambiental, alta resolución espacial y temporal, además de su alta flexibilidad en la adquisición de imágenes (Zhang y Kovacs, 2012; Torres-Sánchez y col., 2014, 2015; Pajares, 2015)

El problema fundamental de los UAV es que hoy en día la tecnología no está suficientemente desarrollada como para permitir una detección en tiempo real y la actuación simultánea sobre el terreno, a este tipo de sistema de visión se le denomina indirecto, ya que necesitan de un procesamiento posterior a la captura de imágenes,

conjuntamente con una supervisión, apoyada por sistemas computacionales, para identificar las zonas a tratar.

Por el contrario, los sistemas de Visión Artificial directos instalados a bordo de los robots terrestres o tractores permiten una acción inmediata sobre el cultivo o los campos de cultivo en general, dependiendo de la aplicación. Gran parte de ellos han sido diseñados para detectar las plantas verdes mediante diferentes sistemas sensoriales en función de sus respuestas espectrales a las distintas longitudes de ondas, por ejemplo los que adquieren las imágenes en el espectro visible en forma de imágenes RGB o los que captan el infrarrojo cercano. En Berge y col. (2012) utilizan un robot equipado con un pulverizador, guiado por un GPS, que capta imágenes en RGB e identifica sus características y la información necesaria para determinar y reconocer los diferentes objetos, incluyendo las texturas vegetales, que se encuentran en los campos de cereal, para ello utilizan una herramienta de software llamada Weedcer, está desarrollada por SINTEF Information and Communication Technology sobre la que la compañía Adigo AS ha introducido posteriores mejoras y desarrollos para abordar aplicaciones en tiempo real.

En el presente trabajo de investigación nos centramos en los sistemas de Visión Artificial directos, que para las tareas agrícolas presentan el siguiente conjunto de requisitos:

1. Actuar en tiempo real, para que no sea necesaria la intervención de un usuario cuando las imágenes hayan sido procesadas.
2. Suficiente resolución, para detectar las estructuras vegetales de interés, incluyendo las malas hierbas y controlar el vehículo con la mayor precisión posible, cuando esta tarea sea necesaria. La detección de malas hierbas para su eliminación, constituye una de las aplicaciones básicas en AP.
3. Robustez suficiente para ser eficiente en diversas situaciones, como las que aparecen en los entornos agrícolas de exterior, con alta variabilidad, principalmente en lo que respecta a las condiciones de iluminación.

A la hora de realizar una aplicación selectiva de tratamientos mediante Visión por Computador necesitamos disponer de una buena materia prima, es decir, unas imágenes de calidad para poder identificar las distintas texturas y realizar la actuación inmediata sobre el cultivo. Uno de los temas clave de trabajo en AP es el del control de malas hierbas, ya

que éstas compiten con los cultivos por los recursos naturales, es decir el espacio, la luz, los nutrientes y el agua, lo que tiene consecuencias directas sobre la calidad y cantidad de la cosecha final. Muchos cultivos agrícolas requieren el uso de herbicidas como técnica esencial de control de infestaciones para mantener la calidad y cantidad de la cosecha. El coste en herbicidas puede representar aproximadamente el 40% del coste de todos los productos químicos aplicados en agricultura en Europa, pero las preocupaciones ambientales y económicas asociados a este tipo de tratamientos han llevado a la creación de la legislación europea sobre el uso sostenible de los plaguicidas (Willians 2011). Esta legislación incluye directrices para la reducción de las aplicaciones y la utilización de dosis adecuadas en función del grado de infestación que sufra el cultivo (Peña y col., 2013).

La identificación de zonas con ciertas densidades de cultivo o la detección automática de malas hierbas ha sido una tarea desalentadora, dada la gran cantidad de variables a tener en cuenta y la dificultad de adaptar las técnicas a la realidad. Aun siendo esto así, algunos trabajos demuestran el potencial de la Visión Artificial aplicada a la AP, y permiten vislumbrar que la aplicación real de dichas técnicas no está lejos en el tiempo, pues la investigación al respecto está muy avanzada (RHEA, 2015). La solución a este problema no es trivial, el crecimiento de la capa vegetal, la humedad, las condiciones cambiantes de iluminación a lo largo del día, las diferentes especies de malas hierbas encontradas, y la similitud que ofrecen éstas con los cultivos hacen que la discriminación sea una tarea compleja. Según Zheng y col. (2010) resulta difícil separar las diferentes texturas en las imágenes en color conteniendo cultivos y malas hierbas, debido a que son imágenes de exterior donde las condiciones de iluminación son muy cambiantes en cortos intervalos de tiempo e incluso las sombras de las plantas puede influir directamente dificultando una correcta clasificación o discriminación de tales texturas.

1.2. Identificación del problema

Ahondando más en la gestión precisa de un cultivo, la AP tiene como objetivo optimizar el rendimiento de las explotaciones agrarias determinando el área del campo donde se ubican los cultivos y ejecutando de forma precisa las decisiones adoptadas. Esta optimización tiene una repercusión muy favorable en los aspectos de rendimiento y medioambientales ligados a la producción agrícola, principalmente por la racionalización del uso de productos fitosanitarios que lleva asociada. Las TIC han puesto a disposición del sector agrario todo un conjunto de nuevas herramientas que hacen posible dicha

optimización, entre las que se encuentran los sistemas de Visión Artificial, que permiten la localización e identificación de las malas hierbas en extensas superficies cultivadas con el fin de realizar una gestión precisa y económicamente asumible.

El presente trabajo de investigación se inicia ante la necesidad de reconocer de forma automática las texturas en un campo de cultivo de cereal y maíz, nuestro objetivo será separar la parte vegetal, que incluye el cultivo y las malas hierbas, del resto (cielo y suelo), tarea compleja por las condiciones de iluminación cambiantes en los entornos de exterior donde se trabaja. A continuación, ya en la capa vegetal, queremos discriminar entre cultivo y malas hierbas. Esto viene determinado por la necesidad de eliminar las malas hierbas para facilitar el crecimiento del cultivo. Esta operación se realiza en el momento en que tanto el cultivo como las malas hierbas se encuentran en proceso de crecimiento y por tanto presentan una tonalidad verdosa, con componentes espectrales en el rango de longitudes de onda correspondientes a dicha parte del espectro visible. Esta similitud, en lo que a la signatura espectral se refiere, añade un componente de dificultad adicional al proceso de identificación de las malas hierbas en campos de cultivo (maíz o cereal) su solución constituye uno de los objetivos de la presente investigación.

En una primera fase de la investigación tendremos dos tipos de imágenes, unas sólo con suelo y otras con suelo y cielo, y en otra segunda fase sólo serán imágenes del suelo con una capa vegetal que incluye cultivo y malas hierbas. La segmentación en las imágenes de cultivos aparece como un reto que se aborda constantemente, como se comprueba a lo largo de la bibliografía referenciada en el siguiente capítulo, con aportaciones significativas pero todavía no concluyentes, que es donde tiene cabida el presente trabajo de investigación. En líneas generales, en esta tesis se aborda la clasificación de las texturas en imágenes de los cultivos mencionados con el objetivo de automatizar el proceso de detección.

Un primer paso será obtener las imágenes para que la aplicación selectiva alcance el éxito deseado. Desde hace más de una década se han utilizado cámaras para capturar imágenes de campos de cultivo con el fin de realizar tratamientos agrícolas en grandes extensiones de terreno (Davies y col., 1998). Una configuración común es el uso de un vehículo robot equipado con sensores de Visión Artificial. En diversas ocasiones dicho vehículo es en realidad un tractor que de forma automática y en tiempo real reconoce la ubicación de la zona de actuación concreta, esto es, en el caso de las malas hierbas dónde

se localizan éstas. Existen en la literatura diversas soluciones al problema, siendo el objetivo de la presente investigación el proporcionar métodos eficientes, basados en técnicas de Visión por Computador, lo suficientemente eficientes, en lo que respecta a la precisión, la identificación de las partes vegetales.

1.3. Motivación y objetivos

El trabajo de investigación que se presenta en esta memoria tiene su fundamento en los proyectos de investigación que se mencionan a continuación, pertenecientes al plan nacional de I+D+i y de la Unión Europea dentro del marco FP7, a saber:

- 1) Desarrollo de Métodos de Identificación de Texturas Naturales y su Localización Espacial mediante métodos de Visión Artificial (AGL2008-04670-C03-02), en el que se plantea justamente la identificación de estructuras naturales en campos de cultivo mediante métodos basados en visión computacional.
- 2) Robot Fleets for Highly Effective Agriculture and Forestry Management" Theme 4 NMP-Nanosciences, Nanotechnologies, Materials and New Production Technologies (CP-IP 245986-2 (RHEA) FP7-NMP-2009-LARGE-3, finalizado con éxito tras la evaluación por parte de técnicos de la Unión Europea en Mayo de 2014. En este proyecto la identificación de los rodales de malas hierbas mediante métodos basados en Visión por Computador constituyó un elemento clave de investigación, posteriormente sustentada en gran medida por los desarrollos realizados en el marco del presente proyecto.
- 3) Integración de Información Multisensorial y Aprendizaje Automático para la Detección, Caracterización y Reconocimiento Preciso de Estructuras Naturales en Campos de Cultivo (AGL2014-52465-C4-3-R), actualmente vigente y en el que se están aplicando las estrategias descritas en la presente memoria de investigación.

Por tanto, una vez identificado el problema derivado de los proyectos anteriores y con el fin de conseguir el desarrollo de métodos automáticos para la detección de las distintas texturas de una imagen de un cultivo de cereal o maíz independientemente de las condiciones de iluminación existentes en el momento de la toma de la imagen, se plantean los siguientes objetivos:

1. Identificar métodos de segmentación de imágenes existentes en la literatura cuya finalidad es la identificación de texturas naturales.
2. Desarrollar métodos automáticos de procesamiento de imágenes agrícolas de exterior para la correcta extracción y discriminación de cultivo y malas hierbas bajo una gran diversidad de condiciones posibles de trabajo.
3. Diseñar estrategias para identificar las distintas clases de verdor y distintas clases de texturas del suelo.
4. Analizar y validar los resultados de las soluciones propuestas.
5. Identificar las líneas de investigación futuras para mejorar en lo posible tanto los métodos desarrollados para el tipo de cultivo actual, como para la aplicación a otros posibles tipos de cultivo diferentes.
6. Difundir los resultados mediante publicaciones científicas u otro tipo de divulgación.

1.4. Metodología

Para llevar a cabo los objetivos anteriores, se han seguido los siguientes pasos:

1. Consultar información bibliográfica de todas las fuentes disponibles, analizando y evaluando el estado del arte.
2. Definir el problema de investigación.
3. Plantear los objetivos.
4. Obtener los datos para la realización de los experimentos.
5. Analizar, clasificar e interpretar los datos obtenidos.
6. Extraer las conclusiones.
7. Publicación de los trabajos con el fin de dar a conocer el trabajo desarrollado a la comunidad científica en el ámbito de la AP y la Visión por Computador.
8. Establecer las futuras líneas de investigación.

Los experimentos se han realizado con Matlab (Mathworks, 2015) se han implementado los métodos con las técnicas propuestas, las imágenes han sido procesadas con las toolboxes Image Processing y Wavelet, específicos para el tratamiento de imágenes y wavelet respectivamente. El proceso seguido para la realización de los experimentos ha sido el siguiente: 1) obtención de datos, 2) desarrollo e implementación de métodos

basados en técnicas de Visión por Computador y 3) visualización y obtención de resultados.

1.5. Aportes a la investigación

La investigación ha estado enfocada a identificar con suficiente eficiencia, y precisión, las distintas texturas en las imágenes de cultivos de cereal y maíz. La importancia de esta identificación de texturas es localizar el lugar exacto donde se encuentran las malas hierbas u otros tipos de estructuras de interés, en el contexto del campo de cultivo. Para ello, tendremos en cuenta los comportamientos espectrales similares que tienen las plantas de cultivo y las malas hierbas, tal y como se ha mencionado previamente, para que en el momento de la detección se realice el tratamiento de forma precisa al objeto de eliminar las malas hierbas, o para la detección de esas estructuras para la navegación de los vehículos autónomos de forma segura. Las siguientes publicaciones constituyen los aportes relevantes al trabajo de investigación recogido en la presente memoria.

- 1) **Guijarro, M., Pajares, G., Riomoros, I., Herrera, P. J., Burgos-Artizzu, X. P., Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture*, 75(1), 75-83.**

En este trabajo se describe un método automático para segmentar las distintas texturas de imágenes de campos de cereal y maíz (suelo, cielo y plantas) mediante la combinación de índices cromáticos y agrupamiento borroso. De los distintos tipos de índice cromáticos que utilizan el espectro infrarrojo y el visible, se eligen estos últimos, puesto que para el tipo de imágenes analizadas y las características a extraer de ellas son los más apropiados dado que acentúan un color (banda espectral) en particular. Al aplicar los índices cromáticos a las imágenes se obtiene una imagen en escala de grises que definen los diferentes verdes. En este trabajo, para identificar el verdor se propone una nueva estrategia combinada de índices cromáticos, determinando la relevancia de cada índice en función de su contribución para la detección de los verdes. La relevancia se determina mediante el histograma de niveles de intensidad en la imagen de grises, a partir de dos medidas: uniformidad y entropía. La uniformidad es máxima cuando los niveles de gris son iguales y disminuye a partir de ahí. La entropía es una medida aplicada en

procesos aleatorios, que es mínima cuando todos los niveles de gris son iguales. En la estrategia propuesta se busca la uniformidad mínima y la máxima entropía, aunque como las dos medidas tienen un comportamiento similar en las imágenes utilizadas se elige la uniformidad que es menos costosa computacionalmente hablando. El valor de la uniformidad calculado para cada una de las imágenes, sobre las que previamente se aplican los índices de verdor, determina la relevancia de cada una de las imágenes a combinar. Una vez establecido el índice combinado, con el que se extraen las texturas verdes, se aplica la técnica de agrupamiento borroso o *fuzzy clustering*, y de esta forma se obtienen tres tipos distintos de verde, que permiten discernir entre plantas. Este método sirve también para discriminar entre varios tipos de suelo. Además, se propone la adaptación de dos índices cromáticos para identificar los rojos y azules predominantes, correspondientes al suelo y al cielo respectivamente existente en estas imágenes panorámicas. La panorámica ofrecida por las imágenes permite identificar posibles obstáculos para una navegación segura de los vehículos autónomos agrícolas. Por otro lado, con la información obtenida a partir de la combinación de verdes y mediante la aplicación de la técnica de agrupamiento borroso se obtienen dos clases diferentes de verdes, que permiten diferenciar plantas de cultivo y malas hierbas.

Este enfoque combinado obtiene mejores resultados que los obtenidos aplicando los métodos por separado. El método es válido tanto para las imágenes que contienen cielo como para las que no. Este sistema resulta útil para ser utilizado por un vehículo que navega por un campo de cultivo de forma autónoma equipado con un sistema de visión con el objeto de aplicar tratamientos agrícolas selectivos. Se ha demostrado que este método es eficiente bajo diferentes condiciones de iluminación en ambientes de exterior, como son los agrícolas. Este método se describe en el capítulo 4, sección 4.1, mientras que los resultados obtenidos aparecen en el capítulo 5, sección 5.1 de la presente memoria.

2) M. Guijarro, I. Riomoros, G. Pajares, P. Zitinski, (2015). Discrete wavelets transform for improving greenness image segmentation in agricultural images. Aceptado para publicación, Septiembre 2015.

En este trabajo se describe un método que combina información espectral con información espacial para identificar texturas vegetales en campos de cultivo de maíz con malas hierbas. Esta combinación constituye una novedad intrínseca en sí misma. De las imágenes de campos de cultivo con iluminación variable, extraemos los verdes mediante la

aplicación del índice de vegetación exceso de verde (ExG). En el trabajo anterior se utilizó una combinación de índices, en este enfoque es suficiente un único índice debido a que la información espacial por sí misma supera los resultados de la combinación. Una vez segmentada la imagen, mediante la utilización del índice de vegetación, se aplica la transformada de wavelets sobre la imagen segmentada para extraer la información espacial de las texturas a través de sucesivos filtrados en baja y alta frecuencia con un filtro perteneciente a la familia de las Daubechies. Las transformadas wavelets son una de las mejores herramientas para separar las componentes de baja frecuencia de la imagen, ligeras variaciones de intensidad, y las de alta frecuencia, bordes y alta variabilidad que proporcionan los detalles. Por lo tanto, permiten una descomposición de la imagen a diferentes escalas (multiresolución). Al aplicar la transformada obtenemos los coeficientes de aproximación, que preservan las características más relevantes de la imagen de partida, y los coeficientes de detalle (horizontales, verticales y diagonales), que contienen la información suficiente para identificar los píxeles como pertenecientes a la cubierta vegetal, ya que estas estructuras vegetales poseen una alta variabilidad de intensidad espacial. De esta forma, para cada píxel de la imagen segmentada se analiza su vecindad con el fin de determinar su variabilidad mediante análisis estadístico a través de descriptores estadísticos de primer orden, más concretamente la varianza. Las varianzas proporcionadas por los coeficientes de detalle se combinan entre sí, permitiendo incrementar el valor del índice de verdor en ese píxel cuando las varianzas son significativas, asumiendo que los valores significativos son aportados por la vegetación. De esta forma obtenemos una imagen realizada en las texturas vegetales, que en un paso posterior se binariza aplicando el método de umbral de Otsu, obteniendo resultados satisfactorios.

El enfoque propuesto es especialmente útil cuando la calidad del verdor de las imágenes es baja. Este estudio tiene por objetivo ser utilizado en tiempo real en el futuro por un sistema de visión utilizado en trabajos agrícolas. Este método se describe en el capítulo 4, sección 4.2, mientras que los resultados obtenidos aparecen en el capítulo 5, sección 5.2 de la presente memoria.

- 3) Riomoros, I., Guijarro, M., Pajares, G., Herrera, P. J., Burgos-Artizzu, X. P., Ribeiro, A. (2010). Automatic image segmentation of greenness in crop fields. IEEE Int Conf. Soft Computing and Pattern Recognition (SoCPaR'10), pp. 462-467, ISBN 978-1-4244-7897-2, 7-10 December, Paris, France.**

En este trabajo el objetivo es identificar varias clases de verde en imágenes procedentes de campos de cultivo. La principal aportación del artículo consiste en el diseño de un método no supervisado de aprendizaje a partir de un método netamente supervisado como es el conocido como Cuantización Vectorial. El proceso de segmentación consta de dos pasos, primero se aplica una técnica de binarización mediante el índice ExG y en segundo lugar se separan los píxeles según su pertenencia a una determinada clase. Se fija el número de clases deseado y mediante el índice de Dunn (Dunn 1974) se valida la partición con el número de clases previamente establecido, de suerte que si el valor del índice no es satisfactorio se varía el número de clases hasta conseguir un valor de índice considerado como válido.

Este método fue el precursor de los métodos de clasificación utilizados durante el proceso de investigación con la finalidad de establecer una partición de los elementos existentes en las imágenes, particularmente del agrupamiento borroso que se aplicó en el primer artículo referenciado en este documento y descrito en la sección 4.1 de esta memoria.

- 4) Arroyo, J., Guijarro, M., Riomoros, I., Pajares, G., (2012). A symbolic solution to a precision agriculture. Proc. 3rd Workshop in Symbolic Data Analysis (Arroyo, J., Maté, C., Brito, P., Noirhomme-Fraiture, Eds.), 7-9 November 2012, Madrid, Spain, pp. 25-26.**

En este trabajo se utiliza un enfoque basado en aprendizaje para establecer el umbral óptimo de una imagen segmentada mediante índices de vegetación. Como paso previo, las instancias (histogramas en escala de grises) de entrenamiento se han seleccionado previamente, cada una con un valor de umbral óptimo que se ha determinado de forma manual. Dada una nueva imagen, su umbral óptimo se determinará teniendo en cuenta los valores límite de los ejemplos de entrenamiento más cercanos. Estos métodos también se conocen como métodos de aprendizaje ponderados localmente. En este trabajo se comparan diferentes estrategias basadas en el aprendizaje localmente ponderado para

binarizar imágenes, separar el suelo de las plantas verdes (cultivo y malas hierbas), métodos que se pueden utilizar de forma automática en un vehículo robot impulsado en tiempo real, previa selección de las instancias de entrenamiento. Los resultados experimentales obtenidos confirman que el método es válido cuando se dispone de una base suficiente de ejemplos, pero que en imágenes de exterior en entornos agrícolas, resulta más conveniente evitar la dependencia de entrenamientos previos, dada la alta variabilidad de las condiciones ambientales, principalmente la iluminación.

Este trabajo, por tanto, confirma el enfoque dado en la presente investigación y orientado hacia la utilización de estrategias con capacidad de adaptación automática a condiciones variables, como es el caso del método de Otsu que es el utilizado en las investigaciones recogidas en la presente memoria, tal y como se menciona en la sección 3.3.1.

1.6. Organización de la memoria

La memoria de esta tesis se presenta estructurada en capítulos, que se organizan siguiendo el orden de la investigación, la secuencia de la presentación es la siguiente:

Primer capítulo. *Introducción*, en este capítulo además de introducir y motivar las propuestas de los planteamientos realizados, como hemos visto hasta ahora, se formulan también los objetivos de la investigación.

Segundo capítulo. *Estado del arte*, se revisan las distintas estrategias existentes en la literatura, que proporcionan los argumentos y la base de la investigación orientados a la identificación de texturas mediante análisis espectral y espacial.

Tercer capítulo. *Fundamentos teóricos de la segmentación de imágenes agrícolas y detección de las malas hierbas*, donde se describen los conceptos teóricos de las técnicas utilizadas en las distintas fases de las investigaciones realizadas.

Cuarto capítulo. *Sistemas Automáticos: segmentación de imágenes agrícolas y detección de las malas hierbas en maíz*, donde se exponen las aportaciones realizadas, que a su vez están avaladas por las publicaciones previamente descritas.

Quinto capítulo. *Análisis de resultados*, describiendo los resultados de forma separada para cada estrategia. Por una parte aparecen los resultados obtenidos por la primera investigación y a continuación los de la segunda, ambas descritas en el capítulo cuatro.

Sexto capítulo. *Conclusiones y trabajo futuro*, se extraen las conclusiones pertinentes, se exponen las líneas seguidas y se plantean las líneas de investigación futura.

2

Estado del arte

En este capítulo se realiza una revisión de diferentes métodos y estrategias existentes en la literatura especializada que han sido estudiados con el fin de desarrollar una estrategia precisa y automática para la segmentación de imágenes orientada a la detección de cultivos/malas hierbas en imágenes agrícolas. Se presenta un análisis bibliográfico sobre el estado del arte actual y los últimos avances en los sistemas de visión en el campo de la AP con el fin de poder resolver el problema planteado.

A partir de este planteamiento, las dos principales líneas de investigación que se analizan en mayor profundidad con referencia a la literatura existente son:

- a) Combinación de índices de vegetación para la obtención de la cubierta vegetal con métodos automáticos de identificación, separación y reclasificación de texturas basadas en índices espectrales.
- b) Descomposición de la imagen mediante la transformada de wavelets con el fin de identificar las estructuras espaciales existentes en las imágenes, que caracterizan las

texturas vegetales mediante discriminación espacial. Que conduce a la inserción de información espacial sobre los índices cromáticos.

2.1. Visión general

Desde muy antiguo se sabe que la cosecha recogida en una explotación agraria no es uniforme en toda ella, a pesar de haber realizado un manejo del cultivo homogéneo y que éste ha estado sometido a las mismas condiciones medioambientales. Existen ciertos factores íntimamente unidos a las propiedades del suelo, los microclimas o las plagas, que afectan de manera selectiva y con distintos grados de intensidad a zonas concretas del campo, dando lugar a lo que se conoce como variabilidad intra-parcelaria.

En grandes extensiones de terreno el tratamiento homogéneo no es aconsejable, sobre todo a nivel de aplicación de herbicidas, siendo conveniente en estos casos aplicar el fitosanitario de forma localizada, para minimizar la contaminación tanto del suelo como de las personas que los manipulan. Por ello, la AP, como se ha mencionado reiteradamente, se orienta a la realización de tratamientos selectivos en un área de cultivo específica, para lo que se utilizan herramientas tecnológicas de distinta índole, donde los sistemas de detección basados en Visión Artificial constituyen una parte importante de las mismas.

Lo anterior unido a que las malas hierbas son una amenaza importante en la producción de cultivos, ya que compiten con ellos, originando una pérdida en la producción, nos conduce a pensar que la identificación de las estructuras vegetales, incluyendo las malas hierbas, juega un papel vital en la agricultura en general y más concretamente en la AP. En el caso que nos ocupa, esta identificación se realiza con la ayuda de la Visión Artificial, tomando imágenes de los cultivos e identificando las texturas pertinentes. El análisis de la imagen digital se ha aplicado a otros muchos aspectos de la AP, tal como guiado o arado de campos; si bien, en la presente investigación los desarrollos se orientan a la identificación de texturas, donde las partes vegetales constituyen una parte esencial del proceso. Pajares y Cruz (2007) describen la Visión Artificial como la capacidad de una máquina para ver el mundo que le rodea, deduciendo la estructura y las propiedades del mundo tridimensional a partir de una o más imágenes bidimensionales.

El problema ha de ser abordado de forma muy distinta a como se haría en entornos controlados, ya que las condiciones ambientales, altamente cambiantes, hacen que los métodos de procesamiento y tratamiento de imágenes deban diseñarse con suficiente robustez para su adaptación a dichas condiciones. Según Zheng y col., (2010) es difícil separar las texturas vegetales con similares componentes espectrales en las imágenes en color, debido a que son imágenes de exterior donde las condiciones de iluminación son ciertamente cambiantes, en su trabajo hacen una distinción entre diferentes tipos de luz, como por ejemplo la luz directa con sombras, oblicua con sombras y la luz difusa de un día nublado. Estas variaciones se deben a las condiciones climatológicas del momento que pueden causar cambios importantes en las componentes espectrales de los objetos que se proyectan sobre el sensor del sistema de visión (Tian y Slaughter, 1998). A modo de ejemplo, considérese un día con una alta densidad de nubes que ocultan el sol intermitentemente o un tractor equipado con un sistema de visión desplazándose cara al sol y en su regreso de espaldas al mismo.

Por tanto, resulta una tarea compleja clasificar las diferentes texturas existentes en las imágenes (partes verdes, suelo, cielo) de forma automática para su tratamiento tanto desde el punto de vista espectral como espacial, según las líneas de investigación seguidas. El principal problema relacionado con la aplicación de métodos de procesamiento de imágenes es el relativo a la segmentación de las propias imágenes. Una segmentación automática y eficiente de la vegetación de las imágenes agrícolas constituye un paso importante para muchas aplicaciones como por ejemplo la detección de malas hierbas, para de esta forma realizar un tratamiento específico en una localización espacial determinada (Onyango y Marchant, 2003; Sainz-Costa y col., 2011; Tellaeche y col., 2008a; Tellaeche y col., 2008b; Burgos- Artizzu y col., 2009; Guijarro y col., 2011; Guerrero y col., 2012; Montalvo y col., 2012). López-Granados (2011) realiza una revisión de métodos donde el tema clave de los mismos es la identificación de las plantas.

La mayoría de las estrategias van encaminadas a identificar el verde, basándose en que las plantas tienen un alto grado de verdor, pero no consideran el hecho de que las plantas pueden haber perdido en parte o en su totalidad su grado de verdor por diferentes razones (Montalvo y col. 2013), entre ellas destaca el hecho de que los materiales arcillosos del suelo, en particular el barro formado por la lluvia y la tierra, pueden cubrir la planta o parte de ella, por lo que la componente espectral verde de los píxeles pertenecientes a las plantas

afectadas por este fenómeno no será la predominante, sino que en tales casos será la componente roja, que proviene de los componentes principales de los materiales existentes en la tierra. Por otra parte, la discriminación de las texturas, no ya por su signatura espectral, sino por su distribución espacial, no ha sido abordada hasta el momento en la literatura con la intensidad de las técnicas basadas en el análisis espectral.

El problema planteado en la investigación que se describe en el presente trabajo es la segmentación de imágenes que contienen cultivos de cebada y maíz, con existencia de malas hierbas, de suerte que el objetivo principal es la discriminación de las partes verdes conteniendo plantas de cultivo y malas hierbas, la cual se realiza mediante técnicas de Visión por Computador. El sistema de visión que captura las imágenes está instalado a bordo de un tractor automatizado, que se desplaza en el campo, y con el que se obtienen dos tipos de imágenes. Un primer tipo en las que sólo aparece el suelo que contiene la tierra de cultivo y plantas, y un segundo tipo de imágenes más panorámicas, conteniendo parte del cielo. En el primer tipo existen dos clases de texturas principales, suelo y vegetación, que en el momento del tratamiento muestran tonalidades predominantemente verdosas. En el segundo tipo habría tres texturas distintas cielo, tierra y vegetación, estas imágenes tienen el propósito de servir simultáneamente como ayuda a la navegación, ya que por su panorámica permiten identificar otros obstáculos de cara a la navegación segura. No obstante, este aspecto queda fuera del ámbito de la presente investigación.

Teniendo en cuenta las anteriores premisas, es necesario diseñar una estrategia automática de Visión por Computador para hacer frente a los tipos de imágenes mencionados, en la que han de tenerse en cuenta dos aspectos muy importantes, que las imágenes pueden ser adquiridas en diferentes condiciones de iluminación y que el proceso ha de ser computacionalmente eficiente.

2.2. Visión Artificial

Los sistemas de Visión Artificial o por Computador (ambos términos usados indistintamente) buscan reproducir cómo el ser humano capta la luz a través de los ojos y esa información viaja por el nervio óptico al cerebro, donde se procesa la imagen e interpreta la escena.

La Visión Artificial tiene como objetivo extraer información del mundo físico a partir de imágenes, utilizando para ello un computador. En ese intento de reproducir el comportamiento de la visión humana, intervienen cuatro fases:

1. Captura: se adquiere la información, es decir, se obtiene la imagen mediante algún tipo de sensor.
2. Procesamiento: tratamiento digital de la imagen donde se eliminan las características no deseadas y se resaltan las que resultan de interés.
3. Segmentación: se aíslan los elementos de interés existentes en la imagen para su tratamiento.
4. Clasificación: en la que se distinguen los objetos segmentados en base a unas características que se han establecido previamente para su diferenciación.

Normalmente, estas fases se siguen de manera secuencial en un sistema de Visión Artificial, es posible incorporar un ciclo de realimentación, de esta forma si la segmentación falla, se pueda regresar a cualquiera de las dos fases anteriores y volver a realizar el procesamiento o la captura de la imagen, según convenga, para intentar solucionar el fallo.



Figura 2.1.- Fases de un sistema de Visión Artificial por computador.

Dentro del esquema de la figura 2.1, el módulo correspondiente a la captura de imágenes depende exclusivamente tanto del dispositivo de captura utilizado (sensor) como de la plataforma que lo soporta (satélite, UAV, tractor, etc.). La fase de procesamiento se refiere al tratamiento previo de la imagen con el fin de prepararla para la siguiente fase. La segmentación, constituye la fase fundamental dentro del planteamiento que se formula en la presente investigación. Su objetivo consiste en identificar estructuras y elementos relevantes en las imágenes con el fin de distinguir las de interés de las que no lo son, llegando así a su clasificación, que constituye la fase final del proceso y que determina las partes de la escena que establecen el objetivo de la actuación.

A continuación veremos algunos aspectos relacionados con las plataformas de captura, después realizaremos la revisión bibliográfica relativa a la segmentación como fase clave, lo referente al procesamiento puede considerarse englobado en la segmentación y la fase de clasificación es el resultado final de la segmentación.

Así, teniendo en cuenta que la investigación se ha orientado a la identificación de texturas por su respuesta espectral y su composición espacial, la revisión se realiza analizando los índices de vegetación y su descomposición en coeficientes, de aproximación y de detalle, respectivamente. Posteriormente, en el capítulo tres se proporcionan los fundamentos teóricos relativos a estos dos mismos planteamientos.

Una imagen digital se trata como una función bidimensional $f(x,y)$, donde el valor de la función en el punto (x,y) representa la intensidad de radiación recibida por el sensor en dicho punto. A los puntos se les llama píxeles y constituyen las unidades elementales de información de la imagen. Generalmente las imágenes se representan como matrices.

Para formar imágenes a color, se necesitan varios canales o planos de color, cada uno de ellos representa valores de intensidad dentro del correspondiente rango espectral visible, al combinar varios canales se forma la imagen en color. El modelo de color más usado en imágenes digitales es el RGB (Rojo, Verde, Azul), que es el que más se asemeja al funcionamiento del ojo humano y propio de los monitores, por lo que es el utilizado en esta investigación. Otro modelo muy usado es el CYMK (Cian, Amarillo, Magenta, Negro) también llamado de pigmentación, modelo propio de las impresoras. 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. En una imagen RGB, cada píxel tiene información de los tres planos, estando la imagen formada por tres matrices distintas que recogen el valor de la intensidad de los píxeles para cada uno de los planos de color R, G y B.

La Visión Artificial está presente en todas las fases de la AP, sobre todo para la localización de las malas hierbas, esto requiere la recopilación de datos y su posterior análisis, en lo que intervienen los sensores fijos o móviles que pueden medir las características espectrales y espaciales de las plantas en tiempo real. Éstos podrían basarse

en sistemas aerotransportados de visión a bordo de satélites (Martín y col., 2011; De Castro y col., 2013), aviones tripulados (Peña-Barragán y col., 2011; De Castro y col., 2012), vehículos aéreos no tripulados (Torres-Sánchez y col., 2013, 2014), tractores robotizados (Meyer y Camargo-Neto, 2008; Guijarro y col., 2011; Romeo y col., 2013), o conectados a otros tipos de robots móviles. El objetivo de estos sensores es registrar información que permita estimar las densidades de las malas hierbas, la altura del cultivo, la reflectancia, el estado de humedad y otras propiedades importantes, para de esta forma tomar decisiones respecto a los tratamientos selectivos propios de la AP. En el caso de las redes fijas de biovigilancia el objetivo de los sensores es ser robustos y, en la medida de lo posible, abastecerse de fuentes de energía renovables, y deben ser capaces de transmitir información a través de redes inalámbricas para envío de la información remotamente (O'Shaughnessy y Evett, 2010; Wang y col., 2006) a las computadoras que pueden realizar los procedimientos necesarios para el tratamiento de la información.

Thorp y Tian (2004), informan sobre la captura de imágenes de tres formas distintas, mediante dispositivos a bordo de aeronaves, a veces tripuladas como no tripuladas (Pajares, 2015), o de satélites (Gómez-Candón y col., 2012; Meng y col., 2013) y desde vehículos en tierra. Los dos primeros se sitúan en el ámbito de lo que se conoce como teledetección. El problema de los sistemas de teledetección es a veces la baja resolución espacial, que para el propósito que nos ocupa no resultan apropiadas por el nivel de detalle que se requiere para la identificación de las texturas, además de su elevado coste económico. Esta es la razón por la que tales sistemas se están sustituyendo por vehículos aéreos remotamente tripulados (Remotely Piloted Aircrafts, RPAs) o incluso no tripulados (UAVs) que pueden proporcionar imágenes con el nivel de detalle espectral y espacial requerido (Berni y col., 2009a; Laliberte, 2010; Göktogan y col., 2010).

Torres-Sánchez y col., (2013), matizan la importancia del tamaño del píxel en la imagen, el área cubierta por cada imagen y el momento del vuelo, siendo estas características muy sensibles en relación a la altitud del vuelo. En una altitud baja, el UAV capturó imágenes de resolución espacial más fina, esto incrementó el número de imágenes necesarias para cubrir todo el campo, lo que puede ser un factor limitante debido a la energía requerida para una mayor duración del vuelo y los requisitos computacionales para el proceso de realizar los mosaicos correspondientes con las imágenes capturadas. El problema de estos vehículos es principalmente su reducido tamaño, que impiden su equipamiento con

sistemas de actuación, siendo un tanto inestables e inoperativos en condiciones meteorológicas de viento adversas. No obstante, Göktoğan y col. (2010) utilizan un UAV para localizar las malas hierbas en lugares inaccesibles, equipado con un sensor de visión, algoritmos de detección de malas hierbas y un mecanismo de pulverización. Una de las ventajas que tiene este sistema es su bajo coste frente a sistemas terrestres.

En los últimos años ha proliferado el uso de vehículos terrestres robotizados, que favorecen la captura de las imágenes desde una altura inferior a los tres metros con fácil instalación de los sistemas de visión en mástiles (Motohka y col., 2010; Sakamoto y col., 2011; Yu y col. y col., 2013), o acopladas en tractores o robots agrícolas (Montalvo y col., 2012, 2013; Tellaache y col. 2008a-b; Burgos-Artizzu y col. 2009, 2010; Guerrero y col., 2012) o manualmente (Thorp y Tian, 2004). Estas imágenes junto a los métodos proporcionados por la Visión por Computador constituyen una buena herramienta para la identificación de texturas por diferenciación espectral o espacial (cultivo y malas hierbas).

2.3. Métodos de segmentación por la respuesta espectral

La segmentación de imágenes trata de descomponer una imagen en las partes que la constituyen, es decir, los objetos de interés y el fondo, esto se realiza a partir de unas características locales. La mayoría de las imágenes están constituidas por regiones o zonas que tienen características homogéneas (textura, nivel de gris, etc.). La umbralización es el método más simple para la segmentación de una imagen. A partir de una imagen en escala de grises se puede utilizar la umbralización para crear imágenes binarias (Shapiro y Stockman, 2001). Un umbral sencillo basado en el valor medio del nivel de gris de la imagen (histograma) se ha propuesto en Gée y col. (2008), tras su aplicación, la parte vegetal (cultivo y mala hierba) aparece en blanco, y el resto (superficie del suelo, piedras, sombras) en negro.

El objetivo de la segmentación es cambiar y/o simplificar la representación de la imagen de forma que la nueva imagen sea más significativa y resulte más fácil de analizar. Se han venido utilizando cámaras para obtener las imágenes de las zonas de cultivo de grandes dimensiones (Davies y col., 1998). En nuestra investigación, las imágenes serán tomadas por un robot que se desplaza y actúa en los cultivos, donde una parte de la

información es suministrada por un sistema de visión. De las imágenes principalmente nos interesará principalmente extraer la vegetación, con el fin de localizar las malas hierbas para un tratamiento selectivo. Los siguientes trabajos cuyo objetivo es la identificación de las plantas verdes (Burgos-Artizzu y col., 2009; Guerrero y col., 2012; Guijarro y col., 2011; Montalvo y col., 2012; Onyango y Marchant, 2003; Tellaeché y col., 2008 o López-Granados, 2011) tienen un primer paso común, la separación de la capa vegetal (cultivo y malas hierbas) con respecto a los otros elementos que aparecen en la imagen (tierra y cielo). Existen métodos de segmentación de imágenes muy eficientes en diferentes aplicaciones (Oliva, 2015), pero estos métodos dejan mucho que desear en las imágenes de cultivos agrícolas.

Existen diversas estrategias que tiene como objetivo segmentar la capa vegetal en las imágenes de cultivos, que pueden agruparse en las siguientes categorías:

1. Basadas en índices cromáticos (relativos al espectro visible).
2. Basadas en una umbralización específica.
3. Basadas en aprendizaje.

2.3.1. Índices de vegetación (Índices cromáticos)

Los índices de vegetación (VI, vegetation index) son medidas empíricas obtenidas al cuantificar y combinar dos o más bandas del espectro electromagnético, normalmente el infrarrojo cercano y la luz visible, que reflejan la vegetación. Con los VI se trata de obtener datos con los que caracterizar el tipo, la cantidad, el estado y la situación de la vegetación en el escenario que se quiere analizar. Dentro de estos índices es posible hacer una diferenciación, por un lado están los que emplean el espectro infrarrojo cercano (NIR) y por otro los que hacen uso únicamente del espectro visible.

No es muy frecuente que las cámaras tengan uno o varios canales destinados a medir alguna de las bandas del espectro que ocupa el NIR, aunque sin duda existen cámaras multispectrales e hiperspectrales que utilizan el NIR con tal propósito (Jurado-Expósito y col., 2003; Peña y col., 2013; Castro y col., 2012). Normalmente, las cámaras digitales están equipadas con un filtro de corte de infrarrojos, aunque se ha llegado a proponer modificar

físicamente la cámara para suprimirlo y equiparla con otros que dejan pasar una amplia gama de radiaciones a partir del rojo (Rabatel y col., 2014).

Por lo general, los datos de los índices en los que se utiliza alguna banda del espectro infrarrojo suelen obtenerse a partir de datos capturados por plataformas remotas, en aplicaciones de teledetección, donde estas imágenes tienen naturaleza multispectral o hiperespectral. NIR es menos intuitivo para el humano, ya que el ojo humano no es sensible a la región NIR. Siendo sólo capaz de discernir plantas a partir del verdor de imágenes en color.

El Índice de Vegetación de Diferencia Normalizada (NDVI), se utiliza habitualmente para determinar cubiertas vegetales, medir el crecimiento de las plantas y controlar la producción de biomasa, fue introducido por Rouse y col. (1973) basándose en los estudios previos de Kriegler y col. (1969). Se define mediante la siguiente fórmula, $NDVI = (NIR - R) / (NIR + R)$, que utiliza la componente espectral R y el NIR. El NDVI ha sido utilizado en diversos trabajos (Zarco-Tejada y col., 2005; Zhang y col., 2006; Perry y Davenport, 2007; Hasegawa y col., 2010; Zarate-Valdez y col., 2012) y se ha demostrado que está muy relacionado con el índice de área foliar y la fracción ocupada por la cobertura vegetal (Carlson y Ripley, 1997). Torres-Sánchez y col. (2013) analizaron las diferencias espectrales entre malezas, cultivos y suelo con los índices de vegetación Exceso de Verde (ExG), Diferencia Normalizada Verde-Rojo (NDGRI) y Vegetación de Diferencia Normalizada (NDVI), a unos 30 m de altitud, obtuvieron una mayor separabilidad espectral entre la vegetación y el suelo con el índice NDVI.

En la literatura se han definido otros muchos índices, algunos derivados del NDVI, que para detectar la vegetación o alguna de sus propiedades utilizan el NIR. El NIR además de detectar la vegetación, a veces se utiliza para diferenciar en la vegetación los distintos tipos de plantas, e incluso para clasificar las plantas según su estado de salud. La vegetación absorbe la mayor parte de la luz visible que le llega y refleja la mayor parte de la luz infrarroja, si posee buena salud; por el contrario si está enferma la vegetación refleja más luz visible y menos infrarroja. Estos índices son: *Perpendicular Vegetation Index* (PVI, Richardson y Wiegand, 1977), *Difference Vegetation Index* (DVI, Tucker, 1979), *Weighted Difference Vegetation Index* (WDVI, Clevers, 1988), *Percentage Vegetation Index* (PVI, Crippen, 1990), *Green Normalized Difference Vegetation Index* (GNDVI, Buschmann y Nagel, 1993), CASI NDVI (Cloutis y col., 1996), NG (Normalized Green), NR (Normalized

Red), NNIR (Normalized Near Infrared), Green-Red NDVI (GRNDVI), *Green Ratio Vegetation Index* (GRVI, Sripada y col., 2006), *Green Difference Vegetation Index* (GDVI, Sripada y col., 2006) y *Green-Blue NDVI* (GBNDVI, Wang y col., 2007), entre otros.

Muchas cámaras capturan imágenes en el espectro visible usando el espacio de color RGB (Pajares y Cruz, 2007; Gonzalez y Woods, 2007), con un canal para el espectro que ocupa el color verde. Una de las formas más sencillas de detectar la vegetación es analizando solamente los verdes presentes en la imagen sin necesidad de considerar los canales rojo y azul del espectro visible. Yang y col. (2003), consideran vegetación todo píxel cuya intensidad en el plano el color verde es mayor que un cierto umbral. Lo normal es combinar más de un canal de los tres del modelo RGB. En Aitkenhead y col. (2003) se usa la proporción de verde frente a los tres canales $G/(R + G + B)$, combinado con el valor de un cierto umbral determinado mediante ensayo y error.

Un índice utilizado en muchos trabajos existentes en la literatura, con buenos resultados, es el exceso de verde (ExG, Woebbecke y col., 1995) que sirve para la discriminación de plantas bien como índice base o bien para su comparación con otros índices (Meyer y col. 1998; Ribeiro y col. 2005; Tellaeche y col., 2008; Meyer y Camargo-Neto, 2008; Montalvo y col., 2012; Ponti, 2013). Torres-Sánchez y col. (2013, 2014), para discriminar entre vegetación y suelo desnudo utilizaron ExG, obteniendo resultados satisfactorios a pesar de que las imágenes no tenían gran resolución espacial al ser tomadas por un vehículo no tripulado volando a distintas alturas. Soontranon y col., (2014) desarrollaron un programa para la detección de los estados de crecimiento del arroz, en el que para hacer el seguimiento y el análisis del crecimiento de la planta utilizaron varios índices de vegetación, inicialmente compararon tres índices ExG, NGRDI y ExG-ExR, para concluir que el más eficiente es el ExG. El índice cromático se obtiene mediante una combinación lineal de los planos de color RGB con constantes $r = -1$, $g = 2$, $b = -1$.

La segmentación basada en ExG es aceptada generalmente como uno de los métodos más comunes y eficaces, a menudo proporciona resultados de segmentación pobres, especialmente cuando la escena de la imagen contiene una diferencia de iluminación extrema causada por sombras oscuras o excesiva iluminación. Para solucionar esto a veces se ha utilizado el exceso de rojo ExR, (Meyer y col, 1998) en el reconocimiento de los falsos

positivos. Shamugam y col., 2015, para separar las plantas del fondo utiliza ExG-ExR, sabiendo que en el fondo apacere un predominio de la componente espectral roja.

En Neto (2004) se combinan estos dos índices (ExG y ExR) para formar un nuevo índice que también obtiene buenos resultados, son el exceso de verde menos exceso de rojo (ExGR), que define como vegetación aquellos píxeles cuya diferencia entre la componente verde y la componente roja (ExG-ExR) es suficientemente grande. En Meyer y Camargo-Neto (2008), podemos encontrar una comparativa entre los índices ExG, NDI y ExGR. Gitelson y col. (2002) estudiaron las reflectancias del color verde, rojo y azul, junto con el NIR, proponiendo en este caso el índice *VARIg*, muy semejante al NDI, pero añadiendo a la suma del denominador el canal azul: $(G-R)/(G+R+B)$. En Hague y col. (2006) se propone un nuevo índice que aprovecha la potencia de la combinación de canales, basándose en el trabajo previo de Marchant y Onyango (2002). Este índice se creó para que fuese invariante a las condiciones lumínicas, además determinan el valor de un parámetro a , que interviene en el cómputo, para el cual el índice es invariable a la temperatura de color que emitiría un cuerpo negro. El índice de Hague y col. (2006) se define como $VEG=G/(R^a B^{1-a})$ con $a = 0,667$ fijado en sus experimentos para cumplir con las propiedades descritas previamente.

En Ribeiro y col. (2005) se propone la combinación lineal de los tres canales, con constantes $r = -0,7$, $g = 0,588$, $b = 0,136$. Kataoka y col., (2003) proponen un nuevo índice, denominado CIVE, como combinación lineal de los tres canales y añaden una constante a la suma, $CIVE=0.441r-0.811g+0.385b+18.78745$. En esta línea se sitúa el trabajo de Bossu y col (2008), en una primera etapa separa la planta a partir del índice de color desarrollado por Søggaard y Olsen (1999) definido como $In = 2 \times G-R-B$. Laursen y col. (2014), realizan un estudio que tiene como objetivo combinar características de color existentes en campos agrícolas con un esquema de clasificación simple para lograr una mejor segmentación que cualquier función que utiliza un único canal de color, para esto utilizan NDVI, ExG, ExR y un exceso de verde modificado.

2.3.2. Umbralización

La umbralización es una técnica de segmentación que trata de definir un umbral, de forma que separe los objetos de interés respecto del fondo. Los índices revisados en el punto anterior transforman una imagen de color a escala de grises, por lo se necesita un umbral a

partir del cual poder decidir si un píxel es considerado como planta o como cualquier otro material, de esta forma podemos conseguir la imagen binaria donde los píxeles que identifican la vegetación se etiquetarán con blancos y el resto con negro. En general, estas técnicas suponen la solución de un problema de segmentación de dos clases. Los enfoques basados en umbrales, incluyen también lo que se conoce como umbral dinámico, que considera regiones de vecindad en lugar de la imagen al completo.

Considerando el problema bi-clase, y asumiendo que las dos clases a identificar en la imagen siguen una distribución Gaussiana, Reid y Searcy (1987) propusieron un método para determinar una función de decisión. Dentro de los métodos de umbralización, el de Otsu (1979) es uno de los más utilizados, tiene una buena respuesta frente a la mayoría de las situaciones del mundo real y no necesita supervisión, pertenece a los considerados como de naturaleza biclase (Ling y Ruzhitsky, 1996; Shrestha y col, 2004).

Los algoritmos de umbralización se aplican generalmente a imágenes en escala de grises, obteniendo un umbral a partir del cual se binariza la imagen en las dos clases existentes. Gebhardt y col. (2007) transformaron una imagen del modelo de color RGB a una imagen en escala de grises y realizaron una umbralización para segmentar una imagen de color, para ello, utilizaron un algoritmo que fue posteriormente mejorado mediante operaciones morfológicas además de utilizar propiedades de homogeneidad local (Gebhardt y Kaühbauch, 2007).

Kirk y col. (2009) aplicaron una combinación de verdes e intensidades derivadas de las bandas espectrales verde y roja, calculando un umbral automático para un problema biclase, vegetación y terreno, asumiendo dos funciones de densidad de probabilidad Gaussianas asociadas a cada una de las clases. Meyer y Camargo-Neto (2008) han empleado el método de umbralización automático de Otsu para binarizar una imagen obtenida mediante el índice de vegetación ExG y NDI, estableciendo una comparativa con la segmentación obtenida mediante el índice de vegetación ExGR, usando como valor para umbralizar el cero, de este modo no existe dependencia con respecto al método de Otsu, ni con ningún otro método para determinar el umbral. Burgos-Artizzu y col. (2011) proponen un método de segmentación automático, que resulta ser clave para conseguir la deseada robustez del sistema en las condiciones de iluminación cambiante en entornos agrícolas, utilizando inicialmente los índices de color para crear una imagen de gris que luego

binarizan fácilmente usando un método de ajuste de umbral automático, simple y rápido, considerando el valor medio estadístico de los niveles de intensidad de los píxeles, resultando ser un método de segmentación robusto.

En Montalvo y col. (2013) se plantea una doble umbralización mediante Otsu para identificar lo que denominan plantas enmascaradas (impregnadas con tierra, que han perdido el verdor) y sin enmascarar. Según Sezgin y Sankur (2004) el método de Otsu obtiene resultados satisfactorios cuando el número de píxeles en cada lóbulo del histograma es similar, si bien no trabaja apropiadamente en otras imágenes, ya que tiende a la infra-segmentación, siendo en estos casos a veces mejor utilizar la media estadística.

En los experimentos descritos en la presente memoria se ha comprobado que el tipo de sensor utilizado en la captura de las imágenes es determinante a la hora de aplicar la media u Otsu, por lo que en los experimentos llevados a cabo durante esta investigación se ha utilizado uno u otro en función del sensor, tal y como se describe posteriormente. Conviene añadir en este sentido, que el método propuesto en Guijarro y col. (2015), basado en la adición de información contextual mediante wavelets, compensa los problemas derivados de la posible infra-segmentación que produce Otsu, por lo que en este caso el uso de Otsu resulta apropiado.

2.3.3. Segmentación basada en aprendizaje

En una imagen los colores tienden a agruparse formando clases que determinan la distinta naturaleza de las texturas. A veces sucede que la distinción espectral no resulta tan clara como cabría esperar, por lo que existen fronteras difusas en lo que respecta a la separación de las clases. No obstante, existen mecanismos suficientemente potentes como para hacer frente a esta situación, tal es el caso de la aplicación de métodos borrosos o fuzzy, que tienen en cuenta el grado de pertenencia de un píxel a distintas clases, lo cual resulta altamente favorable respecto de la problemática considerada. Como se ha mencionado previamente, esto ocurre en las zonas fronterizas de las texturas, donde se producen transiciones graduales en los valores de los píxeles de una zona a otra. A esta categoría fuzzy pertenece el método de segmentación agrupamiento borroso o “*Fuzzy Clustering*”. En la fase de entrenamiento utiliza un método de optimización con el que se consiguen agrupaciones compactas con el fin de resolver así muchas de las ambigüedades existentes.

En esta línea están los trabajos de Pajares y Cruz (2002), Pajares y col. (2002) o Lim y Lee (1990) entre otros.

Mohaddamzadeh y Bourbakis (1997) identifican las agrupaciones de píxeles que poseen similar color como las regiones a segmentar, observando que los cambios de región se manifiestan gradualmente como se ha explicado previamente. Pajares y Cruz (2007), formalizan esta idea definiendo una función de naturaleza fuzzy en el espacio de color RGB. Cheng y col. (2001) por una parte utilizan el concepto de “homogram” para extraer en las imágenes a color regiones homogéneas; y por otra parte definen el concepto de homogeneidad *fuzzy* para de esta forma extraer regiones de similares características mediante la obtención de una serie de umbrales. Tras la segmentación vuelven a aplicar un criterio de similitud para fusionar regiones, este proceso es similar al método de segmentación de regiones según la técnica de árboles cuaternarios (Pajares y Cruz, 2007). En el ámbito de las imágenes de satélite los siguientes trabajos utilizan métodos de clasificación basados en la lógica fuzzy: Seong y Usery (2001), Zhang y Foody (1998).

También se han empleado otras técnicas de aprendizaje no supervisado, incluyendo obviamente la denominada *Fuzzy Clustering*, para segmentar regiones de interés en imágenes en escala de grises, obtenidas a partir de imágenes de color mediante los índices de vegetación *ExR* y *ExG* (Meyer y col., 2004). Romeo y col. (2013) propusieron un algoritmo en dos fases, en la primera se aplican los métodos clásicos de identificación de verdes y en la segunda se utiliza Fuzzy Clustering. También Tian y Slaughter (1998) propusieron un algoritmo de segmentación adaptativo medioambiental (Environmentally Adaptive Segmentation Algorithm, EASA) para detectar plantas mediante un proceso de aprendizaje supervisado. Ruiz-Ruiz y col. (2009) emplearon el algoritmo EASA, utilizando como imagen de partida una transformada al modelo de espacio de color HSI (Hue-Saturation-Intensity) que tiene la ventaja de separar el color de la intensidad, informando sobre la mayor robustez frente a los cambios de iluminación. Zheng y col. (2009, 2010) utilizan un algoritmo de desplazamiento de la media (*mean-shift*) no supervisado basándose en que el problema de la segmentación de la vegetación respecto del suelo o fondo puede ser tratado como un problema de segmentación biclase; la separación en clases fue validada mediante una red neuronal y el discriminante lineal de Fisher, habiendo empleado los espacios de color RGB, LUV y HSI.

Como podemos observar, la mayoría de los métodos aplican enfoques basados en las características espectrales de los píxeles, ya que resulta difícil la identificación de texturas en imágenes agrícolas de exterior, que son con las que trabajamos sobre ellas, influye muy directamente la variabilidad meteorológica como factor determinante impidiendo que los métodos basados en regiones trabajen correctamente en la mayoría de los casos. Estos métodos se centran en la identificación del verde bajo el supuesto de que la componente espectral verde predomina sobre la componente espectral roja.

Recientemente, el método conocido como Máquinas de Vectores Soporte (*Support Vector Machines, SVM*) se está convirtiendo en un método popular para calcular el hiperplano de separación para aplicaciones agrícolas (Mitra y col., 2004 y Nashat y col., 2011). A diferencia del análisis discriminante lineal, las SVM identifican lo que se conoce como vectores soporte, que son puntos con la máxima información disponible situados en las fronteras de las clases, justamente donde aparecen los problemas de pertenencia de los píxeles a una u otra clase. Estos vectores soporte permiten estimar una función de decisión, que separa las clases de forma eficiente. Además poseen lo que se denominan núcleos que permiten proyectar espacios de una determinada dimensión a otros de mayor dimensión, consiguiendo de este modo un margen de separación entre las clases mayor. Debido a estas posibilidades, las SVM aparecen como métodos robustos y fiables. En Guerrero y col. (2012) se utilizan para clasificar la imagen en dos conjuntos distintos, el que contiene las plantas y el que contiene el resto de elementos de la imagen del cultivo. Siguiendo con la misma línea de segmentación biclase, Tellaeche y col. (2011) tras un proceso de segmentación previa, mediante umbralización basada en la técnica 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; se trata de clasificadores biclase con técnicas basadas en la teoría de Bayes, la Teoría de la Decisión Multicriterio y las SVM.

2.4. Métodos de segmentación por la respuesta espacial mediante Wavelets

Las wavelets constituyen una potente herramienta para afrontar problemas significativos en el tratamiento de imágenes, donde destacan temas relacionados con el énfasis en altas frecuencias o suavizado en las bajas. Esta teoría ha experimentado un gran

desarrollo en las dos últimas décadas mostrándose muy eficiente donde otras técnicas, como por ejemplo la transformada rápida de Fourier, no resultaban satisfactorias por trabajar en dominios transformados

La transformada wavelets divide la imagen original en imágenes de menor resolución espacial, obteniendo una que viene a ser una copia de la imagen original a menor resolución suavizada mediante operaciones de filtrado paso bajo y otras cuatro más que contienen información relativa a los cambios locales en la imagen original y por tanto a las altas frecuencias mediante operaciones de filtrado paso alto. El primer resultado se conoce como coeficientes de aproximación, mientras que los otros tres son los denominados coeficientes de detalle, o en otras palabras coeficientes verticales, horizontales y diagonales. El procedimiento base consiste en aplicar filtrados de tipo paso bajo y alto, convenientemente combinados mediante filtros, cuyos coeficientes se obtienen a partir de lo que se conoce como familias de wavelets (Haar, Daubechies, Coiflets,...). La descomposición de la imagen en los mencionados componentes constituye uno de los asuntos clave para su aplicación en la combinación de información en las imágenes. Algo que forma parte del proceso seguido en Guijarro y col. (2015) para mejorar la segmentación de las imágenes de campos de cultivo a través de las texturas de las partes vegetales, como se explicará posteriormente en el capítulo cuatro.

Pajares y Cruz (2004) realizan un trabajo de recopilación sobre la fusión de imágenes basadas en wavelets, en él establecen tres métodos de fusión como puntos de referencia basados en los trabajos de Li y col. (1995), Zhang y Blum (1999) y Ranchin y Wald (2000). En general, el problema que la fusión de imágenes trata de resolver es la combinación de información procedente de varias imágenes (sensores) tomadas de la misma escena con el fin de lograr una nueva imagen fusionada, que contiene la mejor información extraída de las imágenes originales. Por lo tanto, la imagen fusionada posee mejor calidad que cualquiera de las imágenes originales. El enfoque basado en wavelets es apropiado para la realización de tareas de combinación por las siguientes razones:

1. Es una escala múltiple (multiresolución) muy adecuada para gestionar las diferentes resoluciones de una imagen.

2. La *transformada wavelets discreta* (DWT), conservando la información de la imagen original, permite la descomposición de la imagen en diferentes tipos de coeficientes.
3. Estos coeficientes que proceden de diferentes imágenes pueden combinarse para obtener nuevos coeficientes, de forma que la información de las imágenes originales se recopila adecuadamente.
4. Fusionados los coeficientes, la imagen fusionada final se obtienen a través de la *transformada wavelet discreta inversa* (IDWT), en esta imagen también se conserva la información de los coeficientes fusionados. Si bien este último aspecto no se aplica en el procedimiento de fusión diseñado para el realzado de las texturas vegetales en las imágenes a procesar procedentes de los campos de cultivo.

Li y col. (1995), presentan un esquema de fusión de imágenes que se basa en la transformada wavelet, combinan adecuadamente la transformada wavelets de las imágenes de entrada de forma que la nueva imagen la obtienen tomando la transformada wavelet inversa de los coeficientes wavelet fusionados. Realizan varios experimentos con diversos conjuntos de imágenes, entre ellas imágenes multifoco, el objetivo de la fusión consiste en lograr una imagen de alta calidad a partir de varias imágenes degradadas. Zhang y Blum (1999) establecen una descomposición multiescala y diferentes reglas de fusión, incluyendo los métodos existentes en la literatura que no asumen un modelo estadístico para las imágenes origen. En Ranchin y Wald (2000) introducen el concepto ARSIS (Amélioration de la Résolution Spatiale par Injection de Structures) para describir y modelar la información que está ausente en las imágenes que van a ser fusionadas, utilizan un método multiescala. En algunos casos se combina la información mediante diversas estrategias, como ocurre en el método combinado descrito en Zhang y col. (2004), que proponen un procedimiento multi-resolución mediante la descomposición de la imagen original en bandas de frecuencia a través de la transformada de wavelets, llegando a un cierto nivel de descomposición que es donde se utilizan descriptores basados en las matrices de co-ocurrencia (Pajares y Cruz, 2007). En Chen y col. (2005) también se utiliza una estrategia multi-resolución para la descomposición de la imagen en diferentes bandas de frecuencia.

En las imágenes de satélite que se fusionan quieren conseguir una alta resolución espacial junto con un alto contenido espectral de baja resolución, es decir:

- 1) Imágenes con alta calidad en el contenido espectral pero baja calidad en la resolución espacial.
- 2) Imágenes con alta resolución espacial pero con una única banda espectral. Este es también el objetivo en Garguet-Duport y col. (1996).

La aplicación de técnicas de descomposición mediante wavelets, aplicada a imágenes agrícolas, se puede estudiar en Chou y col. (2007), Bossu y col. (2007, 2008) e Ishak y col. (2009), los cuales utilizaron más de 30 familias de wavelets y el filtro de Gabor para diferenciar automáticamente entre cultivo y maleza. Este filtro se caracteriza por realizar una descomposición de la imagen en distintas frecuencias, similar a la descomposición de wavelets. De las 30 familias de wavelets que utilizaron en sus trabajos, se seleccionaron exactamente dos como las más representativas, concretamente la familia que proporcionaba los mejores resultados y a la vez la que generaba los peores, en ambos casos se compararon con el filtro Gabor. Del trabajo anterior concluyeron que todas las familias de wavelets proporcionaron mejores resultados que el filtro de Gabor, y respecto a las 30 familias de wavelets estudiadas, las que obtuvieron los mejores resultados fueron las familias de Daubechies y Meyer, siendo la de Daubechies la que ofrecía la mejor relación entre tiempo de procesamiento y precisión en la clasificación. La familia con la que peores resultados se obtenían resultó ser la biortogonal. Aparte de imágenes de cultivos reales, Bossu y col. (2008) crearon imágenes sintéticas con tres tipos diferentes de distribución espacial de malas hierbas.

Ishak y col. (2009) aplicaron una combinación de la descomposición basada en wavelets y Gabor utilizando una distribución de gradiente de campo. Todas las imágenes fueron tomadas bajo las mismas condiciones de iluminación, y optaron por clasificar las malas hierbas como planta de hoja estrecha, mientras que el cultivo fue identificado como textura vegetal de hoja ancha. Con esta categorización generaron un algoritmo de entrenamiento. Los resultados indicaron una exactitud de diferenciación de un 94%. Aproximadamente el mismo porcentaje de acierto (94.63%) consiguieron Chou y col. (2007), en el que combinaron la transformada de wavelets basada en la familia Daubechies con la distancia bayesiana ponderada; se proporcionan las características de las distintas clases de hojas existentes en las texturas vegetales.

Según la literatura existente podemos concluir que entre las wavelets más populares se encuentra la propuesta por Daubechies (1992), que es la función más compacta para el número de orden utilizado. Estas funciones resultan óptimas para un orden lo suficientemente alto ($n > 15$), si bien no son simétricas. Otra familia interesante es la que se define en Meyer (1990), con una construcción basada en el espacio de Fourier. Esta función puede ser aproximada correctamente con una versión truncada según Bossu y col. (2008), aunque no resulta ser compacta. Las familias de wavelets de orden superior suelen ser generalmente funciones suaves, sin picos fuertemente acentuados, y suelen proporcionar buena localización tanto en el espacio como en la frecuencia.

En lo que respecta al uso de wavelets para resaltar las texturas procedentes de las plantas (cultivo y malas hierbas), según el trabajo propuesto en Guijarro y col. (2015) y descrito en el capítulo cuatro, se ha aplicado la transformada de Haar, que resulta ser el filtro más elemental (más bajo número de coeficientes) de la familia de Daubechies, habiendo comprobado su suficiencia en esta aplicación frente a otros filtros de orden superior de esta misma familia y de otras familias.

2.5. Conclusión general

La mayor o menor precisión en la detección del verdor, objetivo principal del presente trabajo, dependerá de las decisiones que se tomen en cuanto a la captura de imágenes y su posterior procesamiento. Así pues, llegados a este punto del análisis de nuestro problema, podemos señalar que en la búsqueda de la solución existen diversas estrategias orientadas a resolverlo a partir de las cuales se han diseñado las dos propuestas que se formulan en esta investigación, y que suponen una mejora sobre las existentes y bajo la consideración de los siguientes factores fundamentales:

1. Captura de la imagen, que se realice con un tipo de sensor con una resolución espectral suficiente. Como quiera que en el caso que nos ocupa las cámaras operan en el espectro visible, éstas deben garantizar respuestas espectrales y espaciales con las resoluciones apropiadas.
2. Índice de vegetación utilizado en la discriminación entre vegetación y suelo desnudo, éste es justamente el punto que acabamos de revisar previamente en el

estudio bibliográfico realizado. Estudiada la amplia variedad de índices que la literatura propone y teniendo en cuenta el tipo de imágenes de campos agrícolas tratadas, la investigación se ha orientado en dos sentidos. Por un lado, hacia el diseño de una estrategia que aplica índices simples y combinados (ExG, ExGR, CIVE y VEG) para identificar partes relevantes en la imagen, separando cielo, suelo y partes verdes y aún para separar las partes verdes y de suelo mediante un método automático de clasificación basado en el conocido método de clasificación borrosa. Por otro lado, para incluir información espacial mediante una estrategia de fusión de la información con wavelets, cuya finalidad es mejorar los resultados de la segmentación proporcionados por el índice de vegetación ExG.

3. Método de cálculo para encontrar el valor óptimo del umbral para la separación entre clases (suelo-vegetación) a partir de los resultados proporcionados por los métodos de segmentación, paso importante para realizar la identificación final.

3

Fundamentos teóricos

En este capítulo se describen los fundamentos teóricos relativos a las técnicas que se han utilizado en las distintas fases de las que se compone la investigación, cuyo objetivo es el desarrollo de métodos automáticos para la detección de las distintas texturas de una imagen de un cultivo de cereal o maíz en entornos de exterior agrícolas. Estos métodos mejoran los resultados obtenidos por otros métodos existentes en la literatura, tal y como se ha indicado previamente.

La tarea resulta compleja por tratarse de imágenes reales tomadas en entornos de exterior, donde las condiciones de iluminación pueden variar de forma repentina e incluso a lo largo del día, en ocasiones debido a la variabilidad en la nubosidad. Estos cambios de iluminación, pueden crear sombras o reflejos, y producir pérdidas importantes de información. A modo ilustrativo, supongamos por ejemplo que la imagen se toma en un día en condiciones de iluminación normal, ésta puede caracterizarse por poseer valores de intensidad dentro del rango aceptable sin llegar a la saturación, conteniendo partes brillantes, zonas con alto contraste y probablemente con un fuerte contraste verde. Por el contrario si la imagen se toma un día con iluminación escasa, ésta tendrá zonas saturadas o próximas a la saturación, principalmente en las zonas más reflectivas, como puede ser el suelo o cielo.

La iluminación es un factor muy importante a la hora de tomar las imágenes y afecta directamente en lo que respecta a la identificación de las texturas naturales en las imágenes, hecho que ya se ha considerado en la literatura (Zhang y col.1995; El-Faki y col. 1997, 2000).

Los procesos que sustentan la investigación se concretan en los siguientes:

- 1) *Captura de imágenes*. Las imágenes se obtienen a través de los diferentes sensores y bajo distintos modelos de color, en el rango de longitudes de onda correspondientes a la porción del espectro visible, que es lo que se aplica en la presente investigación. Se proporciona una revisión general sobre este asunto.
- 2) *Índices de vegetación*. Constituyen un elemento esencial del proceso, ya que gracias a ellos es posible distinguir las partes vegetales con firmas espectrales verdes frente a las que no lo son. Incluso es posible modificar dichos índices para identificar otros tipos de firmas espectrales involucradas en las imágenes agrícolas, entre las que destacan texturas de suelo o cielo. El resultado es una imagen de grises para su posterior umbralización.
- 3) *Umbralización*. Relativo al proceso de discernir entre las texturas (incluyendo las plantas verdes) generadas por los índices de vegetación anteriores.
- 4) *Clasificación de texturas*: cuya finalidad es la discriminación fina de diferentes tipos de firmas espectrales verdes, lo que permite discernir entre plantas pertenecientes a cultivos y malas hierbas.
- 5) *Transformada de wavelets*. Permite descomponer las imágenes en coeficientes de detalle, para identificar las texturas por su variabilidad espacial, en contraposición a la discriminación por diferencias espectrales.

A continuación se proporcionan los detalles teóricos suficientes para abordar los procesos anteriormente mencionados.

3.1. Imágenes

Una imagen es una representación de la escena que manifiesta la apariencia visual de un objeto real en dicha escena. Esta imagen puede provenir de diferentes fuentes, las imágenes utilizadas en esta investigación han sido tomadas por cámaras digitales equipadas con sensores del tipo CCD. Para la formación de una imagen en color no basta sólo con el sensor, puesto que éste debe recibir el rango espectral adecuado, en nuestro caso en el rango visible. Esto significa que se requiere un sistema de filtros apropiado, de forma que cada píxel, en el sensor, reciba la radiación suficiente en función de las longitudes de onda en el espectro visible y por tanto según la naturaleza de dichas longitudes correspondientes al rojo, verde y azul. Resulta habitual colocar un filtro de infrarrojos capaz de bloquear las longitudes de onda correspondientes a este rango, ya que de lo contrario éstas afectan a la sensibilidad de los píxeles.

En definitiva, las imágenes digitales obtenidas son una representación bidimensional de una escena tridimensional, por ello, al final se dispone de un conjunto de tres matrices numéricas, una por cada canal de color cuyos valores son distintos según el modelo de color que se utilice en cada caso. A continuación se describen los modelos de color, junto con sus posibilidades de aplicación en el ámbito de las investigaciones realizadas para las imágenes agrícolas.

3.3.1. Modelos de color

Un modelo de color es una representación matemática abstracta que describe la forma en que se representan los colores como matrices de números, utilizando normalmente tres componentes de color. El espacio de color es el conjunto de colores obtenidos cuando un modelo se asocia a una descripción precisa de cómo se han de interpretar sus componentes. A continuación se describen distintas formas en las que puede ser representado el color en Visión Artificial.

RGB

El modelo de color RGB (*Red, Green, Blue*; Rojo, Verde, Azul) está basado en la síntesis aditiva, mediante la cual los colores primarios rojo (R), verde (G) y azul (B) se combinan de distintas formas para generar un conjunto de colores. Así, combinando estos

colores primarios en proporciones iguales o distintas se obtienen otros colores mezclados. Por ejemplo, combinando los tres colores primarios con los máximos valores de representación se obtiene el blanco puro. Este modelo de color es uno de los más utilizados en las cámaras a color. Un canal de color está representado por una matriz donde el valor de cada componente se calcula a partir de una combinación de valores digitales en bits, así en una representación de 8 bits por valor, el máximo valor de representación es $2^8 = 256$ y el mínimo es 0. El rango de valores, en este caso está comprendido entre 0 y 255, en el que también están los valores intermedios. Por las características del modelo y la disponibilidad de los índices de vegetación éste es el modelo que hemos utilizado en las investigaciones llevadas a cabo en este trabajo.

HSI

En el modelo HSI (*Hue, Saturation, Intensity*; Matiz, Saturación, Intensidad) los colores se distinguen unos de otros por su tono o matiz, saturación e intensidad. El tono hace referencia al color dominante, la saturación es la cantidad de luz blanca mezclada con el color dominante, de modo que diferencia un color pálido de uno intenso, y en la intensidad se recoge la iluminación percibida. Este modelo se suele utilizar cuando la iluminación de la escena es determinante, y se necesita trabajar con la intensidad de la imagen. Existe una relación de transformación entre los modelos RGB y HSI y viceversa (Pajares y Cruz, 2007).

YUV

Este modelo es el usado por el hardware de gráficos por ordenador, se acerca más a la percepción humana del color que el estándar RGB. Tiene en cuenta que la visión humana es más sensible a cambios en la intensidad de la luz que a los del color. El ojo humano no llega a apreciar el nivel razonable de error que admite. El parámetro Y representa la luminancia (información en blanco y negro), y los parámetros U y V representan la crominancia (información relacionada con el color). Permite enviar información de color independiente de la información de luminancia, este modelo se creó para garantizar el funcionamiento de los televisores en blanco y negro, cuando se codificaba la transmisión de información de los televisores a color. Dawson-Howe (2014) definió una relación de transformación, ya establecida de antemano, entre los modelos RGB y YUV y viceversa.

Y_D D_R

Basado en el modelo YUV, es utilizado en el estándar de televisión en color SECAM. Consta de una componente que representa la luminancia (Y) y dos que representan

la crominancia (D_B y D_R). Es útil para diferenciar en una imagen los píxeles en los que domina la componente roja. Se relaciona con el modelo RGB, los valores se suman para generar una señal simple de luminancia Y , y la señal D_B se obtiene restando Y de la señal azul del modelo RGB, y D_R restando Y de la señal roja, estando las dos multiplicadas por factores de escala diferentes.

3.2. Índices de vegetación: detección del verdor

Los índices de vegetación constituyen una buena herramienta para detectar y resaltar la vegetación existente en una zona en relación con otros elementos existentes en la misma, como por ejemplo el suelo o el cielo. Los índices de vegetación se determinan a partir de las medidas de reflectancia de dos o más longitudes de onda para analizar características específicas de la vegetación, una vez se han combinado de manera conveniente. En el caso de las imágenes procedentes de aplicaciones agrícolas permiten la obtención de la cobertura vegetal, la identificación de las plantas es clave en AP, esto constituye un paso necesario en la mayoría de las técnicas para realizar tratamientos específicos, el éxito del tratamiento depende de esa correcta identificación. En el presente trabajo, las investigaciones realizadas han tenido como objetivo diferenciar las distintas texturas encontradas en las imágenes agrícolas, especialmente en los verdes.

Algunas de las técnicas usadas en AP utilizan los índices para discriminar las capas vegetales, medir la producción de biomasa y la fortaleza de las plantas, estimando y evaluando el estado de salud de las plantas mediante la radiación que éstas emiten o reflejan, por ejemplo el NDVI (*Normalized Difference Vegetation Index*; índice de vegetación diferencial normalizado) propuesto por Rouse y col. (1973). El índice NDVI se utiliza básicamente para medir el crecimiento de las plantas, determinar cubiertas vegetales y controlar la producción de biomasa. Se calcula a partir de la luz visible e infrarroja cercana reflejada por la vegetación como sigue,

$$NDVI = \frac{IRC - R}{IRC + R} \quad (3.1)$$

Donde IRC es la reflectividad del infrarrojo más cercano y R es la reflectividad en el canal rojo. La vegetación sana absorbe la mayor parte de la luz visible que le llega, y refleja

una gran parte de la luz del infrarrojo cercano. La vegetación enfermiza o escasa refleja más luz visible y menos luz del infrarrojo cercano. En consecuencia, cuando una planta está vigorosa, refleja mucha radiación solar en el infrarrojo cercano y poca en el rojo, y en consecuencia, se obtiene un NDVI elevado. En cambio, cuando una planta está enferma, ocurre lo contrario. Por ejemplo, si una planta vigorosa refleja un 50% en el infrarrojo cercano y un 8% en el rojo, el NDVI será 0.72, según la fórmula anterior. En cambio, si una planta está enferma refleja un 40% en el infrarrojo cercano y un 20% en el rojo, el NDVI será 0.33. Por tanto, teniendo en cuenta que el NDVI siempre resulta en un número entre -1 y 1, debido a la forma en que se calcula, se podría decir que un $NDVI < 0$ se corresponde con cubiertas artificiales o zonas de agua, un NDVI entre 0 y 0.3 se corresponde con suelo sin cubierta, y un NDVI elevado se corresponde con zonas de vegetación.

El NDVI, es un buen estimador de la fracción de la radiación fotosintéticamente activa interceptada por la vegetación (fPAR) (Monteith, 1981), la productividad primaria (Paruelo y col., 1997; Tucker y col., 1985) y una variable integradora del funcionamiento del ecosistema (Virginia y Wall, 2001). Posee un gran valor en términos ecológicos, tiene la ventaja de ser fácil de calcular y de facilitar la interpretación directa de los parámetros biofísicos de la vegetación. En cambio, tiene el inconveniente de la poca capacidad para minimizar la influencia del suelo y de la atmósfera. Por esto, resulta ineficiente en el problema relativo a la identificación de texturas vegetales verdes, que se aborda como parte fundamental en el presente trabajo de investigación. Conviene reseñar a este respecto que para nosotros es muy importante poder diferenciar las plantas del suelo y otros elementos con el fin de identificar el cultivo y las malas hierbas con la máxima precisión posible. Otro problema es que se necesita el uso de cámaras multispectrales en lugar de cámaras RGB, esto encarece el sistema a desarrollar, incrementando los posibles costes de producción.

Recientemente se han realizado propuestas que permiten la obtención simultánea de los valores espectrales necesarios para calcular el índice NDVI a partir de la adecuada manipulación de los filtros de cámaras comerciales, removiendo el habitual filtro de bloqueo del infrarrojo y calibrando los tres canales espectrales restantes (Rabatel y col. 2014).

Además del NDVI existen otros índices que utilizan el espectro infrarrojo para realizar una estimación de la capa vegetal de una imagen, vienen dados por las ecuaciones (3.2) a (3.5), tabla 3.1. Aunque no se utilizan en la investigación propuesta, resulta

conveniente una revisión de los mismos por su funcionalidad, que coincide con la desarrollada en la presente investigación.

| | | |
|---|--|-------|
| <i>Environmental Vegetation Index</i> (Birth y McVey, 1968) | $EVI = \frac{IRC}{R}$ | (3.2) |
| <i>Enhanced Vegetation Index (MODIS)</i> (Huete y col., 2002) | $EVI = g \cdot \frac{IRC - R}{IRC + C_1 R - C_2 B + L}$ Con L=1, C ₁ =6, C ₂ =7.5 y g=2.5 | (3.3) |
| <i>Perpendicular Vegetation Index</i> (Richardson y Wiegand, 1977) | $PVI = \sqrt{(R_{suelo} - R_{veg})^2 + (IRC_{suelo} - IRC_{veg})^2}$ | (3.4) |
| <i>Soil-Adjusted Vegetation Index</i> (Huete, 1988) | $SAVI = (1 + L) \frac{(IRC - R)}{(IRC + R + L)}$ | (3.5) |

Tabla 3.1.- Varios índices utilizando el espectro infrarrojo

En el índice EVI (*Enhanced Vegetation Index*, característico de la plataforma satelital MODIS), el parámetro *L* realiza un ajuste de la capa vegetal que relaciona de forma no lineal el infrarrojo cercano con el espectro rojo que se transfiere a través de la capa vegetal, *C₁* y *C₂* son coeficientes que se utilizan para corregir la influencia del espectro azul en el espectro rojo y *g* es el factor de ganancia (Huete y col., 1994, 1997). El índice SAVI (*Soil-Adjusted Vegetation Index*), introduce el valor de una constante *L* en la ecuación, que es el factor de corrección de la reflectancia del suelo, su valor se determina en función de la densidad de la vegetación y varía desde 0 (alta cobertura vegetal) hasta 1 (baja cobertura).

Con estas consideraciones previas, y teniendo en cuenta nuestro problema, sería conveniente disponer de índices que únicamente sean efectivos en el espectro visible, donde las plantas y el suelo aparezcan muy contrastados, en una imagen en escala de grises. Así tenemos, el índice diferencial normalizado NDI (*Normalized Difference Index*), ecuación 3.6, propuesto en Pérez y col. (2000), aunque está basado en el NDVI, sólo utiliza los canales verde (*G*) y rojo (*R*) siendo posible trabajar con una imagen en el modelo RGB.

$$NDI = \frac{G - R}{G + R} \quad (3.6)$$

Existen otros tipos de índices de vegetación basados en el espectro visible. A lo largo de este trabajo se analizan los más relevantes con el fin de proporcionar la base que permite extraer los píxeles de una imagen que pertenecen a la capa vegetal. En la investigación orientada a la identificación de verdes mediante análisis espectral utilizamos una combinación de este tipo de índices. Como se deduce de Meyer y Camargo-Neto (2008), la ventaja de utilizar índices de color es que éstos acentúan un color en particular, en este caso, el verde. Las imágenes panorámicas utilizadas en la investigación están compuestas por tres zonas principales, las cuales pueden asociarse preferentemente con cada canal de la imagen RGB: el verde para las plantas, rojo para el suelo, y en ocasiones azul para el cielo. Teniendo en cuenta estas consideraciones, resultará apropiado utilizar los índices de verde para identificar las plantas y los índices de rojo para la identificación del suelo y en caso de que tengamos cielo utilizaremos el índice azul para identificarlo.

Como los índices utilizados están basados en componentes del espectro visible, es necesario realizar una normalización de los valores de cada canal. Por ello, dada una imagen original en el espacio de color RGB, se aplica la relación de normalización, ecuación (3.7), utilizada habitualmente en la segmentación de imágenes agrícolas (Gée y col., 2008), obteniendo las componentes espectrales r , g y b normalizadas en el intervalo $[0, 1]$.

$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B} \quad (3.7)$$

donde R , G y B son las componentes espectrales en el modelo de color RGB en el rango $[0,255]$ o equivalentemente $[0,1]$ y obtenidas de la siguiente forma:

$$R = \frac{R}{R_{\max}} \quad G = \frac{G}{G_{\max}} \quad B = \frac{B}{B_{\max}} \quad (3.8)$$

donde $R_{\max} = G_{\max} = B_{\max} = 255$ para imágenes de 24-bits, que se corresponde con 8 bits por canal y por píxel.

Los índices de vegetación se pueden utilizar independientemente o combinados, tal como se propone en Guijarro y col. (2008). Para obtener los verdes se combinan cuatro de los índices más utilizados en el tratamiento de imágenes agrícolas ExG, ExGR, CIVE y VEG, tabla 3.2, generando un resultado combinado que es posteriormente aplicado por Guerrero y col, (2012).

| | | |
|---|--|--------|
| Exceso de verde: | $ExG = 2g - r - b$ Woebbecke (1995) y Ribeiro y col., (2005) | (3.9) |
| Exceso de rojo: | $ExR = 1.4r - g$ Meyer y col., (1998) | (3.10) |
| Exceso de azul: | $ExB = 1.4b - g$ derivado de ExR | (3.11) |
| Diferencia del exceso de verde con el exceso de rojo | $ExGR = ExG - ExR$ Neto (2004) | (3.12) |
| Extracción de la Vegetación | $CIVE = 0.441r - 0.811g + 0.385b + 18.78745$ Kataoka y col., (2003) | (3.13) |
| Vegetativo | $VEG = \frac{g}{r^a b^{1-a}}$, con a inicializado a 0.667 como en el trabajo de Hague y col. (2006) | (3.14) |

Tabla 3.2.- Índices de vegetación basados en el espectro visible.

3.3. Umbralización

Los índices anteriores transforman las imágenes en color a escala de grises, por esto necesitamos un umbral que determine a partir de qué valor se considera planta o no, tras este proceso los píxeles que corresponden a la vegetación se etiquetarán con blancos y en negro el resto.

Es muy difícil determinar cuál es el umbral óptimo para realizar una binarización que separe adecuadamente el objeto del fondo. Existen diversas técnicas de umbralización, algunas de ellas están basadas en la información estadística que ofrece el histograma de la imagen, que en

una aproximación clásica se puede considerar formado por la suma de funciones de densidad de probabilidad. En el caso del histograma bimodal, la función global que aproxima el histograma puede ser: $p(z) = P_1p_1(z) + P_2p_2(z)$, donde z es una variable aleatoria que representa la intensidad, $p_1(z)$ y $p_2(z)$ son las funciones de densidad de probabilidad y, P_1 y P_2 son las probabilidades a priori.

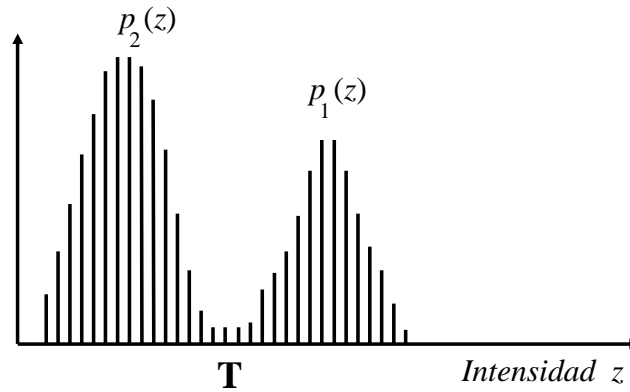


Figura 3.1.- Histograma de intensidad como suma de dos funciones de densidad de probabilidad.

Al número de veces que un determinado valor aparece en la imagen se le denomina frecuencia de ese valor. En la figura 3.1, se muestra un diagrama de barras, en la que la altura de cada barra representa la frecuencia de aparición de cada valor de intensidad. El histograma se utiliza principalmente cuando en las imágenes los objetos tienen una textura o superficie homogénea y el fondo es más o menos uniforme.

Si el histograma de una imagen presenta dos picos suficientemente diferenciados podemos separar dos zonas o regiones y el umbral será aquel valor que se encuentra en el valle entre ambas. Este es un caso ideal, el histograma de intensidad de una imagen (niveles de gris) tendría bien marcado los dos picos para objeto y fondo, y diríamos que el umbral óptimo es aquel valor que separa diversos tipos de regiones. Sin embargo, las imágenes siempre presentan ruido provocado por factores, como puede ser la luz, y, como consecuencia de esto, sus histogramas son difíciles de analizar, es decir se obtienen histogramas sin que la separación entre sus regiones esté claramente diferenciada.

Para que el histograma obtenido resulte útil para establecer una buena umbralización interesa que tenga dos lóbulos cuya distribución sea más o menos equilibrada. Sabemos que, la probabilidad de que un píxel dado se encuentre cercano al fondo suele ser igual a la

probabilidad de que se encuentre al borde de un objeto, así se aumenta la simetría entre los picos del histograma.

Una vez obtenido el histograma, pueden aplicarse diferentes técnicas con el fin de obtener un valor de umbral adecuado (Gonzales-Barron y Butler, 2006; Sahoo y col., 1988). Gonzales-Barron y Butler (2006), compararon los resultados obtenidos con 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 procedimiento basado en *K-means* (Hartigan, 1975) para los valores de umbral óptimo. En esta investigación se evidencia que no existe un único método que pueda ser considerado adecuado para todas las imágenes de forma global. En Sezgin y Sankur (2004), se realiza una amplia revisión bibliográfica en la que se pueden encontrar los principales métodos de umbralización basados en histograma, bajo la cual se analizan, por ejemplo, los picos, valles y curvaturas del histograma suavizado.

3.3.1. Umbralización manual

Una vez obtenido el histograma de la imagen se puede realizar una umbralización manual a partir de un valor fijo, o bien seleccionar un umbral manual identificando los lóbulos del histograma y los valles entre ellos, así se realiza en Gonzales-Barron y Butler (2006). Identificados los dos lóbulos, el umbral más adecuado es aquel que además de estar situado a la izquierda del valor medio de la intensidad, es un mínimo local para el valor de frecuencia.

En cualquiera de los casos, se requiere una intervención por parte del usuario para analizar cada imagen y seleccionar el umbral correcto, como en el método propuesto en Romeo y col. (2012b). En Arroyo y col. (2012) se utiliza un enfoque de aprendizaje para establecer el umbral de una imagen determinada cuyo umbral óptimo es desconocido. A partir de los histogramas de las imágenes en escala de grises, cuyo valor de umbral óptimo se determina de forma manual, dada una nueva imagen, su umbral se determina teniendo en cuenta los valores límite de los ejemplos de entrenamiento más cercanos.

En la investigación que se plantea en el presente trabajo no es posible utilizar aproximaciones manuales dado que no se puede aplicar a tareas automáticas y en entornos

de exterior donde las condiciones de iluminación son muy variables. Por ello es necesario aplicar técnicas de umbralización automática. Algunas de esas técnicas son: algoritmo *K-means* (Hartigan, 1975), método de entropía (Pun, 1980), algoritmo de mínimo error (Kittler y Illingworth, 1986), umbralización *fuzzy* (Huang y Wang, 1995), ISODATA (Ridler y Calvard, 1978), algoritmo de preservación del momento (Tsai, 1985), o el método de Otsu (1979). En la investigación realizada se han utilizado dos métodos automáticos sobre la imagen segmentada, a saber: la media estadística y el método de Otsu.

A continuación se describe este último dada su eficacia y rendimiento en un determinado tipo de imágenes utilizadas en este trabajo y por su extenso uso en umbralización de imágenes agrícolas. La media resulta efectiva en otros tipos de imágenes, ya que Otsu produce en ocasiones infra-segmentación a la hora de determinar el verdor, como se explica en Guijarro y col. (2011), efecto éste que también se compensa mediante la introducción de información espacial de la textura en el caso de los verdes, como se hace en Guijarro y col. (2015), lo que permite aplicar Otsu con resultados satisfactorios. Para la detección de las partes azules correspondientes al cielo en imágenes panorámicas Otsu se ha comportado eficientemente, mientras que para la detección de las tonalidades rojas procedentes del suelo es la media la más efectiva. Estas son las razones por las que en la investigación propuesta se utiliza en unos casos Otsu y en otros la media estadística, tal como se describe en el capítulo cuatro.

En cualquier caso, gracias al trabajo desarrollado en Arroyo y col. (2012) se demostró la eficiencia de métodos de umbralización basados en el análisis del histograma para imágenes agrícolas, lo que afianza la utilización del método de Otsu en la investigación propuesta en la presente memoria.

3.3.2. Método de Otsu

El método de Otsu (1979) es una de las técnicas de binarización de imágenes digitales ampliamente utilizadas en la literatura, sobre todo cuando se trabaja con imágenes reales tomadas en entornos de exterior. Otros métodos en este tipo de imágenes no dan buenos resultados debido a la presencia de ruido, histogramas planos o iluminación altamente cambiante. Este método, que se adapta a diversas condiciones adversas, se basa en conceptos estadísticos, en concreto se utiliza la varianza que es una medida de dispersión de

valores (en el caso que nos ocupa, se trata de la dispersión de los niveles de gris). El objetivo del método es calcular el valor del umbral de forma que la dispersión dentro de cada clase sea lo más pequeña posible, pero al mismo tiempo que la dispersión sea lo más alta posible entre clases diferentes. Otra característica importante de Otsu radica en que es un método no supervisado, en el sentido de que es automático, es decir, no necesita supervisión humana ni información previa de la imagen.

Se supone que la imagen en el histograma bimodal está formada por la suma de dos densidades de probabilidad gaussianas donde cada gaussianiana se aproxima a cada uno de los lóbulos. Esto permite suponer que a medida que las gaussianas se asemejan al histograma real, las desviaciones estándar deben disminuir y como consecuencia de ello se debe elegir el umbral que minimice la suma de las varianzas de los dos lóbulos del histograma.

Una imagen es una función bidimensional de la intensidad del nivel de gris, conteniendo N píxeles cuyos niveles de gris se encuentran en el rango de valores 1 y L . El método tiene como punto de partida una imagen con L niveles de intensidad, en nuestro caso 256, suponemos que T es el umbral buscado, por lo que las probabilidades acumuladas hasta T y desde T hasta L vienen dadas por las expresiones (3.15).

$$w_1(t) = \sum_{z=1}^T P(z) \quad \text{y} \quad w_2(t) = \sum_{z=T+1}^L P(z) \quad (3.15)$$

Una vez obtenidas las probabilidades acumuladas, se obtienen las medias y varianzas asociadas a partir de las ecuaciones (3.16) y (3.17).

$$\mu_1(t) = \sum_{z=1}^T zP(z) \quad \text{y} \quad \mu_2(t) = \sum_{z=T+1}^L zP(z) \quad (3.16)$$

$$\sigma_1^2(t) = \sum_{z=1}^T (z - \mu_1(t))^2 \frac{P(z)}{w_1(t)} \quad \text{y} \quad \sigma_2^2(t) = \sum_{z=T+1}^L (z - \mu_2(t))^2 \frac{P(z)}{w_2(t)} \quad (3.17)$$

Por último se obtiene la varianza ponderada aplicando la expresión (3.18).

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t) \quad (3.18)$$

Para obtener el valor, se elige el valor T que corresponde al nivel de intensidad que proporciona la mínima varianza ponderada establecida en la ecuación (3.18).

3.4. Clasificación: agrupamiento borroso

Uno de los métodos que se proponen en el presente trabajo de investigación, a partir de la imagen segmentada proporcionada por los índices de vegetación enumerados en la sección 3.2, es la utilización de métodos de clasificación, más concretamente de agrupamiento borroso; gracias al cual es posible identificar las texturas correspondientes a las partes vegetales, y el suelo existentes en las imágenes. La identificación de diferentes tipos de plantas y suelo se realiza mediante la técnica de clasificación conocida como *Agrupamiento borroso o Fuzzy Clustering*, que describimos en la sección siguiente, para seguidamente proporcionar los detalles de la estrategia global aplicada en la sección 4.1.1.

El objetivo de la técnica de agrupamiento borroso consiste en dividir n objetos $x \in X$ caracterizados por p propiedades en c clústeres o grupos. Supongamos 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 . Este clasificador puede encontrarse perfectamente especificado en las siguientes referencias clásicas: Bezdek (1981), Duda y Hart, (2000) o Zimmerman (1991).

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 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.

Una distancia entre dos objetos de un universo X es una función $d : X \times X \rightarrow R^+$ que toma valores reales que denotamos por $d(x_k, x_l) = d_{kl} \geq 0$ y que cumple tres propiedades:

1. $d_{kl} = 0 \Leftrightarrow x_k = x_l$ para todo $0 \leq k, j \leq N$

2. Simetría: $d_{kl} = d_{lk}$ para todo $0 \leq k, l \leq N$

3. Desigualdad triangular $d_{kl} \leq d_{kj} + d_{jl}$ para todo $0 \leq k, j, l \leq N$ (3.19)

Cada partición del conjunto $X = \{x_1, x_2, \dots, x_n\}$ puede enfocarse desde dos perspectivas: fuzzy y no fuzzy. Una partición no fuzzy se conoce en terminología inglesa como “crisp”. Si se desea realizar una partición del conjunto X en c clases tendremos $S_i \{i = 1, \dots, c\}$ subconjuntos.

Para cada partición, S_i define un conjunto borroso μ_i sobre el universo X , $\mu_i : X \rightarrow [0,1]$ que asigna lo que se conoce como grado de pertenencia μ_{ik} de cada objeto x_k al subconjunto S_i . Así pues denotamos dicho grado de pertenencia del elemento x_k al clúster S_i como $\mu_i(x_k) = \mu_{ik}$. En el caso de conjuntos “crisp” un objeto x_k se dice que pertenece a un S_i dado y no pertenece al resto. Esto se expresa con los valores discretos $\{0,1\}$ de la siguiente forma $\mu_{ik} = 1$ para indicar que pertenece y $\mu_{ik} = 0$ para expresar que no pertenece. Por el contrario, en el caso de conjuntos fuzzy se dice que un objeto puede pertenecer a diferentes subconjuntos y así se habla por ejemplo de que x_k pertenece a un conjunto S_i con grado de pertenencia μ_{ik} y a S_j con grado de pertenencia μ_{jk} . Como ejemplo, supongamos que se tienen tres conjuntos S_i, S_j y S_h , en este caso podríamos decir que el objeto x_k pertenece a los conjuntos con los siguientes grados de pertenencia $\mu_{ik} = 0.4$, $\mu_{jk} = 0.5$ y $\mu_{hk} = 0.1$. Los valores tomados pertenecen al intervalo continuo $[0,1]$.

Dado $X = \{x_1, x_2, \dots, x_n\}$ 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 representando la partición de la siguiente manera $U = \{\mu_{ik}\} \in V_{cn}$. Tanto en el supuesto “crisp” como en el fuzzy se deben cumplir las condiciones establecidas en (3.20).

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

2) $\sum_{i=1}^c \mu_{ik} = 1 \quad 1 \leq k \leq n$ (3.20)

$$3) 0 < \sum_{k=1}^n \mu_{ik} < n \quad 1 \leq i \leq c$$

Para ilustrar los conceptos anteriores sea $X = \{x_1, x_2, x_3\}$ con él podríamos construir las siguientes particiones considerando $c = 2$,

“crisp”

$$U = \begin{bmatrix} x_1 & x_2 & x_3 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} x_1 & x_2 & x_3 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

“fuzzy”

$$U = \begin{bmatrix} x_1 & x_2 & x_3 \\ 0.3 & 0.5 & 0 \\ 0.7 & 0.5 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} x_1 & x_2 & x_3 \\ 0.9 & 0.4 & 0.2 \\ 0.1 & 0.6 & 0.8 \end{bmatrix}$$

La localización de un clúster S_i se representa por su centro $v_j = \{v_{j_1}, v_{j_2}, \dots, v_{j_p}\} \in \mathfrak{R}^p$ con $j = 1, \dots, c$, alrededor del cual se concentran los objetos.

La definición básica de llevar a cabo el problema de la partición fuzzy para $m > 1$ consiste en minimizar la función objetivo según la ecuación (3.23).

$$\min z_m(U; v) = \sum_{k=1}^n \sum_{j=1}^c \mu_{jk}^m \|x_k - v_j\|_G^2 \quad (3.21)$$

donde G es una matriz de dimensión $p \times p$ que es simétrica y definida positiva. Así se puede definir una norma general del tipo como (3.22).

$$\|x_k - v_j\|_G^2 = (x_k - v_j)^t G (x_k - v_j) \quad (3.22)$$

Diferenciando la función objetivo para v_i (suponiendo constante U) y μ_{jk} (suponiendo constante v) y aplicando la condición de que $\sum_{j=1}^c \mu_{jk} = 1$, se obtiene la expresión (3.23).

$$v_j = \frac{1}{\sum_{k=1}^n (\mu_{jk})^m} \sum_{k=1}^n (\mu_{jk})^m x_k \quad j = 1, \dots, c \quad (3.23)$$

$$\mu_{jk} = \frac{\left(\frac{1}{\|x_k - v_j\|_G^2} \right)^{2/m-1}}{\sum_{h=1}^c \left(\frac{1}{\|x_k - v_h\|_G^2} \right)^{2/m-1}} \quad j = 1, \dots, c; k = 1, \dots, n \quad (3.24)$$

el exponente m se conoce como peso exponencial y disminuye la influencia del ruido al obtener los centros de la clases, reduciendo la influencia de los valores pequeños de μ_{jk} (puntos alejados de v_j) frente a valores altos de μ_{jk} (puntos cercanos a v_j). Cuanto mayor sea $m > 1$ mayor será dicha influencia.

Existen diferentes medidas escalares para validar la partición, siendo el objetivo encontrar el número de clases que obtenga las mejores medidas. De entre todos ellos destaca por su importancia (Balasko y col. 2005; Kim y col. 2003) el *Coefficiente de partición (CP)* que mide la cantidad de solapamiento entre las clases,

$$CP = \frac{1}{n} \sum_{j=1}^c \sum_{i=1}^n (\mu_{ji})^2 \quad (3.25)$$

Cuanto más próximo sea este valor a la unidad, tanto mejor será la clasificación de los datos, llegando a ser una partición pura cuando toma el valor de la unidad.

3.5. Transformada de wavelets: combinación

Desde el punto de vista que nos compete, la transformada de wavelets constituye una potente herramienta para afrontar problemas fundamentales en el tratamiento de imágenes. Entre ellos se encuentran la descomposición de la imagen original en distintas bandas de frecuencia para extraer información relativa a componentes de alta y baja frecuencia, como es el caso de las primeras en las texturas vegetales. Esta teoría ha experimentado un gran desarrollo en las dos últimas décadas, surgió alrededor del 1984 y en 1988 se logró aplicar la

teoría a la práctica, mostrándose muy eficiente donde otras técnicas no resultaban satisfactorias (Chou y col., 2007; Ishak y col., 2009)

En Pajares y Cruz (2007) se describen con detalle las características de esta herramienta matemática que nace de la posibilidad que la misma ofrece de superar algunas de las limitaciones de otras transformaciones, entre las que se destaca la muy conocida transformada de Fourier. Una *wavelet* es una forma de onda de duración limitada que tiene un valor medio cero. En cuanto al análisis de Fourier consiste básicamente en descomponer la señal en ondas sinusoidales de diferentes frecuencias. El estudio de las wavelets consiste en la descomposición de una señal arbitraria f en versiones escaladas y trasladadas de la wavelet original. Es decir, la idea básica de esta transformada reside en representar cualquier función arbitraria f como una superposición de un conjunto de dichas wavelets o funciones base.

La señal es pasada por varios filtros paso bajo y paso alto, los cuales permiten separar las porciones de la señal de alta y baja frecuencia, este procedimiento se repite sobre algunas porciones de la señal correspondiente a aquellas frecuencias que han sido eliminadas de la señal original.

La diferencia entre la teoría de Fourier y las wavelets radica en que las funciones wavelet varían tanto en frecuencia como en escala, además permiten analizar señales que no se comportan de forma estacionaria o que presentan cambios bruscos en intervalos muy pequeños Asimismo, poseen una serie de características que hacen que sean muy atractivas para la resolución del problema planteado en lo que respecta a la introducción de la variabilidad espacial procedente de las texturas vegetales presentes en las imágenes analizadas:

1. Su descomposición jerárquica que permite la caracterización de la señal a distintas escalas (análisis multiresolución).
2. Es en esencia una descomposición de la señal en sub-bandas; lo que está muy relacionado con una gran variedad de técnicas de descomposición multifrecuencia.
3. Existe un algoritmo rápido para calcularla de implementación digital sencilla.

La transformada de wavelets de una señal f , es la familia de coeficientes $C(a,b)$, que dependen de dos índices a y b que se asocian con la escala y la posición de la señal. Estos coeficientes cuando tenemos una única dimensión se obtienen mediante la siguiente función

$$C(a,b) = \int_{-\infty}^{+\infty} f(x) \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) dx \quad a \in R^+ - \{0\} \text{ y } b \in R \quad (3.26)$$

donde R es el conjunto de los números reales.

Teniendo en cuenta la ecuación anterior, la transformada se define como la suma de la señal multiplicada por las versiones escaladas y trasladadas de la función wavelet ψ , sobre todo el intervalo. Multiplicando cada coeficiente por la apropiada wavelet escalada y trasladada obtenemos las wavelets que componen la señal original. Cuando decimos escalar una wavelet nos estamos refiriendo a comprimirla o expandirla, por tanto a es el factor de escala.

Para realizar este proceso podemos seguir los siguientes pasos:

1. Se selecciona una wavelet y se compara con una sección al principio de la señal original.
2. Se calcula un coeficiente C , éste representa el grado de correlación entre la wavelet y la sección correspondiente de la señal original, figura 3.2. El grado de similitud será tanto mayor cuanto mayor sea C . Se puede comprobar que dependiendo de la forma elegida para la wavelet así serán los resultados.

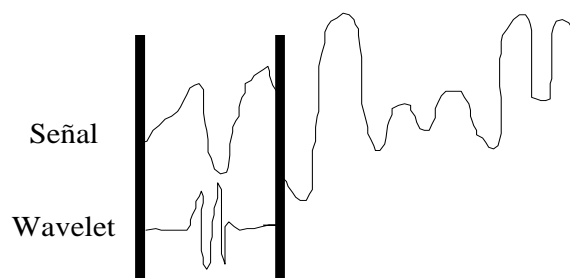


Figura 3.2.- Cálculo del primer coeficiente

3. Se traslada la wavelet a la derecha y se repiten los pasos 1 y 2 hasta cubrir toda la señal, figura 3.3.

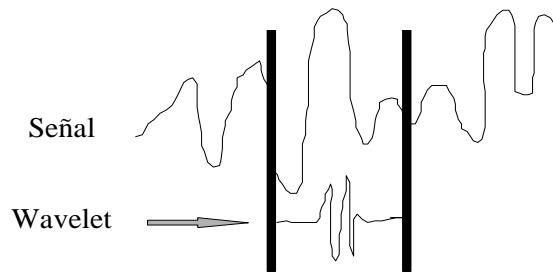


Figura 3.3.- Desplazamiento de la wavelet

4. Se escala la wavelet y se repiten los pasos 1 a 3.
5. Se repiten los pasos 1 a 4 para todas las escalas, figura 3.4.

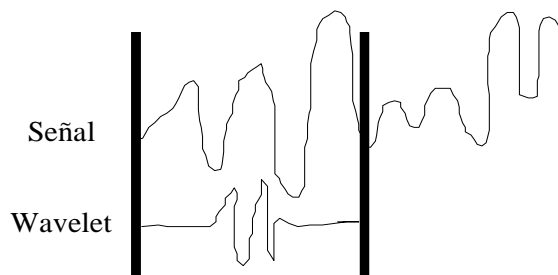


Figura 3.4.- Escalado (expansión) y comienzo de nuevo desde el principio

En el caso bidimensional la familia de wavelets viene dada por la siguiente expresión

$$\frac{1}{\sqrt{a_1 a_2}} \psi\left(\frac{x_1 - b_1}{a_1}, \frac{x_2 - b_2}{a_2}\right) \text{ con } x = (x_1, x_2) \in R^2, a_1 > 0, a_2 > 0, b_1 \in R \text{ y } b_2 \in R. \quad (3.27)$$

Ante la imposibilidad de realizar los cálculos de los coeficientes para todas las escalas y posiciones, se recurre a la versión discreta denominada *transformada discreta de wavelets* (TDW) donde se elige un subconjunto de escalas y posiciones. El análisis resulta mucho más eficiente y preciso cuando el subconjunto de escalas y posiciones son potencias de 2.

$$C(a,b) = C(j,k) = \sum_{x \in Z} f(x) g_{j,k}(x) \text{ con } a = 2^j, b = k2^j, j \in N, k \in Z \quad (3.28)$$

donde Z es el conjunto de los números enteros y N el de los naturales.

Inicialmente Woods y O'Neill (1986) y posteriormente Mallat (1989) implementaron este esquema de una forma eficiente mediante el uso de filtros. El contenido en baja frecuencia es la parte más importante para algunas señales, porque proporciona a la señal su identidad, éste puede ser el caso de las imágenes. Cuando lo que se maneja son imágenes, como es nuestro caso, el filtrado de baja frecuencia suaviza la imagen y ésta sigue siendo visible, y el de alta frecuencia extrae los bordes y en este caso ya no es visible la imagen.

Teniendo en cuenta estos detalles, en la teoría general se dice que se descompone la señal de entrada (imágenes) en dos componentes: (a) *aproximación de baja resolución* y (b) *detalle de la señal*. Generalmente se conoce como *sub-bandas* el resultado de descomponer la señal de entrada en versiones paso bajo y paso alto. Cada una de estas sub-bandas se puede seguir descomponiendo por el mismo procedimiento. De esta manera se dice que la TDW descompone una señal de entrada en un cierto número de bandas de frecuencia. En nuestra investigación, son las componentes de alta frecuencia las que poseen el interés intrínseco en relación a la propuesta realizada.

En nuestro caso las señales son imágenes, por eso esta transformada descompone la imagen original en cuatro imágenes submuestreadas o diezmadas. El submuestreo se realiza cada dos píxeles.

El resultado consta de

- Una imagen que ha sido filtrada mediante un filtro paso alto tanto en la dirección vertical como horizontal.
- Una imagen que ha sido filtrada en paso alto en la vertical y en paso bajo en la horizontal.
- Una imagen que ha sido filtrada en paso bajo en la vertical y en paso alto en la horizontal.
- Una que ha sido filtrada en paso bajo en ambas direcciones.

La TDW se implementa en el dominio espacial utilizando filtros de convolución 1-D. Los filtros deben ser filtros de *reconstrucción perfecta*, para que se satisfagan favorablemente las condiciones de esta transformada. Que los filtros sean de reconstrucción perfecta significa que cualquier distorsión introducida por la transformación directa debe ser

cancelada por la transformación inversa y deben cumplir algunas propiedades tales como respuesta de impulso finita de los filtros directos e inversos o fase lineal.

Existen muchas familias de filtros para implementar la TDW, las más utilizadas son las de Daubechies y en su versión más simplificada la wavelet de Haar, que pertenece a esta misma familia, que es justamente la utilizada en nuestra investigación cuando se incorpora información espacial para la segmentación de las texturas naturales.

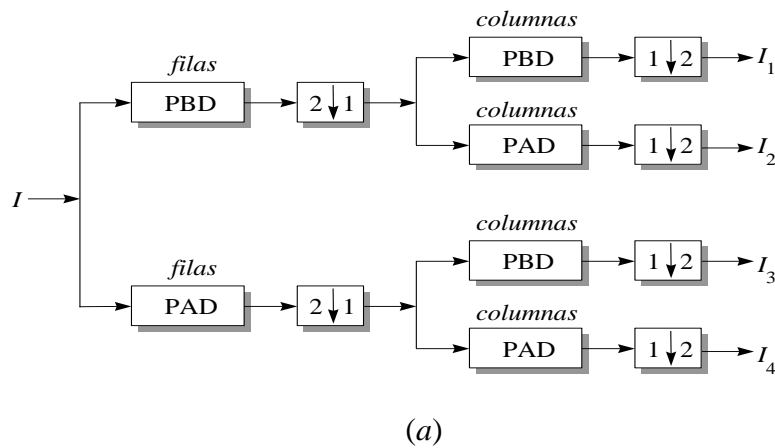
Para realizar las operaciones paso bajo/alto se eligen los vectores base, y para obtener la TDW se siguen los siguientes pasos:

1. Se realiza la convolución de las filas con el filtro paso bajo y se guardan los resultados.
2. Se realiza la convolución de las columnas con el filtro paso bajo, a partir de los resultados del paso 1. Se obtiene una imagen reducida tomando sólo un píxel de cada dos; esto nos genera una versión que denominamos *paso bajo/paso bajo* de la imagen.
3. Se realiza la convolución del resultado del paso 1 con el filtro paso alto en las columnas. Se obtiene una imagen reducida tomando sólo un píxel de cada dos, obteniendo ahora una imagen *paso bajo/paso alto*.
4. Se realiza la convolución de la imagen original con el filtro paso alto en las filas y se guarda el resultado.
5. Se realiza la convolución del resultado del paso 4 con el filtro paso bajo en las columnas. Se obtiene una imagen reducida tomando sólo un píxel de cada dos, obteniendo ahora una imagen *paso alto/paso bajo*.
6. Se realiza la convolución de las columnas del resultado del paso 4 con el filtro paso alto. Se obtiene una imagen reducida tomando sólo un píxel de cada dos, obteniendo ahora una imagen *paso alto/paso alto*.

| | |
|--|--|
| <i>Paso Bajo/ Paso Bajo</i> | <i>Paso Bajo/ Paso Alto</i> |
| <i>Paso Alto/ Paso Bajo</i> | <i>Paso Alto/ Paso Alto</i> |

Figura 3.5.- Localización de las bandas de frecuencia en una TDW con cuatro bandas. La convención es fila/columnna.

En la figura 3.6, se muestra el algoritmo anterior y su simbología. Además PBD significa filtro Paso Bajo Directo y PAD filtro Paso Alto Directo.



- filas*
X : Convolución con el filtro X de las *filas* de la entrada
- columnas*
X : Convolución con el filtro X de las *columnas* de la entrada
- 2 ↓ 1 : Submuestreo de *columnas*, mantener las pares
- 1 ↓ 2 : Submuestreo de *filas*, mantener las pares

(b)

Figura 3.6.- (a) Esquema del algoritmo de filtrado por filas y columnas para la descomposición de la imagen original en cuatro subimágenes. (b) Simbología utilizada

Woods y O’Neil (1986) utilizaron esta combinación de filtros separables unidimensionales para realizar la descomposición en cuatro bandas, conocida como

Quadrature Mirror Filterbanks (QMF). La construcción de familias ortogonales de funciones base de wavelets puede realizarse en el caso continuo, sin embargo, esto mismo se puede obtener a partir de filtros discretos, esto lo describió por primera vez Daubechies (1988), ya que en el caso continuo se pueden usar filtros de respuesta impulso finita para derivar una familia de wavelets. Por esto podemos deducir que la transformación en sub-bandas es un caso particular de la transformada de wavelets.

El proceso inverso trata de reconstruir la señal original a partir de sus elementos previamente descompuestos, la reconstrucción en el caso continuo resulta en la expresión (3.29).

$$f(x) = \frac{1}{K_\psi} \int_{R^+} \int_R C(a,b) \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \frac{dad b}{a^2} \quad (3.29)$$

donde K_ψ es una constante que depende de ψ .

Y en el caso discreto resulta en la expresión (3.30).

$$f(x) = \sum_{j \in Z} \sum_{k \in Z} C(j,k) \psi_{j,k}(x) \quad (3.30)$$

En las imágenes, que es el caso que nos interesa, la TDW inversa se realiza aumentando los datos de la TDW a su dimensión original, para luego insertar ceros entre cada valor, realizar la convolución a cada una de las cuatro subimágenes, según corresponda (*paso bajo* y *paso alto*) con los filtros inversos y sumar los resultados para obtener la imagen original.

Al igual que hicimos con la imagen original y su descomposición, la figura 3.7 muestra por una parte un esquema del procedimiento de (véase figura 3.7a), es decir de la TDW inversa, y por otra la simbología utilizada (véase figura 3.7b). Donde además PBI significa filtro Paso Bajo Inverso y PAI filtro Paso Alto Inverso

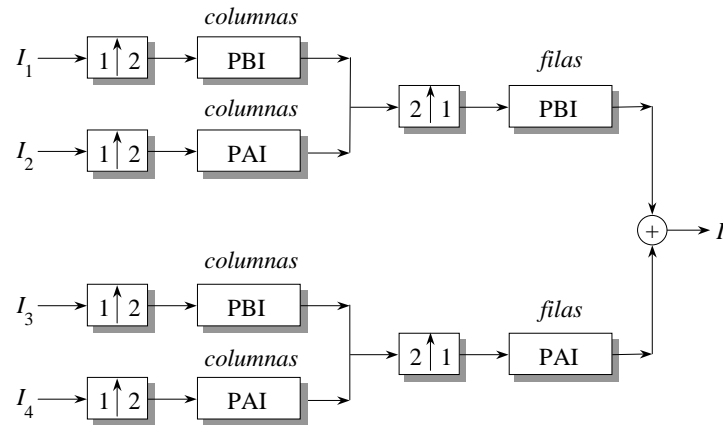


Figura 3.7. a –Esquema del procedimiento de reconstrucción.

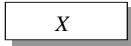
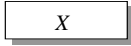



- filas*
 : Convolución con el filtro *X* de las *filas* de la entrada
- columnas*
 : Convolución con el filtro *X* de las *columnas* de la entrada
-  : Extender imagen, dada una imagen con *m* *columnas*, insertar *m*-1 columnas de ceros entre las *m* *columnas* dadas
-  : Extender imagen, dada una imagen con *n* *filas*, insertar *n*-1 filas de ceros entre las *n* *filas* dadas
-  : Cada imagen de entrada I_i genera una imagen extendida por filas y columnas siguiendo el camino correspondiente. Si la imagen extendida es de mayor dimensión que la de la imagen *I*, se elige sólo la parte central de la imagen extendida cuya dimensión coincide exactamente con la de *I*. Las cuatro imágenes resultantes se suman para obtener *I*

Figura 3.7.- b) Simbología utilizada

El uso de la TDW es cada día más popular para la compresión de imágenes, un área de investigación muy activa hoy en día.

3.5.1. Wavelets de Daubechies

La familia de wavelets de Daubechies es una clase de wavelets ortogonales. Se adaptan bien a las imágenes que tienen cierta “suavidad”, siendo ésta una de sus principales ventajas. Se caracterizan por un número máximo de momentos de desvanecimiento denominados wavelet de orden *N*, donde *N* es un entero positivo y denota el número de coeficientes de filtro que posee la wavelet. Dentro de esta familia se encuentran las wavelets de Haar, fue introducida por primera vez por el matemático húngaro Alfred Haar, en 1909. Es la más simple de todas las wavelts ya que viene a ser una secuencia de funciones, que

consisten en un breve impulso positivo seguido de un breve impulso negativo. La función de esta wavelet se describe como:

$$\psi(t) = \begin{cases} 1 & 0 \leq t \leq 1/2 \\ -1 & 1/2 \leq t \leq 1 \\ 0 & e. o. c. \end{cases} \quad (3.31)$$

Su gráfica se muestra en la figura 3.8, donde puede observarse su simplicidad. Cuenta con un único momento de desvanecimiento, un concepto importante que permite conocer la forma de la wavelet, ya que se trata de un parámetro que define la suavidad de la wavelet. El i-ésimo momento de una wavelet se calcula mediante la integral (3.32).

$$\int_{-\infty}^{\infty} h(x)x^i dx = 0 \quad (3.32)$$

donde h(x) es la función madre originaria de la wavelet. Una función tiene v momentos de desvanecimiento si la integral es 0 para i=0,..., v-1.

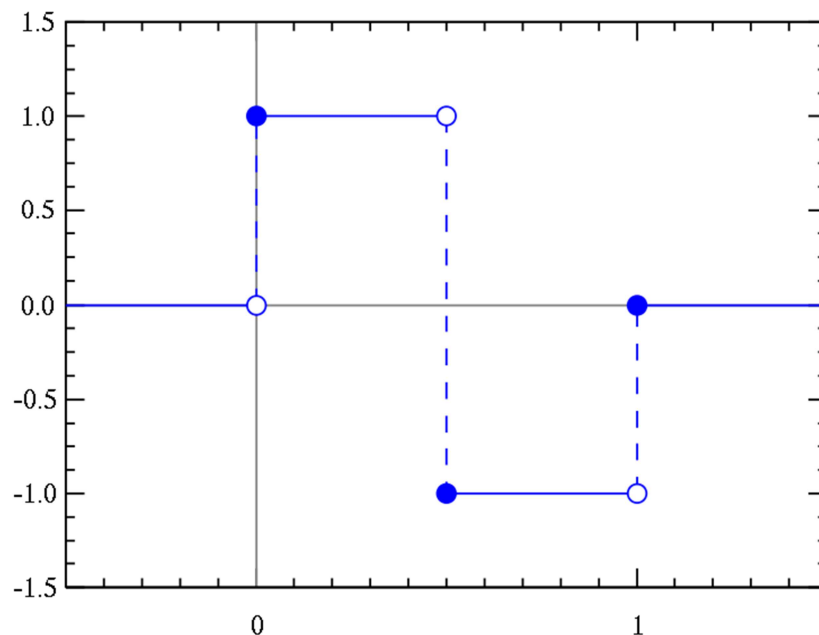


Figura 3.8.- Representación gráfica de la Wavelet de Haar

Como se ha indicado previamente, la wavelet Daubechies de orden N=1 es la wavelet Haar. En la figura 3.9 se pueden ver las distintas representaciones gráficas de esta familia de wavelets para distintos órdenes.

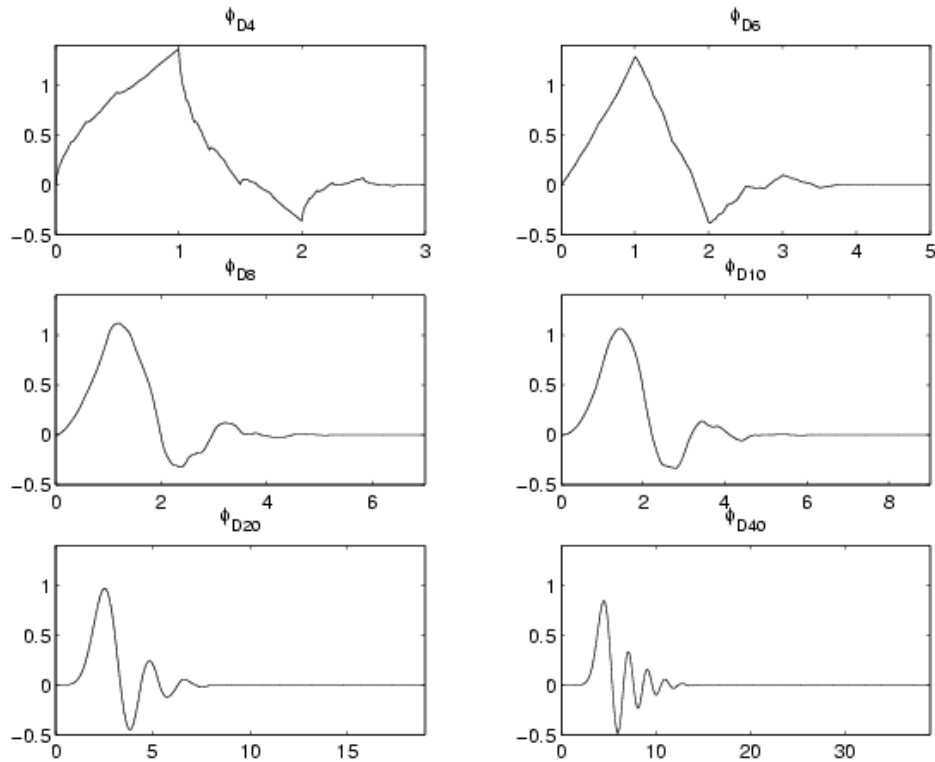


Figura 3.9.- Wavelets de Daubechies de distintos órdenes

3.5.3. Wavelets Coiflet

Este tipo de wavelets también fueron creadas por Daubechies, son ortogonales y buscan mejorar a las anteriores, consiguiendo una mayor aproximación entre los valores de las aproximaciones y los de la señal original. Tiene características de mayor simetría y posee mayor número de momentos de desvanecimiento $2N$.

3.5.4. Wavelets Symlets

La familia de funciones wavelets llamadas *symlets*, nombre que proviene de “symmetrical wavelets”, se diseñaron con el objetivo de aumentar la simetría de las wavelets Daubechies, permite realizar la transformada continua y discreta de wavelets. Las Symlets pueden ser ortogonales, biortogonales, y no son perfectamente simétricas, pero fueron diseñadas para tener la menor asimetría posible y el mayor número de momentos nulos en un soporte relativamente compacto, su número de desvanecimiento es N .

3.5.5. Wavelets Biortogonales

Poseen la característica de reconstrucción exacta de la señal a través de filtros FIR (Finite Impulse Response), su orden viene dado por N_r y por N_d , utilizados son para la reconstrucción y descomposición respectivamente. Sus momentos de desvanecimiento son N_r .

3.5.6. Wavelets Meyer

Una wavelet Meyer es una función de frecuencia limitada a una banda cuya transformada de Fourier es suave, tiene un rápido y asintótico decaimiento en el tiempo. Suaviza mediante interpolación el problema de discontinuidad, presentado en algún tipo de wavelet, las bases de Meyer tienen soporte infinito (Daubechies 1992), con un número infinito de momentos nulos. Si la función de suavizado empleada es una función simple se obtiene la wavelet Meyer de orden 0. Esta familia de wavelet es ortogonal e infinitamente periódica.

4

Sistemas Automáticos: identificación de texturas

Como se ha mencionado reiteradamente, un tema que surge con fuerza en la AP es la automatización de tareas, que se ha hecho realidad debido al creciente desarrollo de los sistemas robotizados equipados con diferentes sensores, incluyendo por supuesto los de Visión Artificial. Estos sensores, proporcionan las imágenes de los campos que deben ser procesadas, para de esta forma optimizar los tratamientos selectivos, como puede ser, el uso y aplicación de los diversos insumos (fertilizantes, fitosanitarios, semillas), y el ajuste en el uso de las semillas y de los agroquímicos considerando la diversidad del medio físico y del medio biológico. De esta forma se consigue reducir el uso de productos fitosanitarios y mejorar la rentabilidad de la producción.

Las malas hierbas que proliferan en los campos de cultivo suponen un grave problema para los agricultores, ya que interfieren en el crecimiento de los cultivos, lo que implica un bajo rendimiento de producción. Además, suele ser habitual en tratamientos no selectivos fuera de la AP la aplicación de fertilizantes, herbicidas, insecticidas, etc..., en las mismas dosis en toda el área del cultivo sin tener en cuenta las necesidades específicas localizadas, ya que la realidad demuestra que en cada campo las condiciones del cultivo

pueden variar considerablemente, debido a diferentes factores, tales como distintas densidades de malas hierbas o carencias de nutrientes en el suelo en zonas muy localizadas.

Un tratamiento selectivo supondría una reducción en el uso de herbicidas y otras aplicaciones, llevando consigo una importante reducción de costes al agricultor y una disminución del daño medioambiental. Como se ha dicho reiteradamente, dentro de los tratamientos selectivos en AP está el control de las malas hierbas, que resulta ser una actividad fundamental en la agricultura si se desea obtener cosechas y rendimientos adecuados. La investigación que se describe en la presente memoria se enmarca en esta línea para ofrecer métodos de respuesta a esta necesidad, planteando como principal objetivo la detección de las diferentes texturas de forma automática con la ayuda de las nuevas tecnologías, en este caso mediante sistemas de Visión por Computador, que proporcionan las imágenes necesarias captadas por dichos sistemas a bordo de un tractor. Pero el equipamiento de los vehículos agrícolas con sistemas de visión, va todavía un paso más allá en el sentido de la autonomía, de suerte que de cara a la navegación de los vehículos sigue siendo necesario determinar las estructuras relevantes existentes en el campo a través del procesamiento de las imágenes, de forma que conjuntamente con otros sistemas de navegación, como puede ser GPS completan el proceso de percepción global y en ocasiones de forma más precisa. Y aún más, dentro de esa autonomía, la navegación segura detectando o evitando obstáculos (objetos, personas, etc.) constituye otro aspecto de interés donde los sistemas de visión constituyen una buena herramienta, conjuntamente con otros sistemas tales como sensores de rango o alcance tipo láseres u otros.

Por tanto y a modo de resumen, se puede decir que desde la perspectiva de los sistemas de Visión Artificial, éstos pueden contribuir a la automatización de las tareas agrícolas mediante la utilización de vehículos automatizados, en ocasiones autónomos (RHEA, 2015), así como para una navegación segura. En el primer caso se trata de identificar estructuras relevantes tales como líneas de cultivo o rodales de malas hierbas para actuaciones específicas, siendo habitual centrar la región de interés de la imagen exclusivamente en zonas especificadas con tal finalidad, mientras que en el segundo caso se requiere la identificación de esas mismas estructuras así como otras derivadas de obstáculos u otros elementos, en cuyo caso las zonas de interés se extienden a toda la imagen de forma panorámica, abarcando tanto la parte del cultivo como elementos del suelo o el cielo.

En cualquier caso, la investigación desarrollada se ha centrado en dar solución a esa problemática mediante el diseño de métodos automáticos de identificación de texturas, que constituye el objetivo del presente capítulo. En relación a esto se han tenido en cuenta hechos relevantes en este tipo de imágenes como es que las firmas espectrales correspondientes a las plantas de cereal, maíz y malas hierbas, en el momento de los tratamientos requeridos para la eliminación de estas últimas, son muy similares en condiciones normales de iluminación, pero pueden diferir bastante en función precisamente de la iluminación, dado que las actuaciones se realizan siempre en entornos agrícolas de exterior.

También serán importantes las variaciones de iluminación debidas a las condiciones climatológicas del momento que podrían producir cambios importantes en las componentes espectrales (Tian y Slaught, 1998), por ejemplo días nublados frente a soleados o días con gran abundancia de nubes, alternando los claros con la ocultación del sol de manera repentina. Bajo las consideraciones anteriores y con el fin de cumplir con el objetivo de segmentar las imágenes para obtener estructuras y texturas relevantes de las mismas, se han diseñado los dos enfoques de segmentación de imágenes que se describen a continuación, basados cada uno de ellos en procesos diferenciados de tratamiento de imágenes que constituyen la entrada de datos en ambos enfoques y cuyo esquema se muestra en la figura 4.1, junto con los procesos individuales de tratamiento de las imágenes, que convenientemente secuenciados conforman los dos procedimientos completos, a saber:

1. Método de segmentación automático combinando índices de vegetación y agrupamiento borroso (Gujarro y col., 2011), que comprende los siguientes subprocesos: 1) selección del índice de vegetación, en este caso una combinación de los índices (ExG, ExGR, CIVE y VEG), convenientemente promediados; 2) aplicación de umbralización mediante la media; 3) agrupamiento borroso; 4) obtención de resultados de la umbralización en forma de imágenes binarias y clasificación distinguiendo dos tipos de plantas (cultivo y malas hierbas) En este proceso, además de los índices de vegetación también se aplican otros índices cromáticos para identificar texturas de suelo y cielo.
2. Método para la extracción de verdes basado en un índice cromático, cuyos resultados se refuerzan mediante la adición de información espacial procedente de la aplicación de la transformada de wavelets y posterior combinación de los

detalles resultantes de la transformación (Guijarro y col., 2015), consistente en los siguientes subprocesos: 1) selección del índice cromático, en este caso ExG; 2) aplicación de la transformada directa de wavelets (TDW), para obtener detalles de las texturas; 3) aplicación de umbralización mediante Otsu; 4) resultados de la umbralización conteniendo las texturas en forma de imágenes binarias.

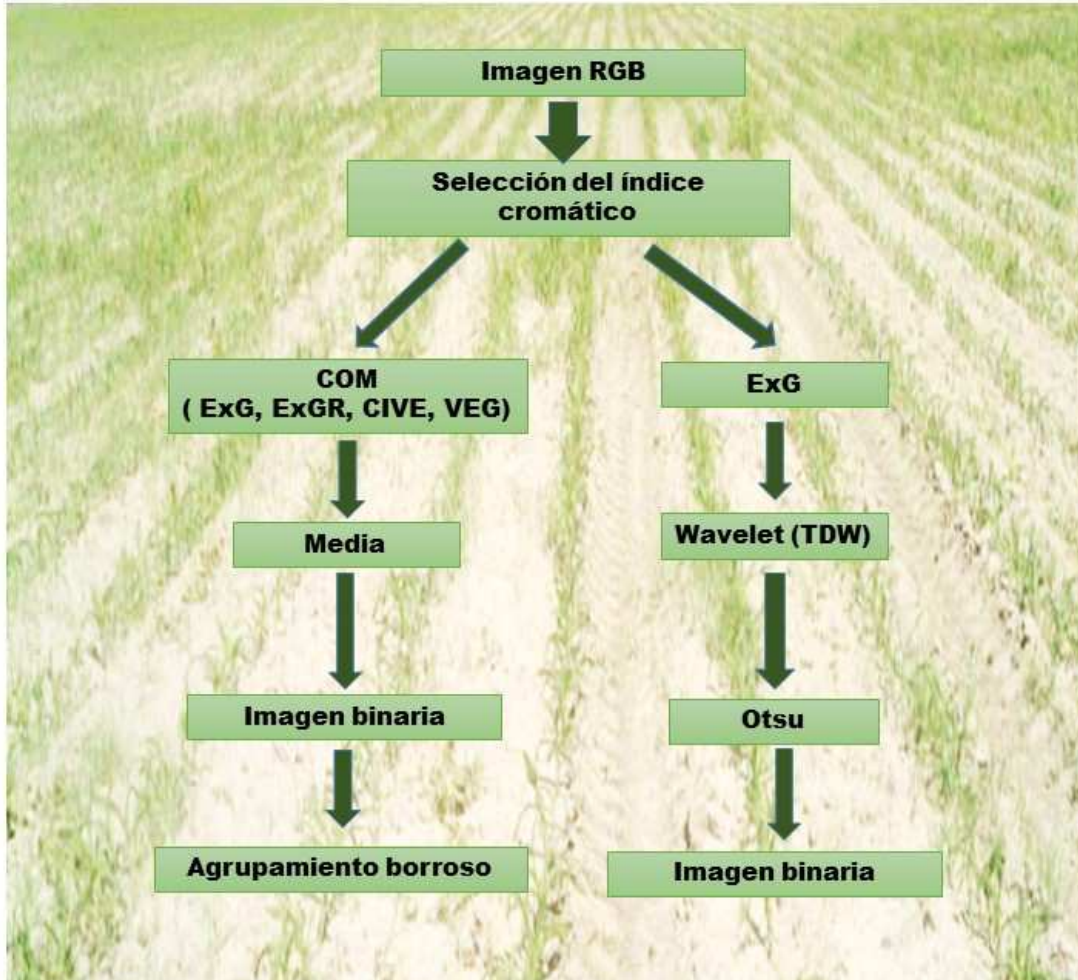


Figura 4.1.-Esquema de los pasos seguidos para las dos aproximaciones propuestas

4.1. Método de segmentación automático basado en la combinación de índices cromáticos y agrupamiento borroso

En este primer enfoque se propone un nuevo método automático para la segmentación de las principales texturas encontradas en la imagen (cielo, suelo y plantas) y también para afinar la identificación de subtexturas dentro de las principales,

particularmente en las plantas. Este método se encuentra explicado con detalle en Guijarro y col., (2011), realizándose tres contribuciones importantes:

- *Identificación de verdes*, se propone un nuevo enfoque que explota el rendimiento de varios índices de vegetación existentes mediante el diseño de una estrategia combinada de esos índices. Para la combinación se tiene en cuenta la relevancia de la información proporcionada por cada índice basada en la variabilidad de las intensidades de los niveles de gris proporcionadas por cada índice a nivel individual. Esta medida se determina mediante el análisis de las texturas utilizando criterios de mínima uniformidad y máxima entropía como descriptores de éstas.
- *Adaptación de los índices de verdor* para la detección de otras componentes espectrales dominantes, como son las correspondientes al suelo y el cielo.
- *Identificación de subtexturas* de forma automática ajustando convenientemente el método *fuzzy clustering* de naturaleza supervisada para discernir entre tres clases de vegetación y dos de suelo, una vez separadas estas texturas del resto. Esto permite principalmente distinguir entre cultivo y malas hierbas por sus diferencias espectrales, captadas por el método de clasificación propuesto.

El rendimiento del método permite verificar su viabilidad para las tareas automáticas en la agricultura basado en procesamiento de imágenes. En la figura 4.2 se muestra un esquema representativo del proceso seguido que comprende los siguientes pasos, cuyos detalles se proporcionan posteriormente:

- 1) Captura de las imágenes de entrada en el espacio de color RGB conteniendo sólo texturas de suelo con plantas (cultivo y malas hierbas) o panorámicas en las que aparece el cielo.
- 2) Normalización de las componentes espectrales R, G y B al rango de valores [0,1] con respecto a sus correspondientes valores máximos.
- 3) Cómputo de los cuatro índices cromáticos de vegetación (ExG, ExGR, CIVE y VEG).
- 4) Adaptación y cómputo de los índices cromáticos para firmas correspondientes a cielo (ExB) y suelo (ExR).

- 5) Combinación de los cuatro índices cromáticos de vegetación, determinando su relevancia ponderada mediante los criterios de mínima uniformidad y máxima entropía.
- 6) Umbralización mediante la media (texturas verdes y suelo) y Otsu (texturas de cielo), para obtener las correspondientes imágenes binarias de verdes (G_b), suelo (R_b) y cielo (B_b).
- 7) Determinación si se trata de una imagen panorámica en función de la existencia o no de partes identificativas del cielo.
- 8) Aplicar la correspondiente regla lógica con las imágenes binarias disponibles para obtener la imagen binaria final que indica la cobertura de suelo, S_b .
- 9) Sobre la imagen binaria S_b aplicar el método de clasificación Fuzzy Clustering para obtener dos tipos de clases de suelo.
- 10) Sobre la imagen binaria de verdes, G_b aplicar el método de clasificación Fuzzy Clustering para obtener tres tipos de clases de verdes.

A continuación se describen detalladamente los diferentes procesos mencionados previamente. El proceso global se inicia con la captura de imágenes en el modelo de color RGB con las características ya reseñadas previamente, y destacando que las mismas son de dos tipos, esto es, por un lado panorámicas conteniendo un amplio campo de vista incluyendo zonas de cultivo, cielo y otras texturas y por otro lado centradas exclusivamente en la zona de cultivo.

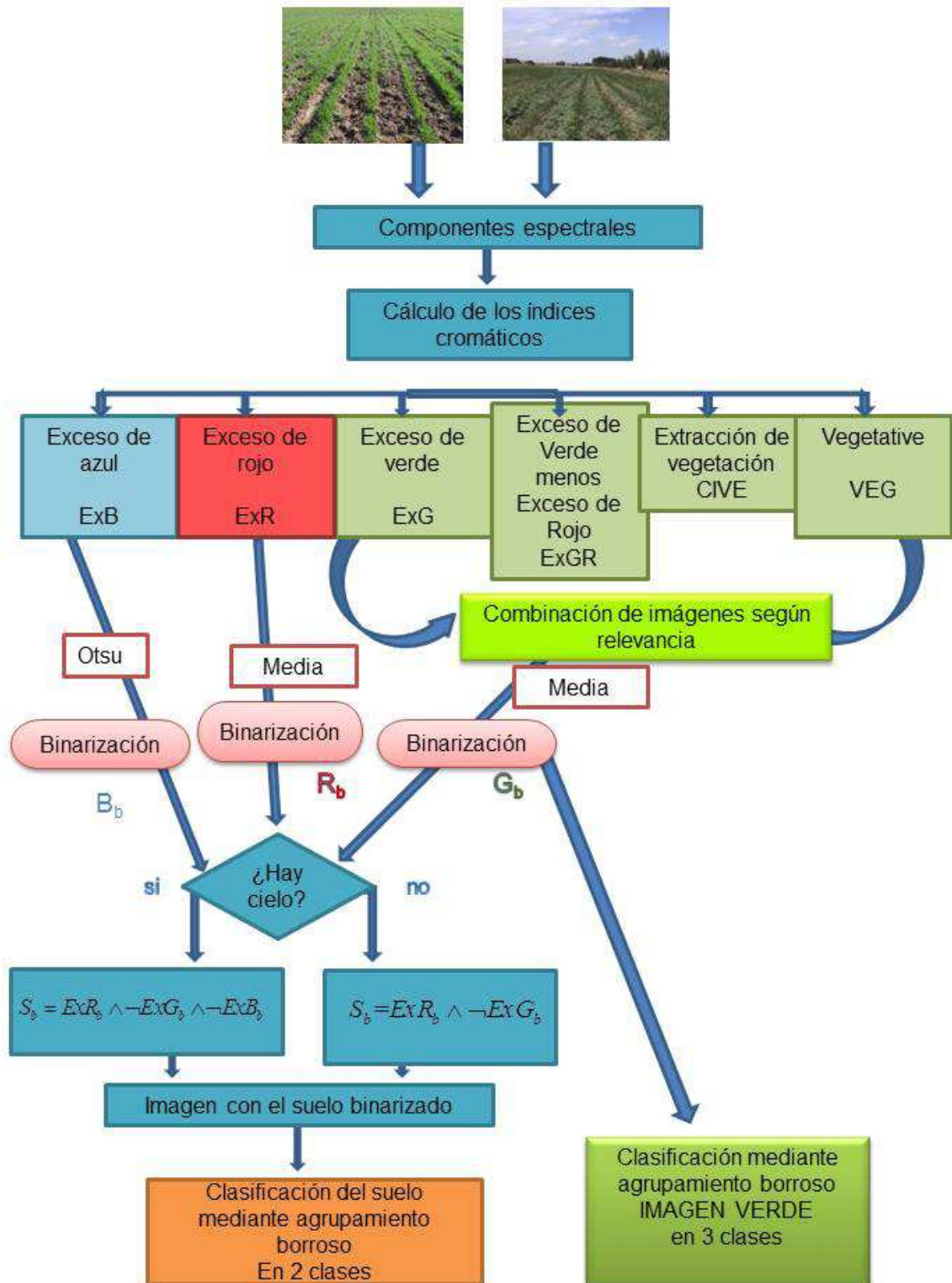


Figura 4.2.- Esquema de los pasos seguidos en la estrategia basada en índices cromáticos

4.1.1. Segmentación automática

La parte clave que se plantea bajo el punto de vista de la segmentación automática lo constituye la transformación de las imágenes originales en el espacio de color RGB a imágenes de grises mediante la aplicación de índices cromáticos, para identificar verdes con la combinación promediada de los índices de verdor ExG, CIVE, ExGR y VEG, azules para las texturas del cielo con ExB y tonalidades rojizas procedentes del suelo con ExR. Esto se completa con la aplicación de métodos de binarización (media y Otsu) y un clasificador para identificación de dos clases de texturas de suelo y tres de verdes. Los procedimientos se describen en detalle en lo que sigue.

4.1.1.1. Normalización del espacio de color

Partimos de una imagen a color para poderla identificar y extraer las texturas de los campos, como se ha mencionado reiteradamente la imagen está tomada en el espacio de color RGB. Cada píxel está bien caracterizado por sus tres valores de intensidad en sus respectivos tres canales R, G y B.

En el caso de las imágenes agrícolas, antes de aplicar un algoritmo para discriminación de las texturas, se requiere una normalización previa de los valores para las tres componentes espectrales de suerte que los valores en dichas componentes contribuyan por igual en el cómputo de los índices cromáticos. Con tal fin, la imagen capturada se normaliza según las ecuaciones (4.1) y (4.2) que se reproducen aquí por simplicidad, tal y como suele ser habitual en los tratamientos de las imágenes agrícolas Gée y col., (2008),

$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B} \quad (4.1)$$

donde R , G y B varían en el rango $[0,255]$ o equivalentemente $[0,1]$ y se obtienen de la siguiente forma:

$$R = \frac{R}{R_{\max}} \quad G = \frac{G}{G_{\max}} \quad B = \frac{B}{B_{\max}} \quad (4.2)$$

donde $R_{\max} = G_{\max} = B_{\max} = 255$ para imágenes de 24-bits.

Las figuras 4.3, 4.4 y 4.5 muestran tres ejemplos de imágenes con las texturas objeto de la identificación, a saber textura vegetal, suelo y cielo. Además se presentan bajo distintas condiciones de iluminación como es el caso de las imágenes de las figuras 4.3 y 4.4 obtenidas con iluminación suficiente, mientras que la imagen de la figura 4.5 se ha capturado con iluminación relativamente escasa.



Figura 4.3.- Imagen original mostrando sólo el suelo



Figura 4.4.- Imagen original mostrando el suelo y el cielo con mucha iluminación



Figura 4.5.- Imagen original mostrando el suelo y el cielo con poca iluminación

4.1.1.2. Índices cromáticos

Como hemos comentado, las imágenes analizadas contienen tres firmas espectrales dominantes, verdes para las plantas, rojas para las tonalidades del suelo y azules para el

cielo. Por lo tanto, utilizamos índices de vegetación para las plantas y con predominio de rojos para la identificación de suelos, ya que acentúan el verde y el rojo respectivamente. Con tal fin, se seleccionaron ExG, CIVE, ExGR y VEG para los verdes y ExR para los rojos, debido a sus buenos resultados demostrados en la literatura y en los diversos experimentos realizados previamente. Se ha trabajado con el índice NDI, que también está basado en componentes cromáticas del espectro visible ya que las plantas y el suelo aparecen muy contrastados, por el hecho de que sólo utiliza los canales verde y rojo, en su definición. No obstante, con dicho índice no se obtuvieron resultados satisfactorios ya que identificaba una gran cantidad de píxeles que corresponden al cielo como verdes. Como se ha mencionado previamente, las tonalidades azules se detectan mediante la adaptación de un índice cromático especialmente diseñado a partir del ExR, utilizado para el suelo. Esto permite la detección de píxeles pertenecientes al cielo con un cierto grado de robustez, donde las imágenes resultantes se denominan ExB por analogía con el nombre de dicho índice. La ventaja de utilizar índices cromáticos es que acentúan un color particular, tal como el verdor de la planta. Estos índices son los definidos en la tabla 3.2.

4.1.1.3. Combinación de los índices cromáticos

Durante los diferentes experimentos llevados a cabo para la extracción de verdes se utilizaron de forma individual los índices de vegetación *ExG*, *ExGR*, *CIVE* y *VEG*, definidos en la tabla 3.2 donde *ExG* proporciona el exceso de verde, *ExGR* el exceso de verde menos el exceso de rojo, *CIVE* es el índice de color de la extracción de vegetación y *VEG* fue diseñado para hacer frente a la variabilidad de la iluminación de la luz del día. Debido a que todos tienen información del verdor y acentúan la banda espectral con verde dominante, las plantas existentes en las imágenes originales aparecen con niveles de intensidad mayor que el resto en las segmentadas y, además, se puede ver que tienen un alto contraste debido a los diferentes grados de verdor mostrados por cultivos y malezas y también debido a las sombras, que varían dependiendo del estado de crecimiento de las plantas. Si bien, los resultados obtenidos con su utilización individual apuntaron la posibilidad de posibles mejoras de cara a una mayor efectividad en la detección de los verdes. Siguiendo esta línea de investigación, la propuesta formulada en este trabajo consiste en extraer de cada uno de ellos la máxima información posible a nivel individual para realizar un proceso de combinación de los mismos otorgando mayor peso específico a aquellos con mayor

efectividad a nivel individual, a la vez que en conjunto se refuerzan esos comportamientos individuales.

Las figuras 4.6. (a) y (b) muestran los resultados correspondientes a la aplicación de los índices *ExG* y *ExGR* respectivamente en la misma imagen.

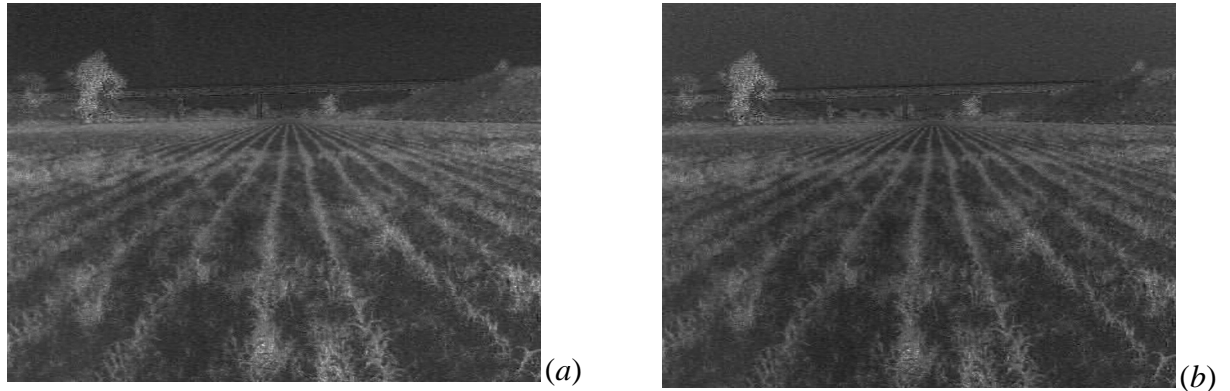


Figura 4.6.- Diferentes grados de verdes (a) *ExG*; (b) *ExGR*

Las imágenes de la figura 4.6 se han obtenido aplicando las ecuaciones (3.9) y (3.12) sobre la imagen original mostrada en la figura 4.4. Como puede observarse, las plantas se distinguen por su mayor intensidad con respecto al resto, quedando tanto el cielo como el suelo con una tonalidad bastante más oscura. Los resultados obtenidos aplicando ambos índices cromáticos son muy similares entre sí, la diferencia estriba en la escala de grises que define los diferentes verdes. No obstante, basándonos en esta observación, podemos determinar que el mejor índice será el que ofrezca mayor contraste, ya que su contribución a la combinación será más relevante.

El concepto de contraste puede ser medido basándonos en el histograma de la escala de grises, que proporciona dos medidas como son la uniformidad y la entropía (Gonzalez y col., 2004 y Gonzalez y Woods, 2008). La uniformidad es máxima cuando todos los niveles de gris son iguales (máximo uniforme) y disminuye a partir de ahí. La entropía es una medida característica de los procesos aleatorios, que es mínima cuando todos los niveles de gris son iguales. Por lo tanto, en nuestra aproximación buscamos uniformidad mínima y máxima entropía, ya que proporcionan un alto poder discriminante. En nuestros experimentos hemos verificado que las dos medidas tienen un comportamiento similar para el conjunto de imágenes utilizadas, por lo que finalmente hemos elegido el menos costoso

computacionalmente hablando, en este caso, la uniformidad, que se calcula según la expresión (4.3),

$$U = \sum_{i=0}^{L-1} p^2(z_i) \quad (4.3)$$

donde z es una variable aleatoria que determina el nivel de gris y $p(z_i)$ representa el histograma, $i = 0,1,2,\dots,L-1$, donde L es el número niveles de gris distintos. Como nuestra representación de la imagen original es de 8 bits en escala de grises, la representación máxima es de $2^8 = 256$, y por tanto $L = 256$. Los valores normalizados de la escala de grises varían de forma continua en el intervalo $[0,1]$, dividimos este intervalo en 256 subintervalos etiquetados como i y asignamos a cada nivel de gris de la imagen de entrada el límite superior del subintervalo correspondiente. Este valor corresponde al z_i necesario para calcular el histograma $p(z_i)$.

4.1.1.4. Cálculo de la relevancia de los índices de color verde (vegetación)

Teniendo en cuenta que la uniformidad mínima es considerada mejor que la máxima, calculamos dicha uniformidad para cada imagen resultante de la aplicación de los índices de vegetación con U_{G_k} , según las ecuaciones (4.4), donde

$$\Omega = \{G_k \mid k \equiv ExG, ExGR, CIVE, VEG\},$$

teniendo en cuenta que las uniformidades se normalizan en el intervalo $[0,1]$. Obtenemos a partir de aquí los pesos asociados a cada imagen resultante, donde el peso máximo se corresponde con la uniformidad mínima, teniendo en cuenta el complemento con respecto a la unidad y una nueva normalización al rango $[0,1]$. Estos pesos determinan la relevancia de cada una de las imágenes a combinar o fusionar.

$$U_{G_k} = (1 - U_{G_k}) / \sum_{G_h \in \Omega} U_{G_h} ; w_{G_k} = U_{G_k} / \sum_{G_h \in \Omega} U_{G_h} \quad (4.4)$$

A partir de las cuatro imágenes disponibles, y sus respectivos pesos asociados, la combinación se lleva a cabo a partir del cómputo de la media ponderada de las imágenes con dichos pesos a través de la expresión (4.5)

$$G = \sum_{\Omega} w_{G_k} G_k \quad (4.5)$$

4.1.1.5. Umbral de la imagen

Llamamos G a la imagen combinada resultante, las tonalidades de gris correspondientes a los verdes en la imagen combinada se han acentuado en relación con las otras dos bandas espectrales (R y B) y con respecto a las imágenes resultantes individuales que han sido combinadas. El próximo objetivo será aislar las partes verdes con el fin de aprovechar esta mejora así obtenida.

Como se mencionó anteriormente, varios de los métodos de umbralización estudiados en la literatura se han aplicado para alcanzar este objetivo. Puesto que la idea es la automatización del proceso, buscamos un enfoque para la umbralización también de naturaleza automática. Uno de los posibles candidatos es el método de Otsu (1979), que ya se ha aplicado en el trabajo de Meyer y Camargo (2008), obteniendo unos buenos resultados. Este método estima el mejor umbral a partir del histograma de intensidades, de la imagen G . Sin embargo, también como se ha mencionado previamente, el método de Otsu en algunas ocasiones tiende a producir una infra-segmentación (Guijarro y col., 2011), ya que proporciona un valor superior al valor medio o media estadística del propio histograma. Esto hace que muchos de los píxeles verdes lleguen a ser descartados del proceso, perdiendo así una valiosa información sobre las texturas vegetales. Esta es la razón por la cual en la propuesta que formulamos proponemos la utilización de la media como umbral para la binarización de las partes correspondientes a los verdes en lugar del umbral proporcionado por Otsu. Las figuras 4.7(a) y (b) muestran el resultado de la binarización de la imagen G a partir de los umbrales de histograma mediante la media y Otsu, respectivamente. La figura 4.7(c) muestra las diferencias entre ambas imágenes, 4.7(a) y 4.7(b), siendo los píxeles blancos los correspondientes a los verdes extraídos por la media y ocultados por Otsu. Calculamos el valor medio de z (media de los niveles de grises) mediante la ecuación (4.6).

$$m = \sum_{i=0}^{L-1} z_i p(z_i) \quad (4.6)$$

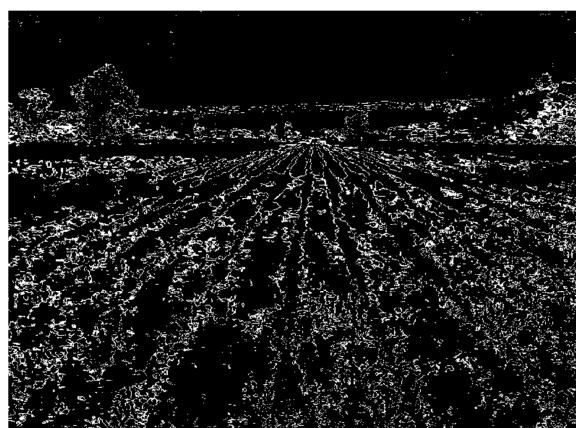
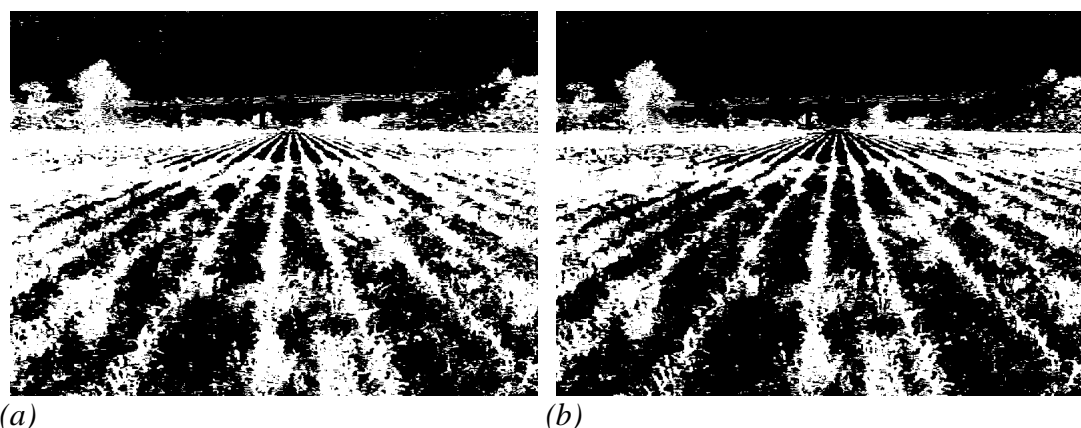


Figura 4.7.- (a)-(b) Imágenes binarias obtenidas mediante la aplicación del histograma de la media y Otsu, (c) imagen obtenida como diferencia entre las imágenes (a) y (b) donde los píxeles blancos corresponden a los verdes ocluidos por Otsu

En las imágenes anteriores, los valores obtenidos para los umbrales proporcionados por la media y por Otsu son 0,2496 y 0,2706, respectivamente, teniendo en cuenta el rango de valores de niveles de gris en el intervalo $[0,1]$. Este comportamiento, es general para el conjunto de imágenes utilizadas, tal y como se ha indicado previamente.

Por último, aplicando la binarización por la media obtenemos la imagen binaria G_b (figura 4.2), imagen resultante de la combinación de verdes binarizada por la media, donde los píxeles asociados con la vegetación se etiquetan con un valor binario de uno en contraposición al resto cuyo valor binario es cero.

4.1.1.6. Detección de los azules

Aplicando el índice cromático ExB (tabla 3.2) el objetivo es obtener una imagen binarizada para los azules. Como se explica más adelante en la siguiente sección, hay que

tener en cuenta si las imágenes tienen o no cielo en la parte superior, puesto que los valores de azul se ven afectados y esto puede cambiar las decisiones tomadas con respecto al suelo o a los verdes. Por lo tanto, el objetivo es separar sólo las zonas con alto contenido de azul, lo cual se logra mediante el enfoque de umbralización de Otsu por las razones anteriormente expuestas.

En la figura 4.8, podemos ver el resultado de esta binarización. En la imagen (a) se muestra el resultado obtenido al aplicar el índice cromático ExB; y tanto en (b) como en (c) se muestran las imágenes binarias obtenidas mediante la aplicación de la media y el método de Otsu. Como se observa en la imagen (b) una parte importante de suelo se ha etiquetado como de color azul lo que no ocurre en (c), tal y como se esperaba. Este es un comportamiento común que se observa en el conjunto de imágenes analizadas.

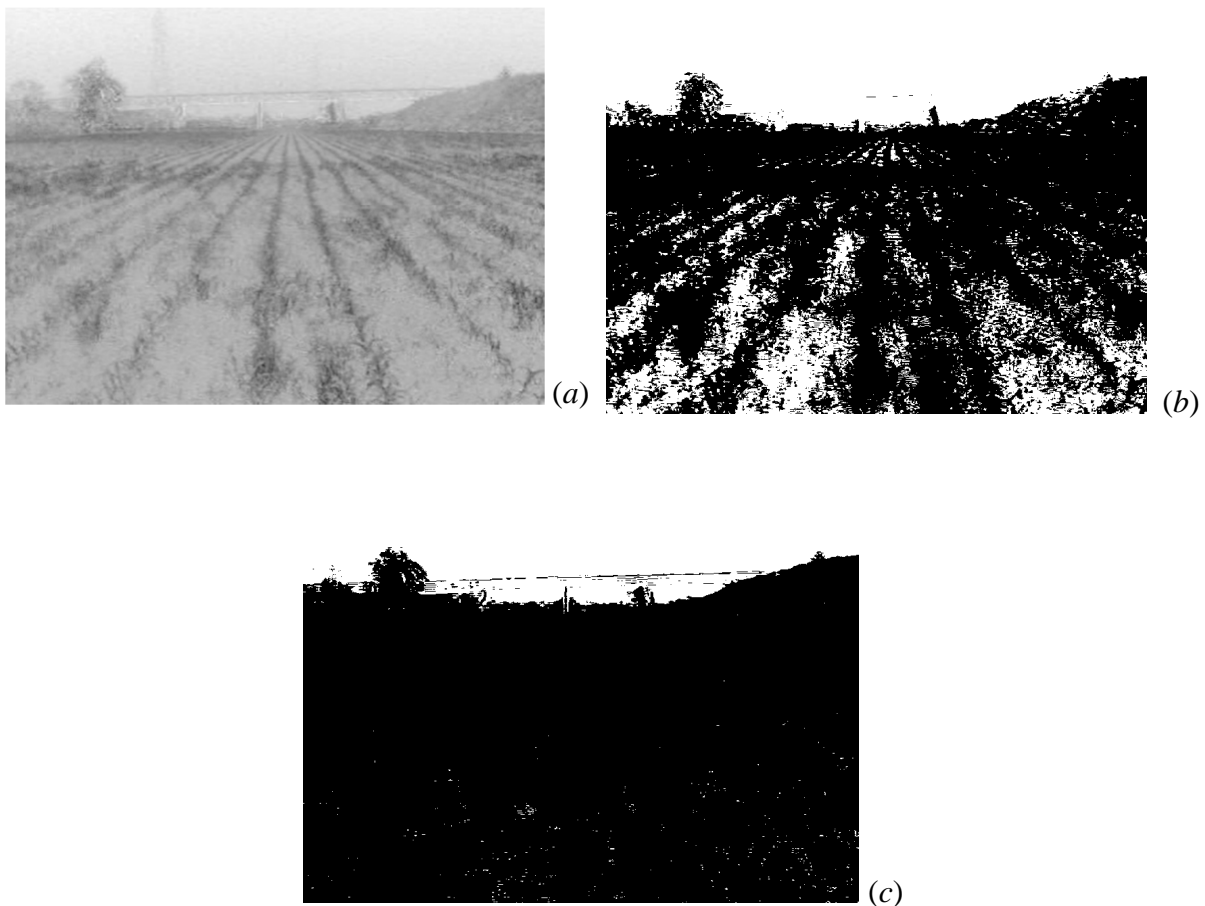


Figura 4.8.- (a) ExB imagen obtenida con la ecuación (3.11); (b) (c) imágenes binarias de la imagen de (a) obtenidas mediante la aplicación de la media y Otsu respectivamente

4.1.1.7. Detección de los rojos

Aplicando el índice cromático ExR (tabla 3.2) obtenemos la imagen binarizada del suelo. Al contrario que con el azul, ahora el objetivo es extraer el rojo, el cual está íntimamente relacionado con el suelo, por tener éste una tendencia cromática dominante hacia el rojo; por lo tanto, ya que el suelo es una parte relevante en los campos de cultivo, ahora tratamos de conservar el mayor número de los píxeles que son candidatos potenciales para el suelo. Esta detección es muy parecida al procedimiento realizado para la vegetación, por lo que aquí de nuevo se aplica la media como el mejor valor umbral para la binarización. En la figura 4.9(a) se muestra la imagen obtenida a partir de ExR y en la figura 4.9(b) la imagen binaria mediante la aplicación del criterio ya establecido.



(a)



(b)



(c)

Figura 4.9.- (a) ExR imagen obtenida de la ecuación (3.10); (b) imagen binaria obtenida aplicando el criterio de la media; (c) diferencias entre las imágenes binarias dadas en la figura 4.8.(b) y 4.9.(b)

4.1.2. Clasificación mediante agrupamiento borroso

El próximo objetivo es la identificación de las diferentes clases de texturas para los verdes y también diferentes clases de texturas pertenecientes al suelo.

Tanto con el azul como con el rojo surge un problema importante que hace necesario aislar los píxeles identificados como pertenecientes al suelo. Para solucionarlo, clasificamos los tipos de imágenes en dos grupos, aquellas que no tienen componente cielo y aquellas otras imágenes que sí lo poseen, tal y como se muestra en las figuras 4.3 y 4.4. Si extraemos los azules y rojos del segundo tipo de imágenes, obtenemos una clara separación entre las componentes del cielo y el suelo, como se muestra en las figuras 4.8(b) y 4.9(b) respectivamente. Pero si el azul y rojo se extraen para el primer tipo de imágenes nos encontramos con una importante cantidad de solapamiento entre ambos colores. Esto se puede observar en las imágenes de las figuras 4.10(a) y (b). Si calculamos la diferencia entre las imágenes binarias obtenidas a partir de los índices ExB y ExR el resultado de las imágenes sin cielo se muestra en la figura 4.10(c), donde podemos observar que la proporción de píxeles blancos es insignificante en comparación con la imagen completa. Por el contrario, si las imágenes contienen cielo, la diferencia entre las imágenes binarias obtenidas a partir de los índices ExB y ExR genera la imagen que aparece en la figura 4.9 (c), donde se puede observar que la relación entre los píxeles blancos en comparación con el total de la imagen es bastante elevada. En base a este estudio, se calcula un factor de proporción según la relación (4.7).

$$r = \frac{N_w}{N_T} \quad (4.7)$$

donde N_w es el número de píxeles con valores comprendidos entre +1 o -1 en la imagen Y obtenida como la diferencia entre las imágenes binarias de ExR y ExB ; N_T es el número total de píxeles de la imagen. Por último, con el fin de aislar los píxeles que se consideran como pertenecientes al suelo se aplica la siguiente regla de decisión (4.8).

$$\begin{aligned} & \text{if } r < T \text{ then} \\ & \quad S_b = ExR_b \wedge \neg ExG_b \\ & \text{else} \\ & \quad S_b = ExR_b \wedge \neg ExG_b \wedge \neg ExB_b \end{aligned} \quad (4.8)$$

donde el subíndice b representa la imagen binaria y los símbolos \wedge y \neg son las operaciones lógicas *and* y *not* sobre las imágenes binarias; S_b es la imagen binaria final que contiene los píxeles del suelo etiquetados como uno (blanco), y por último T es un umbral fijado a 0.6, valor obtenido mediante diversos experimentos de ensayo y error con el conjunto de imágenes disponibles.

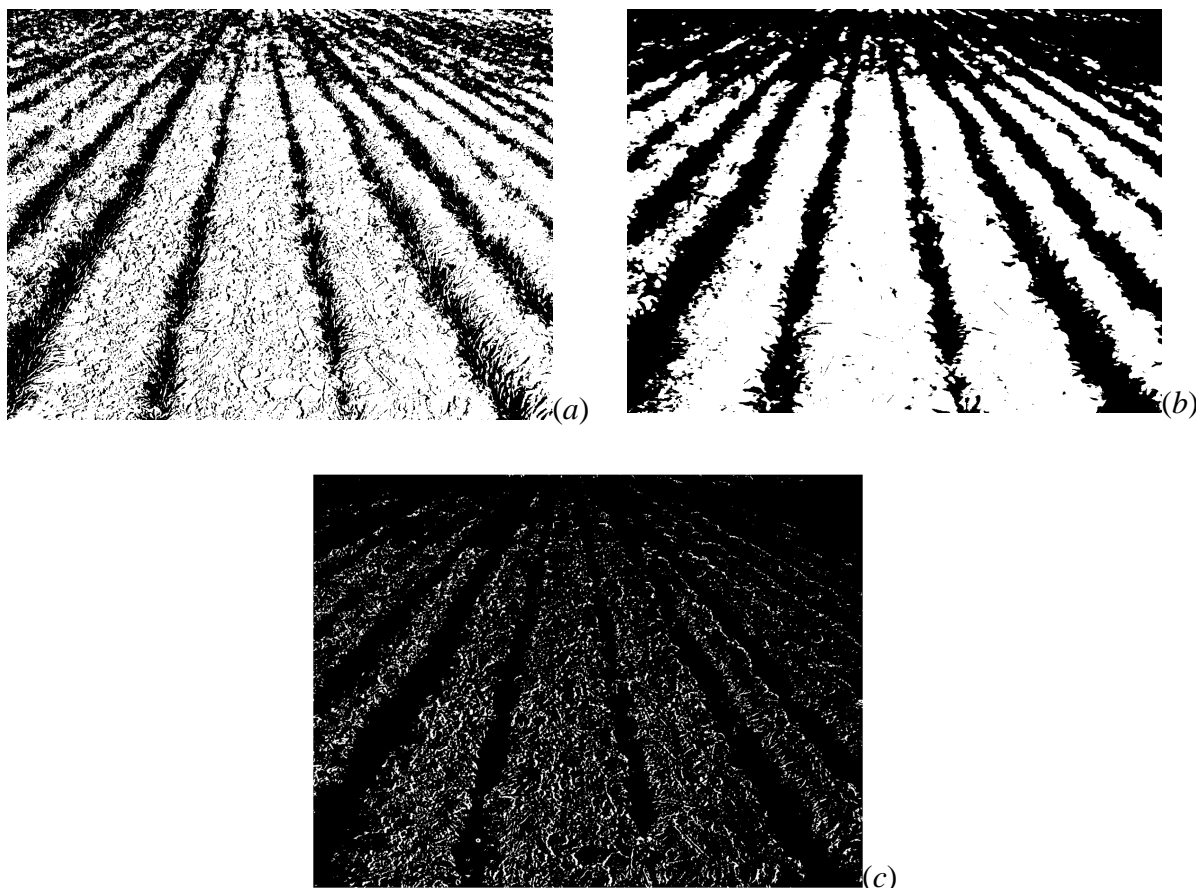


Figura 4.10.- Imágenes binarias obtenidas desde la imagen original de la figura 4.3 (a) ExB ; (b) ExR ; (c) diferencia entre la binarización de imágenes ExB y ExR

Por lo tanto, tenemos disponibles dos imágenes binarias principales, G_b y S_b que representan los grados de verde y de rojo en las imágenes originales, respectivamente. Con base a estas dos imágenes extraemos dos conjuntos de muestras de la imagen original identificándolas como X_G y X_S . El primer conjunto, X_G , contiene muestras representativas de los píxeles de la imagen original que han sido identificados como píxeles verdes en la imagen G_b y el segundo contiene los píxeles que representan al suelo. Sin pérdida de generalidad, representamos los dos conjuntos denominándolos X , con n muestras de entrenamiento, es decir, $X = \{x_1, x_2, \dots, x_n\} \in \mathcal{R}^d$, donde d es la dimensionalidad de los datos. Cada muestra se asigna a una clase determinada w_j , donde el número de clases es c , con, $j =$

1, 2, ..., c. Cada nueva muestra identificada por el vector \mathbf{x}_i representa un píxel de la imagen, donde sus componentes son las componentes RGB de ese píxel en la ubicación de la imagen (x, y) dada por G_b y S_b , por lo que en nuestros experimentos la dimensionalidad de los datos es $d = 3$.

Este proceso recibe las muestras de entrada \mathbf{x}_i para el entrenamiento y establece una partición, suponiendo que el número de clases c es conocido. El proceso calcula para cada muestra \mathbf{x}_i en cada iteración t , su grado de pertenencia a la agrupación w_j (μ_i^j) y actualiza los centros de los clústers de la siguiente manera (Zimmermann, 1991):

$$\mu_i^j(t+1) = \frac{1}{\sum_{r=1}^c (d_{ij}^2(t)/d_{ir}^2(t))^{2/(b-1)}}; \mathbf{v}_j(t+1) = \frac{\sum_{i=1}^n [\mu_i^j(t)]^b \mathbf{x}_i}{\sum_{i=1}^n [\mu_i^j(t)]^b} \quad (4.9)$$

siendo $d_{ij}^2 \equiv d^2(\mathbf{x}_i, \mathbf{v}_j)$ la distancia Euclídea al cuadrado. El número b es el peso exponencial (Bezdek, 1981; Duda y col., 2000), $b > 1$. El criterio de parada viene determinado por $\|\mu_i^j(t+1) - \mu_i^j(t)\| < \varepsilon \forall ij$ o por un número máximo de iteraciones t_{max} .

El método requiere la inicialización de los centros de los clusters, de modo que la ecuación (4.8) se puede aplicar en la iteración $t = 1$. Para ello, aplicamos el siguiente procedimiento:

- 1) **Realizar una transformación lineal** $Y = f(X)$ de la muestra de entrenamiento para que los valores varíen en el intervalo [0,1].
- 2) **Inicializar** $\mathbf{v} = 2D\bar{\mathbf{M}} \circ \mathbf{R} + D\bar{\mathbf{m}}$, donde $\bar{\mathbf{m}}$ es el vector de medias para las muestras de entrenamiento en Y ; $\bar{\mathbf{M}} = \max(abs(Y - \bar{\mathbf{m}}))$, con tamaño $l \times d$; $D = [1 \dots 1]^T$ con tamaño $c \times l$; \mathbf{R} es una matriz aleatoria de tamaño $c \times d$ con valores entre [0,1] y la operación \circ representa la multiplicación elemento por elemento.
- 3) **Proceso de fusión:** una vez que la clasificación se ha realizado mediante el clasificador basado en el agrupamiento borroso, tendremos la partición de las muestras de entrada y para cada clase w_j existirá un centro asociado \mathbf{v}_j . Consideramos como una única clase, aquellas dos clases que estén lo suficientemente próximas entre sí. El criterio para fusionar las clases se establece en base a la proximidad de los centros, por lo que dados

dos centros v_i y v_j representantes de sus respectivas clases, éstas se fusionarán si $d(v_i, v_j) < t_1$; donde d es la distancia Euclídea y t_1 es un umbral fijado en el 20% de la máxima distancia posible entre las muestras dentro de la clase, es decir, como los valores de los píxeles varían desde (0,0,0) (negro) a (1,1,1) (blanco) en el espacio de color RGB y su distancia máxima es la longitud de la diagonal principal en el cubo que representa este espacio. Cuando dos clases se fusionan, se calcula su centro para la nueva clase como la media entre todas las muestras pertenecientes a las dos clases que se han fusionado.

4.2. Método para la extracción de verdes mediante un índice cromático e información espacial

Como se mencionó previamente, el objetivo principal de la investigación reflejada en esta tesis es definir una estrategia automática para la segmentación de imágenes agrícolas con el fin de distinguir correctamente las texturas de las que se compone. Con la estrategia definida en el apartado anterior, basada en los índices cromáticos, se han obtenido muy buenos resultados, particularmente en lo que respecta a la detección de los verdes; si bien, en un avance posterior de investigación se ha determinado que aún es posible mejorar los resultados añadiendo a la información cromática información espacial, lo que se consigue mediante la transformada discreta de wavelets, que posee esta capacidad gracias a sus filtros de naturaleza paso alto. Esta estrategia se describe íntegramente en Guijarro y col. (2015).

La mayoría de las estrategias utilizadas para la segmentación de una imagen con vegetación abordan el problema de la identificación bajo la asunción de que las plantas muestran un alto grado de verdor. Por eso se han utilizado comúnmente los ya mencionados índices de vegetación con el objeto de extraer las texturas correspondientes a las plantas verdes. De hecho en la estrategia descrita en la sección 4.1 se utiliza una combinación de índices. Si bien en este segundo método utilizaremos sólo uno de ellos, en concreto el índice ExG introducido previamente por Woebbecke y col. (1995). La justificación para no utilizar la combinación previamente propuesta se fundamenta en el hecho de que la mejora producida por la combinación se ve ampliamente superada con la incorporación de la información espacial llevada a cabo mediante esta estrategia. Siguiendo con la línea de razonamiento y justificación de la propuesta, conviene reseñar el hecho de que los cultivos y

las malas hierbas en las imágenes agrícolas presentan una alta variabilidad espacial con una distribución irregular y aleatoria. En base a lo anterior, el enfoque propuesto consta de los siguientes pasos:

1. Extracción del verdor basada en ExG por las razones aducidas previamente.
2. Descomposición de la imagen resultante en los correspondientes coeficientes de detalle mediante la transformada de wavelets, aplicando los filtros base de Haar pertenecientes a la familia Daubechies. La aplicación de la transformada de wavelets, a la imagen resultante del paso anterior, permite la extracción de estructuras espaciales en tres bandas (horizontales, verticales y diagonales) que contienen información detallada.
3. Utilización de descriptores de textura para capturar la variabilidad espacial en las tres bandas.
4. Combinación de información de verdor y de textura espacial en la banda correspondiente a los coeficientes de aproximación resultantes del filtrado paso bajo mediante los filtros correspondientes de wavelets. Los coeficientes de aproximación se ven realzados mediante este procedimiento.
5. Aplicación de la transformada inversa de wavelets (TDIW) con la imagen realzada de los coeficientes de aproximación y los coeficientes de detalle, que no han sufrido modificación alguna.
6. Aplicación de un método de umbral, en este caso Otsu, sobre la imagen resultante de aplicar la TDIW para la identificación de las texturas de verdes resultantes.

Con este proceso de combinación de información espectral y espacial, la segmentación de los verdes resulta más efectiva. El enfoque propuesto es especialmente útil cuando la calidad del verdor de las imágenes es baja.

En la imagen de la figura 4.11 se muestra el proceso completo, de forma que tras la descomposición mediante wavelets se lleva a cabo la combinación de la información espectral y espacial sobre los coeficientes de aproximación. Hecho lo cual se aplica la TDIW, tal y como se ha explicado previamente, para obtener la imagen resultante, estos dos procesos se identifican en la mencionada figura con la etiqueta combinación de información, como síntesis de los mismos.

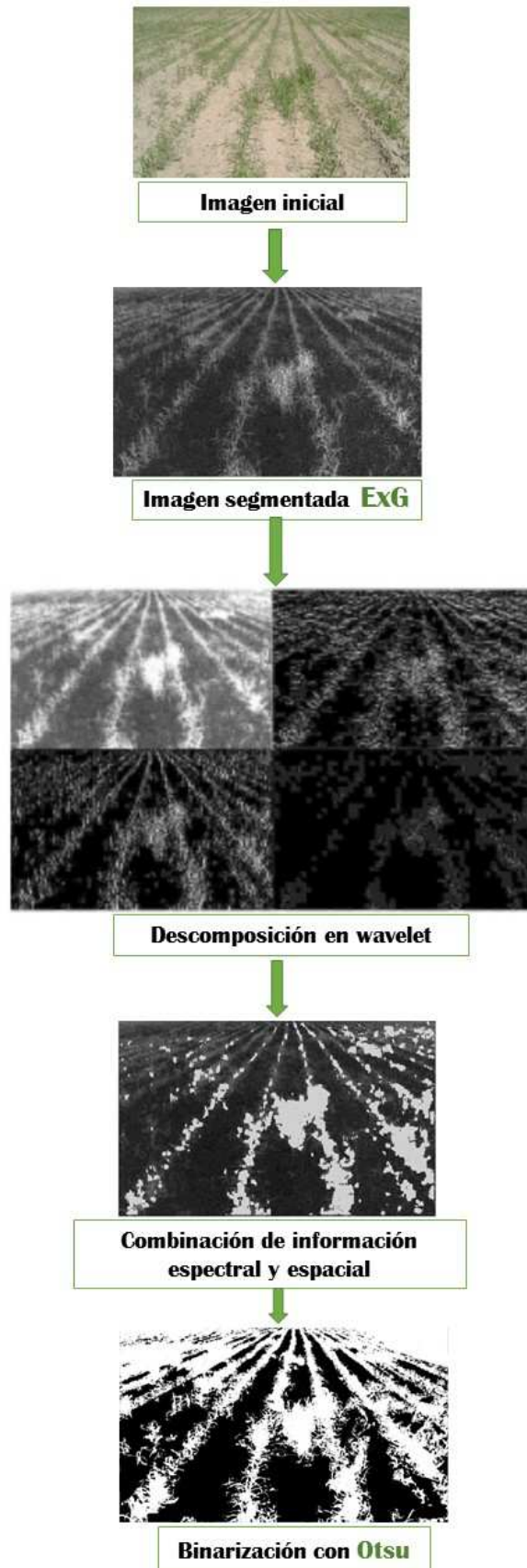


Figura 4.11.- Proceso completo de identificación de la capa vegetal

En la figura 4.12 presentamos tres imágenes seleccionadas del total del conjunto de imágenes utilizadas en los experimentos, de forma que son lo más representativas posibles del conjunto total, habiendo sido obtenidas bajo diferentes condiciones de iluminación sin llegar a la saturación. En dicha figura se muestra:

- a) Una imagen con partes brillantes y con alto contraste del verde, además de tener diversas sombras, posee una alta iluminación sin llegar a la saturación.
- b) Imagen con zonas del suelo saturadas o próximas a la saturación.
- c) Imagen oscura debido a la escasa iluminación del momento.



(a) Imagen con iluminación normal



(b) Imagen con alta iluminación



(c) Imagen iluminación pobre al anochecer

Figura 4.12.- Imagen original capturada con el sensor BASLER

En el procedimiento que aquí se describe, las imágenes para ser procesadas son las que contienen solamente suelo y plantas verdes, estas últimas formadas por cultivos y malas hierbas.

4.2.1. Segmentación de verdes

Al igual que en el procedimiento anterior, nos encontramos con un primer problema que es la segmentación de imágenes, en nuestro caso la segmentación de las partes verdes, uno de los mayores problemas en el campo de la Visión por Computador aplicada a la AP,

sobre todo en este tipo de imágenes reales, imágenes agrícolas, donde las condiciones de iluminación son variables. Tal y como se ha explicado previamente, los distintos autores han aplicado diversas estrategias para resolver el problema, tales como las basadas en el espectro visible, las que utilizan un umbral definido o las que se fundamentan en métodos de aprendizaje.

Partimos de una imagen válida, considerada como tal aquella que no posee niveles de saturación elevados, determinados porque la media del histograma de la imagen de intensidad no es elevada. En este punto se requiere extraer la vegetación, para ello se realizan diversos experimentos mediante la aplicación de los índices cromáticos vistos previamente (ExG, ExGR, CIVE y VEG) tal y como aparece en las imágenes 4.13 y 4.14.



Figura 4.13. Imagen en RGB



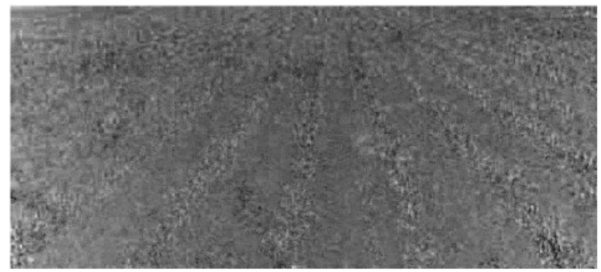
Exceso de verdes ExG



Diferencia entre verdes y rojos ExGR



“Color Index of Vegetation Extraction”-CIVE



Índice vegetativo VEG

Figura 4.14.- Resultado de aplicar los índices cromáticos (ExG, ExGR, CIVE y VEG) a la imagen 4.13.

En estas imágenes se puede apreciar fácilmente que las zonas pertenecientes a texturas vegetales, es decir las partes verdes, están representadas en partes claras y muy claras, mientras que el resto aparece en tonos más oscuros, llegando incluso al negro. En base a este criterio el índice elegido será aquel que ofrezca mayores contrastes. Basándonos en esto podemos observar que los mejores resultados los ofrece ExG, mientras que los peores resultados se obtienen con VEG.

Nos encontramos con dos problemas principales a la hora de determinar si un píxel es vegetación o no, por un lado la existencia de sombras producidas por las plantas en los días soleados y por las nubes en los días nublados, aunque este último hecho no es considerado como sombras sino como zonas de mayor oscuridad. En estas zonas los píxeles tienen unos valores muy bajos de intensidad en los tres canales siendo estos píxeles complicados de clasificar incluso para un experto, que podría concluir sobre su pertenencia a la clase de plantas o no, simplemente a partir de su experiencia o intuición. No podríamos implementar reglas en las que tuviéramos en cuenta todos los casos existentes, así como la aparición de las sombras debidas a las distribuciones irregulares de las plantas a lo largo del campo.

El otro problema al que nos enfrentamos son los reflejos por la incidencia de los rayos del sol, éstos hacen que los píxeles afectados tengan valores muy elevados de intensidad, aunque este problema se suele producir en menor proporción que el descrito previamente relativo a las sombras.

El análisis de grandes conjuntos de imágenes nos permite concluir que los diferentes tipos de enfoques de segmentación basados en los índices de vegetación, incluyendo ExG, producen resultados muy variables debidos a los problemas mencionados, es decir del tipo de imágenes y de las condiciones ambientales, por lo que a veces aparecerá sobre-segmentación (excesiva extracción de verdes) o infra-segmentación (pocos píxeles verdes extraídos), en ambos casos sin ninguna causa clara.

Basándonos en estas observaciones, podemos concluir que la segmentación basada únicamente en índices cromáticos, a veces no resulta suficiente o cuando menos es manifiestamente mejorable. Además, teniendo en cuenta las consideraciones anteriores, los resultados de aplicar los distintos índices y el coste computacional importante que ello

conlleva para sistemas de tiempo real, hemos optado por uno de los más sencillos con bajo coste computacional y relativa eficiencia, reduciendo con esta elección el número de operaciones, el elegido es el exceso de verde ExG, ecuación (3.9). La figura 4.15 muestra el resultado de aplicar ExG a la imagen original mostrada en 4.12.



Figura 4.15.- Índice ExG aplicado a la imagen original, imagen 4.12 a)

4.2.2. Descomposición en wavelets

En la figura 4.16 se muestra el esquema utilizado en la descomposición multiescala de la imagen utilizando la transformada de wavelets. Tomamos como referencia la imagen segmentada ExG a partir de la original $I(x, y)$, tal y como se ha explicado en el apartado anterior, aplicamos horizontalmente a cada fila un filtro paso bajo L y un filtro paso alto H produciendo cada uno de ellos las matrices de coeficientes $I_L(x, y)$ e $I_H(x, y)$, respectivamente. A continuación y verticalmente a cada columna en I_L e I_H se aplica un filtro paso bajo y paso alto, L y H, obteniéndose así las subimágenes AExG, HExG, VExG y DExG. (Mallat, 1989; Pajares y Cruz, 2004)

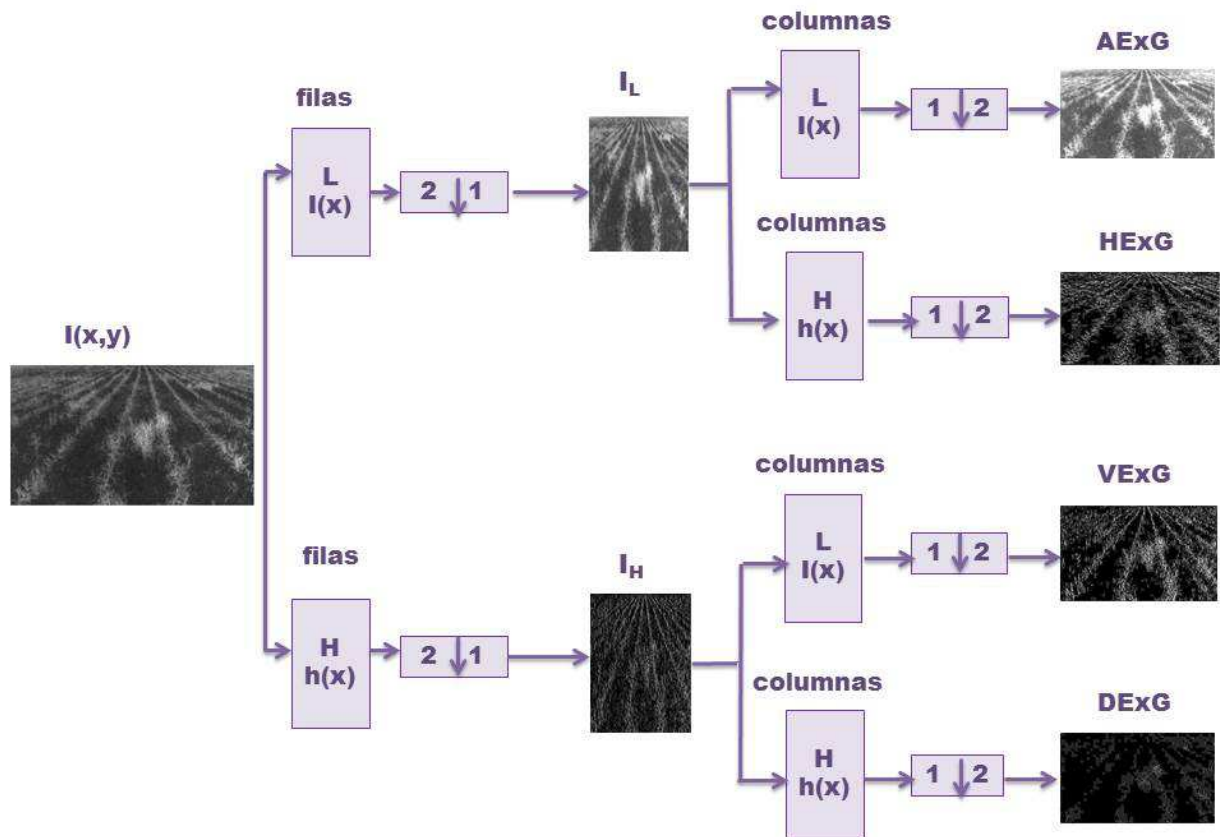


Figura 4.16.- Descomposición multiescala de una imagen mediante la transformada de wavelets directa

Como se ha comentado previamente, existen varios tipos de transformadas wavelets, denominados como familias wavelets, Daubechies (incluyendo Haar), biortogonal, Coiflet, Symlet, biortogonal reversa, etc... Estas familias se encuentran descritas en el tema tres y en la documentación asociada al toolbox de Matlab (The Mathworks, 2015). En base a los resultados obtenidos principalmente por Bossu y col. (2009) y Chou y col. (2007) referentes a los entornos agrícolas, en esta investigación se determina que la familia de Daubechies es la que mejor resultados produce con tamaños de los filtros relativamente bajos con $N=4$ y $N=6$ según la tabla 5.2. Esto está en consonancia con el hecho de Haar con $N=2$ captura bien las texturas según el descriptor SURF (Speeded-Up Robust Features), (Bay y col, 2008), basado en sumas de componentes wavelet Haar, que se demuestra que es simple, rápido y robusto. Si bien en nuestro caso los mejores resultados se obtienen con filtros con órdenes superiores debido a la alta resolución de las imágenes utilizadas.

Una vez que se obtiene la imagen ExG aplicando el mencionado índice cromático a la imagen original, se emplea la transformada discreta wavelet (DWT) para obtener los coeficientes de aproximación y los de detalle horizontales, verticales y diagonales, figura 4.17. De izquierda a derecha y de arriba abajo los coeficientes se identifican, respectivamente, como AExG, HExG, VExG y DExG.

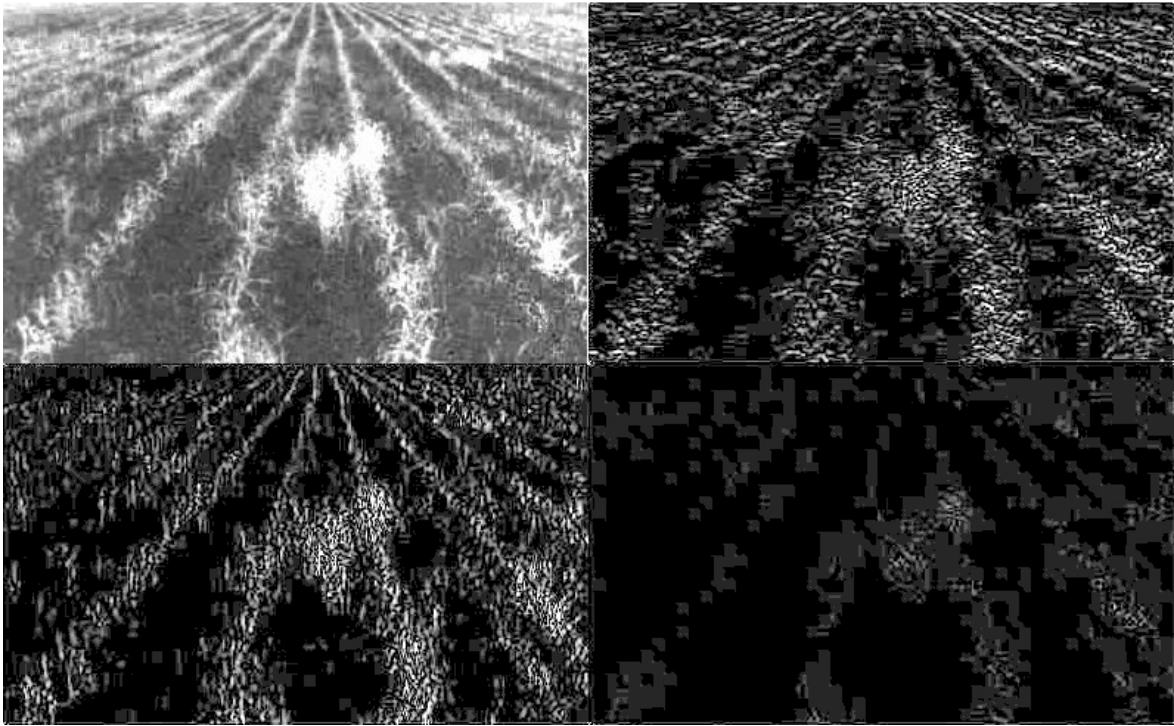


Figura 4.17.- Coeficientes de aproximación y detalle AExG, HExG, VExG y DExG

Los coeficientes de aproximación (AExG) tienen la información referente a las características más importantes obtenidas mediante un filtrado paso-bajo. Esto significa que estos coeficientes poseen información relevante procedente de la imagen original. Los coeficientes de detalle (HExG, VExG y DExG) muestran también variaciones espaciales de textura en diferentes orientaciones, tal y como proporciona la TDW, y que también contienen información importante a nivel de detalles.

4.2.3. Descriptores de textura

Basándonos en el razonamiento anterior, los descriptores estadísticos de primer orden son adecuados para el cálculo de la alta variabilidad de textura, una vez que, la transformada wavelets se ha aplicado para obtener las distribuciones espaciales de las

texturas. Estos descriptores se basan en el análisis de los niveles de gris del histograma de una imagen, en nuestro caso los niveles de gris son los valores de verdor de las imágenes ExG. Por lo tanto, sea z una variable aleatoria que denota la intensidad, es decir los valores de verdor en la imagen ExG, y sea $p(z_i)$ con $i= 0, 1, 2, \dots, L-1$ el histograma, donde L es el número de diferentes valores de verdor. Debido a la similitud de ExG con imágenes grises, definimos L , que puede tomar 256 valores en el rango $[0, 1]$ o equivalentemente en $[0, 255]$.

De lo expuesto anteriormente, es posible calcular el valor medio m de z , es decir, la media del verdor en una región caracterizada por su verdor puntual z_i con sus correspondientes probabilidades. Este valor viene dado por la expresión (4.10).

$$m = \sum_{i=0}^{L-1} z_i p(z_i) \quad (4.10)$$

El momento estadístico de orden n de z sobre la media se define teniendo en cuenta la expresión (4.11).

$$\mu_n(z) = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i) \quad (4.11)$$

A partir de la ecuación (4.11) se deduce que $\mu_0(z) = 1$ cuando todos los niveles de intensidad están involucrados en el cálculo, de forma que al tratarse de una distribución de probabilidad, la suma de todas las probabilidades debe ser la unidad, por ser una propiedad de las funciones de probabilidad. Además, $\mu_1(z) = 0$ como se puede deducir fácilmente a partir de la ecuación (4.11). Así, estos momentos no representan variabilidad de la textura.

Por el contrario, el segundo momento, que define la varianza, $\mu_2(z) = \sigma^2(z)$ es de particular importancia en las descripciones de la textura, ya que representa las desviaciones con respecto a la media de los valores de verdor en una región dada, es decir, mide el contraste de los valores de verdor en esa región. Por lo tanto, en una región homogénea, donde todos los píxeles poseen los mismos valores de verdor, la varianza será cero, mientras que en las regiones con alto contraste, el valor de la varianza será alto. En este trabajo, hemos utilizado la desviación estándar que es la raíz cuadrada de la varianza, porque tiene un comportamiento similar y tiende a ser más intuitiva, como se señala en González y Woods (2008). Los momentos estadísticos de orden superior al dos no muestran evidencias sobre la variabilidad de la textura.

Una vez obtenidos los coeficientes de detalle HExG, VExG, DExG, se calcula la desviación estándar para cada píxel i en su posición espacial (x,y) , obteniéndolos respectivamente para cada coeficiente de detalle $\sigma_i^H(x,y)$, $\sigma_i^V(x,y)$ y $\sigma_i^D(x,y)$, de acuerdo con la ecuación (4.11), identificándose con los superíndices H, V y D, que denotan coeficientes horizontales, verticales y diagonales respectivamente. Supongamos una ventana de tamaño 5×5 centrada en el píxel que se está considerando para el cálculo de la desviación estándar, la utilizada en nuestros experimentos, y también los valores de los coeficientes de detalle que toman valores en el intervalo $[0,1]$, los valores máximos y mínimos para la desviación estándar para cada píxel son 0,5099 y 0 respectivamente. Estos valores se obtienen como sigue: dada la ventana de 5×5 , es decir, con 25 píxeles, aparecen dos situaciones extremas. La primera ocurre cuando 13 de los 25 píxeles de la ventana (5×5) tienen el valor "1" y 12 píxeles poseen el valor "0", o viceversa, en estos casos se obtiene la desviación estándar máxima con el valor de 0.5099. El segundo se produce cuando todos los píxeles tienen los mismos valores, es decir, "1" o "0", la desviación estándar en este caso tiene el valor mínimo, 0.

El problema que se plantea ahora es cómo identificar o clasificar un píxel específico como perteneciente a una textura vegetal. Como se mencionó anteriormente, estos píxeles se caracterizan por una alta variabilidad en la distribución de los valores de verdor. Por lo tanto, los coeficientes de detalle, que en realidad provienen del resultado de aplicar el filtro de alta frecuencia, proporcionan información suficiente para identificar píxeles como pertenecientes a la cubierta vegetal, debido a que los detalles más significativos aparecen donde existe una alta variabilidad, ya sea en HExG, VExG o DExG. En las zonas más o menos homogéneas tales como las áreas asociadas con el suelo, esta variabilidad es mucho menor o incluso inexistente. La alta variabilidad de textura coincide precisamente con una desviación estándar elevada.

Así, bajo la suposición de que la textura de alta variabilidad identifica las áreas con muy alta probabilidad de existencia de plantas verdes, modificamos el valor del píxel i correspondiente al coeficiente de aproximación AExG en la posición (x, y) de la siguiente manera, dado el píxel i en AExG con un valor de verdor $g_i(x,y)$ se calcula $\sigma_i(x,y) = \max\{\sigma_i^H(x,y), \sigma_i^V(x,y), \sigma_i^D(x,y)\}$ y modificamos su valor a un nuevo valor de verdor según la expresión (4.12).

$$g_i^{new}(x, y) = g_i(x, y) + k\sigma_i(x, y) \quad (4.12)$$

Esta ecuación indica que el valor de verdor se incrementa en una magnitud representada por k veces la desviación estándar; por lo que un píxel con alta variabilidad, considerado como la vegetación, se incrementará en su valor de verdor y en los píxeles con baja variabilidad su textura será ligeramente modificada o quedará sin modificar. Tras esta modificación se obtiene un nuevo AExG donde los valores se escalan nuevamente para que tomen valores de nuevo en el rango $[0, 1]$ como en los coeficientes AExG de aproximación originales.

Tal vez, algunas áreas con densa vegetación podrían producir valores bajos en la desviación estándar debido a su homogeneidad y sus valores aparecerán sin modificar. Debido a la operación de escalado estos valores podrían aparecer con valores menores que el original y tras el proceso de binarización pudieran ser desechados como pertenecientes a las plantas verdes. Esta situación rara vez ocurre porque tales zonas originalmente muestran altos valores de verdor siempre con variaciones de textura relevantes debido a las diferencias en la iluminación natural aumentado en días soleados, sin afectar a los procesos de segmentación. Sin embargo para evitar modificaciones drásticas en los valores representativos del verdor, el coeficiente k , se ajusta a 0.5 tras diversos ensayos de prueba y error, la ecuación (4.12) trata de compensar este efecto adverso.

4.2.4. Combinación de información de verdor y textura

Al igual que antes, los píxeles de la nueva AExG siguen representando valores de verdor. Aplicamos la transformada discreta inversa de wavelets, TDIW a esta nueva AExG conjuntamente con los coeficientes de detalle HExG, VExG o DExG, que permanecen sin cambios. Después de esta transformación inversa, obtenemos la imagen ExG que contiene nuevos valores de verdor realizados y mejorados mediante la adición de la información espacial. Esta imagen es la que se utiliza como entrada para la binarización, que describiremos a continuación.

Una vez procesada la imagen completa correspondiente a los coeficientes de aproximación realizados, cuando algunos valores se modifican de acuerdo con el

procedimiento descrito anteriormente, se obtiene finalmente la imagen fusionada, con una cobertura de vegetación mejorada con respecto a la imagen original, tal como se muestra en la figura 4.18 que corresponde a la imagen mostrada en la figura 4.12(a).

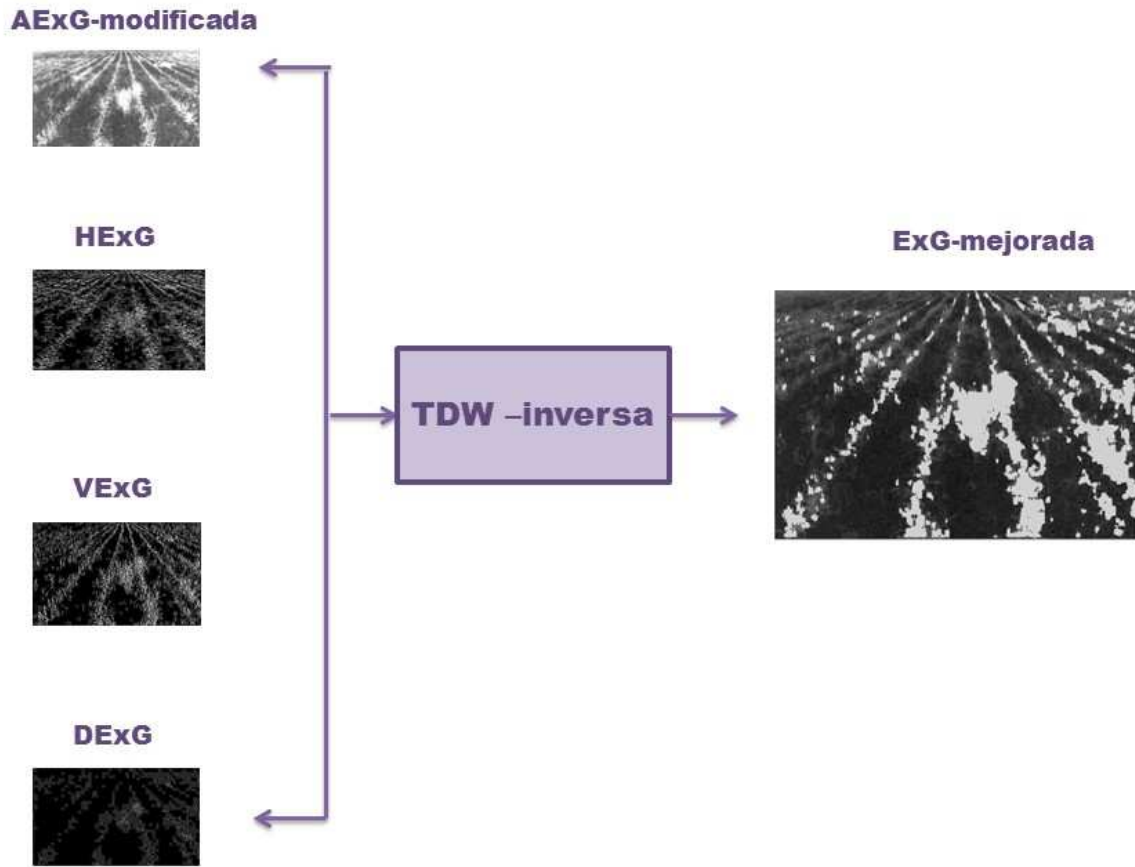


Figura 4.18.- Imagen con los verdes obtenida a partir de la nueva AExG con la TDIW

4.2.5. Binarización

Mejoradas las imágenes ExG mediante la aplicación de la TDIW, sobre la que se ha añadido la información espacial de la textura, se realiza el correspondiente proceso de binarización para distinguir entre plantas y otros elementos en el campo. Como se describió en la sección 2.3 se han propuesto varias técnicas automáticas basadas en el análisis de histograma con tal propósito, incluyendo el método de Otsu (1979) con doble umbralización desarrollado en Montalvo y col. (2012) o la media estadística en Burgos-Artizzu y col. (2011) y Guijarro y col. (2011). Este último trabajo, que da soporte a la investigación contiene los resultados que aparecen en él. Debido a que la incorporación de la información espacial realza las imágenes conteniendo el verdor, el método de Otsu no produce infra-

segmentación, evitando así la problemática referida en el proceso descrito en la sección 4.2, razón por la cual en este caso se ha utilizado finalmente Otsu. En la figura 4.19 se muestran respectivamente los resultados obtenidos tras aplicar Otsu a la imagen ExG sin mejorar y mejorada mediante el enfoque propuesto, observándose en el segundo caso una mejora significativa con respecto a la ExG sin realzado. Este comportamiento es característico en el conjunto de imágenes utilizadas.

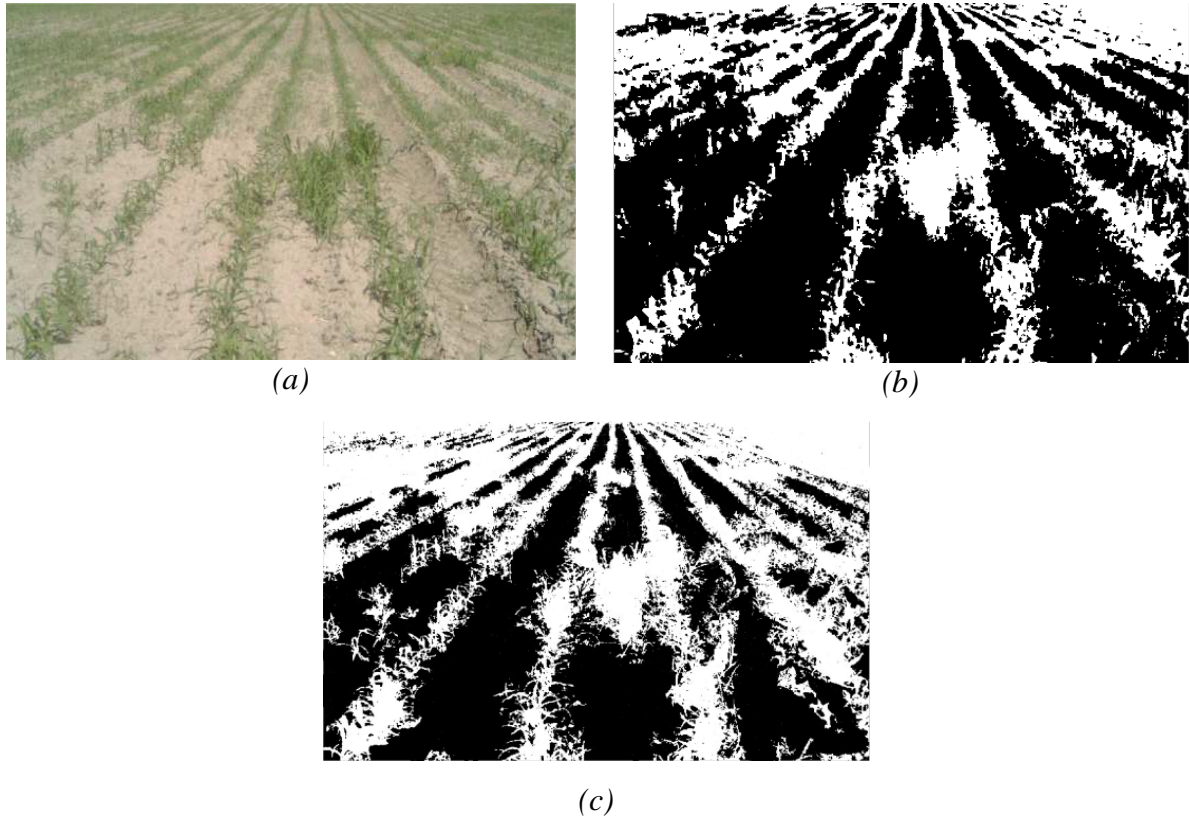


Figura 4.19. (a) *Imagen original;* (b) *Imagen binarizada aplicando el umbral de Otsu sobre la imagen ExG;* (c) *Imagen resultante de la binarización con Otsu, de la imagen ExG mejorada.*

5

Análisis de resultados

Presentadas las dos técnicas relativas a la segmentación de imágenes descritas en el capítulo cuatro, identificando las diferentes texturas tanto del suelo, cielo como de la cubierta vegetal, es indispensable el análisis de resultados obtenidos con los dos métodos propuestos, tanto desde el punto de vista cualitativo como cuantitativo.

5.1. Resultados de la segmentación automática combinando índices cromáticos y agrupamiento borroso

5.1.1. Imágenes reales de campos de cultivo

Las imágenes utilizadas para este trabajo pertenecen a dos tipos de cultivos de cereales y maíz. Para las capturas del primer grupo de imágenes se utilizó la videocámara de la compañía SONY, DCR PC 110E. Las imágenes digitales se almacenaron en el formato de 24 bits y en el espacio de color RGB en JPEG y con resoluciones de 1152×864 píxeles. Estas imágenes fueron tomadas durante los meses de abril y mayo desde 2005 a 2011 en un campo experimental de 1,7 hectáreas de cebada en La Estación de Investigación de Poveda, Arganda del Rey, Madrid. El segundo grupo de imágenes han sido adquiridas con una

cámara digital, HPR817, que captura imágenes de 800x600 píxeles, almacenándolas también en el espacio de color RGB con el formato JPEG, fueron tomadas durante diferentes días en los meses de abril y mayo de 2007 a 2012. Todas las capturas de imágenes fueron espaciadas en el tiempo dejando cinco o seis días entre toma y toma con el fin de disponer de imágenes obtenidas bajo diferentes condiciones de iluminación y diferentes estados de desarrollo de los cultivos y las malas hierbas.

Las imágenes han sido procesadas con el toolbox Image Processing de Matlab (The MathWorks 2015) específico para tratamiento de imágenes. El conjunto de imágenes que finalmente se ha procesado está constituido por 240 imágenes.

La figura 5.1, muestra una de las imágenes originales que pertenece a un campo de cebada (primer tipo), donde sólo se captura el suelo bajo una alta calidad de iluminación. También la figuras 5.2 y 5.3 pertenecen a un campo de maíz (segundo tipo), donde tanto el suelo como el cielo aparecen bajo diferentes condiciones de iluminación, con alto y bajo de contraste respectivamente, con la imagen de la figura 5.3 tomada al atardecer, es decir, con poca iluminación.

La estrategia propuesta consta de las dos partes principales ya mencionadas, a saber:

- a) Extracción de los verdes, rojos y azules,
- b) Identificación de los diferentes tipos de verdes y rojos.



Figura 5.1.- Imagen original capturada con un sensor mostrando sólo el suelo

Figura 5.2.- Imagen original mostrando el suelo y el cielo con mucha iluminación



Figura 5.3.- Imagen original mostrando el suelo y el cielo con poca iluminación

5.1.2. Extracción del verdor

Se trata de verificar la validez de la combinación de los índices *ExG*, *ExGR*, *CIVE* y *VEG*, que se comparan también cuando se utilizan de forma individual, refiriéndonos en este caso a ellos como métodos simples.

Las pruebas correspondientes se llevan a cabo utilizando un conjunto de imágenes de referencia o “ground truth”, que es como técnicamente se les conoce en la terminología especializada. Este conjunto está formado por imágenes ya clasificadas y binarizadas según el objeto a clasificar. Para la elaboración de las mismas procedemos como sigue. Como el objetivo principal de la investigación consiste en obtener los diferentes verdes de la imagen, lo primero que se hace es extraer de la imagen original, representada en el espacio de color RGB, todos los píxeles donde la componente verde es mayor que las restantes al menos en un valor del 10%. De esta forma obtenemos una parte importante de las plantas verdes, que son etiquetadas con un uno lógico, esto es en blanco. Seguidamente se extraen aquellos píxeles donde el rojo y el azul sean dominantes, también al menos en otra cantidad representada por el 10%, etiquetándose en este caso con el valor cero lógico, esto es en negro. El tercer paso consiste en refinar y clasificar de forma manual, bajo el asesoramiento de un experto, aquellos píxeles que aún no hayan sido etiquetados correctamente, comparándolos con la imagen original. La figura 5.4 es una de las imágenes del conjunto “ground truth” obtenida a partir de la imagen original mostrada en la figura 5.2 siguiendo el procedimiento anteriormente descrito.



Figura 5.4.- “Ground truth”

Como se ha mencionado previamente, se han utilizado 240 imágenes originales en los experimentos adquiridas del modo descrito anteriormente, donde 190 son imágenes que contienen cielo y las demás, es decir, 50, sin él. Aproximadamente, el 10% de cada grupo fueron adquiridas en condiciones de escasa iluminación. Se ha construido el “ground truth” a partir de 60 imágenes del conjunto que contienen cielo y 15 sin cielo, en ambos casos bajo diferentes condiciones de iluminación.

La tabla 5.1 muestra el porcentaje medio de error sobre el número de imágenes comparadas con el conjunto “ground truth” para *ExG*, *ExGR*, *CIVE* y *VEG* de forma individualizada y comparándolo con nuestro enfoque que combina los índices anteriores (COM).

| | <i>ExG</i> | <i>ExGR</i> | <i>CIVE</i> | <i>VEG</i> | <i>COM</i> |
|----------------------------|------------|-------------|-------------|------------|-------------|
| % de error para los verdes | 11.19 | 10.71 | 10.37 | 18.23 | 8.31 |

Tabla 5.1.- Porcentaje de error medio comparándolo con el conjunto de imágenes del “ground truth”

Mediante la ecuación (4.4), obtenemos cuatro pesos para cada imagen fusionada que se aplican en la ecuación (4.5) para calcular el resultado de la combinación de los índices cromáticos que extraen el verde. La media de los valores para el conjunto de las 240 imágenes son: $w_{G_{ExG}} = 0.25$, $w_{G_{ExGR}} = 0.30$, $w_{G_{CIVE}} = 0.33$ y $w_{G_{VEG}} = 0.12$. Si observamos estos resultados junto con los resultados mostrados en la tabla 5.1, se puede apreciar una relación directa entre el peso promediado y el porcentaje de error. A partir de estos datos podemos deducir que el método más relevante resulta ser *CIVE*, ya que su contribución a la imagen

fusionada resultante ha sido en promedio del 33%. El menos relevante ha sido VEG con el 12%. Probablemente, esto sea debido al parámetro a , que participa en la ecuación (3.14). Tras diversos experimentos, no se ha encontrado una clara relación entre este comportamiento y su influencia.

Podemos observar que hay una relación distinta entre los umbrales de Otsu y la media según sean las imágenes, con cielo o sin cielo. En las imágenes que contienen cielo, el valor de umbral de Otsu es, en promedio, un 16% mayor al valor de la media utilizada como umbral y en las imágenes sin cielo, disminuye hasta el 12%. Esto hace que si utilizáramos el umbral de Otsu, en lugar de la media, al ser aquel superior, consideraríamos muchos píxeles que son verdes como si no lo fueran, aumentando el porcentaje de error aproximadamente en un 5% en relación al alcanzado con nuestra propuesta de combinación COM.

5.1.3. Identificación de los diferentes tipos de verdes y rojos

Al aplicar el clasificador basado en el *agrupamiento borroso* a la imagen de la figura 5.2 obtenemos las figuras 5.5(a) y 5.5(b), fijando en tres el número de clases ($c=3$) para la imagen (a) y en dos ($c=2$) para la imagen (b). En (a) se observan los resultados obtenidos para la clasificación de las plantas verdes, conteniendo tres colores diferentes que nos sirven para etiquetar tres clases de plantas.

Cualquier observador humano, tras una simple observación le resulta fácil deducir que existen tres clases, donde una de ellas está asociada a una tonalidad verde oscuro que pertenece a los árboles y las sombras en los surcos. Las otras dos corresponden a dos tonalidades diferentes en las plantas. Como se muestra en la figura 5.5(b) el proceso de clasificación se ha llevado a cabo con el fin de obtener las dos clases mostradas que identifican dos clases de suelo.

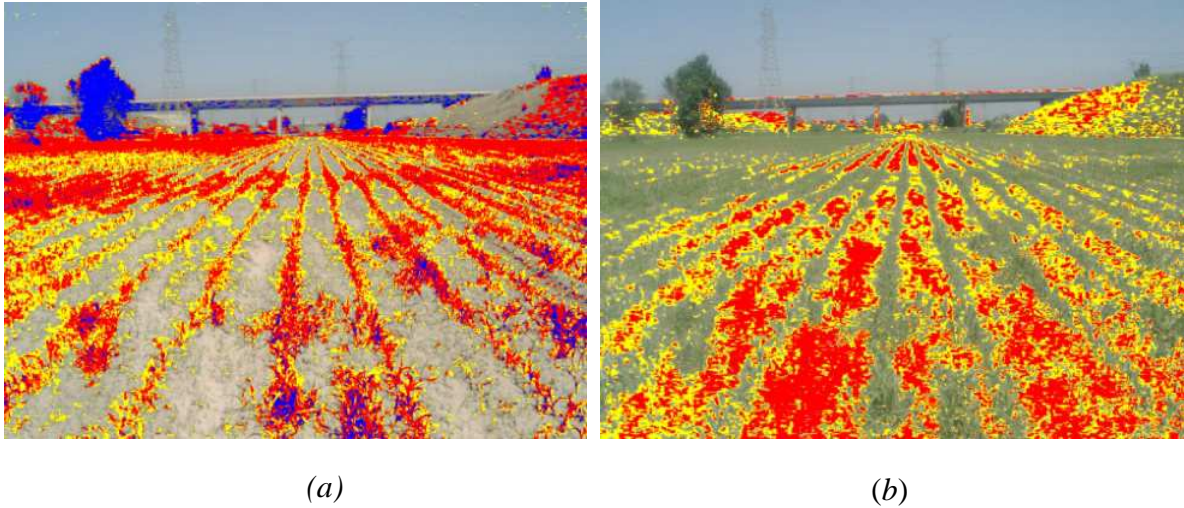


Figura 5.5.- Clasificación de imagen para la figura 5.2.: (a) Tres tipos diferentes de clases de plantas verdes; (b) 2 clases diferentes de suelo

Aplicando el mismo procedimiento a las imágenes sin cielo se obtienen los resultados mostrados en las figuras 5.6(a) y (b) sobre la imagen original de la figura 5.1 con $c = 2$. En (a) vemos los resultados obtenidos para las plantas verdes, donde la clasificación en este caso se ha realizado con las dos clases previstas. En (b) observamos los resultados de la clasificación de las zonas de suelo, también en dos clases. En ambos casos con resultados cualitativamente satisfactorios.

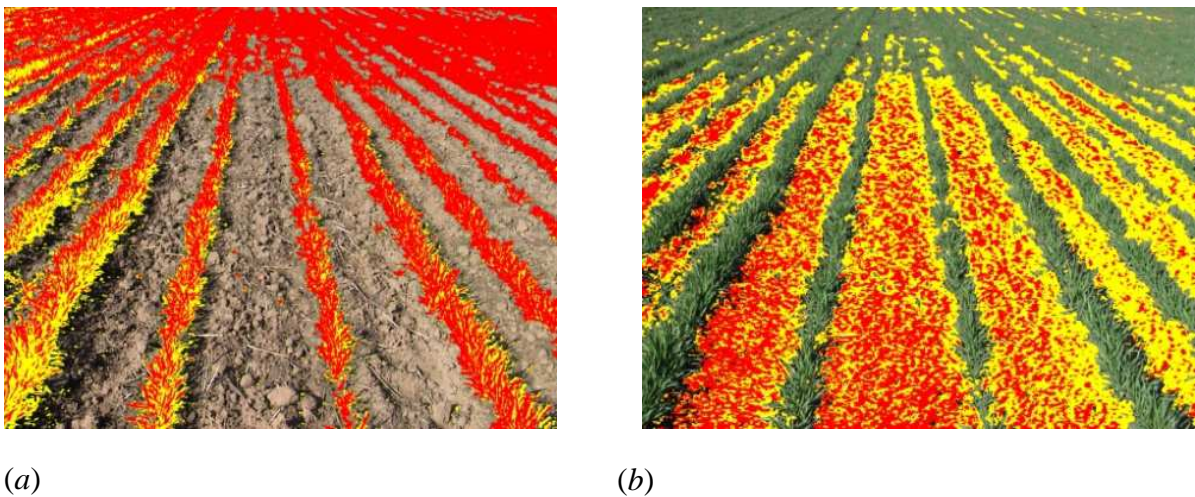


Figura 5.6.- Resultados de la clasificación de la imagen de la figura 5.1: (a) 2 clases distintas de plantas verdes; (b) 2 clases diferentes de suelo

La valoración de los resultados realizada previamente es aplicable al conjunto de imágenes disponibles, para las que se han obtenido similares resultados, confirmando de esta forma la validez de la propuesta para el conjunto de imágenes analizadas.

5.2. Resultados de la extracción de verdes mediante un índice cromático e información espacial

Para cumplir con los objetivos marcados en este procedimiento, cuyos resultados se encuentran detallados en Guijarro y col. (2015), primero realzamos los tonos verdes de la imagen aplicando los índices cromáticos, el resultado de esta etapa es la obtención de una imagen en tonos grises, figura 5.7(b), donde los píxeles verdes se etiquetan como blancos, mientras que los píxeles del suelo, con bajo contenido de verdes, se convertirán en negros. Con esta segmentación se obtiene una imagen en la que cada píxel posee un valor en la escala de grises, dentro del intervalo $[0,1]$, con 1 representando el blanco y el 0 el negro.

En esta sección se realiza un estudio comparativo con los índices cromáticos ExG, ExGR, VEG y CIVE, aplicados para la extracción del verde, definidos en la sección 3.2, validando y justificando la elección de ExG frente al resto. Decidido el índice cromático ExG descomponemos la imagen resultante de aplicar el índice mediante la TDW en sus coeficientes de detalle mediante la transformada de Haar. Con ello se realiza un estudio comparativo de los distintos resultados obtenidos al aplicar distintas familias de wavelets, a saber: Daubechies, Biorthogonal, Coiflets, Symlets y Discrete Meyer. Este análisis nos permite determinar la familia de Daubechies como la mejor por su buen comportamiento sobre el conjunto de imágenes analizadas.

Obtenida la imagen resultante tras la incorporación de la información espacial a la espectral, se procede a la binarización con Otsu, para lo que se proporcionan los resultados obtenidos. Como en el caso anterior, el método se ha contrastado con el “ground-truth”, base de imágenes binarizadas por un experto humano, que ha permitido comparar los resultados obtenidos de forma cuantitativa mediante el cálculo del porcentaje de error.

En la figura 5.7 se muestra un ejemplo ilustrativo de los resultados obtenidos por aplicación directa del ExG y binarización sin información espacial de textura y con ella. Este ejemplo permite visualizar claramente la mejora obtenida cuando se incluye información espacial, apreciándose los detalles de una forma relevante frente al resultado que no incorpora tal información.

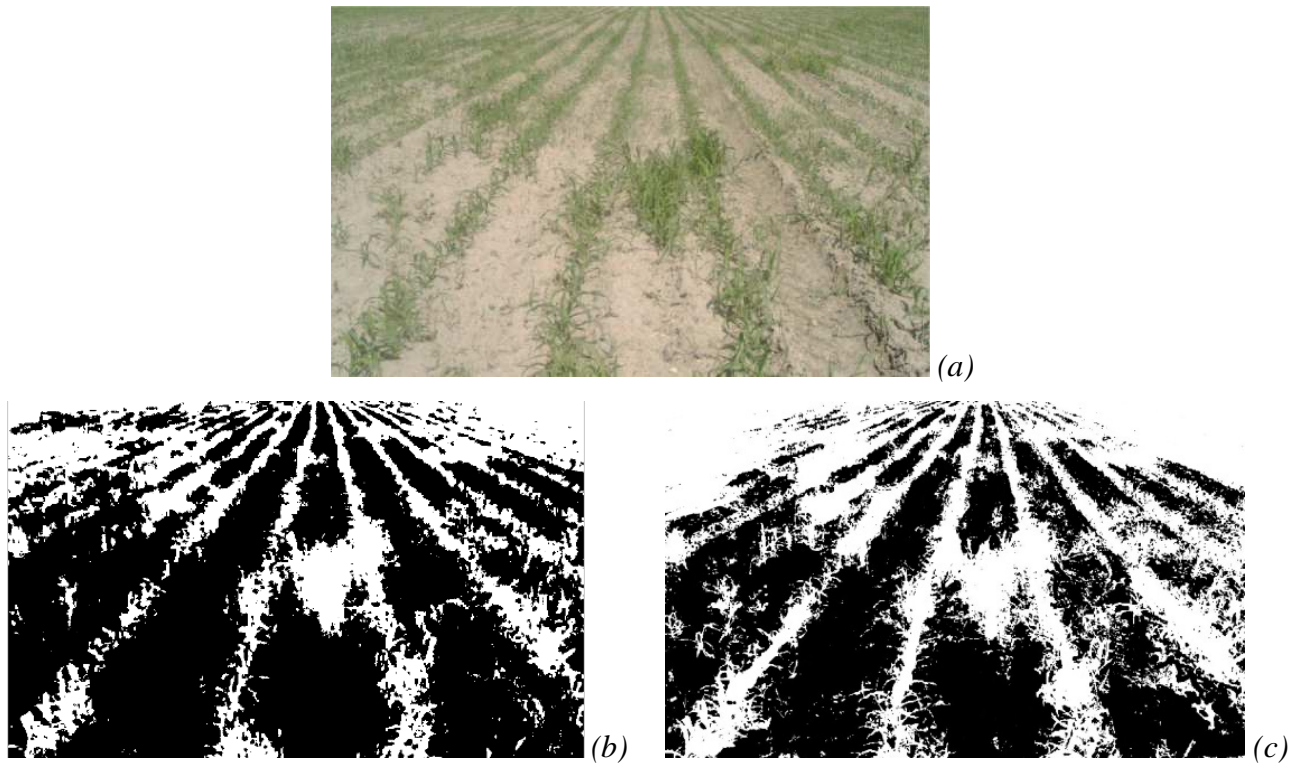


Figura 5.7.- Binarización de la imagen (a) mediante el umbral Otsu. (a)Imagen original (b) Información de los verdes utilizando ExG; (c) ExG con verdes e información de la textura

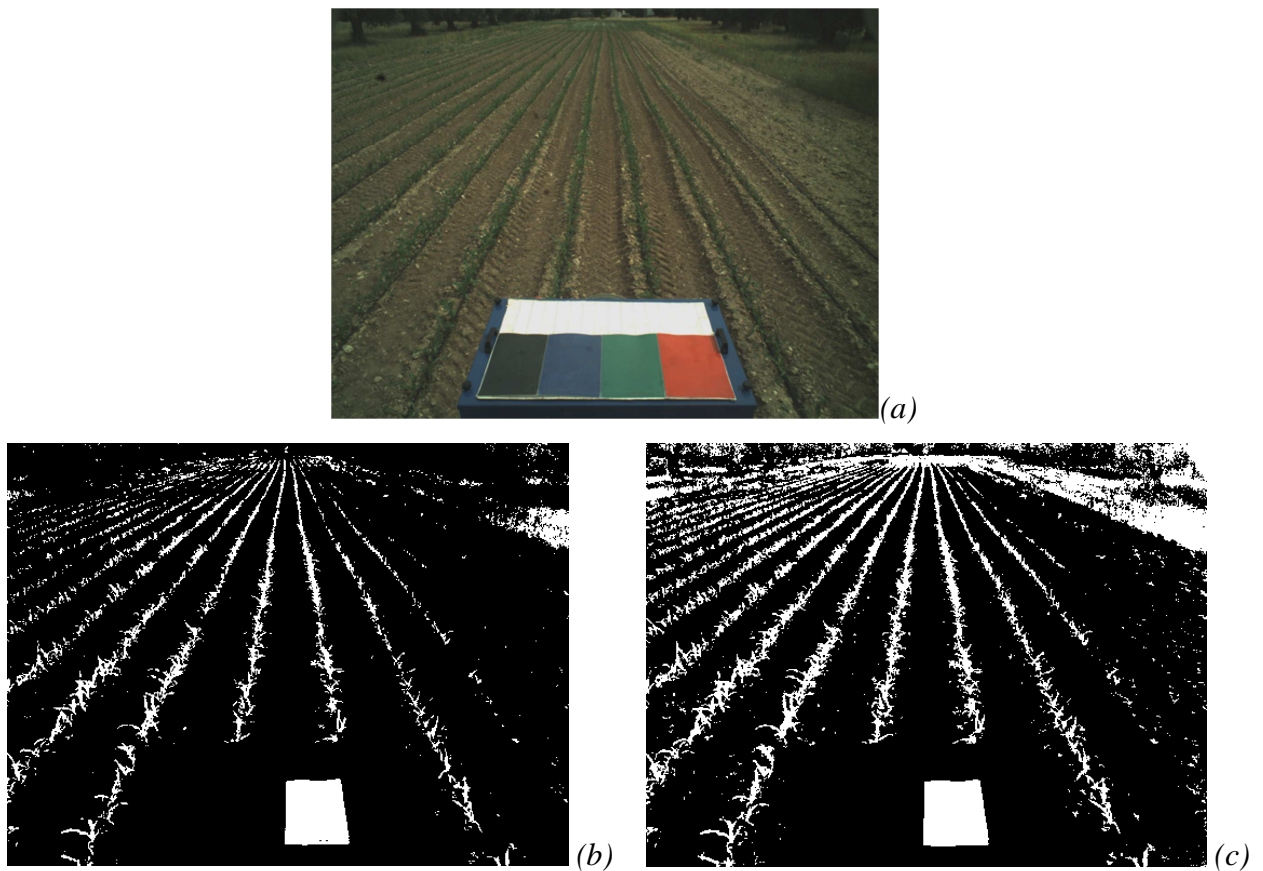


Figura 5.8.-. Binarización de la imagen (a) mediante el umbral de Otsu (a) Imagen original (b) Información de los verdes utilizando ExG (b) ExG con verdes e información de la textura

Las consideraciones y razonamientos que se derivan a continuación del análisis de los resultados son válidos para el conjunto completo de imágenes disponibles y analizadas. Observando los ejemplos de las imágenes originales, figuras 5.7(a) y 5.8(a), se determina que contienen las plantas verdes (malas hierbas y cultivos) distribuidas irregularmente en el campo con diferentes densidades en el espacio existente entre las líneas de cultivo.

También es posible observar los distintos grados de verdor entre las partes verdes. Este hecho hace que algunas partes verdes no aparezcan en la imagen binaria tras la aplicación de la umbralización sobre la imagen ExG mediante Otsu, en la que aún no se ha incorporado información de textura, figuras 5.7(b) y 5.8(b). Por el contrario, cuando se considera la información espacial de textura, algunas partes verdes no detectadas se han recuperado, figuras 5.7(c) y 5.8(c). Esta mejora es muy importante en AP donde resultan de notable relevancia la identificación de los cultivos y las malezas además de las densidades y ubicaciones de cara a la realización de los tratamientos específicos selectivos, siendo también de vital importancia para la navegación de los vehículos autónomos siguiendo las líneas de cultivo en cultivos de calle ancha, por ejemplo maíz o girasol, sin provocar daños en las plantas. Esta mejora cualitativa es el comportamiento general observado para el conjunto de las 1200 imágenes analizadas (sección 5.2.1), por lo que el procedimiento propuesto es adecuado para tareas de segmentación de imágenes agrícolas.

5.2.1. Imágenes de campos de cultivo

Se ha realizado un análisis sobre el conjunto de imágenes disponibles para verificar que la estrategia propuesta obtiene resultados satisfactorios. Con tal fin, se ha analizado un conjunto de 1200 imágenes, de las cuales hemos seleccionado al azar 240, correspondientes al 10% que han sido adquiridas con la cámara Basler y el otro 10% capturadas con el dispositivo SVS-Vistek, cuyas especificaciones se proporcionan a continuación. Para cada una de las imágenes seleccionadas se genera el correspondiente “ground-truth”, sobre una región de interés en el terreno, que es la que determina el espacio para el tratamiento.

5.2.1.1. Cámara Basler

Es una cámara industrial, que permite el procesado de las imágenes directamente al ser adquiridas, no es necesario almacenar las imágenes capturadas, facilitando un ahorro

importante en la memoria del sistema de procesamiento. Las imágenes obtenidas poseen una resolución de 1390×1038 píxeles y la cámara es capaz de capturar fotogramas a una tasa máxima de 17 imágenes por segundo (fps). Las imágenes se pueden adquirir bajo distintos modelos tanto de monocolor como de color, dependiendo del objetivo final, es por ello que para el caso de imágenes de exterior en entornos agrícolas, la cámara se ha configurado para obtener imágenes a color según el modelo Mono8. La cámara cuenta con un sensor CCD Sony ICX 285, cuyo tamaño de píxel es de $6.45 \mu m$ tanto en horizontal (h) como vertical (v). El tamaño de óptica requerida es de $2/3''$ y la comunicación se lleva cabo a través del estándar IEEE 1394b (también conocido como FireWire 800 o i.Link). La alimentación de esta cámara se realiza a través del mismo cable FireWire cuyo voltaje varía en el rango de +8 a +36 VDC. El consumo de este sistema en la configuración reseñada es de 3.0 W a 12 V. El rango de temperatura, para su correcto funcionamiento, debe situarse entre 0° y 50° C. La cámara tiene un peso de 120 g y un tamaño de $73.7 \text{ mm} \times 44 \text{ mm} \times 29 \text{ mm}$ (largo \times ancho \times alto), considerando estas dimensiones sin lente ni elementos de conexión.

5.2.1.2 SVS-VISTEK

También es una cámara industrial permitiendo el procesado de las imágenes directamente. Este modelo incorpora un sensor CCD Kodak KAI 04050M/C con un filtro de color Bayer con patrón GR. El tamaño del CCD es de $1''$. La resolución es de 2336×1752 píxeles con un tamaño de píxel de $5.5 \mu m$ tanto en horizontal (h) como vertical (v) y utiliza una rosca M42 para el acoplamiento de la lente. La cámara es capaz de capturar fotogramas a una tasa máxima de 16 imágenes por segundo, pudiéndose controlar el tiempo de exposición, cuyo rango varía en función de si se controla de forma interna o externa, siendo sus valores $6 \mu s - 60 s$ y $6 \mu s - \infty$ respectivamente. Dispone de un micro-controlador interno que permite diferentes modos de ajuste para el tiempo de exposición y posee distintos modos de disparadores (mecanismos para indicar que se ha de realizar una captura), por ello, la cámara se puede configurar en *Modo Libre*, que permite adquirir imágenes con la máxima tasa de adquisición o en *Modo Disparador*, el cual sincroniza la captura de imágenes en relación a un evento externo. Además, el tiempo de exposición se puede controlar mediante un puerto serie o un ancho de pulso, permitiendo aumentar el tiempo para condiciones de poca iluminación. Otra de las características que tiene esta cámara es el ajuste de las ganancias, permitiendo asignar valores diferentes a cada canal (Rojo, Verde, Azul) o realizar un ajuste automático de la misma. La lente se acopla a través de una montura tipo C y la

comunicación se lleva cabo a través del estándar Gigabit Ethernet (GigE), con dos puertos de conexión de esta naturaleza, con posibilidad de enviar la misma imagen simultáneamente. La alimentación se realiza a través de un conector distinto y el voltaje es de 12 VDC con un consumo de 600 mA. El rango de temperatura para su correcto funcionamiento es de -10° hasta $+45^{\circ}$ C. El tamaño de la cámara es de 55 mm \times 70 mm \times 70 mm (largo \times ancho \times alto) sin ningún tipo de accesorio adicional, tal como lente o conectores.

5.2.1.3. Imagen original

Con el fin de construir la realidad del terreno, seleccionamos una región de interés (ROI, *Region Of Interest*), donde todas las operaciones propias del tratamiento de imágenes aplicadas tienen en consideración solamente esa parte de la imagen. Esto es debido a que no es necesario procesar toda la imagen al ser solamente de interés dicha región, donde se encuentra la zona de cultivo a tratar. De este modo definimos este ámbito de acción prioritaria y en esta zona resulta posible el retoque manual de cara a la construcción del “ground-truth”, siendo inviable en zonas remotas de la imagen. El considerar sólo una parte de la imagen a tratar tiene además como finalidad reducir el tiempo de ejecución o coste computacional del procedimiento, cuyo objetivo último es el desarrollo de un método capaz de operar en tiempo real.

Como en el caso de la generación del “ground-truth” de la sección anterior, dada una imagen original en el espacio de color RGB, extraemos todos los píxeles cuyo valor de la componente espectral verde (G) es un 10% mayor en relación con las otras componentes la roja (R) y la azul (B); este porcentaje se ajustó después de diversas pruebas de ensayo y error. Utilizando esta medida nos aseguramos que se han extraído todas las plantas verdes, aunque así también la mayoría de los píxeles se identifican como plantas sin posiblemente formar parte de una planta, por lo que se crean falsos positivos. De esta manera se proporciona una diferenciación de lo que se consideran las plantas verdes, marcadas como blanco y el resto de la imagen, es decir, las zonas no vegetales cuyos píxeles se etiquetan como negro. Además, se retoca cada imagen manualmente de acuerdo con criterios de expertos humanos comparando esta imagen con la original, hasta obtener el “ground-truth”. Como hemos comentado antes, sin el retoque manual, las imágenes binarias contienen excesivas partes blancas que no se corresponden exactamente con las plantas verdes.

En la figura 5.9 se muestra el “ground-truth” de la región de interés ROI de la imagen de la figura 5.7(a). La subimagen etiquetada como "Greenness" representa los píxeles mal clasificadas (píxeles negros) obtenidos a partir de la comparación de la imagen original procedente de la realidad sobre el terreno con la imagen obtenida sin la mejora, figura 5.7(b). La subimagen etiquetada como "Greenness + texture" también representa los píxeles mal clasificadas (píxeles negros) obtenidos a partir de la comparación de la imagen de la realidad sobre el terreno con la imagen obtenida con la estrategia propuesta, es decir, el verdor combinado con la textura espacial, figura 5.7.(c). Como podemos ver la densidad de errores en la ROI con sólo el verdor es mayor que los obtenidos con verdor combinado con textura espacial, verificando claramente y de forma cualitativa el mejor resultado obtenido mediante el enfoque propuesto.

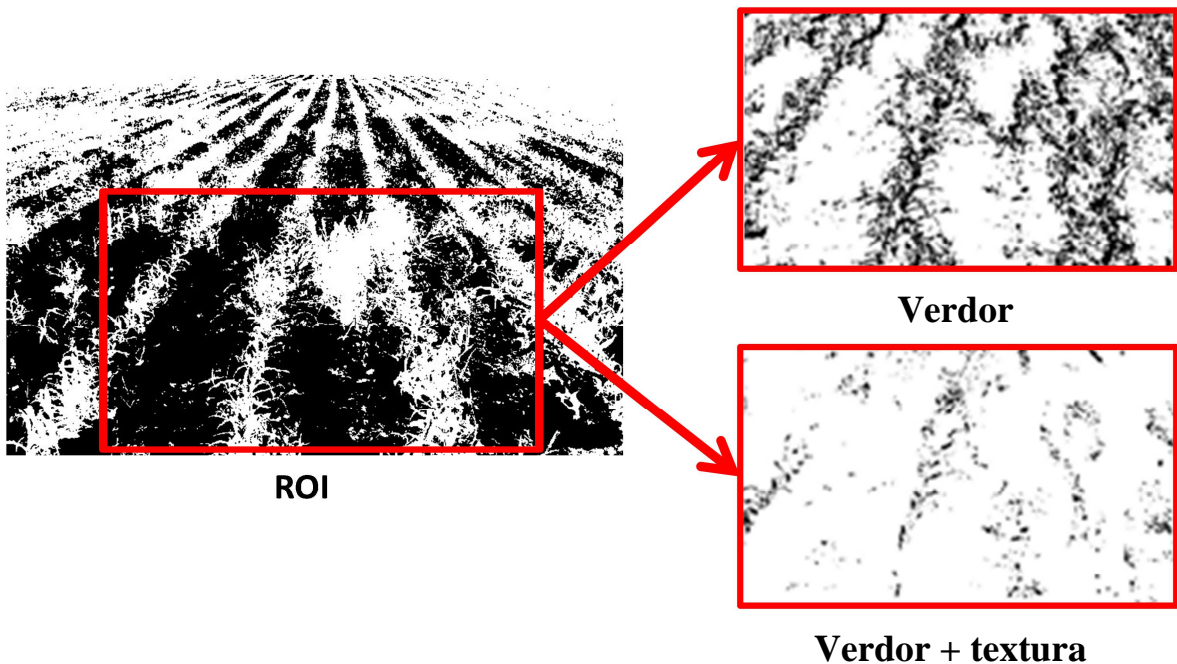


Figura 5.9.- El “ground-truth” de la ROI con clasificación errónea (puntos negros) solamente verdor y con verdor más textura espacial.

Con el fin de evaluar la validez de la estrategia de segmentación combinada, se ha realizado el correspondiente análisis cuantitativo con los dos objetivos siguientes:

- En primer lugar, como la combinación propuesta se basa en la descomposición de la imagen mediante la transformada de wavelets se estudia el comportamiento de las diferentes familias de ondas madre o base con diferentes niveles de resolución, lo que nos permite determinar la mejor familia.

- Una vez se determina la mejor familia, el segundo estudio se centra en determinar el rendimiento alcanzado por el enfoque combinado en comparación con las estrategias existentes basadas exclusivamente en la información cromática de verdor.

5.2.2. Familias de wavelets

Disponemos de un conjunto de 240 imágenes del terreno con su correspondiente región de “ground-truth” seleccionando al azar el 10% con cada uno de los sensores disponibles, cuyo objetivo en este caso es comparar las diferentes familias de wavelets. La tabla 5.2 proporciona el porcentaje medio de error obtenido para las familias estudiadas (Daubechies, 1992).

Para cada familia hemos identificado las escalas de las wavelets madre, que se representan como el tamaño de los bancos de filtros (N). Dependiendo del número N cada filtro digital contiene un número dado de coeficientes. Por lo tanto, teniendo en cuenta la familia Daubechies, si $N = 2, 4, 6, 10$ etc. el número de coeficientes en los filtros 1-D en paso bajo y paso alto correspondientes son 2, 4, 6, 10 y así sucesivamente. Esto significa que cuanto mayor es el número de coeficientes de filtro más píxeles existen en las imágenes o subimágenes involucrados durante el proceso de filtrado, tanto en las direcciones de transformación directa como en la inversa. Así, en las familias biortogonales tanto directa como inversa existen dos funciones de escalado implicadas, cada una con una escala diferente, representadas como n y m . Los tamaños del filtro en estas familias están determinados por la combinación de n y m . Por ejemplo, si $n = 1$ y $m = 1$, ambos filtros paso bajo y paso alto se definen con dos coeficientes; si $n = 1$ y $m = 3$ ambos filtros contienen seis coeficientes o cuando $n = 2$ y $m = 6$ los filtros contienen catorce coeficientes y así sucesivamente. Los diferentes tamaños de los filtros en la tabla han sido ordenados de forma ascendente de izquierda a derecha.

| | | | | | | | | | | |
|-----------------------------|--------------|-------|------------|------------|-------|-------|-------|-------|-------|-------|
| Daubechies | N | 2 | 4 | 6 | 10 | 20 | 24 | 28 | 32 | 40 |
| | % | 8.1 | 7.1 | 7.2 | 12.8 | 18.9 | 28.0 | 30.6 | 34.2 | 49.1 |
| Biorthogonal | N | 2 | 6 | 6 | 8 | 10 | 10 | 14 | 16 | 18 |
| | (n,m) | (1,1) | (1,3) | (2,2) | (3,3) | (1,5) | (4,4) | (2,6) | (3,7) | (6,8) |
| | % | 8.5 | 7.8 | 7.3 | 11.1 | 13.1 | 13.4 | 14.8 | 17.5 | 18.0 |
| Reverse biorthogonal | N | 2 | 6 | 6 | 8 | 10 | 10 | 14 | 16 | 18 |
| | (n,m) | (1,1) | (1,3) | (2,2) | (3,3) | (1,5) | (4,4) | (2,6) | (3,7) | (6,8) |
| | % | 8.3 | 7.3 | 7.9 | 11.5 | 12.9 | 14.1 | 15.0 | 18.8 | 19.8 |
| Coiflets | N | 6 | 12 | 18 | 24 | 30 | | | | |
| | % | 7.5 | 15.9 | 17.3 | 19.1 | 22.2 | | | | |
| Symlets | N | 4 | 8 | 10 | 12 | 14 | 20 | 24 | 28 | 32 |
| | % | 7.2 | 10.7 | 13.6 | 16.3 | 15.7 | 20.5 | 28.5 | 31.1 | 35.8 |
| Discrete Meyer | N | 45 | | | | | | | | |
| | % | 53.2 | | | | | | | | |

Tabla 5.2.- Porcentaje medio de error obtenido para las diferentes familias de wavelets

A partir de los resultados mostrados en la tabla 5.2 se puede inferir que los mejores rendimientos se obtienen con filtros de tamaños 4 y 6. Los de tamaño 4 obtienen los mejores resultados en las imágenes capturadas con el sensor Basler y los de tamaño 6 con el sensor SVS-Vistek. Este comportamiento está directamente relacionado con las resoluciones espaciales de las imágenes y también con la longitud focal en sus respectivos sistemas ópticos, 8 y 10 mm respectivamente. Con tales resoluciones de imagen y longitudes focales, una conclusión general es que a medida que la resolución de la imagen aumenta el tamaño del filtro también se puede aumentar, es decir, las texturas son capturadas de la forma correcta. La familia Daubechies obtiene mejores resultados que el resto, estando en línea con la previsión que se había formulado anteriormente. Así, para las pruebas cuyos resultados se describen a continuación seleccionamos esta familia con filtros de tamaños 4 para las imágenes capturadas con Basler y de tamaño 6 para las imágenes procedentes de SVS-Vistek.

5.2.3. Análisis ExG

Una vez que hemos seleccionado el mejor tamaño y escala de las familias, aplicamos la TDW para el resto de imágenes formadas por el 90% de las 240 imágenes con “ground truth” cuyo objetivo es comparar los resultados del método propuesto que refuerza el ExG con información espacial de textura y se comparan estos resultados con cuatro enfoques de binarización existentes, pero sin información de textura.

Estos métodos se resumen a continuación:

- a) Umbralización Otsu (Otsu, 1979): se aplica directamente sobre la imagen ExG resultante calculada al aplicar la ecuación $ExG=2g-r-b$.
- b) COM (Guijarro y col., 2011): explicado en la sección 4.1, con los valores obtenidos en la sección 5.1.2, donde la binarización de la imagen se realiza mediante la media.
- c) SVM (Guerrero y col., 2012): consiste en un procedimiento de división secuencial en dos pasos. En el primero se aplica el procedimiento COM descrito con la binarización mediante Otsu. En el segundo se aplica un enfoque basado en las conocidas Máquinas de Vectores Soporte para identificar dos tipos de plantas con altos y bajos grados de verdor.
- d) AES: (Montalvo y col., 2013): con un enfoque similar a la estrategia SVM, basado en un doble umbral aplicando Otsu en dos pasos. En el primero se obtiene una imagen binaria y sobre las partes identificadas como las plantas verdes, se aplica un segundo Otsu, distinguiendo así dos tipos de verdes.

La tabla 5.3 muestra los resultados obtenidos por estos métodos frente al método propuesto (WAV, Wavelets) en términos del porcentaje medio de error y desviación estándar, que se obtienen por comparación de los correspondientes “ground-truth”.

| Método | | Otsu | | COM | | SVM | | AES | | WAV | |
|---------|---------------------|-------|------|-------|------|-------|------|-------|------|------|------|
| % error | Desviación estándar | 19.36 | 1.71 | 12.29 | 1.50 | 11.86 | 1.27 | 11.78 | 1.23 | 7.12 | 0.93 |

Tabla 5.3.- Porcentaje de error medio y desviación estándar sobre los ROI's en el conjunto de imágenes “ground truth”.

Los resultados muestran que WAV, el método propuesto, supera a las otras estrategias con eficacia. Esto significa que la combinación de índices de vegetación cromática incorporando información de textura espacial funciona adecuadamente. Con WAV hemos obtenido el valor de error promedio más bajo de 7,12% y de la desviación estándar también más baja. Esto significa que el método propuesto es más estable y robusto que las otras estrategias debido al menor valor de la desviación estándar. Por lo tanto, WAV se comporta bien en este doble sentido, es decir, menos errores con mayor estabilidad. Como podemos observar, COM, SVM y AES superan al método de Otsu simple, esto se debe a que la combinación de varios índices de vegetación aplicados en las otras tres estrategias es más

favorable que la utilización de un único índice. Los resultados obtenidos por COM, SVM y AES son ligeramente peores que los reportados en sus respectivas referencias. Esto es una consecuencia directa de la incorporación de nuevas imágenes procedentes de la cámara SVS-Vistek para el análisis. Por otra parte, la utilización de la media estadística como umbral tras la aplicación de COM se ve mejorado por las estrategias descritas SVM y AES donde se aplica Otsu como primer umbral. Estos resultados nos llevan a la conclusión de que el sensor utilizado para la adquisición de imágenes tiene su propia influencia en los resultados, siendo especialmente relevante en COM, SVM y AES. En este sentido, podemos concluir que la aportación realizada en esta investigación, WAV, no se ve particularmente influenciada por el tipo de sensor, una vez que la mejor escala de la wavelet madre ha sido identificada, como se ha explicado previamente. Además, el problema de usar como umbral la media u Otsu ha sido resuelto en esta estrategia. En resumen, WAV se convierte en una estrategia apropiada, según el análisis comparativo llevado a cabo.

Con el fin de profundizar más en el análisis de WAV, la tabla 5.4. muestra los valores de “precision”, “recall” y ”accuracy” tal y como se definen en las ecuaciones (5.1), (5.2) y (5.3) y que varían en el intervalo [0,1].

$$Precision = TP / (TP+FP) \tag{5.1}$$

$$Recall=TP / (TP+FN) \tag{5.2}$$

$$Accuracy = (TP+TN) / (TP+TN+FP+FN) \tag{5.3}$$

Donde los símbolos T y F significan verdadero y falso respectivamente y P, N positivo y negativo. Siendo TP el término que hace referencia a los verdaderos positivos, es decir, aquellas muestras que han sido clasificadas como verdes y realmente lo son. Los TN, son aquellos píxeles clasificados como suelo siendo realmente suelo. FP hace referencia a las muestras que han sido catalogadas como verde y son suelo. Y, finalmente, los FN son aquellos píxeles que han sido clasificados como suelo siendo verdes. El término de “precision” indica cuántos píxeles seleccionados son verdes, viéndose penalizado por los FP y sin tener en cuenta ni los TN ni los FN. El término “recall” se ve penalizado por los FN, y expresa cuántos píxeles verdes han sido clasificados como tales. Estos datos nos indican que se han obtenido más píxeles considerados como suelo siendo verdes, un FN con valor más alto que el FP. Así vemos que el principal fallo en la clasificación realizada radica en la asignación de píxeles verdes como suelo. El término “accuracy” proporciona la precisión de

la clasificación tanto en los píxeles que deben ser verdes como aquellos que tienen que ser suelo.

| | Precision | Recall | Accuracy |
|-----|-----------|--------|----------|
| WAV | 0.9209 | 0.8952 | 0.9288 |

Tabla 5.4.- Valores de “Precision”, “Recall” y “Accuracy”

También podemos inferir que los tres valores son lo suficientemente altos a la vez en WAV, que para el resto de los métodos se obtienen peores resultados coincidiendo en este caso con las conclusiones derivadas de la tabla 5.3. Además, merece la pena señalar que WAV obtiene un mayor número de FN, que FP, dado el menor valor de “recall” que de “precision”. Esto significa que las partes verdes se pierden durante el proceso de segmentación, debido a que se consideran como suelo cuando en realidad no lo son.

La figura 5.10 muestra las curvas ROC (*Receiver Operating Characteristic*) (Provost y col., 1998). La curva azul corresponde a la identificación realizada por Otsu sin incorporar información espacial de textura. La curva naranja representa los resultados obtenidos por Otsu tras la mejora de ExG con la información espacial de textura. El área bajo la curva ROC azul es 0.8532 y bajo la naranja es 0.9211. Como se deduce fácilmente, la zona con el método propuesto posee un valor más alto, indicando que la segmentación realizada con el método propuesto es mejor debido a que la probabilidad de que un píxel de verde seleccionado al azar sea clasificado como verde es mayor (Zhou y col., 2002).

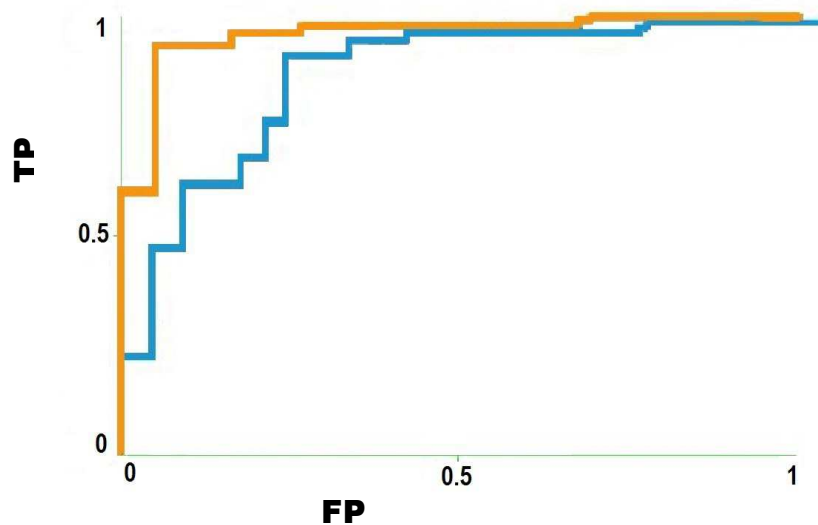


Figura 5.10. Curva ROC: Clasificación por Otsu; el azul utiliza la información del verdor de ExG, y el naranja utiliza la información de verdor del ExG con la información de la textura espacial.

6

Conclusiones y trabajo futuro

Una vez realizados el análisis y la valoración de los resultados obtenidos en las dos propuestas formuladas, llega el momento de plantear el balance global sobre las líneas de investigación seguidas, que han tenido como consecuencia los resultados expresados en el capítulo anterior. Además, es también el momento reflexionar, pensar en los resultados logrados y cómo mejorarlos en la medida de lo posible; en consecuencia, definir las futuras líneas de investigación para dirigir los esfuerzos en el futuro.

En la sección 6.2 se incluyen las conclusiones generales obtenidas para cada uno de los sistemas automáticos desarrollados mientras que en la 6.3 se incluyen las líneas de actuación futura.

6.1. Conclusiones

6.1.1. Segmentación automática combinando índices cromáticos y agrupamiento borroso

La estrategia de segmentación automática propuesta permite identificar las plantas y el suelo. El objetivo ha sido mejorar el rendimiento de algunos procesos de segmentación

existentes, sobre todo en lo referente a la extracción de las partes verdes que contienen los cultivos e incluso las malas hierbas. Para conseguirlo se ha desarrollado una estrategia combinada basada en la información contenida en las imágenes a través de la uniformidad de los histogramas. Se ha verificado que el enfoque combinado obtiene mejores resultados que los métodos utilizados por separado, se ha demostrado que esta estrategia es válida para dos tipos de imágenes, las panorámicas que contienen cielo y las que sólo contienen suelo y partes vegetales. Ambos tipos de imágenes se procesan con el objeto de aplicar tratamientos agrícolas específicos, las primeras contienen un campo más amplio de visión con el propósito de identificar posibles estructuras que podrían ser útiles durante la navegación autónoma de un vehículo equipado con el sistema de visión. También se ha diseñado una nueva estrategia automática para discriminar diferentes texturas en las plantas verdes y en el suelo, que también son útiles para tratamientos específicos a la vez que para recuperar información importante contenida en el ámbito de las aplicaciones y tratamientos futuros.

El rendimiento del método ha sido probado en el conjunto de imágenes disponibles, habiendo verificado su validez y eficacia para aplicaciones de segmentación de imágenes obtenidas bajo diferentes condiciones de iluminación en entornos agrícolas de exterior, también se ha comparado favorablemente con otros métodos ampliamente utilizados por la comunidad científica con la misma finalidad, es decir con los diferentes métodos de extracción de verdes.

6.1.2. Extracción de verdes mediante un índice cromático e información espacial

A partir de la estrategia basada exclusivamente en extracción de texturas mediante índices cromáticos, se han planteado nuevos retos y objetivos para encontrar una solución más eficaz de cara al reconocimiento de las plantas verdes (cultivos y malas hierbas) en campos de cultivo basada en la incorporación de información espacial a la cromática. La finalidad es realizar tratamientos en AP. La estrategia propuesta combina el verdor y la información de textura espacial, mediante el uso de índices de vegetación y la transformada wavelet. La distribución espacial de textura irregular, existente en el suelo y partes verdes (cultivos y malas hierbas), se captura adecuadamente a través de la descomposición de wavelets en coeficientes de detalle (horizontales, verticales y diagonales) y coeficientes de aproximación proporcionados por la propia transformada de wavelets, permitiendo mejoras

importantes en la segmentación de imágenes. Se ha utilizado Otsu como principal estrategia de umbral, una vez que se ha verificado que el uso de la media como umbral no resulta un problema con este enfoque. Se ha determinado que los resultados dependen del sensor utilizado para la captura de las imágenes, pero una vez que se establece la escala de la wavelet madre, el método propuesto con este enfoque supera a otras estrategias existentes, mejorando los resultados en un 4,5% aproximadamente.

Basado en el indicador Recall y el valor asociado FN, discutido previamente, cabe señalar que aún se podría realizar un cierto esfuerzo adicional para poner de relieve las partes verdes descubiertas por dicho índice, ya que el objetivo principal es obtener precisamente esas partes verdes de la imagen.

6.2. Trabajo futuro

Las estrategias propuestas han sido probado en un amplio conjunto de imágenes procedentes de los campos de cebada y maíz, que son cultivos conteniendo plantas verdes y partes del suelo visibles con tonalidades rojas dominantes y quizás cielo con coloración azul dominante. Por lo tanto, se puede aplicar a cualquier tipo de cultivos con tales características, por supuesto en otros cultivos de cereal tales como trigo o centeno.

Por otro lado, los dos métodos propuestos han demostrado su eficacia bajo diferentes condiciones de iluminación en entornos de exterior, con un alto grado de dificultad. En este sentido, una posible mejora sería tratar de minimizar el riesgo de la iluminación mediante la aplicación de técnicas tales como el filtrado homomórfico, que es una técnica probada en entornos de exterior con un alto grado de eficacia (Pajares y col., 2005).

También se puede plantear la utilización de nuevos métodos de combinación de la información espacial en la línea de la fusión de imágenes en el campo de la Visión por Computador, destacando los siguientes (Pajares y Cruz, 2007):

- Selección por máximos.
- Selección por componentes principales.
- Media ponderada general.
- Media ponderada adaptativa.

- Fusión basada en regiones por detección de características.
- Eliminación del fondo.
- Basada en la varianza.

Se pueden plantear otras técnicas de clasificación mediante la combinación de clasificadores que han demostrado su efectividad en otros contextos, como es el caso de los métodos propuestos en Pajares y col, (2009), Guijarro y Pajares (2009) y Guijarro (2010). Finalmente, es necesario ampliar la investigación y comprobar la validación de la propuesta formulada para otros tipos de cultivos tales como arroz o leguminosas, donde las firmas espectrales de las plantas pueden variar.

En el futuro se podría considerar la aplicación de nuevas estrategias de umbralización automática (Avci y Avci, 2009) que tratan de mejorar las prestaciones respecto de las consideradas en este trabajo de investigación.

Finalmente, para simular situaciones reales encontradas en el campo de cultivo, se podría utilizar en el futuro herramientas de modelado y simulación como las propuestas en García-Magariño y col., (2010) y Fuentes-Fernández y col., (2012). Estas herramientas pueden ofrecer la posibilidad de introducir cambios en el campo simulado con el propósito de representar situaciones que pueden ocurrir en la realidad.

Bibliografía

- Ahmed, F., Al-Mamun, H. A., Bari, A. H., Hossain, E., Kwan, P. (2012). Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection*, 40, 98-104.
- Aitkenhead, M. J., Dalgetty, I. A., Mullins, C. E., McDonald, A. J. S., Strachan, N. J. C. (2003). Weed and crop discrimination using image analysis and artificial intelligence methods. *Computers and Electronics in Agriculture*, 39(3), 157-171.
- Alchanatis, V., Ridel, L., Hetzroni, A., Yaroslavsky, L. (2005). Weed detection in multi-spectral images of cotton fields. *Computers and Electronics in Agriculture*, 47(3), 243-260.
- Arroyo, J., Guijarro, M., Riomoros, I., Pajares, G., (2012). A symbolic solution to a precision agriculture. *Proc. 3rd Workshop in Symbolic Data Analysis* (Arroyo, J., Maté, C., Brito, P., Noirhomme-Fraiture, Eds.), 7-9 November 2012, Madrid, Spain, pp. 25-26.
- Ashjari, B. (1982). Singular value decomposition texture measurement for image classification. University of Southern California, Los Angeles, CA.
- Avci, E., Avci, D. (2009). An expert system based on fuzzy entropy for automatic threshold selection in image process. *Expert Systems with Applications*, 36(2), 3077–3085.

-
- Bai, X., Cao, Z., Wang, Y., Yu, Z., Hu, Z., Zhang, X., Li, C. (2014). Vegetation segmentation robust to illumination variations based on clustering and morphology modelling. *Biosystems Engineering*, 125, 80-97.
- Balasko, B., Abonyi, J., Feil, B. (2005). Fuzzy clustering and data analysis toolbox. Department of Process Engineering, University of Veszprem, Veszprem.
- Bay, H., Ess, A., Tuytelaars, T., Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3), 346-359.
- Berge, T. W., Goldberg, S., Kaspersen, K., Netland, J. (2012). Towards machine vision based site-specific weed management in cereals. *Computers and Electronics in Agriculture*, 81, 79-86.
- Berni, J. A. J., Zarco-Tejada, J. A., Suárez, L., González-Dugo, V., Fereres, E. (2009a). Remote sensing of vegetation from UAV platforms using lightweight multispectral and thermal imaging sensors. ISPRS Hannover Workshop High-Resolution Earth Imaging for Geospatial Information: ISPRS Archives – Volume XXXVIII-1-4-7/W5.
- Bezdek, J.C., (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer, Plenum Press, New York.
- Bharati, M. H., Liu, J. J., MacGregor, J. F. (2004). Image texture analysis: methods and comparisons. *Chemometrics and Intelligent Laboratory Systems*, 72(1), 57-71.
- Birth, G.S., McVey, G.R. (1968). Measuring the color of growing turf with a reflectance spectrophotometer. *Journal in Agronomy*. 60, 640-643.
- Bossu, J., Gée, Ch., Jones, G., Truchetet, F. (2009). Wavelet transform to discriminate between crop and weed in perspective agronomic images. *Computers and Electronic in Agriculture*, 133-143.
- Bossu, J., Gée, C., Truchetet, F. (2007). Wavelet Transform to discriminate between crop and weed in agronomic images. *Optics East 2007* (pp. 67630R-67630R). International Society for Optics and Photonics.
- Burgos-Artizzu, X.P., 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.
- Burgos-Artizzu, X. P., Ribeiro, A., Tellaache, A., Pajares, G., Fernández-Quintanilla, C. (2010). Analysis of natural images processing for the extraction of agricultural elements. *Image and Vision Computing*, 28(1), 138-149.
- Burgos-Artizzu, X.P., Ribeiro, A., Tellaache, A., Pajares, G., Fernández-Quintanilla, C. (2009). Improving weed pressure assessment using digital images from an

- experience-based reasoning approach. *Computers and Electronics in Agriculture*, 65, 176-185.
- Buschmann, C., Nagel, E. (1993). In vivo spectroscopy and internal optics of leaves as basis for remote sensing of vegetation. *International Journal of Remote Sensing*, 14(4), 711-722.
- Carlson, T. N., Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote sensing of Environment*, 62(3), 241-252.
- Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J. (2001). Color image segmentation: advances and prospects. *Pattern Recognition*, 34, 2259-2281.
- Chou, J.J., Chen, C.P. Yeh, J.T. (2007). Crop identification with wavelet packet analysis and weighted Bayesian distance. *Computers and electronics in Agriculture*, 57(1), 88-98.
- Clevers, J. G.. (1998). The derivation of a simplified reflectance model for the estimation of leaf area index. *Remote Sens. Environ.*, 25 , 53–69
- Crippen, R.E. (1990). Calculating the vegetation index faster. *Remote Sensing of Environment*, 34, 71-73.
- Cloutis, E. A., Connery, D. R., Major, D. J., Dover, F. J. (1996). Airborne multi-spectral monitoring of agricultural crop status: effect of time of year, crop type and crop condition parameter. *Remote Sensing*, 17(13), 2579-2601.
- Daubechies, I., (1992). Ten Lectures on Wavelets. SIAM, Philadelphia, PA, USA.
- Dawson-Howe, K. (2014). A Practical Introduction to Computer Vision with OpenCV (1^a ed.). ISBN: 978-1-118-84845-6, 234 pages, John Wiley and Sons, United Kingdom.
- Davies, G., Casady, W., Massey, R., (1998). Precision agriculture: an introduction. Water Quality Focus Guide, WQ450, disponible en línea:
<http://extension.missouri.edu/explorepdf/envqual/wq0450.pdf>.
(Accedido octubre 2015)
- De Castro, A. I., López-Granados, F., Jurado-Expósito, M. (2013). Broad-scale cruciferous weed patch classification in winter wheat using QuickBird imagery for in-season site-specific control. *Precision Agriculture*, 14(4), 392-413.
- De Castro, M. Jurado-Expósito, J.M. Peña-Barragán, F. López-Granados (2012). Airborne multi-spectral imagery for mapping cruciferous weeds in cereal and legume crops. *Precision Agriculture*, 13, 302–321.
- De Castro, A. I., Jurado-Expósito, M., Gómez-Casero, M. T., López-Granados, F. (2012a). Applying neural networks to hyperspectral and multispectral field data for

- dicrimination of cruciferos weeds in winter crops. *The Scientific World Journal*, 2012, 1–11.
- Duda, R.O., Hart, P.E., Stork, D.S., (2000). *Pattern Classification*. Wiley.
- Dunn J. C. (1974). Well-separated clusters and optimal fuzzy partitions. *Journal of cybernetics*, 4(1), 95-104.
- Earl, R., Wheeler, P.N., Blackmore, B.S., Godwin, R.J. (1996). *Precision Farming: The management of variability*. *Landwards*, 51(4), 18–23.
- El-Faki, M. S., Zhang, N., Peterson, D. E. (2000). Weed detection using color machine vision. *Transactions of the ASAE-American Society of Agricultural Engineers*, 43(6), 1969-1978.
- El-Faki, M.S., Zhang, N., Peterson, D.E., 1997. Weed detection using color machine vision. ASAE paper 97-3134.
- Fuetes-Fernandez, R., Hassan, S., Pavon, J., Galan, J.M., Lopez-Paredes, A. (2012). Metamodels for role-driven agent-based modelling. *Computational and Mathematical Organization Theory*, 18(1), 91-112.
- García-Magariño, I., Fuentes-Fernández, R., Gomez-Sanz, J. J. (2010). A framework for the definition of metamodels for Computer-Aided Software Engineering tools. *Information and Software Technology*, 52(4), 422-435.
- Garguet-Duport, B., Girel, J., Chassery, J. and Pautou, J. G. (1996). The use of Multiresolution Analysis and Wavelets Transform for Merging SPOT Panchromatic and Multispectral Image Data. *Photogrammetric Engineering and Remote Sensing* 62(9),1057-1066.
- Gebhardt, S., Käuhbauch, 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. *Precision Agriculture*, 8(1), 1–13.
- Gebhardt, S., Schellberg, J., Lock, R., Käuhbauch, W.A., (2006). Identification of broadleaved dock (Rumex obtusifolius L.) on grassland by means of digital image processing. *Precision Agriculture* 7 (3), 165–178.
- Gée, Ch., Bossu, J., Jones, G., Truchetet, F. (2008). Crop/weed discrimination in perspective agronomic images. *Computers and Electronics in Agriculture* 60, 49-59.
- Gitelson, A. A., Zur, Y., Chivkunova, O. B., Merzlyak, M. N. (2002). Assessing Carotenoid Content in Plant Leaves with Reflectance Spectroscopy. *Photochemistry and Photobiology*, 75(3), 272-281.
- Gómez-Candón, D., López-Granados, F., Caballero-Novella, J.J., Peña-Barragán, J.M., García-Torres, L. (2012). Understanding the errors in input prescription maps based

- on high spatial resolution remote sensing images. *Precision Agriculture*, 13(5), 581-593.
- 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. *Journal of food engineering*, 74(2), 268-278.
- Gonzalez, R.C., Woods, R.E., Eddins, S.L., (2004). *Digital Image Processing using Matlab*. Prentice-Hall, Upper Saddle River, NY.
- Gonzalez, R.C., Woods, R.E., (2008). *Digital Image Processing*, Prentice-Hall, Upper Saddle River, NY.
- Gonzalez, R.C., Woods, R.E. (2007). *Digital Image Processing 3^a Ed.* Upper Saddle River, NJ: Prentice Hall.
- Göktoğan, A. H., Sukkarieh, S., Bryson, M., Randle, J., Lupton, T., Hung, C. (2010). A rotary-wing unmanned air vehicle for aquatic weed surveillance and management. In *Selected papers from the 2nd International Symposium on UAVs, Reno, Nevada, USA June 8–10, 2009* (pp. 467-484). Springer Netherlands.
- Guerrero, J. M., Pajares, G., Montalvo, M., Romeo, J., Guijarro, M. (2012). Support vector machines for crop/weeds identification in maize fields. *Expert Systems with Applications*, 39, 11149–11155.
- Guijarro, M., Riomoros, I., Pajares, G., & Zitinski, P. (2015). Discrete wavelets transform for improving greenness image segmentation in agricultural images. *Computers and Electronics in Agriculture*, 118, 396-407.
- Guijarro, M., Pajares, G., Riomoros I., Herrera, P.J., Burgos-Artizzu, X.P., Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture*, 75(1), 75-83
- Guijarro Mata-García, M. (2010). *Combinación de clasificadores para identificación de texturas en imágenes naturales: nuevas estrategias locales y globales*. Tesis Doctoral. Universidad Complutense de Madrid.
- Guijarro, M., Pajares, G. (2009). On combining classifiers through a fuzzy multicriteria decision making approach: Applied to natural textured images. *Expert Systems with Applications*, 36(3), 7262-7269.
- Hague, T., Tillet, N., Wheeler, H. (2006). Automated crop and weed monitoring in widely spaced cereals. *Precision Agriculture*, 1(1), 95–113.
- Haralick, R. M., Shanmugam, K., Dinstein, I. H. (1973). Textural features for image classification. *IEEE Trans., Systems, Man and Cybernetics*, (6), 610-621.

- Hartigan, J. A. (1975). *Clustering algorithms*. John Wiley Sons, Inc..
- Hasegawa, K., Matsuyama, H., Tsuzuki, H., Sweda, T. (2010). Improving the estimation of leaf area index by using remotely sensed NDVI with BRDF signatures. *Remote Sensing of Environment*, 114(3), 514-519.
- Huang, L. K., Wang, M. J. J. (1995). Image thresholding by minimizing the measures of fuzziness. *Pattern recognition*, 28(1), 41-51.
- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G.(2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(2-Jan), 195-213.
- Huete, A.R., Liu, H.Q., Batchily, K., vanLeeuwen, W. (1997). A comparison of vegetation indices global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 59(3), 440-451.
- Huete, A.R., Justice, C., Liu, H.Q. (1994). Development of vegetation and soil indices for MODIS-EOS. *Remote Sensing of Environment*, 49(3), pages 224-234:
- Huete, A.R. (1988). A Soil-adjusted vegetation index. *Remote Sensing of Environment*, 25, 295-309.
- Ishak, A. J., Hussain, A., Mustafa, M.M. (2009). Weed image classification using Gabor wavelet and gradient field distribution. *Computers and Electronics in Agriculture*, 66(1), 53-61.
- Jurado-Expósito, M., López-Granados, F., Atenciano, S., Garcia-Torres, L., González-Andújar, J. L. (2003). Discrimination of weed seedlings, wheat (*Triticum aestivum*) stubble and sunflower (*Helianthus annuus*) by near-infrared reflectance spectroscopy (NIRS). *Crop Protection*, 22(10), 1177-1180.
- Joycy J.S., Prabavathy K. (2012). Survey on automatic segmentation of relevant textures in agricultural images. *Proc. Int. Conf. on Advances in Engineering, Science and Management (ICAESM)*, 26–32.
- Kataoka, T., Kaneko, T., Okamoto, H., Hata, S. (2003). Crop growth estimation system using machine vision. *Proc. IEEE/ASME Internat. Conf. on Advanced Intelligent Mechatronics*, 2, 1079-1083.
- Kirk, K., Andersen, H.J., Thomsen, A.G., Jørgensen, J.R. (2009). Estimation of leaf area index in cereal crops using red-green images. *Biosystems Engineering*, 104, 308–317.
- Kim, D.W., Lee, K.H. and Lee, D (2003). Fuzzy Cluster validation index based on intercluster proximity. *Pattern Recognition Letters*, 24, 2561-2574.

-
- Kittler, J., Illingworth, J. (1986). Minimum error thresholding. *Pattern recognition*, 19(1), 41-47.
- Kriegler, F. J., Malila, W. A., Nalepka, R. F., Richardson, W. (1969). Preprocessing transformations and their effects on multispectral recognition. *Remote Sensing of Environment*, VI (Vol. 1, p. 97).
- Krishna, K. R. (2013). *Precision farming: soil fertility and productivity aspects*. CRC Press.
- Kropff, M. J., Wallinga, J., Lotz, L. A. P. (1997). Crop-weed interactions and weed population dynamics: current knowledge and new research targets. *Proc. 10th EWRS (European Weed Research Society) Symposium, Poznan', Poland, 22-26 June 1997* (pp. 41-48). European Weed Research Society.
- Laliberte, A. S., Herrick, J. E., Rango, A., Winters, C. (2010). Acquisition, orthorectification, and object-based classification of unmanned aerial vehicle (UAV) imagery for rangeland monitoring. *Photogrammetric Engineering Remote Sensing*, 76(6), 661-672.
- Laursen, M. S., Midtiby, H. S., Krüger, N., Jørgensen, R. N. (2014). Statistics-based segmentation using a continuous-scale naive Bayes approach. *Computers and Electronics in Agriculture*, 109, 271-277.
- Li, H., Manjunath, B.S., Mitra, S.K. (1995). Multisensor Image Fusion Using the Wavelet Transform. *Graphical Models and Image Processing* 57(3), 235-245.
- Lim, Y.W., Lee, S.U. (1990). On the color image segmentation algorithm based on the thresholding and the fuzzy c-means technique. *Pattern Recognition*, 23(9), 935-952.
- Ling, P.P., Ruzhitsky, V.N., (1996). Machine vision techniques for measuring the canopy of tomato seedling. *Journal Agricultural Engineering Research*, 65 (2), 85–95.
- López-Granados, F. (2011). Weed detection for site-specific weed management: mapping and real-time approaches. *Weed Research*, 51, 1–11.
- Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60, 91-110.
- Luccheseyz, L., Mitray, S. K. (2001). Color image segmentation: A state-of-the-art survey. *Proc. Indian National Science Academy (INSA-A)*, 67(2), 207-221.
- Luscier, J.D., Thompson, W.L., Wilson, J.M., Gorham, B.E., Dragut, L.D., (2006). Using digital photographs and object-based image analysis to estimate percent ground cover in vegetation plots. *Frontiers in Ecology and the Environment* 4 (8), 408–413.

-
- Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(7), 674-693.
- Marchant, J. A., Onyango, C. M. (2000). Shadow-invariant classification for scenes illuminated by daylight. *JOSA A*, 17(11), 1952-1961.
- Martin, M. P., Barreto, L., Fernández-Quintanilla, C. (2011). Discrimination of sterile oat (*Avena sterilis*) in winter barley (*Hordeum vulgare*) using QuicKBird satellite images. *Crop Protection*, 30, 1363–1369.
- Meng, J.H., Dong, T., Zhang, M., You, X., Wu, B. (2013). Predicting optimal soybean harvesting dates with satellite data. *Precision Agriculture*, 13, 33-39.
- Meyer, G.E., Camargo-Neto, J. (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture*, 63, 282-293.
- 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. *Computers and Electronics in Agriculture*, 42, 161-180.
- Meyer, G., Metha, T., Kocher, M., Mortensen, D., Samal, A., (1998). Textural imaging and discriminant analysis for distinguishing weeds for spot spraying. *Trans. ASAE*, 41(4), 1189–1197.
- Meyer, F., Beucher, S. (1990). Morphological segmentation. *Journal of visual Communication and Image Representation*, 1(1), 21-46.
- Meyer, Y. (1990). Ondelettes, filtres miroirs en quadrature et traitement numérique de l'image (pp. 14-25). Springer Berlin Heidelberg.
- Mitra, P., Shankar, B. U., Pal, S. K. (2004). Segmentation of multispectral remote sensing images using active support vector machines. *Pattern Recognition Letters*, 25(9), 1067-1074.
- Mohaddamzadeh A., Bourbakis, N. (1997). A fuzzy region growing approach for segmentation of color images. *Pattern Recognition*, 30(6), 867-881.
- Montalvo, M., Guerrero, JM., Romeo, J., Emmi, L., Guijarro, M., Pajares, G. (2013). Automatic expert system for weed/crops identification in images from maize fields. *Expert Systems with Applications*, 40(1), 75-82.
- Montalvo, M., Pajares, G., Guerrero, JM., Romeo, J., Guijarro, M., Ribeiro, A., Ruz, J. J., Cruz, J.M. (2012). Automatic detection of crop rows in maize fields with high weeds pressure. *Expert Systems with Applications*, 39(15), 11889-11897.
-

-
- Monteith, J.L. (1981). Evaporation and surface temperature. *Quart. J. Roy. Meteorolog. Soc.*, 107, 1–27.
- Motohka, T., Nasahara, K. N., Oguma, H., Tsuchida, S. (2010). Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sensing*, 2(10), 2369-2387.
- Nashat, S., Abdullah, A., Aramvith, S., Abdullah, M. Z. (2011). Support vector machine approach to real-time inspection of biscuits on moving conveyor belt. *Computers and Electronics in Agriculture*, 75(1), 147-158.
- Neto, J.C., (2004). A Combined Statistical—Soft Computing Approach for Classification and Mapping Weed Species in Minimum Tillage Systems. University of Nebraska, Lincoln, NE.
- O'Shaughnessy, S. A., Evett, S. R. (2010). Canopy temperature based system effectively schedules and controls center pivot irrigation of cotton. *Agricultural Water Management*, 97(9), 1310-1316.
- Oliva Navarro, D. A. (2015). Aplicación de algoritmos metaheurísticos en procesamiento de señales, imágenes y energías alternativas. Tesis Doctoral. Universidad Complutense.
- Onyango, C.M., Marchant, J.A., (2003). Segmentation of row crop plants from weeds using colour and morphology. *Computers and Electronics in Agriculture* 39, 141–155.
- Otsu, N., (1979). A threshold selection method from gray-level histogram. *IEEE Trans. on System Man and Cybernetics* 9, 62–66.
- Pajares, G. (2015). Overview and Current Status of Remote Sensing Applications Based on Unmanned Aerial Vehicles (UAVs). *Photogrammetric Engineering Remote Sensing*, 81(4), 281-330.
- Pajares, G., Guijarro, M., Herrera, P. J., Ribeiro, A. (2009). Combining classifiers through fuzzy cognitive maps in natural images. *Computer Vision, IET*, 3(3), 112-123.
- Pajares, G, Cruz, J.M. (2007). *Visión por Computador: Imágenes Digitales y Aplicaciones*, RA-MA, Madrid.
- Pajares, G., Ruz, J. J., de la Cruz, J. M. (2005). Performance analysis of homomorphic systems for image change detection. *Pattern Recognition and Image Analysis* (563-570). Springer Berlin Heidelberg.
- Pajares, G., Cruz de la, J. M. (2004). A wavelet-based image fusion tutorial. *Pattern Recognition*, 37(9), 1855-1872.
- Pajares, G., De la Cruz, J. M., Moreno, V. (2002). Clasificación de texturas naturales mediante agrupamiento borroso. *Ingeniería Civil*, (127), 83-89.

-
- Pajares, G., Cruz, J. M. (2002). Clasificación de texturas naturales mediante k-means. *Revista electrónica de visión por computador*, 6, 1-18.
- Paruelo, J.M., Epstein, H.E., Lauenroth, W.K., Burke, I.C. (1997). ANPP estimates from NDVI for the central grasslands region of the U.S. *Ecology*, 78, 953-958.
- Peña, J. M., Torres-Sánchez, J., de Castro, A. I., Kelly, M., & López-Granados, F. (2013). Weed mapping in early-season maize fields using object-based analysis of unmanned aerial vehicle (UAV) images. *PLoS One*, 8(10), e77151.
- Peña-Barragán, J. M., Ngugi, M. K., Plant, R. E., Six, J. (2011). Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, 115(6), 1301-1316.
- Peña-Barragán JM, Torres-Sánchez J, de Castro AI, Kelly M, López-Granados F. (2013). Weed Mapping in Early-Season Maize Fields Using Object-Based Analysis of Unmanned Aerial Vehicle (UAV) Images. *PLoS ONE*, 8(10), e77151.
- Pérez, A.J., López, F., Belloch, J.V., Christensen, S. (2000). Color and shape analysis techniques for weed detection in cereal fields. *Computer and Electronics in Agriculture*, 25, 197-212.
- Perry, E. M., Davenport, J. R. (2007). Spectral and spatial differences in response of vegetation indices to nitrogen treatments on apple. *Computers and Electronics in Agriculture*, 59(1), 56-65.
- Persson M., Åstrand B., (2008). Classification of crops and weeds extracted by active shape models. *Biosystems Engineering*, 100 (4), 484–497
- Ponti, M. P. (2013). Segmentation of low-cost remote sensing images combining vegetation indices and mean shift. *Geoscience and Remote Sensing Letters, IEEE*, 10(1), 67-70.
- Provost, F., Fawcett, T., Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. Proc. of the 15th International Conference on Machine Learning (445–453). Morgan Kaufmann, San Francisco, CA.
- Pun, T. (1980). A new method for grey-level picture thresholding using the entropy of the histogram. *Signal processing*, 2(3), 223-237.
- Rabatel, G., Gorretta, N., Labbé, S. (2014). Getting simultaneous red and near-infrared band data from a single digital camera for plant monitoring applications: Theoretical and practical study. *Biosystems Engineering*, 117, 2-14.

-
- Ranchin, T., Wald, L. (2000). Fusion of High Spatial and Spectral Resolution Images: The ARSIS Concept and its Implementation. *Photogrammetric Engineering and Remote Sensing*, 66(1), 49-61.
- Reid, J., Searcy, S. (1987). Vision-based guidance of an agriculture tractor. *IEEE Control Systems Magazine*, 7(2), 39-43.
- RHEA 2015. Robot Fleets for Highly Effective Agriculture and Forestry Management. Disponible en línea: <http://www.rhea-project.eu/> (Accedido octubre 2015)
- Ribeiro A., Fernández-Quintanilla C., Barroso J., García-Alegre, M.C. (2005). Development of an image analysis system for estimation of weed. Proc. 5th European Conf. On Precision Agriculture (5ECPA), 169–174.
- Richardson, A. J., Weigand, C. L. (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing*, 43(12).
- Riomoros, I., Guijarro, M., Pajares, G., Herrera, P. J., Burgos-Artizzu, X. P., Ribeiro, A. (2010). Automatic image segmentation of greenness in crop fields. *IEEE Int Conf. Soft Computing and Pattern Recognition (SoCPaR'10)*, 7-10 December, Paris, France, pp. 462-467.
- Romeo, J., Guerrero, J.M., Montalvo, M., Emmi, L., Guijarro, M., González-de-Santos, P., Pajares, G. (2013). Camera sensor arrangement for crop/weeds detection accuracy in agronomic images. *Sensors* 13, 4348-4366.
- Romeo, J., Pajares, G., Montalvo, M., Guerrero, J. M., Guijarro, M., Ribeiro, A. (2012). Crop row detection in maize fields inspired on the human visual perception. *The Scientific World Journal*.
- Rouse, J.W., Haas, R.H., Schell J.A., Deering, D.W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. Proc. Third ERTS-1 Symposium, NASA Goddard, NASA SP-351 pp. 309-317.
- Ruiz-Ruiz, G., Gomez-Gil, J., Navas-Gracia, L.M. (2009). Testing different color spaces based on hue for the environmentally adaptive segmentation algorithm (EASA). *Computer and Electronics in Agriculture*, 68(1). 88-96.
- Russ, T. D., Reeves, A. P. (1999). High accuracy depth measurement using multi-view stereo. Proc. of the Vision Interface Conference, p.p. 103-110.
- Sahoo, P. K., Soltani, S. A. K. C., Wong, A. K. (1988). A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*, 41(2), 233-260.

-
- Sainz-Costa, N., Ribeiro, A., Burgos-Artizzu, X., Guijarro, M., Pajares, G. (2011). Mapping wide row crops with video sequences acquired from a tractor moving at treatment speed. *Sensors*, 11, 7095–7109.
- Sakamoto, T., Shibayama, M., Kimura, A., Takada, E. (2011). Assessment of digital camera-derived vegetation indices in quantitative monitoring of seasonal rice growth. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(6), 872-882.
- Samal, A., Brandle, J. R., Zhang, D. (2006). Texture as the basis for individual tree identification. *Information Sciences*, 176(5), 565-576.
- Shanmugam, M., Asokan, R. (2015). A Machine-Vision-Based Real-Time Sensor System to Control Weeds in Agricultural Fields. *Sensor Letters*, 13(6), 489-495.
- Shapiro, L.G., Stockman, G.C. (2001) *Computer Vision*. Prentice Hall. First edition. ISBN 0-13-030796-3.
- Sezgin, M. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*, 13(1), 146-168.
- Selvarajah, S., Kodituwakku, S. R. (2011). Analysis and comparison of texture features for content based image retrieval. *International Journal of Latest Trends in Computing*, 2(1).
- Seong, J.C., Usery, E.L. (2001). Fuzzy image classification for continental scale multitemporal NDVI series images using invariant pixels and an image stratification method. *Photogrammetric Engineering and Remote Sensing*, 67(3), 287-294.
- Shanmugam, M., Asokan, R. (2015). A Machine-Vision-Based Real-Time Sensor System to Control Weeds in Agricultural Fields. *Sensor Letters*, 13(6), 489-495.
- Shapiro, L. G., Stockman, G. C. *Computer Vision: Theory and Applications*. 2001.
- Shrestha, D.S., Steward, B.L., Birrell, S.J., (2004). Video processing for early stage maize plant detection. *Biosystems Engineering* 89 (2), 119–129.
- Soontranon, N., Srestasathien, P., Rakwatin, P. (2014, May). Rice growing stage monitoring in small-scale region using ExG vegetation index. *Proc. IEEE Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2014 11th International Conference on* (pp. 1-5).
- Sripada, R. P., Heiniger, R. W., White, J. G., Meijer, A. D. (2006). Aerial color infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal*, 98(4), 968-977.
- Tang, X. (1998). Texture information in run-length matrices. *IEEE Trans. on Image Processing*, 7, 1602-1609.
-

- Tang, L., Tian, L. F., Steward, B. L. (2000). Color image segmentation with genetic algorithm for in-field weed sensing. *Transactions of the ASAE*, 43(4), 1019.
- Tang, X. (1998). Texture information in run-length matrices. *IEEE Trans. on Image Processing*, 7 (1998) 1602-1609.
- Tellaache, 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(1), 908–915.
- Tellaache, A., Burgos-Artizzu, X., Pajares, G., Ribeiro, A., Fernández-Quintanilla, C., (2008a.) A new vision-based approach to differential spraying in precision agriculture. *Computers and Electronics in Agriculture*, 60 (2), 144–155.
- Tellaache, A., Burgos-Artizzu, X.P., Pajares, G., Ribeiro, A., (2008b). A vision-based method for weeds identification through the Bayesian decision theory. *Pattern Recognition*, 41, 521–530.
- TheMathworks. 2015. (<http://www.mathworks.com/>) (Accedido en octubre 2015)
- Thorp, K. R., Tian, L. F. (2004). A review on remote sensing of weeds in agriculture. *Precision Agriculture*, 5(5), 477-508.
- Tian, LF., Slaughter, DC. (1998). Environmentally adaptive segmentation algorithm for outdoor image segmentation. *Computer and Electronics in Agriculture*. 21(3). 153-168.
- Tian, Y. W., Li, C. H. (2004). Color image segmentation method based on statistical pattern recognition for plant disease diagnose. *Journal of Jilin University of Technology* (Natural Science Edition), 2, 028.
- Tomita, F., Tsuji, S. (1990). Computer analysis of visual textures. Kluwer Academic Publishers.
- Torres-Sánchez, J., López-Granados, F., Peña, J. M. (2015). An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. *Computers and Electronics in Agriculture*, 114, 43-52.
- Torres-Sánchez, J., Peña, J.M., de Castro, A.I., López-Granados, F. (2014). Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. *Computers and Electronics in Agriculture* 103, 104–113.

-
- Torres-Sánchez J., López-Granados F., De Castro A.I., Peña-Barragán J.M. (2013a). Configuration and Specifications of an Unmanned Aerial Vehicle (UAV) for Early Site Specific Weed Management. *PLoS ONE* 8(3), e58210.
- Torres-Sánchez, J.; Peña-Barragán, J.M.; Gómez-Candón, D.; de Castro, A.I.; López-Granados, F. (2013b). Imagery from unmanned aerial vehicles for early site specific weed management. *Precision Agriculture* (Stafford, J.V.; Ed.), vol. 13, pp. 193-199.
- Tsai, W. H. (1985). Moment-preserving thresholding: A new approach. *Computer Vision, Graphics, and Image Processing*, 29(3), 377-393.
- Tucker, C. J., Justice, C. O., & Prince, S. D. (1986). Monitoring the grasslands of the Sahel 1984-1985. *International Journal of Remote Sensing*, 7(11), 1571-1581.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation, *Remote Sensing of Environment*, 8,127-150.
- Virginia, R.A., Wall, D.H. (2001). Ecosystem Function, Principles of. In *Encyclopedia of Biodiversity*, (ed. Simon A. Levin), pp. 345-352. New York; Elsevier.
- Wang, F. M., Huang, J. F., Tang, Y. L., & Wang, X. Z. (2007). New vegetation index and its application in estimating leaf area index of rice. *Rice Science*, 14(3), 195-203.
- Williams, J. C., Orson, J., Bush, M., Cook, S., Boys, E., Cussans, J. (2011). New EU Pesticide Legislation—the View of a Manufacturer. *Aspects of Applied Biology*, 106, 269-274.
- Woebbecke, D.M., Meyer, G.E., von Bargen, K. and Mortensen, D.A. (1995). Shape features for identifying young weeds using image analysis. *Trans. Am. Soc. Agric. Eng.* 38 (1), 271–281.
- Woods, J.W. and O’Neil, S.D. (1986). Subband Coding of Images. *IEEE Trans. Acustics, Speech. Signal Processing*, 34(5), 1278-1288.
- Yang, L., Xian, G., Klaver, J. M., Deal, B. (2003). Urban land-cover change detection through sub-pixel imperviousness mapping using remotely sensed data. *Photogrammetric Engineering Remote Sensing*, 69(9), 1003-1010.
- Yu, Z., Cao, Z., Wu, X., Bai, X., Qin, Y., Zhuo, W., Xue, H. (2013). Automatic image-based detection technology for two critical growth stages of maize: Emergence and three-leaf stage. *Agricultural and forest meteorology*, 174, 65-84.
- Zarate-Valdez, J. L., Whiting, M. L., Lampinen, B. D., Metcalf, S., Ustin, S. L. Brown, P. H. (2012). Prediction of leaf area index in almonds by vegetation indexes. *Computers and electronics in agriculture*, 85, 24-32.

-
- Zarco-Tejada, P.J., González-Dugo, V., Berni, J.A.J. (2012). Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sensing of Environment*, 117, 322–337.
- Zhang, C., Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agriculture*, 13(6), 693-712.
- Zhang, J., Foody, G.M. (1998). A fuzzy classification of sub-urban land cover from remotely sensed imagery, *Int. J. Remote Sensing*, 19(14), 2721-2738.
- Zhang, Y. J. (1996). A survey on evaluation methods for image segmentation. *Pattern Recognition*, 29(8), 1335-1346.
- Zhang, N., Chaisattapagon, C. (1995). Effective criteria for weed identification in wheat fields using machine vision. *Transactions of the ASAE*, 38(3), 965-974.
- Zhang, Y., Wang, R., Munroe, K. (2004, July). Multi-resolution and multi-spectral image fusion for urban object extraction. *Proc. of ISPRS XXth Congress*, 12-23.
- Zhang, X., Yan, G., Li, Q., Li, Z. L., Wan, H., Guo, Z. (2006). Evaluating the fraction of vegetation cover based on NDVI spatial scale correction model. *International Journal of Remote Sensing*, 27(24), 5359-5372.
- Zheng, L. Shi, D., Zhang, J. (2010). Segmentation of green vegetation of crop canopy images based on mean shift and Fisher linear discriminate. *Pattern Recognition Letters*. 31(9), 920-925.
- Zheng, L.; Zhang, J. Wang, Q. (2009). Mean-shift-based color segmentation of images containing green vegetation. *Computers and Electronics in Agriculture*, 65, 93–98.
- Zhou XH, Obuchowski NA, McClish DK (2002) Statistical methods in diagnostic medicine. Wiley-Interscience.
- Zimmermann, H.J., (1991). Fuzzy Set Theory and its Applications. Kluwer Academic Publishers, Norwell.