

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS FÍSICAS
DEPARTAMENTO DE ARQUITECTURA DE COMPUTADORES Y AUTOMÁTICA



TESIS DOCTORAL

**Técnicas de visión artificial para la segmentación y detección de líneas de cultivo
en imágenes agrícolas**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Martín Montalvo Martínez

Directores

Gonzalo Pajares Martinsanz
Jesús Manuel de la Cruz García
María Guijarro Mata-García

Madrid, 2015

TESIS DOCTORAL

TÉCNICAS DE VISIÓN ARTIFICIAL

PARA LA SEGMENTACIÓN Y

DETECCIÓN DE LÍNEAS DE CULTIVO

EN IMÁGENES AGRÍCOLAS



Martín Montalvo Martínez

Departamento de Arquitectura de Computadores y Automática
Facultad de Informática
Universidad Complutense de Madrid

Madrid 2014

AGRADECIMIENTOS

En primer lugar quiero agradecer a mis tres directores la dedicación, apoyo y ayuda que me han ofrecido a lo largo del tiempo que he estado elaborando la presente tesis. A Gonzalo porque los primeros pasos en el mundo de la investigación fueron gracias a él, consiguió que lo que era complicado pareciera más sencillo. Siempre me ha asesorado por donde debía seguir y me introdujo en el importante proyecto de investigación RHEA (Robot Fleets for Highly Effective Agriculture and Forestry Management) sobre el que se basa el trabajo de tesis. A Jesús Manuel por su buena predisposición, porque siempre que he necesitado algo en lo que me pudiera ayudar siempre he recibido su ayuda. A María porque en los momentos de desaliento una conversación con ella me ha animado a seguir hacia adelante o ver las cosas de otro modo.

Al Departamento de Arquitectura de Computadores y Automática y al grupo de investigación ISCAR.

A la Universidad Complutense por haberme concedido una beca predoctoral que me ha facilitado el desarrollar el trabajo de investigación que ha conducido a la presente tesis doctoral.

A los miembros del Centro de Automática y Robótica, UPM-CSIC, por su apoyo y ayuda.

A todos los amigos que me han acompañado durante estos últimos años y con los que he pasado buenos momentos, tanto a los que conozco desde hace mucho tiempo como a los que he conocido recientemente.

A mi hermano porque siempre está ahí para lo que necesite, me aguanta y me consigue sacar una sonrisa hasta cuando me cuesta sonreír. A mis padres porque todo lo que soy se lo debo a ellos, los esfuerzos que han realizado por mí, la educación que me han dado y los valores que me han inculcado. Porque siendo la familia algo que nos es impuesto, algo que no podemos elegir, he tenido la inmensa suerte de tener una familia excelente.

RESUMEN

El presente trabajo de tesis se enmarca dentro del área conocida como Agricultura de Precisión (AP), cuyo principal soporte es la investigación desarrollada en el proyecto RHEA, perteneciente al VII Programa Marco de la Unión Europea, cuyo principal objetivo es optimizar los recursos en agricultura mediante una flota de robots. Más específicamente, se centra en el uso de la visión artificial como sistema de percepción.

Se presentan dos Sistemas Expertos Automáticos (SEA), basados en técnicas de visión por computador, cuya principal finalidad es la detección de líneas de cultivo y rodales de malas hierba en campos de maíz bajo condiciones ambientales de iluminación adversas y altamente cambiantes, además de abordar sendas problemáticas específicas que se presentan en estos campos de cultivo.

El primero de los SEA está especialmente diseñado para trabajar en campos agrícolas aunque éstos posean una alta densidad de malas hierbas. El segundo, se plantea para dar solución a situaciones complejas en las que las plantas se encuentran enmascaradas por estar cubiertas de una película de barro debido a la lluvia o a la acción de regado artificial.

Los principales objetivos en ambos SEA son la detección de las plantas y su posterior clasificación como malas hierbas o cultivo, determinando la localización de las líneas de cultivo y la posición exacta de los rodales de mala hierba para proceder a la eliminación de estos últimos de forma precisa.

Para alcanzar los objetivos, en el ámbito de la visión por computador, se han desarrollado: *a)* técnicas para regular el tiempo de exposición en la cámara, con el fin de controlar automáticamente los cambios de iluminación natural; *b)* estrategias robustas para la segmentación de imágenes de exterior en tiempo real, válidas para otros entornos de la misma naturaleza; *c)* métodos de umbralización con los que diferenciar entre cultivo y malas hierbas; *d)* métodos para establecer la localización de las líneas de cultivo, que sirven a la vez para un guiado preciso del tractor.

Se describe con detalle la implementación de los sistemas propuestos y su integración en una de las tres unidades de robots que conforman la flota de vehículos en RHEA.

Palabras clave: agricultura de precisión, visión computacional, segmentación de imágenes, índice de vegetación, umbralización de imágenes, discriminación malas hierbas/cultivo, detección de líneas de cultivo, tiempo real, tiempo de exposición, iluminación no controlada, sistema experto automático.

ABSTRACT

The work in this thesis belongs to the area known as Precision Agriculture (PA), whose main support is the research developed under the RHEA project, belonging to the European Union's Seventh Framework Programme, where the main goal is the optimization of resources in agriculture through a fleet of robots. More specifically, it is focused on the use of artificial vision as perception system.

Two Automatic Expert Systems (AES), based on computer vision techniques, are proposed, with the aim of detecting crop rows and weed patches in maize fields under adverse environmental illumination conditions with high variability. Additionally, both systems address two specific problems that occur in these maize fields.

The first SEA is specially designed to work in agricultural maize fields even if they contain high weeds densities. The second is focused to solve complex situations when plants are masked with materials coming from the soil due to rainfall or artificial irrigation.

The main objectives in both AES are the plants detection for their posterior classification as weeds and crops, determining the crop rows location and accurate positioning of weeds patches, for removing these last.

To achieve the objectives, within the scope of computer vision, we have developed: a) automatic techniques to adjust the exposure time of the camera sensor to control the variability under natural illumination conditions; b) robust strategies for outdoor image segmentation in real time; c) image segmentation methods to subsequently develop thresholding methods to discriminate between maize and weeds; d) methods to establish the location of crop rows, useful for precise guidance of the robot.

Detailed implementation of the proposed systems is provided and also their integration in one of the three units belonging to the fleet of agricultural vehicles in RHEA.

Keywords: precision agriculture, computer vision, image segmentation, vegetation index, image thresholding, weeds/crop discrimination, crop line detection, real time, exposure time, uncontrolled illumination, automatic expert system.

ÍNDICE

Índice de Figuras	1
Índice de Tablas.....	3
Lista de abreviaturas	5
1. Introducción	7
1.1. Antecedentes	7
1.2. Descripción del problema	13
1.3. Motivación y objetivos.....	16
1.3.1. Motivación	16
1.3.2. Objetivos	19
1.4. Metodología	19
1.5. Aportaciones de la investigación	21
1.6. Estructuración de la memoria.....	26
2. Estado del arte	29
2.1. Segmentación de la imagen orientada a la obtención de verdes.....	29
2.1.1. Segmentación basada en índices espectrales.....	30
2.1.2. Segmentación basada en umbralizaciones	36
2.1.3. Segmentación basada en aprendizaje	37
2.2. Detección y localización de líneas en imágenes agrícolas	37
2.2.1. Métodos basados en la exploración horizontal de bandas.....	38
2.2.2. Métodos basados en la transformada de Hough.....	39
2.2.3. Métodos basados en el punto de fuga.....	40
2.2.4. Métodos basados en regresión lineal	41
2.2.5. Métodos basados en estereovisión	41
2.2.6. Métodos basados en el análisis de gotas.....	42
2.2.7. Métodos basados en la acumulación de plantas verdes	42
2.2.8. Métodos basados en el análisis de la frecuencia.....	43
3. Fundamentos teóricos en la detección de líneas de cultivo y malas hierbas	45
3.1. Introducción	45
3.2. Tiempo de exposición	46
3.3. Índices de vegetación: detección del verdor	51
3.3.1. Índices de vegetación simples.....	54
3.3.2. Índices de vegetación compuestos.....	57

3.4.	Umbralización.....	58
3.4.1.	Transformación de escala.....	58
3.4.2.	Método de Otsu	59
3.5.	Ajuste de líneas	62
3.5.1.	Consideraciones generales.....	62
3.5.1.1.	Mínimos cuadrados totales.....	64
4.	Sistemas Expertos Automáticos: detección de líneas de cultivo y discriminación entre plantas de cultivo y malas hierbas en maíz.....	67
4.1.	Introducción	67
4.2.	SEA para detección de líneas de cultivo y malas hierbas	69
4.2.1.	Control del tiempo exposición	71
4.2.2.	Segmentación de la imagen	81
4.2.3.	Doble umbralización mediante el método de Otsu	85
4.2.4.	Detección de las líneas de cultivo	90
4.2.5.	Cálculo de las coberturas de malas hierbas	97
4.3.	SEA para detección de plantas de cultivo y malas hierbas enmascaradas	104
5.	Equipamiento y materiales.....	111
5.1.	Introducción	111
5.2.	Descripción del material.....	111
5.2.1.	Cámara.....	112
5.2.2.	Sistema óptico	116
5.2.3.	Inertial Measurement Unit (IMU)	118
5.2.4.	Carcasa	118
5.2.5.	Computadores	119
5.2.6.	Prototipo RHEA.....	122
6.	Análisis de resultados.....	125
6.1.	Introducción	125
6.1.1.	SEA para detección de líneas de cultivo y malas hierbas	125
6.1.2.	SEA para detección de plantas de cultivo y malas hierbas enmascaradas	132
6.1.3.	Implementación del sistema RHEA	141
7.	Conclusiones y trabajo futuro	147
7.1.	Introducción	147
7.2.	Conclusiones.....	147
7.2.1.	Detección de líneas de cultivo y malas hierbas.....	147

7.2.2.	Detección de plantas de cultivo y malas hierbas enmascaradas	148
7.2.3.	Sistema RHEA	149
7.3.	Trabajo futuro	150
ENGLISH SUMMARY		153
8.	English summary	155
8.1.	Introduction	155
8.1.1.	Antecedents	155
8.1.2.	Problem description	159
8.1.3.	Motivation and goals	161
8.1.3.1.	Motivation	161
8.1.3.2.	Goals	163
8.2.	Automatic Expert System for crop row and weed detection	163
8.2.1.	Introduction	163
8.2.1.1.	Problem statement	163
8.2.1.2.	Revision of methods	165
8.2.1.3.	Motivational research and design of the proposed strategy	169
8.2.2.	Materials and methods	170
8.2.2.1.	Images	170
8.2.2.2.	Image acquisition and segmentation	171
8.2.2.3.	Double thresholding	172
8.2.2.4.	Crop row identification	173
8.2.3.	Results	176
8.3.	Automatic Expert System for masked crop and weeds detection	178
8.3.1.	Introduction	178
8.3.1.1.	Problem statement	178
8.3.1.2.	Revision of methods	179
8.3.1.3.	Motivational research of the proposed strategy	181
8.3.2.	Expert system design	182
8.3.2.1.	Reasoning for knowledge extraction	182
8.3.2.2.	Automatic image processing modules	182
8.3.3.	Results	184
8.4.	Conclusions and future work	188
8.4.1.	Introduction	188
8.4.2.	Conclusions	188

8.4.2.1. Crop rows and weeds detection.....	188
8.4.2.2. Masked plants detection.....	188
8.4.2.3. RHEA system.....	189
8.4.3. Future work.....	190
9. Bibliografía/References	193

Índice de Figuras

Figura 1-1. Mapa de prescripción (Imagen cortesía de F. López-Granados y J.M. Peña; Instituto de Agricultura Sostenible, CSIC-Córdoba, España; Peña y col., 2012, 2013).....	10
Figura 1-2. Sistema basado en sensores que no requiere de mapa de prescripción.....	11
Figura 3-1. Sistema de visión con cámara, óptica y filtro de corte: (a) integrado; (b) con sus elementos separados.	48
Figura 3-2. (a) Objetivo con un número-f grande, (b) el mismo objetivo con un número-f más pequeño y (c) objetivo visto desde un plano cenital	49
Figura 3-3. Respuesta espectral en el rango visible del espectro.....	52
Figura 3-4. (a) Imagen de líneas de cultivo tomada sin filtro de corte UV/IR; (b) la misma imagen con filtro de bloqueo UV/IR.....	52
Figura 3-5. (a) Imagen en escala de grises; (b) Histograma asociado a la imagen en (a); (c) Imagen binarizada resultante tras aplicar Otsu.....	62
Figura 3-6. Distintas formas de contabilizar el error producido con respecto a la recta: (a) como distancia vertical; (b) como distancia perpendicular.	65
Figura 4-1. Esquema global del proceso.	69
Figura 4-2: Panel con rectángulos de distintos colores para controlar el tiempo de exposición.	74
Figura 4-3. Cultivo en distintos estados de crecimiento y mostrando ROI: (a) y (b) ROI representada con un rectángulo rojo; (c) y (d) ROI representada mediante proyección de perspectiva	78
Figura 4-4. Esquema de la fase de captura de la imagen	81
Figura 4-5. Esquema para la fase de segmentación	85
Figura 4-6. (a) Imagen segmentada; (b) Imagen binarizada mediante el valor 152; (c) Histograma de la imagen segmentada.....	87
Figura 4-7. (a) Imagen binaria mostrando la vegetación; (b) Imagen binaria mostrando solo el cultivo; (c) Imagen mostrando en verde el cultivo y en rojo las malas hierbas.....	89
Figura 4-8. Esquema para la fase de umbralización	90
Figura 4-9. (a) Imagen binarizada con solo cultivo; (b) plantilla utilizada para la detección de líneas de cultivo.	93
Figura 4-10. Esquema para la fase de detección de líneas.	97
Figura 4-11. Líneas para el mallado: (a) líneas entre dos líneas; (b) región de exclusión.....	100
Figura 4-12. Mallado utilizado para la cobertura vegetal.	102
Figura 4-13. Esquema general de la fase "Cálculo de coberturas".	103
Figura 4-14. Arquitectura modular para el sistema experto diseñado para tratar con plantas enmascaradas.	106
Figura 5-1. (a) Cámara Basler scA 1400-17fc; (b) imagen capturada por la cámara Basler.....	113
Figura 5-2. Curvas de la respuesta espectral para la cámara scA 1400-17fc.	113

<i>Figura 5-3. Cámara AXIS 211M: (a) Vista frontal, (b) vista trasera de la cámara AXIS 211M.....</i>	<i>114</i>
<i>Figura 5-4. Cámara SVS 4050: (a) vista lateral y (b) vista trasera.....</i>	<i>116</i>
<i>Figura 5-5. Curva de operación del filtro pasa banda.....</i>	<i>117</i>
<i>Figura 5-6. Sensor inercial con receptor GPS 3DM-GX3-35.....</i>	<i>118</i>
<i>Figura 5-7. (a) Vista exterior de la carcasa; (b) Vista interior de la carcasa.....</i>	<i>119</i>
<i>Figura 5-8. (a) Esquema con las conexiones del panel frontal de la cRIO 9082; (b) Vista general de la cRIO.</i>	<i>121</i>
<i>Figura 5-9. Vista de la carcasa abierta.....</i>	<i>123</i>
<i>Figura 5-10. Integración de la cRIO en el sistema del tractor en RHEA.....</i>	<i>123</i>
<i>Figura 6-1. Dos ejemplos representativos de los tipos de imágenes: (a) Tipo-1; (b) Tipo-2.....</i>	<i>126</i>
<i>Figura 6-2. Umbralizaciones para la imagen de la Figura 6-1(a): (a) Primera umbralización; (b) Segunda umbralización.....</i>	<i>127</i>
<i>Figura 6-3. Umbralizaciones para la imagen de la Figura 6-1(b): (a) Primera umbralización; (b) Segunda umbralización.....</i>	<i>127</i>
<i>Figura 6-4. Líneas de cultivo detectadas (rojo) a partir de las imágenes mostradas en la Figura 6-1.</i>	<i>128</i>
<i>Figura 6-5. Comparativa de porcentaje de éxito entre RL y Theil-Sen.....</i>	<i>131</i>
<i>Figura 6-6. Imagen representativa del Conjunto-1.....</i>	<i>133</i>
<i>Figura 6-7. Imagen representativa del Conjunto-2.....</i>	<i>134</i>
<i>Figura 6-8. (a) Imagen obtenida al aplicar el índice COM a la Figura 6-6; (b) histograma de la imagen en (a) con cuatro umbrales.....</i>	<i>135</i>
<i>Figura 6-9. (a) Imagen binaria obtenida con el umbral c de la Tabla 6-3 conteniendo plantas enmascaradas y sin enmascarar en blanco; (b) imagen segmentada con SEA con plantas enmascaradas en azul y sin enmascarar en rojo; (c) imagen segmentada con DEM con plantas enmascaradas en azul y sin enmascarar en rojo.....</i>	<i>137</i>
<i>Figura 6-10. (a) Imagen obtenida a partir de la Figura 6-7 aplicando el COM; (b) histograma de (a). ...</i>	<i>138</i>
<i>Figura 6-11. (a) Imagen binaria obtenida con el umbral c de la Tabla 6-3 conteniendo plantas enmascaradas y sin enmascarar en blanco; (b) imagen segmentada con SEA con plantas enmascaradas en azul y sin enmascarar en rojo; (c) imagen segmentada con DEM con plantas enmascaradas en azul y sin enmascarar en rojo.....</i>	<i>139</i>
<i>Figura 6-12. Imágenes con distinto nivel de infestación y crecimiento.....</i>	<i>142</i>
<i>Figura 6-13. Imágenes en las que se han detectado las líneas de cultivo.....</i>	<i>143</i>
<i>Figura 6-14. Imagen con la identificación de las líneas de cultivo, las malas hierbas y el mallado.....</i>	<i>144</i>

Índice de Tablas

<i>Tabla 3-1: Longitudes de onda del espectro visible y luz infrarroja</i>	<i>51</i>
<i>Tabla 4-1: Valores medios de los canales RGB en los rectángulos de color del panel.</i>	<i>74</i>
<i>Tabla 4-2: Media de valores para el primer día</i>	<i>79</i>
<i>Tabla 4-3: Media de valores para el segundo día</i>	<i>79</i>
<i>Tabla 4-4: Media de valores para el tercer día</i>	<i>79</i>
<i>Tabla 4-5: Media de valores para el cuarto día</i>	<i>79</i>
<i>Tabla 4-6: Media de valores para el quinto día</i>	<i>79</i>
<i>Tabla 4-7: Valores de los coeficientes para la obtención del nuevo índice COM</i>	<i>107</i>
<i>Tabla 6-1: Rendimiento de los enfoques HOU y RL medido en términos de efectividad y tiempo de proceso</i>	<i>129</i>
<i>Tabla 6-2: Porcentaje de efectividad y tiempos de proceso para diferentes aproximaciones</i>	<i>131</i>
<i>Tabla 6-3: Valores de los umbrales obtenidos mediante SEA y DEM para los dos conjuntos analizados</i>	<i>135</i>
<i>Tabla 6-4: Media de las componentes RGB para plantas enmascaradas y sin enmascarar según SEA, DEM y SVM</i>	<i>139</i>
<i>Tabla 6-5: Porcentaje de éxito obtenido por los distintos enfoques para los dos conjuntos de imágenes</i>	<i>140</i>

Lista de abreviaturas

AP	Agricultura de Precisión
SSCM	<i>Site-Specific Crop Management</i> (Gestión del cultivo dependiente del lugar de tratamiento)
VRA	<i>Variable-Rate Application</i> (Aplicación de tasa variable)
LPG	<i>Liquefied Petroleum Gas</i> (Gas licuado del petróleo)
cRIO	Sistema de procesamiento CompactRIO
SEA	Sistema Experto Automático
IMU	<i>Inertial Measurement Unit</i> (Unidad de medición inercial)
LVQ	<i>Learning Vector Quantization</i> (Aprendizaje mediante cuantización vectorial)
VI	<i>Vegetation Index</i> (Índice de vegetación)
NIR	<i>Near-Infrared</i> (Infrarrojo cercano)
ExG	Índice de vegetación Excess Green
CCD	<i>Charge-Couple Device</i> (Dispositivo de carga acoplada)
EV	<i>Exposure Value</i> (Valor de exposición)
ROI	<i>Region Of Interest</i> (Región de interés)
ROIT	<i>Region Of Interest for Treatment</i> (Región de interés a tratar)
SVM	<i>Support Vector Machine</i> (Maquina de Vectores Soporte)
RL	<i>Regresión Lineal</i>

Capítulo 1

1. Introducción

1.1. Antecedentes

El continuo auge y desarrollo de los computadores ha permitido la automatización de muchos procesos de forma eficaz. El computador ha pasado de ser un elemento de lujo al alcance de una minoría, destinado para unos fines muy concretos y relacionados con la investigación, a ser un elemento imprescindible en la vida diaria a todos los niveles.

La agricultura no ha sido menos, siendo uno de los sectores donde el uso de computadores se está imponiendo paulatinamente. La maquinaria agrícola moderna está equipada con dispositivos computacionales de altas prestaciones, logrando así una mayor eficiencia en la producción final, a la vez que se consigue un menor esfuerzo por parte de los agricultores en la realización de las tareas agrícolas.

Tradicionalmente, cuando ha sido necesaria la realización de una tarea agrícola, ésta se ha basado en la aplicación de las técnicas correspondientes en toda la superficie de terreno disponible para ello. Pensemos por ejemplo en la aplicación de herbicida en un campo de cultivo con el fin de eliminar las malas hierbas existentes en el campo para favorecer el crecimiento y desarrollo del cultivo. Es habitual ver en estos campos un tractor aplicando herbicida de forma homogénea y uniforme por toda la parcela mediante la correspondiente maquinaria. Pensemos en este sentido que en la propia parcela, pueden existir rodales de malas hierbas con distintas densidades de éstas o incluso zonas carentes de malas hierbas. Mediante aplicaciones los tratamientos se realizan por igual en todas las zonas de la parcela, independientemente de la necesidad que requiera cada zona de forma particular. Lo ideal sería pensar que las zonas con alta densidad necesitarían un tratamiento más agresivo que las de menor densidad, mientras que las zonas carentes de malas hierbas no requerirían la aplicación de ninguna dosis de herbicida.

Volviendo a los tratamientos no específicos de herbicidas, un modo de tratar de forma uniforme todo el campo, sin aplicar una tasa variable de producto, es mediante una barra pulverizadora convencional que aplica un compuesto químico, mezclado con agua desde un depósito, utilizando boquillas de pulverización instaladas sobre la barra y una válvula con la que se regula la presión del líquido. El control de esta válvula no se modifica mientras se realiza la labor, por lo que la dosis resulta uniforme para todo el campo siempre que se mantenga una velocidad constante del vehículo. Es habitual aprovechar el control de la válvula por la velocidad del vehículo o alternativamente por las revoluciones del motor, para aplicar una dosis variable del tratamiento. El cambio de velocidad o revoluciones lo realiza el operario de forma manual para incrementar o disminuir la cantidad de dosis aplicada en función de lo que él mismo aprecie visualmente sobre el campo. Lo dicho hasta aquí es extrapolable a cualquier otra actividad agrícola en esta línea, véase abonado, riego u otros tratamientos.

El hecho de no diferenciar áreas de cultivo con distintas necesidades de tratamiento implica una serie de desventajas que se sintetizan en: *a)* un mayor coste económico de producción al desperdiciar herbicidas, abonos, agua u otros elementos utilizados en el tratamiento del cultivo, con el agravante de que alguno de ellos resulta excesivamente caro desde el punto de vista económico (herbicidas, abonos) y otros a veces escasos (agua); *b)* incremento del tiempo de trabajo, consumido en parte en tratamientos innecesarios; *c)* un mayor impacto medioambiental, debido a que algunos compuestos químicos utilizados en los tratamientos del cultivo son dañinos para el medio ambiente, filtrándose por las capas del terreno, llegando a contaminar las corrientes de agua subterráneas o acuíferos.

El hecho de reducir la cantidad de agroquímicos, utilizados en los tratamientos, mediante su aplicación selectiva no conlleva una merma en la efectividad del tratamiento y por tanto en la calidad y cantidad del producto final que con ello se pretende, ya que la idea es aplicar dosis correctas allí donde se necesite. (Earl y col., 1996; Krishna, 2013). Esto, unido a las ventajas y beneficios antes mencionados, hace que se tienda a tratamientos específicos en el ámbito de lo que se conoce como Agricultura de Precisión (AP), donde los computadores juegan un extraordinario papel en la automatización de los procesos involucrados.

La nueva forma de cultivar la tierra, mediante técnicas de AP, nos ofrece nuevas posibilidades que vienen siendo consideradas desde hace ya algunos años (Srinivasan, 2006 y Stafford, 2000), encontrándose en continua evolución y desarrollo como se expondrá a lo largo de este trabajo. Mientras que la agricultura tradicional se basa en tratamientos no específicos, la AP realiza una gestión del cultivo dependiente del lugar de tratamiento (SSCM, *site-specific crop management*). Las tareas desarrolladas bajo la perspectiva de la AP facilitan la labor de los operarios, incidiendo positivamente en la mejora de su calidad de vida sin mermar la productividad.

Para conseguir una SSCM se pueden utilizar aplicaciones basadas en lo que se conoce como tasa variable (VRA, Variable-Rate Application) (Zhang y col., 2010; Dammer, 2010; Dillon y Kusunose, 2013; Liakos y col., 2013) que son capaces de diferenciar zonas en las que se debe actuar de distintos modos. Existen distintos métodos VRA disponibles, algunos de los cuales utilizan información GPS. Básicamente se pueden agrupar en las tres categorías siguientes: manuales, tecnologías basadas en mapas y tecnologías basadas en sensores.

En las tecnologías manuales el control sobre la tasa o dosis de tratamiento a realizar o aplicar es regulado por un operador. Este operador es el encargado de variar la tasa mientras se desarrolla la tarea. Las decisiones que toma el operador sobre cómo regular la actuación pueden estar basadas únicamente en su criterio de percepción visual o tomarse con la ayuda de otros elementos auxiliares tales como mapas o sensores.

Las tecnologías basadas en mapas ajustan la cantidad de dosis a administrar o el volumen de trabajo a realizar basándose en un mapa digital, también llamado mapa de *prescripción*. Haciendo uso de la posición actual del vehículo en el campo, obtenida mediante un receptor GPS, y de un mapa de prescripción donde figuren las dosis o tareas específicas deseadas para cada zona, se cambia la concentración de la dosis o la actuación en función de la posición donde se encuentre el sistema de actuación. Los mapas se crean mediante mediciones realizadas sobre el terreno con anterioridad al tratamiento o tarea a realizar. Estas mediciones tienen como objetivo obtener valores sobre distintas propiedades o características, en función de la finalidad para la que se quiera el mapa de prescripción. Algunas de las propiedades o características que se pueden medir en función del tipo de elemento o estructura son las siguientes (Christensen y col., 2005; Piikki y col., 2013; Serrano y col., 2013; Silva y Molin, 2013;

Weis y col., 2013): *a)* tipos y densidad de plantas, color y textura; *b)* tipo de suelo, humedad, pH o conductividad eléctrica del suelo; *d)* rendimiento de las cosechas; *e)* nutrientes o nivel de nitrógeno. Algunas estrategias se centran solamente en una característica para crear el mapa de prescripción, mientras que otras combinan la información de distintas fuentes de información, incluyendo imágenes aéreas tomadas desde vehículos aéreos no tripulados (Berni y col., 2009a; Torres-Sánchez y col., 2013) o imágenes de satélite (Gómez-Candón y col., 2012; Meng y col., 2013).

Independientemente del número de características que se utilicen, el sistema debe de ser capaz de determinar la posición del vehículo, con el que se realiza el tratamiento, dentro del campo y asociar a esa posición una dosis de aplicación. Con estas asociaciones de dosis-posición se crea el mapa de prescripción que se utilizará posteriormente. En la Figura 1-1 se muestra un ejemplo de un mapa de prescripción con las diferentes zonas a tratar, indicando mediante las diferentes tonalidades de color la tasa de aplicación que requiere cada zona.

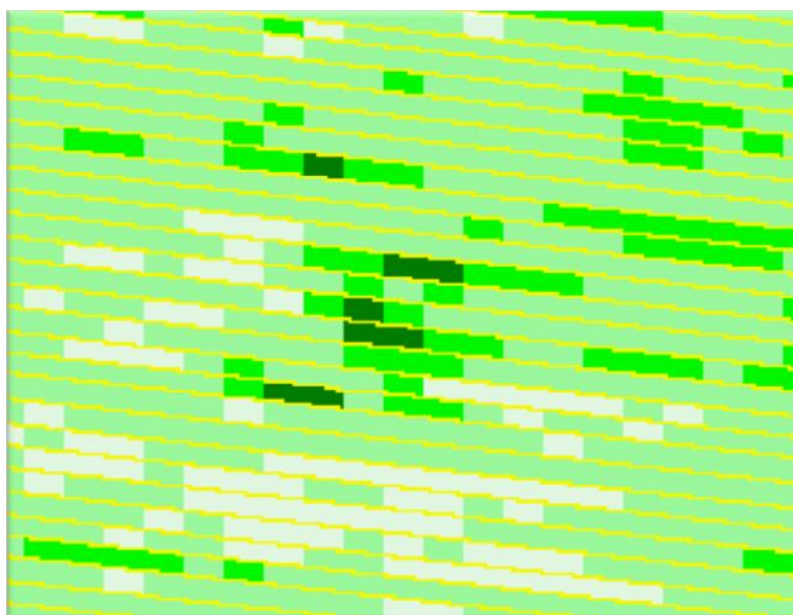


Figura 1-1. Mapa de prescripción (Imagen cortesía de F. López-Granados y J.M. Peña; Instituto de Agricultura Sostenible, CSIC-Córdoba, España; Peña y col., 2012, 2013)

Las tecnologías basadas en sensores, que captan y procesan la información en los instantes previos al tratamiento, no requieren de un mapa o de un sistema de posicionamiento para aplicar las dosis de producto necesarias para dicho tratamiento. Los sensores situados en el vehículo son capaces de medir propiedades del terreno o características del cultivo sobre la marcha. Esta información se proporciona al

controlador, previo paso por un computador donde se procesa la misma a la vez que se toman las decisiones pertinentes. Mediante esta re-alimentación continua de información, el sistema de decisión y control se encargará de seleccionar las actuaciones a realizar sobre el campo y transmitirá estas órdenes a las herramientas que efectuarán el tratamiento. Si bien, en general los sistemas basados en sensores no requieren de un sistema de posicionamiento ni de información previa para realizar su función, si se dispone de un sistema de posicionamiento los datos se pueden almacenar georeferenciados. Esta información puede emplearse para crear un mapa de prescripción que se utilice en futuras operaciones, ya sean diferentes o como una base de datos que registre la evolución a lo largo de los años con fines de análisis y series históricas. En la Figura 1-2 se muestra un sistema basado en sensores que no requiere de un mapa de prescripción. Este sistema, constituido entre otros sensores por un sistema de visión, se encuentra instalado en la parte superior de un tractor agrícola. El sistema de visión será el encargado de detectar las malas hierbas para posteriormente, tras el procesamiento de las imágenes y las decisiones tomadas al respecto, indicar al actuador dónde y cómo debe activarse en función de las densidades de malas hierbas encontradas.

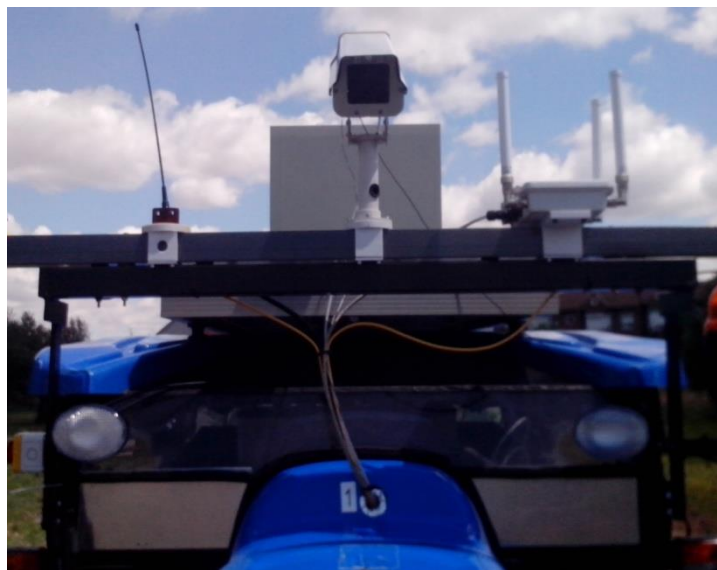


Figura 1-2. Sistema basado en sensores que no requiere de mapa de prescripción

Debido a que las tecnologías basadas en mapas poseen ventajas que no disponen las basadas en sensores y viceversa se propone, como una idea posible, para futuros desarrollos combinar las dos tecnologías en un único sistema SSCM con el que explotar los beneficios de ambas.

Dependiendo de la actividad agrícola que se quiera realizar, las operaciones basadas en VRA se pueden clasificar entre otras en: sembrado (Dillon, 2013; Heege, 2013*a,b*; Shearer y Pitla, 2014), control de malas hierbas, mohos u hongos (Dammer, 2010; Sökefeld, 2010; Wu y col., 2011; Andújar y col., 2013; Montalvo y col., 2013), administración de cal (Adamchuk y col., 2004; Havlin y Heininger, 2009), administración de fósforo (Mallarino y Wittry, 2006), administración de nitrógeno (Scharf y col., 2009; Chan, 2012; Dillon y Kusunose, 2013; Liakos y col., 2013) o riego (Al-Kufaishi y col., 2006).

Todas estas operaciones, basadas en sistemas sensoriales, permiten disponer de la información necesaria con fines de actuación específica. Los sensores en AP son elementos esenciales para operaciones y tratamientos selectivos. En base a la información disponible identifican y detectan los elementos y las zonas sobre las que actuar. Dentro de los sensores, el sistema de visión artificial ha sido ampliamente utilizado en AP con fines de detección, siendo por otra parte el núcleo de los trabajos de investigación que se describen en el presente trabajo de tesis.

Además de los sistemas de detección para tratamientos selectivos, otra de las aportaciones y desarrollos dentro de la AP son los sistemas de guiado automático para vehículos. Mediante un sistema de guiado de tractores automático (Al-Hamed, 1996; Wilson, 2000; Shockley y Dillon, 2008; Heege, 2013; Zhou y col., 2013) se podría eliminar la necesidad de una persona para conducir el vehículo suprimiéndose de este modo restricciones horarias, límites de tiempo o incluso de visibilidad. Con el guiado automático se podría trabajar durante largas jornadas de forma continua sin requerir el descanso del conductor o su relevo por otra persona. El trabajo no se vería afectado por condicionantes ambientales adversas, salvo que éstas impidieran la realización de las tareas agrícolas, ya que se puede utilizar un sistema de navegación por satélite (GNSS) como GPS (2014), GLONAS (2014) o el futuro sistema civil europeo Galileo (2014). Con estos sistemas se puede delimitar el campo y trazar la trayectoria a seguir, junto con sus puntos críticos (lugar de entrada y salida del campo, localización de obstáculos tales como torres eléctricas, grandes rocas dentro del campo, elementos de riego, etc.). Estas consideraciones son necesarias para conseguir una navegación segura, así como la realización de las tareas agrícolas de forma eficiente. El posicionamiento del vehículo puede realizarse mediante visión artificial, que puede servir de complemento a los

sistemas GNSS (Spacek y Burbridge, 2007; Subramanian y col., 2006; Ming y col., 2012).

Otra motivación para desarrollar sistemas de guiado autónomo o semiautónomo surge de la necesidad de evitar que el operador realice correcciones continuas en la dirección de avance del tractor mientras remolca el instrumento de trabajo para conseguir un nivel de rendimiento aceptable. El operador realiza dos funciones básicas, a saber: el guiado del vehículo y el manejo de los utensilios o maquinaria acoplados al propio vehículo. La función del operador no ha cambiado de forma sustancial en décadas. Sin embargo tanto los tractores como la maquinaria empleada han sufrido grandes desarrollos durante ese mismo periodo de tiempo. Los tractores han aumentado su potencia, fuerza, velocidad, a la vez que han incrementado sus prestaciones. Debido a lo anterior y al uso de utensilios cada vez más grandes, potentes y complejos, el control de éstos y del tractor en general es una tarea que entraña dificultad para un operario instalado en la cabina del tractor. Esto se debe a la dificultad encontrada a la hora de determinar visualmente por donde circula el tractor y el utensilio, si se están solapando zonas ya tratadas o si se deja de cubrir una zona de la parcela. Se constata así la necesidad de incluir dispositivos capaces de abordar esta problemática de forma eficiente, a la vez que desde el punto de vista de la AP se investigan técnicas y procedimientos con fines de detección y navegación eficaces, que pasan necesariamente por la incorporación de sistemas sensoriales y de actuación, en conjunción con los procesadores necesarios para el tratamiento de la información y la toma de decisiones.

1.2.Descripción del problema

Dentro del amplio mundo de la AP en este trabajo de investigación nos centramos en los sistemas de visión con fines de detección de cultivos y malas hierbas en campos de maíz para identificación de zonas de tratamiento selectivo. El desarrollo de métodos de procesamiento, con la finalidad expresada previamente, constituye el núcleo central de la investigación. Estas tareas se enmarcan dentro de los proyectos RHEA y AGL2011, que se describen en detalle más adelante en la sección 1.3.1.

La investigación se desarrolla en campos de cultivos de maíz de dimensiones suficientes. Conviene destacar que desde el punto de vista de la eficacia de los sistemas

automáticos, la dimensión de los campos debe garantizar la operatividad de aquellos, a la vez que resulten rentables desde el punto de vista de la inversión y la productividad. Téngase en cuenta también a este respecto que dichos sistemas están pensados para actuaciones prolongadas en el tiempo con la mínima supervisión humana posible.

En las etapas iniciales de crecimiento del maíz, su desarrollo se ve influenciado de manera muy importante por la existencia de malas hierbas. Éstas suponen una competencia muy fuerte porque se aprovechan de los nutrientes del suelo en detrimento del maíz. Si no se eliminan continuarán con tal aprovechamiento a la vez que van impidiendo la recepción de luz solar por parte de las plantas de maíz. Además, en esos estados iniciales, las malas hierbas crecen generalmente a una velocidad mayor que el maíz y pueden evitar que el cultivo se desarrolle como lo haría en ausencia de las mismas. Para evitar los problemas antes mencionados es preciso eliminar las malas hierbas en etapas tempranas del crecimiento, a la vez que se preservan las plantas de maíz.

En los proyectos RHEA y AGL2011 se plantea el problema de identificar las malas hierbas existentes en el campo de maíz para aplicar un tratamiento selectivo sobre ellas en la dosis requerida en función de su densidad en el campo. El cultivo de maíz se encuadra dentro de los denominados cultivos de línea ancha, de suerte que la separación entre líneas de cultivo es de 75cm. Los espacios existentes entre cada dos líneas pueden estar cubiertos o no con plantas de malas hierbas. El sistema de visión es responsable de su detección, localización y cuantificación con el fin de aplicar la dosis precisa allí donde se necesite. En este proceso, la identificación de las líneas de cultivo resulta crucial, ya que las malas hierbas se sitúan entre ellas, por lo que su identificación constituye el primer paso del proceso. Una vez localizadas, el siguiente paso consiste en identificar las malas hierbas existentes entre las líneas y su cuantificación. Adicionalmente, las líneas de cultivo, una vez detectadas, sirven para el guiado preciso del vehículo en los momentos en que éste se desvíe, bien por errores en la trayectoria debidos a la imprecisión del GPS o por otras circunstancias tales como que el propio vehículo tienda a desviarse por la tracción de los elementos de actuación mecánica.

La forma física de eliminar las malas hierbas en RHEA se realiza de dos formas, una mediante un procedimiento de arado actuando entre las líneas de cultivo y otra mediante llamaradas que queman y debilitan en gran medida las malas hierbas sin

apenas afectar a la planta de maíz (Fontanelli y col., 2011). En función de la cantidad de malas hierbas en un área localizada, se aplicará una llamarada con mayor o menor intensidad. Estas llamaradas se generan mediante quemadores alimentados con Gas Licuado del Petróleo (LPG). Cuanto mayor sea la concentración de malas hierbas en un área específica, mayor será la presión administrada a los quemadores y por tanto la intensidad de la llamarada. La aplicación de las llamas no tienen ningún efecto negativo para el maíz, sin embargo, la mala hierba es destruida o termina muy debilitada. El proceso de detección, mediante un sistema de visión a bordo de un tractor, tanto de las líneas de cultivo como de las malas hierbas, no está exento de problemas y dificultades, máxime teniendo en cuenta que se trabaja en entornos de exterior donde existen diversos problemas relacionados con la iluminación y otros factores derivados; así como por el hecho de que el tractor se encuentra en continuo movimiento realizando tareas de tracción mecánica. Los procedimientos basados en visión por computador deben ser lo suficientemente robustos para trabajar en condiciones adversas de iluminación. En este sentido conviene señalar que existen días con diferentes niveles de iluminación (soleados, nubosos), iluminación variable a lo largo del día (mañana, tarde, mediodía) e incluso altamente variable con alternancia repentina de nubes y claros. A lo anterior, se añaden las radiaciones infrarrojas y ultravioletas que influyen en la calidad de la imagen. Además, dichos procedimientos de visión por computador deben estar diseñados de forma suficientemente flexible como para asumir movimientos producidos por la irregularidad de los terrenos donde se mueve el tractor, así como desviaciones de éste respecto de su planificación de trayectoria inicialmente prevista.

Por otro lado, en relación a la detección de vegetación, sucede que en condiciones normales durante el proceso del tratamiento, tanto el cultivo como las malas hierbas son verdes, lo que significa que para el sensor de visión, basado en la captura de las componentes espectrales procedentes de la escena, resultan indistinguibles, salvo por determinados matices que se deben aprovechar. Otro problema, también de naturaleza espectral, al que nos enfrentamos en la detección de plantas de cultivo y malas hierbas se debe al hecho de que las hojas de maíz y malas hierbas a veces aparecen cubiertas parcialmente por una película de barro procedente del suelo debido a una fuerte lluvia o como consecuencia del regado del terreno. Ante esta situación la identificación de las plantas se complica enormemente debido a que parte de la vegetación se enmascara frente al sensor, presentando una tonalidad más

próxima al rojizo, asociado con la naturaleza arcillosa del suelo, que a los verdes que en teoría les corresponden.

En la investigación desarrollada se han considerado los anteriores factores, diseñando estrategias y procedimientos para solucionar tal problemática, que constituyen el eje central de la investigación para la detección de líneas de cultivo y malas hierbas.

Adicionalmente, el proceso que aborda la identificación de plantas enmascaradas frente a las no enmascaradas, se ha diseñado de forma que es capaz de discriminar entre plantas con alta actividad clorofílica y aquellas otras con su actividad clorofílica mermada o inexistente. Este último hecho ocurre cuando tras un tratamiento de eliminación de malas hierbas, éstas entran en proceso de mortalidad, disminuyendo su actividad clorofílica, lo que conduce a modificar su verde inicial hacia otras tonalidades propias de las plantas secas. Gracias a esta discriminación es posible determinar la eficacia del tratamiento efectuado, lo que constituye un aporte extra al resultado de la investigación.

1.3.Motivación y objetivos

1.3.1. Motivación

El trabajo de investigación desarrollado tiene su origen en los proyectos de investigación que se mencionan a continuación, enmarcados dentro de la investigación del grupo ISCAR (2014), oficialmente reconocido por la Universidad Complutense de Madrid desde el año 2006. Algunos de ellos en colaboración con otros grupos de investigación tanto de carácter nacional como internacional, tal y como se refleja a continuación.

1. La actividad investigadora se inicia en el proyecto RHEA (2014) financiado por la Comunidad Europea dentro del VII Programa Marco de Investigación y Desarrollo tecnológico (FP7) cuya duración se extiende en el cuatrienio 2010-2014, con el título *Robot Fleets for Highly Effective Agriculture and Forestry Management* NMP-CP-IP 245986-2 RHEA. El FP7 es el principal instrumento de la Unión Europea en materia de financiación de la investigación en Europa para el periodo 2007-2013. En el FP7 destacan como objetivos estratégicos: *a)* el refuerzo de la base científica y tecnológica de la industria europea; *b)* el favorecer la

competitividad internacional mediante la formación de consorcios transnacionales entre la industria y centros de investigación. RHEA está constituido por un consorcio formado por diecinueve grupos de trabajo pertenecientes a quince organizaciones de ocho países europeos con un presupuesto de 8.69 millones de euros.

El proyecto RHEA se enmarca en agricultura dentro del área temática de alimentación, agricultura y pesca y biotecnología. El objetivo consiste en diseñar tres prototipos que forman parte de una flota de tractores robotizados y coordinados, en conjunción con vehículos aéreos no tripulados “drones”. Cada uno de los prototipos robotizados tiene prevista su actuación en tres campos de cultivo, a saber: maíz, cereal y olivar. Los tres tractores se encuentran equipados con distintos tipos de sensores con el objetivo de realizar navegación autónoma y localización de las áreas de interés, también de forma autónoma, para tratamientos selectivos con el fin de eliminar malas hierbas en cereal y maíz y combatir plagas en el olivar, minimizando el uso de agroquímicos, a la vez que se reducen los costes de producción ahorrando cantidades de agroquímicos, tiempo y energía por los tratamientos selectivos.

La investigación desarrollada y expuesta en esta memoria de tesis se centra en el tractor dedicado al maíz, que está equipado con un sistema de visión instalado en su parte superior mediante el cual se llevan a cabo procesos tanto de identificación de líneas de cultivo como de densidades de malas hierbas entre las citadas líneas (Figura 1-2). La investigación desarrollada en este proyecto se realiza mediante un contrato de trabajo asociado al mismo con una duración de seis meses, continuada mediante el disfrute de una beca pre-doctoral concedida por la Universidad Complutense de Madrid, uno de cuyos avales para su concesión fue este mismo proyecto.

Los diseños finales tienen en cuenta la simplicidad de los procesos para facilitar su aplicación por parte de los agricultores como usuarios finales, para quienes las cuestiones técnicas derivadas de la investigación son transparentes. Además, quedan abiertos para su adaptación hacia otras tareas no contempladas en el proyecto como es el caso de la evaluación de la eficacia de los tratamientos aplicados, para lo cual se ha desarrollado un procedimiento específico.

2. La investigación se desarrolla simultáneamente en el marco del proyecto titulado *Sistema de inspección terrestre en vehículos autónomos y su aplicación efectiva a la detección de malas hierbas y su control localizado* (AGL2011-30442-C02-02), perteneciente igualmente al Plan Nacional de I+D+i para el trienio 2011/2014, AGL (2011). Se trata de un contrato de colaboración entre la Universidad Complutense, el Centro de Automática y Robótica, centro mixto CSIC-Universidad Politécnica de Madrid, titular del proyecto coordinado conjuntamente, así como con el Instituto de Agricultura Sostenible, también del CSIC. El objetivo principal estriba en la localización de malas hierbas dentro de los campos de cultivo de maíz y cereal. Como en el caso anterior, la finalidad es similar a la planteada en RHEA, si bien las técnicas relativas al procesamiento de imágenes se centran más en la localización espacial de estructuras que en la identificación. La localización es relativa a los sistemas de referencia locales en el vehículo. La actividad dentro de este proyecto se circunscribe a las actividades asignadas dentro de las tareas previstas como consecuencia de la participación como miembro oficial del mismo, aprobado por el Ministerio de Economía y Competitividad en la concesión del proyecto.

3. Proyecto del Plan Nacional de I+D+i titulado *Sistema de Vigilancia, Búsqueda y Rescate en el Mar mediante colaboración de Vehículos Autónomos Marinos y Aéreos* (DPI2009-14552-C02-01), DPI (2009), prorrogado hasta finales de 2013. En este proyecto se plantea como una de las tareas fundamentales la colaboración de vehículos tanto aéreos como marinos. Estos vehículos están dotados con sistemas de visión artificial operando en entornos de exterior adversos, haciéndose necesario que tales sistemas sean capaces de reconocer texturas y objetos en la superficie del agua y sus proximidades tanto para el posicionamiento de los barcos como para la localización de elementos de interés por parte de los vehículos aéreos. La actividad investigadora desarrollada en este proyecto se centra en la asesoría para la aplicación de los métodos investigados en RHEA y AGL2011, relativos al tratamiento de la iluminación, para su adaptación a estos entornos, que por otra parte poseen características similares a los agrícolas en este sentido.

1.3.2. Objetivos

De acuerdo a todo lo expresado en los apartados anteriores y sin perder de vista la necesidad de cubrir las expectativas y necesidades requeridas en los proyectos anteriores, se plantean los siguientes objetivos de investigación:

1. Estudio bibliográfico de la literatura existente sobre agricultura de precisión usando como medio de detección la visión por computador.
2. Diseñar métodos para controlar el tiempo de exposición del sensor de modo que se consigan capturar imágenes de calidad suficiente para su posterior análisis y procesamiento.
3. Diseñar estrategias robustas para la segmentación de imágenes de exterior en tiempo real, con especial énfasis en las imágenes agrícolas.
4. Analizar los resultados de la segmentación de imágenes para posteriormente diseñar métodos de umbralización con los que diferenciar fácilmente entre el cultivo de maíz y las malas hierbas que se quieren eliminar.
5. Diseñar métodos para establecer la localización de las líneas de cultivo.
6. Comprobar que todos los procesos anteriores se pueden llevar a cabo en tiempo real, a medida que avanza el vehículo agrícola con su sistema de visión.
7. Diseñar procedimientos para identificar la eficacia de los métodos utilizados en la eliminación de las malas hierbas mediante el análisis de las plantas que se están secando en la etapa del post-tratamiento.
8. 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.
9. Divulgar y difundir los resultados de la investigación mediante publicaciones científicas u otros mecanismos.

1.4. Metodología

En el desarrollo del trabajo se han seguido las directrices marcadas por el método científico. Inicialmente, se ha realizado una búsqueda de información bibliográfica en todas las fuentes disponibles, utilizando principalmente para ello el acceso a través de la red del Consorcio Madroño al que están suscritas las Universidades de la Comunidad de Madrid. Se han estudiado los antecedentes metodológicos y se ha analizado y evaluado el estado del arte en este área. Seguidamente, se han dado los

pasos que se indican a continuación, a la vez que se han revisado de manera continua las fuentes bibliográficas hasta el final de la investigación e incluso en el momento de la redacción de este trabajo. Las etapas seguidas han sido:

- Definición del problema de investigación en función tanto de las necesidades identificadas como de la motivación establecida previamente.
- Planteamiento de los objetivos de la investigación.
- Formulación de planteamientos verificables.
- Obtención de datos experimentales.
- Clasificación, análisis e interpretación de los datos.
- Conclusiones generales.
- Descripción de líneas futuras de investigación.

Para la realización de los experimentos se han implementado los métodos basados en las técnicas propuestas. Una primera aproximación de todo lo investigado se desarrolló en Matlab (MathWorks, 2014) ya que es un programa muy útil como herramienta de prototipado relativamente rápido. Una vez que se realizan las primeras pruebas y resultan satisfactorias se implementan los algoritmos que tratan las imágenes en el lenguaje de programación C. Estos algoritmos se encapsulan como una librería de enlace dinámico. Finalmente todo el sistema se diseña y programa en el entorno de programación LabView haciendo uso de las librerías de enlace dinámico anteriores. LabView es un entorno de desarrollo para un lenguaje de programación visual propiedad de National Instruments (2014) especialmente pensado para adquisición de datos y automatización industrial. El sistema final corre en un sistema embebido CompactRIO (cRIO) también de la empresa National Instruments. El proceso de realización de los experimentos abarca desde la adquisición de datos para la creación de una base suficientemente amplia con la que experimentar, hasta la obtención de resultados, pasando por su visualización.

1.5. Aportaciones de la investigación

La contribución de la presente tesis es un proceso que engloba una serie de tareas que abarcan desde la adquisición de imágenes de calidad en exteriores, detección de la vegetación existente, diferenciación entre la vegetación de distintos tipos de plantas, extracción de las líneas de cultivo, determinación de la cobertura vegetal de mala hierba existente en el campo de cultivo, hasta la estimación del éxito conseguido en la etapa de eliminación de las malas hierbas.

La importancia de las aportaciones está respaldada por las siguientes publicaciones en distintas revistas y congresos internacionales.

Revistas

Martín Montalvo, G. Pajares, J.M. Guerrero, J. Romeo, M. Guijarro, A. Ribeiro, J.J. Ruz, J.M. Cruz. 2012 Automatic detection of crop rows in maize fields with high weeds pressure. *Expert Systems with Applications*, 39(15), 11889-11897.

Se propone un método orientado a la detección de líneas de cultivo en imágenes capturadas en campos de maíz. Los campos a los que está dirigida la investigación presentan la característica de poseer una alta densidad de malas hierbas, lo que en trabajos previos habituales suponía una dificultad añadida. El sistema consiste en una cámara montada en la parte superior de un tractor. El método permite trabajar en tiempo real, mientras el tractor navega por el campo sin tener que modificar su velocidad de laboreo habitual. El procesamiento de la imagen se divide en tres etapas principales: la segmentación de la imagen, la umbralización y la detección de las líneas de cultivo. La principal contribución consiste en la umbralización y la detección de líneas. En la umbralización se obtienen tres umbrales, mediante el método de Otsu, con los que se diferencia entre cultivo, malas hierbas y el resto de elementos presentes en el campo. La detección de las líneas de cultivo se lleva a cabo una vez identificados todos los píxeles que constituyen el cultivo. Mediante técnicas de estimación de parámetros basadas en mínimos cuadrados totales se obtienen los parámetros de las ecuaciones de las rectas en la imagen que representan las líneas de cultivo en la realidad. Este sistema se describen en su totalidad en el capítulo 4, Sistemas Expertos Automáticos: detección de líneas de cultivo y discriminación entre plantas de cultivo y malas hierbas en maíz, sección 4.2 SEA para detección de líneas de cultivo y malas hierbas.

Martín Montalvo, J.M. Guerrero, J. Romeo, L. Emmi, M. Guijarro, G. Pajares. 2013. Automatic expert system for weeds/crops identification in images from maize fields. *Expert Systems with Applications*, 40(1), 75-82.

Se propone un método orientado a la detección de plantas en campos de maíz donde la tarea se complica debido a que las plantas se encuentran impregnadas por materiales procedentes del suelo. Este fenómeno se suele producir después de que se haya inundado el campo de cultivo para su regado o debido a una lluvia intensa. A las plantas que se encuentran impregnadas total o parcialmente se las denomina plantas enmascaradas. Tras una primera clasificación de los píxeles de la imagen se obtienen dos grupos: los píxeles pertenecientes a las plantas y al fondo. Las plantas enmascaradas han perdido su verdor natural, razón por la cual se dificulta su identificación y no son catalogadas como plantas verdes. Para su detección es necesario seleccionar todos los píxeles clasificados como fondo y realizar una segunda clasificación con dichos píxeles, tras lo cual se identifican satisfactoriamente como plantas enmascaradas.

También se puede emplear el método con una finalidad distinta, evaluar el éxito del tratamiento realizado en el cultivo para la eliminación de malas hierbas. Tras el tratamiento con herbicida, las malas hierbas comienzan un proceso de secado donde pierden parte de su verdor y por lo tanto no se identifican como plantas verdes. Este sistema se describe en su totalidad en el capítulo 4, *Sistemas Expertos Automáticos: detección de líneas de cultivo y discriminación entre plantas de cultivo y malas hierbas en maíz*, sección 4.3 SEA para detección de plantas de cultivo y malas hierbas enmascaradas.

J.M. Guerrero, G. Pajares, **Martín Montalvo**, J. Romeo, M. Guijarro. 2012. Support Vector Machines for crop/weeds identification in maize fields. *Expert Systems with Applications*, 39(12), 11149-11155.

En este trabajo se colaboró mediante la identificación de los píxeles pertenecientes a las plantas verdes. Se diseñó un índice de vegetación para identificar los verdes presentes en la imagen, el cual es una combinación de varios índices de vegetación existentes en la literatura en el que cada uno de ellos contribuye a la decisión final con un peso específico determinado por los correspondientes coeficientes que se estiman convenientemente. Este aporte se encuentra detallado en la etapa 1 de la sección

4.3 y relacionado con la sección 4.2.2. También se realizó un estudio comparativo del método utilizado para clasificar los píxeles como pertenecientes a plantas enmascaradas y sin enmascarar basado en las máquinas de vectores soporte con el método descrito en la sección 4.3, mostrando los resultados en la sección 6.1.2.

J. Romeo, G. Pajares, **Martín Montalvo**, J.M. Guerrero, M. Guijarro, A. Ribeiro. 2012. Crop row detection in maize fields inspired on the human visual perception. *The ScientificWorld Journal*, 2012, 484390-484398.

En este trabajo se colaboró mediante la implementación de un método para la detección de líneas haciendo uso de la transformada de Hough y la comparativa con distintas técnicas, entre ellas la presentada en la sección 4.2, para comprobar cuál producía mejores resultados.

J.M. Guerrero, M. Guijarro, **Martín Montalvo**, J. Romeo, L. Emmi, A. Ribeiro, G. Pajares. 2013. Automatic expert system based on images for accuracy crop row detection in maize fields. *Expert Systems with Applications*, 40(2), 11149-11155.

En este trabajo se colaboró en la umbralización de las imágenes una vez se encontraba segmentada. También se realizó una comparativa del método propuesto en el artículo para la detección de las líneas de cultivo, basado en el método de Theil-Sen, con el detallado en la sección 4.2, mostrando los resultados en la sección 6.1.1.

J. Romeo, G. Pajares, **Martín Montalvo**, J.M. Guerrero, M. Guijarro, J.M. de la Cruz. 2013. A new Expert System for greenness identification in agricultural image. *Expert Systems with Applications*, 40(6), 2275-2286.

En este trabajo se colaboró ideando el estudio de la validez del contraste de las imágenes en comparación con los métodos clásicos.

J. Romeo, J. M. Guerrero, **Martín Montalvo**, L. Emmi, M. Guijarro, P. Gonzalez-de-Santos, G. Pajares. 2013. Camera Sensor Arrangement for Crop/Weed Detection Accuracy in Agronomic Image. *Sensors*, 13(4), 4348-4366.

Se colaboró comprobando y evaluando la precisión en la detección de las líneas de cultivo al variar los parámetros extrínsecos del sistema de visión, así como en el diseño del método para el cálculo del tiempo de exposición. Es una continuación del

trabajo “Camera System geometry for site specific treatment in precision agriculture” y uno de los precursor del trabajo “Acquisition of Agronomic Images with Sufficient Quality by Automatic Exposure Time Control and Histogram Matching”.

Congresos

Martín Montalvo, J.M. Guerrero, J. Romeo, P.J. Herrera, M. Guijarro, A. Ribeiro. 2011. Camera System geometry for site specific treatment in precision agriculture. Proceedings of the First International Workshop on Robotics And Associated High Technologies And Equipment For Agriculture RHEA Workshop (G. Rabatel and P. González de Santos), September 9, Montpellier (France).

Se analizó y estudió la influencia del modelo geométrico del sistema de visión, considerando tanto los parámetros intrínsecos como los extrínsecos, en la detección de las líneas de cultivo en un campo de maíz. Se utilizaron dos modelos de cámaras con tres sistemas ópticos distintos a diferentes alturas con el fin de decidir cuál de todas las configuraciones era la más adecuada. Mediante la geometría del sistema se puede orientar la búsqueda de las líneas de cultivo y de este modo restringir el espacio de búsqueda para optimizar el sistema. En la sección 4.2.4 se hace referencia a cómo se emplea la geometría del sistema para realizar la búsqueda de las líneas de cultivo mediante una unidad de medida inercial (IMU, Inertial Measurement Unit).

Martín Montalvo, J. M. Guerrero, J. Romeo, D. Oliva, M. Guijarro, G. Pajares (2012). Unsupervised learning for crop/weeds discrimination in maize fields with high weed densities. Int. Conf. of Agricultural Engineering (CIGR-AgEng2012), 5 pages, July 8-12, Valencia (Spain), (<http://cigr.ageng2012.org>).

Se propone un método para la diferenciación de las plantas presentes en un cultivo de maíz y su clasificación en dos conjuntos según sean malas hierbas o plantas de maíz. Para la clasificación se empleó un método de aprendizaje basado en LVQ (Learning Vector Quantization).

Martín Montalvo, J.M. Guerrero, J. Romeo, M. Guijarro, J.M. de la Cruz, G. Pajares (2013). Acquisition of Agronomic Images with Sufficient Quality by Automatic Exposure Time Control and Histogram Matching. Advanced Concepts for Intelligent Vision Systems (ACIVS'13) (J. Blanc-Talon, A. Kasinski, W. Philips, D. Popescu, P. Scheunders, Eds.), Lecture Notes in Computer Science, vol. 8192, pp. 37-48, Springer International Publishing Switzerland. October 28-31, Poznan, Poland.

Se propone un método para la adquisición de imágenes en entornos agrícolas de exterior con la suficiente calidad mediante el control del tiempo de exposición y modificación del histograma. En el método se decide si la imagen es inválida, en cuyo caso se debe descartar para capturar otra con un nuevo tiempo de exposición. Una vez que se tiene una imagen capturada con un tiempo de exposición apropiado se estudia y modifica el histograma para mejorar la calidad de la imagen. La sección 4.2.1 hace referencia a la obtención de una imagen mediante el control del tiempo de exposición.

J. M. Guerrero, **Martín Montalvo**, J. Romeo, D. Oliva, M. Guijarro, G. Pajares (2012). Detection of images with high CCD charge level due to indirect sunlight in maize fields. Int. Conf. of Agricultural Engineering (CIGR-AgEng2012), July 8-12, Valencia (Spain), (<http://cigr.ageng2012.org>).

Se colaboró realizando el estudio de los métodos para la identificación de los verdes presentes en la imagen y la umbralización para su posterior binarización.

J. Romeo, D. Oliva, J. M. Guerrero, **Martín Montalvo**, M. Guijarro, G. Pajares 2012. Plants discrimination in maize fields applying human perception to digital images. Proc. Int. Conf. of Agricultural Engineering (CIGR-AgEng2012), July 8-12, Valencia (Spain), (<http://cigr.ageng2012.org>).

Se colaboró realizando el diseño del método para la comprobación de la validez del algoritmo de extracción en diferentes bases de imágenes disponibles.

J. M. Guerrero, J. Romeo, L. Emmi, **Martín Montalvo**, M. Guijarro, G. Pajares, P. González-de-Santos. 2012. Influence of the vision system pitch angle on crop and weeds detection accuracy. First RHEA Int. Conf. on Robotics and associated High technologies and Equipment for Agriculture, September 19-21, Pisa (Italy), (<http://www.rhea-conference.eu/2012/>).

Este trabajo es la ampliación del estudio realizado en “Camera System geometry for site specific treatment in precision agriculture”.

M. Guijarro, J.M. Guerrero, **Martín Montalvo**, J. Romeo, G. Pajares (2013). A new approach to solve image thresholding in precision agriculture. 9th European Conference on Precision Agriculture. Lérida, Spain, July 7-11, poster. <http://www.ecpa2013.udl.cat/>

Se realizó la comparativa de distintos métodos para la umbralización de imágenes agrícolas en campos de maíz segmentadas, atendiendo al criterio de porcentaje de acierto.

1.6. Estructuración de la memoria

La memoria de tesis se presenta estructurada en capítulos. Estos capítulos se organizan siguiendo el orden natural de la investigación, cuya distribución es la que se expone a continuación:

Primer capítulo. *Introducción*, donde se estudian y plantean los objetivos de la tesis, así como la motivación e introducción de las propuestas y planteamientos realizados, tal y como se ha expuesto en las secciones precedentes.

Segundo capítulo. *Estado del arte*, donde se revisan diferentes estrategias existentes en la literatura, que proporcionan los argumentos y la base de la investigación.

Tercer capítulo. *Fundamentos teóricos en la detección de líneas de cultivo y malas hierbas*, donde se describen los conceptos de las técnicas que se utilizan en las distintas fases de las que constan las aportaciones realizadas.

Cuarto capítulo. *Sistemas Expertos Automáticos: detección de líneas de cultivo y discriminación entre plantas de cultivo y malas hierbas en maíz*, donde se exponen las aportaciones mediante dos técnicas que están formadas de varias fases. La primera técnica se basa en la detección de líneas de cultivo y malas hierbas, mientras que la segunda técnica se basa en la detección de plantas cubiertas por barro (enmascaradas) o en proceso de secado por alguna circunstancia.

Quinto capítulo. *Equipamiento y materiales* , donde se describe el material utilizado en las diversas pruebas realizadas a lo largo de la tesis y el diseño del sistema de visión en el prototipo RHEA y los resultados obtenidos.

Sexto capítulo. *Análisis de resultados*, donde los resultados se muestran de forma individual, por una parte aparecen los resultados obtenidos por la primera técnica y a continuación los resultados de la segunda de las técnicas, ambas descritas en el capítulo cuatro. Para terminar se muestran los resultados finales obtenidos en el proyecto RHEA.

Séptimo capítulo. *Conclusiones y trabajo futuro*, donde se exponen las líneas generales seguidas, realizando un balance general del trabajo. Se plantean las líneas de investigación futuras así como posibles mejoras o extensiones de los métodos propuestos.

Capítulo 2

2. Estado del arte

En el capítulo uno se describió el principal problema que se quiere resolver, la detección de plantas de cultivo, incluyendo las líneas asociadas que definen y malas hierbas en campos de maíz con el objetivo de realizar la identificación de zonas de tratamiento selectivo así como para tareas de guiado, mediante el uso de sistemas de visión acoplados en el vehículo autónomo. En el presente capítulo se realiza un análisis bibliográfico sobre el estado del arte actual y últimos avances en los sistemas de visión dentro del campo de la AP para resolver el problema planteado.

Las dos principales líneas de investigación que se analizan en mayor profundidad explorando en la literatura posibles enfoques para su solución son: (a) segmentación de la imagen para detectar los distintos elementos de interés presentes en la escena y (b) detección y localización de las líneas de cultivo sobre el terreno.

2.1.Segmentación de la imagen orientada a la obtención de verdes

Los métodos proporcionados por la visión por computador constituyen en su conjunto una excelente herramienta para la obtención de verdes a partir de imágenes proporcionadas por los sensores instalados en tractores, considerando también los tractores autónomos. Desde hace ya más de una década se han venido utilizando cámaras para la captura de imágenes con el fin de realizar tratamientos agrícolas sobre áreas específicas en grandes extensiones de campos de cultivo (Davies y col., 1998). Centrándonos en campos de maíz, uno de los tratamientos más importantes consiste en eliminar las malas hierbas, donde las plantas (cultivo y malas hierbas) deben ser identificadas como paso previo a su eliminación. Diferentes métodos y estrategias para la identificación de las plantas han sido aplicados en diferentes trabajos (Burgos-Artizzu y col., 2009; Guerrero y col., 2012; Guijarro y col., 2011; Montalvo y col., 2012; Onyango y Marchant, 2003; Tellaeche y col., 2008b; Tellaeche, Burgos-Artizzu y col., 2008a). López-Granados (2011) realiza una revisión de métodos donde la identificación de las plantas resulta ser el paso clave en el proceso. La mayoría de las estrategias

abordan el problema de la identificación del verde bajo la asunción de que las plantas muestran 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.

Ciertamente, el maíz es un cultivo de regadío, el cual está desprotegido de la lluvia. Cuando la capa de agua es abundante o su caída sobre el suelo del campo de cultivo es relativamente fuerte, los materiales del suelo (particularmente barro) impregnan la cobertura vegetal, concretamente aquellas partes cercanas o próximas a la tierra. En este caso, la componente espectral verde de los píxeles pertenecientes a las plantas es enmascarada por la componente espectral dominante, en este caso la componente roja proveniente de los materiales existentes en la tierra; la Figura 6-6 muestra una imagen donde aparece claramente este fenómeno, concretamente en la parte central de la mitad de la imagen y también en los bordes de las hojas del maíz orientadas hacia la tierra. Esto genera que los métodos basados exclusivamente en la identificación del verdor existente en la literatura fallen ante tales situaciones, llegando a identificar las plantas enmascaradas como partes del terreno. Este es uno de los problemas que se abordan en el presente trabajo de investigación.

Independientemente de lo anterior, se han propuesto diferentes estrategias para segmentar plantas en imágenes agrícolas, concretamente orientadas a la obtención de las partes verdes de la escena. A continuación se realiza una revisión de los principales métodos utilizados con tal propósito, agrupados en las categorías que se indican según su naturaleza.

2.1.1. Segmentación basada en índices espectrales

Uno de los intereses de querer observar el planeta tierra es estudiar el papel de la vegetación terrestre en procesos globales a gran escala con el objetivo de entender cómo funciona el planeta como sistema (EOS, 2014). Esto requiere un entendimiento de la distribución a nivel global de la vegetación en la Tierra, así como sus propiedades biofísicas y estructurales junto a sus variaciones espacio-temporales. 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. A partir de las medidas de reflectancia emitida por

los objetos existentes en la escena se construyen los VI. Los sensores que proporcionan la información con la que se calculan los VI pueden encontrarse instalados en distintos lugares o dispositivos, tales como satélites que orbitan alrededor de la tierra, aeronaves que sobrevuelan la zona que se quiera estudiar o en vehículos terrestres que se desplazan por el terreno (Thorp y Tian, 2004). En la sección 3.3 se detalla de manera pormenorizada el concepto de VI y las distintas bandas espectrales utilizadas más frecuentemente para su obtención. Podemos hacer una diferenciación dentro de los VI, por un lado están los que emplean el espectro infrarrojo cercano (NIR) y por otro los que hacen uso únicamente del espectro visible. Se realiza esta diferenciación debido a que la mayoría de las cámaras comerciales y de uso industrial utilizan el modelo de color RGB para la representación de las imágenes, pudiendo descomponer el espectro visible en los tres canales (rojo, verde y azul) y sin embargo es menos 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 multiespectrales e hiperespectrales que utilizan el NIR con tal propósito (Jurado-Expósito y col., 2003; Peña y col., 2013; Castro y col., 2012). Incluso se han propuesto modificaciones físicas en cámaras digitales comerciales para suprimir el filtro de corte de infra-rojos, con el que están equipadas, sustituyéndolo por otro que deja pasar una amplia gama de radiaciones a partir del rojo (Rabatel y col., 2014).

Los VI han sido utilizados desde hace bastante tiempo para la detección de alguna propiedad de la vegetación, siendo uno de los primeros el índice de vegetación ratio (RVI, *Ratio Vegetation Index*) o ratio simple (SR, *Simple Ratio*), usado por Jordan (1969) para obtener un índice que calculara el área foliar de un bosque basándose en el hecho de que las hojas absorben mejor la luz roja del espectro visible que la infrarroja. El RVI es el cociente entre el NIR y el rojo visible. Un problema asociado a este índice es su falta de sensibilidad cuando la vegetación está dispersa y su densidad foliar es escasa. También presenta problemas cuando la zona tiene una densidad de vegetación elevada ya que la cantidad de luz roja se aproxima a cero y consecuentemente el resultado del cociente para su cálculo tiende hacia infinito. Estos problemas pueden ser solucionados mediante el índice vegetativo de diferencia normalizada (NDVI, *Normalized Differential Vegetation Index*) atribuido a Rouse y col. (1973) basándose en los estudios previos de Kriegler y col. (1969). El índice NDVI utiliza también la componente espectral roja y el NIR. El NDVI ha sido utilizado en una amplia gama de

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 fuertemente relacionado con el índice de área foliar y la fracción ocupada por la cobertura vegetal (Carlson y Ripley, 1997). Existen otros muchos índices, algunos derivados del NDVI, que utilizan el NIR para detectar la existencia de la vegetación o alguna de sus propiedades como son los que se enumeran a continuación: DVI (*Difference Vegetation Index*) (Tucker, 1979), GDVI (Green Difference Vegetation Index) (Sripada y col., 2006), GNDVI (Green Normalized Difference Vegetation Index) (Buschmann y Nagel, 1993), PVI (*Perpendicular Vegetation Index*) (Richardson y Wiegand, 1977), WDVI (*Weighted Difference Vegetation Index*) (Clevers, 1988), NG (Normalized Green), NR (Normalized Red), NNIR (Normalized Near Infrared), GRVI (Green Ratio Vegetation Index) (Sripada y col., 2006), PVI (*Percentage Vegetation Index*) (Crippen, 1990), GRNDVI (Green-Red NDVI), GBNDVI (Green-Blue NDVI) (Wang y col., 2007), CASI NDVI (Cloutis y col., 1996). El NIR además de detectar la vegetación también se puede utilizar para clasificar la vegetación en distintos tipos de plantas o incluso según su estado de salud. En un estado bueno de salud la vegetación absorbe la mayor parte de la luz visible que le alcanza y refleja la mayor parte de la luz infrarroja, mientras que la vegetación enferma refleja más luz visible y menos infrarroja que la vegetación saludable.

Los VI en los que se utiliza alguna banda del espectro infrarrojo, mencionados previamente, suelen obtenerse a partir de datos capturados por satélites o aeronaves, incluyendo los vehículos aéreos no tripulados o “drones”. A este modo de obtener imágenes desde plataformas remotas, la cuales suelen ser de naturaleza multispectral o hiperspectral, se le denomina teledetección. La teledetección ha sido ampliamente estudiada y utilizada desde hace tiempo como herramienta para la detección y localización de malas hierbas en cultivos agrícolas (Lamb y Brown, 2001; Moran y col., 1997; Zwiggelaar 1998; Wilfong y col., 2011; Shaw, 2005). El procedimiento de teledetección utilizado para detectar y analizar la vegetación suele ser de naturaleza pasiva, en la que el sensor, operando en un determinado rango del espectro, recibe la radiación emitida por la vegetación sin intervención sobre la escena; mientras que en la activa el sensor emite ondas, recibiendo de nuevo los reflejos de éstas producidos por la vegetación o los objetos existentes en la escena debido a que no se absorbe por completo la onda.

En contraposición a la detección se sitúan los sistemas que captan imágenes en las proximidades del terreno, las imágenes obtenidas con este tipo de sistemas suelen poseer una mayor resolución que las obtenidas con los sistemas a bordo de aeronaves o satélites, lo que permite detectar la aparición de malas hierbas que se encuentran en pequeñas concentraciones. En este sentido cabe mencionar que usando vehículos aéreos para capturar imágenes de la vegetación se puede cubrir extensas zonas de cultivo en poco tiempo, a cambio de una menor resolución espacial. Lamb y col. (1999) obtuvieron imágenes multiespectrales de avena desde un vehículo aéreo con una resolución que variaba entre los 2 metros y los 50 centímetros. En Herrmann y col. (2013) se emplea una cámara hiperespectral con 849 bandas cubriendo una amplia gama de bandas espectrales, entre las que se encuentran el infrarrojo y el visible. Situando la cámara sobre un trípode, a una altura de 1,35 metros respecto del suelo se consigue una resolución espacial de 0,5 mm. Torres-Sánchez y col. (2013a,b, 2014) y Peña y col. (2012, 2013) utilizaron un cuatri-rotor equipado con un sensor multiespectral con seis canales individuales, así como una cámara comercial operando en el espectro visible para la detección de cultivos y malas hierbas para generar mapas de vegetación aplicando varios índices de vegetación.

Debido a que las imágenes aéreas tiene una menor resolución que las tomadas in situ, que utilizar un satélite aumenta considerablemente el coste de adquisición de imágenes, y que el uso del satélite está condicionado a la climatología, el cielo debe de estar muy despejado para obtener una imágenes de calidad, en ocasiones los esfuerzos se centran en el uso de vehículos aéreos no tripulados o vehículos terrestres, en ambos casos equipados con sensores de captura de imágenes, para detectar e identificar el cultivo y las malas hierbas. La mayor parte de los VI mencionados previamente y que hacen uso del NIR se obtienen a partir de imágenes obtenidas desde aeronaves. Sin embargo, cuando se utilizan vehículos terrestres se pueden conseguir altas resoluciones de imágenes con elevada precisión, llegando incluso a tratamientos a nivel de planta en vez de las grandes extensiones cubiertas por las imágenes de satélite y en menor medida por las obtenidas mediante diferentes aeronaves. En el caso terrestre es menos habitual el uso del infrarrojo para el cálculo de los VI, siendo más frecuente la utilización de canales espectrales cercanos o incluidos en el espectro visible. Como se menciona en Meyer y Camargo-Neto (2008) una de las ventajas de usar VI que hacen uso únicamente de la banda que ocupa el espectro de color es que acentúan un color particular,

centrándose en el que resulte de interés en cada caso, por ejemplo en el caso del maíz para la detección de líneas de cultivo y malas hierbas, dicho canal es el verde.

Actualmente la mayoría de las cámaras digitales, operando en el espectro visible, utilizan el modelo de color RGB (Pajares y Cruz, 2007; Gonzalez y Woods, 2007), donde se destina un canal al espectro que ocupa el color verde. La obtención del modelo de color conocido como RGB se suele llevar a cabo mediante un filtro Bayer (denominado así en honor a su creador), en el que no se tiene un sensor que cuantifique la radiación de cada uno de los tres canales (R, G y B), sino que en cada píxel sólo se cuantifica la cantidad de radiación perteneciente a la frecuencia de uno de los canales, obteniendo el valor de los otros dos canales mediante interpolación de los píxeles circundantes. El 50% de los sensores identifican la radiación perteneciente al color verde, el 25% al color rojo y el 25% restante al color azul. El hecho de que el número de píxeles verdes sea el doble que el de rojos y azules es una consecuencia de la evolución natural del hombre, asumiendo que en los orígenes éste se debía desenvolver en un entorno natural predominantemente verde. Esta circunstancia y el uso de este tipo de filtros favorecen la detección de las plantas verdes que constituyen el objeto de interés.

Mediante una adecuada resolución espectral y espacial de las imágenes se puede llegar a detectar las plantas mediante el análisis únicamente de los verdes presentes en la imagen sin considerar los canales rojo y azul del espectro visible, éste resulta ser uno de los modos más sencillos para detectar la vegetación. Yang y col. (2003) siguen el enfoque anterior, se fijan exclusivamente en el canal verde de los píxeles y si su valor excede un valor predeterminado, el píxel es considerado como vegetación. Si bien, lo normal es combinar más de un canal de los tres disponibles en el modelo de color RGB. Woebbecke y col. (1995a) realizan una comparativa de cuatro índices, a saber: $R-G$, $G-B$, $(G-B)/(R-G)$ y $2G-R-B$. En dicho trabajo concluyeron que el índice con mejor comportamiento fue el último de ellos, ampliamente conocido en la literatura como exceso de verde ($ExG = 2G - R - B$). A raíz de este estudio otros muchos autores han utilizado este índice bien como índice base o para su comparación con otros índices (Meyer y col., 1998; Ribeiro y col., 2005; Meyer y Neto, 2008; Montalvo y col., 2012; Ponti, 2013).

Intuitivamente el mayor interés en el uso de los índices de vegetación se centra en detectar los colores verdes, ya que las plantas poseen este color, sin embargo también

puede ser interesante detectar las partes rojas o rojez presente en la imagen, que principalmente representan el suelo del terreno, para descartarlas y facilitar la detección de los verdes. Con tal propósito Meyer y col. (1998) propusieron un VI para detectar las partes rojas de la imagen. Este índice, llamado exceso de rojo, hace uso de los canales rojo y verde ($ExR = 1.4R - G$). En Neto (2004) se combinan el ExG y el ExR para crear un nuevo índice, en este caso conocido como el exceso de verde menos el exceso de rojo ($ExGR = ExG - ExR$). Este índice presenta la ventaja de no necesitar un proceso con el que determinar un umbral a partir del cual binarizar la imagen, a diferencia de lo que sucede con el ExG que sí lo necesita, se utiliza como umbral el valor cero. Pérez y col. (2004) crearon el índice de diferencias normalizadas (NDI) que utiliza un cociente para distinguir las plantas del resto, $NDI = (G - R) / (G + R)$. Una comparativa entre los índices ExG , NDI y $ExGR$ puede encontrarse en Meyer y Camargo-Neto (2008). En Aitkenhead y col. (2008) se emplea también un cociente, pero en este caso se trata de la proporción de verde existente frente a los tres canales, $G / (G + R + B)$. 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: $(G - R) / (G + R + B)$. Un nuevo índice que aprovecha la potencia de la combinación de canales fue descrita en Hague y col. (2006) basándose en Marchant y Onyango (2002). Este índice fue concebido para que fuese invariante a las condiciones lumínicas, además determinan la existencia de un valor para un parámetro a , que incluye, para el cual el índice es invariable a la temperatura de color que emitiría un cuerpo negro. El VI de Hague y col. (2006) es como sigue: $VEG = G / (R^a B^{1-a})$ con $a = 0.667$ fijado en sus experimentos con las propiedades descritas previamente.

Existen también índices que emplean una combinación lineal de los tres canales RGB para identificar las plantas verdes, estos índices poseen la siguiente expresión: $xR + yG + zB$, donde x , y y z son constantes que establecen la contribución o peso específico de cada canal. En Ribeiro y col. (2005) a las constantes se les asigna los valores -0.7, 0.588 y 0.136 para los parámetros x , y y z respectivamente mediante el uso de algoritmos genéticos. Kataoka y col. (2003) realizan también una combinación lineal

de los tres canales, añadiendo una constante a la suma global, resultando finalmente la siguiente expresión: $CIVE = 0.444R - 0.811G + 0.385B + 18.78745$.

Los índices *ExG*, *ExGR*, *CIVE* y *VEG* han sido combinados en Guijarro y col. (2011) con la siguiente contribución a la combinación 0.25, 0.30 0.33 y 0.12 respectivamente para las imágenes objeto de estudio.

2.1.2. Segmentación basada en umbralizaciones

Generalmente estas técnicas asumen que el problema a resolver es de naturaleza biclase donde las plantas y el terreno constituyen los dos elementos esenciales a ser identificados. Reid y Searcy (1987) crearon una función de decisión basándose en la asunción de que las dos clases a identificar en la escena siguen una función de distribución Gaussiana. El método de Otsu (1979) también es utilizado considerando la naturaleza del problema biclase (Ling y Ruzhitsky, 1996; Shrestha, Steward y Birrel, 1994). Estos algoritmos son aplicados a imágenes en escala de grises, obteniendo un umbral a partir del cual se binariza la imagen en dos clases. Gebhardt y col. (2006) también realizaron una umbralización para segmentar una imagen de color, para ello transformaron una imagen del modelo de color RGB a una imagen en escala de grises. El algoritmo que utilizaron fue posteriormente mejorado mediante operaciones morfológicas y aplicando propiedades de homogeneidad local en un trabajo posterior (Gebhardt y Kaihbauch, 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 asumiendo dos funciones de densidad de probabilidad Gaussianas asociadas cada una de ellas a la vegetación y al terreno. 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 un umbral. Guijarro y col. (2011) y Burgos-Artizzu y col. (2011) han utilizado el valor medio estadístico de los niveles de intensidad de los píxeles de la imagen transformada obtenida con los VI en vez de utilizar un método de umbralización automática, tal como Otsu o similares. En los trabajos se justificó la elección de la media como umbral porque el método de Otsu proporciona un valor umbral mayor que

la media, produciéndose una infra segmentación, lo que genera que algunas plantas no sean identificadas convenientemente.

2.1.3. Segmentación basada en aprendizaje

Meyer y col. (2004) han empleado técnicas de aprendizaje no supervisado, incluyendo el denominado agrupamiento borroso o *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*. 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 de segmentación EASA utilizando como imagen de partida o referencia una transformada al modelo de espacio de color HSI (Hue-Saturation-Intensity) para enfrentarse al problema de la variabilidad en la iluminación. Zeng y col. (2009, 2010) usan 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 con un problema de segmentación biclase; la separación en clases fue validada mediante una red neuronal y el discriminante lineal de Fisher, habiendo empleado como espacios de color los siguientes: RGB, LUV y HSI. En Guerrero y col. (2012) y Tellaeche y col. (2011) se utilizan las máquinas de vectores soporte para clasificar la imagen en dos conjuntos distintos, el que contiene las plantas y el que contiene el resto de elementos de la escena, particularmente suelo.

2.2. Detección y localización de líneas en imágenes agrícolas

Se han propuesto diferentes estrategias para la detección de las líneas de cultivo. Fontaine y Crowe (2006) comprobaron la eficiencia de cuatro algoritmos para la detección de líneas y su localización proporcionando la posición y el ángulo de inclinación con respecto a los ejes de referencia en la imagen. Los algoritmos se evaluaron tanto en conjuntos de imágenes con líneas de cultivo simuladas como reales, con y sin malas hierbas. Los conceptos desarrollados y los procedimientos utilizados en estos algoritmos fueron: *a)* exploración de bandas geométricas horizontales; *b)* análisis de regiones continuas o gotas, *c)* regresión lineal y *d)* transformada de Hough.

Las estrategias anteriores no son exclusivas para resolver la problemática relativa a la detección y localización de las líneas de cultivo. Una revisión sobre diferentes técnicas existentes en la literatura nos conduce a la agrupación de estos métodos en las ocho categorías que se incluyen a continuación.

2.2.1. Métodos basados en la exploración horizontal de bandas

Sogaard y Olsen (2003) utilizaron una transformación desde el modelo de color RGB a una imagen en escala de grises. El proceso se realiza dividiendo primeramente la imagen de color en sus canales rojo, azul y verde y aplicando posteriormente los bien probados VI para la segmentación de las plantas verdes descritos en Woebbecke y col. (1995). Después de esto, la imagen en escala de grises es dividida en bandas horizontales donde los valores máximos indican la presencia de una línea candidata a ser una línea de cultivo. Cada máximo determina una línea de cultivo segmentada y el centro de gravedad del segmento se marca en esta posición de la banda. Las líneas de cultivo se identifican uniendo los puntos marcados a través de métodos similares a los utilizados en la transformada de Hough o mediante regresión lineal. Sainz-Costa y col. (2011) han desarrollado una estrategia basada en el análisis de secuencias de videos para identificar las líneas de cultivo. Las líneas de cultivo persisten a lo largo de las direcciones definidas por la proyección de perspectiva con respecto a la escena 3D en el campo. Explotando este hecho, se aplica la transformación a escala de gris basada en el método propuesto por Ribeiro y col. (2005) y se consigue una imagen binarizada aplicando una técnica de umbralización. Cada imagen se divide en cuatro bandas horizontales, sobre las que se trazan regiones rectangulares sobre la imagen ya binarizada para identificar zonas de cultivo. Los centros de gravedad de estas regiones se utilizan como los puntos que definen las líneas de cultivo, ajustándose una línea definida por dichos puntos. La primera imagen en la secuencia se usa como referencia para guiar el proceso global posterior. Las posiciones de los rectángulos en esta primera imagen sirven de base para determinar las posiciones de las nuevas regiones en las subsiguientes imágenes de la secuencia de video. Hague y col. (2006) transforman una imagen originalmente en el formato RGB a escala de grises. La imagen transformada se divide en 8 bandas geométricas horizontales. La intensidad de los píxeles a través de estas bandas muestra una variación periódica, debido a que las líneas de cultivo son paralelas. Ya que las características de la cámara, su localización y la distancia entre las líneas de cultivo son conocidas a priori, la distancia entre las líneas de cultivo, en

píxeles de la imagen, puede calcularse para cada una de las bandas horizontales usando el modelo ideal de la cámara conocido como pinhole, que determina la geometría de la cámara y su óptica (Tsai, 1987). Como mejora al proceso anterior, se propone la utilización de un filtro pasa banda con respuesta en el dominio de la frecuencia. En imágenes con una alta densidad de malas hierbas, éstas y el cultivo aparecen unidos formando una única región espacial, donde los centros de gravedad y otras características de las bandas geométricas pueden identificarse erróneamente.

2.2.2. Métodos basados en la transformada de Hough

De acuerdo a Slaughter y col. (2008) uno de los métodos de visión artificial más comúnmente usados para la identificación de líneas de cultivo está basado en la transformada de Hough (1962). Este método fue diseñado para tratar discontinuidades en las líneas, donde el cultivo tiene algunos saltos en las líneas de cultivo debido a una pobre germinación o a otros factores que resultan en la pérdida de plantas en las líneas de cultivo. La transformada de Hough ha servido como procedimiento para el guiado automático en tiempo real de los vehículos agrícolas en algunos desarrollos (Astrand y Baerveldt, 2005; Hague y col., 1997; Leemans y Destain, 2006; Marchant, 1996). Estos métodos se aplican sobre imágenes binarias, las cuales se han obtenido mediante la aplicación de alguna de las técnicas explicadas anteriormente o similares, concretamente por transformaciones desde el espacio de color RGB a una imagen en escala de grises y su posterior binarización (Tellaache y col., 2011; Tellaache y col., 2008a, 2008b). Gee y col. (2008) aplican la transformada de Hough dos veces basándose en la asunción de que las líneas de cultivo son las únicas líneas en la imagen que convergen en el punto de fuga, el resto de estructuras en las imágenes que no siguen esta filosofía se descartan; adicionalmente la aplicación de algunas restricciones tales como el espaciado entre líneas y conceptos como la perspectiva de geometría ayudan a identificar las líneas. Se requiere establecer un umbral para poder aplicar la transformada de Hough y así establecer los valores máximos pico (Jones y col., 2009a, 2009b) o los picos predominantes (Rovira-Más y col., 2005). Dependiendo de la densidad del cultivo es posible que más de una línea detectada se solape con una misma línea de cultivo, dificultad ésta que se soluciona mediante un proceso posterior que fusiona líneas cuyos parámetros son similares (Tellaache y col., 2008a, 2008b, 2011). Ji y Qi (2011) informan sobre el hecho de que la transformada Hough es computacionalmente lenta debido a la enorme cantidad de datos que se deben procesar,

proponiendo una variante de la transformada de Hough de naturaleza aleatoria para reducir el tiempo de cómputo. La propuesta consiste en evitar cálculos redundantes, se opera iterativamente de forma aleatoria muestreando un conjunto de puntos para calcular una única localización en el espacio de Hough. Ya que dos puntos son trivialmente colineales, los parámetros de la línea en la cual caen los puntos pueden estimarse. Estos parámetros se usan para incrementar el valor de la celda del acumulador en el espacio de Hough. La idea consiste en seleccionar pares de puntos evidentes que pertenezcan a una línea de cultivo para reducir el coste. Esto llega a ser inviable en imágenes con una alta densidad de malas hierbas porque a menudo las líneas de cultivo son indistinguibles de aquellas y los pares de puntos no pueden ser identificados como pertenecientes a una línea de cultivo.

Se han propuesto algunas modificaciones para mejorar la efectividad de la transformada de Hough como la propuesta por Asif y col. (2010), la cual se aplica solamente a los puntos que están situados en los puntos de borde a lo largo del cultivo. No obstante, esto requiere la aplicación de técnicas para la extracción de bordes que incrementarían el coste computacional.

2.2.3. Métodos basados en el punto de fuga

Pla y col. (1997) proponen una estrategia en la que se identifican regiones (cultivo/malas hierbas y tierra) mediante la aplicación de una segmentación de la imagen basada en el color. En su trabajo utilizan el esqueleto de cada región definida como una característica para encontrar las líneas que constituyen el cultivo. Los esqueletos resultantes de cada región pueden usarse como curvas que definen la estructura subyacente del cultivo y por lo tanto para extraer las líneas rectas que representan las filas de plantas y la tierra de los surcos. Los segmentos en los esqueletos son definidos como cadenas de puntos de contorno conectados y de una longitud definida. Esto permite seleccionar líneas que confluyen en el punto de fuga como candidatas a representar las líneas de cultivo. El punto de fuga se determina utilizando información previa sobre la localización del mismo en las imágenes anteriores ya procesadas, realizando un seguimiento del punto de fuga. Este proceso es altamente dependiente de los esqueletos que se extraen, los cuales no son siempre fáciles de segmentar, particularmente en imágenes con una alta densidad de malas hierbas.

2.2.4. Métodos basados en regresión lineal

Algunas de las técnicas anteriores utilizan este enfoque. Billingsley y Schoenfisch (1997) plantearon un sistema para la detección de las líneas de cultivo que es relativamente insensible al ruido visual proveniente de la existencia de las malas hierbas. Usaron regresión lineal en cada uno de los tres segmentos en los que caían las filas de cultivo y una función de coste similar al momento de la línea de mejor ajuste para detectar líneas ajustadas a valores atípicos, es decir ruido y malas hierbas. El objetivo final del proceso era la obtención de información para realizar el guiado. Como se ha mencionado con anterioridad, *Søgaard y Olsen* (2003) también aplicaron regresión lineal, un enfoque que es factible cuando los píxeles pertenecientes a las líneas de cultivo se encuentran bien separados de los pertenecientes a las malas hierbas. Ya que la efectividad de la regresión lineal se encuentra altamente afectada por los píxeles que pertenecen a las malas hierbas, aquella se considera que solamente es recomendable si los píxeles pertenecientes a las malas hierbas y al cultivo han sido previamente separados. Como se verá más adelante en el Capítulo 0, en el enfoque seguido en la aportación descrita en la sección 4.2, hemos realizado una separación de las malas hierbas y cultivo gracias al método de umbralización de Otsu, logrando que la aplicación de la regresión lineal sea una estrategia viable para resolver el problema propuesto en la investigación desarrollada y reflejada en el presente trabajo.

2.2.5. Métodos basados en estereovisión

Kise y col. (2005) o Kise y Zhang (2008) desarrollaron un sistema de navegación agrícola mediante el seguimiento de líneas de cultivo basado en estereovisión. El procesado estéreo se usa para determinar la localización 3D de los puntos de la escena para los objetos de interés a partir de una imagen estereoscópica. Las posiciones 3D, determinadas por medio del cálculo de disparidades de las imágenes, proporcionan la información necesaria para crear un mapa de elevaciones el cual utiliza una matriz bidimensional con valores variables para indicar la altura del cultivo. Esta estrategia requiere que el cultivo posea una altura significativa con respecto al terreno. Debido a que durante las épocas de tratamiento de las malas hierbas en los campos de maíz ni éstas ni los cultivos tienen una altura relevante, este enfoque se considera inefectivo para nuestra aplicación. Rovira-Más y col. (2008) han aplicado y ampliado técnicas de estereovisión a otras áreas dentro de la AP. Al igual que antes,

éstas técnicas sólo son viables si el cultivo y las malas hierbas presentan una altura significativa en la escena 3D y si las diferencias de alturas entre el cultivo y las malas hierbas son apreciables entre sí.

2.2.6. Métodos basados en el análisis de gotas

Estos métodos localizan y caracterizan regiones de píxeles contiguos con los mismos valores en imágenes binarias (Fontaine y Crowe, 2006), las agrupaciones de estos píxeles conectados en regiones es lo que identifican con el concepto de gota. El algoritmo busca gotas o regiones blancas (espacio entre filas) de más de 200 píxeles, se establece un mínimo de píxeles porque un número menor podría representar ruido proveniente de las líneas de cultivo. Una vez que las regiones han sido identificadas, el algoritmo determina el ángulo de sus ejes principales cuando dichas regiones se circunscriben a una elipse, así como la localización de sus centros de gravedad. Para una franja blanca perfectamente recta, el centro de gravedad de la región se sitúa sobre la línea central de la franja blanca y como ángulo representativo se considera el ángulo de los espacios entre filas. El algoritmo devuelve el ángulo y el centro de gravedad de la región más cercana al centro de la imagen. La identificación de las regiones en imágenes infestadas con una alta densidad de malas hierbas se vuelve inviable.

2.2.7. Métodos basados en la acumulación de plantas verdes

Olsen (1995) propuso un método basado en la consideración de que a lo largo de las filas de cultivo presentes en la imagen aparece una importante acumulación de plantas verdes. La imagen está representada con una escala de grises donde las plantas verdes aparecen más claras que el resto de elementos. Se obtiene una acumulación de valores mediante la suma de los niveles de grises de las columnas pertenecientes a una región rectangular dada, lo cual se realiza explorando todas las columnas en el rectángulo. Se asume que las líneas verticales siguen esta dirección en la imagen. En esta aplicación concreta las imágenes están libres de proyección de perspectiva debido a que fueron obtenidas con la cámara en una posición ortogonal. Posteriormente se ajusta una curva sinusoidal mediante mínimos cuadrados a la curva previamente obtenida. Los máximos locales de la senoide proporcionan la localización de los centros de las líneas buscadas. De nuevo la existencia de elevadas densidades de malas hierbas provoca que este enfoque no sea factible en tales circunstancias.

2.2.8. Métodos basados en el análisis de la frecuencia

Debido a que las líneas de cultivo son paralelas en la escena 3D, al proyectarlas en la imagen, mediante proyección de perspectiva, se evidencia un claro comportamiento en el dominio de la frecuencia. Vioix y col. (2002) explotan esta característica y aplican un filtro de Gabor bidimensional, el cual se define como una modulación de una función gaussiana por una señal sinusoidal. El parámetro de la frecuencia requerido por el filtro de Gabor se deduce empíricamente de la transformada rápida de Fourier 2D (Bossu y col., 2006). Bossu y col. (2009) aplican la transformada de wavelets para discriminar líneas de cultivo basándose en el análisis de la frecuencia. Los autores explotan el hecho de que las líneas de cultivo están bien descritas en el dominio de la frecuencia. De esta forma seleccionan una función wavelets madre con una frecuencia adecuada a las líneas de cultivo que pretenden identificar. En aplicaciones reales el cultivo no muestra una clara frecuencia contenida en el espacio de Fourier, particularmente en imágenes con una alta densidad de malas hierbas, por lo tanto la aplicación de filtros en el dominio de la frecuencia resulta ser una tarea complicada para la detección de las líneas de cultivo por este procedimiento.

Capítulo 3

3. Fundamentos teóricos en la detección de líneas de cultivo y malas hierbas

3.1.Introducción

El presente capítulo describe las técnicas y fundamentos teóricos que se han utilizado en las distintas fases del proceso de investigación general con el objetivo de identificar y localizar las malas hierbas y las líneas de cultivo, desde la toma de imágenes hasta la detección de las líneas de cultivo con el objetivo de proporcionar un porcentaje de infestación de malas hierbas en el sembrado, con fines de actuación específica.

Las distintas fases en las que se desglosa el proceso son cinco:

- a) *Obtención de una imagen de calidad.* Mediante el control del tiempo de exposición del sensor a la entrada de iluminación y la radiación procedente de los elementos en la escena.
- b) *Segmentación de la imagen.* Consistente en la descripción de los diferentes VI y técnicas orientadas a realzar las plantas (cultivo y malas hierbas) con respecto al resto de elementos presentes en la imagen.
- c) *Umbralización.* Relativa al tratamiento de la imagen segmentada previamente, mediante métodos que obtienen un umbral de forma automática, que permita separar las plantas del resto de la imagen, obteniendo una imagen binaria en la que las plantas se identifican con una etiqueta lógica y el resto con otra etiqueta diferente.
- d) *Identificación de las líneas de cultivo.* Tras la binarización es necesario identificar distribuciones espaciales específicas de píxeles, más concretamente, alineaciones de píxeles asociados con la ubicación de las plantas pertenecientes al cultivo, que determinan la localización de las líneas de cultivo.

- e) *Cobertura de malas hierbas*. El objetivo final del proceso consiste en la ubicación de las malas hierbas y la obtención de la densidad de las mismas en la zona de actuación específica a partir de la imagen dada.

La obtención de una imagen de calidad mediante el control del tiempo de exposición se aborda en la sección 3.2; la segmentación de la imagen en la sección 3.3; la umbralización de la imagen se describe en la sección 3.4; la extracción de las líneas de cultivo se trata en la sección 3.5. Finalmente, el cálculo de la cobertura de malas hierbas se deriva al capítulo cuatro, como la parte previa al proceso de decisión relativo a la aplicación del tratamiento específico. En el capítulo cuatro se detalla cómo se adaptan para nuestro caso específico los métodos seleccionados en cada una de las fases.

3.2. Tiempo de exposición

En fotografía el término exposición se refiere a la cantidad de luz que incide en el elemento encargado de capturar la escena que se encuentra en el campo de visión de la cámara. En las cámaras fotográficas antiguas, este elemento era la película fotográfica que se impresionaba con la llegada de la luz. Aunque la película fotográfica todavía se utiliza en algunas aplicaciones, actualmente se encuentra en desuso dando paso a nuevos sensores instalados en cámaras digitales, que son las utilizadas en la presente investigación. Dentro de estos sensores se encuentra el conocido CCD (dispositivo de carga acoplada). El término tiempo de exposición se refiere al tiempo que la luz y radiación están incidiendo en el sensor CCD. A veces se utiliza de forma errónea el término velocidad de obturación como sinónimo de tiempo de exposición. Una forma de comprobar el mal uso de ambos conceptos es que el tiempo de exposición se mide en unidades de tiempo mientras la velocidad de obturación se mide en unidades de distancia dividido por unidades de tiempo, tal y como corresponde a la magnitud de una velocidad.

Otro término relacionado con la cantidad de luz incidente en un sensor fotográfico es el que tiene que ver con la apertura o abertura del diafragma de la lente. El diafragma del sistema óptico, al que pertenece la lente, posee un orificio de tamaño ajustable que permite la entrada de una mayor o menor cantidad de luz por unidad de tiempo, lo que se consigue aumentando o disminuyendo, respectivamente, el tamaño del

orificio. El mecanismo para permitir una mayor entrada de luz a través de la lente es muy parecido al mecanismo que utiliza el ojo humano. El diafragma sería el equivalente al iris humano y la abertura a la pupila del ojo. Una forma de cuantificar la apertura de una lente es mediante el *número-f* (N), que es una medida adimensional que resulta de dividir la focal de la lente entre el diámetro de orificio por el que viaja la luz hacia el sensor CCD, ecuación (3.1),

$$N = \frac{f}{D} \quad (3.1)$$

donde f es la distancia focal de la lente, D es el diámetro de la apertura del diafragma. Otra forma de expresar un *número-f* es mediante la relación f/N .

Existen dos tipos de mecanismos para controlar la apertura, los que disponen de un control automático y los controlados manualmente. En los automáticos el grado de apertura se determina mediante el análisis del histograma de la imagen de entrada, de forma que si los valores de intensidad de la imagen, en función de valores medios de intensidad, son altos, la apertura del diafragma se disminuye, aumentándose en caso contrario. Por otra parte, en el caso de los manuales aún se pueden distinguir dos tipos, los que permiten una conexión electrónica para realizar el ajuste a cierta distancia y los que el ajuste se realiza de forma directa por el usuario mediante la manipulación de los anillos que controlan la apertura. Las dificultades intrínsecas de esta manipulación directa provienen de que el propio operario puede interferir en la entrada de iluminación al sistema por su proximidad al mismo, con resultados no deseados y también porque el sistema de visión puede encontrarse en movimiento, resultando muy complicada la actuación directa. Esta situación no se da en los sistemas electrónicos puesto que el operario lo manipula a distancia.

En el proyecto RHEA se han utilizado los dos tipos de sistemas, automático y manual. Así los tractores específicamente destinados a los campos de cereal y olivos están equipados con un sistema de visión con diafragma de apertura automática, controlado desde la propia cámara. En estos vehículos la misión principal del sistema de visión es la detección de obstáculos durante la navegación, cuyo sistema óptico está equipado con una lente de focal variable en el rango 3.5 a 8 mm ya que el objetivo principal es conseguir un campo de vista suficiente de cara a la detección de los obstáculos. Por el contrario, el vehículo dedicado al campo de maíz está equipado con

un sistema óptico de diafragma con manipulación directa de tipo Schneider Cinegon (Schneider, 2014) y longitud focal fija de 10 mm. En su momento se optó por este diseño debido a que la cámara SVS4050CFLGEA (SVS-VISTEK, 2014) permite el control automático del tiempo de exposición por software, de suerte que mediante un análisis de histograma basado en técnicas desarrolladas en el marco de esta investigación se puede controlar aquél sin necesidad de depender de las técnicas desarrolladas por los fabricantes, que a veces no resultan totalmente apropiadas. La Figura 3-1(a) muestra el sistema de visión con sus componentes integrados, cuyo desglose aparece en la Figura 3-1(b), aquí se observa el sensor CCD de tipo Kodak KAI 04050M/C con filtro Bayer para color y rejilla GR, sobre él incide la luz y las radiaciones externas. Aparece también el sistema óptico de diafragma manual mencionado previamente. Las roscas de adaptación son de tipo F con conversores de tipo C para acoplamiento entre el sistema óptico y la cámara. En dicha figura se muestra un filtro de corte de la radiación ultravioleta e infrarroja para evitar su entrada al sensor, ya que afectan grandemente a la calidad del color de la imagen, particularmente la infrarroja por su mayor incidencia en las tonalidades rojas y verdes existentes en los campos de cultivo de maíz. Este filtro es de tipo Schneider UV/IR 486 (Schneider Kreuznach, 2014).



Figura 3-1. Sistema de visión con cámara, óptica y filtro de corte: (a) integrado; (b) con sus elementos separados.

En la Figura 3-2(a) se muestra el sistema óptico desde su parte frontal, con un diámetro pequeño de apertura y en la Figura 3-2(b) se muestra el mismo sistema con un diámetro mayor de apertura. La Figura 3-2(c) muestra una vista cenital del mismo sistema donde se puede apreciar la palanca utilizada para regular manualmente la

apertura cuya escala de graduación está referida al *número-f* variando desde el 1.9 a 16 que se corresponden a mínima y máxima apertura respectivamente.

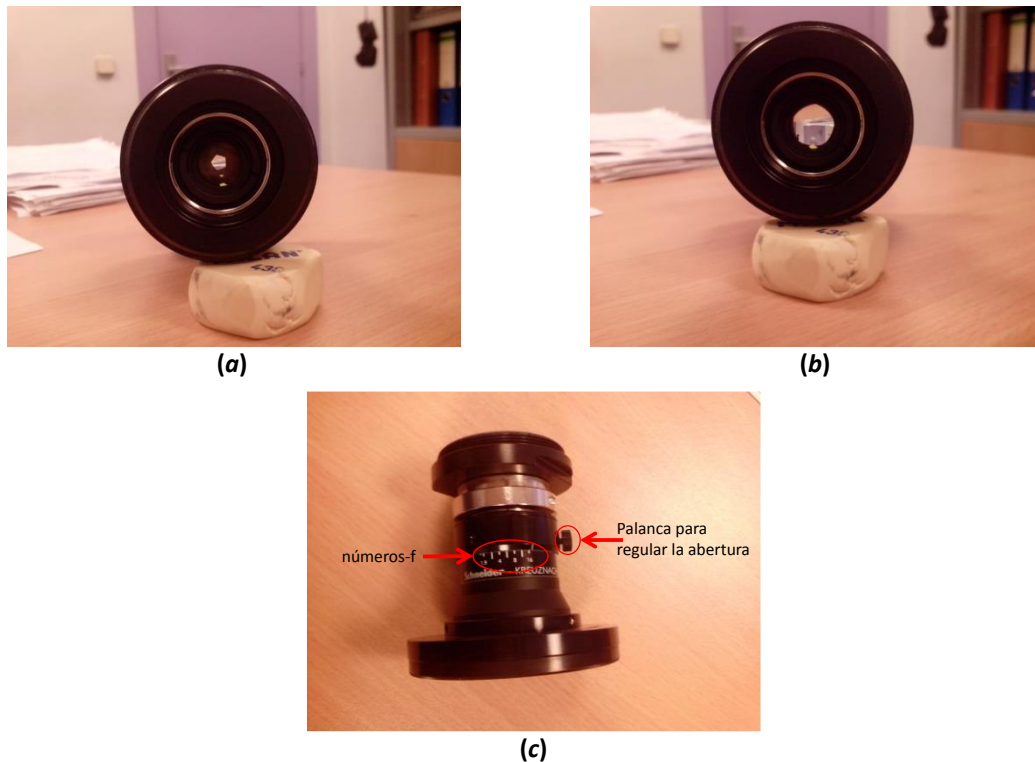


Figura 3-2. (a) Objetivo con un número-f grande, (b) el mismo objetivo con un número-f más pequeño y (c) objetivo visto desde un plano cenital

El tiempo de exposición junto con la apertura determinan lo que se conoce como exposición total del sistema. Habitualmente la exposición se expresa mediante el valor de exposición, EV (*Exposure Value*). Una expresión apropiada para calcular el EV a partir del tiempo de exposición y la apertura de la lente se puede hallar en Jacobson y col. (2000). Dicho autor utiliza la ecuación (3.2) para hallar un valor del EV,

$$EV = 3.32 \times \log \left(\frac{N^2}{t} \right) \quad (3.2)$$

así, con un tiempo de exposición de 1/125 segundos y una apertura dada por el *número-f* de 11, el EV será de 13.88.

Estas relaciones se usan normalmente en fotografía profesional para obtener imágenes de calidad mediante una la combinación del EV con uno cualquiera de los otros dos parámetros involucrados en la ecuación (3.2). En el ámbito profesional existe un general conocimiento sobre los valores más apropiados para un determinado tipo de escena, de suerte que dado un valor EV y fijado t se puede determinar N para

conseguir una calidad de imagen deseada, que ya ha sido suficientemente contrastada. De hecho, existen tabulaciones y representaciones gráficas para determinar N frente a EV y t conocidos o fijados de antemano, algo por otra parte fácil de obtener a partir de la ecuación (3.2). Supongamos que para una escena con luz de sol clara se sabe que un valor apropiado de EV es 12, pues bien, si fijamos N a 14 obtenemos un tiempo de exposición de acuerdo con la expresión (3.2) de 1/16 segundos. En los diagramas de exposición típicos el resultado es un valor de 1/15 segundos, obteniéndose un error entre ambos valores de aproximadamente un 6%.

Lo anterior resulta muy apropiado para la obtención de imágenes de alta calidad visual, si bien desde el punto de vista de la aplicación agrícola no es éste precisamente el interés, pues de lo que se trata es de obtener buenos resultados en la segmentación de las imágenes de cara al tratamiento que se pretende. En los entornos agrícolas y en los momentos del tratamiento de las malas hierbas de cara a su eliminación, la variabilidad de las condiciones ambientales, en lo que a la iluminación se refiere, puede ser considerable. Por ejemplo se puede pasar en cuestión de segundos de una situación con elevada intensidad luminosa a otra con abundantes sombras por la existencia de una alta densidad de nubes en movimiento. Ello conllevaría en todo momento determinar de forma automática estas situaciones para elegir los mejores valores posibles de EV que se llevarían tabulados, algo que no parece resultar apropiado en el sistema automático desarrollado. Además, los expertos en fotografía a la hora de determinar el EV tienen muy en cuenta la nitidez o no de las sombras, si existen, un factor determinante en el contraste de la imagen.

Por las razones antes expuestas, en el sistema de visión utilizado en la presente investigación se ha optado por controlar directamente el tiempo de exposición, ya que la cámara utilizada, Figura 3-1, lo permite. El control se realiza en función de la variación de la intensidad de la imagen, que se determina mediante análisis del histograma para cada imagen capturada, de suerte que si los valores resultantes del análisis no son satisfactorios, el tiempo de exposición se modifica convenientemente, tal y como se explica con detalle en el capítulo cuatro, hasta conseguir una imagen con la calidad suficiente de cara a los procesos automáticos posteriores.

3.3. Índices de vegetación: detección del verdor

Los VI 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. Los VI 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 aplicaciones agrícolas permiten la obtención de la cobertura vegetal.

Un índice de vegetación permite obtener una medida que caracteriza propiamente la vegetación, calculada a partir de los valores de reflectancia producidos por la radiación solar al incidir sobre los elementos presentes en la escena y captada por el sensor CCD de la cámara. Existen diferentes índices que se obtienen a través de la combinación de bandas espectrales abarcando un amplio rango de longitudes de onda del espectro comprendido aproximadamente entre los 400nm y los 2500nm (Asner, 1998). Este ancho de banda comprende tanto el espectro visible como la radiación infrarroja. La Tabla 3-1 muestra los distintos rangos del espectro en función de las longitudes de onda, λ , asociadas.

Espectro	λ (nm)	Espectro	λ (nm)
Luz visible	400-790	Infrarrojo	800-2500
Azul	450-500	Infrarrojo cercano (NIR)	750-1300
Verde	500-600	Onda corta del infrarrojo 1	1500-1900
Rojo	600-700	Onda corta del infrarrojo 2	1900-2450

Tabla 3-1. Longitudes de onda del espectro visible y luz infrarroja

Existen distintos dispositivos sensibles a las diferentes longitudes de onda. Los dispositivos de captura de imágenes en color, tal como el sensor mostrado en la Figura 3-1 y utilizado en la investigación desarrollada, es sensible a las longitudes de onda del espectro visible, pudiendo verse influido también por radiaciones infrarrojas, razón por la cual se utiliza el filtro UV/IR 486 mencionado previamente, con el fin de evitar la incidencia de este tipo de radiación por las razones expuestas con anterioridad. En este caso, el dispositivo proporciona una imagen de salida compuesta por los tres canales espectrales visibles R, G y B correspondientes al Rojo, Verde y Azul respectivamente, cuya respuesta espectral se muestra en la Figura 3-3, cada una de las curvas representa la correspondiente respuesta espectral del canal de su color en función de la longitud de onda en nanómetros (*nm*).

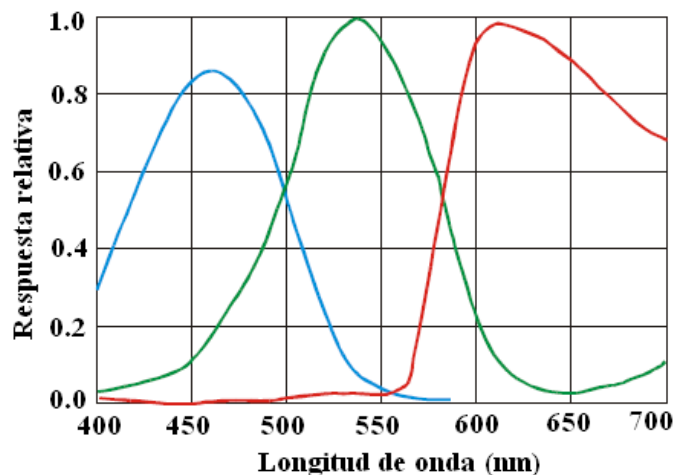


Figura 3-3. Respuesta espectral en el rango visible del espectro.

A la vista de las curvas representadas, se deduce que la máxima respuesta para el canal verde se produce en torno a los 540 nm, mientras que en el caso del rojo la respuesta máxima se alcanza alrededor de los 620 nm. Se observa además un fenómeno en este último caso, y es la fuerte respuesta del canal rojo para longitudes de onda que superan los 700 nm, llegando a penetrar en el infrarrojo cercano. Esto provoca una distorsión de colores en la imagen tal y como aparece en la imagen de la Figura 3-4(a), cuyo efecto debe compensarse con la utilización del filtro de corte UV/IR 486 (Schneider 2014) mostrado en la Figura 3-1, que bloquea la radiación infrarroja, obteniéndose la imagen que se muestra en la Figura 3-4(b)

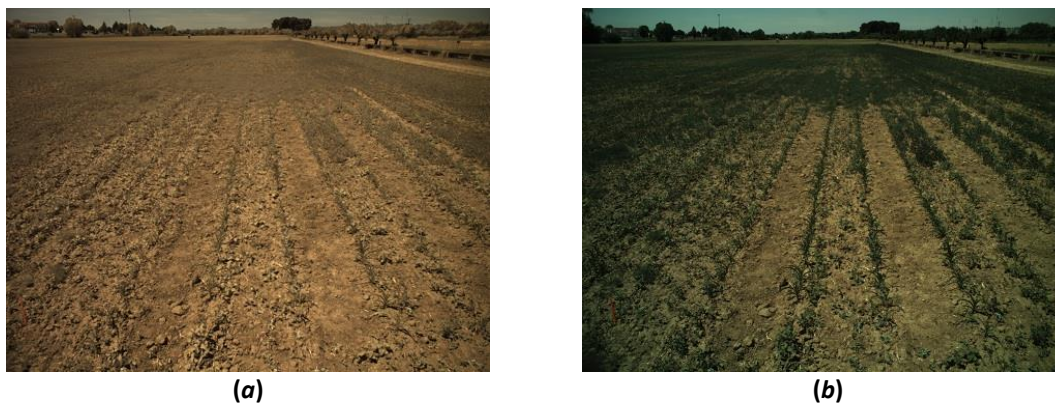


Figura 3-4. (a) Imagen de líneas de cultivo tomada sin filtro de corte UV/IR; (b) la misma imagen con filtro de bloqueo UV/IR.

Existen dispositivos sensibles a un amplio rango de longitudes de onda, proporcionando imágenes correspondientes a diferentes bandas, dependiendo del número de bandas utilizadas se habla de sensor multispectral o hiperespectral.

Dependiendo sobre qué tipo de material incida la radiación solar se produce una distinta interacción con la materia, así por ejemplo la tierra emite una reflectancia distinta a la vegetación. De igual modo, distintos tipos de plantas y partes de éstas tienen un comportamiento diferente a la hora de reflejar la energía que proviene de la radiación solar. Una forma de observar y analizar el fenómeno de la reflectancia es mediante dispositivos generadores de imágenes sensibles a las distintas longitudes de onda incidentes. Según los sensores que se utilicen para detectar estas longitudes de onda se obtienen imágenes multiespectrales, como las del Landsat (2013) que puede contener hasta 11 bandas mediante los “Landsat 8 Operational Land Imager (OLI) y Thermal Infrared Sensor (TIRS)”, o hiperespectrales como las de AVIRIS (2013) que utiliza 224 bandas con longitudes de onda que varían desde los 400nm hasta los 2500nm. Existen dispositivos hiperespectrales con un número de bandas aún mayor. La absorción y reflectancia de la radiación solar es el resultado de las interacciones con los materiales y varía significativamente según la longitud de onda. Los elementos que componen la vegetación tales como agua, pigmentos, nutrientes o carbono se pueden detectar analizando el espectro 400nm-2500nm, existiendo solapamientos en las bandas que se detectan, si bien cada uno de ellos es espectralmente distinto y posee reflectancias bien diferenciadas. Estas características hacen único a cada elemento y en su conjunto se conocen como firma espectral. Mediante esta firma se pueden combinar las medidas de reflectancia en las diferentes longitudes de onda y resaltar el elemento o característica de la vegetación que se busca en una imagen dada. Determinadas combinaciones de reflectancias dan lugar a los VI mencionados previamente.

El éxito para detectar y analizar los elementos que se desean, requiere conocimiento sobre las medidas de las bandas espectrales que se utilizan para cada índice, y un entendimiento de la forma en la que dichas bandas son detectadas por los sensores.

Los sensores de los que se ha hablado hasta el momento hacen uso de varias bandas y están situados bien en satélites (Landsat) o en aviones (AVIRIS) que sobrevuelan la zona de la que se quiere obtener información sobre la vegetación. Existen sistemas multi e hiper-espectrales más modestos que se han utilizado en agricultura y silvicultura a bordo de vehículos aéreos no tripulados, con la finalidad de analizar igualmente la vegetación. (Zarco-Tejada y col., 2012; Peña y col., 2013; Berni y col., 2009b).

En los proyectos bajo los que se ha desarrollado la investigación que se presenta se ha optado por la cámara a color mostrada en la Figura 3-1, por resultar suficiente para el propósito pretendido, siendo a la vez técnicamente un sensor más sencillo para su integración en el conjunto del equipo de cara a la futura comercialización. No obstante, como se verá más adelante en el capítulo cuatro, los métodos y procedimientos desarrollados poseen la suficiente flexibilidad como para considerar los sensores multi o hiper-espectrales si en algún momento son necesarios en función de la aplicación prevista.

A continuación se realiza una revisión de los índices de vegetación que se han estudiado como posibles candidatos a la hora de ser utilizados para detectar la vegetación. Todos se basan en el modelo de color RGB, existiendo un canal para las diferentes respuestas espectrales correspondientes a las longitudes de onda del espectro visible en los rangos Rojo (R), Verde (G) y Azul (B), de suerte que cada píxel de la imagen viene caracterizado por tres valores, que se corresponden con dichos canales. Al utilizar VI basados en el color se puede aprovechar el conocimiento de cómo percibe el ser humano el verdor de la vegetación o la información del resto de bandas que contienen los diferentes colores, de esta forma puede llegar a ser intuitivo el crear un índice de vegetación basado en el color. En el cómputo del índice NDVI, que se incluye a continuación, se utiliza la banda del infrarrojo, algo que no es posible en nuestro caso, al no disponer de esa banda en el sensor.

Se ha realizado una sencilla clasificación de los VI según dos categorías distintas. En la primera se engloban los índices de vegetación simples, donde se combinan los canales R, G y B de cada píxel para obtener un valor con el cual poder decidir más adelante si se trata de un píxel que pertenece a una planta o no. En la segunda categoría se encuentran los índices de vegetación compuestos, los cuales están diseñados como combinación de otros índices de vegetación simples.

3.3.1. Índices de vegetación simples

Como paso previo al cálculo de cualquier índice de vegetación es conveniente realizar una normalización de los valores en cada canal para que todos ellos se sitúen en el mismo rango. Estas operaciones son bastante corrientes en la literatura (Woebbecke y col., 1995; Gée y col., 2008; Meyer y Camargo-Neto, 2008), con ellas se consigue que los valores que intervienen en el índice se encuentren en el intervalo [0,1] y la suma de

los valores de los tres canales sea la unidad. La primera operación de normalización consiste en dividir cada uno de los valores de los canales de los píxeles, que constituyen la imagen, por un valor. Este valor puede ser constante o variable para cada imagen. En el caso de ser constante se trata del máximo valor que puede tomar un canal y en nuestro caso al ser imágenes de 24 bits por píxel es 255. En el caso de que se opte por utilizar un valor variable se elige para cada canal el máximo valor existente en dicho canal. En la práctica el máximo valor en todos los canales suele ser el 255, con lo que no existe diferencia entre una u otra normalización. Si denominamos mediante R , G y B a los valores de intensidad en los canales rojo, verde y azul respectivamente para cada píxel, la operación de normalización es la que aparece en la ecuación (3.3)

$$\hat{R} = R/R_m, \quad \hat{G} = G/G_m, \quad \hat{B} = B/B_m \quad (3.3)$$

donde R_m , G_m y B_m son respectivamente, los máximos valores en su correspondiente canal, cuyos valores normalizados son \hat{R} , \hat{G} y \hat{B} .

La siguiente operación en la ecuación (3.4) tiene como finalidad que la suma de las componentes de los tres canales sume uno, es decir $r + g + b = 1$.

$$r = \frac{\hat{R}}{\hat{R} + \hat{G} + \hat{B}}, \quad g = \frac{\hat{G}}{\hat{R} + \hat{G} + \hat{B}}, \quad b = \frac{\hat{B}}{\hat{R} + \hat{G} + \hat{B}} \quad (3.4)$$

A partir de este momento, cuando se haga referencia a los valores de los píxeles de los canales RGB se entenderá por tal los valores normalizados de las coordenadas cromáticas r , g y b de la ecuación (3.4).

Los siguientes seis índices constituyen la lista de los analizados en el marco de la presente investigación:

1. *Exceso de Verde* (ExG, *Excess of Green*) (Woebbecke y col., 1995a; Ribeiro y col., 2005; Shi y col., 2007), hace uso de la idea intuitiva y evidente ante el ojo humano de que la vegetación es verde, así cuanto más verde sea un elemento de la imagen más probabilidad existirá de que se trate de vegetación, mientras que si las componentes que predominan son la roja o la azul, es muy probable de que no se trate de una planta verde. Debido a que tiene mayor relevancia la componente verde que las otras dos, roja y azul, y a que existe una desigualdad al tratarse de un canal frente a dos, se le dota de mayor importancia al canal

verde multiplicándole por un factor de dos. El índice ExG se obtiene como el valor del canal verde menos el valor del canal azul y menos el valor del canal rojo, ecuación (3.5)

$$ExG = 2g - b - r \quad (3.5)$$

2. *Exceso de Rojo (ExR, Excess of Red)* (Meyer y col., 1998), cuando lo que se quiere enfatizar es el rojo frente a los otros dos, por ejemplo para descartar ahora elementos no verdes, se utiliza este índice o su homólogo en el azul (ExB) para eliminar partes rojizas (suelo del campo de cultivo) o azuladas (cielo en la imagen).

$$ExR = 1.4r - g \quad \text{y} \quad ExB = 1.4b - g \quad (3.6)$$

3. *CIVE (Color Index of Vegetation Extraction)* (Kataoka y col., 2003), se basa en la combinación lineal de los tres canales R,G,B, incluyendo a su vez un término aditivo independiente, proporcionado por los autores y obtenido de forma experimental, ecuación (3.7)

$$CIVE = 0.441r - 0.811g + 0.385b + 18.78745 \quad (3.7)$$

4. *VEG* (Hague y col., 2006). Este índice hace uso de la constante a , cuya finalidad es tratar de ser invariante a la temperatura de color de la iluminación de un cuerpo negro, es decir, trata de ser insensible a la amplitud de la iluminación. El valor propuesto para dicha constante por sus autores es de 0.667, viene definido según la ecuación (3.8) que es el utilizado en este estudio

$$VEG = \frac{g}{r^a b^{1-a}} \quad (3.8)$$

5. *NDI (Normalized Difference Index)*, Pérez, y col. (2000) propusieron este índice. Observaron que si bien en las plantas los valores de las componentes espectrales roja y azul son pequeños en contraposición con los de la verde que es relativamente elevada, utilizar los valores RGB no normalizados no resulta suficiente para diferenciar entre plantas y tierra del suelo. Este hecho puede explicarse teniendo en cuenta que los valores R, G, B no normalizados son

directamente proporcionales a la luz total reflejada por cualquier superficie y por lo tanto, son muy sensibles a la intensidad de la fuente de iluminación, así como al ángulo de visión y al ángulo de procedencia de la iluminación. La solución fue proponer una diferencia normalizada entre los canales rojo y verde, ecuación (3.9).

$$NDI = \frac{g - r}{g + r} \quad (3.9)$$

1. *NDVI (Normalized Differential Vegetation Index)* (Rouse y col., 1973), que utiliza la región del infrarrojo cercano (NIR) y la banda verde normalizada del espectro visible. Fue un índice pionero para la detección de vegetación a partir de imágenes, utilizado en el satélite Landsat-1, se define según la ecuación (3.10).

$$NDVI = \frac{NIR - r}{NIR + r} \quad (3.10)$$

3.3.2. Índices de vegetación compuestos

Los índices de vegetación compuestos surgen como combinaciones de índices simples, descritos previamente. El objetivo que se persigue con la combinación es aprovechar la bondad que manifiestan por separado cada uno de los índices simples. Destacan como más relevantes los siguientes,

1. *ExGR (Neto, 2004) (Excess Green minus Excess Red)*, definido según la ecuación (3.11), como su nombre indica consiste en la diferencia entre ExG menos ExR. La idea consiste en enfatizar los verdes en detrimento de los rojos sobre el supuesto de que en los campos de cultivo la clave reside en destacar los verdes de las plantas frente a los tonos rojizos del suelo.

$$ExGR = ExG - ExR \quad (3.11)$$

2. *COM (Guijarro y col., 2011)* combina linealmente los cuatro índices siguientes según la ecuación (3.12): ExG, ExGR, CIVE y VEG. Los coeficientes utilizados en la combinación se determinan experimentalmente. Los estudios realizados por dichos autores se encaminaron a determinar la eficiencia de cada uno de los

índices que participan en la combinación. En función del grado de eficiencia obtenido, se asigna un peso específico a cada índice individual con el que participa finalmente en la combinación para la obtención del índice compuesto.

$$COM = 0.25 \times ExG + 0.30 \times ExGR + 0.33 \times CIVE + 0.12 \times VEG \quad (3.12)$$

Los métodos propuestos en el capítulo cuatro, como resultado de la investigación desarrollada, hacen uso de índices de vegetación para la segmentación de las imágenes. De acuerdo con los resultados experimentales obtenidos, para las imágenes captadas con el sistema de visión de la Figura 3-1, así como por los informes proporcionados por los diferentes autores en relación al comportamiento de los mencionados índices se utilizan los índices ExG y una versión del COM. El ExG se utilizará en el primer SEA, cuya finalidad es la detección de líneas de cultivo y malas hierbas, mientras que el COM se utilizará en el segundo SEA, cuya finalidad es la detección de plantas de cultivo y malas hierbas enmascaradas.

3.4.Umbralización

3.4.1. Transformación de escala

La umbralización es en realidad una parte del proceso global de segmentación de imágenes. Consiste en transformar una imagen en escala de grises a una imagen en la que aparecen diferentes elementos etiquetados según sus niveles de intensidad en dicha escala. Dada la distribución de los valores en la escala de grises, o si se prefiere el histograma de la misma, la operación de umbralización selecciona ciertos valores característicos como valores umbral que separarán los píxeles de la imagen en grupos de acuerdo con sus valores de gris en la distribución. En el caso más simple, la umbralización consiste en seleccionar un único valor umbral u de forma que todos los píxeles cuyos valores sean mayores o iguales que u se considerarán píxeles de primer plano y el resto como píxeles de fondo. Esta umbralización se conoce como umbralización superior. Si la relación se establece con el criterio del menor, la umbralización se conoce como inferior.

Pues bien, dada la imagen en escala de grises, obtenida por la aplicación de los VI descritos previamente, el proceso de binarización, en el marco de la investigación propuesta, tiene como objetivo discriminar entre plantas verdes (maíz y malas hierbas) por lo que para tal propósito con un único umbral resulta suficiente.

Para una binarización efectiva, el rango de valores de la imagen de entrada, en nuestro caso las imágenes en escala de grises, requieren un escalado de sus valores para restringirlos a un rango de valores efectivo, siendo apropiado el intervalo [0,255]. Los valores de las imágenes en escala de grises pueden tomar valores en muy diferentes rangos. Así, por ejemplo las imágenes captadas directamente por los sensores en escala de grises o la de intensidad, I , obtenida por la transformación desde el modelo de color RGB al HSI (Pajares y Cruz, 2007) varían en el rango [0,255] cuando su representación es en 8 bits/píxel. La aplicación del índice ExG genera imágenes en el rango [-2,2], mientras que en el caso de CIVE, el rango es [17.97645, 19.61345].

En cualquiera de estos dos últimos casos la transformación de una escala dada a la deseada se realiza mediante la ecuación (3.13)

$$N = \frac{A - Min_i}{Max_i - Min_i} \times (Max_f - Min_f) + Min_f \quad (3.13)$$

donde A representa el valor de un píxel en la escala de grises inicial; Min_i y Max_i son los valores mínimo y máximo, respectivamente, en la escala inicial; Min_f y Max_f son los valores mínimo y máximo, respectivamente, que se desea tomen los píxeles en la escala resultante, que en el caso de su transformación al rango [0,255] Min_f y Max_f toman exactamente esos valores extremos.

3.4.2. Método de Otsu

El método de Otsu (1979) tiene como finalidad la obtención del pretendido umbral de binarización a partir de la imagen de grises normalizada, previa aplicación de los VI.

Este método toma como referencia el histograma de la imagen de grises normalizada y determina un único umbral, para ello se recorre el histograma desde el valor mínimo al máximo, asignando un valor de umbral según el valor del histograma analizado. Este valor divide el histograma en dos modos o clases, cuyas distribuciones de intensidades asume que siguen una distribución Gaussiana. En esta distribución se evalúa la varianza intraclase y la varianza interclase junto con la varianza total. El objetivo es minimizar la varianza intraclase y maximizar la varianza interclase. El umbral que consiga optimizar estos dos objetivos se selecciona como el más apropiado.

El número de píxeles con intensidad o nivel de gris i en la imagen se denota por n_i , generalmente con 256 valores en el rango $[0,1]$, el número total de píxeles en la imagen es N , la probabilidad de ocurrencia de cada nivel de gris es $p_i = n_i/N$, finalmente el valor de umbral se denota por k .

Para cada umbral k se establecen dos clases C_1 y C_2 , si el valor de un píxel dado es mayor que k el píxel pertenecerá a la clase C_2 y si es menor o igual pertenecerá a la clase C_1 . Por lo tanto las probabilidades de ocurrencia de cada clase quedan como:

$$\omega_1^k = \sum_{i=0}^k p_i, \quad \omega_2^k = \sum_{i=k+1}^{255} p_i = 1 - \omega_1^k \quad (3.14)$$

donde ω_j^k es la probabilidad de ocurrencia acumulada para un píxel perteneciente a la clase j de acuerdo al umbral k y p_i es la probabilidad de ocurrencia para el nivel de gris i . De forma similar los momentos estadísticos de primer orden sobre el origen de las dos clases son:

$$\mu_1^k = \sum_{i=0}^k ip_i, \quad \mu_2^k = \sum_{i=k+1}^{255} ip_i \quad (3.15)$$

donde μ_1^k y μ_2^k representan respectivamente las medias de los valores de los píxeles para las clases C_1 y C_2 , utilizando en ambos casos como umbral el valor k . A partir de las ecuaciones (3.14) y (3.15) se obtiene la expresión (3.16) que se obtiene multiplicando los valores de probabilidad de pertenencia a las clases C_1 y C_2 por la diferencia de los momentos de primer orden de ambas clases al cuadrado.

$$\sigma_B^2(k) = \omega_1^k \omega_2^k (\mu_1^k - \mu_2^k)^2 \quad (3.16)$$

El valor dado por $\sigma_B^2(k)$ determina la bondad del umbral k . De todos los posibles valores k , candidatos a umbral, se selecciona el que corresponda al mayor $\sigma_B^2(k)$. Cuanto mayor sean las probabilidades de ocurrencia ω_1^k y ω_2^k , mayor será el valor $\sigma_B^2(k)$, si bien, a partir de las ecuaciones (3.14) se deduce que cuando una probabilidad se incrementa, la otra se reduce, ya que ambas deben sumar la unidad, el

método busca un equilibrio entre ambas. El hecho de que el producto de las probabilidades de ocurrencia sea elevado refleja que cuanto mayor sea la probabilidad de pertenecer a una clase mejor será la clasificación, recogiendo el criterio de maximización de la varianza interclase. La minimización de la varianza intraclase se refleja en el cuadrado de la resta de las medias de las clases. Cuanto mayor sea la diferencia entre las medias de los elementos de las clases mayor será la diferencia entre las dos clases, que es lo que se persigue. La decisión final para elegir el mayor $\sigma_B^2(k)$ se determina mediante el criterio expresado en la ecuación (3.17),

$$\sigma_B^2(K_0) = \max_{0 \leq k \leq 255} \sigma_B^2(k) \quad (3.17)$$

donde el valor que finalmente se escoge como umbral es K_0 .

El siguiente ejemplo pedagógico ilustra intuitivamente el funcionamiento del método de Otsu. El objetivo es encontrar un umbral óptimo mediante Otsu en una imagen en escala de grises con valores comprendidos en el intervalo $[0, 255]$. En la Figura 3-5(a) se muestra una imagen original en escala de grises, cuyos píxeles toman valores en ese intervalo. El histograma asociado a la imagen se muestra en la Figura 3-5 (b). Fijándose en dicho histograma se deduce a simple vista que se podrían realizar dos divisiones distintas del mismo eligiendo umbrales diferentes, con valores en torno a 40 y 160.

El segundo valor se ratifica mediante el algoritmo de Otsu, al elegir 160. A la vista del histograma parece lógico que haya escogido el segundo umbral ya que incluso desde un punto de vista lógico-estadístico resulta mucho más evidente que el primero. En efecto, tanto en las proximidades del 40 y del 160 se produce un descenso en la ocurrencia de los píxeles bastante clara, pero en el caso del segundo umbral además existe una gran cantidad de píxeles con valores inferiores y otra gran cantidad con valores superiores, mientras que existen muy pocos píxeles con valores inferiores a 40. En la Figura 3-5(c) se muestra la binarización de la imagen de la Figura 3-5(a) mediante el valor de 160 obtenido por Otsu.

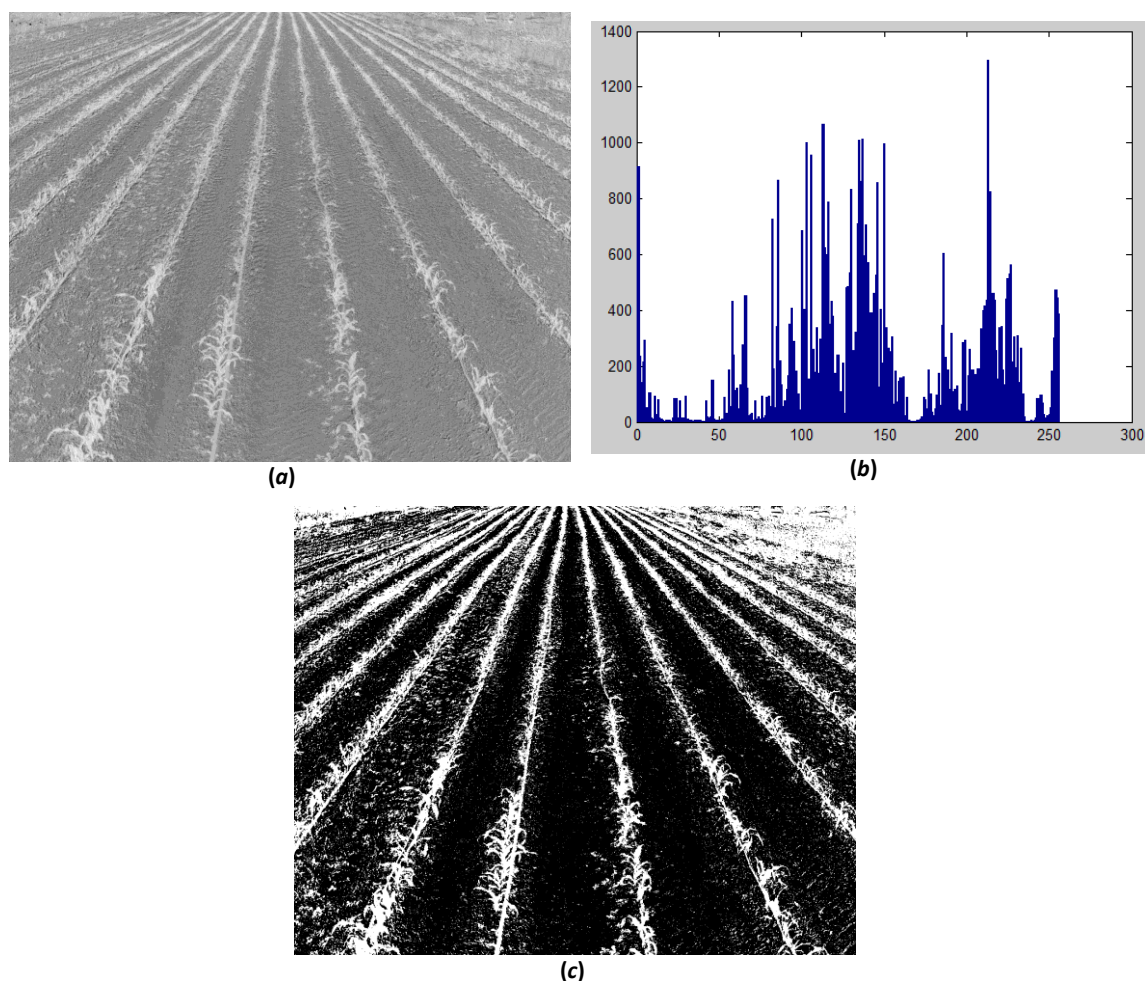


Figura 3-5. (a) Imagen en escala de grises; (b) Histograma asociado a la imagen en (a); (c) Imagen binarizada resultante tras aplicar Otsu.

3.5. Ajuste de líneas

3.5.1. Consideraciones generales

Dada la imagen binaria, obtenida siguiendo los procesos anteriores, existen en ella distribuciones de puntos que pertenecen a plantas, etiquetados con un determinado valor, y otros que no son plantas, etiquetados con un valor diferente. Dentro de las agrupaciones de píxeles identificando plantas, existirán alineaciones que se corresponden con las líneas de cultivo, mientras que aquellos otros pertenecientes a plantas de malas hierbas probablemente no sigan ninguna alineación, por su aleatoriedad en el campo. Pues bien, el objetivo que se persigue en este momento, es tratar de identificar las alineaciones de píxeles relacionadas con las líneas de cultivo mediante el ajuste de líneas rectas.

El término ajuste de líneas hace referencia al proceso de generar un modelo o función matemática que mejor se adapte a la distribución de una nube de puntos

existente, pudiéndose aplicar una serie de restricciones. El modelo o función que caracteriza los puntos se puede crear de dos formas distintas dependiendo de la naturaleza de los puntos y la finalidad que se persiga.

Una forma de crear el modelo es mediante interpolación de los datos de partida. Si se pretende utilizar este enfoque no pueden existir dos puntos que compartiendo el valor de la coordenada x tengan distinto el valor de la coordenada y , ya que entonces dado un valor para la coordenada x no se podría saber con exactitud qué valor tomaría la coordenada y , al existir más de una posibilidad. Con este enfoque el número de funciones que se obtienen es igual al número de puntos menos uno. Se asociará una función a cada dos puntos que se encuentren uno a continuación del otro, uniendo esta función ambos puntos. En caso de querer estimar la coordenada y de un punto conociendo la coordenada x situada entre dos puntos conocidos, basta con sustituir en la función asociada a dichos puntos conocidos el valor de la coordenada x . El tipo de estas funciones asociadas a dos puntos contiguos de la distribución es variable, se pueden utilizar el que se desee, desde una simple línea recta, pasando por polinomios de distinto grado o ecuaciones exponenciales, todo depende de si se conoce aproximadamente cómo podría variar la distribución de los píxeles que se pretenden ajustar entre dos datos conocidos.

Otro procedimiento para crear una función que represente, con una cierta fidelidad, una determinada distribución de puntos consiste en capturar la tendencia general de los puntos sin exigir que la función obtenida se ajuste de forma exacta a dicha distribución, ya que esto en algunos casos llega a ser imposible debido a la existencia de puntos con el mismo valor en la coordenada x pero distinto valor en la coordenada y . De este modo se pretende estimar dónde se encontraría un punto sabiendo su coordenada x , pero sin conocer su localización exacta.

En el problema de ajuste de líneas a las de cultivo, resulta más apropiado el segundo enfoque, abordando el problema mediante ajuste por mínimos cuadrados que se expone a continuación.

3.5.1.1. Mínimos cuadrados totales

El ajuste de mínimos cuadrados lineal es probablemente la forma más simple de regresión lineal, la regresión lineal simple. La regresión lineal es un enfoque con el que se intenta modelar la relación entre una variable dependiente y una o más variables independientes. En caso de que la variable dependiente se vea influida únicamente por una variable independiente se estará ante una regresión lineal simple.

Mínimos cuadrados es un enfoque general para aproximar la solución de un sistema sobredeterminado. Los sistemas sobredeterminados son aquellos en los que el número de ecuaciones o restricciones es superior al número de incógnitas. Se trata de sistemas sobredeterminados porque se intenta encontrar la ecuación de una línea recta que se aproxime a la distribución de la nube de puntos, pero no existe ninguna línea recta sobre la que se superpongan todos los puntos. La solución se obtiene minimizando la suma de los cuadrados de los errores o desplazamientos acumulados al resolver cada una de las ecuaciones de forma individual. La diferencia entre el valor observado y el valor esperado por el modelo se denomina residuo. Los problemas que se pueden resolver con mínimos cuadrados se pueden clasificar como lineales y no lineales dependiendo de si los residuos son lineales o no en todas las incógnitas. Debido a que en nuestro caso, identificación de líneas de cultivo, las funciones que se quieren estimar son siempre líneas rectas y que se quiere alcanzar un bajo coste computacional con fines de implementación en tiempo real, se considera exclusivamente la versión lineal, teniendo en cuenta que existe la versión no lineal por si en alguna aplicación fuera necesaria debido a la existencia de distribuciones agrícolas con disposiciones de este tipo (Bates y Watts, 1988; Johnson y col., 1981; Marquardt, 1963).

Se emplea la suma de los errores al cuadrado en vez de la suma de los errores en valor absoluto debido a que la suma de los cuadrados puede ser tratada como una función diferenciable continua. Sin embargo, debido a que el valor resultante de elevar al cuadrado un número mayor que uno es superior que el valor obtenido al realizar la operación del valor absoluto al mismo número, se puede producir un problema con valores periféricos que se alejen de la distribución seguida por la mayoría de los puntos. Este problema se traduce en un efecto desproporcionado en el ajuste de la línea recta, algo que no es deseable en ningún momento. Para nuestro caso se explicará en la

sección 4.2.4 cómo se evitan estos efectos no deseados mediante el uso de una plantilla o de una Unidad de Medida Inercial (IMU).

Los residuos pueden ser distancias verticales o perpendiculares respecto a la línea recta que se pretende estimar. En las Figura 3-6(a) y (b) se muestra cómo se emplean las distancias verticales y perpendiculares como residuos respectivamente. El método de mínimos cuadrados ordinario de Gauss y Legendre considera las distancias verticales y se suelen utilizar los errores como el cuadrado de la distancia desde los puntos a la función que se quiere ajustar a lo largo de la dirección del eje vertical. El método de mínimos cuadrados totales (Golub, 1973; Golub y Van Loan, 1980, Van Huffel y Vanderwalle, 1991) es una generalización natural del enfoque seguido en el método tradicional de mínimos cuadrados, utilizado cuando tanto la componente x como la componente y de los datos están afectados por perturbaciones (ruido).

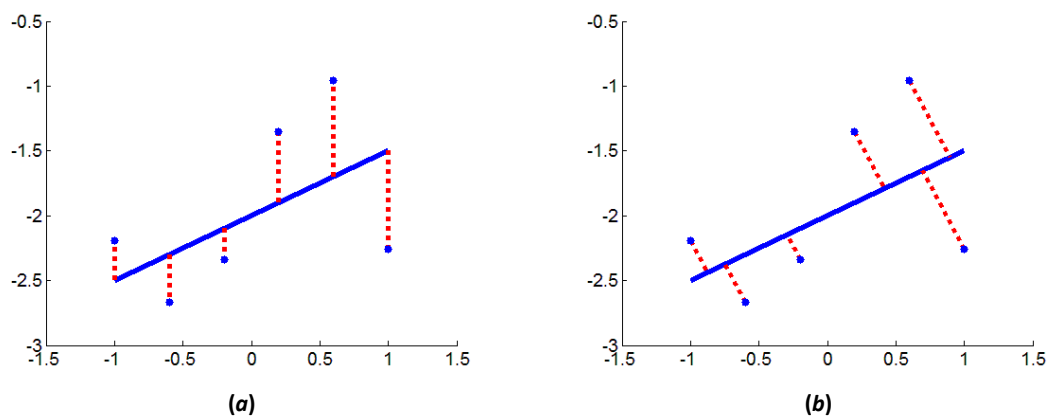


Figura 3-6. Distintas formas de contabilizar el error producido con respecto a la recta: (a) como distancia vertical; (b) como distancia perpendicular.

El problema del ajuste se plantea de la siguiente manera. Dado un conjunto de n puntos $X \equiv \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, el objetivo consiste en estimar la línea recta que mejor se ajuste a la distribución seguida por dichos puntos. La ecuación de la línea recta se especifica como $y = mx + a$, donde los parámetros a estimar son m y a (Markovsky y Van Huffel, 2007), que representan la pendiente de la recta y el término independiente respectivamente. De este modo, la expresión a minimizar se corresponde con el error total E , que es en definitiva la suma de los desplazamientos entre los distintos puntos y la línea recta a estimar, cuya expresión viene dada por la ecuación (3.18)

$$E = \sum_{i=1}^n d_i = \sum_{i=1}^n \frac{|y_i - (a + mx_i)|}{\sqrt{1 + m^2}} \quad (3.18)$$

donde d_i es el desplazamiento para el punto i cuyas coordenadas son (x_i, y_i) .

Minimizando la ecuación (3.18) se obtiene la ecuación (3.19):

$$a = \frac{\sum_{i=1}^n y_i - m \sum_{i=1}^n x_i}{n} = \bar{y} - m\bar{x}; \quad m = -B\sqrt{B^2 + 1} \quad (3.19)$$

El término B se define como sigue:

$$B = \frac{\left(\sum_{i=1}^n y_i^2 - n\bar{y}^2 \right) - \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)}{2 \left(n\bar{xy} - \sum_{i=1}^n x_i y_i \right)} \quad (3.20)$$

Capítulo 4

4. Sistemas Expertos Automáticos: detección de líneas de cultivo y discriminación entre plantas de cultivo y malas hierbas en maíz

4.1.Introducción

El presente capítulo describe los dos temas clave de la investigación, a saber:

- Detección de líneas de cultivo e identificación con posterior discriminación entre malas hierbas.
- Cultivos mediante un sistema de visión artificial.

Se describen exactamente dos procedimientos, con tal propósito. Cada uno de ellos consta de una secuencia de procesos elementales, convenientemente estructurados, de suerte que la información extraída en cada uno de los procesos elementales constituye la base para el siguiente. De esta forma, se completa un proceso general en cada uno de los dos casos. Para la secuenciación y estructuración de los dos procedimientos generales se ha aplicado conocimiento humano, siguiendo las pautas que un experto realiza para la solución del problema. Esto ha desembocado en sendos sistemas automáticos conteniendo el conocimiento del experto, de ahí su denominación como Sistemas Expertos Automáticos (SEA). En su diseño se ha tenido en cuenta su implantación en un tractor robotizado para tareas agrícolas en campos de maíz donde la identificación de plantas de cultivo, malas hierbas y las propias líneas de cultivo constituyen la base de la aplicación, naturalmente con las restricciones de tiempo real que se requieren.

La base teórica de los principales procesos elementales que conforman los SEA se encuentra descrita en el capítulo tres. En síntesis se concretan en lo siguiente:

- SEA para la detección automática de líneas de cultivo en campos de maíz (Montalvo y col., 2012). El sistema tiene como finalidad la identificación y diferenciación de las malas hierbas frente al cultivo presente en el campo. Además, se plantea como objetivo la determinación de las líneas de cultivo y la estimación del grado de infestación de malas hierbas en el campo para su posterior tratamiento. Uno de los avances clave del sistema frente a otros ya existentes de la misma naturaleza es su capacidad de realizar las discriminaciones e identificaciones señaladas incluso en campos con una alta densidad de malas hierbas y por tanto con una gran cobertura vegetal.
- SEA para la identificación de plantas tanto de maíz como de malas hierbas, en situaciones donde las plantas aparecen cubiertas parcialmente por el barro (plantas enmascaradas, que han perdido su tonalidad verdosa) debido a la lluvia o tras una sesión de regado, ya que el maíz es un cultivo típico de regadío (Montalvo y col., 2013). El avance que supone esta técnica estriba en la capacidad de discriminación del sistema para solucionar un problema real consistente en la identificación de plantas que de otro modo se confundirían con el suelo. Tal confusión es debida a la capa de barro de similares características espectrales que el mismo terreno de donde procede dicha capa. Hasta donde se conoce, este problema no había sido abordado previamente. Adicionalmente, gracias a la potente capacidad discriminante del sistema, también es posible diferenciar plantas en un estado de verdor saludable frente a otras cuyo estado no es tan favorable debido a que la función clorofílica no es tan intensa. Esta situación aparece tras un tratamiento de las malas hierbas para su eliminación, de suerte que al cabo de un espacio breve de tiempo, aquellas inician un proceso de secado, disminuyendo paulatinamente su función clorofílica a medida que van muriendo. Gracias al método propuesto es posible valorar la eficacia del tratamiento obteniendo información de la cantidad de verde que ha desaparecido en comparación a los días previos al tratamiento.

El primer SEA se describe en la sección 4.2 mientras que en la sección 4.3 se describe el segundo.

4.2.SEA para detección de líneas de cultivo y malas hierbas

La estructura del SEA que se describe en esta sección se muestra de forma esquemática en la Figura 4-1. Se observa que consta de cinco fases:

- a) Obtención de una imagen de calidad.
- b) Segmentación de la imagen.
- c) Procesos de umbralización.
- d) Extracción de las líneas de cultivo.
- e) Cálculo de las coberturas de malas hierbas.

En el capítulo tres se han explicado de forma detallada los fundamentos teóricos de los métodos principales que se utilizan en cada una de las distintas fases contempladas en dicho esquema. Ahora, cada una de las fases se describe en los correspondientes apartados. Se ha optado por dividir el esquema general en varios subesquemas más simples para facilitar su comprensión y visualización. De este modo se incluye la Figura 4-1 aquí para mostrar de forma global el proceso y en cada uno de los apartados, donde se detalla cada fase, se incluye un diagrama específico, con su correspondiente desglose y descripción.

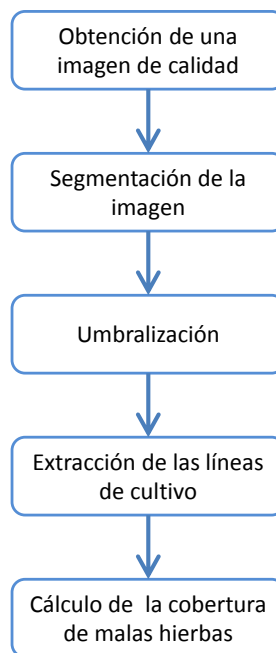


Figura 4-1. Esquema global del proceso.

En el apartado 4.2.1, *Control del tiempo exposición*, se explican las restricciones físicas de operación tales como el tipo de objetivo disponible o las condiciones lumínicas de la escena, para posteriormente continuar con el algoritmo utilizado para controlar el tiempo de exposición de forma que éste se adapte a las condiciones de iluminación variables del entorno. Este primer apartado, que constituye a su vez el primer paso es de vital importancia en el proceso global, ya que sobre la imagen capturada con el valor del tiempo de exposición hallado se realizarán todas las operaciones subsiguientes. Se trata en todo caso de conseguir imágenes con la calidad suficiente para los tratamientos posteriores, de forma que se eviten imágenes de baja calidad, que conducirían a resultados poco satisfactorios. En el apartado 4.2.2, *Segmentación de la imagen*, se explica el índice de vegetación utilizado para segmentar la imagen de forma que se resalte la vegetación presente en la imagen sobre el resto de elementos de la escena (suelo, cielo, edificaciones). En el apartado 4.2.3, *Doble umbralización mediante el método de Otsu*, se describe dicho método de umbralización utilizado por su buen comportamiento en las imágenes analizadas. Con este método se lleva a cabo la diferenciación entre vegetación y resto de elementos existentes en la imagen procedentes de la escena. Posteriormente dentro de la vegetación detectada se separan las plantas de cultivo y las malas hierbas existentes en la parcela. En el apartado 4.2.4, *Detección de las líneas de cultivo*, se describe el método utilizado para obtener las líneas de cultivo a partir de las imágenes binarias obtenidas previamente. El método que se ha utilizado para obtener estas líneas rectas, se basa en un ajuste mediante mínimos cuadrados para obtener finalmente la ecuación de la recta que mejor se ajusta a las mismas. El apartado 4.2.5, *Cálculo de las coberturas de malas hierbas*, versa sobre cómo se contabilizan las malas hierbas existentes en el terreno, separándolas del cultivo, para obtener una matriz de cobertura con la que determinar el nivel de infestación del cultivo por malas hierbas. En realidad, dicha matriz expresa el porcentaje de malas hierbas existentes en las correspondientes cuadrículas o celdas diseñadas específicamente, que se corresponden con celdas imaginarias sobre el terreno. Este último paso requiere conocer tanto las posiciones exactas de las líneas de cultivo como de las malas hierbas ubicadas en el campo. El objetivo final consiste en aplicar un tratamiento selectivo a las posiciones del terreno que se corresponden con las mencionadas celdas. Dependiendo de la densidad obtenida, así será la intensidad del tratamiento.

Todas las operaciones tienen en consideración solamente una parte de la imagen, esto se debe a que no es necesario procesarla al completo al ser solamente de interés una región (ROI, *Region Of Interest*), donde se encuentra el área de cultivo que se quiere tratar. El hecho de considerar sólo la parte de la imagen a tratar tiene como finalidad reducir considerablemente el tiempo de ejecución o coste computacional del procedimiento, cuya finalidad última es el desarrollo de un método capaz de operar en tiempo real.

4.2.1. Control del tiempo exposición

En la sección 3.2 se explicó el concepto relativo al valor de exposición y cómo éste influye de forma determinante en la calidad de las imágenes, con el fin de obtener imágenes con la suficiente calidad para conseguir resultados satisfactorios. En dicha sección se establecieron las relaciones entre el valor de exposición, el tiempo de exposición y la apertura, concluyendo que en el sistema de visión utilizado en los experimentos, la variable a controlar es exclusivamente el tiempo de exposición.

Antes de explicar los algoritmos que se utilizan para controlar el tiempo de exposición, se enumeran las restricciones de tipo físico que afectan a su control, a saber: a) objetivo del que se dispone para la cámara; b) las condiciones lumínicas del entorno en el momento de realizar las labores agrícolas y por tanto en el momento de captura de las imágenes.

Debido a que la apertura de la lente se regula de forma mecánica, mediante la palanca que se observa en la Figura 3-2(c), no es posible automatizar este mecanismo por ser de regulación manual y no se puede adaptar dinámicamente según las necesidades y cambios que impongan las condiciones lumínicas del entorno en el que se encuentra trabajando la cámara. Tal y como se mencionó en el capítulo tres existen objetivos en los que sí es posible regular la apertura de la lente mediante comandos software o señales eléctricas externas que permiten su control automático en tiempo de ejecución. Esta restricción nos obliga a fijar de antemano una apertura dada para el objetivo, de forma que con dicha apertura la cámara trabajará durante todo el tiempo de operación. Esto que en principio podría parecer una desventaja no es realmente así, ya que como también se indicó en el capítulo tres, va a permitir que la única variable a controlar sea el tiempo de exposición, algo que sin duda facilitará la labor de cara a la obtención de imágenes de calidad.

Como se debe establecer una apertura en la lente antes de trabajar, habrá que determinar cuál es una buena opción, no se puede establecer una cualquiera de forma aleatoria ya que podría no alcanzarse la exposición deseada y correcta para trabajar variando únicamente el tiempo de exposición. Para determinar la apertura es necesario analizar previamente el entorno en el que se trabaja, con sus características de luminosidad. El entorno de trabajo es un entorno de exterior, exactamente un campo de maíz y la época en la que se realiza el tratamiento del cultivo suele ser el mes de mayo de cada año. En esta época del año existe una alta luminosidad debido a las condiciones solares, siendo conveniente fijar una pequeña apertura, o lo que es lo mismo un *número-f* grande, para no saturar el sensor CCD por exceso de luminosidad. Si se utilizase un *número-f* pequeño el sensor podría verse impactado por un exceso de iluminación procedente del entorno de forma que aunque se redujese el tiempo de exposición no se conseguiría tener una imagen con la calidad adecuada, todos o la mayoría de los píxeles del CCD proporcionarían unos niveles de intensidad elevada una vez realizada la correspondiente conversión a imagen, perdiendo parte importante de la información de la escena. Cuando esto ocurre, la imagen aparece deteriorada, conociéndose vulgarmente como “imagen quemada”. Por otro lado, es posible obtener un beneficio adicional al trabajar con valores grandes de apertura, al ser el cuadrado del *número-f* (apertura) directamente proporcional al tiempo de exposición para un mismo valor de exposición (ecuación (3.2)), el tiempo de exposición será pequeño, lo que reduce el tiempo necesario para que la cámara capture una imagen y la trasmita. Esto es importante, ya que como se ha mencionado previamente se trata de una aplicación de tiempo real y cuanto menos tiempo se emplee en el proceso de captura y envío al elemento de procesamiento, de mayor tiempo dispondremos para el tratamiento de la imagen. En resumen, debemos trabajar con un *número-f* lo suficientemente grande para que la imagen no resulte quemada, llegando a un compromiso de mínimos para que el tiempo de exposición sea lo más pequeño posible, pudiendo así dedicar más tiempo al procesado de la imagen. En nuestro caso, con el equipo de visión disponible se alcanzó este equilibrio mediante una apertura con valor $f/16$.

Partiendo de estas restricciones es necesario desarrollar un método capaz de controlar el tiempo de exposición de forma automática y adecuada, que sea capaz de adaptarse a las condiciones lumínicas del día y a los cambios en la escena que se produzcan durante el tiempo que la cámara esté capturando imágenes. Además, el

método debe ser válido para todo tipo de condiciones lumínicas y medioambientales que se producen en los días en los que se realizan los tratamientos, como pueden ser los días nublados o soleados y no debe verse afectado por hechos tales como que una nube intercepte los rayos solares durante un periodo de tiempo y por lo tanto se reduzca la cantidad de luz que recibe la escena momentáneamente. A pesar de estos cambios, por cierto muy frecuentes en la época del tratamiento, el tiempo de exposición se debe ajustar de forma automática para obtener una imagen con la calidad deseada y por consiguiente sin un exceso o defecto de impacto de luminosidad en el sensor, que produciría imágenes saturadas o muy oscuras respectivamente.

Al aumentar la cantidad de luz que incide sobre el sensor de la cámara el valor de intensidad de los tres canales espectrales R, G y B, para todos los píxeles se incrementa correlativamente y si lo que sucede es que disminuye la cantidad de luz, se produce el efecto contrario, los valores de intensidad de los tres canales se reducen. Este es el principio básico elemental que tomamos como premisa para diseñar una estrategia con la que controlar el tiempo de exposición. Este principio formulado de forma cualitativa requiere su materialización en la correspondiente expresión cuantitativa; por tanto se trata de obtener unos valores sobre los que decidir si aumentar o reducir el tiempo de exposición y en qué cantidad ha de hacerse.

En un primer momento se decidió situar un panel que fuese visible desde la cámara. El lugar escogido fue sobre la caja que protege el equipo de energía disponible en el tractor en la parte delantera, sobrepasando el morro del mismo. Este dispositivo de energía cae fuera del objetivo de este estudio y por lo tanto no requiere de mayor explicación. El mencionado panel sirve para estudiar la intensidad de los píxeles mediante la observación del mismo, ya que posee unos colores controlados por nosotros y por tanto con valores teóricos conocidos, que deberían aparecer sobre la imagen respecto de cada color. El panel posee cinco rectángulos de colores, negro, azul, verde, rojo y blanco como se puede ver en la Figura 4-2. La idea consiste en analizar la intensidad en los tres canales espectrales R, G, B de los píxeles contenidos en uno de los cinco rectángulos. Si los valores de dichas intensidades son elevados se debería reducir el tiempo de exposición, ya que la imagen aparece quemada por sobresaturación del sensor, mientras que si el valor es demasiado bajo se debería aumentar el tiempo de exposición, ya que la imagen resulta muy oscura por el efecto contrario. Los rectángulos

son cartulinas sobre las que se pintó lo que se trataba que fueran colores puros, es decir el máximo exponente de cada canal.



Figura 4-2: Panel con rectángulos de distintos colores para controlar el tiempo de exposición.

Se hicieron pruebas con los cinco rectángulos de color y se llegó a la conclusión de que el color más adecuado era el blanco, por ser siempre el que más se aproximaba a su valor teórico y en el que las diferencias de intensidades que se producían eran menos significativas que en el resto. En la Tabla 4-1 se muestran los valores medios que toman los tres canales de los píxeles pertenecientes a cada uno de los cinco rectángulos. Para los cálculos se utilizó un conjunto de 950 imágenes, tomadas durante los meses de mayo, junio y julio de los años 2013 y 2014, mostrándose como ejemplo representativo una de ellas en la Figura 4-2.

Color	Valor teórico	Valor real
Negro	(0,0,0)	(13,16,15)
Azul	(0,0,255)	(24,39,56)
Verde	(0,255,0)	(28,65,48)
Rojo	(255,0,0)	(143,39,20)
Blanco	(255,255,255)	(246,251,199)

Tabla 4-1: Valores medios de los canales RGB en los rectángulos de color del panel.

Una vez se decidió centrarse en el rectángulo de color blanco para controlar el tiempo de exposición se fijó una zona de interés contenida totalmente en dicho rectángulo. También se necesitan dos umbrales por canal, un umbral superior y otro inferior. Con estos dos umbrales se establece el intervalo dentro del cual se permite variar la media de valores de intensidad de los canales dentro del rectángulo sin variar el tiempo de exposición, al considerar que la imagen fue tomada con una iluminación

adecuada. De este modo, para cada canal se tienen los umbrales um_i^x y um_s^x . Los umbrales inferior y superior para un canal x son um_i^x y um_s^x respectivamente, donde $x \in \{R, G, B\}$, identificándose R , G y B con los canales rojo, verde y azul respectivamente. Para establecer estos umbrales hay que estudiar los valores medios de los tres canales en el rectángulo blanco μ_R, μ_G y μ_B con el fin de escoger un valor de tolerancia t , para ello nos fijamos en los valores de la Tabla 4-1. El objetivo de esta tolerancia es que el procedimiento de ajuste del tiempo de exposición esté dotado de una cierta estabilidad con el fin de evitar incrementos y decrementos del tiempo de exposición de forma continua. Este valor de tolerancia es experimental, determinándose empíricamente mediante prueba de ensayo y error que el tiempo de exposición se adapta de forma satisfactoria a la vez que se consigue la estabilidad deseada. Una forma de estimar los umbrales se basa en centrarse en los valores medios y aumentar y disminuir éstos con el valor de la tolerancia, resultando las expresiones siguientes, $um_i^x = \mu_x - t$ para el umbral inferior y $um_s^x = \mu_x + t$ para el superior, en ambos casos para un canal dado x . Tras diversos experimentos se llegó a la conclusión de que el anterior enfoque no resulta satisfactorio, ya que los valores medios tanto para el canal rojo como el verde son muy elevados (246 y 251), de forma que con pequeños valores de tolerancia los valores para los canales terminan por saturarse, sobrepasando el máximo valor de 255 niveles disponibles para representar los valores en cada canal. Por este motivo y tras la realización de diversas pruebas adicionales, se llegó a la conclusión de que una opción apropiada para establecer los umbrales es $um_i^x = \mu_x - t$ para el umbral inferior y $um_s^x = \mu_x$ para el umbral superior para el canal x . Mediante estas dos reglas los umbrales resultantes son $um_i^R = 216$, $um_s^R = 246$, $um_i^G = 221$, $um_s^G = 251$, $um_i^B = 169$ y $um_s^B = 199$.

Este esquema experimental para el cálculo del tiempo de exposición proporcionó resultados satisfactorios, si bien tras la elevada casuística existente en los entornos naturales de operación y tras varios días de pruebas se observó una situación problemática de fallo, que no sucedía en otras ocasiones debido a la posición del tractor respecto de los rayos solares. La situación que se produjo fue que la sombra de la cabina

del tractor se proyectaba sobre el rectángulo blanco utilizado para calcular el tiempo de exposición. El efecto de la proyección de la sombra sobre el panel blanco es que los valores medios de intensidad sobre la imagen disminuían mientras que en el resto de la imagen se mantenían con valores altos, la consecuencia es que los niveles de intensidad de toda la imagen se incrementaban al aumentar el tiempo de exposición para que los valores medios del panel blanco se mantuvieran de nuevo entre los umbrales. Es decir, se aumentaba el tiempo de exposición sin ser necesario. El resultado es que se capturaban imágenes con un valor muy elevado de intensidad, con las que después se hacía una incorrecta detección de la vegetación por exceso de iluminación.

Ahondando en el problema se ve claramente que el error se encuentra en intentar calcular el tiempo de exposición centrándose en una zona (panel blanco) que no es de interés desde el punto de vista de la aplicación agrícola y sobre la que no se realiza ningún procesamiento posterior. En efecto, la zona de interés sobre la que se aplican los algoritmos es exactamente la zona de cultivo, por tanto, no tiene sentido fijarse en una zona distinta para una operación de calibración, que afectará a toda la imagen.

Para solucionar este problema se pensó en analizar la zona de cultivo para comprobar si se encontraba algún tipo de patrón, o intervalo uniforme y no muy extenso en el que poder situar los valores medios de los tres canales, en cuyo caso se solucionaría el problema de las sombras, ya que aunque se proyecte la sombra del tractor o de cualquier otro elemento cercano a la zona de estudio (árboles, poste de la luz, edificios, etc.), éste no influiría en el resultado, debido a que el tiempo de exposición se adapta para una correcta visualización de la zona de estudio y no como en el caso anterior, que se adaptaba el tiempo de exposición para la correcta visualización del panel blanco.

Para obtener estos nuevos umbrales se deben capturar imágenes con un valor del tiempo de exposición adecuado, determinado por el usuario manualmente, de modo que la imagen fuese válida para su posterior procesamiento. Previamente se realizó este ajuste manual del tiempo de exposición para obtener imágenes similares a las que se quieren capturar con el método de exposición automático. Con estas imágenes se calcula la media de los valores de intensidad de la zona de interés para los tres canales. A diferencia de lo que ocurría antes, que independientemente del estado de crecimiento del maíz y las malas hierbas el rectángulo blanco tenía siempre la misma información,

ahora podría darse el caso de que la media de valores de los tres canales dependa fuertemente del estado de crecimiento de las plantas, por lo que es recomendable utilizar imágenes donde dicho estado de crecimiento se adecúe lo máximo posible al momento de su tratamiento. Se analizaron cinco conjuntos de imágenes con diferentes estados de crecimiento, las imágenes de cada conjunto se capturaron en días diferentes, espaciándose cada captura en el tiempo al menos una semana. Como se ha indicado previamente las imágenes fueron tomadas durante los meses de mayo, junio y julio de los años 2013 y 2014. En promedio, cada conjunto constaba de 1500 imágenes. En el mismo día se tomaron imágenes bajo diferentes condiciones de luminosidad, en diferentes zonas del campo de cultivo y con distintas orientaciones respecto de la dirección de iluminación solar. A pesar de la variabilidad que cada conjunto pueda presentar, debido a las diferentes condiciones de captura en el tiempo y en el estado de crecimiento de la vegetación (maíz y malas hierbas), todas las imágenes deben ser válidas para realizar la posterior segmentación de la vegetación mediante el ajuste del tiempo de exposición. La zona de estudio o ROI que se ha definido sobre la imagen comprende la zona que transcurre desde la columna 400 a la columna 1950 y de la fila 500 a la fila 1180. Sobre el terreno dicho rectángulo corresponde a una zona del cultivo en la que el lado inferior del rectángulo se sitúa a 3m por delante del eje vertical que pasa por el centro de la cámara con un ancho de 3m (cuatro líneas de cultivo separadas 75 cm entre sí). La altura del rectángulo son 2m. En la Figura 4-3(a) y (b) se muestran dos imágenes representativas del conjunto de imágenes disponibles, tomadas durante dos días distintos, con la ROI marcada sobre las imágenes con un rectángulo rojo. La Figura 4-3(c) y (d) muestra la zona de interés sobre el terreno mediante la proyección de perspectiva. El cuadrado naranja representa el plano ortogonal a la superficie del terreno, situado sobre un hipotético eje vertical que sostiene la cámara. La intersección de este plano con el plano del suelo define la línea recta a partir de la cual, situada a 3 metros y en perpendicular, comienza la ROI. Obsérvese en las imágenes mostradas el distinto grado de crecimiento de la vegetación, que resulta un factor determinante para el ajuste del tiempo de exposición.

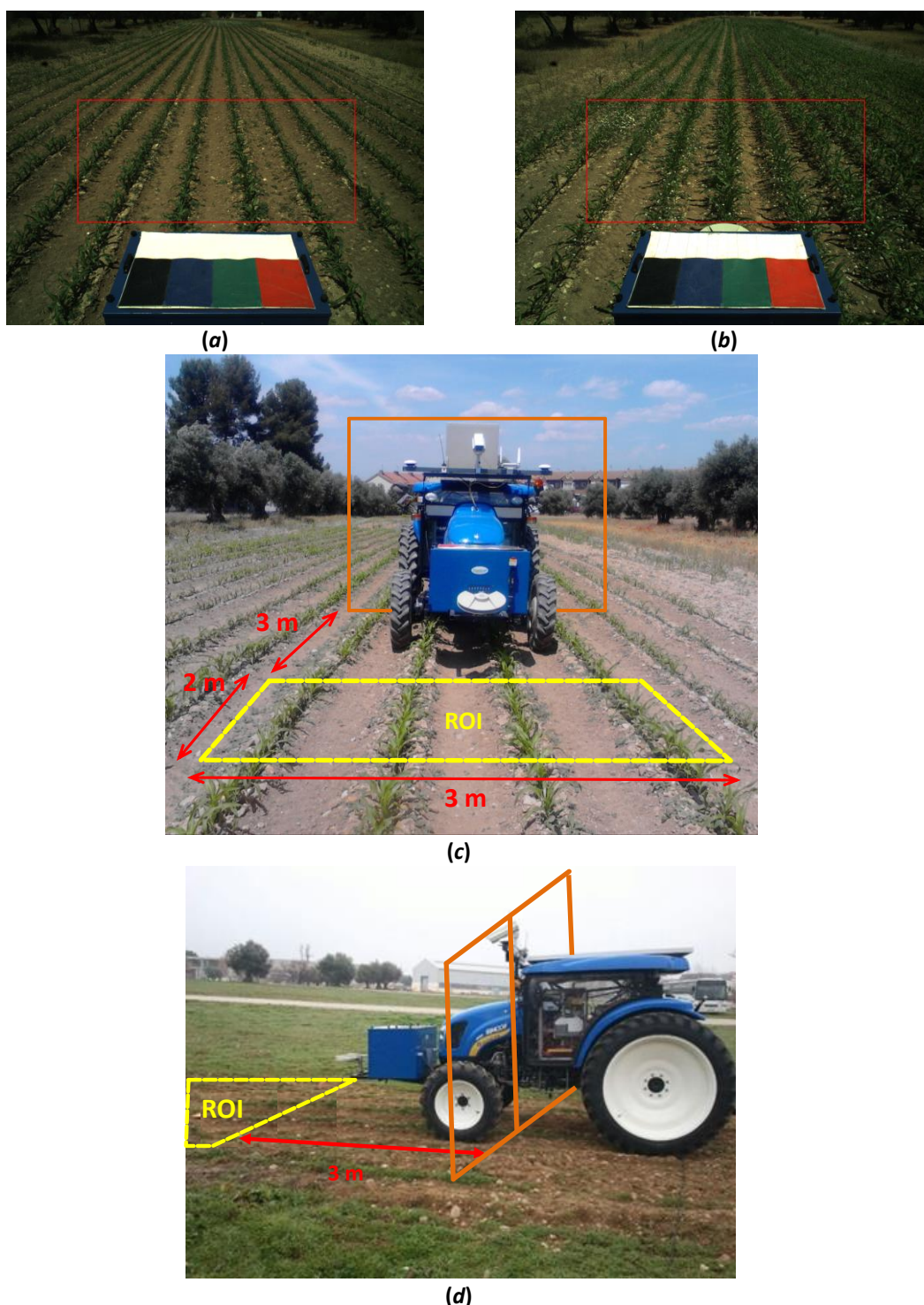


Figura 4-3. Cultivo en distintos estados de crecimiento y mostrando ROI: (a) y (b) ROI representada con un rectángulo rojo; (c) y (d) ROI representada mediante proyección de perspectiva.

Para cada imagen analizada del conjunto disponible, se obtuvieron los valores medios de intensidad en los tres canales R, G y B correspondientes a la ROI. Para cada uno de los cinco días de estudio se calculó el promedio de los valores medios de cada imagen, a partir de los cuales se obtuvieron el máximo y mínimo valor para los tres canales. En las tablas Tabla 4-2 a Tabla 4-6 se muestran los valores obtenidos para cada uno de los cinco días. En las columnas “Valores medios”, “Media máxima” y “Media

mínima” aparece respectivamente el valor medio, máximo y mínimo relativos a las medias de los canales para las imágenes capturadas en ese mismo día.

Canal	Valores medios	Media máxima	Media mínima
R	79	86	23
G	68	79	21
B	39	46	12

Tabla 4-2: Media de valores para el primer día

Canal	Valores medios	Media máxima	Media mínima
R	59	71	48
G	58	66	49
B	33	38	27

Tabla 4-3: Media de valores para el segundo día

Canal	Valores medios	Media máxima	Media mínima
R	61	74	50
G	61	72	50
B	34	40	28

Tabla 4-4: Media de valores para el tercer día

Canal	Valores medios	Media máxima	Media mínima
R	81	84	81
G	76	78	75
B	45	46	44

Tabla 4-5: Media de valores para el cuarto día

Canal	Valores medios	Media máxima	Media mínima
R	44	51	33
G	40	46	31
B	24	27	18

Tabla 4-6: Media de valores para el quinto día

A partir de la información reflejada en las tablas se pueden realizar las siguientes observaciones. Tanto el canal rojo como el verde toman valores semejantes de media y valores pico (máximo y mínimo). El canal azul toma valores menores con respecto a los otros dos canales, tanto en la media como en el rango de los valores pico. Estas apreciaciones, junto con el hecho de que la vegetación tiene como componente predominante el canal verde y el suelo el canal rojo, justifican el intento de usar los canales verde y rojo para ajustar el tiempo de exposición, pudiendo prescindir del canal azul. Por otro lado, como resulta que tanto los píxeles del canal verde como los del rojo toman valores próximos entre sí, se puede simplificar el número de umbrales de cuatro a dos agrupando ambos canales de forma que estos dos canales puedan compartir los mismos umbrales.

Tras el análisis realizado a partir de los datos reflejados en las tablas anteriores y mediante numerosas pruebas se llega a la conclusión de que los valores de umbral 40 y 85, constituyen dos valores de referencia para el valor mínimo y máximo respectivamente. De este modo, cuando el valor medio del canal rojo y verde en la región de interés sea mayor que 85, la siguiente imagen se captura con un tiempo de exposición menor respecto al tiempo con el que se tomó la anterior. Si por el contrario, el valor medio del canal rojo y verde en la región de interés es menor que 40, la próxima imagen se captura con un tiempo de exposición mayor respecto al tiempo de captura de la última imagen.

Tal y como se ha mencionado previamente, la ROI en la imagen está definida entre las columnas 400 a 1950 y entre las filas 500 a 1180. Con el fin de reducir el tiempo de cómputo del proceso, sin que se vea afectada la calidad final de la imagen por un incorrecto ajuste del tiempo de exposición, se realiza un submuestreo de la ROI. El paso de este submuestreo es de cinco, lo que significa que de cada cinco píxeles se contabiliza únicamente uno para el cálculo de las medias de los valores para los canales verde y rojo. Este submuestreo se efectúa tanto en las filas como en las columnas de las imágenes. En relación al submuestreo se ha aplicado la transformada de wavelets hasta el nivel de descomposición cinco, como en el caso anterior, con las mismas familias que las utilizadas en (Pajares y Cruz, 2004), por su buen comportamiento, aunque en otro contexto. El objetivo era verificar si la influencia del filtrado paso bajo, que da origen a los coeficientes de aproximación mejoraba los resultados. Las mejoras obtenidas no son significativas, mientras que por el contrario se incrementa el tiempo computacional al tener que realizar las operaciones de filtrado inherentes a esta transformada, razón por la cual no se aplica dicha transformada.

En Montalvo y col. (2013b) se planteó otro método para la adquisición de imágenes en entornos agrícolas de exterior con la suficiente calidad mediante el control del tiempo de exposición y modificación del histograma. Este método captura una imagen y comprueba mediante el análisis del histograma y los parámetros estadísticos de la media y asimetría estadística (*skewness*) si ha sido adquirida con un tiempo de exposición adecuado, pudiendo ocurrir dos situaciones distintas, bien que se tenga que capturar una nueva imagen antes de realizar el procesado porque el tiempo de exposición no fuera adecuado o bien que el tiempo de exposición fuese adecuado y se continúe con el procesamiento de la imagen.

En la Figura 4-4 se muestra el esquema general propuesto para capturar una imagen de calidad en función del tiempo de exposición adecuado. El primer paso consiste en capturar una imagen en crudo mediante la cámara con el último tiempo de exposición utilizado en la captura de la imagen anterior. En caso de que se trate de la primera captura se realiza con un tiempo de exposición aleatorio. El proceso que captura la imagen se denomina en el esquema “Captura de imagen en crudo”. Tras la captura se obtiene una imagen en formato RGB. Esta imagen, junto con el tiempo de exposición, constituyen las entradas al método “Cálculo del tiempo de exposición” que ajusta el tiempo de exposición cuando es necesario. Este método proporciona el nuevo valor del tiempo de exposición que se utilizará en la captura de la siguiente imagen. Esto se refleja en el diagrama mediante la flecha que entra de nuevo en el proceso “Captura de imagen en crudo”. La fase denominada como “Obtención de una imagen de calidad” proporciona como salida la imagen RGB capturada, la cual constituye la entrada a la fase siguiente “Segmentación de la imagen”.

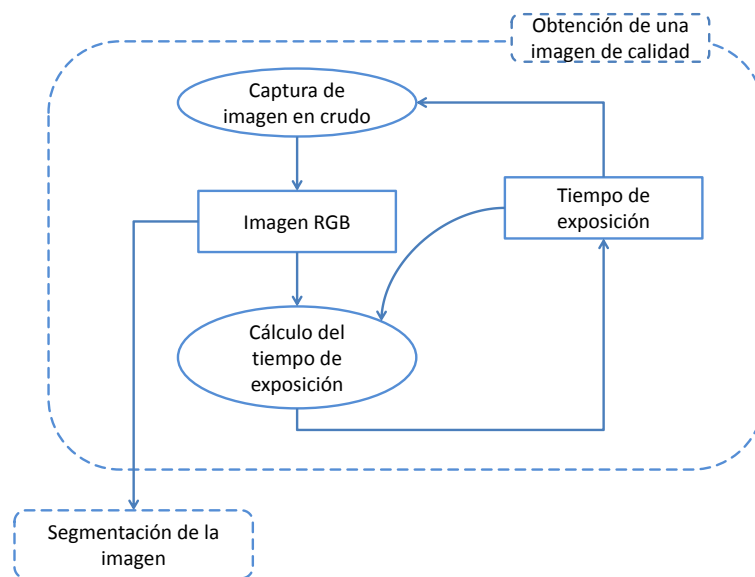


Figura 4-4. Esquema de la fase de captura de la imagen

4.2.2. Segmentación de la imagen

Partiendo de una imagen válida tras la fase anterior, el objetivo siguiente se orienta hacia la extracción de la vegetación. Se trata de identificar los píxeles de la imagen que se corresponden con las plantas (malas hierbas y cultivo) verdes existentes en la imagen sobre una nueva ROI. Debido a que existen distintos tipos de verde, desde el punto de vista computacional, es necesario definir un criterio para su extracción. El

verde ideal teórico sería el que tuviera un valor de 255 en el canal verde y un valor de 0 tanto en el canal rojo como en el azul. No obstante, existe un amplio abanico de colores con distintas tonalidades de verde, así como azules-verdosos que se podrían situar a medio camino entre el azul y el verde. Las plantas tampoco poseen un verde ideal teórico y tampoco resulta uniforme a lo largo de toda la planta, hojas y tallo, sino que varía tanto por la tonalidad de la propia planta como por los reflejos producidos por la incidencia de los rayos solares dependiendo de la dirección de éstos. Por tanto se necesita un método que clasifique los píxeles de la imagen según que su color se aproxime más o menos al verde.

Para esta clasificación se utilizan los índices cromáticos explicados en la sección 3.3. Con estos índices se dota a cada píxel de un valor numérico en función de los valores de sus tres canales, convenientemente combinados, con la finalidad de cuantificar su verdor. Después de analizar varios índices se optó por uno de los más sencillos de calcular desde el punto de vista de su bajo coste computacional de cara al procesamiento en tiempo real, éste es el exceso de verde ExG, ecuación (3.5).

Antes de aplicar el ExG es necesario realizar una serie de consideraciones y aplicar algunas restricciones para mejorar la identificación de los verdes, que constituye el objetivo perseguido.

En las imágenes agrícolas, es habitual la existencia de sombras, particularmente en días soleados, ya que son producidas por las propias plantas, siendo tanto mayores cuanto más crecida está la vegetación. En días nublados no existen sombras como tal, pero sí zonas de mayor oscuridad por las mismas razones. Resulta de fácil observación el hecho de que en esas zonas los píxeles poseen unos valores muy bajos de intensidad en los tres canales. Estos píxeles son complicados de clasificar incluso para un experto, ya que no se aprecia con claridad si se trata de plantas o no. Las hojas de las plantas de maíz cuando éste está crecido, producen sombras significativas. Para un experto resulta complicado discernir a qué clase pertenecen los píxeles afectados por la sombra, pudiendo llegar a la conclusión de que pertenecen a la clase de las plantas o no basándose en la experiencia y una serie de reglas como puede ser intuir que se trata de la prolongación del tallo u hoja de una planta. Debido a la inmensa casuística existente en este sentido, así como lo caprichoso de la aparición de las sombras por tratarse de plantas con distribuciones irregulares, no se dispone del tiempo

de proceso suficiente para realizar este tipo de comprobaciones basadas en reglas, tal y como podría realizar un sistema experto. Por este motivo, se llega a la conclusión de que una opción práctica consiste en descartar estos píxeles puesto que por el propio hecho de que se produzca sombra sobre ellos suelen ser píxeles que en la escena se sitúan bajo la cobertura vegetal, siendo generalmente suelo. Con tal propósito, se realiza la suma de los valores de intensidad de los tres canales para cada uno de los píxeles de la imagen, los píxeles cuya suma no sobrepase un cierto valor se descartan, siendo considerados desde este momento como no vegetación. Analizando las imágenes tomadas durante las pruebas experimentales y los posteriores resultados mediante diferentes pruebas de ensayo y error se ha establecido 40 como valor apropiado para el sistema de visión disponible, así los píxeles cuya suma de los canales rojo, verde y azul no sobrepase dicho valor son considerados como terreno.

Al igual que se producen sombras, también se suelen producir reflejos por la incidencia de los rayos solares. Estos reflejos hacen que los píxeles afectados muestren valores muy elevados de intensidad y llegan a confundir al ojo de un experto, no sabiendo discernir a qué clase pertenece ese píxel. Si bien es cierto que los reflejos se suelen producir en menor proporción que las sombras, cuando el maíz está algo crecido, también es deseable eliminar estos píxeles con reflejos y tratarlos como píxeles que no son vegetación. En esta ocasión se eliminan aquellos píxeles cuya suma de los valores de los tres canales sobrepase un valor, establecido en 400 para nuestras imágenes.

Realizados los filtrados previos, en este momento estamos en disposición de aplicar los índices de vegetación para asignar un valor de verde a cada píxel. En la sección 3.3.1 se explicaron algunos VI simples y se comentaron algunas operaciones de normalización habituales cuando se usan en aplicaciones agrícolas, tal y como se recoge en la literatura. Estas operaciones tienen como finalidad trasladar los valores que toman los canales rojo, verde y azul del intervalo $[0,255]$ al intervalo $[0,1]$ y además que la suma de las tres componentes sea igual a la unidad.

Como se comentó previamente, el índice de vegetación seleccionado fue el exceso de verde, que además de su buen comportamiento cualitativo resulta eficaz desde el punto de vista computacional. Debido a que la aplicación final requiere el procesamiento en tiempo real se trata de reducir al máximo el número de operaciones. Las operaciones de normalización constan de sumas y divisiones, siendo las divisiones

altamente costosas. Por tal motivo, una vez conocido el índice que se va a utilizar se opera y simplifican todas las operaciones que engloban las normalizaciones y el índice de vegetación de modo que se consigue reducir el número de operaciones.

Tras aplicar el exceso de verde se obtiene una imagen de la misma dimensión que la imagen original cuyos valores se sitúan en el intervalo $[-1, 2]$. Es necesario transformar esta imagen en una imagen en escala de grises con los valores de los píxeles en el intervalo $[0, 255]$ ya que la siguiente etapa de umbralización requiere los datos en este rango. Para realizar esta operación se realiza la siguiente transformación lineal:

$$z = \frac{x - \min}{\max - \min} (MAX - MIN) = \frac{x + 1}{3} \times 255 = (x + 1) * 85 \quad (4.1)$$

donde x representa el valor del píxel antes de la transformación y z tras su transformación a la escala de grises con valores en el rango 0 a 255, $\min = -1$, $\max = 2$, $MIN = 0$ y $MAX = 255$. La imagen formada con los valores obtenidos de la ecuación (4.1) es la que se utiliza en la fase de umbralización.

A modo de resumen, en la Figura 4-5 se muestra el esquema general de esta fase. La entrada es la imagen capturada por la cámara en el modelo de color RGB, “Imagen RGB”. A partir de esta imagen se realiza el proceso de “Segmentación” en el cual se computa el índice vegetativo exceso de verde y se aplica la transformación a una imagen en escala de grises cuyos píxeles se encuentran en el rango $[0, 255]$, la imagen resultante es la denominada “Imagen en escala de grises”, que será la suministrada a la fase de “Umbralización”.

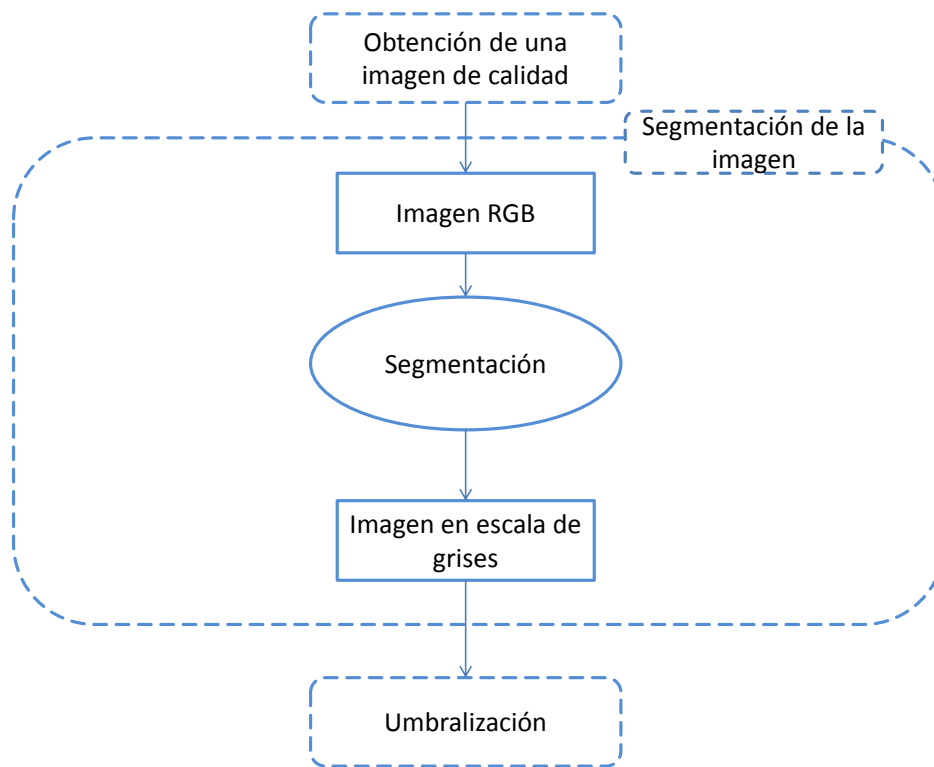


Figura 4-5. Esquema para la fase de segmentación.

4.2.3. Doble umbralización mediante el método de Otsu

En este apartado se describe la aportación realizada para la umbralización a partir de imágenes en escala de grises usando doblemente el método de Otsu en dos etapas diferentes. Uno de los objetivos que se plantean en el diseño del SEA es la discriminación de malas hierbas frente al cultivo de maíz incluso en campos con alta densidad de aquellas, que como se indicó previamente y hasta donde hemos podido conocer, este problema no se ha tratado en la literatura.

En la sección 3.4.2 se explicó de forma detallada el método de umbralización propuesto por Otsu (1979). Ahora se trata de explicar cómo se ha utilizado y adaptado en nuestro caso para determinar qué píxeles de la imagen pertenecen a elementos del cultivo, cuáles a malas hierbas y cuáles al resto de elementos presentes en la imagen, estos últimos carecen de relevancia desde el punto de vista de la aplicación agrícola, son partes tales como el cielo o construcciones que aparecen en el horizonte de la imagen.

En las imágenes en escala de grises, obtenidas tras el proceso de segmentación, se pueden diferenciar distintos niveles de intensidad, de forma que a mayor intensidad

mayor verdor de la imagen en ese píxel. Esto significa que las plantas, tanto las malas hierbas como el cultivo, aparecerán como puntos cercanos al color blanco y el resto (tierra, piedras, sombras, pajas y otros materiales) aparecerá como zonas oscuras o casi negras. Esta correspondencia en la escala de grises se debe a la propia transformación realizada por el índice de vegetación. De esta forma se requiere de un método automático para detectar el mejor valor que separe las plantas del resto. Esto se convierte en un problema de umbralización. En el transcurso de la investigación se han analizado distintas estrategias de umbralización existentes en la literatura a partir de una imagen de grises (Sezgin y Sankur, 2004; Rosin y Loannidis, 2003; Kapur y col., 1985). De entre todos ellos, el método adaptativo de Otsu (1979), que maximiza la varianza interclase de los píxeles pertenecientes a las plantas y al fondo, produce unos resultados más que aceptables en las imágenes agrícolas. De hecho ha sido aplicado en diferentes estudios relacionados con la detección del cultivo y de las malas hierbas (Burgos-Artizzu y col., 2011; Jiang y Zhao, 2010; Meyer y Neto, 2008; Wei y Shangfeng, 2006; Guijarro y col., 2011; Romeo y col., 2013; Guerrero y col., 2013).

Tal y como se explicó en la sección 3.4.2, este método asume que los histogramas de las imágenes son bimodales, representando dos distribuciones de intensidad gaussianas, una de las cuales está asociada a las plantas y la otra al resto de elementos de la imagen. Si bien, en general, las distribuciones gaussianas son teóricas sin que lleguen a cumplirse totalmente en la práctica, la suposición es válida con resultados satisfactorios. Como se muestra en la Figura 4-6(c), el histograma de una de las imágenes tomadas en el campo y ya segmentada, figura 4.6(a), no presenta con claridad dos distribuciones gaussianas teóricas. Para dicho caso el algoritmo de Otsu establece el umbral de corte en 152. La Figura 4-6(a) muestra el recorte de una imagen segmentada, se ha realizado el recorte porque se quiere tener sólo en cuenta los valores de los píxeles de la ROI, de otra forma el histograma quedaría influenciado por otros píxeles y el método de Otsu proporcionaría un valor diferente. La Figura 4-6(b) muestra la binarización de la imagen según el valor umbral 152, siendo blancos o negros los píxeles cuyos valores son mayores que dicho umbral en el primer caso y menores o iguales en el segundo.

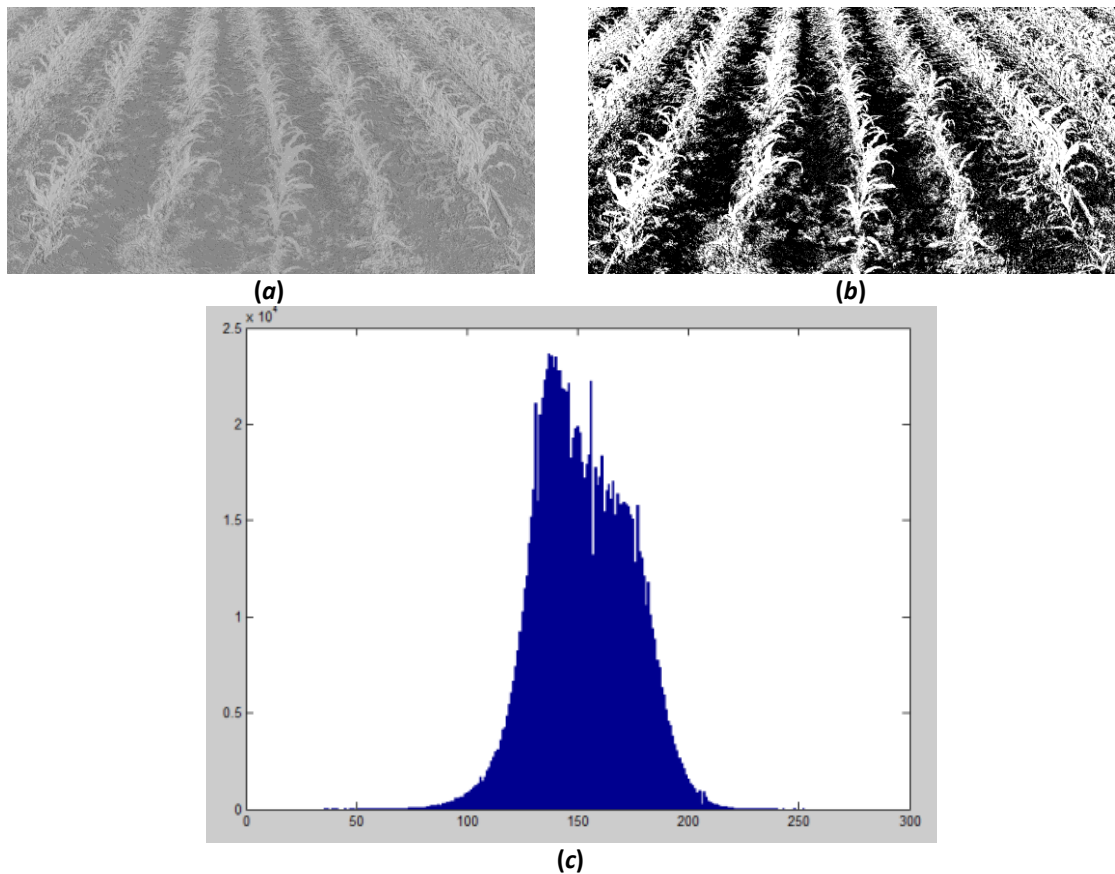


Figura 4-6. (a) Imagen segmentada; (b) Imagen binarizada mediante el valor 152; (c) Histograma de la imagen segmentada.

Considerando sólo la ROI se evitan distorsiones del histograma debidas a otros elementos presentes en la imagen tales como el panel en la parte frontal del tractor, el cielo u otras estructuras o edificaciones.

Tras esta primera umbralización, los píxeles que pertenecen a las plantas verdes se encontrarán separados del resto de píxeles en la imagen, como se puede observar en la Figura 4-6(b). Ahora, debido a que las plantas incluyen tanto a las malas hierbas como al cultivo, podemos asumir que incluso si tienen componentes espectrales similares es posible aún distinguir alguna pequeña variación espectral, que podría permitir la separación de las plantas en diferentes tipos de verdes, conduciendo finalmente a la separación entre las plantas de maíz propiamente dichas y las malas hierbas. Una segunda umbralización con Otsu permite tal distinción. Esta segunda umbralización se aplica únicamente sobre los valores de los píxeles pertenecientes a la imagen en escala de grises que han sido ya identificados como plantas verdes en el primer proceso de umbralización. En definitiva, con estos píxeles se genera un nuevo histograma que es considerado también bimodal. En este segundo caso las dos distribuciones gaussianas obedecen a la existencia de malas hierbas y de maíz. Gracias a

la suposición de existencia de las dos nuevas distribuciones gaussianas se puede aplicar nuevamente el método de umbralización de Otsu, obteniendo un segundo umbral con el objetivo de discernir los dos tipos de plantas inicialmente separadas, cultivo y malas hierbas.

Mediante estos dos umbrales se pueden clasificar los píxeles de la imagen en tres categorías: maíz (cultivo), malas hierbas y resto de la imagen. Sean el primer y segundo umbral obtenidos t_1 y t_2 respectivamente; $g(i, j)$ el valor de intensidad de la imagen de grises en la localización de píxel (i, j) obtenido mediante el índice de vegetación donde i y j se corresponden con la fila y columna del píxel $p(i, j)$ en la imagen, la regla de clasificación se expresa de la siguiente forma:

$$g(i, j) = \begin{cases} g(i, j) < t_1 & \rightarrow p(i, j) \in F \\ t_1 < g(i, j) < t_2 & \rightarrow p(i, j) \in M \\ g(i, j) > t_2 & \rightarrow p(i, j) \in C \end{cases} \quad (4.2)$$

siendo C el conjunto de los píxeles pertenecientes al cultivo, M el que identifica las malas hierbas y F el que engloba al resto de los píxeles que no forman parte de la vegetación.

Esta estrategia de realizar una doble umbralización constituye el paso fundamental del enfoque propuesto en esta estrategia. La Figura 4-7(a) muestra un ejemplo de una imagen binarizada tras aplicar la primera umbralización, es decir resaltando en verde la vegetación detectada; en la Figura 4-7(b) se muestra la imagen binarizada aplicando el segundo umbral, es decir resaltando solamente lo que se considera cultivo; por último, en la Figura 4-7(c) se han resaltado en la misma imagen tanto las malas hierbas como el cultivo, generando una fusión de ambas imágenes, donde el cultivo se muestra en color verde y las malas hierbas en rojo.

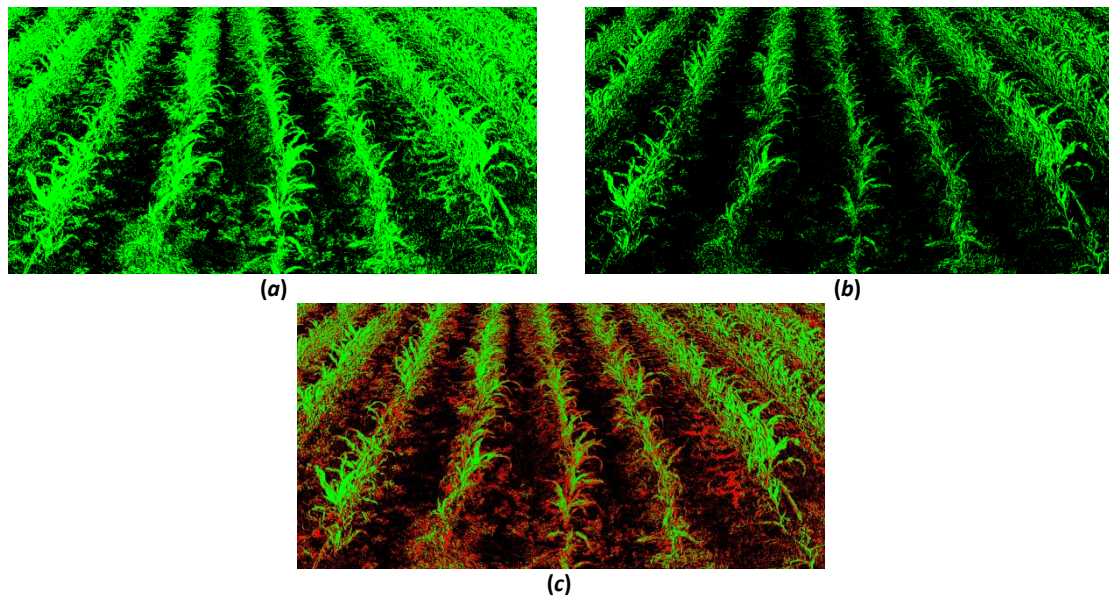


Figura 4-7. (a) Imagen binaria mostrando la vegetación; (b) Imagen binaria mostrando solo el cultivo; (c) Imagen mostrando en verde el cultivo y en rojo las malas hierbas.

En la Figura 4-8 se muestra el diagrama completo correspondiente a la fase de umbralización que permite visualizar el proceso en su totalidad. Procedente de la fase anterior (Segmentación de la imagen) se recibe la “Imagen en escala de grises” sobre la que se aplica un primer proceso de umbralización mediante el método de Otsu, obteniendo un primer umbral, “Umbral verdes”, que permite diferenciar la vegetación del resto de la imagen. Con este primer umbral se obtienen los píxeles considerados como vegetación (malas hierbas y cultivo). A continuación, con los píxeles de la “Imagen en escala de grises” considerados vegetación se realiza una segunda umbralización, también con el método de Otsu, obteniendo un segundo umbral, “Umbral cultivo”, mediante el cual se puede diferenciar entre el cultivo y las malas hierbas. Con los dos umbrales obtenidos se realizan dos binarizaciones: la primera genera la “Imagen binaria cultivo”, que muestra únicamente los píxeles considerados cultivo, cuyos valores de intensidad son mayores que el segundo umbral; la segunda binarización genera la “Imagen binaria malas hierbas”, que muestra únicamente los píxeles considerados malas hierbas, cuyos valores se encuentran entre los dos umbrales calculados.

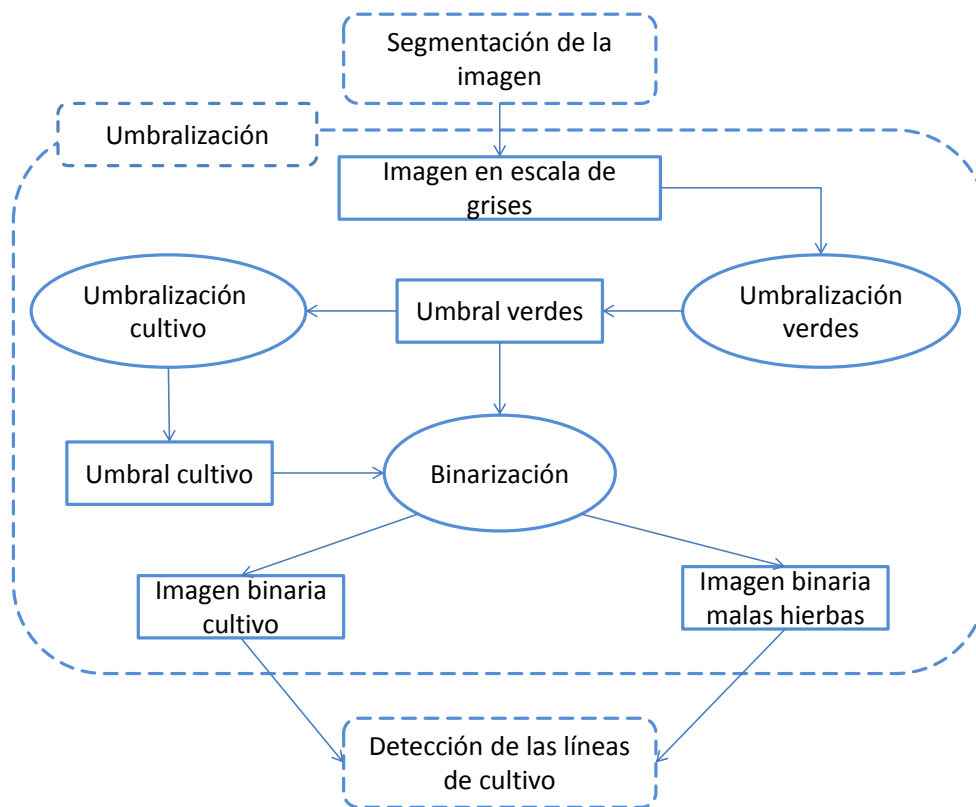


Figura 4-8. Esquema para la fase de umbralización

4.2.4. Detección de las líneas de cultivo

Tras la segunda umbralización el siguiente objetivo es asignar la correspondiente ecuación de línea recta a cada una de las líneas de cultivo. Se trata, en definitiva, de identificar o detectar las líneas de cultivo. Con tal propósito recurrimos a los métodos basados en regresión lineal, donde los parámetros de la ecuación de la línea recta se estiman a partir de los píxeles considerados como parte de las líneas del cultivo. En este paso, el problema es determinar qué píxeles participan en la estimación de los parámetros para cada una de las líneas.

Una vez resaltado el cultivo existente en la imagen en forma de imagen binaria, donde los píxeles que forman parte de las plantas de maíz tienen un valor de 255 y el resto de 0, se requiere el análisis de esta información para detectar las líneas establecidas por la distribución de las plantas de maíz.

Debido a que existen varias posibles líneas en cada una de las imágenes se deben imponer ciertas restricciones en la búsqueda de estas líneas. Existen varias razones que justifican tal imposición:

- a) Pueden existir líneas rectas en la imagen, algunas de las cuales no son realmente líneas de cultivo. Una de las razones por las que se puede identificar una línea recta en la imagen que no coincida con ninguna línea de cultivo se debe a que el proceso de binarización no haya resultado lo suficientemente efectivo, con alineaciones de píxeles que no coinciden precisamente con las líneas de cultivo.
- b) También se pueden detectar varias líneas rectas muy próximas entre sí, situándose sobre la misma línea de cultivo. En la práctica este caso es altamente probable ya que las líneas de cultivo siguen alineaciones aproximadamente rectas con una cierta anchura, siendo además no uniformes, pudiendo aparecer casos extremos tales como desaparecer en algunas zonas o llegar a unirse con otra línea diferente, que se consideran seguidamente.
- c) Las líneas de cultivo pueden llegar a desaparecer en algunas zonas por varios motivos. Es posible que la siembra no se haya realizado correctamente en esas zonas o no haya crecido el cultivo por diversas razones. En este sentido cabe mencionar la desaparición de las semillas por habérselas comido las aves o roedores, las malas hierbas hayan dificultado su crecimiento o las condiciones de humedad y temperatura hayan sido adversas.
- d) El caso de la fusión de dos o más líneas puede aparecer por motivos diversos, achacables tanto a cuestiones derivadas del sistema de visión y el procesamiento de imágenes, como a razones de índole puramente agrícolas, tales como: a) la proyección de perspectiva geométrica del sistema de visión con el que se adquiere la imagen; b) el cultivo se ha desarrollado rápidamente y su estado es tal que las plantas de maíz presentan un estado de crecimiento avanzado, pudiendo entrar en contacto hojas de plantas de distintas líneas de cultivo; c) la existencia de malas hierbas en los surcos que se hayan identificado como plantas de maíz o que durante el proceso de binarización se haya introducido ruido en forma de píxeles de maíz.

Para solventar estos problemas, debemos hacer uso del conocimiento subyacente sobre la disposición del cultivo en el campo de cultivo de maíz, aprovechando a la vez las restricciones geométricas del sistema de visión. El objetivo es diseñar un método lo suficientemente robusto capaz de dar solución a esta problemática. Con tal propósito se realizan las siguientes consideraciones que han de tenerse en cuenta durante el proceso y que se supone su conocimiento:

- 1) El número de líneas de cultivo.
- 2) La localización teórica esperada de cada una de las líneas en las imágenes.
- 3) El área a explorar en la imagen, limitada por sus valores máximos y mínimos en las coordenadas de la imagen, tanto en el eje horizontal como en el vertical (filas y columnas de imagen).

Para evitar que se detecten líneas donde la probabilidad real de encontrarlas es mínima, en base a las consideraciones anteriores, se proponen las dos alternativas siguientes: a) hacer uso de una plantilla genérica que delimite la búsqueda; b) utilizar la información proporcionada por una unidad de medida inercial (IMU, Inertial Measurement Unit). A continuación se esbozan ambos enfoques. En Montalvo y col. (2011) se realizó un primer estudio sobre la geometría del sistema de visión, gracias al cual se determinó la posibilidad de utilizar la IMU, y que a su vez sirvió de base para la generación de la plantilla.

a) Plantilla

Mediante el uso de una plantilla se delimitan las regiones donde se deben situar las líneas de cultivo, permitiendo delimitar la región donde se han de buscar dichas líneas para su localización. Gracias a esta plantilla no se situarán líneas rectas en lugares donde su ubicación resulte altamente improbable, restringiendo a la vez la posibilidad de situar dos líneas muy próximas entre sí, que en realidad representarían la misma línea de cultivo en la realidad.

Una plantilla no es más que una imagen con las mismas dimensiones que la imagen original que se está procesando. Dicha imagen es binaria, al igual que la imagen de la que se deben obtener las líneas que representan el cultivo. Los píxeles blancos (con valor 255) de la plantilla indican las zonas donde puede existir una línea de cultivo, mientras que los negros (con valor 0) indican zonas prohibidas. La plantilla dispone de

tantas zonas blancas como surcos se quieran detectar. Estas zonas sobre el campo de cultivo tienen una forma rectangular, incluyendo cada una de ellas el ancho del cultivo con un cierto margen de tolerancia para contemplar la posibilidad de desviaciones debidas al movimiento del tractor, que indudablemente afecta a la captura de la imagen. La altura de tales rectángulos viene determinada por el alcance de la zona de aplicación del tratamiento, a fijar en todo momento por el usuario, en el caso que nos ocupa se trata de la ROI. La proyección de perspectiva de tales rectángulos sobre el plano de la imagen hace que éstos aparezcan como trapecios escalenos, donde los lados paralelos representan sus bases y equivalen a la zona más cercana y más alejada de la región que se quiere tratar en la imagen. La Figura 4-9(a) muestra una imagen binarizada donde las líneas de cultivo de maíz se representan con píxeles blancos y en la que el objetivo consiste en la detección de los tres surcos centrales. Por tal motivo, la plantilla se diseña con tres zonas blancas como la mostrada en la Figura 4-9(b). En cada una de las tres regiones se buscan las tres líneas de cultivo deseadas.



Figura 4-9. (a) Imagen binarizada con solo cultivo; (b) plantilla utilizada para la detección de líneas de cultivo.

De este modo, para estimar la localización de una línea de cultivo, solamente se tienen en consideración los píxeles correspondientes a plantas de cultivo que caigan dentro de la respectiva zona blanca de la plantilla. Con tal propósito, superponemos la imagen binaria y la plantilla. En consecuencia, para determinar las posiciones de una línea sólo se considerarán los píxeles blancos de la imagen binaria que se sitúen sobre la correspondiente zona blanca de la plantilla.

Desde el punto de vista computacional este proceso equivale a realizar una operación lógica “AND” con sendas imágenes binarias; esto es, la correspondiente a la

que contiene las plantas y la propia plantilla. Tras dicha operación los píxeles blancos resultantes son los que se utilizan para el ajuste de la recta a la línea de cultivo.

La plantilla se genera de forma manual, a partir de una serie de imágenes capturadas previamente por el sistema de visión. El objetivo es tratar de incluir el mayor número posible de posiciones de las líneas dentro de las zonas blancas de la plantilla. La plantilla, así creada, sirve para todo campo de cultivo que se haya sembrado con el mismo espaciado entre surcos y en terrenos de similares características. Además como es lógico, el sistema de visión debe poseer la misma configuración en relación a sus parámetros intrínsecos (distancia focal, tamaño y resolución del sensor) y extrínsecos (geometría, posición y orientación). Conocemos que el vehículo, equipado con el sistema de visión, navega sobre un terreno real que presenta irregularidades y desigualdades. Esto produce vibraciones y balanceos principalmente en los ángulos de cabeceo y alabeo. El ángulo de guiñada se asume que es correcto, ya que de otra forma el tractor estaría navegando de forma errónea y esto derivaría en un problema de control del guiado, que cae fuera del ámbito de la detección de líneas de cultivo propiamente dichas. Basándonos en este hecho y tras la captura de tres secuencias de video consistentes en 1400 imágenes cada una, con el vehículo navegando a 4 km/h, se han determinado los máximos desplazamientos producidos tanto hacia la izquierda como hacia la derecha alrededor de las líneas de cultivo esperadas en la imagen. A partir del estudio de estos tres vídeos, se construye la plantilla o máscara para determinar los límites alcanzables por los píxeles pertenecientes a las líneas de cultivo.

b) IMU

Otro modo para determinar los píxeles binarios con las mayores posibilidades de ser o no candidatos para la detección de las líneas de cultivo es hacer uso de la información proporcionada por una IMU. Una IMU es un dispositivo electrónico que realiza mediciones sobre la velocidad, orientación y fuerzas gravitacionales sobre el sistema en el que se encuentra situada, para ello emplea acelerómetros y giróscopos, pudiéndose utilizar en algunas ocasiones magnetómetros. Una IMU detecta la aceleración y los cambios en los parámetros rotacionales como son los ángulos de cabeceo, alabeo y guiñada. En nuestro caso se utiliza para conocer cuál es la posición y orientación de la cámara, ya que debido a las irregularidades del terreno, con sus correspondientes baches, sobre el que navega el tractor se producen vaivenes y

vibraciones indeseados. Si la superficie del terreno por la que navega el tractor fuese uniforme, sin obstáculos o irregularidades (piedras o baches) conoceríamos con la exactitud que permite el sistema cuál es la posición y orientación del sistema de visión en particular y de la cámara en general, pudiéndose proyectar las líneas de cultivo sobre la imagen con los únicos errores inherentes al sistema, ya que conocemos tanto la geometría del sistema como la distribución teórica de las plantas de cultivo en el campo, esto es, la formación de las líneas rectas, espaciadas entre sí una distancia también conocida durante el proceso de siembra del cultivo (Montalvo y col., 2011).

El sistema de visión consta de una cámara calibrada basada en la tecnología CCD con todos los parámetros intrínsecos conocidos, es decir la distancia focal de la lente, la distorsión radial de la lente, las dimensiones del CCD y la resolución dada en píxeles. La cámara y el sistema óptico, junto con el resto de elementos empleados en los estudios y en el prototipo del proyecto RHEA, se describen con detalle en la sección 5.2, parte de los cuales se introdujeron en el capítulo tres. La cámara se encuentra situada en la parte superior delantera del tractor, rotada un ángulo de inclinación (cabeceo) y a una altura desde la superficie del terreno. Los ángulos de guiñada y alabeo son también conocidos. Esto permite determinar las matrices de rotación y traslación con respecto a un sistema de referencia base situado sobre el suelo, definiendo así los parámetros extrínsecos del sistema. De este modo, cualquier área localizada en el campo puede ser perfectamente identificada en el plano de la imagen y viceversa. Esto significa que dado cualquier elemento en la imagen, en las condiciones de operación del vehículo agrícola, es posible determinar su posición relativa en el campo. En relación al sistema de referencia, el plano definido por el origen y los ejes de coordenadas OXZ se sitúa sobre el suelo y el eje Y pasa exactamente por el centro del plano de la imagen.

Dado que se conoce la ubicación de las líneas de cultivo sobre el terreno y se dispone de las matrices de transformación geométrica (rotación y traslación), a partir de los correspondientes ángulos de alabeo y cabeceo proporcionados por la IMU, además de la distancia focal de la lente, es posible determinar la proyección de dichas líneas sobre la imagen en cada instante (Fu y col., 1988). Conocida la posición teórica de las líneas de cultivo en la imagen, el siguiente paso consiste en delimitar alrededor de cada una de ellas unas áreas de búsqueda semejantes a las zonas blancas de la plantilla descrita previamente. Estas áreas se crean considerando la línea de cultivo teórica y un margen a cada lado de estas líneas, de forma que la región delimitada por las dos líneas

trazadas a ambos lados de la línea teórica determina el espacio de búsqueda de cada línea. Como se ha mencionado previamente es importante tener en cuenta que sólo los píxeles de la imagen binarizada que caigan dentro de este espacio de búsqueda serán tenidos en cuenta para localizar la correspondiente línea de cultivo.

Una vez delimitadas las áreas de búsqueda, tanto mediante el uso de la plantilla como con la IMU, el siguiente paso consiste en determinar la recta que mejor se ajusta a la distribución de píxeles blancos procedentes de la imagen binarizada en cada área de búsqueda.

El método utilizado es el de mínimos cuadrados totales (Golub, 1973; Golub y Van Loan, 1980; Markovsky y Van Huffel, 2007) explicado en la sección 3.5.1.1. Aplicando la técnica de ajuste, mediante dicho método, a las coordenadas de los píxeles se obtienen dos parámetros para cada línea, concretamente la pendiente y el término independiente de la ecuación de la recta que la representa. En la mencionada sección, cuando se explicó el método de mínimos cuadrados se comentó que determinados puntos de la periferia, muy alejados de la distribución general que siguen la mayoría de los píxeles, podrían producir efectos no deseados en el ajuste de las rectas, estos problemas se evitan mediante el uso de la plantilla o la IMU tal y como se ha explicado previamente. Conocidas las ecuaciones de las rectas, que identifican las líneas de cultivo, ya es posible trabajar directamente con dichas ecuaciones, concretamente para hallar puntos de intersección entre cada línea y una serie de líneas horizontales con el fin de establecer las cuadrículas necesarias sobre la imagen para el cálculo de las densidades de malas hierbas. Dichas cuadrículas delimitan las zonas de actuación específica sobre la ROI.

En el esquema mostrado en la Figura 4-10, se representa de forma esquemática los pasos seguidos en el proceso de detección de líneas. Inicialmente se dispone de una “Imagen binaria del cultivo” obtenida en la etapa anterior, “Umbralización”. Para conocer qué píxeles de la imagen binaria forman parte de las líneas de cultivo, así como las propias líneas, se utiliza o bien una plantilla establecida al efecto o bien la IMU, según el planteamiento expresado anteriormente. El empleo de esta plantilla o IMU se representa en el esquema bajo el epígrafe “Plantilla o IMU”. Conociendo los píxeles candidatos a formar parte de cada línea, se realiza un “Ajuste de líneas” mediante mínimos cuadrados con el que se obtienen las ecuaciones de las correspondientes rectas

“Líneas de cultivo. Las ecuaciones así calculadas se proporcionan al módulo correspondiente encargado de la última fase del proceso global, “Cálculo de coberturas”.

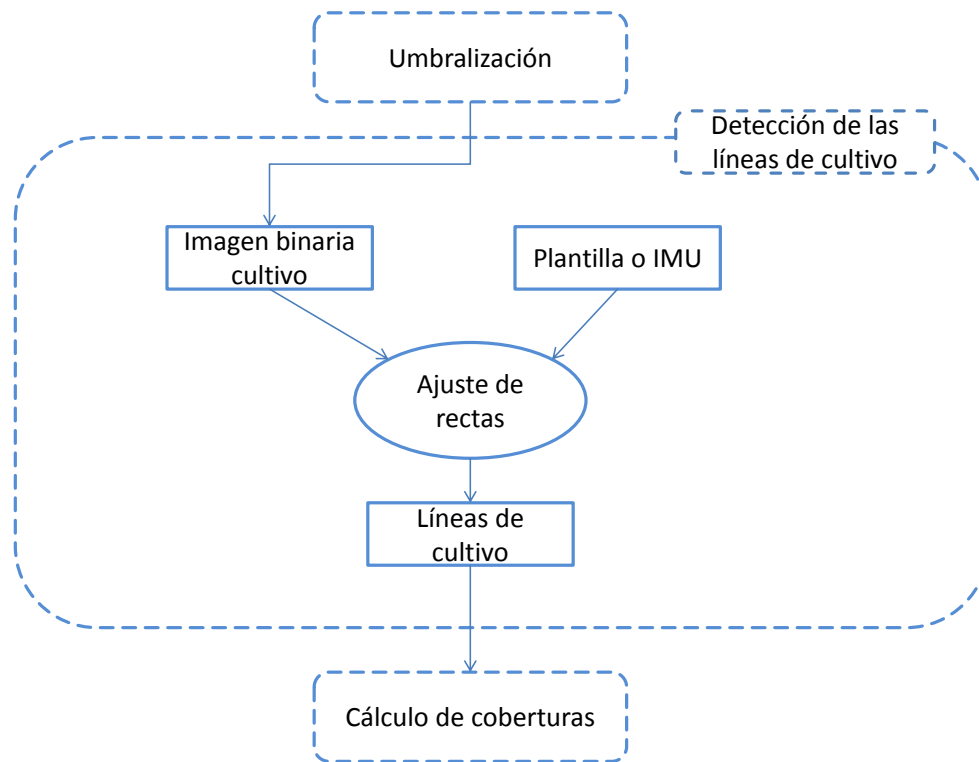


Figura 4-10. Esquema para la fase de detección de líneas.

4.2.5. Cálculo de las coberturas de malas hierbas

Nos encontramos ante la última fase del proceso global. En las etapas anteriores se identificó la vegetación existente en las imágenes, realizando posteriormente un proceso de clasificación o separación de la vegetación en dos conjuntos, el grupo de las malas hierbas y el grupo correspondiente al cultivo. A partir de los píxeles identificados como maíz, se extraen las líneas de cultivo, expresadas en forma de ecuación de una línea recta.

Ahora resta crear un mallado sobre la ROI, con el fin de aplicar un tratamiento específico, y calcular para cada una de las celdas que constituyen el mallado un porcentaje de verde, ese porcentaje representa la proporción de píxeles verdes respecto del total que integran la superficie de la celda. A dicho porcentaje asignado a una celda en concreto lo denominaremos cobertura de mala hierba para dicha celda.

La zona de la imagen que se quiere tratar es rectangular en la realidad, es decir sobre el terreno, debido a la proyección de perspectiva de la imagen, en ésta aparece en forma de trapecio. A la región que se desea tratar la denominamos región de interés a tratar, ROIT (*Region Of Interest for Treatment*). En nuestro caso debido a que la investigación se desarrolla en el marco de los proyectos AGL2011 (2011) y RHEA (2014), como ya se mencionó en el capítulo uno, especialmente en este último y también debido a que existen dependencias con el resto de sistemas involucrados, las dimensiones de esta zona vienen preestablecidas. El rectángulo que define la ROIT es de dos metros de alto y tres metros de ancho, cuyo eje central se sitúa entre las dos líneas centrales de cultivo. Los tres metros de ancho equivalen a tratar cuatro líneas de cultivo, separadas 75 centímetros entre sí. Cada línea de cultivo determina una línea central del trapecio en que se proyecta la ROIT sobre la imagen, de suerte que la mitad de los dos anchos entre dos líneas adyacentes determina a su vez el ancho de la celda. En las Figura 4-3(a) y (c) se muestra la ubicación de la ROIT y su posición respecto de la cámara. La ROIT se encuentra centrada según el eje longitudinal asociado al tractor que pasa por el centro de la cámara.

A pesar de que durante el proceso de umbralización se realiza una diferenciación entre cultivo y malas hierbas, para conseguir una mayor fiabilidad en el cálculo de las coberturas de malas hierbas se considera una región próxima a los surcos que no será tomada en cuenta a la hora de hallar dichas coberturas. Estas regiones se calculan a partir de las líneas rectas que representan el cultivo, teniendo en cuenta la proyección de perspectiva con la que se obtiene la imagen. Esta región se denomina región de exclusión, ya que se excluye en el cómputo de la cobertura vegetal de malas hierbas. Para determinar la zona de exclusión alrededor de cada línea recta que describe la correspondiente línea de cultivo, se parte de dicha línea y siguiendo ésta en el sentido de avance del tractor se considera un cierto número de píxeles a ambos lados de la misma, que equivale a un ancho fijo sobre el terreno. El desplazamiento desde la línea que representa el cultivo se denomina margen de seguridad, que permite excluir el cultivo en el cálculo de la cobertura de malas hierbas. El número de píxeles que determina el desplazamiento a izquierda y derecha a partir de la línea recta y siguiendo la misma debe ser variable en función de las siguientes consideraciones:

- La distancia a la que se encuentra la zona a tratar del tractor.

- La línea de cultivo a la que corresponda la celda.
- La posición y orientación de la cámara.
- El estado de crecimiento del maíz.

La distancia a la que se encuentra la zona que deseamos tratar del tractor varía de forma inversa al margen de seguridad, cuanto mayor sea la distancia a la que se encuentra un punto del terreno respecto del tractor menor será dicho margen en esa zona y viceversa.

Debido a la proyección de perspectiva con la que se capturan las imágenes, en los extremos de la ROIT se acentúa la convergencia hacia el punto de fuga según dicha proyección. Si bien todas las líneas de cultivo están sembradas siguiendo líneas paralelas sobre el terreno e igualmente distanciadas entre sí, en el centro de la región sobre la imagen, los surcos aparecen con una mayor distancia en cuanto al número de píxeles de separación entre sí y en relación a las celdas de los extremos de la ROIT, esto se traduce en el hecho de que las celdas son consecuentemente de menor tamaño, al igual que el ancho correspondiente al margen de seguridad.

Puesto que los parámetros intrínsecos y extrínsecos del sistema de visión, una vez calibrado, permanecen fijos, los tamaños de celdas y márgenes de seguridad permanecen invariantes respecto de sus posiciones en la imagen.

El estado de crecimiento del maíz determina también en gran medida el margen de seguridad, ya que si el maíz está muy crecido, la densidad de sus hojas será alta, ocupando una mayor superficie de la imagen. Esto implica que el margen de seguridad deberá ser mayor que cuando el maíz presenta un grado de crecimiento bajo. Esto es así para evitar considerar las plantas de maíz como malas hierbas en el cómputo de la cobertura de éstas. El crecimiento del maíz depende de diversos factores climatológicos que determinan distintos valores de temperatura, luz solar, humedad, etc. Dependiendo del estado de crecimiento del maíz así se fijará el margen de seguridad. Como bien es sabido, los tratamientos se suelen realizar siempre en un estado de crecimiento similar por lo que dicho margen no requerirá ajustes importantes a lo largo de los años.

El tamaño de cada una de las celdas que constituyen el mallado está determinado por la siembra (distancia entre líneas o ancho del surco) y la precisión de los implementos utilizados para el tratamiento de las malas hierbas. Dicho ancho resulta

ser aproximadamente la mitad del ancho de un surco sin descontar el margen de seguridad, que en el caso de los campos donde se han desarrollado los experimentos resulta ser de 37'5 cm sobre el terreno. Para su transformación a píxeles de la imagen basta con hallar las bisectrices de las líneas rectas ajustadas a las líneas de cultivo y descontar el margen de seguridad. Trazando las bisectrices y teniendo en cuenta las rectas ajustadas que representan las líneas de cultivo se delimita el ancho de cada celda, excepto en los extremos en los que se debe trazar la bisectriz entre dos líneas sin que se haya identificado una de ellas. En nuestro caso sólo se identifican las cuatro líneas centrales. Estas líneas que delimitarán la matriz de celdas se obtienen por simetría. En el caso de la línea izquierda se procede de forma que dicha línea se toma como eje de simetría y a continuación se realiza una reflexión de la primera bisectriz, entre la primera (la de la izquierda) y la segunda línea. En el caso de la línea derecha se considera la línea derecha como eje de simetría y se realiza una reflexión de la última bisectriz, entre la tercera y cuarta línea (la de la derecha).

En la Figura 4-11(a) se observa la línea verde, que representa una línea de cultivo. Dicha línea se toma como eje de simetría para realizar la reflexión de la línea roja, que representa el centro del primer surco o bisectriz de la primera y segunda líneas, generándose de este modo la línea rosa, que determina la línea extrema izquierda de la matriz de coberturas. Siguiendo cualquiera de las líneas trazadas en azul existen tantos píxeles entre la línea roja y la verde como entre la línea verde y la rosa.

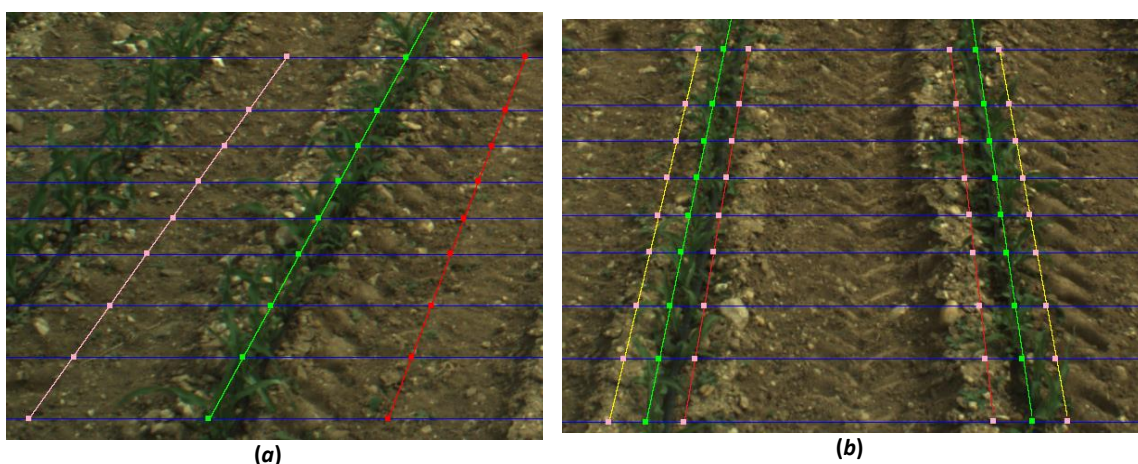


Figura 4-11. Líneas para el mallado: (a) líneas entre dos líneas; (b) región de exclusión.

El margen de seguridad se halla como un porcentaje de la distancia entre dos líneas de cultivo, de forma que la línea teórica, que define dicho margen, se traza

situándola en paralelo con la línea calculada y a una distancia de aquella dada por el margen de seguridad. Esta operación se realiza a ambos lados de cada una de las líneas de cultivo. En nuestro caso se comprobó de forma empírica que un valor adecuado para hallar el margen de seguridad resulta ser un 10% de la distancia entre líneas de cultivo en la imagen. Esta forma de calcular el margen de seguridad cumple con las condiciones enunciadas más arriba respecto del mismo. En la Figura 4-11(b) se pueden apreciar las líneas, en color rojo y amarillo, que delimitan las regiones de exclusión. En dicha figura se observan dos líneas verdes (que representan las líneas de cultivo) de forma que para calcular el margen de seguridad situándonos sobre una fila azul cualquiera, hallamos la distancia existente en píxeles entre las dos líneas verdes en esa fila y calculamos el 10% de dicha distancia. Ese porcentaje determina el valor del margen de seguridad a la derecha y a la izquierda de cada línea verde. Realizando esta misma operación para todas las líneas azules horizontales se obtienen los diferentes márgenes de seguridad, que unidos convenientemente proporcionan los márgenes globales delimitados por las líneas situadas a la izquierda y derecha de las respectivas líneas verdes identificativas del cultivo. En las líneas de cultivo de los extremos, debido a que no se dispone de otra línea de referencia se aplica el criterio de simetría explicado previamente.

Con lo expuesto hasta aquí se dispone de todas las líneas verticales, que definen la malla, que nos permitirá calcular la matriz de coberturas, habiendo delimitado el ancho de las celdas. Ahora resta por situar las líneas horizontales para completar la malla. Como se ha mencionado previamente, el alto de la ROIT sobre el terreno es de dos metros y cada una de las celdas tiene una altura de 25 centímetros. Este tamaño de celdas viene determinado por el implemento que se utiliza en RHEA (2014) para el tratamiento selectivo de las malas hierbas, concretamente un dispositivo de quemado de las mismas, desarrollado por el equipo de la Universidad de Pisa (Italia) participante en el mencionado proyecto. El área mínima sobre la que el implemento actúa es de 25 centímetros y de ahí el tamaño definitivo de las celdas. El equivalente de esos 25 centímetros a píxeles, una vez que las áreas son proyectadas sobre la imagen, no es uniforme en toda la imagen por lo que es necesario realizar la correspondiente calibración para su determinación. Con tal propósito, se extendieron nueve cintas rojas en dirección perpendicular al sentido de la marcha del vehículo, la primera a 3 metros de la vertical de la cámara sobre el suelo y el resto cada una espaciada 25 centímetros de la anterior. Se tomaron cuarenta imágenes y se observó la posición de las filas de los

píxeles en los que caía cada una de las cintas, los valores promedios resultantes de estas mediciones son los utilizados como referencia para trazar las líneas horizontales. Estos valores son válidos siempre y cuando no se cambie la configuración de la cámara, posición, orientación y focal de la misma. Conviene añadir, que dichos valores coinciden con un margen de error por debajo del 1% con los obtenidos teóricamente mediante la correspondiente transformación entre sistemas de coordenadas, incluyendo la transformación de perspectiva, una vez conocidos los valores nominales tanto los parámetros intrínsecos como extrínsecos del sistema de visión.

Una vez obtenidas tanto las líneas horizontales como las verticales, se está en disposición de establecer el mallado de celdas sobre la imagen, que permite calcular los porcentajes de malas hierbas. En la Figura 4-12 se muestra dicho mallado sobre la imagen, donde las celdas están delimitadas por líneas rojas y azules, también se muestran las líneas ajustadas sobre los surcos mediante líneas verdes para poder identificar fácilmente cuál es la zona de exclusión determinada por los surcos.

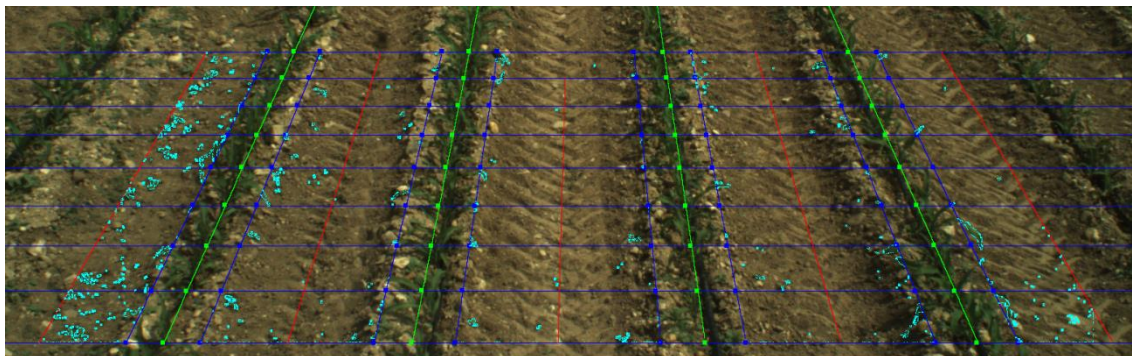


Figura 4-12. Mallado utilizado para la cobertura vegetal.

Como se puede ver en la figura anterior se han etiquetado los píxeles identificados como malas hierbas en color turquesa, lo cual permite comprobar a simple vista y de un modo sencillo qué píxeles se han clasificado como malas hierbas. Las regiones comprendidas entre dos líneas verticales azules y que son atravesadas por las líneas verdes, delimitan la región de exclusión y por lo tanto no son tenidas en cuenta a la hora de calcular la densidad de malas hierbas. Como puede apreciarse, aparecen celdas a ambos lados de los cuatro surcos. A la izquierda de los surcos las celdas están delimitadas por una línea vertical roja en el lado izquierdo y por una línea vertical azul en el lado derecho, mientras que las celdas que se encuentran a la derecha de un surco

están delimitadas por una línea vertical azul y una roja, por el lado izquierdo y derecho respectivamente.

Sólo resta obtener el valor de cobertura para cada celda, para ello basta con contabilizar el número de píxeles totales de cada celda y de esos píxeles identificar cuáles han sido etiquetados como de malas hierbas. El cociente entre ambos valores proporciona el porcentaje o cobertura de malas hierbas en la celda correspondiente.

En la Figura 4-13 se muestra el esquema del proceso propio de esta última fase que se acaba de explicar de forma detallada.

En la fase de “Detección de las líneas de cultivo” se obtienen las “Líneas de cultivo” que sirven de base para trazar el mallado sobre la imagen. En la acción “Conteo píxeles mala hierba” realizamos el mallado en la zona con las “Líneas de cultivo” y utilizamos la “Imagen binaria malas hierbas”, proveniente de la fase “Umbralización”, para hallar los porcentajes de malas hierbas existentes en cada celda, que finalmente se proporcionan en una matriz, “Matriz densidad malas hierbas”.

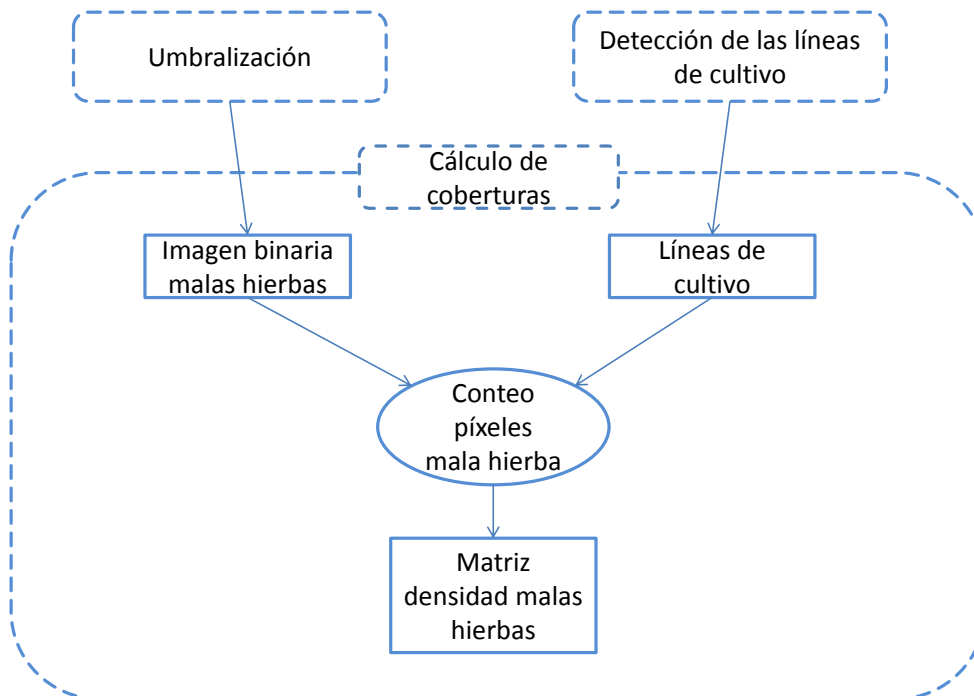


Figura 4-13. Esquema general de la fase “Cálculo de coberturas”.

Del proceso anterior se deduce que el ajuste de las rectas a las líneas de cultivo constituye un paso fundamental en el proceso. Por ello, en el marco de la investigación sobre este tema, en Guerrero y col. (2013) se propone un SEA orientado a la detección de líneas. Básicamente consiste en un proceso de segmentación aplicando el índice de vegetación COM (Guijarro y col., 2011), seguido de un proceso de binarización basado en el método de Otsu. Tras este proceso se realiza un ajuste de líneas rectas mediante dos técnicas de ajuste: a) coeficiente de correlación de Pearson y b) Theil-Sen.

4.3.SEA para detección de plantas de cultivo y malas hierbas enmascaradas

Hasta este momento se ha abordado el problema de la identificación de los verdes suponiendo que las plantas muestran un alto grado de verdor. No obstante, existe la posibilidad de que las plantas pierdan parte de su verdor por distintas razones. En efecto, el maíz es un cultivo de regadío, además crece y se desarrolla a la intemperie. Tanto durante el regadío como en épocas de lluvia, cuando la cantidad de agua es abundante o cae sobre el suelo con relativa fuerza, los materiales provenientes de la tierra (principalmente barro) impregnan la cobertura vegetal, en especial aquellas partes próximas a la tierra. En este caso la componente espectral verde de los píxeles pertenecientes a las plantas es enmascarada por la componente espectral roja proveniente de dichos materiales que las impregnan. Esto produce que las plantas enmascaradas y la tierra sean identificadas ambas como suelo, resultando su clasificación errónea.

El nuevo enfoque que se propone en esta sección se basa en un nuevo método automático con etapas secuenciales, donde la conexión de estas etapas y el proceso de segmentación de imágenes, aplicado a cada etapa, están basados en la aplicación de conocimiento humano experto. Esto conduce al diseño del SEA propuesto, descrito en Montalvo y col. (2013), que constituye la principal aportación de este enfoque gracias al cual se consigue una importante ventaja respecto a otros sistemas existentes (Guerrero y col., 2012; Guijarro y col., 2011; Meyer y col., 2004). Dicha ventaja se materializa principalmente en el hecho de que no se requiere de una etapa de entrenamiento, pudiéndose aplicar directamente a la imagen que se está procesando, llegando a ser totalmente independiente de otras imágenes que deban estar procesadas o almacenadas

previamente en una base de datos. Uno de los procesos que intervienen en dos etapas diferentes del método es la umbralización de la imagen, basada también en el método de Otsu (1979), ya utilizado y descrito previamente junto con las ventajas de su aplicación en este tipo de imágenes. De forma adicional, ya que este sistema ha sido diseñado para identificar plantas que han perdido su verdor, se puede aplicar para evaluar la efectividad de cualquier tratamiento aplicado con la intención de eliminar las malas hierbas del cultivo. Como la finalidad de este tipo de tratamiento es matar las malas hierbas presentes en el campo de maíz, cuando aquellas comienzan a morir, antes de que se hayan secado completamente, presentan una pérdida de verdor si se comparan con su estado saludable debido a la pérdida de actividad relativa a la función clorofílica. El SEA propuesto puede usarse para identificar tales plantas y de este modo valorar la efectividad del tratamiento.

El SEA se diseña de acuerdo a la arquitectura modular mostrada en la Figura 4-14. La cual consta de tres etapas conectadas secuencialmente para formar el SEA como un todo. Cada etapa contiene los módulos necesarios para el procesamiento automático de la imagen.

- Etapa 1: las plantas sin enmascarar no presentan ninguna dificultad en su identificación, se extraen directamente en esta primera etapa.
- Etapa 2: las plantas enmascaradas y las partes visibles del terreno se clasifican como pertenecientes a la misma clase, el siguiente paso consiste en su separación. Esta acción se lleva a cabo mediante la identificación de los píxeles que poseen cierto grado de verdor, los cuales se asocian a las plantas enmascaradas. Los píxeles restantes se consideran como pertenecientes al suelo del terreno.
- Etapa 3: bajo la suposición de que los píxeles pertenecientes a plantas están agrupados formando áreas de tamaños más o menos grandes y rara vez aparecen aislados, se aplica un procedimiento basado en operaciones morfológicas para conseguir mejores resultados.

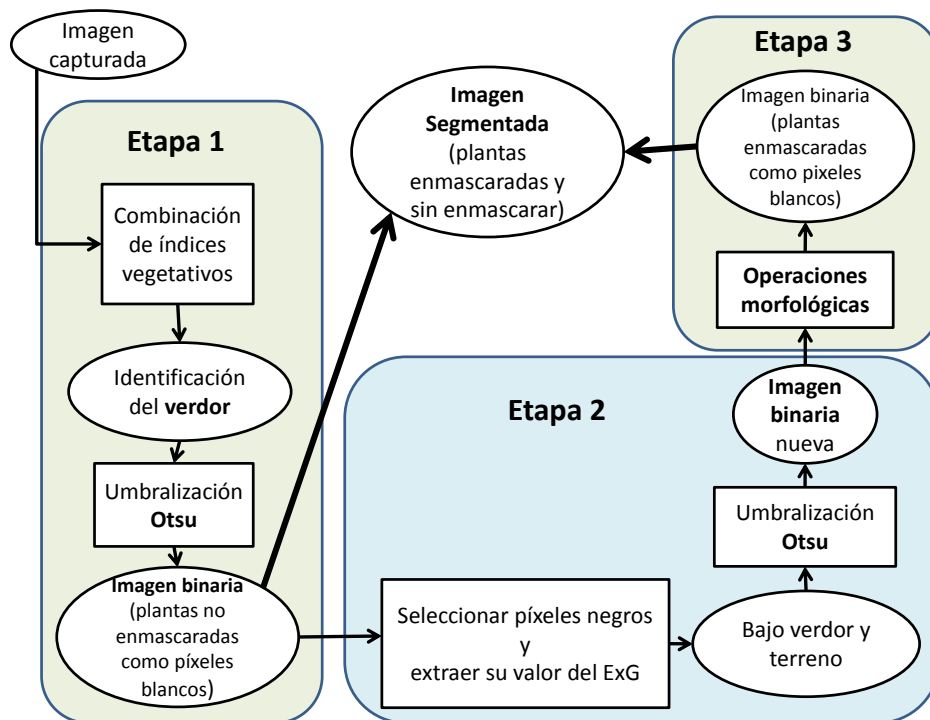


Figura 4-14. Arquitectura modular para el sistema experto diseñado para tratar con plantas enmascaradas.

A continuación se describen de forma detallada las tres etapas que conforman el proceso global.

Etapa 1: Combinación de índices de vegetación y primera umbralización

Dada una imagen original como entrada en el espacio de color RGB se desea obtener una imagen binaria en la que se disponga de los verdes existentes en la imagen. Para lo cual se utiliza una versión del ya conocido índice de vegetación COM (Guijarro, y col., 2011), descrito en el capítulo tres sección 3.3.2. Al igual que en la sección 4.2.2, donde se detalla la fase de segmentación relativa al SEA allí propuesto, antes de aplicar el índice de vegetación es necesario realizar una serie de consideraciones y restricciones para mejorar la identificación de los verdes, que constituye el objetivo perseguido. Estas restricciones obligan a que la suma de los valores de los tres canales de un píxel se encuentren en un rango dado (en la sección 4.2.2 se describe en detalle este rango) para que dicho píxel pueda ser clasificado como perteneciente al conjunto de plantas verdes, en caso de que la suma anterior no se encuentre en dicho rango el píxel es considerado como perteneciente al conjunto de los píxeles que no son vegetación.

Una vez que se han descartado los píxeles que no aportan la suficiente información en base a las restricciones anteriores, se realiza la serie de operaciones de

normalización, por otra parte muy extendidas en la literatura, descritas en el capítulo tres, sección 3.3.1.

El siguiente paso es utilizar un índice vegetativo para detectar la vegetación existente en la imagen. El índice utilizado para extraer los verdes es una versión del índice combinado COM, descrito en Guijarro y col. (2011), que combina tres de los cuatro índices utilizados en el índice original de acuerdo a la ecuación (4.3):

$$COM = W_{ExG} \times ExG + W_{CIVE} \times CIVE + W_{VEG} \times VEG \quad (4.3)$$

donde W_{ExG} , W_{CIVE} y W_{VEG} son los pesos para cada índice, representando su relevancia en la combinación. Guijarro y col., (2011) proporcionan los cuatro valores de los pesos que participan en la combinación, debido a que en este trabajo se ha excluido el índice ExGR, los valores de los pesos para los tres índices se recalculan de forma proporcional, obteniendo los siguientes valores:

Coeficiente	W_{ExG}	W_{CIVE}	W_{VEG}
Valor	0.36	0.47	0.17

Tabla 4-7: Valores de los coeficientes para la obtención del nuevo índice COM

Los valores de los píxeles de la imagen obtenida al aplicar el índice COM, se proyectan linealmente al rango $[0, 255]$, después de lo cual se realiza una umbralización aplicando el método de Otsu, obteniendo una imagen binaria donde los píxeles blancos identifican las plantas sin enmascarar, es decir, plantas no contaminadas con materiales provenientes del terreno. La componente espectral verde de estos píxeles domina sobre las componentes roja y azul.

Etapa 2: Selección de los píxeles blancos y segunda umbralización

Una vez que se han identificado los píxeles que forman las plantas sin enmascarar, los restantes píxeles se extraen con el valor obtenido mediante el índice COM. Estos píxeles restantes se asume que forman parte de la tierra y las plantas enmascaradas, siendo estas últimas el objetivo que se persigue para su identificación. A partir del conjunto de píxeles, ya separados, que constituyen la tierra y las plantas enmascaradas, se obtiene el correspondiente histograma a partir del cual y mediante la aplicación del método de Otsu por segunda vez se obtiene un segundo umbral. Tras esta segunda umbralización se identifican dos tipos nuevos de píxeles, aquellos que

pertenecen exclusivamente a la tierra y a diferentes materiales provenientes del terreno y aquellos otros pertenecientes a las plantas enmascaradas. Estos últimos, junto con los identificados previamente como plantas sin enmascarar, forman la imagen binaria final o imagen segmentada, de acuerdo al esquema de la Figura 4-14. Cuando el método se utiliza para evaluar la efectividad del tratamiento en el proceso de post-tratamiento, los píxeles identificados después de la segunda umbralización como plantas enmascaradas son aquellos pertenecientes a las plantas afectadas por el tratamiento, las cuales han comenzado su proceso de secado por disminución de la actividad clorofílica.

Etapas 3: operaciones morfológicas e identificación de las plantas enmascaradas y sin enmascarar

Una vez que la imagen se ha binarizado, se aplican las tres operaciones morfológicas siguientes en el orden mostrado:

- *Apertura*: para eliminar pequeñas zonas irrelevantes en la imagen.
- *Mayoría*: un píxel se establece a cero si cinco o más píxeles en su vecindad de 3x3 son ceros.
- *Limpieza*: cuya finalidad es eliminar píxeles aislados resultantes de las dos operaciones anteriores, es decir, píxeles blancos rodeados de píxeles negros.

Una vez que el proceso completo ha concluido, de acuerdo a las tres etapas previas, se obtiene la imagen segmentada final, donde ambos tipos de plantas, enmascaradas y sin enmascarar, son identificadas.

En Guerrero y col. (2012) se propone también un método basado en la aplicación de las conocidas Máquinas de Vectores Soporte (SVM, *Support Vector Machines*) para la identificación de plantas enmascaradas y sin enmascarar. El procedimiento consiste en aplicar una versión del consabido índice COM (Guijarro y col., 2011) de segmentación para separar las plantas del suelo, tras lo cual se obtiene una imagen binaria mediante Otsu. De esta forma se determinan dos clases, de suerte que los vectores soporte asociados a cada una de las clases se sitúan por definición en los límites del hiperplano de separación de las clases. Pues bien, se trata de identificar los vectores soporte, de forma que aquellos vectores soporte pertenecientes a la clase del suelo, definen en realidad las plantas enmascaradas, ya que están en el límite

mencionado y próximos a los vectores soporte de la clase asociada con las plantas verdes. Con los valores espectrales R, G, B de los vectores soporte asociados a las plantas se obtiene un valor promedio para cada componente, R_p , G_p y B_p conformando los vectores $x_p = (R_p, G_p, B_p)$, que constituyen las entradas al sistema de los patrones, tanto de entrenamiento como para la clasificación. Igualmente, con los vectores soporte asociados a la clase del suelo se obtiene un vector promedio de sus valores espectrales, $x_s = (R_s, G_s, B_s)$. Simultáneamente se obtienen las correspondientes desviaciones estándar σ_{pR} , σ_{pG} y σ_{pB} , así como las σ_{sR} , σ_{sG} y σ_{sB} .

El proceso consta de una fase de entrenamiento y una de decisión. En la primera se obtienen los vectores soporte y en la segunda se utilizan éstos para la discriminación, cuyo proceso se resume como sigue. Dado un píxel con componentes espectrales R, G y B, es decir $x = (R, G, B)$, el objetivo consiste en identificar aquellos píxeles pertenecientes a cada clase de plantas. Puesto que las plantas se caracterizan por su componente espectral verde, el porcentaje de verde se obtiene a partir de x_p y x_s como sigue,

$$r_{pG} = G_p / (R_p + G_p + B_p) \text{ y } r_{sG} = G_s / (R_s + G_s + B_s) \quad (4.4)$$

Mientras la desviación relativa con respecto a las componentes espectrales verdes resulta ser,

$$t_{pG} = \sigma_{pG} / G_p \text{ y } t_{sG} = \sigma_{sG} / G_s \quad (4.5)$$

Finalmente, dado un píxel con sus correspondientes componentes espectrales, $x = (R, G, B)$, procedente de una nueva imagen, las reglas de decisión se establecen como sigue,

$$\text{Regla-1: } x \text{ es planta enmascarada si } r_{sG} - t_{sG} \leq G / (R + G + B) \leq r_{pG} - t_{pG} \quad (4.6)$$

$$\text{Regla-2: es no enmascarada si } r_{pG} - t_{pG} < G / (R + G + B)$$

Estas reglas definen los límites de verdor. La Regla-1 define una región de verdor entre los vectores soporte de ambos conjuntos que permite identificar píxeles pertenecientes al conjunto de plantas que se identifican como enmascaradas o afectadas

por el tratamiento. La Regla-2 define una región con valores espectrales altos de la componente verde, y por tanto relacionada con las plantas enmascaradas.

Capítulo 5

5. Equipamiento y materiales

5.1.Introducción

El presente capítulo se estructura de forma que describe los diferentes elementos físicos empleados durante las pruebas realizadas y que han permitido la aplicación de los métodos desarrollados en los capítulos previos. Todos los elementos están enmarcados dentro del sistema de visión, tanto desde el punto de vista de la captura de imágenes como de su procesamiento y las decisiones que en base a estos dos procesos se tomen. La descripción del sistema de visión se corresponde con el prototipo desarrollado en el marco del proyecto RHEA (2014) que estuvo operativo hasta la finalización de dicho proyecto.

El capítulo se organiza de forma que en la sección 5.2 se describen los dispositivos utilizados, que forman en su conjunto el material del sistema completo de visión. En la sección 5.2.1 se describen las cámaras utilizadas, en la sección 5.2.2 se detallan los distintos sistemas ópticos utilizados con cada una de las distintas cámaras, en la sección 5.2.3 se especifica el modelo de IMU que se empleó, en la sección 5.2.4 se describe la carcasa que alberga la cámara y la IMU para su protección, en la sección 5.2.5 se indican los distintos computadores utilizados junto con sus características y por último en la sección 5.2.6 se detalla cómo se conjugaron los sistemas descritos en la secciones anteriores para lograr el prototipo empleado en el proyecto RHEA.

5.2.Descripción del material

Los elementos básicos que componen el sistema completo de visión encargado de detectar las líneas de cultivo, malas hierbas y cultivo son los siguientes: cámara, sistema óptico, IMU, carcasa de la cámara y computador. A continuación, se detallan las especificaciones técnicas más relevantes de cada uno de los distintos elementos de forma individualizada. Esta sección finaliza con la descripción del sistema de visión en su conjunto, con dichos elementos integrados y utilizados en el prototipo de RHEA

(2014). Los diferentes equipos se han instalado a bordo de los vehículos que constituyen la flota de tractores en RHEA. Los tractores, proporcionados por Case New Holland, son vehículos comerciales adaptados específicamente para la realización de las tareas previstas en RHEA, incluyendo la navegación autónoma.

5.2.1. Cámara

La elección de la cámara utilizada en el sistema de visión no fue arbitraria, en su momento se estudiaron varias posibilidades que se comentan a continuación.

Para las capturas de las primeras imágenes se empleó la videocámara de la compañía SONY, DCR PC 110E (Sony, 2014). Las imágenes que captura esta cámara poseen una resolución de 640x480 píxeles, almacenándose en el espacio de color RGB mediante el formato JPEG. Esta cámara se utilizó en la primera fase del proyecto correspondiente a la definición de los requisitos del sistema. El objetivo era capturar imágenes en el entorno de trabajo para conseguir una primera aproximación al tipo de imágenes con las que se iba a enfrentar el sistema, las dificultades intrínsecas y sus posibilidades de trabajo en tiempo real. Al carecer de la posibilidad para su procesamiento en tiempo real, las imágenes se almacenaban en una unidad Memory Stick y posteriormente se procesaban en un PC de sobremesa. En esta primera etapa también se utilizó la cámara comercial HP R817 (Hewlett Packard, 2014) que captura imágenes de 800x600 píxeles, almacenándolas también en el espacio de color RGB con el formato JPEG. Al igual que sucede con la cámara SONY, la HP se utilizó para la captura de imágenes y su posterior procesamiento en un PC de sobremesa.

La primer cámara industrial que se utilizó fue la scA 1400-17fc de la compañía Basler (2014). En la Figura 5-1(a) se muestra dicha cámara y una imagen tomada con ella en (b). La resolución de las imágenes proporcionadas por esta cámara es de 1392 x 1038 píxeles a una tasa máxima de 17 imágenes por segundo a color. El CCD es el sensor Sony ICX 285, cuyo tamaño de píxel es de $6.45\mu m$ en Horizontal (H) y Vertical (V). El interfaz de comunicación es a través del estándar IEEE 1394b, también conocido como FireWire 800. La alimentación se realiza mediante un cable de conexión independiente con un consumo de alimentación típico de 12 V DC y 3.0 W. Una mínima variación en el voltaje de entrada superando los 36 V resulta altamente probable dañar la cámara, mientras que si el voltaje es inferior a 8V la cámara puede funcionar erráticamente. El rango de temperatura nominal de funcionamiento con seguridad se

encuentra entre 0° y los 50° C, mientras que el rango en el cuál su almacenamiento es seguro se encuentra entre -20° y 80° C. Las dimensiones de la cámara son 44, 29 y 85.8 mm de ancho, alto y largo respectivamente.



Figura 5-1. (a) Cámara Basler scA 1400-17fc; (b) imagen capturada por la cámara Basler.

Las curvas relativas a la respuesta espectral de la cámara scA 1400-17fc son las que se muestran en la Figura 5-2.

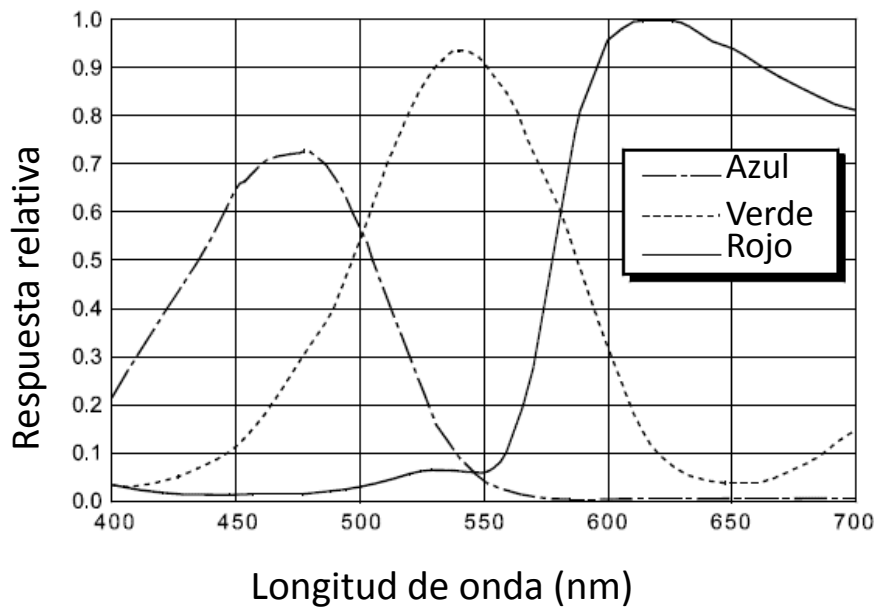


Figura 5-2. Curvas de la respuesta espectral para la cámara scA 1400-17fc.

Otra de las cámaras utilizada es la AXIS 211M de la compañía AXIS (2014). Es una cámara de red diseñada para propósitos de videovigilancia, adaptándose bien para su uso en agricultura. Esta cámara es capaz de capturar imágenes con una resolución que varía desde los 160 x 90 hasta los 1280 x 1024 píxeles. La cámara transmite video tanto en el formato MPEG-4 a una tasa de 8 imágenes por segundo

como en el formato Motion JPEG a razón de 12 imágenes por segundo a la máxima resolución. El sensor es de tipo CMOS con barrido progresivo de 1/3" de la compañía Micron. La velocidad de obturación puede modificarse, tomando valores entre 1/15000 y 1/4 de segundo. El objetivo con el que viene de fábrica es variable tanto en la apertura del iris como en la focal pudiendo variar ésta entre 3 y 8 mm. Al ser un iris de tipo DC, la apertura puede modificarse vía software, disponiendo de un pequeño motor que modifica la apertura de la lente en función de la luminosidad proveniente del entorno. Algunos de los ajustes que se pueden realizar en la imágenes son: compresión de las imágenes, color, brillo, nitidez, contraste, balance de blancos control de la exposición, zonas de exposición, rotación, efecto espejo y superposición de texto en la imagen. El rango de temperatura para un correcto funcionamiento se sitúa entre los 0° y 45° C. Las dimensiones son 44 x 88 x 172 mm y el peso es de 244 g sin la fuente de alimentación. La conexión para la transmisión de las imágenes es a través de un puerto Ethernet RJ-45 10BaseT o 100BaseTX PoE. También dispone de conexiones para la entrada y salida de audio y un micrófono integrado. En la Figura 5-3 se muestra la parte frontal de la cámara (a) y la parte trasera con las conexiones (b).



Figura 5-3. Cámara AXIS 211M: (a) Vista frontal, (b) vista trasera de la cámara AXIS 211M.

Por último, se ha utilizado la cámara industrial svS-4050CFLGEA de la compañía SVS-VISTEK (2014). La cámara puede capturar imágenes con una resolución máxima de 2336 x 1752 píxeles a una tasa máxima de imágenes de 16 Hz. El CCD es el sensor KAI 04050M/C de la compañía Kodak con un filtro de color Bayer que sigue el patrón GR8, siendo el tamaño de cada uno de los píxeles 5.5 μm , tanto en H como en V. El interfaz de comunicación es a través del estándar GigE Vision que utiliza Gigabit Ethernet y para el que posee dos puertos disponibles, realizándose la conexión mediante el uso de cables de par trenzado de categoría 5e. La alimentación

nominal es de 12 V y 700 mA y se suministra a través de otro cable distinto del Ethernet, éste utilizado para la transmisión de datos. Esta cámara resulta ser altamente robusta a las variaciones de tensión. El rango de temperatura en el que la cámara puede funcionar con seguridad se sitúa entre -10°C y 45°C . Las dimensiones de la cámara son 70, 71 y 50 mm en ancho, alto y largo respectivamente.

Existen cuatro modos de operación admitidos por la cámara SVS, a saber:

- a) *Frecuencia fija*: la cámara entrega imágenes de forma automática, no se requiere ninguna señal proveniente de un disparador externo. El tiempo de exposición (en μs) se establece mediante el software instalado en el computador. El último valor utilizado se puede grabar en la memoria EEPROM de la cámara, quedando disponible para la siguiente operación.
- b) *Disparador externo y tiempo de exposición interno*: en este modo, la cámara entrega imágenes cada vez que se le indique a través del disparador. La tasa de imágenes se determina fijando el número de pulsos del disparador por unidad de tiempo. Con cada transición positiva la cámara entrega una imagen al PC. El tiempo de exposición se establece de la misma forma que en el modo frecuencia fija. El tiempo de exposición se puede cambiar incluso con la cámara funcionando.
- c) *Disparador y tiempo de exposición externos*: en este modo, la cámara se encuentra en estado de espera a la llegada de una señal procedente del disparador para comenzar el proceso de captura y entrega de la imagen. El tiempo de exposición se determina variando la longitud del pulso del disparador (el tiempo entre el flanco de subida y de bajada). El tiempo de exposición almacenado por el usuario en la cámara se ignora.
- d) *Disparador software*: en este modo, la cámara entrega las imágenes cada vez que el PC se lo indica mediante comandos software. El tiempo de exposición que se emplea para su captura es el almacenado en la cámara, que puede variarse mientras ésta se encuentre en funcionamiento.

El modo que se utilizó durante los experimentos desarrollados con esta cámara fue el disparador software ya que permite variar el tiempo de exposición en tiempo real para ajustarlo a las variaciones que se producen en la escena y reduce el cableado necesario para la comunicación.

En la Figura 3-3 se muestra el comportamiento espectral de la cámara svS-4050CFLGEA para las distintas longitudes de onda. El comportamiento para los canales azul, verde y rojo está representado por las curvas roja, azul y verde respectivamente. Las consecuencias derivadas de esta respuesta se analizaron convenientemente en el capítulo tres. En la Figura 5-4 se muestra un vista lateral (a) y trasera con las conexiones de la cámara (b).



Figura 5-4. Cámara SVS 4050: (a) vista lateral y (b) vista trasera.

5.2.2. Sistema óptico

El sistema óptico empleado con la cámara Basler está constituido únicamente por una lente de la empresa PENTAX (2014). El modelo de la lente es el C30811, cuya longitud focal es fija, siendo de 8.5 mm. Dispone de una apertura manual del iris que se puede regular con valores para el *número-f* desde el 2.8 a 22 con enfoque manual bloqueable. Se puede utilizar con sensores cuyo tamaño de diagonal máximo sea de $2/3''$ y está equipada con un montura tipo C. El campo de visión es de $56^{\circ}28'$ en horizontal y la distancia mínima a la que se puede enfocar correctamente un objeto es de 20 cm.

El sistema óptico que se utilizó con la cámara SVS-4050 está constituido por una lente y un filtro, ambos de la compañía Schneider (2014). Las especificaciones técnicas de este sistema se introdujeron en el capítulo tres, sección 3.2, a continuación se amplía lo expuesto anteriormente. La lente utilizada es el modelo Cinegon 1.9/10-0901 con una apertura manual del iris que se puede regular con valores para el *número-f* desde el 1.9 al 16 con enfoque manual bloqueable. Es válida para sensores con un tamaño de diagonal máximo de $1''$ y está equipada con una montura de tipo F, pudiendo ser adaptada a una montura M42. La longitud focal es fija y tiene un valor de 10 mm.

Su campo de visión se sitúa por encima de 50° , siendo la distancia imagen de un objeto situado en el infinito $7'5$ mm, lo que permite trabajar con una anchura de tres metros como requiere la aplicación en el campo de maíz para la que ha sido utilizada en los experimentos desarrollados. El rango espectral admisible varía desde los 400 a los 1000nm, lo que equivale al espectro visible y al infrarrojo cercano. Con este sistema óptico las imágenes se capturan bajo proyección de perspectiva (Fu y col., 1988).

Como se ha mencionado anteriormente, el sistema trabaja en entornos adversos de exterior donde la iluminación natural contiene una alta componente infrarroja. El sensor de la cámara es muy sensible a la radiación NIR y en menor medida a la ultravioleta (UV). La radiación NIR contamina altamente los tres canales espectrales (Rojo, Verde y Azul), en especial el rojo por su proximidad al infrarrojo dentro del espectro, produciendo imágenes con colores cálidos, tal como se explica en el capítulo tres. Esto produce que la identificación de las líneas de cultivo y las malas hierbas sea inviable debido a que durante el tratamiento estas estructuras son básicamente verdes. Para evitar este efecto no deseable, al sistema óptico se le ha añadido un filtro pasa banda Schneider, cuyo modelo es el UV/IR 486 (Schneider Kreuznach, 2014). La Figura 5-5 muestra la curva de respuesta del filtro en función de las longitudes de onda de trabajo. Se observa cómo el rango de operación se encuentra entre los 370 y 700 nm, permitiendo únicamente el paso de longitudes de onda correspondientes al espectro visible, bloqueando las radiaciones con longitudes de onda NIR y UV.

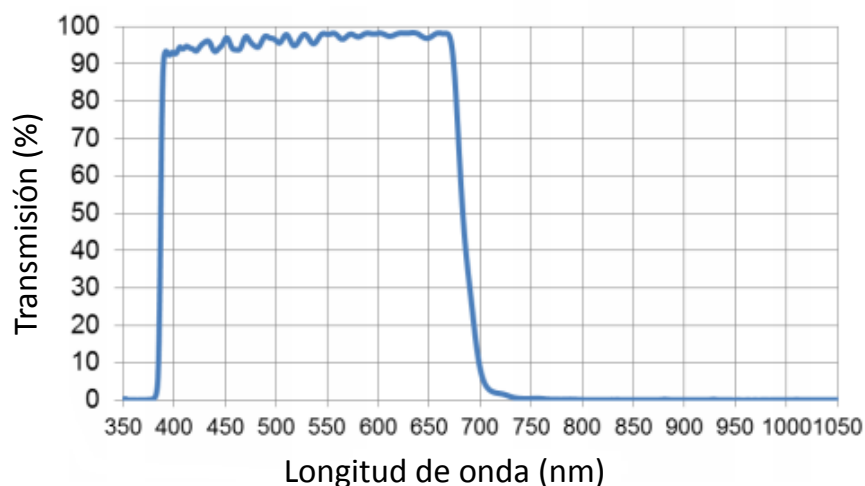


Figura 5-5. Curva de operación del filtro pasa banda.

5.2.3. Inertial Measurement Unit (IMU)

La IMU utilizada es un dispositivo de MicroStrain (2014), consistente en un sensor inercial. El modelo utilizado es el 3DM-GX3-35, un pequeño AHRS (sistema de referencia de actitud y rumbo) con GPS, que combina la tecnología de un sensor micro-electro-mecánico MEMS (*Micro-electro-mechanical Systems*) y un receptor GPS embebido de alta sensibilidad. El sensor incorpora un acelerómetro, un giróscopo y un magnetómetro capaces de medir en los tres ejes del espacio 3D, y un receptor GPS, gracias a los cuales proporciona información inercial, de orientación y de posición. El sensor puede suministrar además información tal como: aceleración, velocidad angular, ángulo girado, campo magnético, ángulos de Euler o la posición y el instante de tiempo según GPS.

El rango de temperatura a la que puede trabajar el sensor sin dañarse varía desde los -40° hasta los 65° C. Las dimensiones son $40 \times 24 \times 14$ mm sin considerar las pestañas para su montaje con la que el ancho pasaría de los 24 a los 37mm, su peso total es de 23 g. La interfaz de conexión se realiza a través de un puerto RS-232. En la Figura 5-6 se muestra una imagen del sensor inercial descrito.



Figura 5-6. Sensor inercial con receptor GPS 3DM-GX3-35.

5.2.4. Carcasa

Debido a que tanto la cámara como la IMU no están diseñadas para trabajar en entornos de exterior sin ninguna protección se les ha protegido mediante una carcasa. Tanto la cámara como la IMU están instaladas sobre un tractor trabajando en un campo de cultivo de maíz, existiendo agentes externos capaces de afectar negativamente a ambos componentes que constituyen el sistema de detección de líneas de cultivo y

malas hierbas. El polvo, objetos extraños y temperaturas extremas (tanto negativas como positivas) son algunos de tales agentes.

Los requisitos del proyecto RHEA establecen que la protección debe ser al menos con certificación IP65 (International Electrotechnical Commission). La carcasa finalmente elegida posee grado de protección IP66, que supera los propios requisitos de RHEA. Esta carcasa viene equipada con un ventilador y un calefactor en su interior, alimentados ambos a 12V y regulados por un termostato, cuya finalidad es mantener la temperatura dentro de los rangos operativos de los dispositivos en su interior. La carcasa está construida de aluminio con un pequeño parasol situado en la parte superior del cristal tras el cual se sitúa el objetivo de la cámara. Dispone de dos salidas en la parte inferior para el cableado. En su interior posee una plataforma con dos raíles que permiten el deslizamiento de la cámara a lo largo de su interior pudiéndola ajustar en la posición deseada. Las dimensiones de la carcasa son 107x141x370 y el peso es de 600g.

En la Figura 5-7 se muestran sendas vistas exterior e interior de la carcasa. En la vista interior se puede apreciar en la parte izquierda de la carcasa (en el extremo donde se encuentra el cristal) un trozo de metal que desempeña la función de calefactor al permitir su calentamiento; en el extremo opuesto se encuentra el ventilador. Próximo al ventilador se sitúa el termostato, con la circuitería asociada, encargado de controlar el funcionamiento del calefactor y del ventilador.



Figura 5-7. (a) Vista exterior de la carcasa; (b) Vista interior de la carcasa.

5.2.5. Computadores

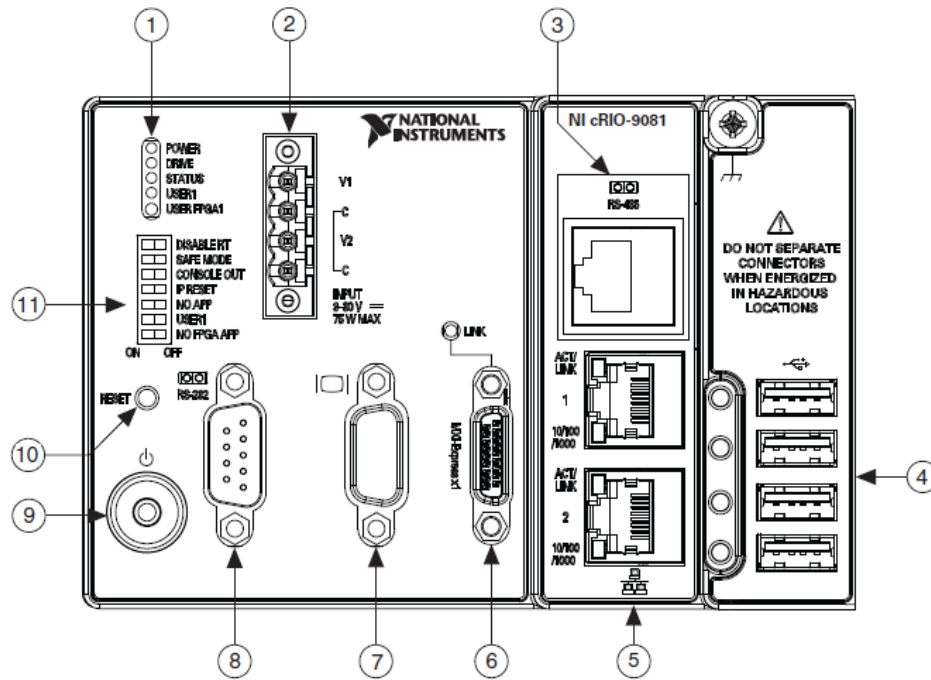
El computador con el que se han procesado off-line las imágenes capturadas con las cámaras de uso comercial, SONY y HP, y con la cámara de uso industrial Basler ha sido un portátil HP Pavilion dv6 Notebook PC, con un procesador Intel Core i7-

2670QM, 4Gb de memoria RAM y una tarjeta gráfica ATI RADEON con 2 Gb de memoria dedicada.

Las imágenes capturadas con las cámaras SVS y AXIS son procesadas en tiempo real utilizando como elemento de procesamiento la plataforma CompactRIO (cRIO) de la empresa National Instruments (2014). cRIO es un sistema embebido y reconfigurable de control y adquisición. La arquitectura robusta del hardware en la cRIO incluye módulos de E/S, un chasis FPGA reconfigurable y un controlador embebido. cRIO se programa con la herramienta gráfica LabVIEW, que facilita la labor de programación a ingenieros sin conocimiento en los lenguajes tradicionales como C y permite además la integración de librerías dinámicas DLL como un módulo más de la herramienta.

El modelo cRIO utilizado es el 9082 cuya CPU es un Intel Core i7-660UE, con dos núcleos, cuya frecuencia base es de 1.33GHz, frecuencia turbo de 2.44GHz, memoria caché de segundo nivel de 256 KB por núcleo y caché de tercer nivel de 4MB compartida por los dos núcleos. El sistema de memoria posee 2 GB de memoria RAM DDR3 y 32GB de memoria no volátil. En la Figura 5-8(a) se muestra un esquema del panel frontal, donde se encuentran las interfaces de conexión, los interruptores y los indicadores. En la Figura 5-8(b) se muestra una imagen frontal de la propia cRIO utilizada.

Las dimensiones de la cRIO son 406.4 x 138.94 x 121.9 mm de largo, ancho y alto respectivamente. Su montaje se realiza sobre un rail DIN pudiéndose montar y desmontar de forma sencilla sin necesidad de utensilios específicos.



(a)



(b)

Figura 5-8. (a) Esquema con las conexiones del panel frontal de la cRIO 9082; (b) Vista general de la cRIO.

A continuación se describen las conexiones mostradas en la Figura 5-8(a):

- 1) Leds que muestran el estado e información de la cRIO.
- 2) Conector que suministra la energía.
- 3) Puerto serie RS-485 accesible a través de un conector RJ-50.
- 4) Puertos USB (4)
- 5) Puerto Ethernet RJ-45 (2)
- 6) Conector MXI Express
- 7) Conector de video VGA
- 8) Puerto serie RS-232
- 9) Botón de encendido
- 10) Botón de reinicio
- 11) Interruptores DIP para configuración

5.2.6. Prototipo RHEA

Como se ha mencionado previamente, las cámaras utilizadas para el prototipo utilizado en el proyecto RHEA son la cámara industrial SVS junto con el sistema óptico Schneider explicado en la sección 5.2.2 y la cámara de la compañía AXIS. De los tres vehículos que componen la flota de RHEA, el primero de ellos está equipado con la cámara SVS y los otros dos con la cámara AXIS. Las cámaras se encuentran situadas en la parte superior de los tractores a una altura de 2,10 metros desde el suelo y con una inclinación de 21° respecto de la vertical, cuya configuración se muestra en la figura 4.3. El computador empleado es la solución cRIO de National Instruments. Al sistema también se le añade la IMU y la carcasa anteriormente descritas.

Ambos modelos de cámara se conectan a la cRIO a través del primer puerto Ethernet RJ-45, mediante una conexión GigE. La IMU se conecta a través del único puerto RS-232 al computador. A través del puerto RJ-45 que queda libre se realiza una conexión a un conmutador al que se conectan el resto de elementos a bordo del vehículo como son el GPS, los sistemas de control del implemento trasero para el tratamiento de las malas hierbas y el sistema de guiado.

Tanto la cámara como la IMU se encuentran ubicadas dentro de la carcasa. Como se puede observar en la Figura 5-9 la cámara se encuentra en la parte delantera de la carcasa en contacto con el cristal. La IMU se encuentra situada a continuación de la cámara, lo más próxima a la misma con el fin de captar los mismos movimientos y variaciones que sufre la cámara, ya que los valores proporcionados por la IMU se utilizan para la detección de las líneas de cultivo por la determinación de los parámetros extrínsecos de la cámara. En la parte trasera de la carcasa y en el extremo opuesto de la cámara se encuentra el ventilador encargado de la refrigeración. Entre el ventilador y la IMU se encuentra el termostato y la circuitería encargada de activar tanto el ventilador como el calefactor, localizado en la parte inferior del sistema óptico.

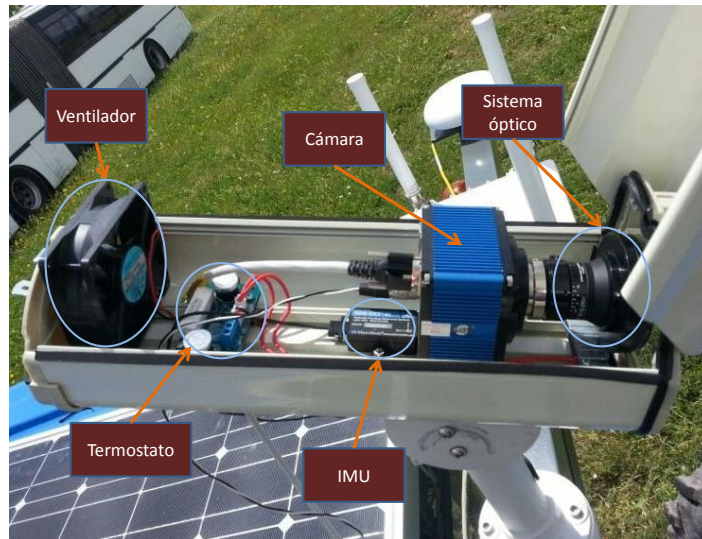


Figura 5-9. Vista de la carcasa abierta.

La cRIO se encuentra dentro de la cabina del tractor, de modo que no necesita de ninguna protección especial al estar ya protegida de la intemperie. Su ubicación integrada en el tractor se muestra en la Figura 5-10.

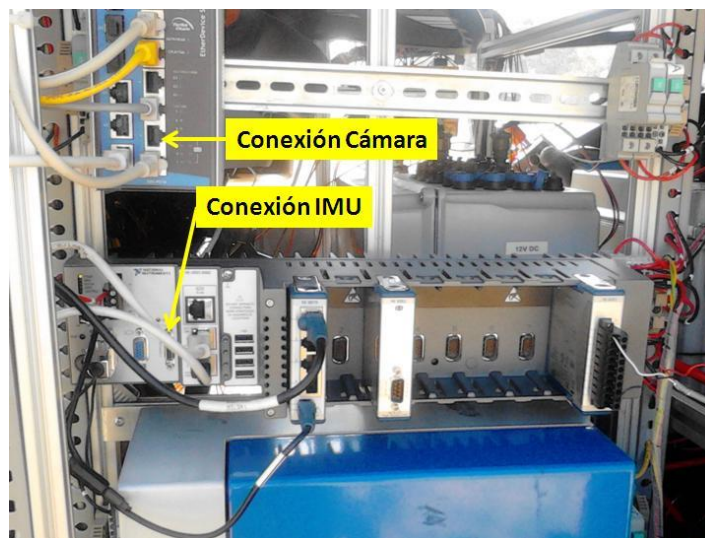


Figura 5-10. Integración de la cRIO en el sistema del tractor en RHEA.

Como se explicó en la sección 1.2 *Descripción del problema* el método elegido para eliminar las malas hierbas una vez detectadas, es mediante un procedimiento de arado, actuando entre las líneas de cultivo, simultaneado con otro sistema de quemado que lanza llamaradas que debilitan en gran medida a las malas hierbas sin apenas afectar a las plantas de maíz.

Capítulo 6

6. Análisis de resultados

6.1.Introducción

En el presente capítulo se describen y analizan los resultados obtenidos por los métodos propuestos en el capítulo cuatro. Como quiera que son dos los procedimientos desarrollados bajo el paradigma de los sistemas expertos, en las secciones 6.1.1 y 6.1.2 se incluyen los resultados de ambos SEA, distinguiéndose entre el SEA para detección de líneas de cultivo y malas hierbas y el SEA para detección de plantas de cultivo y malas hierbas enmascaradas. Se han desarrollado los anteriores dos SEA pensando en dos situaciones relativamente comunes en las que se puede encontrar el campo. La primera se refiere a la existencia de una elevada concentración de malas hierbas y la segunda está relacionada con el hecho de que las plantas hayan sido impregnadas y salpicadas con barro y distintos elementos procedentes del terreno debido a la lluvia o regado del cultivo. En cualquier caso, los SEA están diseñados para trabajar en campos de cultivo que no poseen tales particularidades. Con tal propósito, en el apartado 6.1.3 se explica cómo se extiende y adapta el primer SEA para su aplicación en el proyecto RHEA, ya que el tipo de imágenes que se han analizado en el RHEA son imágenes con alta o baja densidad de malas hierbas pero casi nunca se ha estado ante la situación de analizar imágenes donde las plantas hubieran sido impregnadas con elementos provenientes del terreno.

6.1.1. SEA para detección de líneas de cultivo y malas hierbas

Todas las imágenes utilizadas en este estudio pertenecen a campos de maíz. Las imágenes analizadas se han clasificado en dos conjuntos. Un primer conjunto formado por 1200 imágenes reales capturadas con la cámara Basler scA 1400-17fc durante los meses de Abril y Mayo del año 2011, estas imágenes se incluyen en lo que denominamos Tipo-1 a efectos de exposición de los resultados. Las imágenes digitales fueron capturadas bajo proyección de perspectiva y almacenadas como imágenes a color de 24 bits con una resolución de 1392×1038 píxeles y guardadas en el espacio de color

RGB en el formato TIFF. La configuración de la cámara fue la siguiente: los ángulos de cabeceo, alabeo y deriva fueron 20° , 0° y 0° respectivamente y la altura a la que se encontraba respecto del terreno fue de 1.5 m. Las capturas de las imágenes se realizaron durante varios días, obteniendo conjuntos de imágenes con distintos estados de crecimiento y diferentes condiciones de iluminación. En las imágenes aparecen áreas con distintas densidades de malas hierbas. Un segundo conjunto de 1400 imágenes digitales, identificadas como pertenecientes al Tipo-2, capturadas con la cámara comercial Sony DCR PC110e con una resolución de 720×576 píxeles, se tomaron durante los meses de abril y mayo de los años 2007 a 2009. Las imágenes pertenecientes al Tipo-2 también fueron almacenadas en el espacio de color RGB con el formato JPEG. La configuración de la cámara para este tipo de imágenes fue: ángulos de cabeceo, alabeo y deriva establecidos a 20° , 0° y 0° respectivamente y la altura de la cámara a 1.5 m respecto del terreno. Todas las imágenes, tanto las pertenecientes al Tipo-1 como al Tipo-2, fueron adquiridas en la finca experimental de 1.7 hectáreas situada en La Poveda, Arganda del Rey (Madrid), pertenecientes al Centro de Investigaciones Agrarias del Consejo Superior de Investigaciones Científicas (CSIC). Dicha finca se sitúa próxima al Centro de Automática y Robótica (Centro Mixto CSIC-UPM), al que pertenecen los grupos de investigación con quienes se ha colaborado en los proyectos RHEA y AGL2011.

La Figura 6-1(a) y (b) muestran dos ejemplos ilustrativos de ambos tipos de imágenes, identificadas como Tipo-1 y Tipo-2 respectivamente. Ambas contienen una alta densidad de malas hierbas, resultando en el caso de (b) la concentración excepcionalmente alta. Se utilizarán estas dos imágenes para ilustrar las conclusiones derivadas de los resultados, ya que son dos imágenes representativas de los dos tipos de imágenes adquiridas, presentando el resto de imágenes comportamientos similares



Figura 6-1. Dos ejemplos representativos de los tipos de imágenes: (a) Tipo-1; (b) Tipo-2.

Observando las imágenes que aparecen en las Figura 6-2(a) y Figura 6-3(a), podemos ver cómo tras realizar la primera umbralización, tal y como se contempla en este SEA, las plantas verdes, englobando tanto las malas hierbas como el cultivo, son recuperadas e identificadas a partir de las imágenes originales, mostradas en la Figura 6-1. Después de la segunda umbralización, únicamente las plantas correspondientes al cultivo son identificadas. Esta identificación se realiza con éxito en ambos tipos de imágenes, Figura 6-2(b) y Figura 6-3(b), a pesar de existir una alta densidad de malas hierbas, con especial relevancia en el caso de la imagen de la Figura 6-1(b). Esto significa que la estrategia propuesta aplicando de forma escalonada un doble Otsu resulta apropiada para imágenes de campos que contengan una alta densidad de malas hierbas. Resulta a todas luces evidente el hecho de que tras este doble proceso de umbralización con Otsu, aparentemente las líneas de cultivo se visualizan lo suficientemente bien como para facilitar la posterior detección de malas hierbas basada en el ajuste de líneas. Un hecho destacable es que este comportamiento del doble Otsu resulta ser la conducta general para el conjunto de imágenes analizadas, lo que determina su utilidad en este tipo de imágenes.

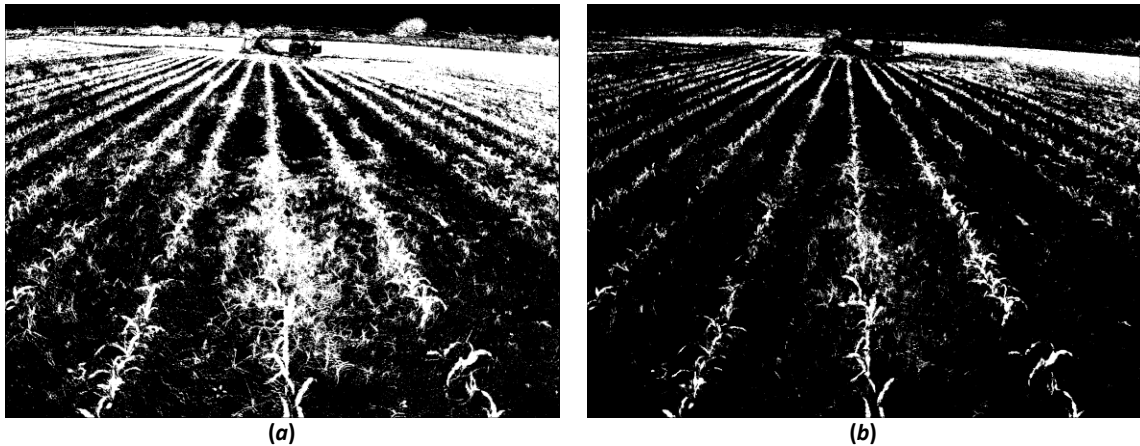


Figura 6-2. Umbralizaciones para la imagen de la Figura 6-1(a): (a) Primera umbralización; (b) Segunda umbralización.

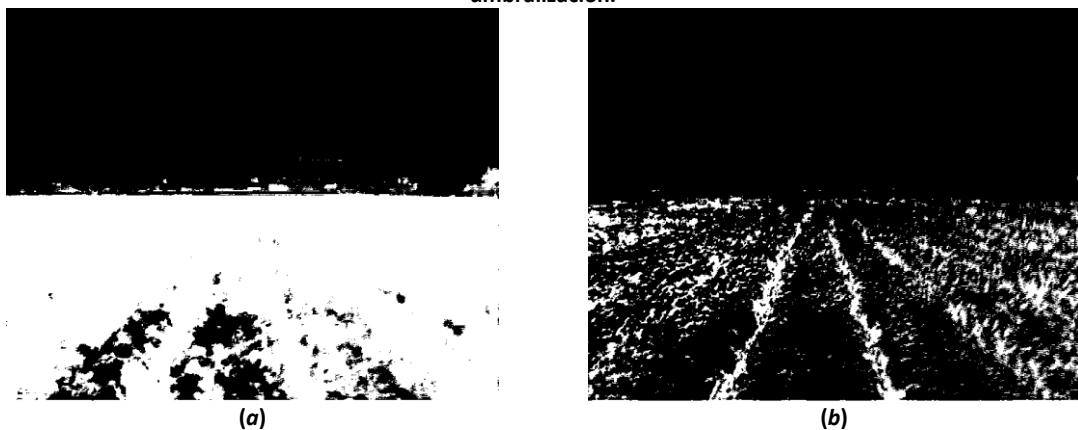


Figura 6-3. Umbralizaciones para la imagen de la Figura 6-1(b): (a) Primera umbralización; (b) Segunda umbralización.

En las Figura 6-4 se muestran las tres líneas de cultivo, identificadas en cada imagen, de forma que cada línea recta viene identificada por su correspondiente ecuación. Como se puede ver en estas imágenes, las líneas rojas se ajustan a las líneas de cultivo inclusive cuando éstas aparecen sin una distribución simétrica respecto a un eje central imaginario que supuestamente pasa por el centro de la imagen. Esto último es especialmente apreciable en la imagen mostrada en la Figura 6-4(b).



Figura 6-4. Líneas de cultivo detectadas (rojo) a partir de las imágenes mostradas en la Figura 6-1.

Como se ha mencionado en el capítulo dos, la transformada de Hough ha sido aplicada en muchos métodos para identificar líneas de cultivo, de aquí que nosotros hayamos comparado el rendimiento de nuestro enfoque basado en regresión lineal (RL) respecto a la transformada de Hough (HOU). Ambas comparativas han sido realizadas después de aplicar el método de la doble umbralización basado en Otsu con el fin de que los resultados sean equiparables. La transformada de Hough se ha implementado en su versión de coordenadas polares, (Gonzalez y col., 2009) donde la variación del parámetro que representa el ángulo se realiza en incrementos de un grado sexagesimal.

La comparación se establece en términos de efectividad y tiempo de proceso o coste computacional medido en unidades de tiempo. La efectividad se mide basándose en el criterio de un experto humano, donde una línea, previamente detectada, se considera como correcta si se solapa con la línea de cultivo establecida de acuerdo a la observación del experto. Sobre el conjunto de 2600 imágenes analizadas, se calcula el porcentaje de la media de líneas rectas solapadas de forma correcta sobre el cultivo para ambas aproximaciones, RL y HOU. Debido a que el principal objetivo del método propuesto en este SEA es su implementación para aplicaciones de tiempo real, se ha medido el coste computacional. También, con el objetivo de conseguir una aplicación

en tiempo real, se realiza un análisis comparativo de la eficiencia de los métodos para imágenes con distintas resoluciones.

Para obtener estas nuevas imágenes, de distinta resolución, se ha realizado un submuestreo por dos, en las filas y columnas, para todas y cada una de las imágenes originales disponibles pertenecientes a los dos conjuntos identificados como Tipo-1 y Tipo-2 descritos previamente. Procediendo de este modo, obtenemos dos nuevos conjuntos de imágenes cuya resolución en tamaño de píxeles resulta ser la cuarta parte de las originales.

Bajo esta acción de reducción de la dimensión de las imágenes, subyace la idea de que, es posible reducir la dimensión de las imágenes, manteniendo en las mismas la información principal que nos permita detectar las líneas de los surcos sin perder efectividad, a la vez que el tiempo de procesamiento disminuye. La Tabla 6-1 muestra los resultados para los distintos conjuntos de imágenes. La primera columna de la tabla contiene la resolución de las imágenes, existiendo dos resoluciones para cada tipo de imagen (Tipo-1 y Tipo-2), que se corresponden con las resoluciones original y reducida. La segunda y tercera columna contienen la media de los porcentajes de efectividad, en términos de solapamiento correcto sobre los conjuntos de imágenes estudiadas; la cuarta y quinta columnas muestran el tiempo medio de ejecución en milisegundos por imagen. Como se ha mencionado previamente, estos tiempos corresponden al proceso completo descrito en el SEA, y por tanto, segmentación de la imagen, doble umbralización mediante Otsu y la regresión lineal (RL) o la transformada de Hough (HOU) en su caso.

Los algoritmos se desarrollaron en el lenguaje de programación orientado a objetos C++ mediante el entorno de desarrollo integrado Visual Studio 2008 y haciendo uso de la librería para tratamiento de imágenes OpenCV (2014) bajo licencia BSD. Los tiempos fueron medidos en un computador con un procesador Intel Core i7 y 4 GB de memoria RAM.

Resolución en píxeles	Porcentaje de efectividad (%)		Tiempo de procesamiento (ms)	
	HOU	RL	HOU	RL
1392×1044 (Tipo-1)	89.3	95.5	4678	297
696×522 (Tipo-1)	86.4	92.9	2567	74
720×576 (Tipo-2)	82.4	89.6	1935	81
360×288 (Tipo-2)	80.1	88.1	1342	28

Tabla 6-1: Rendimiento de los enfoques HOU y RL medido en términos de efectividad y tiempo de proceso

A partir de los resultados mostrados en la Tabla 6-1, podemos inferir que el mejor rendimiento, en términos de efectividad, se consigue mediante RL con la mayor resolución y para imágenes del Tipo-1. Esto se justifica por el hecho de que imágenes

con una mayor resolución contienen más información que las mismas imágenes con una resolución inferior. En referencia a los tipos de imágenes, está claro que las imágenes del Tipo-2 contienen una mayor complejidad que las imágenes del Tipo-1, en base a esto se explican los peores resultados obtenidos, en términos de efectividad, para las imágenes pertenecientes al Tipo-2.

Respecto a los tiempos de procesamiento, RL supera a HOU, esto se debe a que HOU requiere mayor potencia de cálculo, a pesar de que el espacio de búsqueda para el ángulo que involucra la transformada de Hough se restringió, estableciéndose los límites del rango ente 52° y 130° . Así, las líneas rectas cuyas pendientes se sitúen dentro del rango de las pendientes esperadas para las líneas son consideradas como candidatas por el algoritmo. Como se mencionó con anterioridad, el paso para incrementar el ángulo en las celdas de acumulación fue establecido a 1° sexagesimal. En cualquier caso, desde el punto de vista de su aplicabilidad en tiempo real, RL obtiene unos resultados aceptables, incluso en las imágenes de alta resolución del Tipo-1.

En Guerrero y col. (2013) se realiza un análisis cuantitativo sobre tiempos de cómputo para el mismo tipo de imágenes, si bien sin considerar específicamente las altas densidades de malas hierbas, que se requieren en el presente SEA. Tal y como se mencionó en el capítulo cuatro, sección 4.2, el SEA realiza un proceso de segmentación mediante el índice de vegetación ExG, seguido de un proceso de binarización basado en el método de Otsu. Tras este proceso se realiza un ajuste de líneas rectas mediante la técnica de ajuste Theil-Sen. Los resultados obtenidos mediante este SEA se muestran en la Tabla 6-2. En ella aparecen los resultados relativos al porcentaje de efectividad bajo el mismo criterio que el utilizado para obtener los resultados de la Tabla 6-1. En la primera fila se muestra el porcentaje de reducción de las imágenes en relación al número de píxeles. La segunda fila muestra el margen de tolerancia en píxeles para el realizar el ajuste de las líneas según Theil-Sen, tal y como se explica en Guerrero y col. (2013). En los resultados mostrados en esta tabla se ha separado el proceso de segmentación del de ajuste, de suerte que el primero se identifica como SSA, para indicar segmentación sin ajuste, tal y como aparecen en la tercera y quinta fila de la tabla. El proceso de ajuste de líneas se identifica con el nombre Theil-Shen. El sistema de Guerrero y col. (2013) fue desarrollado en LabVIEW Real-Time y ejecutado en una CompactRIO-9082.

	Reducción	0% (imagen original)			67%			84%		
	Margen de tolerancia	50	100	150	50	100	150	50	100	150
Porcentaje de efectividad	SSA	70	70	70	70	70	70	70	70	70
	Theil-Sen	86.1	93.1	94.1	85.2	91.3	84.6	83.8	83.1	77.8
Tiempo (ms)	SSA	105	105	105	11.3	11.3	11.3	3.3	3.3	3.3
	Theil-Sen	5354	7506	9568	92	476	1440	30	98	230

Tabla 6-2: Porcentaje de efectividad y tiempos de proceso para diferentes aproximaciones

Si bien los resultados obtenidos en la Tabla 6-2 no son exactamente comparables con los reflejados en la Tabla 6-1, sí que resulta posible extrapolar los resultados más relevantes para su análisis. Con el procedimiento RL se obtiene un porcentaje del 95.5% frente al mejor resultado obtenido en la Tabla 6-1, que aparece con un 94.1%, en ambos casos para imágenes sin disminución de la resolución. Si ahora comparamos el porcentaje de éxito obtenemos un 92.9% con RL y un 91.3% con Theil-Sen. La reducción que se produce no es la misma en los dos casos ya que en el caso de RL se realiza una reducción del 50% de la resolución y en el caso del Theil-Sen del 67%, siendo dos reducciones de órdenes similares para realizar la comparativa. Para visualizar más fácilmente los resultados de la comparación se ha utilizado una diagrama de barras en la Figura 6-5.

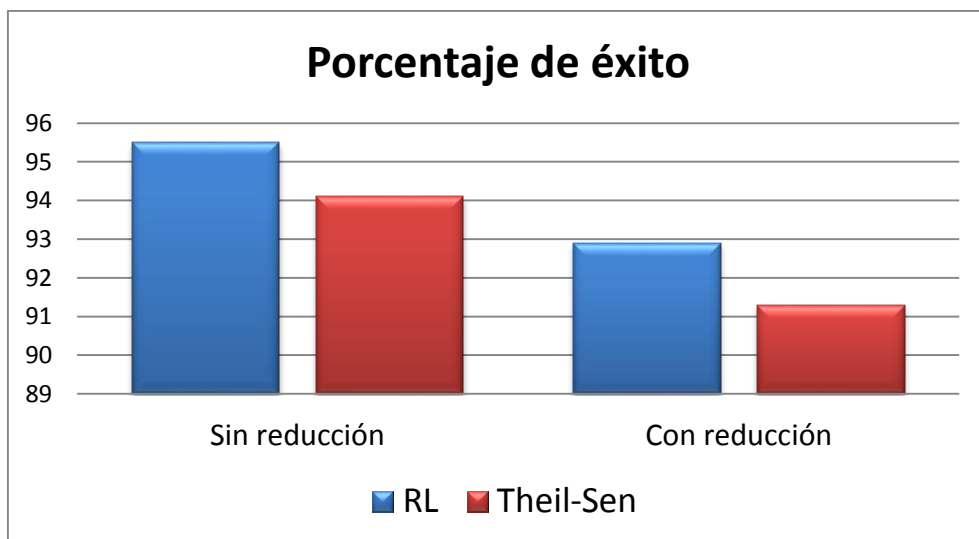


Figura 6-5. Comparativa de porcentaje de éxito entre RL y Theil-Sen.

Por otra parte, en relación a los tiempos de proceso que aparecen en la Tabla 6-2, se observa que SSA junto con el ajuste de líneas suma en el mejor de los casos 5459 ms para una imagen sin reducir, que es muy superior al tiempo obtenido con el SEA

descrito en el presente trabajo de investigación para una imagen también sin reducir, que resulta ser de 297 ms. Si comparamos los tiempos cuando se reduce la resolución de la imagen se comprueba que el mejor de los casos de la Tabla 6-2 reduciendo al 67% es de 103.3 ms mientras que para que reduciendo al 50% nuestro SEA tarda 74 ms. Esto significa que el SEA propuesto sigue siendo más eficiente en términos de coste computacional a la vez que el porcentaje de éxito también sigue siendo mayor. Si bien, la disminución de tiempos de cómputo al reducir la resolución ha sido algo mayor en el caso del Theil-Sen, en el RL se ha reducido al 25% y en el Theil-Sen al 6% en el caso medio.

En cualquier caso, los tiempos obtenidos por el SEA propuesto son perfectamente asumibles desde el punto de vista de los procesos en tiempo real previstos en RHEA. En efecto, el tractor navega a velocidades de entre 3-4 km/h, esto significa que como la ROIT posee 2 m de largo, si el tractor viaja a una velocidad de 4 km/h recorre esa distancia en 1.8s, que sería el tiempo máximo de proceso de cada imagen. Es evidente que los tiempos referidos previamente para el procesamiento de las imágenes se sitúan por debajo de ese límite, excepto en el caso del Theil-Sen con resoluciones por debajo del 84%.

Un problema adicional relativo a la transformada de Hough es la elección del máximo local en el acumulador para cada línea recta. Ciertamente, la intención es obtener una única ecuación para cada alineación de píxeles, si bien en el acumulador que genera dicha transformada para detectar las líneas aparecen diferentes valores para una única línea recta. Esto ocasiona que se encuentren varias líneas rectas asociadas a una misma línea de cultivo de las que se pretende localizar. Podríamos seleccionar la línea recta con el máximo valor pico en el acumulador o podríamos calcular la media de los valores involucrados en una misma línea de cultivo. Ambas soluciones suelen conducir a errores en la identificación de la línea recta correcta para la alineación de los píxeles. Estas son las principales razones que justifican la superioridad del método RL frente HOU para el conjunto de imágenes analizadas.

6.1.2. SEA para detección de plantas de cultivo y malas hierbas enmascaradas

Las imágenes procesadas mediante este SEA fueron adquiridas con la cámara digital HP R817 durante cuatro días diferentes entre los meses de Abril y Mayo del año

2007. Cada uno de los días en los que fueron capturadas las imágenes fue espaciado al menos en cinco o seis días. De todas las imágenes tomadas se crearon dos conjuntos distintos. El primer conjunto fue obtenido durante la fase de pre-tratamiento, después de que el campo fuera artificialmente regado, con una alta dosis de agua y también después de que el campo recibiera distintas cantidades de agua proveniente de fuertes aguaceros de lluvia. La Figura 6-6 muestra una imagen representativa de este primer conjunto de imágenes. En esta imagen se aprecia claramente cómo una parte importante de la vegetación ha sido impregnada de barro procedente del suelo, lo que conduce a que parte de las plantas han perdido su color verde en detrimento de una tonalidad más próxima al color del suelo. Este hecho es el que en condiciones normales, a la hora de la discriminación entre plantas y suelo, provoca que las plantas impregnadas se confundan con el suelo.



Figura 6-6. Imagen representativa del Conjunto-1.

Un segundo conjunto de imágenes fue obtenido en la fase de post-tratamiento, después de aplicar una dosis de herbicida sobre el campo de cultivo; en esta fase las malas hierbas han comenzado su proceso de secado. La imagen mostrada en la Figura 6-7 es una clara representante de este segundo conjunto de imágenes. En dicha imagen se aprecia cómo parte de la vegetación correspondiente a las malas hierbas ha perdido su verdor debido al tratamiento de herbicida. Debido a que las imágenes fueron adquiridas durante diferentes días, una consecuencia lógica de ello ha sido que todas ellas fueron adquiridas bajo diferentes condiciones de iluminación. Esta circunstancia no afecta al rendimiento del método diseñado en el SEA porque los métodos encargados de procesar las imágenes son independientes de esta circunstancia; en particular el método de Otsu que hace uso del histograma. Por lo tanto no se requiere de ningún estudio que dependa de las condiciones de luminosidad. Esto representa una clara ventaja con respecto a los métodos basados en una o varias etapas de aprendizaje

(Guijarro y col., 2011). Esta característica representa una justificación adicional sobre la eficiencia del SEA propuesto.



Figura 6-7. Imagen representativa del Conjunto-2.

Las imágenes digitales usadas fueron capturadas bajo proyección de perspectiva conteniendo únicamente la tierra del terreno y plantas, y por tanto excluyendo partes correspondientes al cielo o el horizonte, donde podrían aparecer elementos ajenos al cultivo y el campo. Estas imágenes han sido almacenadas como imágenes de 24 bits de color con resoluciones de 800×600 píxeles y salvadas en el espacio de color RGB (Red-Green-Blue) en el formato JPG. El Sistema Experto Automático (SEA) propuesto ha sido implementado mediante Matlab R2009a (MathWorks, 2014).

Se ha utilizado un conjunto de 850 imágenes, de las cuales 600 contenían plantas enmascaradas, que constituyen el Conjunto-1, y 250 fueron capturadas después del tratamiento de herbicida para evaluar su efectividad, que en este caso pertenecen al Conjunto-2. Sobra decir que todas las imágenes contienen plantas sin enmascarar.

El núcleo del SEA propuesto consiste en la identificación de las plantas enmascaradas y sin enmascarar basándose para ello en la división del histograma en tres clases separadas por dos umbrales. Por consiguiente el análisis de resultados se centra en el estudio de los dos umbrales obtenidos en las etapas 1 y 2, según el esquema de la Figura 4-14, mediante el método de Otsu. En ambos casos se toma como imagen de partida la obtenida con el índice COM, tal y como se explica en los capítulos tres y cuatro. El rendimiento del SEA se compara con el enfoque basado en una doble umbralización propuesto en Demirkaya y col. (2008), que denominamos DEM de aquí en adelante.

Con respecto a las imágenes pertenecientes al Conjunto-1, con el fin de comentar los resultados obtenidos nos basamos en la imagen mostrada en la Figura 6-6, ya que como hemos indicado se trata de una imagen representativa de este conjunto,

conviene señalar que el resto de las imágenes, pertenecientes a este conjunto, muestra un comportamiento similar frente al SEA. La Figura 6-8(a) muestra la imagen en escala de grises que resulta de aplicar el índice COM a la imagen mostrada en la Figura 6-6; el histograma de dicha imagen se muestra en la Figura 6-8(b). Los valores de dicho histograma se sitúan en el rango [0,255], por tanto los valores originales obtenidos mediante COM se escalan a dicho rango multiplicándolos por 255. En el eje horizontal del histograma aparecen cuatro identificadores, *a*, *b*, *c* y *d*, que representan cuatro umbrales; *a* y *c* son los umbrales obtenidos mediante el SEA en la segunda y primera etapa de umbralización respectivamente; *b* y *d* son los umbrales obtenidos mediante DEM.

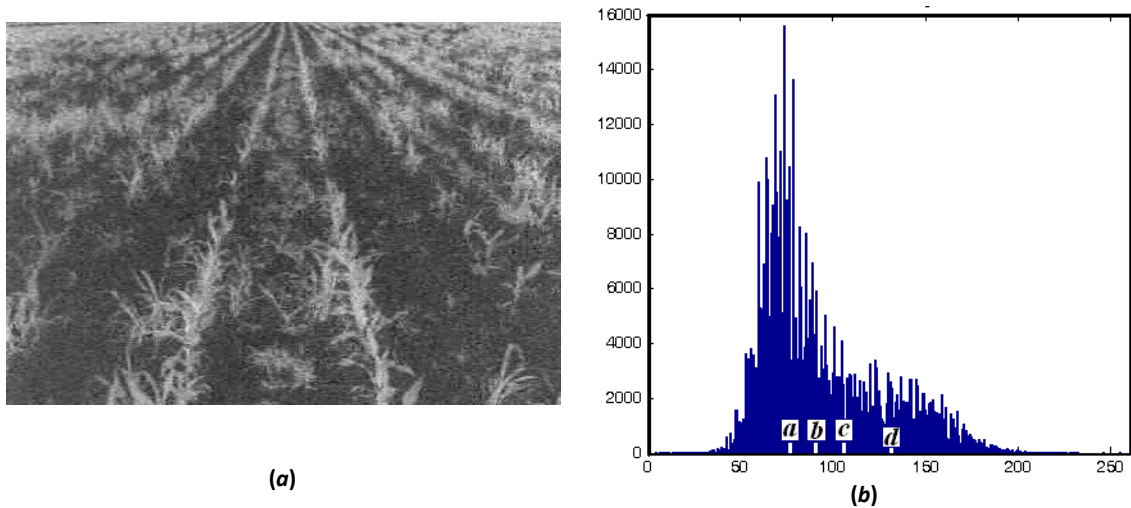


Figura 6-8. (a) Imagen obtenida al aplicar el índice COM a la Figura 6-6; (b) histograma de la imagen en (a) con cuatro umbrales.

La Tabla 6-3 muestra los cuatro valores para cada enfoque (SEA y DEM) para las imágenes del Conjunto-1.

	SEA		DEM	
	<i>a</i>	<i>c</i>	<i>B</i>	<i>d</i>
Conjunto-1	83.00	111.99	97.13	139.94
Conjunto-2	87.01	117.98	104.09	146.93

Tabla 6-3: Valores de los umbrales obtenidos mediante SEA y DEM para los dos conjuntos analizados

La Figura 6-9(a) muestra la imagen binaria obtenida con el umbral *c* mostrado en el histograma de la Figura 6-8(b) y en la Tabla 6-3 para el Conjunto-1. Los píxeles blancos son identificados como píxeles de las plantas sin enmascarar en la primer etapa del SEA y etiquetados como píxeles rojos en la Figura 6-9(b). Ahora, considerando solamente los píxeles negros en la imagen binaria, se procesa la imagen de acuerdo a los procedimientos descritos en las etapas 2 y 3 del SEA; el umbral *a* se obtiene en la etapa

2, este nuevo umbral permite identificar las plantas enmascaradas, etiquetadas como píxeles azules en la imagen de la Figura 6-9(b). La Figura 6-9(c) muestra los píxeles etiquetados usando los umbrales d y b obtenidos mediante el método DEM. El umbral d se emplea para identificar los píxeles de las plantas sin enmascarar, los cuales se etiquetan mediante el color rojo. El umbral b se emplea para identificar los píxeles de las plantas enmascaradas, que se etiquetan mediante el color azul.

Basándonos en las etiquetas mostradas en la Figura 6-9(b) y (c) podemos inferir que DEM identifica un número menor de plantas sin enmascarar que SEA a la vez que SEA identifica áreas de plantas enmascaradas que no es capaz de identificar DEM. Mediante las líneas circulares y elípticas que aparecen en la Figura 6-9(c) se identifican áreas relevantes donde se verifican las dos afirmaciones anteriores. Los mejores resultados obtenidos con SEA se explican considerando los valores umbrales del histograma. Ciertamente, las plantas sin enmascarar se extraen mediante los umbrales c y d de la Tabla 6-3; debido a que d es mayor que c , DEM identifica un número menor de píxeles provenientes de plantas sin enmascarar, los cuales se etiquetan posteriormente como píxeles de plantas enmascaradas. También, debido a que b es mayor que a y a que ambos umbrales están relacionados con la identificación de las plantas enmascaradas, algunas de las plantas enmascaradas o parte de ellas no son identificadas como tal por DEM. A partir del histograma de la Figura 6-8(b) se puede ver que no existe una separación clara entre las clases supuestamente representadas en dicho histograma. Esto significa que no es posible asignar umbrales de forma manual con una cierta claridad, mientras SEA obtiene unos umbrales satisfactorios y a la vez mejores que DEM, a juzgar por los resultados. Llegados a este punto, conviene mencionar que el uso de la media aritmética estadística como umbral, utilizada en Guijarro y col. (2011) y Burgos-Artizzu y col. (2011), no resulta apropiada para nuestro problema, lo cual se explica porque el valor de la media es siempre menor que el umbral obtenido con el método de Otsu, en cuyo caso la mayoría de las plantas enmascaradas y parte de los píxeles de suelo son identificados como plantas sin enmascarar.

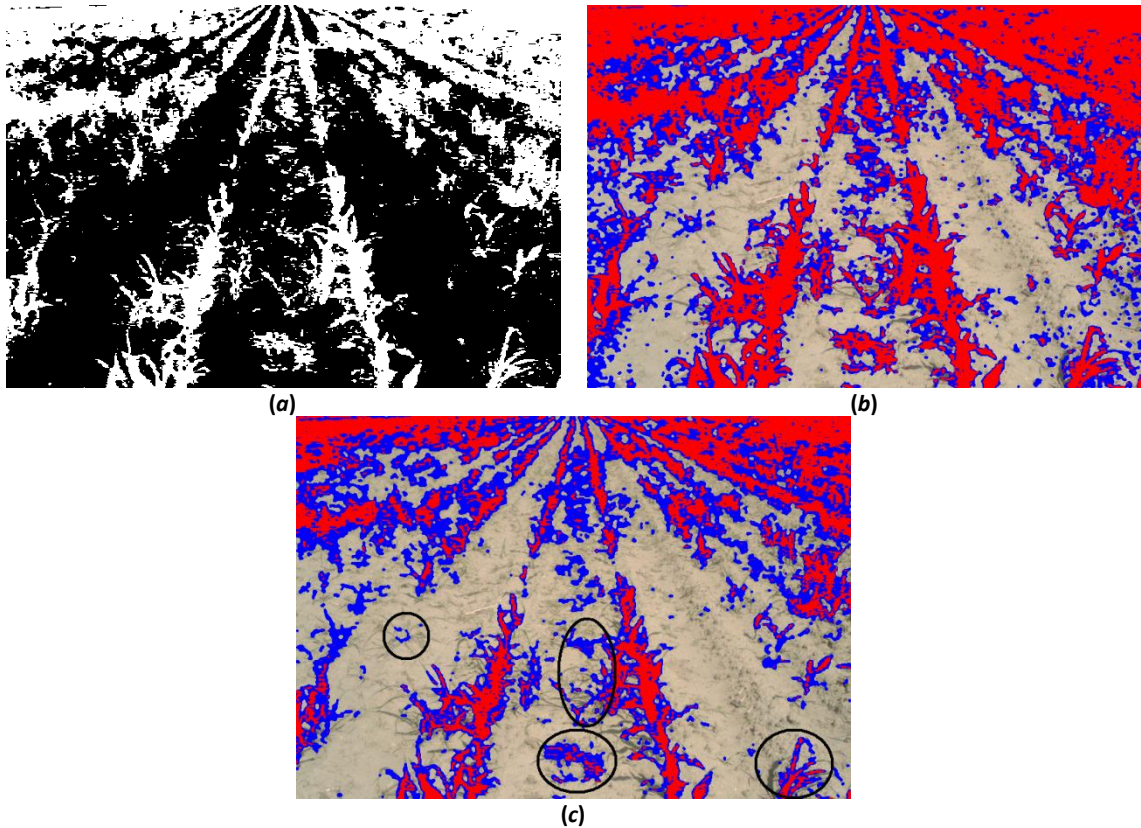


Figura 6-9. (a) Imagen binaria obtenida con el umbral c de la Tabla 6-3 conteniendo plantas enmascaradas y sin enmascarar en blanco; (b) imagen segmentada con SEA con plantas enmascaradas en azul y sin enmascarar en rojo; (c) imagen segmentada con DEM con plantas enmascaradas en azul y sin enmascarar en rojo.

Las imágenes que constituyen el Conjunto-2 contienen áreas con malas hierbas que han comenzado el proceso de secado como resultado del tratamiento de herbicida aplicado anteriormente. La imagen de la Figura 6-7 constituye un buen ejemplo como representante de este conjunto, donde podemos ver zonas importantes con malas hierbas entre las líneas de cultivo, las cuales aparecen afectadas por el tratamiento; si bien todavía existen pequeñas zonas que preservan un relativo alto grado de verdor, lo que significa que no han sido afectadas por el tratamiento. La Figura 6-10(a) muestra la imagen resultante de aplicar el índice de vegetación COM a la imagen de la Figura 6-7; la Figura 6-10(b) muestra el histograma de la imagen en (a) en el rango $[0,255]$. De nuevo, sobre la base del gráfico del histograma aparecen cuatro identificadores, a , b , c y d , que representan los umbrales obtenidos por los dos enfoques (SEA y DEM) para la imagen representante del Conjunto-2. La Tabla 6-3 muestra los cuatro valores para cada enfoque (SEA y DEM) para las imágenes del Conjunto-2.

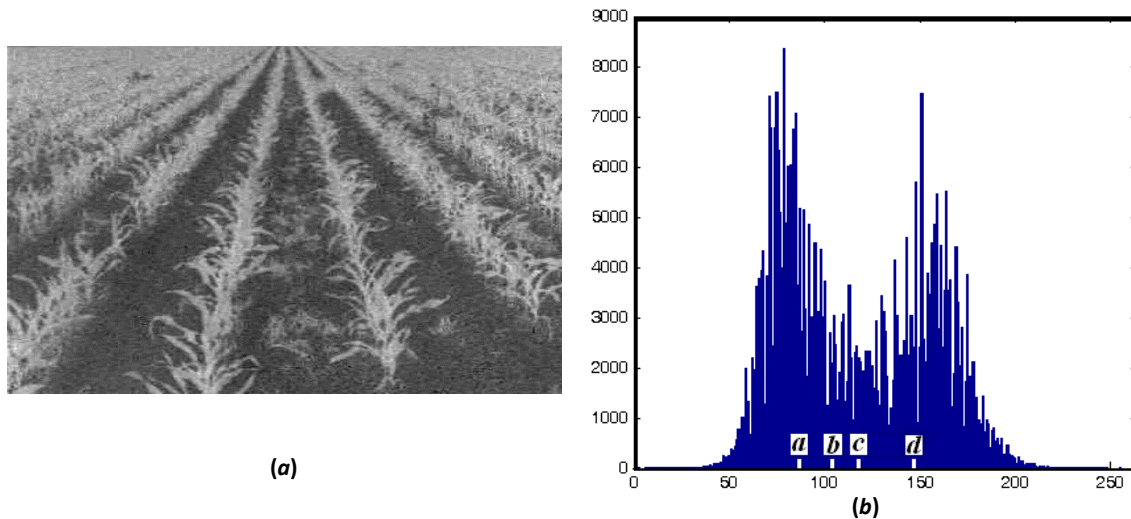


Figura 6-10. (a) Imagen obtenida a partir de la Figura 6-7 aplicando el COM; (b) histograma de (a).

A diferencia del histograma de la Figura 6-8(b), éste resulta ser más o menos bimodal, donde el umbral c es el usado en SEA para obtener la imagen binaria mostrada en la Figura 6-11(a) y sirve para identificar las plantas no enmascaradas, cuyos píxeles son etiquetados mediante el color azul en la Figura 6-11(b). El umbral d se usa en DEM para identificar también las plantas sin enmascarar, Figura 6-11(c). Todos los píxeles con valores superiores a c o d , para el enfoque SEA o DEM respectivamente, se consideran píxeles pertenecientes a plantas sin enmascarar. Muchos píxeles pertenecientes a plantas sin enmascarar son identificados por SEA pero no lo son por DEM, esto puede explicarse por el hecho de que el umbral c divide el histograma mejor que d en dos partes que coinciden con las dos regiones teóricamente modales. Además, existen varias zonas de malas hierbas identificadas por DEM, señaladas mediante líneas elípticas en la Figura 6-11(c), como afectadas por el tratamiento cuando realmente no lo están. Estas reflexiones y conclusiones son perfectamente aplicables a todas las imágenes que constituyen el Conjunto-2, que muestran un comportamiento similar al expuesto previamente con respecto a los enfoques seguidos por SEA y DEM.

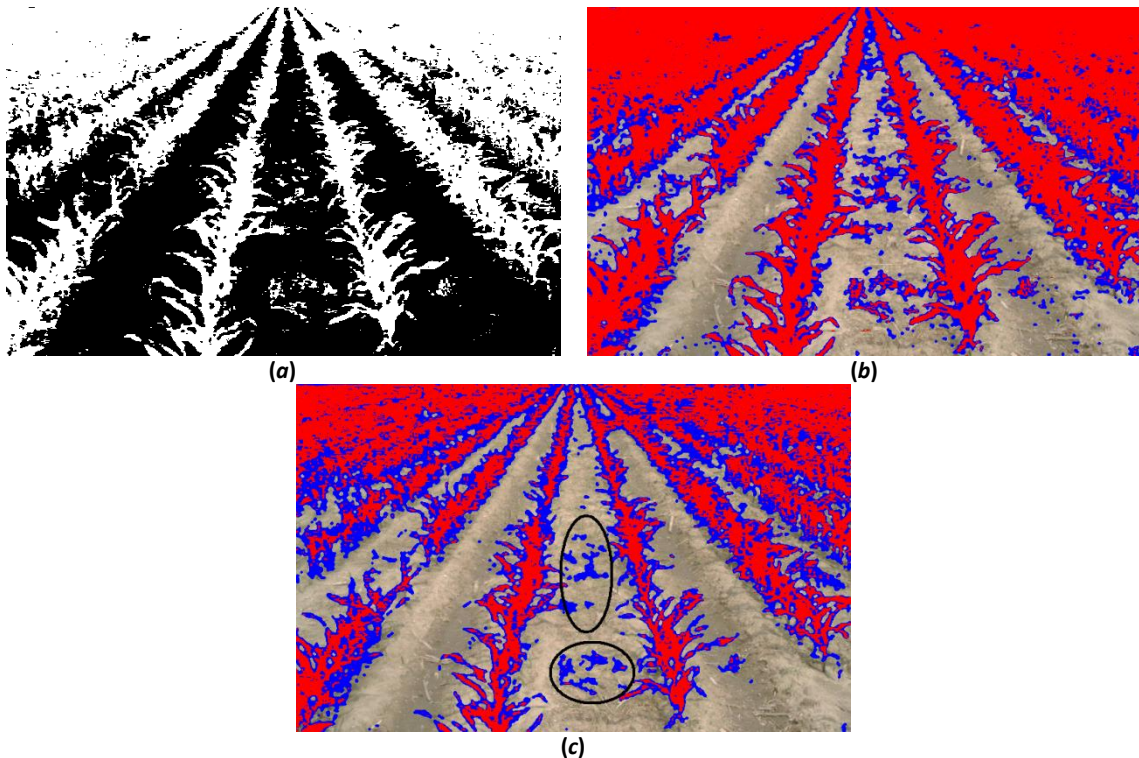


Figura 6-11. (a) Imagen binaria obtenida con el umbral c de la Tabla 6-3 conteniendo plantas enmascaradas y sin enmascarar en blanco; (b) imagen segmentada con SEA con plantas enmascaradas en azul y sin enmascarar en rojo; (c) imagen segmentada con DEM con plantas enmascaradas en azul y sin enmascarar en rojo.

Considerando todos los píxeles etiquetados como parte de plantas enmascarados y sin enmascarar, tanto para el Conjunto-1 como para el Conjunto-2, se han calculado los valores promedio para los tres canales espectrales RGB de la imagen de partida, obteniendo de este modo tres valores, uno para cada canal, representando cada uno la media de los valores de esos píxeles. Estos valores se obtienen tanto para el enfoque SEA y DEM como para el enfoque propuesto en Guerrero y col. (2012) consistente en el uso de Máquinas de Vectores Soporte (SVM) brevemente descrito al final de la sección 4.3; dichos valores muestran en la Tabla 6-4.

Métodos	SEA			DEM			SVM		
	R	G	B	R	G	B	R	G	B
Valores RGB									
Sin enmascarar	129.6	135.5	96.6	131.3	143.5	99.1	130.4	136.1	98.1
Enmascarados	146.1	142.9	111.1	147.8	144.7	109.3	147.0	143.6	112.0

Tabla 6-4: Media de las componentes RGB para plantas enmascaradas y sin enmascarar según SEA, DEM y SVM

En la tabla se muestra la media de los valores que toman los píxeles para cada canal (R, G y B) y cada uno de los métodos (SEA, DEM y SVM), tanto para los píxeles de las plantas enmascaradas, fila denominada ‘Enmascarados’, como para los píxeles de las plantas sin enmascarar, fila denominada ‘Sin enmascarar’. A partir de los resultados mostrados en la tabla se deduce que la componente espectral verde para las plantas sin enmascarar es mayor que la componente roja en los tres enfoques (SEA, DEM y SVM).

Esto significa que el nivel de verde es claro en este tipo de plantas, como era de esperar. En el caso de las plantas enmascaradas no sucede lo mismo, en ellas la componente espectral verde es menor que la componente roja en una pequeña cantidad. Esto verifica y respalda la hipótesis inicial respecto de las consecuencias derivadas de su enmascaramiento.

Con el fin de evaluar el SEA propuesto y determinar su rendimiento en comparación a DEM y SVM, se ha seleccionado de forma aleatoria el 20% de cada uno de los dos conjuntos disponibles (Conjunto-1 y Conjunto-2). Cada imagen fue analizada visualmente por un experto para identificar las malas hierbas y las plantas que constituyen el cultivo. El experto concentra su esfuerzo en identificar las plantas problemáticas, por tanto, aquellas que hemos denominado plantas enmascaradas o plantas ya afectadas por el tratamiento de herbicida. Las asignaciones incorrectas se marcan de forma manual, siendo corregidas o eliminadas de esta forma, generando así nuevas imágenes segmentadas, que pueden considerarse como mapas de verdad. La Tabla 6-5 muestra el porcentaje de éxito obtenido por cada uno de los enfoques para los dos tipos de imágenes analizadas.

Tasa de éxito (%)	SEA	DEM	SVM
Conjunto-1	93.4	86.2	93.1
Conjunto-2	91.8	85.3	89.9

Tabla 6-5: Porcentaje de éxito obtenido por los distintos enfoques para los dos conjuntos de imágenes

A partir de los resultados mostrados en la Tabla 6-5 podemos establecer que SEA presenta un mejor comportamiento que DEM y SVM, ya que el porcentaje de aciertos es mayor en los dos conjuntos analizados. Estos resultados pueden explicarse por el hecho de que los umbrales obtenidos por DEM muestran el comportamiento explicado anteriormente, donde DEM no identifica correctamente algunas de las plantas sin enmascarar, etiquetándolas como plantas enmascaradas cuando en realidad no lo son o no es capaz de identificar plantas enmascaradas, considerándolas parte del terreno. Además, SVM requiere una etapa de entrenamiento y aunque en nuestros experimentos se obtienen unos resultados similares a los obtenidos con SEA, el hecho de involucrar aprendizaje podría variar según el tipo de imágenes, consiguiendo resultados peores ya que depende fuertemente de las muestras usadas para el entrenamiento. Por el contrario SEA no requiere de una etapa de entrenamiento, que implica que se dedique un tiempo previo a la operación que se desea realizar, ni tampoco está sujeto a la fuerte dependencia de seleccionar unas imágenes lo suficientemente buenas para garantizar un

buen entrenamiento y que además proporcionen un número suficiente de muestras. A la vista de los resultados obtenidos se puede establecer que el enfoque SEA resulta apropiado para resolver el problema de identificar las plantas enmascaradas y sin enmascarar en entornos de exterior en campos de maíz, pudiéndose extender probablemente a otro tipo de cultivos, no necesariamente de maíz lo que le otorga un importante valor añadido.

6.1.3. Implementación del sistema RHEA

Como ya se comentó en el capítulo uno, el proyecto RHEA es uno de los proyectos de referencia dentro de los que se enmarca la presente investigación. Gracias al cual se ha partido de un problema real al que buscar una solución con la suficiente eficiencia para realizar tratamientos selectivos en campos de maíz mediante la identificación de las líneas de cultivo y las malas hierbas. En esta sección se analizan los resultados específicos relacionados directamente con RHEA.

Las imágenes fueron adquiridas durante los meses de abril, mayo y junio de los años 2013 y 2014 por las cámaras svx-4050CFLGEA y AXIS 211M, que forman parte del conjunto de sistemas instalados en los tractores autónomos utilizados en RHEA. En cada uno de los vehículos la cámara, dentro de su correspondiente carcasa, se encuentra situada en la parte superior del tractor a una altura de 2,10 metros desde el suelo y con una inclinación de 21° respecto de la vertical. La adquisición de las imágenes se realizó a lo largo de distintos días, obteniendo imágenes con diferentes estados de crecimiento del cultivo y bajo diferentes condiciones de iluminación. Además, debido a que el crecimiento de las malas hierbas no es uniforme en todo el campo se disponen imágenes conteniendo áreas con distintos niveles de infestación de malas hierbas. La Figura 6-12 muestra tres imágenes con tales características donde se identifican los distintos niveles de crecimiento del cultivo y los diferentes niveles de infestación de malas hierbas que sirven como ejemplo ilustrativo para mostrar la diversidad de situaciones existentes en los campos de cultivo utilizados durante los experimentos en el proyecto RHEA. En la Figura 6-12(a) se muestra una imagen donde el cultivo aparece en un estado de crecimiento inicial, mientras que en la Figura 6-12(b) se muestra un cultivo donde las plantas de maíz aparecen con un grado de crecimiento mayor. Tanto en la Figura 6-12(a) como en la (b) el nivel de infestación es relativamente reducido; por el contrario, en la Figura 6-12(c) se muestra una zona del campo donde el nivel de

infestación es elevado. Debido a que el procesamiento se realiza en tiempo real no es necesario el almacenamiento de las imágenes, si bien tanto las imágenes originales como las procesadas se almacenan con el fin de poder proceder a su posterior análisis de cara a comprobar la efectividad de los métodos de procesamiento desarrollados. Las imágenes originales se almacenan a color de 24 bits con una resolución de 2336 x 1752 píxeles para la cámara SVS con una resolución de 1280 x 1024 píxeles para la cámara AXIS, siendo en ambos casos guardadas en el espacio de color RGB en el formato BMP. Las imágenes ya procesadas por el algoritmo son almacenadas como imágenes a color de 24 bits y la misma resolución que las crudas pero guardadas en el espacio de color RGB en el formato JPEG. Todas las imágenes fueron adquiridas en distintas parcelas pertenecientes a la finca experimental de La Poveda, en Arganda del Rey (Madrid), propiedad del Centro de Investigaciones Agrarias del Consejo Superior de Investigaciones Científicas (CSIC).

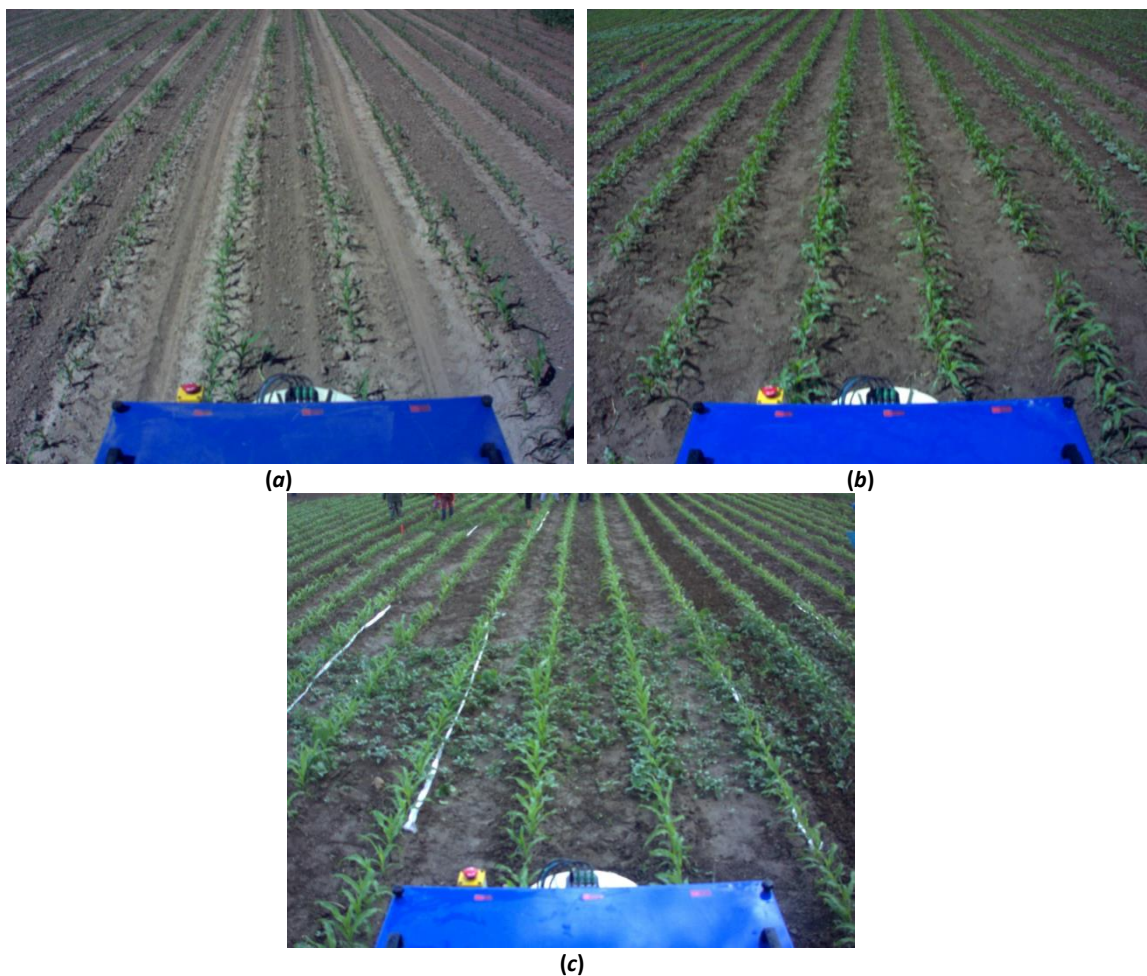


Figura 6-12. Imágenes con distinto nivel de infestación y crecimiento.

El método correspondiente al SEA para el tratamiento de las imágenes se desarrolló en el lenguaje de programación C mediante el entorno de programación

Visual Studio 2008 generándose una librería de enlace dinámico (DLL). Dado que el SEA para la detección de malas hierbas y cultivo forma parte del conjunto de procesos integrados en el sistema de procesamiento de alto nivel de RHEA, la librería se integró como tal bajo Labview. El computador para el procesamiento de las imágenes en tiempo real fue la plataforma cRIO, más exactamente el modelo 9082, descrito previamente.

La ROI definida en el capítulo cuatro, en la sección 4.2.2, abarca desde la columna 360 hasta la 1890 y desde la fila 650 hasta la 1269 de la imagen para la cámara SVS y desde la columna 120 hasta la 1160 y desde la fila 300 hasta la 719 de la imagen para la cámara AXIS.

El objetivo es identificar cuatro líneas de cultivo en cada una de las imágenes y la densidad de malas hierbas en la ROIT. En la Figura 6-13 se observa cómo se han detectado las líneas de cultivo mediante el SEA para las imágenes mostradas en la Figura 6-12.

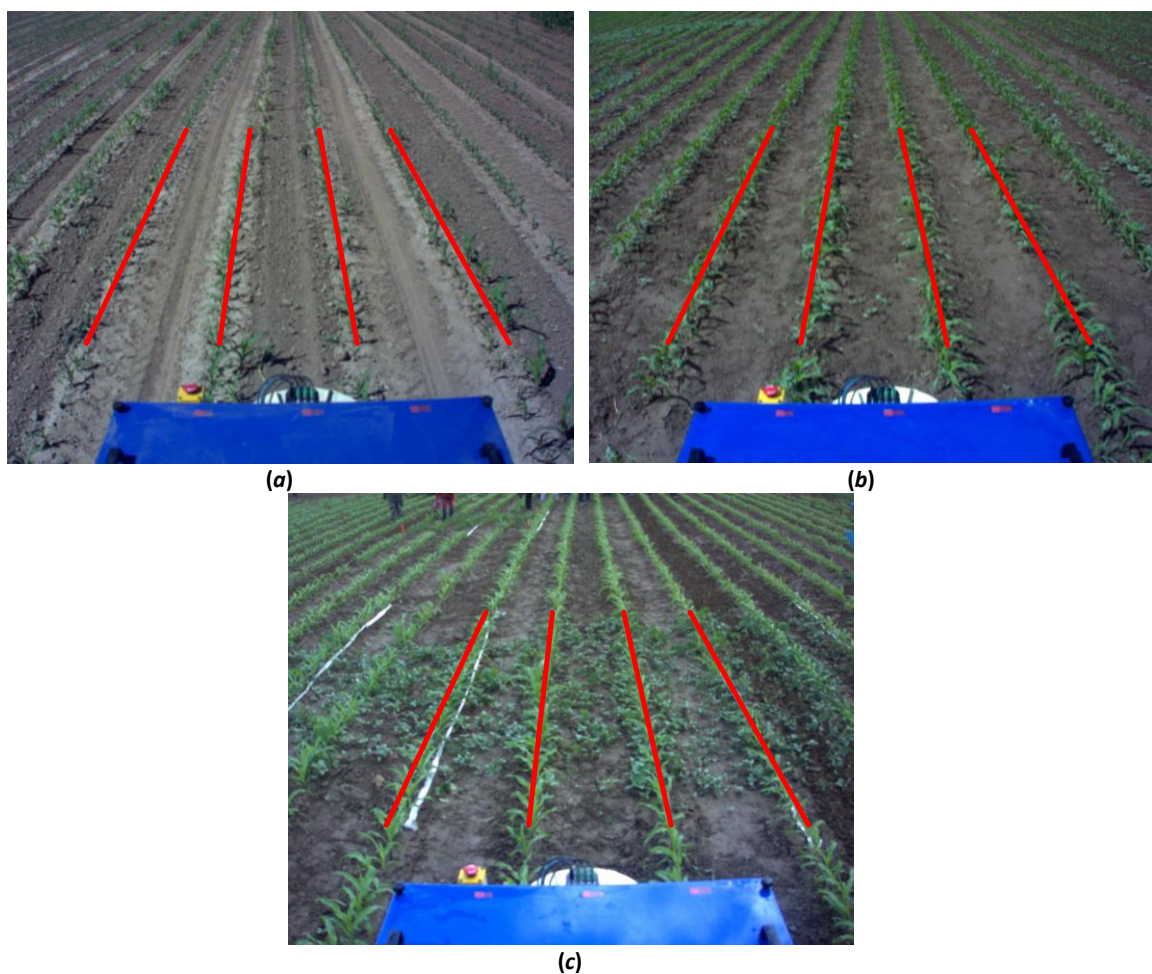


Figura 6-13. Imágenes en las que se han detectado las líneas de cultivo.

En la Figura 6-13 las líneas de cultivo están representadas como líneas rectas rojas. El modo elegido para delimitar la región en la que encontrar las líneas de cultivo fue mediante una IMU, en vez de usar una plantilla, que detecta los movimientos de la cámara debido a la navegación del tractor por el terreno no uniforme que es el campo de cultivo.

Tras detectar las líneas de cultivo, resta trazar el mallado que nos delimitará la ROIT proporcionando una matriz de valores asociada a cada una de las celdas del mallado. Como ya se indicó previamente en la sección 4.2.5, la ROIT se encuentra situada a 3,30 y 4,10 metros de la vertical de la cámara para los vehículos que llevan instaladas las cámaras SVS y AXIS respectivamente. En ambos casos es rectangular, el ancho cubre cuatro líneas de cultivo, equivalente a tres metros, el largo son dos metros y se encuentra centrada entre dos líneas de cultivo. El tamaño de cada una de las celdas que forman el mallado es de 25 centímetros de largo y 37,5 centímetros de ancho (la mitad de la distancia entre dos líneas de cultivo). En la Figura 6-14 se muestran las líneas de cultivo representadas como líneas rectas verdes, las malas hierbas como píxeles de color cian y el mallado delimitado por las líneas rectas azules y rojas. Las líneas rojas se encuentran situadas en el centro de los surcos (equidistante a dos líneas de cultivo consecutivas). Las líneas azules son paralelas a las líneas de cultivo sobre el terreno y desplazadas una distancia igual al margen de seguridad, mirar sección 4.2.5, y su finalidad es la de excluir el cultivo de las celdas que constituyen el mallado.

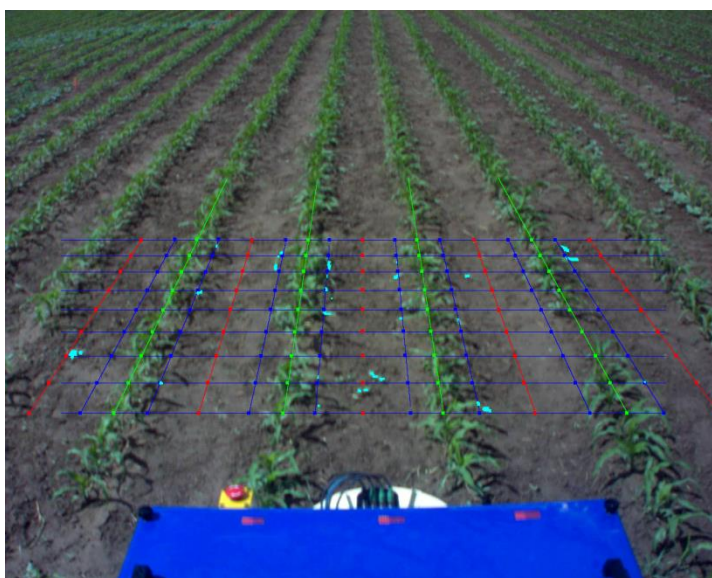


Figura 6-14. Imagen con la identificación de las líneas de cultivo, las malas hierbas y el mallado.

Debido a que se debe realizar un tratamiento de todo el campo, la ROIT de una imagen debe situarse justo a continuación sobre el terreno de la ROIT de la imagen anterior. Las ROIT de imágenes consecutivas no se deben solapar para evitar un doble tratamiento en la zona solapada ni tampoco deben existir zonas del campo fuera de las diferentes ROIT, ya que dichas zonas entre ROIT quedarían sin tratar. Gracias al GPS diferencial que se utiliza en las tareas de guiado se consiguió solapar las ROIT de imágenes consecutivas con errores de 5 cm de media. Suponiendo que la velocidad del vehículo es de 3 km/h, una velocidad más que razonable para realizar el tratamiento con los quemadores y conociendo que el largo de la ROIT es de 2 metros se dispone de 2,4 segundos para procesar cada una de las imágenes, a este tiempo hay que descontar 1 segundo necesario para encender y apagar los quemadores, quedando finalmente 1.4 segundos para el procesamiento de cada imagen. El tiempo que tarda el SEA en procesar la imagen varía dependiendo principalmente de dos factores, que son el estado de crecimiento del cultivo y el nivel de infestación de las malas hierbas existente. El valor medio del tiempo requerido para procesar una imagen se sitúa en torno a los 400 milisegundos, siendo la ejecución en el peor de los casos de 570 milisegundos. Por lo que se concluye que se dispone de un tiempo más que suficiente para que el SEA trabaje correctamente mientras el vehículo navega a lo largo del campo de cultivo realizando las correspondientes labores de tratamiento.

Capítulo 7

7. Conclusiones y trabajo futuro

7.1.Introducción

Tras el análisis y valoración de los resultados obtenidos se hace necesario plantear un balance global sobre las líneas de investigación seguidas, que desembocan en las conclusiones expuestas. Además, llegados a este punto se hace necesario considerar las líneas futuras hacia donde se dirigirán a partir de ahora los esfuerzos para mejorar en lo posible los desarrollos logrados con vistas a la mejora de los resultados y en la medida de lo posible, para la aplicación de los métodos en otros tipos de cultivo.

En la sección 7.2 se incluyen las conclusiones obtenidas para cada uno de los SEA desarrollados así como en lo que respecta a su adaptación en el proyecto RHEA. En la sección 6.3 se incluyen las líneas de actuación futura.

7.2.Conclusiones

7.2.1. Detección de líneas de cultivo y malas hierbas

Se ha propuesto un método para la detección de líneas de cultivo en campos de maíz con una alta densidad de malas hierbas por unidad de superficie. La estrategia propuesta se basa en tres procesos principales: a) segmentación de la imagen para detectar las plantas verdes existentes en la escena, b) umbralización doble basada en el método de Otsu para discriminar entre las malas hierbas y el cultivo de maíz y c) regresión lineal mediante mínimos cuadrados totales para estimar las líneas de cultivo. Las imágenes originales han sido capturadas en el espacio de color RGB y transformadas posteriormente a una imagen en escala de grises usando el índice de vegetación ExG. Después de esta transformación en la que la imagen se convierte en una imagen de un solo canal partiendo de una imagen de tres canales, el método de umbralización basado en Otsu permite separar las plantas de maíz y las malas hierbas del resto de elementos presentes en la imagen, diferenciándose a la vez entre el cultivo de maíz y las malas hierbas. Este proceso de diferenciación constituye una importante

contribución para los procesos en los que se quiere detectar y aislar en un dos conjuntos las malas hierbas y el cultivo. Por último se emplea regresión lineal, basada en mínimos cuadrados totales, que considera los píxeles que constituyen las plantas de maíz, con la finalidad de calcular las ecuaciones que representarán las líneas de cultivo. Se emplea el conocimiento previo sobre las restricciones geométricas impuestas por el sistema de visión embebido en vehículo agrónomo que navega por el terreno para construir una plantilla que guiará y facilitará el cálculo de las anteriores ecuaciones. El rendimiento del método ha sido probado en un amplio conjunto de imágenes, habiendo verificado su validez y eficacia para aplicaciones de tiempo real, a la vez que se ha comparado favorablemente con otros métodos ampliamente utilizados por la comunidad científica con la misma finalidad, como es el caso de la transformada de Hough.

Desde el punto de vista computacional el método ha sido diseñado mediante una arquitectura abierta, de este modo es posible añadir o remplazar nuevos módulos en un futuro.

7.2.2. Detección de plantas de cultivo y malas hierbas enmascaradas

Se ha propuesto un nuevo SEA para la segmentación de imágenes también en campos de maíz. El SEA se basa en tres principales etapas consecutivas donde la idea subyacente es la aplicación sucesiva de tareas para procesar la imagen de forma automática basándose en el conocimiento experto.

El SEA es capaz de identificar plantas (malas hierbas y cultivo) incluso cuando han sido impregnadas con materiales provenientes del terreno, tales como barro, pajas o pequeñas partículas sólidas. Esta impregnación ocurre por la irrigación artificial o por la caída del agua proveniente de las precipitaciones naturales de lluvia. En tales circunstancias la identificación de las plantas resulta ser una tarea difícil. Este método también sirve para la identificación y el monitoreo de las plantas en el post-tratamiento; basándonos en que tras el tratamiento, tanto en el caso de que sea mecánico como químico, las malas hierbas inician un progresivo proceso de degradación por pérdida de la actividad derivada de la función clorofílica, perdiendo su verdor natural a lo largo de la etapa de post-tratamiento. Los resultados obtenidos demuestran la validez del método para el objetivo pretendido, su eficacia ha sido también demostrada mediante la comparación con otros métodos existentes en la literatura.

Además, una vez identificadas las plantas tanto degradadas como no degradadas, las restantes partes existentes en la ROIT pertenecen al terreno y cubren un número de categorías relevantes ecológicamente hablando (Luscier y col., 2006). El SEA propuesto resulta fácilmente ampliable para abordar el problema del análisis de los materiales existentes en el terreno, y por consiguiente para análisis de suelos en campos de cultivo.

Como en el caso anterior, desde el punto de vista computacional el SEA también ha sido diseñado mediante una arquitectura abierta con la posibilidad de añadir o reemplazar nuevos módulos en un futuro.

7.2.3. Sistema RHEA

Se diseñó y desarrolló un sistema automático para AP con la finalidad de aplicar tratamiento selectivo de las malas hierbas en campos de maíz mediante la detección de líneas de cultivo y malas hierbas a través un sistema de visión encargado del procesamiento de las imágenes. Este sistema, basado en cámaras AXIS 211M y SVS-4050CFLGEA, se encuentra instalado en los tractores que conforman la flota de vehículos autónomos en RHEA junto con el resto de sistemas sensoriales para navegación, detección y actuación. El sistema de detección desarrollado es capaz de trabajar en tiempo real, lo que implica que captura las imágenes y procesa las imágenes bajo las especificaciones del proyecto RHEA, con los vehículos navegando a razón de 3 km/h de media. Los resultados han sido verificados y validados dentro del proyecto RHEA mediante los procedimientos establecidos al efecto dentro de los proyectos europeos de esta naturaleza habiéndose realizado la demostración final el 21 de Mayo de 2014 de forma satisfactoria.

Las etapas del procesado son básicamente tres: segmentación, umbralización y detección de líneas. En la segmentación se tenían como posibles opciones dos índices de vegetación, el ExG y el COM, cada uno de ellos usado en uno de los SEA detallados en el capítulo cuatro. Se optó por utilizar el ExG por ser el que menos operaciones costosas (multiplicaciones, divisiones y potencias) requiere y por lo tanto el que menos tiempo exige, ya que ambos índices de vegetación mostraron comportamientos similares en los SEA.

7.3.Trabajo futuro

Si bien los resultados de la presente investigación han sido satisfactorios, tanto a nivel de métodos desarrollados como en lo que respecta a su integración en RHEA, todo sistema es susceptible de mejora. De este modo el trabajo futuro se presenta como posibles mejoras a realizar sobre el sistema global considerando cada una de sus partes por separado.

Las siguientes partes constituyen un objetivo de mejora de cara a un trabajo futuro: captura de imágenes de calidad, segmentación de la imagen y detección de las líneas de cultivo.

Para la captura de una imagen de calidad mediante el control del tiempo de exposición es posible desarrollar otro método que no utilice como referencia unos valores fijos para los tres canales espectrales ya que, si bien el algoritmo actual trabaja adecuadamente capturando imágenes de calidad, al usar unos valores fijos como referencia, se ve ligeramente afectado por el estado de crecimiento de las plantas, así como el tipo y estado de la tierra. En este caso se necesitaría garantizar que el procedimiento tuviese una carga computacional baja, para tiempo real, se determina que el tiempo de proceso debe ser inferior a 100 ms, ya que se deben capturar al menos 10 imágenes por segundo para cubrir todo el campo de cultivo con las ROIT evitando al máximo el solapamiento.

En lo relativo a la mejora de la calidad de la imagen otra posibilidad consiste en mantener el mismo algoritmo si bien analizando los valores espectrales en cada jornada de trabajo mediante un nuevo sistema experto, de modo que cuando el tractor comience la jornada de trabajo, situado ante el campo de cultivo y antes de comenzar a trabajar, capture una secuencia de imágenes mediante las cuales el nuevo sistema experto calcule unos valores espectrales de referencia para las condiciones actuales. Si las condiciones actuales cambian durante la jornada de trabajo, se repetiría esta misma operación, dotando al sistema de una capacidad inteligente para realizar tal proceso de forma automática.

Otro aspecto en el que es posible la mejora del SEA es la detección de las plantas verdes presentes en la escena. El paso relativo a la segmentación de las imágenes resulta de vital importancia en el proceso ya que a partir de esta

transformación se identifican las líneas de cultivo y se obtienen las matrices con los valores que representan la densidad de mala hierba en la zona a tratar. En la presente investigación se abordó el uso de índices de vegetación para segmentar la imagen, descartándose el uso de la banda espectral NIR. La inclusión de esta banda espectral para la detección, podría aportar información relevante, tal y como se constata en la bibliografía especializada. No obstante, habría que buscar el equilibrio necesario para lograrlo sin aumentar los costes de procesamiento ni de producción del sistema de cara a posibles futuras comercializaciones.

En los SEA desarrollados se distingue solamente entre dos tipos de plantas, por un lado las plantas de maíz y por el otro las correspondientes a malas hierbas. Además esta diferenciación se realiza atendiendo al criterio del color, empleando los tres canales del espacio de color RGB. El objetivo en el proyecto RHEA era únicamente detectar las malas hierbas, sin llegar a realizar una clasificación de los tipos de plantas detectadas. En este sentido, la mejora de futuro vendría por desarrollar nuevos métodos capaces de discriminar los tipos de plantas existentes, bien por diferencia de tonalidades de color, textura, forma u otras características relevantes. Esto permitiría aplicar tratamientos todavía más específicos en función de la naturaleza de las plantas detectadas. De este modo se conseguirían distinguir varios tipos de malas hierbas y también se aumentaría el porcentaje de acierto al poseer una mayor información durante la identificación. Como el sistema debe seguir operando en tiempo real, lo más probable es que haya que paralelizar algoritmos o incluso utilizar alguna tarjeta gráfica.

Un parte de especial importancia que se podría estudiar para obtener una mejora sería la umbralización, obteniendo un nuevo método automático con el que generar los umbrales (Avci y Avci, 2009). Por otra parte también se podría añadir una base de datos con la que realizar una segmentación de la imagen basándose en el conocimiento previo acumulado (González-Andújar, 2009).

Finalmente, cabe la posibilidad de adaptar los SEA para que sean operativos en otros tipos de cultivo diferentes al maíz. Se podrían utilizar en cultivos donde la planta tuviera un color distinto al verde, de modo que se deberían detectar otros colores distintos del verde. Asimismo se podría aplicar a cultivos donde las plantas no han sido sembradas siguiendo líneas rectas, debido principalmente al intento de optimizar el espacio de laboreo o la topografía del terreno. En algunos lugares, principalmente

lugares montañosos, debido a que el terreno cultivable se encuentra situado en laderas, es posible que los cultivos se siembren describiendo líneas curvas para tratar de evitar en la medida de lo posible que el vehículo avance en el sentido de la máxima pendiente. En estos casos, sería necesario desarrollar métodos para la detección de líneas curvas, probablemente polinómicas donde los parámetros a determinar sean los coeficientes del polinomio.

En cualquier caso, toda modificación debe tener en cuenta tanto los requisitos de tiempo real de los sistemas y su adaptabilidad a otras aplicaciones, así como mantener los costes de producción dentro de los límites razonables de cara a la comercialización de los desarrollos.

ENGLISH SUMMARY

Chapter 8

8. English summary

8.1.Introduction

8.1.1. Antecedents

The continued rise and development of computers has allowed the automation of many processes effectively. The computer has evolved from a luxury item available to a minority, intended for very specific purposes and related to research, to be an essential part of daily life at all levels.

Agriculture has been no less, one of the sectors where the use of computers progress gradually. Modern farm machinery is equipped with high performance computing devices, achieving greater efficiency in the final production, while less effort is achieved by farmers in agricultural tasks.

Traditionally, when an agricultural task has been necessary to perform, it has been based on the application of techniques throughout the whole area of land available for it. Consider for instance the application of herbicide in a crop field in order to eliminate existing weeds in the field. It is usual to see a tractor applying herbicide uniformly throughout the fields with the corresponding equipment. Consider that in the field can be patches of weeds with different densities or even areas without weeds. So the treatments are performed equally in all areas of the field, regardless of need required for each particular area. Ideally, the high-density areas need a more aggressive treatment than low-density areas, while the areas without weeds no require the application of any herbicide dose.

Returning nonspecific herbicide treatments, a way to treat the field uniformly, without a variable rate application, is through a conventional boom applying a chemical product, mixed with water from a reservoir, using nozzles mounted in the boom and a valve, which regulates the fluid pressure. The control of this valve is not changed while

the work is done, so the dose is uniform throughout the field whenever a constant vehicle speed is maintained. It is normal to take advantage of the vehicle speed or by the revolution of the motor, by controlling the valve to apply a variable dose treatment as required. The operator manually modifies the speed of the vehicle or revolutions to increase or decrease the amount of dose, which is applied, depending on his visual opinion. This is extrapolated to any other similar agricultural activity, like fertilization, irrigation or other treatments.

The non-differentiation of areas requiring different treatments in the field involve a number of shortcomings which are summarized as: a) higher cost of production wasting herbicides, fertilizers, water or other items used in the treatments, becoming worsen when they are extremely expensive; b) increase the working time spent on unnecessary treatments; c) greater negative environmental impact, because some chemicals used in crop treatments are harmful to the environment, seeping through the soil layers, reaching to contaminate streams or groundwater aquifers.

The fact of reducing the amount of chemicals used in the treatment, by selective application, does not imply a decrease in the effectiveness of treatment and hence on the quality and quantity of the final product, since the idea is to apply the correct doses where they are needed (Earl et al., 1996; Krishna, 2013). The advantages and benefits listed above makes it tends to specific treatments in the field of what is known as Precision Agriculture (PA), where computers play an extraordinary role in the automation of the processes involved.

The new way of farming, by PA techniques, offers new possibilities that have been considered for some years (Srinivasan, 2006; Stafford, 2000) and now they are in constant evolution and development as it will be discussed throughout this work. While the traditional agriculture was based on no specific treatments, the PA performs a Site-Specific Crop Management (SSCM). The tasks performed under the perspective of the PA facilitate the work of operators, impacting positively on improving their quality of life without compromising productivity.

Variable-Rate Applications (VRA) can be used for achieving SSCM (Zhang et al., 2010; Dammer, 2010; Dillon and Kusunose, 2013; Liakos et al., 2013), that are able to differentiate areas which must act in different ways. There are several available VRA

methods, some of them use GPS information. Basically they can be grouped into the following three categories: manual, map-based and sensor-based technologies.

In manual technologies, the control of the variable rate dose is performed by the operator. This operator is responsible for varying the rate dose while the vehicle is working in the corresponding task. The decisions made by the operator about how to regulate the rate dose may be based only on his visual criterion or taken with the help of other elements like maps or sensors.

Maps based technologies adjust the amount of administered dose or work volume to be done based on a digital map, also called prescription map. Using the current vehicle position in the field, obtained by a GPS receiver, and a prescription map containing the desired dose or specific tasks for each zone, the concentration of the dose or the kind of action is set depending on the position where the working system is found. The maps are created by field measurements prior to treatment or task. These measurements are intended to obtain values for different properties or characteristics, depending on the purpose for which you want the prescription map. Some of these properties or characteristics that can be measured depending on the kind of goal to be achieved, they are the following (Christensen et al., 2005; Piikki et al., 2013; Serrano et al., 2013; Silva and Molin, 2013; Weis et al., 2013): a) color and texture of the plants or soil; b) kind of soil, moisture, pH, electrical conductivity; d) crop yield; e) nutrients nitrogen level. Some approaches focus only one feature to create the prescription map, while others approaches combine information from different sources, including images taken from aerial vehicles in specific areas (Berni et al., 2009a, Torres-Sánchez et al., 2013) or satellite images (Gómez-Candon et al., 2012; Meng et al., 2013).

Regardless of the number of the characteristics which are used, the system must be able to determine the position of the vehicle applying the treatment, associating the rate dose with that position. With these associations rate dose and position, a prescription map is created which will be used later.

Vehicles applying treatments in-situ and equipped with sensor-based technologies do not require a map or a positioning system. In this case, sensors onboard of the vehicles are capable of measuring soil properties or crop characteristics. This information is provided to the controller, after passing through a computer where it is

processed. Through this continuous feedback, the control system is responsible for deciding the action to perform on the field and transmit orders to the tools that finally apply treatment. Nevertheless, although these sensor-based systems do not require a positioning system or prior information, when a positioning system is available, data can be stored georeferenced. This information can be used to create a prescription map to be used in future operations as a database that records the evolution over the years.

Depending on the agricultural activity to be performed, VRA-based operations can be classified amongst others in: seeding (Dillon, 2013; Heege, 2013; Pitla and Shearer, 2014), weed control, molds or fungi (Dammer, 2010; Sökefeld, 2010; Wu et al., 2011; Andújar et al., 2013; Montalvo et al., 2013), application of lime (Adamchuk et al., 2004, Havlin and Heininger, 2009), application of phosphorus (Mallarino and Wittry, 2006), application of nitrogen (Scharf et al., 2009; Chan, 2012; Kusunose and Dillon, 2013; Liakos et al., 2013) or irrigation (Al-Kufaishi et al., 2006).

All these operations, based on sensory systems, allow to access to the required information with the aim of specific treatment. Sensors in AP are essential for operations and selective treatments. These sensors identify and detect the elements and areas on which to act based on the available information. Within sensors, machine vision system has been widely used in AP for different purposes, including detection; moreover it is the core of the research described in this thesis.

In addition to detection systems for selective treatments, one of the contributions and developments within the AP are automatic guidance systems for vehicles. An automatic guidance system for agricultural vehicles (Al-Hamed, 1996; Wilson, 2000; Shockley and Dillon, 2008; Heege, 2013; Zhou et al., 2013) could eliminate the need for a farmer driving the vehicle without visibility restrictions or time limits. The vehicle could work for hours continuously without requiring the rest of the driver or his replacement by another farmer. The work would not be affected by environmental factors, unless they impede the realization of agricultural work, because it could use a navigation satellite system (GNSS) such as GPS (2014), GLONASS (2014) or the future European Galileo (2014) civil system. These systems could delimit the field, tracing the path to follow, along with its critical points (the place where tractor entries and exits, pylons, large rocks on the field, elements of irrigation, etc.) which are necessary for safe navigation and to perform agricultural tasks efficiently. The

positioning of the vehicle can be accomplished by artificial vision, which can complement the GNSS systems (Spacek and Burbridge, 2007; Subramanian et al., 2006; Ming et al., 2012.).

Another motivation for developing systems with autonomy or semi-autonomy guidance arises from the need to avoid the operator to make continual adjustments in the direction of the tractor while towing the working tool to achieve an acceptable level of performance. The operator performs two basic operations namely: guidance of the vehicle and the handling of the tools or machinery that holds together. Operator function has not changed substantially in decades. However both, tractors and machinery, used have evolved during the same period. The tractors have increased power, strength, speed, as well as performance. Because of this and the use of ever larger and more complex tools, the control of the tractor and machinery is becoming more and more complex for an operator in the cab of the tractor, since it is difficult to visually determine where the tractor and the implement are positioned, if there are overlapped or untreated areas.

Thus the need to include devices capable of addressing this problem efficiently is established, while from the point of view of the AP new techniques and procedures for the effective detection and navigation are required, involving sensors and actuation systems working together with high efficient computers, which are required for data processing and decision making.

8.1.2. Problem description

Within the wide world of AP in this research we focus on vision systems for detecting crop rows and weeds in maize fields for identifying areas of selective treatment, as well as vehicle guidance tasks. The development of image processing methods constitutes the core of the research. These tasks are part of the RHEA and AGL2011 projects, which are described in detail below in section 8.1.3.1.

The research is conducted in large enough maize fields. It should be noted that from the point of view of the effectiveness of automatic systems, the dimension of the fields must ensure the operability, while it must be worthwhile from the point of view of investment and productivity. Have in mind that such systems are designed for longer performances in time with no or minimal human supervision.

In the initial stages of maize growth, its development is influenced very significantly by the presence of weeds. These pose strong competition because weeds take advantage of soil nutrients at the expense of maize. If they are not removed they continue growing while preventing maize plants received sunlight. Moreover, in these early stages weeds grow faster than maize and avoid the crop grow as it would develop in the absence of weeds. To avoid the above problems is necessary to eliminate weeds in early growth stages, while maize plants are preserved.

In the projects AGL2011 and RHEA the main problem is identifying existing weeds in the maize field to apply a selective treatment with the required dose according to their density in the field. The maize crop falls within the so called wide row crops, so that the spacing between crop rows is 75cm. The spaces between every two lines can be covered with weeds. The vision system is responsible for the detection, localization and quantification in order to apply the precise dose where it is needed. In this process, the identification of crop row is crucial because weeds are placed between them, so that the identification is the first step of the process. Once they are located, the next step is to identify and quantify the weeds between the crop lines. Additionally, crop rows, once detected, are used for precise vehicle guidance correcting errors in the path following due to GPS inaccuracies or other circumstances, such as when the vehicle tends to accumulate deviations by tensile strength mechanical actuation elements.

The way to eliminate weeds in RHEA is done in two ways, one of them by a process of ploughing between rows and by flames burning and weakening weeds without affecting the maize plants (Fontanelli et al., 2011). Depending on the weeds densities in a localized specific area, a flame with different intensity is applied. These flames are produced by burners fueled with Liquefied Petroleum Gas (LPG). The higher the concentration of weeds in a specific area, the greater the pressure delivered to the burner and therefore the intensity of the flame. These flames have no negative effect on maize but the weed is destroyed or weakened. The detection process through a vision system aboard a tractor, for detecting both crop lines and weeds, is not exempt of problems and difficulties, especially considering the work is carried out in outdoor environments where there are several problems related to lighting and other additional factors, as well as the fact that the tractor is in continuous movement performing mechanical traction tasks. Procedures based on computer vision should be robust enough to work in adverse lighting conditions, having days with different levels of

illumination (sunny, cloudy), variable lighting throughout the day (morning, noon, afternoon) and even highly variable with sudden clouds appearance. Moreover, infrared and ultraviolet radiations affect the image quality. Such systems must also be designed to deal with movements produced by the irregularity of the land where the tractor moves, and deviations from its initial planning path.

Furthermore, regarding the detection of vegetation, which normally occurs during the treatment process, both crop and weeds are green, which means that for the vision sensor, based on the capture of the spectrum components from the scene are indistinguishable except for certain nuances that can be exploited. Another problem, also derived from the spectral nature, is the fact that the leaves of maize plants and weeds sometimes appear partially covered by a film of mud from the ground due to heavy rainfall or as a result of watered terrain. In this situation the plant identification becomes difficult because of the vegetation is masked for the sensor, having hue closest to red, associated with the clayey nature of the soil.

In the current research has been considered the above factors, developing strategies and procedures to solve such problems, which are the backbone of the research to detect crop rows and weeds.

Additionally, the process that addresses the identification of masked plants against non-masked has been designed to be able to discriminate between plants with high chlorophyll activity and those with their chlorophyll activity diminished or non-existent. This occurs after a treatment is applied for removing weeds. The weeds start a dying process, decreasing its chlorophyll activity, which leads to alter its original greenness to other tonalities of the dried plants. With this discrimination is possible to determine the efficiency of the treatment performed, which is an extra contribution of the investigation result.

8.1.3. Motivation and goals

8.1.3.1. Motivation

The current research has its origin in the research projects listed below, framed within the research group ISCAR (2014), officially recognized by the Complutense University of Madrid since 2006. Few of them in collaboration with other research groups, both national and international, as reflected below.

1. Research activity begins in the RHEA (2014) funded by the European Community under the Seventh Framework Programme for research and technological development (FP7) covering four years 2010-2014 under the title Robot Fleets for Highly Effective Agriculture and Forestry Management NMP-CP-IP 245986-2 RHEA. RHEA is constituted by a consortium of nineteen working groups belonging to fifteen organizations from eight European countries with a budget of 8.69 million euros.

RHEA is focused on the design, development, and testing of a new generation of automatic and robotic systems for both chemical and physical effective weed management in agriculture. The objective is to design three prototypes that are part of a fleet of robotic tractors and coordinated in conjunction with UAVs. Each one of robotic prototype is scheduled to acts in three fields: maize, cereals and olive groves.

The research developed and presented in this dissertation focuses on the tractor dedicated to maize, which is equipped with a vision system installed on its top which carries out the identification process of crop rows as well the estimation of the weed density existing between the crop rows.

2. The research is conducted simultaneously in the framework of the project titled *Sistema de inspección terrestre en vehículos autónomos y su aplicación efectiva a la detección de malas hierbas y su control localizado* (AGL2011-30442-C02-02), which belongs to the National Plan R + D + i for the triennium 2011/2014, AGL (2011). It is a coordinated project in collaboration with the Centre for Automation and Robotics, CSIC-Technical University of Madrid (UPM) mixed center, as well as the Institute for Sustainable Agriculture, belonging to CSIC also. The primary objective is the location of the weeds in the fields of maize and cereal.
3. R+D+i National Plan project entitled *Sistema de Vigilancia, Búsqueda y Rescate en el Mar mediante colaboración de Vehículos Autónomos Marinos y Aéreos* (*System for Surveillance, Search and rescue in the sea by means of collaboration of autonomous marine and air vehicles*) (DPI2009-14552-C02-01), DPI (2009), extended to end of 2013. This project is about coordination of unmanned surface vessel (USVs) and unmanned aerial vehicles (UAVs) to get a fast response in case of contingencies in the sea. The scenario supposes search and tracking with the

following vehicles: a) surface vessels, b) unmanned aerial vehicles and c) a mother ship that supports to the other vehicles.

8.1.3.2. Goals

From the considerations expressed in the previous section, the expectations and needs required in previous projects, the following research objectives are proposed:

1. State of the art review about precision agriculture using methods and procedures based on computer vision systems.
2. Design of methods to control the exposure time of the camera sensor so as to achieve capture images of sufficient quality for further analysis and processing.
3. Design of robust strategies for outdoor image segmentation in real time, with special emphasis on agricultural images.
4. Design image segmentation methods to subsequently develop thresholding methods to discriminate between maize and weeds.
5. Design methods to establish the location of crop rows.
6. Check that all the above processes can be carried out in real time, with the vision system onboard of the agricultural vehicle.
7. Design procedures to identify the effectiveness of the methods used in the elimination of weeds by the analysis of the plants being dried in a post-treatment stage.
8. Identify future lines of research to greatly improve both the methods developed for the current crop and for application to other possible types of different crops.
9. Disseminate research results through scientific publications or other mechanisms.

8.2. Automatic Expert System for crop row and weed detection

8.2.1. Introduction

8.2.1.1. Problem statement

The increasing development of robotics equipped with machine vision sensors applied to precision agriculture (PA) is demanding solutions for several problems. The robot navigates and acts over a site-specific area of a larger farm (Davies et al., 1998), where one important part of the information is supplied by the vision system.

An important issue related with the application of machine vision methods is that concerning the crop row and weed detection, which has attracted numerous studies in this area (Burgos-Artizzu et al., 2009; Onyango & Marchant, 2003; Sainz-Costa et al., 2011; Tellaeché et al., 2008a, 2008b). This will allow site-specific treatments trying to eliminate weeds and to favor the growth of crops. The problem becomes highly complex when high weed pressure exists.

The vision system consists of a CCD-based calibrated camera with known intrinsic parameters, i.e. focal length, radial lens distortion, CCD sensor sizes and pixel resolutions. The camera is located in front of the agricultural vehicle, inclined with a tilt angle (pitch) and at a high from the ground. Yaw and roll angles are also known. This allows determining the rotation and translation matrices defining the extrinsic parameters. Thus, areas in the field can be identified onto the image plane and vice-versa. This means that given any element in the image we can determine its relative positioning on the field.

The vehicle navigates on a real terrain presenting irregularities and roughness where the spacing of crop rows is known. This produces vibrations and also swinging in the pitch, yaw and roll angles. Because of the above, the crop rows are not projected onto the image at the expected locations according to the system geometry. On the other hand the discrimination of crops and weeds in the image is a very difficult task because their Red, Green and Blue spectral components display similar values. Thus, high weed pressure, similar spectral components and unexpected projections, all together make the crop row detection a complex problem that is specifically addressed in this study.

Thus, the problem is to locate accurately the crop rows in images of maize fields with high weed pressure. These images are captured by a camera aboard a vehicle in movement. To achieve this goal, we propose a new strategy that makes use of previous knowledge: (a) the specific arrangement of crops (maize) in the field and (b) image perspective projection and system geometry. We apply an automatic image segmentation process for greenness extraction and a double thresholding, based on the Otsu's method. After this, a total least-squares linear regression approach allows the automatic adjusting of the corresponding straight line equation to each crop row.

Crop row location has been considered already in the literature, the method proposed in this study gains advantage over existing approaches because it has been designed to achieve high effectiveness in real-time applications in PA for maize fields with high weed pressure.

8.2.1.2. Revision of methods

Several strategies have been proposed for crop row detection. Fontaine and Crowe (2006) tested the abilities of fourth line detection algorithms to determine the position and the angle of the camera with respect to a set of artificial rows with and without simulated weeds. These were stripe analysis, blob analysis, linear regression and Hough transform. Below is a revision of methods grouped into eight categories.

Methods based on the exploration of horizontal strips: Søggaard and Olsen (2003) apply RGB color image transformation to grayscale. This is done by first dividing the color image into its red, green and blue channels and then by applying the well-tested methods to extract living plant tissue described in Woebbecke et al. (1995). After this, the gray-scale image is divided into horizontal strips where maximum gray values indicate the presence of a candidate row, each maximum determines a row segment and the center of gravity of the segment is marked at this strip position. Crop rows are identified by joining marked points through a similar method to the one utilized in the Hough transform or by applying linear regression. Sainz-Costa et al. (2011) have developed a strategy based on analysis of video sequences for identifying crop rows. Crop rows persist along the directions defined by the perspective projection with respect the 3D scene in the field. Exploiting this fact, they apply gray scale transformation based on the approach proposed by Ribeiro et al. (2005) and then the image is binarized applying a thresholding technique. Each image is divided into four horizontal strips. Rectangular patches are drawn over the binary image to identify patches of crops and rows. The gravity centers of these patches are used as the points defining the crop rows and a line is adjusted considering these points. The first frame in the sequence is used as a lookup table that guides the full process for determining positions where the next patches in subsequent frames are to be identified. Hague et al. (2006) transform the original RGB image to gray scale. The transformed image is then divided into eight horizontal bands. The intensity of the pixels across these bands exhibits a periodic variation, due to the parallel crop rows. Since the camera

characteristics, pose and the crop row spacing are known a priori, the row spacing in image pixels can be calculated for each of the horizontal bands using a pinhole model of the camera optics. A band-pass filter can then be constructed which will enhance this pattern, and has a given frequency domain response. In images with high weed pressure, weeds and crop appear together forming a unique patch and gravity centers and other strip features are incorrectly identified.

Methods based on the Hough transformation: According to Slaughter et al. (2008), one of the most commonly used machine vision methods for identifying crop rows is based upon the Hough (1962) transform. It was intended to deal with discontinuous lines, where the crop stand is incomplete with gaps in crop rows due to poor germination or other factors that result in missing crop plants in the row. It has been intended for real-time automatic guidance of agricultural vehicles (Astrand & Baerveldt, 2005; Hague et al., 1997; Leemans & Destain, 2006; Marchant, 1996). It is applied to binary images, which are obtained by applying similar technique to the ones explained above, i.e. RGB image transformation to grayscale and binarization (Tellaeché et al., 2008a, 2008b, 2011). Gée et al. (2008) apply a double Hough transform under the assumption that crop rows are the only lines of the image converging to the vanishing point, the remainder lines are rejected, additional constraints such as inter-row spacing and perspective geometry concepts help to identify the lines. It is required to determine the threshold required by the Hough transform to determine maximum peaks values (Jones et al., 2009a, 2009b) or predominant peaks (Rovira-Más et al., 2005). Depending on the crop densities several lines could be feasible and a posterior merging process is applied to lines with similar parameters (Tellaeché et al., 2008a, 2008b, 2011). Ji and Qi (2011) report that Hough transform is slow due to the huge computation, they propose a randomized Hough transform to reduce computational time. Some modifications have been proposed to improve the Hough transformation such as the one proposed in Asif et al. (2010), which is applied only to those points which are edge points along the crops. But this requires the application of techniques for edge extraction. Also the randomized Hough transformation has been proposed with this goal (Ji & Qi, 2011). It is intended to avoid redundant computations in the Hough transform. It operates iteratively by randomly sampling a set of points to compute a single localization in the Hough space. Since two pixels are trivially collinear, the parameters of the line on which they lie can be

estimated. These parameters are used to increment the accumulator cell in the Hough space. The Hough transform is computationally expensive and the randomized Hough transform requires selecting pairs of points to be considered as forming a line, i.e. pairs of points belonging to a crop row. This becomes unfeasible in images with a high weed pressure because crop rows are often indistinguishable and pairs of points cannot be identified as belonging to a given crop row.

Vanishing point-based: Pla et al. (1997) propose an approach which identifies regions (crops/weeds and soil) by applying color image segmentation. They use the skeleton of each defined region as a feature to work out the lines which define the crop. The resulting skeletons of each region can be used as curves which define the underlying structure of the crop and to extract the straight lines where the plants and soil rows lie. Segments in the skeletons are defined as chains of connected contour points and they must be of a defined length. This allows selecting candidate lines that meet the vanishing point for crop row detection. The vanishing point is detected using previous information about the vanishing point found in the previous images, performing a sort tracking on the vanishing point. This process is highly dependent of skeletons, which are not always easy to extract, particularly in images with a high weed pressure.

Methods based on linear regression: Some of the techniques above apply this approach. Billingsley and Schoenfisch (1997) reported a crop detection system that is relatively insensitive to additional visual ‘noise’ from weeds. They used linear regression in each of three crop row segments and a cost function analogous to the moment of the best fit line to detect lines fitted to outliers (i.e., noise and weeds) as a means of identifying row guidance information. As mentioned above, Søggaard and Olsen (2003) also apply linear regression. A feasible approach applicable when pixels belonging to crop rows are well separated from weeds. Its effectiveness is highly affected by pixels belonging to weeds. Therefore, linear regression is only feasible if weeds and crops have been separated. In our proposed approach we separate weed and crops thank to the double Otsu’s approach, making the linear regression a feasible strategy.

Stereo-based approaches: Kise et al. (2005) or Kise and Zhang (2008) developed a stereovision-based agricultural machinery crop-row tracking navigation system. Stereo-image processing is used to determine 3D locations of the scene points

of the objects of interest from the obtained stereo image. Those 3D positions, determined by means of stereo image disparity computation, provide the base information to create an elevation map which uses a 2D array with varying intensity to indicate the height of the crop. This approach requires crops with significant heights with respect the ground. Because in maize fields during the treatment stage the heights are not relevant, this approach becomes ineffective in our application. Rovira-Mas et al. (2008) have applied and extended stereovision techniques to other areas inside PA. As before, only feasible if crops or weeds in the 3D scene display relevant heights and these heights are different in crops and weeds.

Methods based on blob analysis: This method finds and characterizes regions of contiguous pixels of the same value in a binarized image (Fontaine & Crowe, 2006). The algorithm searches for white blobs (inter-row spaces) of more than 200 pixels, as smaller blobs could represent noise in the crop rows. Once the blobs were identified, the algorithm determines the angle of their principal axes and the location of their center of gravity. For a perfectly straight white stripe, the center of gravity of the blob was over the centerline of the white stripe, and the angle was representative of the angle of the inter-row spaces. The algorithm returned the angle and center of gravity of the blob closest to the center of the image. Identification of blobs in images infested with high weeds pressure becomes unfeasible.

Methods based on the accumulation of green plants: Olsen (1995) proposed a method based on the consideration that along the crop row appear an important accumulation of green parts in the image. The image is grayscale transformed where green parts appear clearer that the rest. A sum-curve of gray levels is obtained for a given rectangular region exploring all columns in the rectangle. It is assumed that vertical lines follow this direction in the image. Obviously, the images are free of perspective projection because they are acquired with the camera in orthogonal position. A sinusoidal curve is fitted by mean of least squares to the sum-curve previously obtained. Local maxima of the sinusoid provide row centers locations. Again in images with high weed pressure this approach becomes unfeasible.

Methods based on frequency analysis: Because crop rows are vertical in the 3D scene, they are mapped under perspective projection onto the image displaying some behavior in frequency domain. Vioix et al. (2002) exploits this feature and apply a bi-

dimensional Gabor filter, defined as a modulation of a Gaussian function by a cosine signal. The frequency parameter required by the Gabor filter is empirically deduced from the 2D-Fast Fourier Transform (Bossu et al., 2006). Bossu et al. (2009) apply wavelets to discriminate crop rows based on the frequency analysis. They exploit the fact that crop rows are well localized in the frequency domain; thus selecting a mother wavelet function with this frequency the crop rows can be extracted. Crops do not display clear frequency content in the Fourier space, particularly in images with a high weed pressure, therefore the application of filters based on the frequency becomes a difficult task.

8.2.1.3. Motivational research and design of the proposed strategy

The proposed strategy is focused on crop row detection in maize fields, for specific treatments according to the PA requirements; this requires discrimination among crops and weeds. This means that crop rows must be identified and located in the image conveniently. Some of the requirements proposed by Astrand (2005) and reported in Slaughter et al. (2008) for guidance systems can be considered for crop row detection where, in essence, the problem is similar. They are:

- Ability to locate crop rows with the maximum accuracy as possible.
- Ability to work on real-time.
- Ability to work on sown crops, not planted, which means that the time treatment and the size of the crops can vary in the field. This also means that crops and weeds have about the same size during the treatment displaying similar spectral components, i.e. discrimination between crops and weed cannot be made by size only.
- Ability to work when plants are missing in the row.
- Ability to work when there is high weed pressure.

The advantages of the existing methods introduced above are exploited with the aim of designing a new strategy for crop row detection inspired on the human visual perception abilities that can cope with the above requirements, where the emphasis is put on the weed pressure treatment. This makes the main difference with respect to the

above approaches, which base its design on the assumption that weed pressure is relatively low and the crop rows can be identified among crops and weeds. Moreover, because the vehicle navigates on uneven terrains with perhaps abundant irregularities, the application of geometric constraints does not suffice because the camera is continuously involved in a permanent swinging. The main contribution of the proposed method is focused on the solution of these two important problems from an automatic point of view that applies the underlying knowledge about the system and performs favorably under the above adverse conditions and also when crop plants are missing in the crop rows.

The image is captured in the RGB color model. An image segmentation approach, based on vegetation index, is applied. A first image thresholding method identifies green parts (weeds/crop). By applying a second thresholding approach, crop and weeds are separated. Based on geometrical considerations and on the knowledge about the system configuration, a template is previously built, containing those parts where crop lines are expected. By combining the information provided by this template and the crop plants, already separated after the second thresholding, crop lines are identified. By applying a total least-squares linear regression method, an adjusted line to the crop is extracted and identified by its equation.

8.2.2. Materials and methods

8.2.2.1. Images

All images used for this study belong to maize crops. Two types of images have been used for image analysis. A first set of 1200 real images captured with a Basler scA 1400-17fc camera during April/ May 2011, belonging to Type-1, Figura 6-1(a). The digital images were captured under perspective projection and stored as 24-bit color images with resolutions of 1392 x 1044 pixels saved in RGB (Red, Green and Blue) color space in the TIFF format. Camera setting was: pitch and roll angles of 20° and 0° respectively at a height of 1.5 m from the ground. There were several days of acquisitions, providing some sets of images with different growth state and different conditions of illumination. They contain areas with different denseness. A second set of 1400 digital images, belonging to Type-2, Figura 6-1(b), captured with a Sony commercial DCR PC110E camera, with resolution of 720 x 576 pixels during April/May 2007 and saved also in RGB color space in the JPEG format. All were

acquired in a 1.7-ha experimental field of maize on La Poveda Research Station, Arganda del Rey, Madrid. Camera setting for this type was: pitch and roll angles of 10° and 0° respectively and at a height of 1.5 m from the ground.

All methods were implemented in C++ under the IDE Visual Studio 2008 and using the free computer vision library OpenCV (2014).

Figura 6-1(a) and (b) displays two illustrative examples of both types of images, identified as Type-1 and Type-2 respectively. Both contain high weed densities, but the weed pressure in (b) is exceptionally high.

8.2.2.2. Image acquisition and segmentation

Once the image is captured in the RGB color model, the first step consists in the plant extraction or greenness image segmentation. A pixel where the predominant spectral component is the green is considered vegetation. In our images this assumption is acceptable due to the nature of the images and because they are acquired in the middle of maize fields where only soil, weed and crop plants are present. In order to discriminate between vegetation and non-vegetation we have used color vegetation indices (Ribeiro et al., 2005; Woebbecke et al., 1995; Xiang & Tian, 2011), exactly the Excess Green (ExG). The choice was made after comparing the following indices and to observe how plants were identified: ExG (Meyer et al., 1998; Woebbecke et al., 1995), excess green minus excess red index or ExGR (Neto, 2004), color index of vegetation or CIVE (Kataoka et al., 2003) and NDI (Perez et al., 2000). The previous color vegetation indices (also called color indices in the literature) used are color space transformations from RGB to a space one-dimensional. Woebbecke et al. (1995) tested five indices and found the ExG index to perform well. Next we go into details about all needed operations for extracting the ExG color index. The three spectral bands R, G and B are used as sources. First we normalize the values of each spectral band, this normalization is performed dividing by the highest value within spectral band for the image being analyzed. The normalized pixel values range in $[0, 1]$ according to (8.1).

$$\hat{R} = R/R_m, \quad \hat{G} = G/G_m, \quad \hat{B} = B/B_m \quad (8.1)$$

The actual pixel values from the images based on each *RGB* channel are R, G and B for the channels red, green and blue respectively, C_m is the highest pixel value for

the channel C in the image and \hat{C} is the normalized pixels values for the channel C , where $C \in \{R, G, B\}$. The images are 24-bit color images and the pixel belongs to natural numbers set $[0, 255]$.

A second normalization is performed, which is usually applied in agronomic image segmentation (Gée et al., 2008; Meyer & Neto, 2008), with this new operation the sum of the three channels' values for a given pixel is equals to unity, as follows:

$$r = \frac{\hat{R}}{\hat{R} + \hat{G} + \hat{B}} \quad g = \frac{\hat{G}}{\hat{R} + \hat{G} + \hat{B}} \quad b = \frac{\hat{B}}{\hat{R} + \hat{G} + \hat{B}} \quad (8.2)$$

Finally the index ExG is computed as follows:

$$ExG = 2g - b - r \quad (8.3)$$

8.2.2.3. Double thresholding

In ExG we can see different levels of values so that the higher the value the higher the greenness. This means that the living plant material (crop or weed) appears as white spots and the rest (soil surface, stones, shadows, straw and other debris) is black. Thus, we need an automatic approach to detect the best level separating plants from the rest. This becomes a thresholding problem. We have investigated different thresholding approaches and the adaptive Otsu's (1979) method that maximizes the interclass variance of the plant and background pixels becomes acceptable. It has been applied in different studies related to crop and weed detection (Burgos-Artizzu et al., 2011; Jiang & Zhao, 2010; Meyer & Neto, 2008; Wei & Shangfeng, 2006). It assumes bimodal histograms, representing two Gaussian intensity distributions, one of them associated to plants and the other to the rest. After this first thresholding, pixels belonging to plants are separated from the rest. Now, because plants include both weeds and crops, we can assume that even if they have similar spectral components is still possible to distinguish some spectral variation which is mapped on the ExG. Hence, we can apply a second thresholding to separate different kinds of green plants. Now, this thresholding is applied only over pixels belonging to ExG which have been identified as plants. These pixel produce a new histogram that considered as bimodal allow to apply again the

Otsu's thresholding method, a second threshold is obtained that allows distinguish two different types of plants. This double thresholding approach is the key step in the proposed approach. Figura 6-2 and Figura 6-3 display the binarized images after the first and second thresholding applied to ExG images of Figura 6-1.

From images displayed in Figura 6-2(a) and Figura 6-3(a), after the first thresholding, green plants are identified and marked as white pixels, obviously displaying a high density derived from the also high weed pressure, particularly the image in Figura 6-3(a). After applying the second thresholding a large number of pixels identified as weeds disappear, Figura 6-2(b) and Figura 6-3(b). Discussion about these results is provided in section 8.2.3.

8.2.2.4. Crop row identification

After the second thresholding the next goal is to assign the corresponding straight line equation to each crop row. With such purpose we resort to methods based on linear regression where the parameters of the straight line are to be estimated from pixels considered as belonging to crop rows. At this step the problem is to determine which pixels participate in the parameter estimation for each crop row. To solve this problem, we make use of the underlying knowledge about the arrangement of crops in the maize field and also based on the geometric constraints of the camera system. They are:

- (a) The number of crop rows to be detected is known.
- (b) The expected location of each crop row in the image is known.
- (c) The area to be explored in the image is also known and determines minimum and maximum values in the y-coordinates over the image.

On the other hand we know that the vehicle, equipped with the vision system, navigates on a real terrain presenting irregularities and roughness. This produces vibrations and also swinging mainly in the pitch and roll angles. The yaw angle is assumed to be correct because otherwise the robot will navigate erroneously and this derives in a control problem out the scope of crop row detection. Based on this fact and after capturing three video sequences consisting of 1400 images each, with the vehicle navigating at 4 km/ h, we have verified maximum right and left displacements around

the expected crop lines in the image. Based on this fact and on the observation from the three videos, we build the corresponding template or mask to determine limits where pixels belonging to crop rows are expected. Obviously, for building these templates, we apply the knowledge about the camera settings and system geometry. Figura 4-9(b) displays a template for the image displayed in Figura 4-9(b). This kind of templates are binary images, manually built assuming previous knowledge; they contain the same number of strips than crop lines are to be estimated. Sizes of these templates are identical to the images under processing. The strips are based on the perspective projection assuming the crop rows in the field are parallel. An additional advantage of using templates is that they contain as many areas as crop lines to be detected. This means that only pixels belonging to a given area will participate on the line estimation for crop row identification.

Now, given the binary image after the second thresholding and considering the corresponding template, we apply the logical operation “and”, which allows filtering those pixels that are considered as potential candidates to belong to the crop lines, i.e. pixels which participate in the estimation of the parameters involved in the straight line equations defining the crop rows. The remaining pixels are excluded and do not participate in the parameter estimation. The use of templates allows a more accurate straight line adjustment because pixels, considered as outliers, probably belonging to weeds, are excluded from the estimation.

Now the problem has been reduced to find the best fitting straight line through a set of points. The linear least squares fitting is probably the simplest form of linear regression. Least square is a general approach to the approximate solution of over-determined systems (with more equations than unknowns). The solution minimizes the sum of the squares of the errors made in solving every single equation. The difference between an observed value and the value provided by the model is named residual. Least squares problems can be classified into linear and nonlinear depending on whether or not the residuals are linear in all unknowns. As the functions to be estimated in our approach are always straight lines, we used linear least squares, but nonlinear could be also applied (Bates & Watts, 1988; Johnson et al., 1981; Marquardt, 1963). This choice is also justified on the basis of achieving low computational cost, making the method suitable for real-time applications. In this study least squares is used to find the best-fitting straight line to a given set of points by minimizing the sum of the squares of the

offsets (“the residuals”) of the points from the straight line. The sum of the squares of the offsets is used, instead of the offset absolute values, because this allows the residuals to be treated as a continuous differentiable quantity. Outlying points can have undesirable effects on the fit due to squares of the offsets, which can be harmful. In our case, we have carried out several tests and in all of them there are not problems with results of the fitting because the use of the templates avoids the presence of outliers, as expected.

The offsets can be vertical or perpendicular with respect the straight line under estimation. With vertical offsets (classical least square) usually the errors are defined as the squared distance to fitted function from points along the vertical axis. The total leastsquares method (Golub, 1973; Golub & Van Loan, 1980) is a natural generalization of the least-squares approximation method when the data in both X and Y coordinates is perturbed. For the pixels belonging to the crop row there is not an assignment of “dependent” and “independent” variables and it makes more sense to measure errors as the squared perpendicular distance to the fitted function.

Given a set of points (pixels) $X \equiv \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ from the binary images after the second thresholding and belonging to an isolated zone in the image template, the goal is to estimate the best fitted straight line with equation $y = mx + a$, where the slope m and the y -intercept a are the parameters to be estimated (Markovsky & Van Huffel, 2007). The expression to be minimized corresponds to the error or sum of offsets given by,

$$E = \sum_{i=1}^n d_i = \sum_{i=1}^n \frac{|y_i - (a + mx_i)|}{\sqrt{1 + m^2}} \quad (8.4)$$

where d_i is the offset for the point i whose coordinates are. Minimizing (4) we obtain:

$$a = \frac{\sum_{i=1}^n y_i - m \sum_{i=1}^n x_i}{n} = \bar{y} - m\bar{x}; \quad m = -B\sqrt{B^2 + 1} \quad (8.5)$$

The term B above is defined as follows:

$$B = \frac{\left(\sum_{i=1}^n y_i^2 - n\bar{y}^2 \right) - \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)}{2 \left(n\bar{xy} - \sum_{i=1}^n x_i y_i \right)} \quad (8.6)$$

Figura 6-4 displays the final result for identification crop row for the original images of Figura 6-1. We can see the straight lines, which represented the crop rows, obtained by the algorithm drawn in red on the original image. Figura 6-4(a) and (b) display the crop row detected (red lines) for the images displayed in Figura 6-1(a) and (b) respectively.

8.2.3. Results

From images displayed in Figura 6-2 and Figura 6-3, we can see that after the first thresholding the green plants, including crop and weeds, are recovered and identified from the original images in the corresponding images in Figura 6-1. After the second thresholding, only crop plants are identified even if in the image in Figura 6-1(b) where weed pressure is extremely high. This means that the proposed double Otsu strategy becomes appropriate for high weeds densities and facilitates the posterior crop row detection based on line adjustment. The performance of the double Otsu is verified for the set of images analyzed. Figura 6-4 displays the three crop rows which have been identified at each image, where a straight line equation is associated to each crop row. As we can see in these images, the red lines follow the crop rows even when the crop rows appear without a symmetric distribution with respect an imaginary central axis. This is particularly appreciable in the image displayed in Figura 6-4(a).

As mentioned during the introduction, the Hough transform has been applied in several methods for crop row detection, hence we compare the performance of our linear regression (RL) approach against the Hough (HOU) transform. Both comparisons are studied after applying the double Otsu-based thresholding method. The Hough transform is implemented to work under the normal representation, polar coordinates, (Gonzalez et al., 2009) where increment in the parameter representing the angle is set to one.

The comparison is established in terms of effectiveness and processing times or computational cost measured in time units. The effectiveness is measured based on the expert human criterion, where a line which has been detected, is considered as correct if it overlaps with the real crop row alignment according to the expert's observation. Over the set of 2600 images analyzed, we compute the average percentage of correct

overlapping for both RL and HOU. Because the main goal of the proposed approach is its profit for real-time applications, we measure computational times. Also, with the goal of real-time, we have tested these performances for different image resolutions. Image resolutions differ from the ones in the original images, these resolutions have been obtained by applying a down-sampling by two to each image in both sets of available images, represented by images in Figura 6-1(a) and (b) and identified as Type-1 and Type-2 respectively.

This is intended under the idea that it is possible reduce the image dimension retaining the main information without affecting the effectiveness in the crop row detection at the same time that the processing time is reduced. Tabla 6-1 displays the results. The first column contains two resolutions for each type of images (Type-1 and Type-2). Columns second and third contain the averaged percentage of effectiveness, in terms of correct overlapping, over the sets of images studied; columns fourth and fifth display the averaged processing times, measured in milliseconds, over the sets of images analyzed. These times corresponds to the full process, i.e. image segmentation, thresholding by double Otsu and linear regression (RL) or Hough transform (HOU).

From results in Tabla 6-1, we can infer that the best performance in terms of effectiveness is achieved by RL with high resolution and also for images of Type-1. This is justified on the basis that images with high resolution contain more information than the ones with less resolution. With regard the type of images, it is clear that Type-2 images are much more complex than Type-1, hence the worst results obtained in terms of effectiveness for images of Type-2.

With respect the processing times, RL outperforms HOU, this is because HOU requires more computational operations although the range of search for the angle involved in the Hough transformation was limited to range between 52° and 130° , so only lines with the slopes defined by the expected crop lines are considered. As mentioned before, the step for indexing the angle in the cells accumulator was set to 1° .

Another problem concerning the Hough transform is the choice of the local maxima in the accumulator for each straight line. Indeed, the intention is to obtain a unique equation for each alignment of pixels, but in the accumulator appear different values for a unique straight line. This causes that several lines attached to each crop row are found. We can select the one with the maximum peak value in the accumulator or

we can compute an averaged value among all lines involved. Both solutions often lead to errors in identifying the correct line associated to the pixels alignment. These are the main reasons justifying the outperformance of RL against HOU.

8.3. Automatic Expert System for masked crop and weeds detection

8.3.1. Introduction

8.3.1.1. Problem statement

Machine vision is an excellent sensor, which is being currently incorporated in autonomous tractors, for treatments over site-specific areas in a larger field (Davies et al., 1998). Focusing on maize fields, one of the most important treatments is weeds killing, where plants (weeds and crops) must be identified as a previous step. Different methods and strategies for plant identification have been applied in different works (Burgos-Artizzu et al., 2009; Guerrero et al., 2012; Guijarro et al., 2011; Montalvo et al., 2012; Onyango & Marchant, 2003; Tellaeche et al., 2008*a*, 2008*b*). López-Granados (2011) makes a revision of methods where plant identification is a key step in the process. Most existing strategies address the problem of green identification under the assumption that plants display a high degree of greenness, but they do not consider the fact that plants may have lost their degree of greenness for different reasons.

Indeed, maize is an irrigated crop, which is also unprotected from the rainfall. When the layer of water is abundant or its fall on the ground is relatively strong, soil materials (particularly clays) impregnate the vegetative cover, particularly those parts close or near to the soil. In this case, the green spectral component of pixels belonging to the plants is masked by the dominant red spectral component coming from materials existing in the soil; Figura 6-6 displays an image where this appears clearly in the middle central part of the image and also at the ends of the leaves in the maize that are oriented toward the soil. This makes methods based exclusively on the greenness identification, i.e. plant coverage based on the computation of vegetation indices, fail under such situations. Indeed, soil and masked plants are both identified as soil.

In Guerrero et al. (2012) we have already addressed this problem by applying a learning approach based on support vector machines. As all learning strategies, this method requires a training phase where samples are conveniently provided for estimating the required parameters (support vectors), and then the posterior decision

phase is highly dependent of the samples supplied, i.e. from the images which have been used for training. Moreover, the learning phase requires a certain number of images previously selected.

We propose a new automatic method based on several sequential stages, where the linking of these stages and the image segmentation processes, applied at each stage, are based on the application of the human expert knowledge. This leads to the design of the proposed expert system, gaining an important advantage with regard to the one described in Guerrero et al. (2012) because no training is required and it can be directly applied to the unique image under processing becoming independent from other images, which are to be selected. One of the processes involved into two different stages is image thresholding, based on the Otsu's method, which is self-adjustable, dealing well with images captured under different conditions such as sunny or cloudy days, affecting illumination variability, (Tian & Slaughter, 1998); they are typical situations in agricultural images coming from outdoor environments. The design of this automatic expert system makes the main contribution of this study.

Additionally, because this system is designed to identify plants that have lost the greenness, it can be applied to evaluate the treatment effectiveness. Indeed, as mentioned before, some site-specific treatments are intended to kill weeds in maize crops, when weeds are in the dying process, before reaching dry completely; they also have lost their greenness, compared to their healthy state. The proposed expert system can be used for identifying such plants and hence the treatment effectiveness.

Figura 6-7 displays at its central inter-row crop, weeds evolving toward a dry stage after the chemical treatment with herbicide applied two days ago. This image has also received direct rainfall and some parts are impregnated with materials coming from the soil as before.

8.3.1.2. Revision of methods

Several strategies have been proposed for segmenting crop canopy images, specifically oriented towards green segmentation:

- (1) Visible spectral-index based, including the excess green index (ExG, Ribeiro et al., 2005; Woebbecke et al., 1995), the excess red index (ExR,

Meyer et al., 1998), the color index of vegetation extraction (CIVE, Kataoka et al., 2003), the excess green minus excess red index (ExGR, Neto, 2004) and the vegetative index (VEG) described in Hague et al. (2006), which is designed to cope with the variability of natural daylight illumination. ExG, ExGR, CIVE and VEG have been applied under a combined form in Guijarro et al. (2011) gaining in performance with respect to their individual application. All these approaches need to fix a threshold for final segmentation, i.e. to discriminate between plants and other parts (soil, sky).

(2) Specific threshold-based approaches, including dynamic thresholding. Generally, these techniques assume a twoclass problem where plants and soil are to be identified. Reid and Searcy (1987) estimate a decision function under the assumption that the classes follow Gaussian distributions. The Otsu's method (Otsu, 1979) is also applied considering a bi-class problem (Ling & Ruzhitsky, 1996; Shrestha et al., 2004). These algorithms are applied to gray images. Gebhardt et al. (2006) apply also thresholding for segmentation transforming the images from RGB to gray scale intensity. This algorithm was later improved using local homogeneity and morphological operations in Gebhardt and Kaihbauch (2007). Kirk et al. (2009) apply a combination of greenness and intensity derived from the red and green spectral bands and compute an automatic threshold for a two-class problem assuming two Gaussian probability density functions associated to soil and vegetation respectively; this procedure requires the previous estimation of an angle to rotate the hypothetical greenness axis. Meyer and Camargo-Neto (2008) have applied the automatic Otsu's thresholding method for binarizing ExG and the normalized difference index (NDI), where a comparison is established against the segmentation obtained from ExGR determining that in this last case, a value of zero suffices for the threshold, therefore the Otsu's method is not required. Guijarro et al. (2011) and Burgos-Artizzu et al. (2011) have applied the statistical mean value of the transformed image obtained with the vegetation indices instead of automatic thresholding such as Otsu. They justify its choice because Otsu's method gives a threshold value higher than the mean and produces infra-segmentation, i.e. some plants are not conveniently identified.

(3) Learning-based, Meyer et al. (2004) have applied unsupervised approaches, including fuzzy clustering, for segmenting regions of interest from ExR and ExG. Tian and Slaughter (1998) proposed the environmentally adaptive segmentation algorithm (EASA) for detecting plants through a supervised learning process. Ruiz-Ruiz et al. (2009) applied the EASA under the HSI (hue-saturation-intensity) color space to deal with the illumination variability. Zheng et al. (2009) and Zheng et al. (2010) use a supervised mean-shift algorithm under the assumption that the segmentation of green vegetation from a background can be treated as a two-class segmentation problem; the class separability is validated through a neural network and the Fisher linear discriminant respectively, the color spaces used were RGB, LUV and HSI. Guerrero et al. (2012) apply Support vector machines as the learning strategy applied with identical purpose that the one proposed in this study.

8.3.1.3. Motivational research of the proposed strategy

The above methods are intended for plant identification through their greenness, based on the accentuation of the green color (Meyer & Camargo-Neto, 2008), but their effectiveness drops when the green spectral component becomes less important in favor of the red one as occurs in our case. Based on the images we call unmasked plants to the ones where the green spectral component is dominant and masked plants to these that have lost greenness in favor of the red spectral component.

Based on the above and considering a gray image containing information about the greenness, we could view the problem as a three-clustering approach, where the goal is to identify three classes, from the histogram of the gray image, by computing two thresholds. Demirkaya et al. (2008) have proposed an iterative method for this purpose. The histogram is divided into three regions with two pseudo-random thresholds and both thresholds are dynamically adjusted based on the inter-class mean values until no more adjustments are required. The main drawback of this approach is that in this kind of images, the histogram does not display three well separated classes and this method provides two thresholds which are not valid for solving satisfactorily our problem, as we will see later.

Thus, the idea is to apply an automatic strategy for image segmentation similar to that a human expert would apply to a similar problematic situation where the images contain classes which identify masked and unmasked plants due to different facts. This reasoning or knowledge, based on several stages, is the kernel of the proposed expert system. Each stage is designed for a given purpose and specific image segmentation approaches are applied for achieving the goal at each stage.

8.3.2. Expert system design

8.3.2.1. Reasoning for knowledge extraction

Based on a logical expert reasoning, the proposed expert system is designed according to the modular architecture displayed in Figura 4-14. It contains three stages, which are sequentially linked to form the expert system as a whole. Each stage contains the required automatic image processing modules.

(1) Stage 1: Unmasked plants do not offer any difficulty in their distinction, they are to be extracted at this first stage.

(2) Stage 2: Masked plants and soil remain together, the next step consists in their separation, this is carried out by identifying those pixels, with a certain degree of greenness, which are associated to masked plants. The remainder ones are considered as belonging to the soil.

(3) Stage 3: Based on the assumption that pixels belonging to plants are grouped together forming patches and they rarely appear isolated, a procedure to remove small patches and isolated pixels is to be applied.

8.3.2.2. Automatic image processing modules

Following the three previous stages, at each stage a sequence of image processing techniques are applied for automatic purposes, being grouped and linked conveniently.

(1) Stage 1, combination of vegetation indices and application a first Otsu thresholding: given an original input image in the RGB color space, we apply the following normalization scheme, which is usually applied in agronomic image segmentation (Gée et al., 2008),

$$r = \frac{R_n}{R_n + G_n + B_n}, \quad g = \frac{G_n}{R_n + G_n + B_n}, \quad b = \frac{B_n}{R_n + G_n + B_n} \quad (8.7)$$

where R , G and B are the normalized RGB coordinates ranging from 0 to 1 and are obtained as follows:

$$R_n = \frac{R}{R_{\max}}, \quad G_n = \frac{G}{G_{\max}}, \quad B_n = \frac{B}{B_{\max}} \quad (8.8)$$

where $R_{\max} = G_{\max} = B_{\max} = 255$ for our 24-bit color images.

Vegetation indices to be combined are computed as follows (see references above in Section 8.3.1.2),

$$ExG = 2g - b - r \quad (8.9)$$

$$CIVE = 0.441r - 0.811g + 0.385b + 18.78745 \quad (8.10)$$

$$VEG = \frac{g}{r^a b^{1-a}} \quad (8.11)$$

with $a=0.667$ as in Hague et al.(2006)

Based on Guijarro et al. (2011) the above three indices are combined to obtain the resulting value COM as follows,

$$COM = w_{ExG} \times ExG + w_{CIVE} \times CIVE + w_{VEG} \times VEG \quad (8.12)$$

where w_{ExG} , w_{CIVE} and w_{VEG} are the weights for each index, representing their relative relevance in the combination. Guijarro et al. (2011) provide the four weight values participating in the combination, because in this work we have excluded the $ExGR$ index, the weight values for the three indices are proportionally recalculated, i.e. $w_{ExG} = 0.36$; $w_{CIVE} = 0.47$ and $w_{VEG} = 0.17$.

The resulting combined image COM, is linearly mapped to range into [0,1], after which is thresholded by applying the Otsu's method, obtaining a binary image, where white pixels identify unmasked plants, i.e. plants not contaminated due to materials coming from the soil. The green spectral component of these pixels is dominant with respect the red and blue ones.

(2) Stage 2, select black pixels and apply a second Otsu thresholding: once pixels are identified as unmasked plants, the remainder pixels in COM are extracted; they are assumed to belong to soil and masked plants, these last are now our interest. The histogram obtained from pixels belonging to soil and masked plants, is thresholded by applying again Otsu. After this second thresholding two kinds of pixels are identified, those belonging exclusively to soil, debris and different materials and those belonging to masked plants. These last ones together with those identified previously as unmasked pixels form the final binary image or segmented image. When the method is applied for evaluating the effectiveness in the post-treatment process, all pixels identified after the second thresholding are those belonging to plants affected by the treatment, which have already started their dying process.

(3) Stage 3, morphological operations and identification of masked and unmasked plants: once the image has been binarized, the following three morphological operations are applied in the order expressed below:

(a) Opening: To remove small patches.

(b) Majority: A pixel is set to zero if five or more pixels in its 3x3 neighborhood are zeros.

(c) Cleaning: To remove isolated pixels resulting from the above two operations, white pixels surrounded by black pixels.

Once the full process is finished, according to the three stages, we obtain the final segmented image, where both unmasked and masked plants are identified. Therefore we have sufficient knowledge about their distribution in the field, which was the objective of this work.

8.3.3. Results

The images used for this study were acquired with a HPR817 digital camera device in four different days in April/May 2007. All acquisitions were spaced by five/six days. A set of them were obtained in a pre-treatment phase after the field was artificially watered and also when the field received different amounts of rainfall. Figura 6-6, described in the introduction, is a representative image of this set. A second set of images was acquired in a post-treatment phase after applying a dose of herbicide; in this

phase weeds have started its decease process, Figura 6-7. Because of the difference of the days, different groups of images were acquired under different illumination conditions. This circumstance does not affect the performance of the proposed process because all image processing methods are independent of this circumstance; particularly the Otsu's method which is relative to each image histogram. Therefore it is not required any further study with regard to lighting conditions. This represents also another advantage with regard learning-based methods (Guijarro et al., 2011). This is an additional justification for the choice of proposed expert system.

These digital images were captured under perspective projection containing only soil and plants, i.e. without panoramic sky. They were stored as 24-bit color images with resolutions of 800x600 pixels, and saved in RGB (Red, Green and Blue) color space in the JPG format. The expert system was implemented in Matlab R2009a (MathWorks, 2014) and the images were processed with its Image Processing Toolbox.

A set of 850 images were captured and processed, from which 600 contain masked plants (SET-1) and 250 were captured for post-treatment evaluation (SET-2). Of course, all images contain unmasked plants.

The kernel of the proposed Automatic Expert System (AES) consists in the identification of masked and unmasked plants based on the histogram separation into three classes and two thresholds; hence we focus our analysis on the study of the two thresholds obtained in stages 1 and 2 through the Otsu's method to the COM image. The performance of AES is compared with the double thresholding approach proposed in Demirkaya et al. (2008) (DEM).

With respect images in SET-1, we base the analysis on the image displayed in Figura 6-6. This is because all the images on this set behave similarly with respect to the AES. Figura 6-8(a) displays the COM gray image obtained from image in Figura 6-6; Figura 6-8(b) displays its corresponding histogram ranging in $[0,255]$, i.e. the original values in COM are multiplied by 255.

In the basis of the histogram appear four identifiers, a, b, c and d, indicating four thresholds; a and c are the ones obtained by AES in stages 2 and 1, respectively; b and d are the ones obtained by DEM. Tabla 6-3 displays the four values for each approach (SET-1).

Figura 6-9(a) displays the binary image obtained with threshold c , see histogram in Figura 6-8(b) and Tabla 6-3 for SET-1. White pixels are identified as unmasked pixels at the first stage of the AES and labeled in red, Figura 6-9(b). Now, considering black pixels in the binary image, they are processed according to the image processing procedures defined in stages 2 and 3 in the AES; the threshold a is obtained at stage 2, which allows the identification of masked plants, labeled as blue in the image of Figura 6-9(b). Figura 6-9(c) displays the labeled pixels obtained by DEM with thresholds d (unmasked plants in red) and b (masked plants in blue).

Based on the labels displayed in Figura 6-9(b) and (c) we can see that DEM identifies a smaller number of unmasked plants than AES; also AES identifies patches of masked plants that are not identified by DEM. Elliptical and circular lines in Figura 6-9(c) identify relevant parts and patches verifying the above assertions. The best performance obtained by AES is explained considering the thresholds values from the histogram. Indeed, un-masked plants are extracted with thresholds c and d , Tabla 6-3; because d is greater than c , DEM identifies less number of unmasked plants, which are posteriorly labeled as masked plants. Also, because b is greater than a and both thresholds are related to masked plants, some of these last plants are not identified by DEM. From the histogram in Figura 6-8(b) we can see that there is not clear class separation, this means that clear thresholds cannot be assigned, but AES gets better thresholds than DEM judging by the results. At this point, it is worth mentioning that the use of the statistical mean value as threshold, applied in Guijarro et al. (2011) and Burgos-Artizzu et al. (2011) is not appropriate for our problem because the mean value is always less than the threshold obtained with Otsu and most masked plants and also soil plants are identified as unmasked.

Images in SET-2 contain weeds patches which have started the drying process as a result of a previous treatment. The image displayed in Figura 6-7 is a representative element of this set, where we can see important weeds patches in the central inter-row crops, which appear affected by the treatment, but still there are small patches preserving a relative high degree of greenness, i.e. this means that they are unaffected by the treatment. Figura 6-10(a) displays the resulting image obtained by the application of the COM vegetation index to the image in Figura 6-7; Figura 6-10(b) displays its corresponding histogram ranging in $[0,255]$ with identifiers similar to those above for the thresholds. Tabla 6-3 displays the four values for each approach (SET-2).

Unlike the previous histogram, this is bimodal, where threshold c is the one used in AES for obtaining the binary image displayed in Figura 6-11(a) and also to identify unmasked plants, which are labeled in red, Figura 6-11(b). Threshold d is used in DEM for identifying also unmasked plants, Figura 6-11(c). Many pixels belonging to unmasked plants are identified by AES but not by DEM, this can be explained because threshold c divides the histogram into two parts coinciding better than d with the two modal regions. Moreover, all weeds patches identified by DEM are affected by the treatment, elliptical lines in Figura 6-11(c), when in fact they are not. Discussions above are applicable to all images in SET-2, i.e. they all display similar behavior with respect to both AES and DEM approaches.

Considering all pixels labeled as unmasked and masked plants for SET-1 and SET-2, we have computed the averaged values for the three RGB spectral components from the original images. These values are obtained for AES and DEM and also for the Support Vector Machines (SVM) approach in Guerrero et al. (2012), obtaining the values displayed in Tabla 6-4.

Results display that the green spectral component for unmasked plants is always greater than the red one for the three approaches (AES, DEM and SVM). This means that greenness is clear on these plants, as expected. On the contrary, the green spectral component is less than the red one in a small amount for masked plants. This verifies and supports the initial hypothesis.

In order to assess the validity of the proposed AES and to determine its performance as compared to DEM and SVM, we have randomly selected the 20% from each one of the two sets of images analyzed (SET-1 and SET-2). Each image was visually analyzed by an expert to identify weeds and crop plants. The human visual observation is carried out for each image guided by the segmented image through the approach proposed in this study. The expert concentrates his major effort in identifying the most troubled plants, i.e. those we call masked plants or plants already affected by the treatment. Incorrect assignments are manually marked, corrected or removed, generating new-segmented images, which are considered as ground-truth. Tabla 6-5 displays the percentage of successes obtained for the three approaches.

From results in Tabla 6-5 we can see AES outperforms DEM and SVM. This behavior is explained on the fact that thresholds obtained by DEM behave as explained

before. Moreover, SVM requires training and although in our experiments we obtain similar results than AES, this could vary, getting worse, depending on whether samples used for training becomes insufficient; AES is free of this circumstance, i.e. it is a desirable approach for solving the problem of identifying masked and unmasked plants in outdoor images coming from maize fields.

8.4. Conclusions and future work

8.4.1. Introduction

After the discussion of the above results it is necessary to raise an overall balance on the lines of investigation followed, leading to the conclusions below. Also we are going to express where focus efforts to improve the work outlined for achieve better results.

8.4.2. Conclusions

8.4.2.1. Crop rows and weeds detection

A method for crop rows detection in maize fields with high weeds pressure has been proposed. The proposed strategy is based on three main processes: a) image segmentation; b) double thresholding based on the Otsu's approach; c) least squares linear regression for computing the equations associated to the identified crop rows. Original images are captured in the color space RGB and transformed to grey images by applying the ExG vegetation index. After this transformation, a double Otsu approach allows to separate weeds and crops. A linear regression approach based on total least squares is applied for computing equations of straight lines associated to the crop rows. We apply the underlying knowledge about geometrical constraints embedded in the vision system to build a template that guides the process for computing the equations. The performance of the proposed approach is tested with a broad set of images, verifying its outperformance against the classical Hough approach.

8.4.2.2. Masked plants detection

An AES is proposed for image segmentation in maize fields. It is based on three consecutive stages where the main underlying idea is the successive application of automatic image processing tasks mapping the expert knowledge.

The AES is able to identify plants (weeds and crops) when they have been contaminated with materials coming from the soil, due to artificial irrigation or natural rainfall. It is also valid for monitoring post-treatments; this is based on the assumption that weeds, after chemical or mechanical treatments, must initiate a progressive degradation expressed by the loss of the greenness displayed during the pre-treatment stage. The damage in the crop, when it occurs, can also be analyzed based on the same criterion because of loss of greenness.

In addition, once green plants are identified, the remainder parts belong to the soil and cover a number of ecologically relevant categories (Luscier et al., 2006), thus the proposed AES could be extended to deal with the analysis of soil materials.

The AES has been designed with an open architecture, so that in the future it is possible to replace or add new modules, being of particular interest to study new automatic thresholding methods (Avci & Avci, 2009) or add a knowledge-base for improving image segmentation based on the accumulated knowledge (Gonzalez-Andujar, 2009).

8.4.2.3. RHEA system

An autonomous system in AP has been designed and developed for the selective treatment of weeds in maize fields. The machine vision system is installed on board an adapted tractor provided by Case New Holland Company, a camera (there are two possible models, AXIS 211M and SVS-4050CFLGEA) with the corresponding processing elements. All together are integrated in the whole system for working in real time with the vehicle navigating at 3 km/h on average.

The processing steps are basically three: segmentation, thresholding and line detection. For the segmentation, there were two possible options, two vegetation indices, COM and ExG, each used in one of the AES detailed in chapters 9 and 0. We chose to use the ExG because it has the least expensive operations (multiplication, division and powers) and therefore the least time required, since both vegetation indices showed similar behavior in AES.

The prototype, which is part of RHEA project (2014), has been developed over 4 years with favorable results, being the successful final demo test on May 21, 2014.

One goal of the project was to develop a prototype which will be a commercial product in the near future.

8.4.3. Future work

Although the results were satisfactory, the systems are still susceptible of improvement. Thus future work is presented as possible overall system improvements in each of their parts. It has been thought that these sections could be improved in future work: to capture quality images, image segmentation and crop row detection.

To capture an image with sufficient quality by controlling the exposure time was thought to look for another method that does not use reference fixed values for the three spectral channels. This is because although the algorithm works properly capturing images with sufficient quality, using fixed values as reference, the performance can be slightly affected by the growth stage of the plants and the type and condition of the land. Another possibility would be to maintain the same algorithm but the spectral values should be calculated by a new expert system for each working day, so when the tractor starts the workday, located to the crop field and before starting work, the vision system captures a sequence of images and the new expert system calculates the spectral reference values for the current conditions. When the environmental conditions change drastically, this expert system will recalculate the new values again. In this last case the process must be computationally fast to fulfill real time requirements. It is estimated below 100 ms for the system in mobile units like in the RHEA project, i.e. capturing and processing ten different ROITs per second.

Another aspect that could be improved is the detection of green plants present in the scene by the AES. The step of segmenting the images is the most important because from this transformation the crop rows are identified and the coverage matrix with weeds densities in the ROIT obtained. The greenness was obtained by applying vegetative indices based on the three spectral visible channels, discarding the use of NIR. The use of this spectral band is a matter of study, because it has been broadly used with such purpose. Nevertheless, a trade-off must be achieved to preserve the real time requirements without increasing the production cost, considering that this could be a future commercial product.

In the proposed AES only maize and weeds are distinguished based on the three spectral RGB visible channels and geometrical arrangement. The potentiality of the detection system can be increased by including new methods to distinguish different types of weeds by considering properties such as color hue, texture, shape or different relevant features. To fulfill the real time requirements perhaps parallel architectures or graphics cards should be considered.

Finally, the AES could be adapted to work in other kind of crops. It could be used in crops with a different color to green. It could also be applied to crops where plants have not been planted in a straight line, mainly due to the attempt to optimize arable space or topography. In some places, particularly highland areas, because the field is a hillside and the crops describe curves to try to avoid as far as possible that the vehicle moves in the maximum slope direction.

9. Bibliografía/References

- Adamchuk, V.I., Morgan, M.T., Lowenberg-Deboer, J.M. 2004. A Model for Agro-Economic Analysis of Soil pH Mapping. *Precision Agriculture*, 5(2), 111-129.
- AGL. 2011. Ground inspection systems in unmanned ground vehicles and their effective application in Weed detection for site-specific control (GroW). AGL 2011-30442-C02-02. Plan Nacional de I+D+I, Ministerio de Economía y Competitividad, España.
- Al-Hamed, Saad A.I. 1996 Automatic tractor guidance system based on global positioning system. The University of Nebraska - Lincoln, ProQuest, UMI Dissertations Publishing.
- Al-Kufaishi, S.A., Blackmore, B.S., Sourell, H. 2006. The feasibility of using variable rate water application under a central pivot irrigation system. *Irrigation and Drainage Systems*, 20 (2-3), 317-327.
- Andújar, D., Escolà, A., Rosell-Polo, J.R., Fernández-Quintanilla, C., Dorado, J. 2013. Potential of a terrestrial LiDAR-based system to characterise weed vegetation in maize crops. *Computers and Electronics in Agriculture*, 92, 11-15.
- Asif, M., Amir, S., Israr, A., y Faraz, M. 2010. A vision system for autonomous weed detection robot. *Int. J. Comput. Electr. Eng.*, 2(3), 486-491.
- Asner, G.P. (1998). Biophysical and Biochemical Sources of Variability in Canopy Reflectance. *Remote Sensing of Environment*. 64(3), pp 234-253.
- Astrand, B. 2005. Vision based perception or mechatronic weed control. Doctor of Philosophy Thesis, Chalmers and Halmstad Universities, Sweden.
- Astrand, B., y Baerveldt, A. J. 2005. A vision based row-following system for agricultural field machinery. *Mechatronics*, 15, 251-269.
- Avci, E., y 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.
- AVIRIS. 2013. Airbone Visible Infrared Imaging Spectrometer. <http://aviris.jpl.nasa.gov/> (accedido 2014-07-01).
- AXIS. 2014. Axis communications. <http://www.axis.com> (accedido 2014-07-01).
- Basler. 2014. Basler the power of sight. <http://www.baslerweb.com/> (accedido 2014-07-01).

- Bates, D. y Watts, D. 1988. *Nonlinear regression and its applications*. New York: Wiley-Interscience.
- 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*.
- Berni, J. A. J., Zarco-Tejada, P. J., Suárez, L., Fereres, E. 2009b. Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Transactions on Geoscience and Remote Sensing*, 47(3), 722-738.
- Billingsley, J., y Schoenfisch, M. 1997. The successful development of a vision guidance system for agriculture. *Computers and Electronics in Agriculture*, 16, 147–163.
- Bossu, J., Gée, Ch., Guillemain, J. P., y Truchetet, F. 2006. Development of methods based on double Hough transform and Gabor filtering to discriminate crop and weeds in agronomic images. In *Proc. SPIE 1 8th annual symposium electronic imaging science and technology*, paper No. 23, San Jose, USA (Vol. 6070).
- Bossu, J., Gée, Ch., Jones, G., y Truchetet, F. 2009. Wavelet transform to discriminate between crop and weed in perspective agronomic images. *Computers and Electronics in Agriculture*, 65, 133–143.
- 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.
- 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.
- Chan, C.W. 2012. Variable Rate Application of Fertilizer in Citrus. *Advances in Citrus Nutrition*, 257-269. Capítulo 18.
- Christensen, L.K., Upadhyaya, S.K., Jahn, B., Slaughter, D.C., Tan, E., Hills, D. 2005. Determining the Influence of Water Deficiency on NPK Stress Discrimination in Maize using Spectral and Spatial Information. *Precision Agriculture* 6(6), pp. 539-550.
- Dammer, K.H., Wartenberg, G. 2007. Sensor-based weed detection and application of variable herbicide rates in real time. *Crop protection*, 26 (3), pp. 270-277.

- Dammer, K.H. 2010. Variable Rate Application of Fungicides. Precision Crop Protection – the Challenge and Use of Heterogeneity. Capítulo 22.
- Dammer, K.R. 2013. Precision Farming: Soil Fertility and Productivity Aspects. Apple Academic Press.
- Davies, G., Casady, W., Massey, R., 1998. Precision agriculture: an introduction. Water Quality Focus Guide (WQ450). University of Missouri. Disponible en: <http://extension.missouri.edu/explorepdf/envqual/wq0450.pdf> (accedido 2014-07-01).
- Demirkaya, O., Asyali, M. H., y Sahoo, P. K. 2008. Image processing with MATLAB: Applications in medicine and biology. Boca Raton: CRC Press – Taylor and Francis Group.
- Dillon, C.R. 2013. Heuristic optimization for variable rate nitrogen and seeding decisions. Precision agriculture '13, 761-768
- Dillon, C.R., Kusunose, Y. 2013. Dispelling misperceptions regarding variable rate application. Precision agriculture'13, 769-776
- DPI. 2009. *Sistema de Vigilancia, Búsqueda y Rescate en el Mar mediante colaboración de Vehículos Autónomos Marinos y Aéreos* (DPI2009-14552-C02-01). Plan nacional de I+D+I, Ministerio de ciencia e Innovación.
- Earl, R., Wheeler, P.N., Blackmore, B.S., Godwin, R.J. 1996. Precision Farming: The management of variability. Landwards, 51(4), 18–23.
- Fontaine, V., y Crowe, T. G. 2006. Development of line-detection algorithms for local positioning in densely seeded crops. Canadian Biosystems Engineering, 48(7), 19–29.
- Fontanelli, M., Frasconi, C., Raffaelli, M., Martelloni, L., Peruzzi, A. 2011. Effect of flaming at different LPG doses on maize plants. Proceedings of the first international workshop on Robotics And Associated High-Tecnologies and Equipment for Agriculture (RHEA 2011), 23-32, Montepiller, France.
- Fu, K.S., Gonzalez, R.C., Lee, C.S.G. 1998. Robótica: Control, Detección, Visión e Inteligencia. McGraw Hill, Madrid, Spain, pp. 306–336.
- GALILEO. 2014. European GNSS Agency. Disponible en: <http://www.gsa.europa.eu/galileo-0> (accedido 2014-07-01).
- Gebhardt, S., Schellberg, J., Lock, R., Kaühbauch, W.A. 2006. Identification of broad-leaved dock (*Rumex obtusifolius* L.) on grassland by means of digital image processing. Precision Agriculture, 7 (3), 165–178.

- Gebhardt, S., Kauhbauch, 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.
- Gée, Ch., Bossu, J., Jones, G., Truchetet, F. 2008. Crop/weed discrimination in perspective agronomic images. *Computers and Electronics in Agriculture*, 60 (1), 49-59.
- Giltelson, A. A., Kaufman, Y. J., Stark, R., Rundquist, D. 2002. Novel algorithm for remote estimation of vegetation fraction. *Remote Sensing of Environment*. Elsevier, 80, 76–87.
- GLONASS. 2014. Federal space agency. Information-Analitical Centre(IAC). Disponible en: <http://glonass-iac.ru/en/> (accedido 2014-07-01).
- Golub, G., 1973. Some modified matrix eigenvalue problems. *SIAM Review*, 15, 318-344.
- Golub, G., Van Loan, C., 1980. An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, 17, 883-893.
- 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.
- Gonzalez, R.C., Woods, R.E., Eddins, S.L. 2009. *Digital Image Processing using MATLAB*. Upper Saddle River, NJ: Prentice Hall.
- Gonzalez, R.C., Woods, R.E.. 2007. *Digital Image Processing 3ª Ed.* Upper Saddle River, NJ: Prentice Hall.
- Gonzalez-Andujar, J.L. 2009. Expert system for pests, diseases and weeds identification in olive crops. *Expert Systems with Applications*, 36, 3278-3283.
- GPS. 2014. Official U.S. Government information about the Global Positioning System (GPS) and related topics. Disponible en: <http://www.gps.gov/> (accedido 2014-07-01).
- 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(12), 11149-11155.
- Guerrero, J.M., Guijarro, M., Montalvo, M., Romeo, J., Emmi, L., Ribeiro, A., Pajares, G. 2013. Automatic expert system base on images for accuracy crop row detection in maize fields. *Expert Systems with Applications*, 40(2), 656-664.

- 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, 75-83.
- Hague, T., Marchant, J. A., y Tillett, D. 1997. A system for plant scale husbandry. *Precision Agriculture*, 635–642.
- Hague, T., Tillet, N., Wheeler, H. 2006. Automated crop and weed monitoring in widely spaced cereals. *Precision Agriculture*, 7(1), 95–113.
- Havlin, J.L., Heininger, R.W. 2009. A variable-rate decision support tool. *Precision Agriculture*, 10(4), 356-369.
- Heege, H.J. 2013a. Precision in Guidance of Farm Machinery. *Precision in Crop Farming*, 35-50. Capítulo 4.
- Heege, H.J. 2013b. Site-Specific Sowing. *Precision in Crop Farming*, 171-192. Capítulo 8.
- Hewlett Packard. 2014. HP Store particulares. Disponible en: <http://h10025.www1.hp.com/ewfrf/wc/manualCategory?cc=es&destPage=product&dlc=es&lc=es&product=437509&> (accedido 2014-07-01).
- Hough, P. V. C. 1962. A method and means for recognizing complex patterns. US Patent Office No. 3069654.
- ISCAR. 2014. Ingeniería de Sistemas, Control, Automática y Robótica. Grupo de investigación perteneciente al Departamento de Arquitectura de Computadores y Automática de la Universidad Complutense de Madrid. Número de referencia 910793. Disponible en: <http://www.dacya.ucm.es/area-isa/index.php?page=home> (accedido 2014-07-01).
- Jacobson, R.E., Ray, S.F., Attridge, G.G., Axford, N.R. 2000. The manual of photography: photographic and digital images. Novena edición. Oxford. Focal Press.
- Ji, R., y Qi, L. 2011. Crop-row detection algorithm based on random Hough transformation. *Mathematical and Computer Modelling*, 54, 1016–1020.
- Jiang, G., Zhao, C., 2010. A vision system based crop rows for agricultural mobile robot. *International conference on computer application and system modeling (ICCAS)*, 11, V11-142 - V11-145.
- Johnson, M. L., Correia, J. J., Yphantis, D. A., Halvorson, H. R. 1981. Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophysical Journal*, 36(3), 575-588.

- Jones, G., Gée, Ch., y Truchetet, F. 2009*a*. Modelling agronomic images for weed detection and comparison of crop/weed discrimination algorithm performance. *Precision Agriculture*, 10, 1–15.
- Jones, G., Gée, Ch., y Truchetet, F. 2009*b*. Assessment of an inter-row weed infestation rate on simulated agronomic images. *Computers and Electronics in Agriculture*, 67, 43–50.
- Jurado-Expósito, M., López-Granados, F., Atenciano, S., García-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*, Volume 22(10), 1177-1180.
- Kapur, J.N., Sahoo, P.K., Wong, A.K.C. 1985. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer vision graphics and image processing*, 29(3), 273-285.
- 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.
- Kise, M., y Zhang, Q. 2008. Development of a stereovision sensing system for 3D crop row structure mapping and tractor guidance. *Biosystems Engineering*, 101, 191–198.
- Kise, M., Zhang, Q., y Rovira-Más, F. 2005. A stereovision-based crop row detection method for tractor-automated guidance. *Biosystems Engineering*, 90(4), 357–367.
- Lamm, R. D., Slaughter, D. C., Giles, D. K., 2002. Precision weed control for cotton. *Transactions of the ASAE*, 45, 231–238.
- Landsat. 2013. The Landsat program. Disponible en: <http://landsat.gsfc.nasa.gov/> (accedido 2014-07-01).
- Leemans, V., y Destain, M. F. 2006. Application of the Hough transform for seed row location using machine vision. *Biosystems Engineering*, 94(3), 325–336.
- Li, Y., Chen, D., Walker, C.N., Angus, J.F. 2010. Estimating the nitrogen status of crops using a digital camera. *Field Crops Research*, 118(3), 221–227.
- Liakos, V., Vellidis, G., Harris, G., Hill, R., Henry, H. 2013. Variable rate application of side-dress nitrogen on cotton in Georgia, USA. *Precision agriculture '13*, 435-442.
- 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.
- 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.
- Mallarino, A.P., Wittry, D.J. 2006. Variable-rate application for phosphorus and potassium: Impacts on yield and nutrient management. *Integrated Crop Management Conference*, Iowa State University, 219-224.
- Mao, W., Wang, Y., Wang, Y. 2003. Real-time detection of between-row weeds using machine vision. Paper number 031004. *American Society of Agricultural and Biological Engineers Annual Meeting*.
- Marchant, J. 1996. Tracking of row structure in three crops using image analysis. *Computers and Electronics in Agriculture*, 15, 161–179.
- Marchant, J., Onyago, C. 2000. Shadows-invariant classification for scenes illuminated by daylight. *J. Opt. Soc. Am. A*, 17(11), 1952.
- Markovsky, I., Van Huffel, S. 2007. Overview of total least-squares methods. *Signal Processing*, 87(10), 2283-2302.
- Marquardt, D. W. 1963. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2), 431-441.
- MathWorks. 2014. Accelerating the pace of engineering and science Disponible en: <http://www.mathworks.com/> (accedido 2014-07-01).
- 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., Hindman, T.W., Lakshmi, K., 1998. *Machine Vision Detection Parameters for Plant Species Identification*. SPIE, Bellingham, WA.
- 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.E., Camargo-Neto, J. 2008. Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture* 63 (2), 282-293.

- MicroStrain. 2014. Lord MicroStrain SENSING SYSTEMS. Disponible en: <http://www.microstrain.com/> (accedido 2014-07-01).
- Ming, L., Liu Zh.H., Huang, J.A., Daid, S.H., Wakabayashi, K., Imou K. 2012. Artificial Landmark Positioning System Using Omnidirectional Vision for Agricultural Vehicle Navigation. International Conference on Intelligent System Design and Engineering Application, 665-669.
- Montalvo M., Guerrero J.M., Guijarro M., Romeo J., Herrera P.J., Ribeiro A., Pajares, G. 2011. Camera System geometry for site specific treatment in precision agriculture. In Proc. First RHEA Workshop (G. Rabatel and P. González de Santos), September 9, Montpellier (France).
- Montalvo, M., Pajares, G., Guerrero, J.M., 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.
- Montalvo, M., Guerrero, J.M., Romeo, J., Emmi, L., Guijarro, M., Pajares, G. 2013a. Automatic expert system for weeds/crops identification in images from maize fields. *Expert Systems with Applications*, 40(1), 75-82.
- Montalvo, M., Guerrero, J.M., Romeo, J., Guijarro, M., de la Cruz J.M., Pajares, G. 2013b. Acquisition of Agronomic Images with Sufficient Quality by Automatic Exposure Time Control and Histogram Matching. *Advanced Concepts for Intelligent Vision Systems (ACIVS'13)*. Lecture Notes in Computer Science, 8192, 37-48, Springer International Publishing Switzerland. October 28-31, Poznan, Poland.
- National Instruments. 2014. National Instruments España. Disponible en: <http://spain.ni.com/> (accedido 2014-07-01).
- 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.
- Olsen, H. J. 1995. Determination of row position in small-grain crops by analysis of video images. *Computers and Electronics in Agriculture*, 12, 147–162.
- 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.
- OpenCV. 2014. OpenCV (Open source Computer Vision). Disponible en: <http://opencv.org/> (accedido 2014-07-01).
- Otsu, N. 1979. A threshold selection method from gray-level histogram. *IEEE Trans. System Man and Cybernetics*, 9, 62-66.

- Pajares, G., de la Cruz, J.M. 2004. A wavelet-based image fusion tutorial. *Pattern recognition*, 37, 1855-1872.
- Pajares, G., de la Cruz, J.M. 2007. *Visión por Computador: imágenes digitales y aplicaciones*. RA-MA, Madrid.
- PENTAX. 2014. RICO.image.change. Disponible en: <http://www.ricoh-imaging.co.jp/english/index.html> (accedido 2014-07-01).
- Peña-Barragán, J. M.; Kelly, M.; de Castro, A.I.; Lopez-Granados, F. 2012 Discrimination of crop rows using object-based approaches in UAV images for early site-specific weed management in maize fields. *Proc. of the 1st International Conference on Robotics and Associated High Technologies and Equipment for Agriculture*, Pisa, Italy, 19–21 September 2012, 249–254.
- 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., Benlloch, J. V., Christensen, S. 2000. Colour and shape analysis techniques for weed detection in cereal fields. *Computers and Electronics in Agriculture*, 25, 197-212.
- Piikki, K., Söderström, M., Wetterlind, J., Stenberg, B. 2013. Three-layered soil maps based on sensor measurements. *Precision agriculture '13*, 33-39.
- Pla, F., Sanchiz, J. M., Marchant, J. A., y Brivot, R. 1997. Building perspective models to guide a row crop navigation vehicle. *Image and Vision Computing*, 15, 465–473.
- 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.
- Reid, J.F., Searcy, S.W. 1987. Vision-based guidance of an agricultural tractor. *IEEE Control Systems* 7 (12), 39–43.
- RHEA. 2014. Robot Fleets for Highly Effective Agriculture and Forestry Management (RHEA). European Commission's Seventh Framework Programme. NMP-CP-IP 245986-2 RHEA. Disponible en: <http://www.rhea-project.eu/index.php> (accedido 2014-07-01).
- Ribeiro A., Fernández-Quintanilla C., Barroso J., García-Alegre, M.C. 2005. Development of an image analysis system for estimation of weed. In *Proc. 5th European Conf. On Precision Agriculture (5ECPA)*, 169–174.

- Romeo, J., Pajares, G., Montalvo, M., Guerrero, J.M., Guijarro, M., de la Cruz, J.M. 2013. A new Expert System for greenness identification in agricultural image. *Expert System with Applications*, 40(6), 2275-2286.
- Rosin, P.L., Loannidis, E. 2003. Evaluation of global image thresholding for change detection. *Pattern Recognition Letters*, 24(14), 2345-2356.
- Rouse, J. W., Haas, R.H., Schell, J. A., Deering, D. W. 1973. Monitoring vegetation systems in the Great Plains with ERTS. *Third Earth Resources Technology Satellite-1 Symposium*, NASA SP-351, 309-317.
- Rovira-Más, F., Zhang, Q., y Reid, J. F. 2008. Stereo vision three-dimensional terrain maps for precision agriculture. *Computers and Electronics in Agriculture*, 60, 133–143.
- Rovira-Más, F., Zhang, Q., Reid, J. F., y Will, J. D. 2005. Hough-transform-based vision algorithm for crop row detection of an automated agricultural vehicle. *Journal Automobile Engineering*, 219, 999–1010. Part D.
- Ruiz-Ruiz, G., Gómez-Gil, J., Navas-Gracia, L.M. 2009. Testing different color spaces based on hue for the environmentally adaptive segmentation algorithm (EASA). *Computers and Electronics in Agriculture* 68, 88-96.
- Sainz-Costa, N., Ribeiro, A., Burgos-Artizzu, X., Guijarro, M., y Pajares, G. 2011. Mapping wide row crops with video sequences acquired from a tractor moving at treatment speed. *Sensors*, 11, 7095–7109.
- Scharf, P.C., Oliveira, L., Vories, E.D., Dunn, D., Stevens, G. 2009. Crop sensors for variable-rate nitrogen application to cotton. *ASA-CSSA-SSSA Annual Meeting Abstracts*. ASA-CSSA-SSSA Annual International Meeting, October 5–9, 2008, Houston, Texas, USA. p. 671–672.
- Schneider Kreuznach. 2014. Schneider Kreuznach cath the light. Disponible en: <http://www.schneiderkreuznach.com/> (accedido 2014-07-01).
- Schneider. 2014. *Schneider Optics*. Disponible en: <http://www.schneideroptics.com/>. (accedido 2014-07-01).
- Serrano, J., Shahidian, S., Silva, J.Marques. 2013. Comparing the DUALEM and VÉRIS sensors for mapping soil properties. *Precision agriculture '13*, 25-32.
- Sezgin, M., Sankur, B. 2004. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), 146-165.
- Shearer, S.A., Pitla S.K. 2014. Precision Planting and Crop Thinning. *Automation: The future of Weed Control in Cropping System*, 99-124.

- Shockley, J. M., Dillon, C. R. 2008. Cost Savings for Multiple Inputs with Swath Control and Auto-guidance Technologies. 9th International Conference on Precision Agriculture. Denver, Colorado.
- Shrestha, D.S., Steward, B.L., Birrell, S.J., 2004. Video processing for early stage maize plant detection. *Biosyst. Eng.* 89 (2), 119–129.
- Silva, F.C.S., Molin, J.P. 2013. Real time soil sensing for determination of tropical soils Ph. *Precision agriculture '13*, 41-48.
- Slaughter, D. C., Giles, D. K., y Downey, D. 2008. Autonomous robotic weed control systems: A review. *Computers and Electronics in Agriculture*, 61, 63–78.
- Søgaard, H. T., y Olsen, H. J. 2003. Determination of crop rows by image analysis without segmentation. *Computers and Electronics in Agriculture*, 38, 141–158.
- Sökefeld, M. 2010. Variable Rate Technology for Herbicide Application. *Precision Crop Protection - the Challenge and Use of Heterogeneity*, 335-347.
- Sony. 2014. SONY make.believe. Disponible en: <http://www.sony.co.uk/support/en/product/DCR-PC110E> (accedido 2014-07-01).
- Spacek, L., Burbridge, C. 2007. Instantaneous robot self-localization and motion estimation with omnidirectional vision. *Robotics and Autonomous Systems* 55:667-674.
- Srinivasan, A. 2006. *Handbook of precision agriculture: principles and applications*. CRC Press.
- Stafford, J.V. 2000. Implementing Precision Agriculture in the 21st Century. *Journal of Agricultural Engineering Research*, 76(3), 267-275.
- Subramanian, V., Burks, T.F., Arroyo, A.A. 2006. Development of Machine Vision and Laser Radar Based Autonomous Vehicle Guidance Systems for Citrus Grove Navigation. *Computers and Electronics in Agriculture*, 53(2), 130-143.
- SVS-VISTEK. 2014. *The Focal Point of Machine Vision*. Disponible en: <http://www.svs-vistek.com/> (accedido 2014-07-01).
- Tellaache, A., Burgos-Artizzu, X.P. 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.

- Tellaeché, A., Pajares, G., Burgos-Artizzu, X. P., y Ribeiro, A. 2011. A computer vision approach for weeds identification through support vector machines. *Applied Soft Computing*, 11, 908–915.
- Tian, L.F., Slaughter, D.C., 1998. Environmentally adaptive segmentation algorithm for outdoor image segmentation. *Computers and Electronics in Agriculture*, 21(3), 153–168.
- Torres-Sánchez J., López-Granados F., De Castro A.I., Peña-Barragán J.M. 2013*a*. 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. 2013*b*. Imagery from unmanned aerial vehicles for early site specific weed management. *Precision Agriculture* (Stafford, J.V.; Ed.), vol. 13, pp. 193-199.
- 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.
- Tsai, R. 1987. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4), 323–344.
- Van Huffel, S., Vanderwalle, J. 1991. *The total least squares problem: computational aspects and Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, PA.
- Vioix, J. B., Douzals, J. P., Truchetet, F., Assemat, L., y Guillemain, J. P. 2002. Spatial and spectral method for weeds detection and localization. *EURASIP JASP*, 7, 679–685.
- Vrindts, E., De Baerdemaeker, J., Ramon, H. 2002. Weed detection using canopy reflection. *Precision Agriculture*, 3(1), 63-80.
- Wei, Z., Shangfeng, D. 2006. Machine vision recognizing position baseline in cropland. *Journal of China Agricultural University*, 11(4), 75-77.
- Weis, M., Andújar, D., Peteinatos, G.G., Gerhards, R. 2013. Improving the determination of plant characteristics by fusion of four different sensors. *Precision agriculture '13*, 63-69.
- Wilson, J.N. 2000. Guidance of agricultural vehicles – a historical perspective. *Computers and Electronics in Agriculture*, 25(1–2), 3-9.

- Woebbecke, D.M., Meyer, G.E., Von Bargaen, K., Mortensen, D.A. 1995*a*. Color indices for weed identification under various soil, residue and lightning conditions. *Transactions of ASAE*, 38(1), 259-269.
- Woebbecke, D.M., Meyer, G.E., Von Bargaen, K., Mortensen, D.A. 1995*b*. Shape features for identifying young weeds using image analysis. *Trans. Am. Soc. Agric. Eng.* 38 (1), 271–281.
- Wu, X., Xu, W., Song, Y., Cai, M. 2011. A Detection Method of Weed in Wheat Field on Machine Vision. *Procedia Engineering*. 15, 1998-2003.
- Xiang, H., y Tian, L. 2011. An automated stand-alone in-field remote sensing system (SIRSS) for in-season crop monitoring. *Computers and Electronics in Agriculture*, 78(1), 1-8.
- 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, X., Shi, L., Jia, X., Seielstad, G., Helgason, C. 2010. Zone mapping application for precision-farming: a decision support tool for variable rate application. *Precision Agriculture* 11 (2), 103-114.
- 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.
- 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.
- Zhou, J., Wang, X., Zhang, R., Feng, Q., Ma, W. 2013. Automatic Navigation Based on Navigation Map of Agricultural Machine. *Computer and Computing Technologies in Agriculture VI. IFIP Advances in Information and Communication Technology*, 304-311.