

Máster en Ingeniería de Sistemas y Control

Generación de mapas de cultivo a partir de secuencias de vídeo adquiridas desde un tractor en movimiento con velocidad de tratamiento

Nadir Sainz Costa

Directores: Ángela Ribeiro Seijas
Gonzalo Pajares Martinsanz

Curso 2010/2011

Convocatoria de Septiembre

Máster en Ingeniería de Sistemas y Control

Generación de mapas de cultivo a partir de secuencias de vídeo adquiridas desde un tractor en movimiento con velocidad de tratamiento

Proyecto Fin de Máster tipo B: proyecto específico propuesto por el alumno.

Estudiante: Nadir Sainz Costa

Directores: Ángela Ribeiro Seijas

Gonzalo Pajares Martinsanz



Autorización

Autorizamos a la Universidad Complutense y a la UNED a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a sus autores, tanto la memoria de este Trabajo Fin de Máster, como el código, la documentación y/o el prototipo desarrollado.

Firmado:

Firma del alumno

Resumen

En este proyecto se presenta un método automático de generación de mapas de campos de cultivo empleando para ello secuencias de video grabadas mediante el uso de una cámara montada sobre el techo de un tractor. Este tractor realizaba operaciones reales de aplicación de herbicida en campos cultivados de forma comercial, sin requerimientos especiales para experimentación, y sobre terreno natural e irregular. Dicha irregularidad hace necesario el diseño de un método de estabilización de imagen adaptado a las características concretas de esta situación: ausencia de puntos fijos o línea del horizonte en los fotogramas, patrón constante de hileras de cultivo, iluminación no controlada. El método expuesto aquí utiliza los centros de las propias hileras de cultivo como puntos de interés a detectar y seguir en todos los fotogramas del video, lo que permite calcular los desplazamientos laterales indeseados debidos a las oscilaciones de la cámara y compensarlos. A la par que se detectan los puntos mencionados, se centra el interés en una región delimitada por éstos y se elimina la perspectiva de los fotogramas en esa zona, lográndose una vista plana y cenital de la misma. Utilizando esta nueva disposición de los fotogramas se pueden medir distancias en ellos y situar precisamente los cultivos y los rodales de malas hierbas presentes sobre un mapa. Este mapa muestra el nivel de certeza de que cada celda en él contenga vegetación y por tanto es útil para estimar la cobertura de malas hierbas en el campo y adecuar los tratamientos a ello. Los mapas así generados pueden ser utilizados como fuente de información para la automatización de tareas agrícolas como la aplicación selectiva de herbicidas, reduciéndose con esto tanto los costes económicos asociados con la agricultura como su impacto ecológico. El algoritmo desarrollado ha sido probado en diferentes videos, grabados en momentos y localizaciones distintas, con resultados satisfactorios. Se logran corregir oscilaciones que suponen hasta el 66% del espacio entre hileras de cultivo y éstas, rectas y paralelas sobre el terreno, se muestran de la misma forma en los mapas generados.

Palabras clave

Mapeo de cultivos, estabilización de video, agricultura de precisión, procesamiento de video, transformación de perspectiva.

Índice

Índice de figuras	2
1. Introducción	4
2. Material y métodos	9
2.1. Adquisición de las imágenes.....	9
2.2. Segmentación de las imágenes	11
2.3. Operaciones morfológicas	13
2.4. Detección de los cultivos	16
2.5. Seguimiento de los cultivos	18
2.6. Transformación de perspectiva	19
2.7. Generación de mapas del terreno	23
3. Resultados	24
3.1. Comentarios generales	24
3.2. Justificación del proceso estabilizador	24
3.3. Mapas generados.....	26
4. Conclusiones	29
5. Agradecimientos	31
6. Referencias	32
Anexo I: Secuencias de fotogramas	35

Índice de figuras

- Figura 1. (a)** Movimientos de oscilación y giro de la cámara desencadenados por la irregularidad del terreno. **(b)** Los cabeceos de la cámara se deben al movimiento longitudinal del tractor debido a los baches, surcos, etc. 7
- Figura 2.** Cámara montada sobre el techo de un tractor utilizada para la adquisición de las secuencias de vídeo empleadas en este trabajo. 9
- Figura 3. (a)** Fotograma tipo original de la secuencias empleadas en el trabajo. Los cultivos no se distinguen claramente en la zona superior de la imagen. **(b)** Mitad seleccionada para el procesamiento del fotograma original. Los cultivos se distinguen bien en esta zona de los fotogramas. 10
- Figura 4. (a)** Fotograma original en espacio de color RGB que muestra las tres hileras de cultivo de interés. **(b)** Fotograma construido mediante la aplicación de la Ecuación 1 al fotograma original de la figura (a). La vegetación se muestra en distintos niveles de gris. 12
- Figura 5. (a)** Fotograma en escala de grises construido a partir del fotograma RGB inicial mediante el método propuesto. **(b)** Fotograma después de la aplicación de la función umbral recogida en la Ecuación 2. La capa vegetal (cultivos y malas hierbas) aparece en blanco sobre fondo negro (resto de la imagen). 13
- Figura 6.** Erosión **(a)** Imagen A de partida. **(b)** Elemento estructural B. **(c)** Resultado de la erosión de A por el elemento B. 14
- Figura 7.** Dilatación **(a)** Imagen A de partida. **(b)** Elemento estructural B. **(c)** Resultado de la dilatación de A por el elemento B. 15
- Figura 8. (a)** Fotograma segmentado binario con cultivos y malas hierbas visibles. **(b)** El mismo fotograma tras la aplicación de la operación morfológica de apertura. La mala hierba ha desaparecido prácticamente y las líneas de cultivo están más definidas. 15
- Figura 9. (a)** Fotograma dividido en cuatro bandas y con la intensidad media vertical mostrada como niveles de gris. Más claro implica mayor vegetación y más oscuro menor cobertura. **(b)** El mismo fotograma después de aplicar una función umbral para retener únicamente las zonas con mayor presencia de vegetación (cultivos) y eliminar las de menor cobertura (malas hierbas). 17
- Figura 10. (a)** Fotograma de bandas (medias de intensidad umbralizadas) mostrando las posiciones de los puntos de interés buscados. **(b)** Fotograma segmentado binario con las posiciones de los puntos de interés calculados por el algoritmo. 18
- Figura 11.** Muestra de la búsqueda de centros limitada a bandas de 120 píxeles de anchura en torno a las últimas posiciones conocidas de los puntos de interés. 19
- Figura 12.** Esquema del modelo de proyección central o cónica empleado para representar en dos dimensiones un mundo en tres dimensiones. Las líneas paralelas en la dirección del eje óptico convergen en el *punto de fuga*. 20
- Figura 13.** Correspondencias puntuales elegidas en la imagen en perspectiva para computar la matriz de homografía H . Los cuatro puntos elegidos son colineales dos a dos por lo que cumplen los requisitos. 22

Figura 14. (a) Fotograma en perspectiva con el trapecio seleccionado por el algoritmo como región de interés. (b) Vista cenital de la región de interés seleccionada y mostrada en (a). La perspectiva ha sido eliminada y se recupera el paralelismo de los cultivos.....	22
Figura 15. Gráfica de la distancia entre el centro horizontal de los fotogramas y la hilera central del cultivo a una altura de imagen de 215 píxeles (desde la parte superior) para una secuencia de 306 fotogramas.....	26
Figura 16. Comparación de los mapas generados antes (a) y después (b) de la estabilización para un video de 306 fotogramas grabado en un día parcialmente cubierto.	28
Figura 18. Ampliación del mapa generado (mostrado en la Figura 16b) para un fragmento de unos dos metros de longitud de terreno. Los centros del cultivo son blancos (máxima certeza de su posición correcta) y los rodales aparecen en distintos tonos de gris.	29
Figura 17. Comparación de mapas generados antes (a) y después (b) del proceso de estabilización para un video de 275 fotogramas grabado en un día soleado.	28
Figura 19. Secuencia de fotogramas pertenecientes a una grabación en un campo en día parcialmente cubierto. La irregularidad del terreno provoca significativas oscilaciones laterales en las posiciones de las hileras de cultivo que requieren estabilización.	35
Figura 20. Secuencia de fotogramas en vista cenital y estabilizados correspondientes a los mostrados en la Figura 14. Ahora las hileras de cultivo permanecen en posiciones fijas y estables en todos los fotogramas.....	36
Figura 21. Secuencia de fotogramas en la que se ve como el algoritmo es capaz de detectar el cultivo pese a que éste desaparece en uno de los fotogramas. (a) Fotograma 18 con el cultivo presente y detectado. (b) Fotograma en el que el cultivo es inexistente en la parte inferior de la línea central. El algoritmo no puede por tanto detectarlo. (c) Fotograma en el que el cultivo reaparece y es detectado de nuevo por el algoritmo.....	37
Figura 22. Secuencia de fotogramas en la que se aprecia un error en la detección de los cultivos debido a la presencia de un rodal de malas hierbas cerca de la posición donde debería estar el cultivo. El sistema falla en la detección en el fotograma 140 y se mantiene dicho error durante varios fotogramas, hasta que en el número 183 se recupera la normalidad al detectarse de nuevo el cultivo.....	38

1. Introducción

La Agricultura de Precisión es un conjunto de técnicas de manejo del suelo agrícola que pretende cuantificar y gestionar los cultivos de una zona de forma diferenciada y en función de la variabilidad natural de cada unidad de terreno. Los objetivos principales de la Agricultura de Precisión son la optimización del manejo del terreno, aumentando la sostenibilidad y la eficiencia. Con esta estrategia se pretende reducir los costes económicos y la huella ecológica tradicionalmente asociados a la agricultura.

La Agricultura de Precisión comenzó a desarrollarse en la década de los setenta con estudios sobre la automatización de máquinas agrícolas. Posteriormente, ya en las décadas de los ochenta y noventa, la liberalización del sistema de posicionamiento global por satélite (GPS) permitió la utilización de la geolocalización con fines civiles, lo que supuso el desarrollo de todo un nuevo conjunto de técnicas de análisis y manejo localizado de las prácticas agrícolas que permitió aumentar la eficiencia en el uso de recursos y la reducción de costes. Todo esto ha conducido a lo que hoy en día es la Agricultura de Precisión, la optimización del uso de insumos (ya sean semillas, herbicidas, fertilizantes o maquinaria agrícola) para adecuarlo a la variabilidad espacial de la producción [1].

En cuanto a los tratamientos, la Agricultura de Precisión pretende evitar la aplicación homogénea de un tratamiento al campo completo independientemente de las características particulares que presente el cultivo en cada zona. El tratamiento homogéneo con dosis fijas es una práctica habitual en la agricultura tradicional pese a que las pruebas muestran que las malas hierbas más perniciosas tienden a aparecer concentradas en rodales y no distribuidas de forma homogénea por todo el campo. Marshall y *col.* investigaron la presencia de tres tipos distintos de malas hierbas en cultivos arables y mostraron que entre el 24 % y el 80 % del área muestreada estaba libre de malas hierbas [2]. En [3] se muestra que una media del 30 % del área estudiada en 12 campos de cultivo (siete de maíz y cinco de soja) se mantenía libre de plantas dicotiledóneas, mientras que un 70 % de dicha área no contenía gramíneas en el espacio entre hileras de cultivo donde no se había aplicado previamente herbicida. En estas situaciones es donde los mapas precisos que muestren no sólo la localización, sino también la densidad de los rodales de malas hierbas pueden tener, y de hecho tienen, numerosas aplicaciones. Éstas incluyen la monitorización de la efectividad de las estrategias de tratamiento de la mala hierba, pasadas o presentes, la comprensión de las dinámicas de población de las distintas especies de malas hierbas o la verificación de las predicciones dadas por modelos existentes. Un uso concreto es su utilización como fuente de información para los sistemas de pulverización de herbicida, que mediante un receptor GPS puedan determinar la posición del equipo en la parcela y aplicar los tratamientos donde el mapa de distribución de infestaciones así lo aconseje. Estos sistemas de información espacial deben tener la capacidad de permitir que los agricultores ajusten con precisión las cantidades y las características de las zonas de aplicación [4].

Existen numerosos trabajos de investigación que muestran las grandes ventajas de la aplicación de tratamientos selectivos, así, se han documentado reducciones en las cantidades de herbicidas aplicados de entre un 42 % (para campos de maíz y soja) y un 84 % (para campos de maíz únicamente), dependiendo de la distribución en rodales y su cantidad [4, 5]. Estos ahorros de herbicidas se traducen en una media de 33 €/ha (Euros por hectárea) por año; cantidad que quedaría disponible para destinar a los costes adicionales del muestreo, el procesado de la información obtenida de los muestreos y los sistemas de pulverización selectiva del herbicida [6].

Suelen diferenciarse dos etapas principales dentro del proceso englobado en la Agricultura de Precisión: la fase de adquisición de información y la fase de aplicación de tratamientos. La automatización de la segunda etapa se basa en la información recogida durante la primera fase y por tanto ésta debe ser lo más precisa posible. Para lograr el objetivo de la Agricultura de Precisión a un coste aceptable que no comprometa la viabilidad económica de ésta, es fundamental la automatización de la etapa de adquisición de información que actualmente es la etapa más costosa del proceso completo.

Desde los inicios de la Agricultura de Precisión la obtención de información visual del terreno se ha realizado de varias formas distintas, dependiendo de las tecnologías disponibles en cada momento. Las técnicas empleadas para la adquisición de información se pueden dividir en dos grupos principales: aquellas que emplean medios aéreos (satélites o aeronaves) y las que tiene lugar a nivel de suelo. El uso de aviones comenzó en los años ochenta, cuando el procedimiento habitual era la toma de imágenes mediante cámaras que fotografiaban en la banda del infrarrojo cercano (conocidas como cámaras NIR) y su posterior análisis para elaborar mapas aproximados de la distribución de malas hierbas [7, 8]. En los noventa se generalizó el uso de imágenes procedentes de satélites. A finales de los años noventa los métodos aéreos cayeron en desuso debido a la aparición de ordenadores más potentes que permitían el análisis directo de fotografías tomadas a nivel de suelo, aunque hoy en día están viviendo un resurgimiento gracias al uso de cámaras híper y multi-espectrales que permiten obtener mayor cantidad de información del cultivo y favorece la discriminación de especies en algunos casos [9]. Sin embargo estos métodos siguen presentando el problema del elevado coste y la baja resolución, con píxeles que representan más de un metro cuadrado de superficie. Además los muestreos aéreos son fuertemente dependientes de las condiciones meteorológicas, ya que sólo pueden tomarse imágenes aéreas en días despejados.

Otra forma de muestreo consiste en adquirir información desde el suelo lo que permite obtener imágenes con una resolución inferior al centímetro. En este caso se distinguen dos formas de muestreo: a pie y mediante dispositivos embarcados en vehículos. El muestreo a pie puede consistir, por ejemplo, en la toma de imágenes del terreno en posiciones previamente fijadas para su posterior interpretación. Requiere el trabajo de varios operarios cualificados dadas las grandes superficies de los campos de cultivo y es por tanto una actividad muy costosa tanto en tiempo como en dinero. Incluso obviando este factor, la información se obtiene mediante rejillas de muestreo (celdas de cierto

tamaño) y es, por tanto, discreta [10]. Colliver y *col.* calcularon el tiempo necesario para muestrear manualmente y elaborar un mapa de la presencia de avena loca en un campo en 3.75 h/ha (horas por hectárea) [11]. En función del tamaño de las celdas de muestreo empleadas estos valores pueden variar bastante. En [12] se utilizó una rejilla de 20 m x 20 m resultando en un tiempo estimado por hectárea de 4.36 horas, mientras que en [13] el tamaño de las celdas era de 36 m x 50 m, lo que redujo el tiempo necesario de muestreo a 2.5 horas por hectárea. Con estos tiempos la reducción de costes derivados de la menor cantidad de herbicida aplicado queda anulada por el excesivo coste del muestreo e incluso incrementa el coste del tratamiento selectivo frente al convencional.

Por otro lado, la adquisición de información mediante plataformas móviles requiere un único trabajador a la par que permite una obtención continua de datos. El muestreo continuo consiste en la adquisición de datos de toda el área muestreada, en oposición al muestreo discreto a pie donde la información es adquirida en ciertas posiciones predefinidas y se emplean métodos de interpolación para generalizar dicha información al total del terreno. El muestreo continuo genera una información cualitativa de presencia o ausencia y nivel de cobertura de las plantas (bajo, medio, alto) en contraste con los conteos de plantas habituales del muestreo a pie. Yendo un poco más allá, la adquisición de video desde una plataforma móvil puede ser una gran oportunidad para generar mapas precisos de los cultivos y los rodales de malas hierbas, lo que constituye el objetivo primordial de este trabajo. También a través de estos sistemas de detección visual se puede pensar en la discriminación en tiempo real de líneas de cultivo, un objetivo importante para el guiado autónomo de vehículos en el cultivo [14] que aumenta el interés de la aproximación de nivel de suelo a la obtención de información.

El montaje de cámaras sobre plataformas móviles presenta un problema asociado al movimiento sobre terrenos irregulares. Esta irregularidad se transfiere a través del sistema de montaje de la cámara a las imágenes que se están adquiriendo de modo que éstas aparecen movidas haciendo muy complicada la discriminación de elementos de interés, no sólo para los procedimientos computacionales sino en muchas ocasiones incluso para el ojo humano. En otras palabras, las vibraciones y oscilaciones de la cámara reducen la calidad de las imágenes y hacen que el seguimiento de elementos presentes en ellas sea una tarea compleja.

La estabilización de imagen es el proceso de suprimir, total o parcialmente, estos movimientos no deseados de las imágenes grabadas con anterioridad. Es un proceso muy importante y que en ocasiones debe aplicarse al inicio del procesamiento si las imágenes han sido adquiridas mediante cámaras en movimiento. En el caso que se trata aquí los movimientos indeseados son de varios tipos: oscilaciones, giros, cabeceos y vibraciones. Las oscilaciones se generan al variar la altura de un lado del tractor respecto al otro y desplazando por tanto la cámara de lado a lado. Estas suelen ir acompañadas de giros debido a que la cámara va montada sobre el techo, a cierta altura del eje de las oscilaciones. Los cabeceos se deben a la variación de altura del frontal de vehículo respecto a la parte trasera. No son fácilmente perceptibles en las secuencias de video empleadas por la simetría longitudinal del campo de cultivo. Por último, las

vibraciones se producen por el funcionamiento del motor del tractor y la actuación del sistema de dispersión de herbicidas acoplado a la parte trasera del vehículo mediante una toma de fuerza de 700 rpm. En la Figura 1 se representan tres de estos movimientos indeseados.

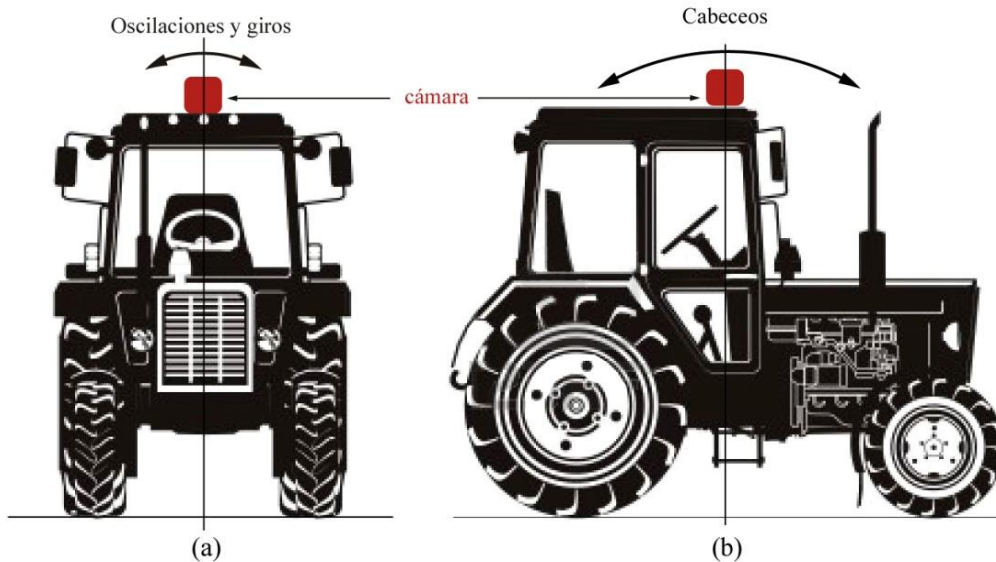


Figura 1. (a) Movimientos de oscilación y giro de la cámara desencadenados por la irregularidad del terreno. (b) Los cabeceos de la cámara se deben al movimiento longitudinal del tractor debido a los baches, surcos, etc.

El problema de la estabilización de imagen se ha tratado ampliamente en la literatura [15 -]. Existen multitud de estudios y técnicas, siendo común la utilización de características aisladas presentes en las imágenes para calcular la magnitud de la inestabilidad. Cuando se trabaja con rasgos cuya posición se supone estable a lo largo de una secuencia de fotogramas (es decir puntos que deben permanecer siempre en lugares fijos) se pueden emplear filtros de Kalmann. Un filtro de Kalmann es un algoritmo matemático que permite producir estimaciones de ciertas medidas a partir de las observaciones hechas en tiempos anteriores y que generalmente son imprecisas por ruido u otros factores. En [15] se extraen ciertas características de las imágenes (vértices de zonas en imágenes aéreas de un mismo lugar) y se predice su trayectoria mediante un filtro de Kalmann lineal. Dado que los puntos extraídos no se deberían mover, los cambios de posición que se miden corresponden al movimiento no deseado. Los autores de [16] proponen un modelo de movimiento de la cámara en el que velocidad y aceleración son lineales. Estos dos parámetros se estiman mediante el uso de un filtro de Kalmann que permite aplicar la corrección necesaria.

Otra técnica frecuente es la estimación del flujo óptico, que se define como el patrón de movimiento aparente de los objetos, superficies y bordes de una escena visual por el movimiento relativo del observador respecto a la escena. Este flujo óptico se puede estimar mediante muchas y variadas técnicas, y una vez hecho los autores de [17 y 18] ajustan un modelo de transformaciones afines entre fotogramas para calcular los parámetros necesarios. En ambos artículos se utiliza un método basado en la

descomposición de las imágenes en pirámides Laplacianas, es decir versiones a distintas escalas de la imagen original, en las que se va restando de cada imagen una copia borrosa de la misma, manteniendo así únicamente las bajas frecuencias, o bordes de la imagen. Es sobre estas imágenes simplificadas que se calcula el flujo óptico y se estima el movimiento entre fotogramas.

Otros trabajos, como [19], calculan primero el desplazamiento en dos dimensiones de una región plana entre dos fotogramas y utilizan este dato para registrar dichas imágenes. Este registro permite cancelar la componente rotacional (en tres dimensiones) de la imagen completa. La traslación de la cámara se calcula buscando la posición del foco de expansión, que es el punto desde el que los objetos parecen fluir y que al depender de la dirección del movimiento permite determinar las traslaciones.

En los estudios mencionados hasta ahora se trabaja con escenas estáticas en las que la cámara se mueve, generalmente de forma ligera. El caso que nos ocupa es muy diferente, pues la cámara se mueve de forma que la escena captada no es estática sino dinámica, cambiante. En otras palabras, la mayoría de los métodos reseñados utilizan características discretas presentes en las imágenes (como pueden ser bordes, vértices, líneas del horizonte o puntos de fuga) y tratan de mantenerlos fijos respecto a un fotograma que se toma como referencia. Este tipo de soluciones no son posibles en el caso que nos ocupa, puesto que en las imágenes no hay características comunes que se mantengan fijas, es decir rasgos que estén presentes en todos los fotogramas y cuyas posiciones se supongan estables. En el caso del muestreo del campo, una porción de terreno se va desplazando en la imagen hasta que desaparece según el tractor recorre el cultivo. Sin embargo, en el contexto de los campos de cultivo sembrados en hileras podemos explotar algunas características del tipo de imágenes, concretamente el hecho de que los cultivos siguen un patrón aproximadamente constante de hileras (o líneas rectas) paralelas y distribuidas uniformemente en el terreno. Algunos investigadores han tratado problemas con características comunes con éste. En [20] se emplean los segmentos, que se pueden considerar casi verticales, presentes en las imágenes para alinearlos con la vertical real (establecida mediante un sensor inercial, por ejemplo) y así estabilizar al menos un tipo de los movimientos no deseados (rotaciones, por ejemplo). En este trabajo se presenta un método para la estabilización de las rotaciones y oscilaciones de secuencias de vídeo grabadas en campos de cultivo de maíz utilizando las propias hileras del cultivo como característica y eliminando la perspectiva presente en las imágenes con el fin de permitir la generación de un mapa de la cobertura vegetal del terreno explorado. El método se ha implementado utilizando el software de código abierto OpenCV [22].

Tras el proceso estabilizador, se genera un mapa del terreno que muestra la cobertura vegetal de éste (cultivos y malas hierbas). Este mapa debe mostrar las hileras de cultivo paralelas y bastante rectas, independientemente de los movimientos indeseados de la cámara debidos a su montaje en un vehículo móvil. La generación de mapas de cultivos ha sido tratada en otros estudios. En [23] se mapean automáticamente las malas hierbas empleando un sistema de tres cámaras bi-espectrales montadas en un vehículo prototipo.

Este es un sistema muy caro debido al uso de cámaras especializadas y a la plataforma móvil de uso específico. La idea detrás del presente trabajo es emplear una cámara doméstica de alta calidad montada en un vehículo agrícola destinado a otras funciones sobre el terreno, de forma que se obtenga una solución de coste más moderado. Otro grupo citado antes, [4], desarrolló un sistema de aplicación selectiva de herbicida en tiempo real que utilizaba la información obtenida mediante el uso de dos a cuatro cámaras instaladas en el frontal de un vehículo agrícola, sin generar un mapa del terreno. Pese a las posibles ventajas del sistema desarrollado por este equipo, un fuerte argumento a favor de la elaboración de un mapa es la posibilidad de utilizarlo para evaluar la efectividad de los tratamientos de una temporada a otra, la extracción de información global sobre la cobertura vegetal del campo o para la distribución de tareas específicas entre una flota de robots.

Por desgracia, ninguno de los intentos reseñados ha logrado, hasta la fecha, concluir con la comercialización de las tecnologías desarrolladas. Los grandes impedimentos son los altos costes económicos y de computación implicados, así como las dificultades para tratar correctamente todas las situaciones posibles en ambientes reales y de exteriores [24].

2. Material y métodos

2.1. Adquisición de las imágenes

Las secuencias de imágenes o fotogramas que se han empleado en este trabajo proceden de grabaciones hechas mediante una videocámara comercial Sony DCR PC110E montada sobre el techo de un tractor, tal y como se muestra en la Figura 2. Este tractor realizaba durante las grabaciones operaciones reales de aplicación de herbicida, moviéndose a una velocidad estándar de tratamiento de unos 6 km/h.



Cámara de vídeo

Figura 2. Cámara montada sobre el techo de un tractor utilizada para la adquisición de las secuencias de vídeo empleadas en este trabajo.

La aplicación de herbicida se lleva a cabo mediante una barra de 10 metros de longitud acoplada a la toma de fuerza del tractor. Esta toma gira a 700 rpm y provoca unas vibraciones aún mayores de las que ya de por sí son habituales en un vehículo agrícola moviéndose en un terreno accidentado, lo que explica la baja calidad de imagen apreciable en los fotogramas y que dificulta el proceso de discriminación de elementos de interés en la imagen.

Las secuencias de vídeo empleadas se grabaron en la época de aplicación de tratamientos herbicidas posteriores a la aparición de malas hierbas (mayo), a lo largo de un periodo de cuatro años. Los vídeos se grabaron en tres campos de maíz diferentes ubicados en las proximidades de Arganda del Rey (Madrid), siendo todos ellos cultivados de forma comercial, sin requerimientos especiales para el trabajo experimental desarrollado. Los campos tienen un área promedio de unas 3 hectáreas, estando las hileras de cultivos separadas unos 75 centímetros y con una densidad de unas 90.000 plantas por hectárea.

La disposición de la cámara sobre el vehículo y el ángulo de ésta enfocando ligeramente hacia el suelo (aproximadamente 10° con la horizontal) provocan que las imágenes adquiridas no contengan, durante la mayor parte del tiempo, la línea del horizonte sino que solamente aparezcan cultivos en ellas. Además la parte superior de las imágenes que muestra las partes lejanas del cultivo no tiene la necesaria definición y en ella las hileras de cultivo no se distinguen unas de otras con claridad. Por este motivo, y para reducir el tamaño de la zona de trabajo y el tiempo de procesamiento, los fotogramas originales se han dividido por la mitad verticalmente y se ha descartado la mitad superior, como se muestra en la Figura 3.

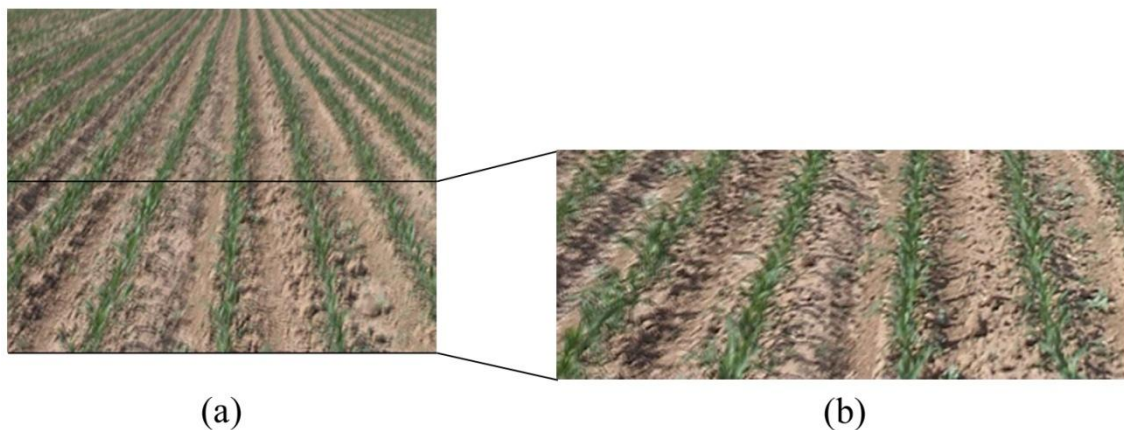


Figura 3. (a) Fotograma tipo original de la secuencias empleadas en el trabajo. Los cultivos no se distinguen claramente en la zona superior de la imagen. (b) Mitad seleccionada para el procesamiento del fotograma original. Los cultivos se distinguen bien en esta zona de los fotogramas.

De esta manera se han obtenido imágenes con unas dimensiones de 720 x 288 píxeles en el espacio de color RGB en las que pueden distinguirse claramente las tres hileras centrales del cultivo sobre las que se centra el procesamiento. A partir de aquí llamaremos fotogramas a estas imágenes de menor tamaño.

2.2. Segmentación de las imágenes

El primer paso del proceso de estabilización que se describe en este trabajo consiste en la separación de la vegetación presente en las imágenes del resto de componentes visibles, por tanto en la segmentación de la capa vegetal. El objetivo es convertir las imágenes de partida (en color) en imágenes binarias (en blanco y negro) en las que la vegetación aparezca en blanco y el resto (suelo desnudo, sombras, paja, etc.) en negro.

En general los píxeles correspondientes a zonas de vegetación, ya sean cultivo o malas hierbas, se caracterizan por tener una componente verde de mayor intensidad frente a las componentes roja o azul. Esta característica se puede emplear para identificar los píxeles de vegetación y separarlos del resto. En nuestro caso logramos este resultado generando un índice de color que representa la intensidad de verde de cada píxel y que se construye a partir de la combinación lineal de los valores de los planos R (rojo), G (verde) y B (azul) de la imagen original tal como se muestra en la ecuación 1.

$$\forall i \in \text{filas}_{\text{imagen}} \wedge \forall j \in \text{columnas}_{\text{imagen}} \rightarrow$$
$$\text{imagen}_{\text{gris}}(i, j) = r \times R_{\text{imagen}}(i, j) + g \times G_{\text{imagen}}(i, j) + b \times B_{\text{imagen}}(i, j) \quad (1)$$

Donde $R_{\text{imagen}}(i, j)$, $G_{\text{imagen}}(i, j)$ y $B_{\text{imagen}}(i, j)$ son los valores de intensidad de los píxeles en la posición (i, j) de los diferentes planos de color de la imagen RGB original. Los coeficientes r , g y b permiten ajustar la combinación obteniendo distintas imágenes en escala de grises donde se resaltan elementos en función del color. Para resaltar la capa vegetal existen distintas propuestas de valores para los coeficientes. Los valores empíricos propuestos en [25] ($r = -1$, $g = 2$, $b = -1$) han demostrado buenos resultados y se utilizan en muchos trabajos. En este caso se han empleado los valores propuestos en [27], obtenidos utilizando un proceso de optimización basado en un algoritmo genético ya que han demostrado discriminar en un gran número de situaciones mejor la capa vegetal que los valores propuestos por Woebbecke y *col.* [25]. En concreto los valores son: $r = -0.884$, $g = 1.262$, $b = -0.311$.

En resumen, la ecuación 1 permite construir una imagen en escala de grises a partir de la imagen original RGB donde los píxeles pertenecientes a la capa vegetal destacan frente al resto. Una muestra del resultado del procesamiento propuesto se muestra en la Figura 4.

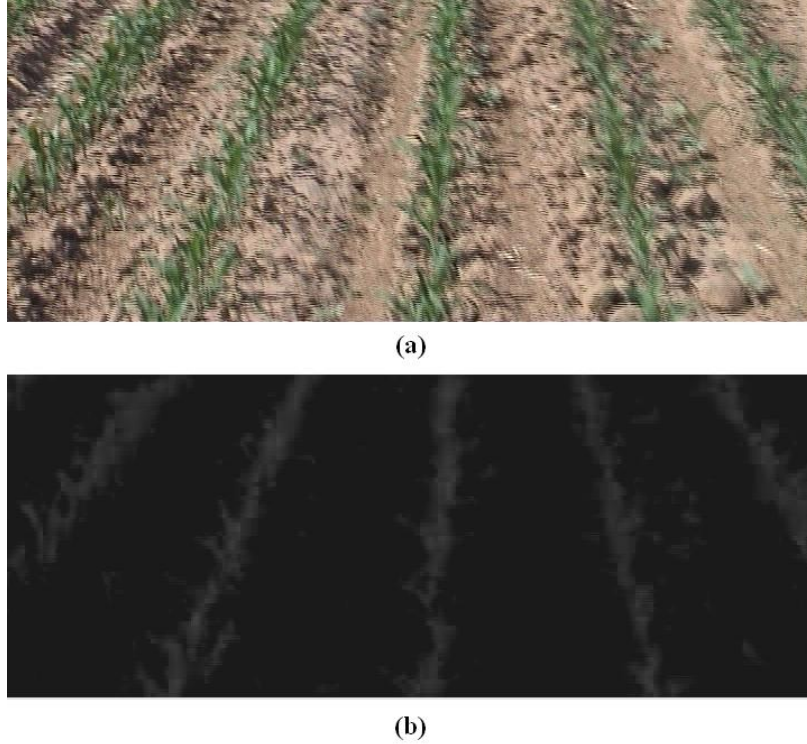


Figura 4. (a) Fotograma original en espacio de color RGB que muestra las tres hileras de cultivo de interés. (b) Fotograma construido mediante la aplicación de la ecuación 1 al fotograma original de la figura (a). La vegetación se muestra en distintos niveles de gris.

Finalmente, la segmentación de la capa vegetal se obtiene aplicando a la imagen en escala de grises una función de umbral (ecuación 2) para retener únicamente los píxeles cuyo nivel de gris es superior a un cierto valor.

$$imagen_{binaria}(i,j) = \begin{cases} 0 & \text{si } imagen_{gris}(i,j) < umbral \\ 1 & \text{si } imagen_{gris}(i,j) > umbral \end{cases} \quad (2)$$

Para determinar el mejor valor para el umbral se han considerado dos métodos: el método de Otsu [28] y el valor medio de intensidad de la imagen.

El método de Otsu permite calcular automáticamente un valor para el umbral que convierte una imagen en escala de grises en una binaria, es decir que separa los píxeles de una imagen en escala de grises en dos regiones diferentes en función de sus valores de intensidad. Su funcionamiento se basa en la dispersión estadística de los valores de intensidad de la imagen o varianza. El método tiene como objetivo que la dispersión en los valores de gris dentro de cada región segmentada (lo que será luego el blanco o el negro) sea lo menor posible a la vez que la varianza (media) entre las dos regiones se mantiene en un valor máximo. Es decir, consiste en maximizar el cociente $Q(t)$:

$$Q(t) = \frac{\sigma_{ext}^2(t)}{\sigma_{int}^2(t)} \quad (3)$$

Este método tiende a fijar un umbral ligeramente superior a la media de intensidad provocando una sobreestimación de la vegetación en la imagen y además su aplicación requiere más tiempo que la intensidad media. El cálculo de este valor medio se muestra en la ecuación 4.

$$umbral = \frac{\sum_{i=1}^{filas} \sum_{j=1}^{columnas} r \times R_{imagen}(i,j) + g \times G_{imagen}(i,j) + b \times B_{imagen}(i,j)}{filas \times columnas} \quad (4)$$

En resumen, una vez tenemos fijado el valor para el umbral, podemos convertir la imagen en escala de grises en una imagen binaria donde la capa vegetal aparece aislada (blanco) del resto de elementos de la imagen (en negro), tal como se muestra en la Figura 5b.

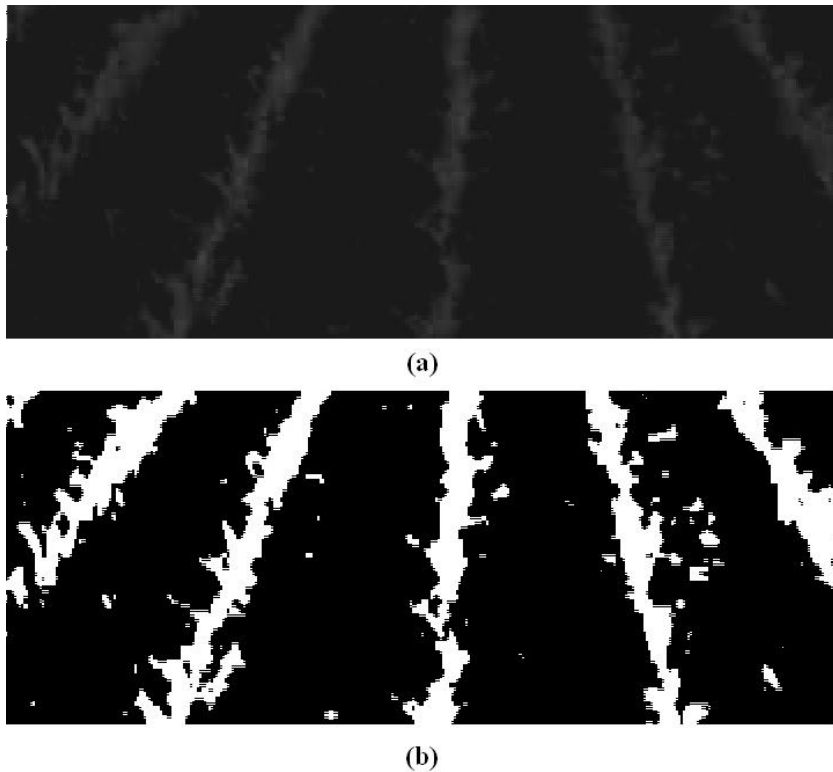


Figura 5. (a) Fotograma en escala de grises construido a partir del fotograma RGB inicial mediante el método propuesto. (b) Fotograma después de la aplicación de la función umbral recogida en la ecuación 2. La capa vegetal (cultivos y malas hierbas) aparece en blanco sobre fondo negro (resto de la imagen).

2.3. Operaciones morfológicas

En este punto del trabajo el objetivo es la estabilización del vídeo y el interés se centra en la detección de las líneas de cultivo. Observando el fotograma de la Figura 5b vemos que éste contiene toda la vegetación presente en la imagen, tanto cultivos como malas hierbas. Para facilitar la tarea de detección de las hileras de cultivo parece interesante reducir o eliminar la presencia de rodales de malas hierbas así como densificar la imagen de las líneas del cultivo, lo que puede conseguirse mediante la aplicación de operaciones morfológicas.

Las operaciones morfológicas son herramientas matemáticas muy útiles en el procesamiento de imagen y su uso está muy extendido en este campo. Básicamente las operaciones morfológicas en el procesamiento de imagen consisten en la convolución de una imagen binaria tomada como matriz de valores de intensidad (0 ó 1) con un cierto elemento estructural, en sí mismo también una matriz binaria, de forma que obtenemos como resultado una imagen transformada en cierta manera respecto a la original. Es posible aplicar un gran espectro de operadores diferentes, cada uno de ellos con unos efectos y características distintos, pero pueden identificarse dos operaciones principales: erosión y dilatación. La erosión produce la eliminación de los píxeles blancos aislados o la reducción de las áreas blancas mientras que la dilatación tiene el efecto contrario, es decir la compactación y expansión de las zonas blancas.

Matemáticamente podemos ver una operación morfológica como la traslación del elemento estructural elegido (una pequeña matriz binaria) a lo largo y ancho de la imagen que se quiere transformar. En forma de expresión matemática, la erosión de una imagen A mediante un elemento estructural B (definidos ambos en un espacio euclídeo E) resulta en la ecuación siguiente:

$$A \ominus B = \{z = (x, y) \in E | B_z \subseteq A\} \quad (5)$$

Donde B_z es la traslación del elemento estructural B por todos los vectores z contenidos en el espacio euclídeo E. En palabras esto quiere decir que la erosión de A por B es el conjunto de puntos $z = (x, y)$ tales que B trasladado el vector z está contenido o es igual a A. Una muestra del resultado de esta operación morfológica se puede ver en la Figura 6.

El caso contrario, la dilatación, puede expresarse con la siguiente ecuación:

$$A \oplus B = \{z = (x, y) \in E | (B^s)_z \cap A \neq \emptyset\} \quad (6)$$

Donde B^s denota el simétrico del elemento B, es decir $B^s = \{x \in E | -x \in B\}$, o los puntos de B cuyas coordenadas (x,y) han sido reemplazadas por (-x,-y). El efecto de la dilatación puede verse en la Figura 7.

La combinación de las operaciones de erosión y dilatación da lugar a dos operaciones con nombre propio: la apertura (opening) y el cierre (closing). Apertura es la aplicación de erosión seguida de dilatación y el cierre la secuencia inversa.

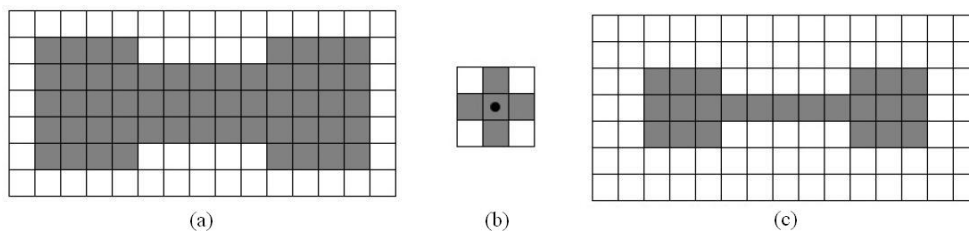


Figura 6. Erosión (a) Imagen A de partida. (b) Elemento estructural B. (c) Resultado de la erosión de A por el elemento B.

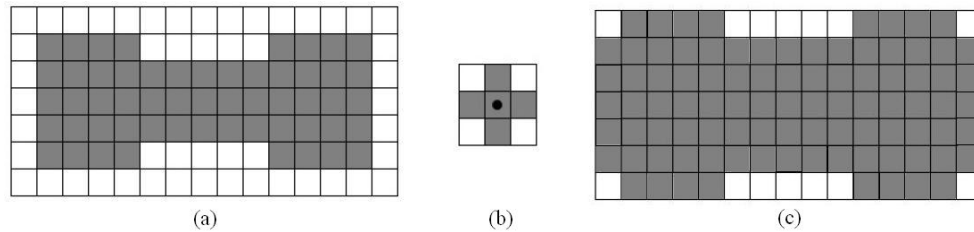


Figura 7. Dilatación (a) Imagen A de partida. (b) Elemento estructural B. (c) Resultado de la dilatación de A por el elemento B.

En este trabajo se ha empleado la apertura para alcanzar el objetivo descrito de reducir o eliminar la presencia de malas hierbas aisladas a la vez que se densifican las zonas blancas correspondientes a líneas de cultivo. Los elementos estructurales empleados aquí son matrices de unos de dimensiones diferentes para la erosión y la dilatación. En el primer caso se utiliza una matriz rectangular de 10 x 15 píxeles y en el segundo una de 12 x 20 píxeles. Estos elementos de forma rectangular y orientación vertical fueron elegidos para coincidir en dirección aproximadamente con las hileras de cultivo. Sus tamaños se fijaron mediante sucesivas pruebas para lograr erosionar lo suficiente como para eliminar la mayor parte de las malas hierbas sin llegar a afectar en exceso al cultivo, y para conseguir, con la dilatación, una expansión de las áreas de cultivo en la dirección vertical que no supusiera un aumento de la anchura de dichos cultivos. Los resultados de la aplicación de las operaciones morfológicas descritas sobre los fotogramas de este trabajo se muestran en la Figura 8.

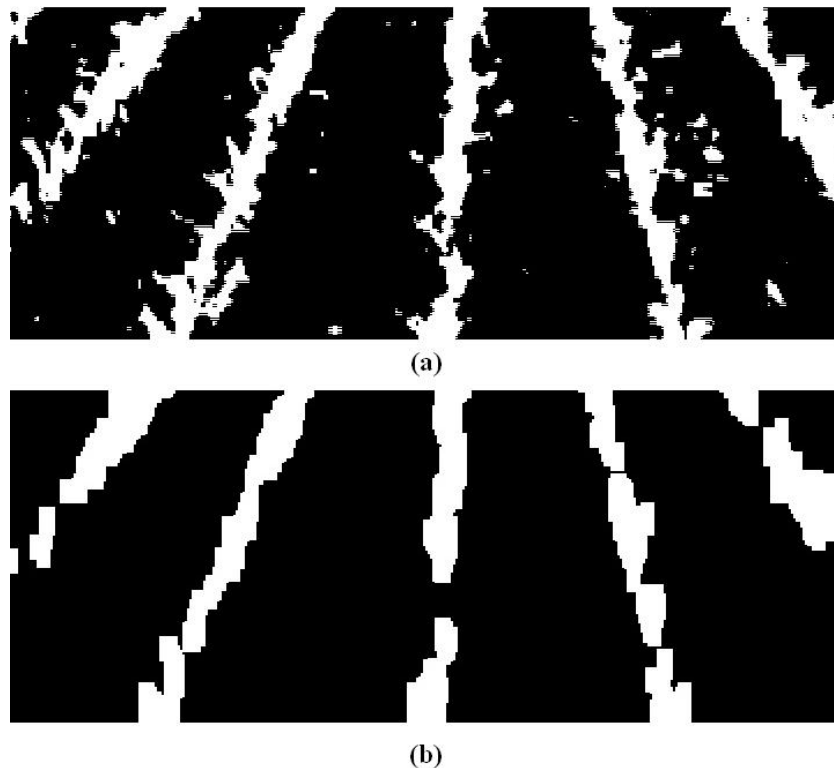


Figura 8. (a) Fotograma segmentado binario con cultivos y malas hierbas visibles. (b) El mismo fotograma tras la aplicación de la operación morfológica de apertura. La mala hierba ha desaparecido prácticamente y las líneas de cultivo están más definidas..

2.4. Detección de los cultivos

Los procesos de estabilización de los movimientos laterales de la cámara se basan generalmente en la compensación de los desplazamientos indeseados que se producen en las posiciones de ciertas características entre fotogramas consecutivos. Para poder medir dichos desplazamientos es necesario conocer dónde están las características seleccionadas en cada fotograma del vídeo de forma que se pueda calcular su cambio de posición debido a las oscilaciones del tractor.

En los fotogramas tipo extraídos de las secuencias de vídeo de este trabajo no suele aparecer la línea del horizonte (frontera entre la tierra y el inicio del cielo), por lo que no se puede contar con esa característica permanente para la estabilización. Debido al constante desplazamiento del tractor (y la cámara sobre su techo), no existen objetos fijos que aparezcan en todos los fotogramas y que puedan por tanto ser tomados como referencias para el cálculo del movimiento entre fotogramas.

Sin embargo se puede hacer uso de las propiedades de simetría y repetitividad del patrón que forman los cultivos para hallar los necesarios puntos a seguir. Las hileras de cultivo, y en especial las tres centrales, están presentes en todos los fotogramas de las secuencias de vídeo. Además sus posiciones laterales deben permanecer fijas si se desea un video estable, por lo que constituyen unas perfectas candidatas como características a detectar y seguir a lo largo de las secuencias. El proceso propuesto para esto se describe a continuación.

Los puntos de interés elegidos son los centros de los cultivos para cuatro alturas en el eje y de la imagen fijas ($y_1 = 36$, $y_2 = 108$, $y_3 = 180$, $y_4 = 252$). Para hallar la coordenada x de estos puntos dividimos los fotogramas que tenemos (Figura 8) en cuatro bandas de ancho fijo (72 píxeles) para las cuales las coordenadas y_i quedan centradas verticalmente. Para cada una de estas bandas de la imagen calculamos la intensidad media de los píxeles por columnas, es decir sumamos los valores de intensidad (0 o 1) de todos los píxeles en cada columna de las bandas y dividimos el total por los 72 píxeles de altura de las bandas. Esta operación da lugar a una intensidad media para cada columna de cada banda del fotograma, que se corresponde gráficamente con un cierto nivel de gris. Una muestra de este resultado intermedio se recoge en la Figura 9a.

Una vez tenemos este resultado intermedio, podemos apreciar que los distintos niveles de gris indican diferentes porcentajes de cobertura vegetal: las columnas más oscuras indican menor presencia de vegetación en ellas mientras que las más blancas corresponden a zonas de mayor cobertura vegetal. Puesto que el interés actual radica en la separación de cultivos (mayor cobertura, próxima al 100%) de los rodales de malas hierbas (baja cobertura vegetal, cercana al cero) parece razonable aplicar una función umbral sobre esta imagen intermedia. De esta forma recuperamos de nuevo una imagen binaria en la que los bloques blancos más anchos representan hileras de cultivo y los más estrechos, si están presentes, corresponden a rodales de mala hierba que raramente

se extienden sobre una distancia vertical significativa. Todo esto lo podemos ver en la Figura 9b.

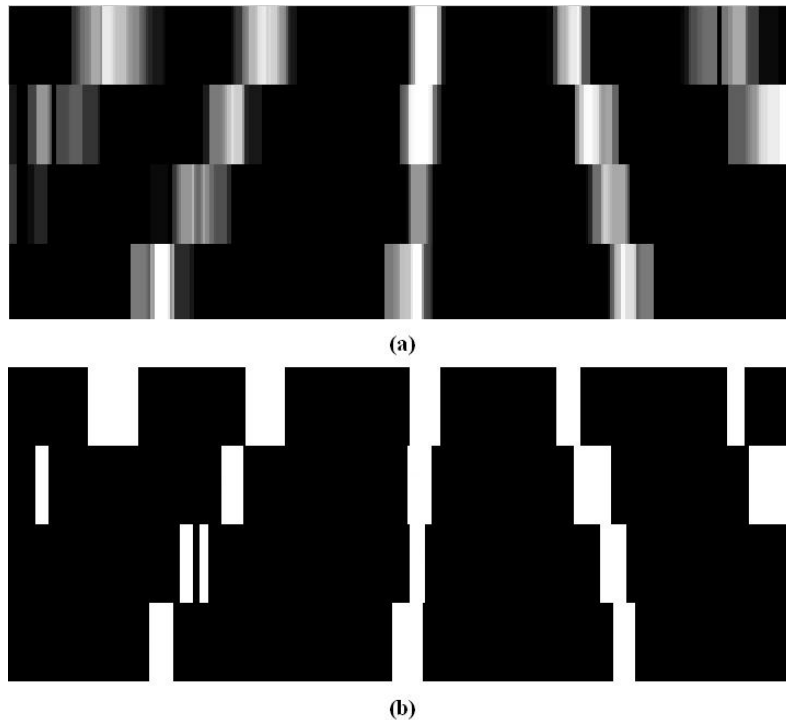


Figura 9. (a) Fotograma dividido en cuatro bandas y con la intensidad media vertical mostrada como niveles de gris. Más claro implica mayor vegetación y más oscuro menor cobertura. (b) El mismo fotograma después de aplicar una función umbral para retener únicamente las zonas con mayor presencia de vegetación (cultivos) y eliminar las de menor cobertura (malas hierbas).

Una vez hemos aplicado estas operaciones sobre los fotogramas tenemos disponibles imágenes en las que las hileras de cultivos están muy bien identificadas y sus centros horizontales son sencillos de calcular. Primero el algoritmo selecciona aquellos bloques que por su anchura y posición en la imagen tienen mayor probabilidad de corresponder a las tres hileras centrales del cultivo y calcula sus centros horizontales, que constituyen la coordenada x de los puntos de interés que necesitamos. En la Figura 10 se han representado los puntos de interés sobre la imagen de bandas (Figura 10a) y sobre la imagen segmentada (Figura 10b) de los cultivos.

Con el conjunto de puntos generados, cuatro sobre cada una de las tres hileras centrales de cultivo, podemos representar las rectas correspondientes a dichos cultivos de forma que cada una de estas tres rectas se superpone a una hilera de cultivos. Estas son las características presentes en todos los fotogramas de las secuencias de vídeo que utiliza el algoritmo para la estabilización.

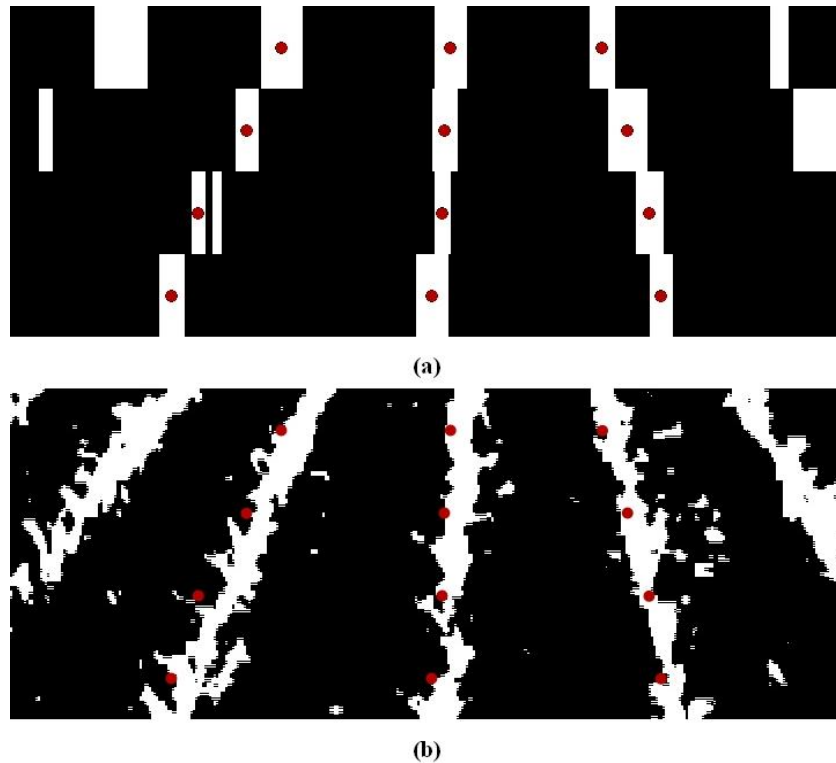


Figura 10. (a) Fotograma de bandas (medias de intensidad umbralizadas) mostrando las posiciones de los puntos de interés buscados. (b) Fotograma segmentado binario con las posiciones de los puntos de interés calculados por el algoritmo.

2.5. Seguimiento de los cultivos

Una vez se han completado los cálculos que se han descrito en el primer fotograma de una secuencia, es preciso establecer un sistema de seguimiento de las hileras de cultivo para todos los fotogramas sucesivos de manera que siempre se detecten las tres centrales pese a los movimientos oscilatorios de la cámara. El método propuesto es sencillo: las posiciones de los centros hallados en el primer fotograma se almacenan en un vector y se emplean en el segundo fotograma para limitar la búsqueda de los nuevos centros en una banda de anchura fija (120 píxeles) en torno a las posiciones conocidas del fotograma anterior. Esto permite evitar falsas identificaciones de líneas de cultivo como centrales cuando no lo son, por ejemplo, en los momentos en que debido a las oscilaciones de la cámara las tres hileras de interés se han desplazado de su posición centrada inicial. También se limita la posibilidad de detectar rodales de malas hierbas y clasificarlos como cultivos. Puede verse un ejemplo de este método en la Figura 11.

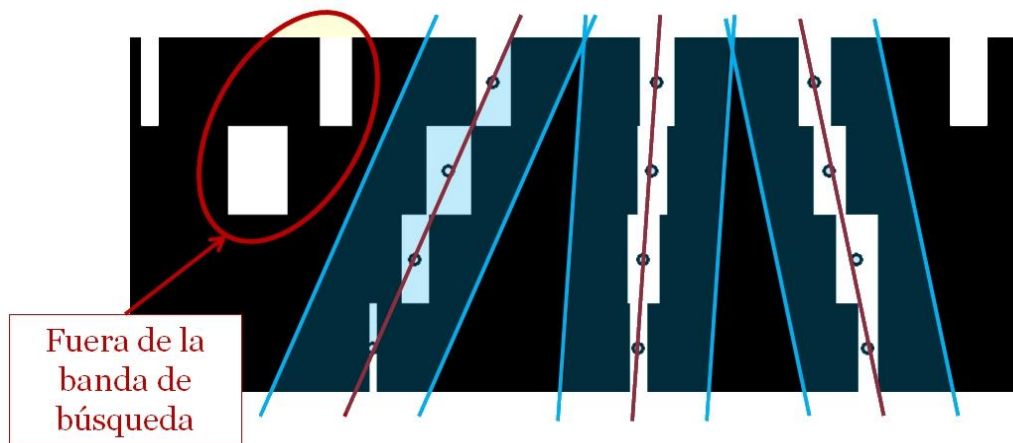


Figura 11. Muestra de la búsqueda de centros limitada a bandas de 120 píxeles de anchura en torno a las últimas posiciones conocidas de los puntos de interés.

El vector de centros se actualiza tras cada fotograma con los nuevos centros encontrados. Sin embargo, si no se detecta alguno de los puntos de interés en un fotograma concreto, en el vector no se modifica la posición almacenada de forma que una ausencia temporal de un punto de interés no perjudica la detección del punto correspondiente en el fotograma siguiente. Este método permite evitar ciertos errores debidos a la ausencia de plantas en las hileras de cultivo durante algunos fotogramas siempre que esta ausencia no se mantenga durante demasiado tiempo (el hueco en la hilera no sea muy extenso) y no venga unida a la presencia de rodales extensos de malas hierbas dentro de la banda de búsqueda que puedan ser clasificados como cultivos y su posición almacenada erróneamente. En el Anexo I (Figura 19 a Figura 22) se muestran secuencias de fotogramas en las que el algoritmo ha tenido éxito y en las que ha fracasado en el seguimiento de los cultivos.

2.6. Transformación de perspectiva

El siguiente paso en el algoritmo propuesto es la eliminación de la perspectiva de los fotogramas y su combinación con la estabilización de las secuencias de video. Las imágenes tomadas mediante cámaras son proyecciones en dos dimensiones de un mundo real de tres dimensiones que responden, por tanto, a un cierto modelo proyectivo. El modelo adecuado para la visión humana o las cámaras fotográficas es la proyección central o cónica, una de cuyas consecuencias es la transformación de las líneas longitudinales paralelas en líneas convergentes en el horizonte. Esto, unido a la no linealidad de la perspectiva, impide la medición directa de profundidades en las imágenes y es la razón que justifica una transformación de perspectiva [30 – 33].

El modelo de proyección cónica o central se caracteriza porque todos los rayos luminosos procedentes de diferentes lugares convergen en un único punto conocido como centro óptico o foco. La lente de la cámara está situada en este centro óptico y se separa del plano de formación de la imagen una distancia f (distancia focal). Sobre este plano se forma la imagen del mundo real mediante la intersección de los rayos luminosos convergentes. La Figura 12 muestra un diagrama de este modelo.

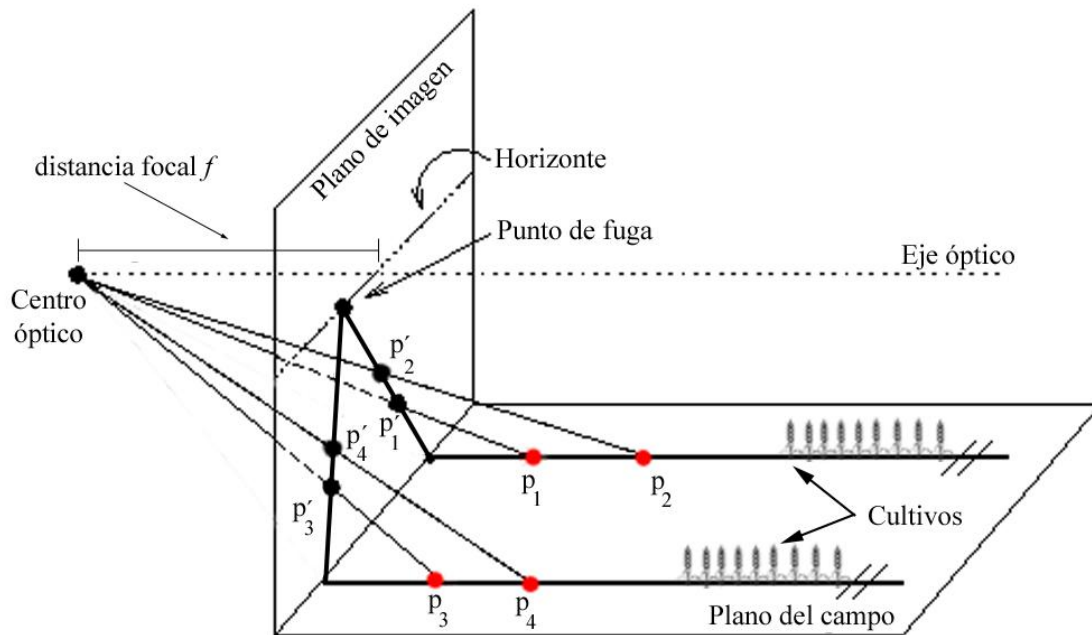


Figura 12. Esquema del modelo de proyección central o cónica empleado para representar en dos dimensiones un mundo en tres dimensiones. Las líneas paralelas en la dirección del eje óptico convergen en el *punto de fuga*.

Los puntos p_1, p_2, p_3 y p_4 situados sobre dos hileras de cultivo paralelas en el campo se proyectan sobre el plano de la imagen siguiendo líneas que convergen en el centro óptico. Sus proyecciones p'_1, p'_2, p'_3 y p'_4 definen hileras de cultivo no paralelas sino convergentes en el plano de la imagen.

Las habituales coordenadas cartesianas presentan un problema cuando se trabaja con imágenes en perspectiva. Todos los puntos situados a lo largo de una línea que pasa por el centro óptico tiene la misma proyección sobre el plano de la imagen y por tanto las mismas coordenadas cartesianas. Sin embargo, la introducción de coordenadas homogéneas [30] permite evitar este problema. En coordenadas homogéneas un punto viene representado por las coordenadas (x, y, z, w) que corresponden a coordenadas cartesianas $(x/w, y/w, z/w)$. De esta manera los puntos en el infinito se representan como el conjunto con $w = 0$.

Puesto que todos los puntos de interés en las imágenes de trabajo se encuentran contenidos en el mismo plano (el del campo de cultivo) es posible emplear coordenadas homogéneas para realizar una correspondencia lineal (invertible) entre dos planos: el del campo y el de la imagen [31]. Esta transformación recibe el nombre de homografía o transformación proyectiva, y se caracteriza por conservar el paralelismo de las líneas. Su expresión matemática se muestra en la ecuación siguiente, para la proyección de un punto \mathbf{P} del campo en el plano de la imagen:

$$\mathbf{u}' = sH\mathbf{u} \quad (7)$$

Aquí \mathbf{u}' son las coordenadas homogéneas (u', v', w') de la proyección \mathbf{P}' en el plano de la imagen, \mathbf{u} las coordenadas homogéneas (u, v, w) de dicho punto \mathbf{P} en el campo, s un

factor de escala y H la matriz de homografía de dimensión 3×3 , es decir la matriz de la transformación que permite obtener la imagen deseada sin perspectiva a partir de la imagen original.

Para el cómputo de esta matriz de homografía existen distintos métodos, pero en visión por computador es muy frecuente emplear el cálculo mediante correspondencias puntuales. Este sistema consiste en determinar un conjunto de n puntos de la imagen en perspectiva cuyas posiciones en una hipotética imagen sin perspectiva sean conocidas. De esta manera es posible plantear un sistema de ecuaciones de la forma recogida en la ecuación 8 que relaciona las coordenadas homogéneas en ambas representaciones gráficas (con y sin perspectiva).

$$s_i \mathbf{u}'_i = H \mathbf{u}_i, (i = 1, \dots, n) \quad (8)$$

Si se fija la dimensión del espacio vectorial de representación en dos (se trabaja con imágenes planas), este sistema cuenta con $(2+1)n = 3n$ ecuaciones y $n + 9 - 1 = n + 8$ incógnitas. Estas incógnitas vienen de las n incógnitas asociadas a los s_i factores de escala, los 9 componentes de la matriz H y el -1 permite determinar H manteniendo un factor global de escala.

El número de pares de correspondencias puntuales necesarias para establecer una correspondencia biunívoca (hasta un factor global de escala) entre los puntos de la imagen en perspectiva y su representación plana (imagen cenital) se puede calcular de forma sencilla.

$$\begin{cases} 3n & \text{ecuaciones} \\ n + 8 & \text{incógnitas} \end{cases} \Rightarrow 3n = n + 8 \Rightarrow n = 8/2 \Rightarrow \mathbf{n = 4} \quad (9)$$

Es preciso aclarar que estas cuatro correspondencias sólo son válidas cuando no más de dos de los puntos elegidos son colineales, es decir que pertenecen a la misma recta.

Para poder establecer cuatro correspondencias de forma automática, es necesario que el algoritmo propuesto seleccione los puntos adecuados de entre los centros de las hileras que se calculan en el paso anteriormente descrito. Se mencionó en el apartado 2.4 que el vector de puntos de interés calculado en cada fotograma de la secuencia contiene 12 centros, cuatro para cada una de las tres hileras centrales de cultivos. Para garantizar que el conjunto de puntos seleccionados no contiene más de dos que sean colineales, a la vez que para facilitar la elección de sus correspondencias en la imagen cenital, el algoritmo extrae del conjunto aquellos que se encuentran en los vértices del trapecio que forman las hileras de cultivo exteriores y que se muestra en la Figura 13.

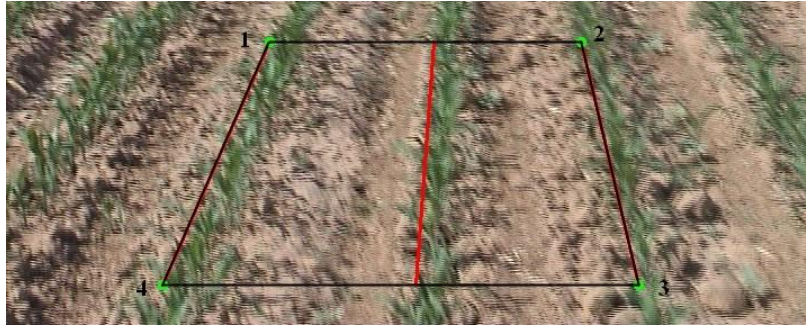


Figura 13. Correspondencias puntuales elegidas en la imagen en perspectiva para computar la matriz de homografía H . Los cuatro puntos elegidos son colineales dos a dos por lo que cumplen los requisitos.

Un trapecio como el que aparece en la imagen corresponde a un rectángulo sobre el terreno, cuya dimensión horizontal es conocida e igual al doble de la separación entre hileras de cultivo (aproximadamente 150 cm) y cuya dimensión vertical se debe establecer para poder calibrar la escala en dicha dirección. Dicho esto, parece interesante establecer los puntos de correspondencia de los vértices del trapecio en las esquinas de un rectángulo de dimensiones conocidas, en nuestro caso de 720 x 576 píxeles. De esta manera, con las correspondencias puntuales necesarias ya fijadas, el algoritmo computa la matriz de homografía adecuada y se posibilita así la conversión de la imagen en perspectiva en una vista cenital del terreno incluido en el trapecio de la Figura 13. Esta vista cenital se muestra en la Figura 14.



Figura 14. (a) Fotograma en perspectiva con el trapecio seleccionado por el algoritmo como región de interés. (b) Vista cenital de la región de interés seleccionada y mostrada en (a). La perspectiva ha sido eliminada y se recupera el paralelismo de los cultivos.

Gracias a la transformación de perspectiva se recupera el paralelismo entre las hileras de cultivo y es posible realizar medidas en las imágenes en todas direcciones, tanto transversales como longitudinales (profundidad). Esto constituye el objetivo de la eliminación de la perspectiva de las imágenes. Además, puesto que los vértices del trapecio seleccionado siempre se hacen corresponder con las mismas cuatro esquinas de la imagen cenital, se logra en una misma operación la eliminación de la perspectiva y la estabilización de los movimientos indeseados de la cámara. Los puntos de interés elegidos siempre están situados sobre las mismas hileras centrales de cultivo pese a que

éstas oscilen lateralmente y cambien de posición, y al ser mapeados a unos puntos fijos para todos los fotogramas, el movimiento lateral de los cultivos se suprime.

2.7. Generación de mapas del terreno

El objetivo de este trabajo es la generación de un mapa a escala del terreno que muestre la presencia de cultivos y malas hierbas con el fin de poder adecuar los tratamientos herbicidas necesarios. Ahora que ya se dispone de una secuencia de fotogramas en vista cenital, que permite la toma de medidas y la estimación de posiciones en un sistema de referencia cartesiano, se puede integrar la información contenida en cada fotograma en un mapa completo del terreno recorrido por el vehículo portador de la cámara.

En este caso el mapa es de tipo rejilla y está formado por una matriz de dimensión adecuada al tamaño del terreno a explorar y cuyos elementos corresponden a celdas de tamaño determinado sobre el terreno. Esto permite adecuar el factor de escala de dichas celdas a la precisión requerida manteniendo un tamaño de mapa manejable. Se ha elegido rellenar las celdas del mapa mediante un histograma de ocupación, es decir realizando un conteo de la cantidad de fotogramas de la secuencia que presentan vegetación para cada celda. De esta forma, aquellas celdas que contienen un número entero más alto implican una mayor seguridad o certeza sobre la presencia de vegetación en el área que representan. Para cada fotograma segmentado (mostrando cultivos y malas hierbas), cuando se encuentra un píxel blanco se incrementa el valor de la celda correspondiente a esa posición del terreno. Debido al avance continuo del vehículo que porta la cámara es necesario ir actualizando la posición del origen de coordenadas respecto al cual se calcula la celda del mapa que corresponde a cada píxel. La cantidad de píxeles a avanzar el sistema de referencia depende de la velocidad del vehículo. Puesto que aquí se trabaja con secuencias ya grabadas y por tanto dicha velocidad no es modificable, se ha tomado la velocidad estándar de tratamiento (6 km/h ó 1,667 m/s) y se ha estimado que el avance necesario es de 6 píxeles por cada fotograma.

Una vez se tiene un mapa completo de una zona del terreno, la matriz correspondiente contiene valores enteros que van desde el cero (celdas donde no se ha detectado vegetación en ningún fotograma) hasta un valor máximo variable. Esta matriz puede tratarse como una imagen en escala de grises donde los valores más altos se convierten en tonos más próximos al blanco y los ceros se convierten en negro. Una forma de asegurar un mínimo de certeza en la presencia de vegetación en el mapa y reducir posibles errores de segmentación en fotogramas aislados es aplicar una función de umbral a los valores contenidos en las celdas del mapa. Los valores típicos de las celdas del mapa se mueven entre 0 (ningún píxel con vegetación) y 17. Un valor que se ha considerado adecuado para el umbral de certeza es aproximadamente el 25% del máximo. Así, todas aquellas celdas con valores inferiores a cuatro se han tomado como posibles errores y aquellas celdas con valores inferiores se han mostrado como zonas sin vegetación, es decir en negro. Esta representación ofrece una vista rápida de la totalidad del terreno cubierto por el mapa y detallado en la matriz asociada.

3. Resultados

3.1. Comentarios generales

El algoritmo descrito en las secciones anteriores ha sido probado con tres secuencias de video diferentes, que como se ha mencionado en la sección 2.1 se grabaron en campos distintos aunque de similares características, en días distintos y con diferentes condiciones de iluminación (días soleados y parcialmente nublados). Todos estos campos eran cultivos comerciales, sin requerimientos específicos para la experimentación, con maíz y se encontraban en la época de aplicación de tratamientos posteriores a la aparición de malas hierbas, en torno al mes de mayo.

Las secuencias tomadas muestran una calidad de imagen mejorable debido a las vibraciones, muy significativas, generadas por el funcionamiento del mecanismo de aplicación de herbicida (barra de pulverización) acoplado a la parte trasera del tractor sobre el que se fijó la cámara. Esto se une a los ya de por sí acusados movimientos del vehículo causados por la irregularidad del terreno típico de un campo de cultivo.

Pese a estos inconvenientes, el algoritmo ha demostrado un funcionamiento bastante satisfactorio en todas las situaciones probadas. En general el algoritmo es capaz de detectar las hileras de cultivo, extraer los puntos de interés necesarios y estabilizar los videos (eliminando la perspectiva) para la mayoría de fotogramas de las secuencias disponibles. Se han producido algunos errores de detección y seguimiento de los cultivos que dan lugar a errores en la estabilización y conversión a vista cenital de los fotogramas, pero del total de fotogramas probados (alrededor de 1000) sólo un 7% se han visto afectados por estos errores.

En las figuras del Anexo I se muestran fragmentos de las secuencias de video empleadas durante este trabajo antes y después del proceso de estabilización. La secuencia mostrada en la Figura 19 pertenece a una grabación realizada durante un día parcialmente cubierto y sobre un terreno muy irregular. Se ve que las oscilaciones de las hileras de cultivo en este video son muy apreciables antes de la estabilización. Los cultivos se desplazan fuera de la posición central, especialmente hacia la derecha de la imagen, una distancia que prácticamente hace que la hilera izquierda se sitúe en el centro de los fotogramas. La estabilización es claramente necesaria. En la Figura 20 se recoge la misma secuencia una vez aplicado el proceso de estabilización y eliminación de perspectiva. En este caso las hileras de cultivo se mantienen siempre centradas en la imagen, puesto que la región de interés se sitúa perfectamente encima de ellas. La distancia de la hilera central de cultivo al centro de los fotogramas se mantiene constante.

3.2. Justificación del proceso estabilizador

Como medida de la importancia de la estabilización propuesta en este tipo de secuencias de vídeo, se ha medido la distancia horizontal entre la hilera central del cultivo y el

centro de los fotogramas en una de las secuencias empleadas en la que la inestabilidad era más evidente. Idealmente, si los videos fueran perfectamente estables, esta distancia debería permanecer constante y ser cercana a cero durante todo el video, incluso en las imágenes en perspectiva. Para tomar siempre la misma referencia de distancia, ésta se ha medido a una altura fija de las imágenes e igual a 215 píxeles desde la esquina superior izquierda. En la Tabla 1 se recogen los valores estadísticos derivados de estas medidas. Se ve que los cultivos se desplazan de sus posiciones ideales hasta un máximo de 93 píxeles hacia la derecha y 26 hacia la izquierda. Siendo la separación entre líneas de cultivo de aproximadamente 140 píxeles en la parte superior y 220 píxeles en la parte inferior, el desplazamiento máximo de 93 píxeles hacia la derecha supone entre un 42% y un 66% de dicha separación entre cultivos. A la hora de elaborar un mapa del terreno estos desplazamientos son claramente excesivos y dan lugar a grandes errores.

Variable	Valor en píxeles
Media de distancia	20
Desviación estándar	28.5
Máxima desviación hacia la derecha	93
Máxima desviación hacia la izquierda	26

Tabla 1. Datos estadísticos de la medida de distancia entre el centro de los fotogramas y la hilera central de cultivos para una secuencia de 306 fotogramas.

En la Figura 15 se muestra una gráfica de los valores para esta medida de distancia entre el centro y la hilera central de cultivo para los 306 fotogramas de una de las secuencias de video de la colección de trabajo.

A la vista de estos resultados gráficos y estadísticos parece claro que la estabilización diseñada y aplicada constituye una media de importancia suficiente para justificar su utilización incluso teniendo en cuenta el aumento de costes computacionales asociados al proceso. Es un paso previo necesario para poder llevar a cabo el mapeo del terreno comprendido entre los cultivos con garantías.

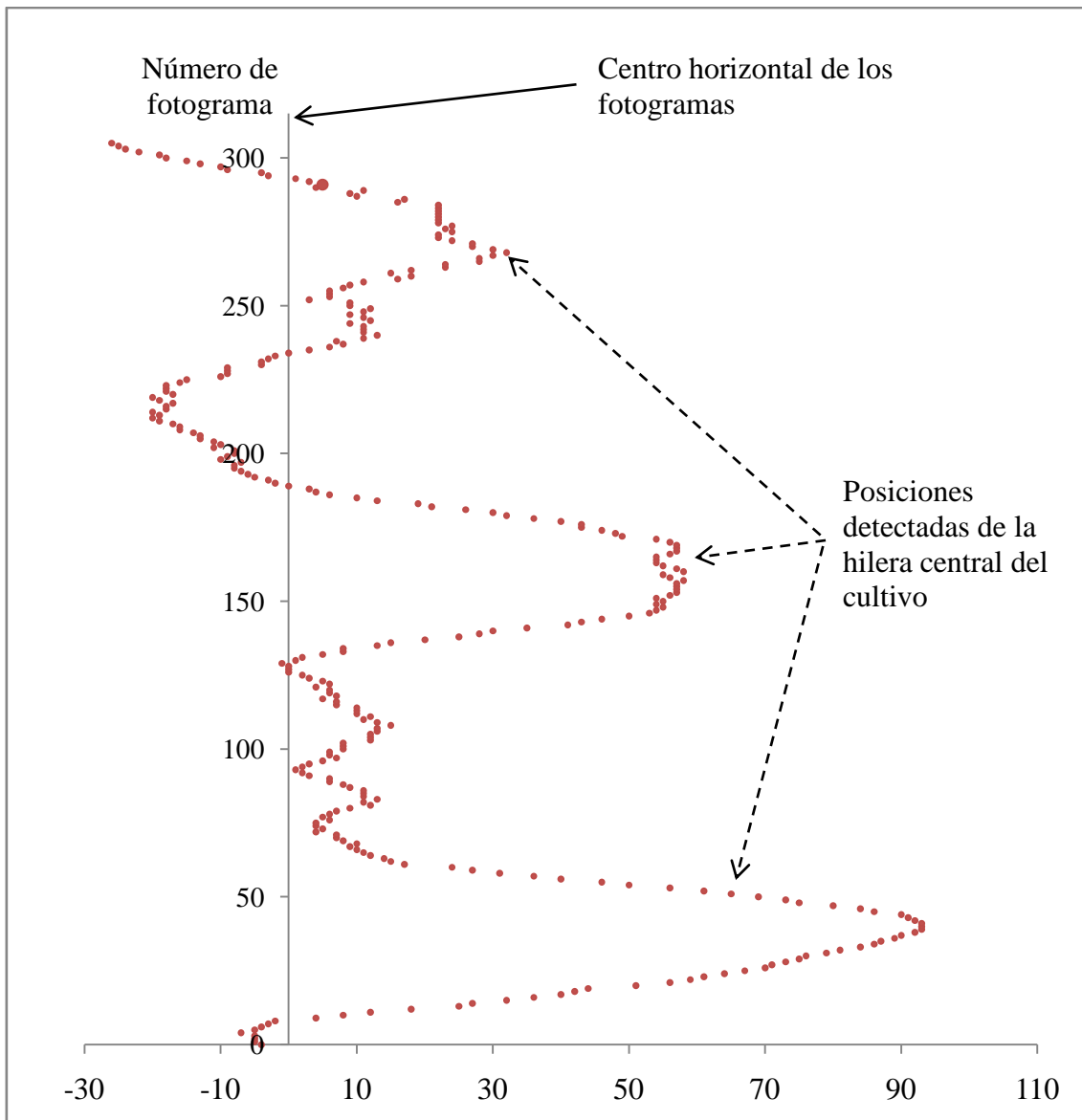


Figura 15. Gráfica de la distancia entre el centro horizontal de los fotogramas y la hilera central del cultivo a una altura de imagen de 215 píxeles (desde la parte superior) para una secuencia de 306 fotogramas.

3.3. Mapas generados

Una vez ya se ha aplicado el proceso de estabilización y eliminación de perspectiva el algoritmo genera la matriz de mapa del terreno y la convierte a una imagen en escala de grises equivalente. Esta imagen muestra, como se ha comentado en la sección 2.7, el nivel de certeza sobre la existencia de vegetación (cultivos o malas hierbas) en las diferentes zonas contenidas en el mapa. En la Figura 17 se muestra una comparación de los mapas generados para una secuencia de 306 fotogramas antes y después de la estabilización. Esta secuencia corresponde a una situación extrema con un día parcialmente cubierto y un terreno muy irregular, de ahí las grandes oscilaciones antes de la estabilización (Figura 17a, página 28).

Una vez se aplica la estabilización propuesta se obtiene un mapa en el que las hileras de cultivo, paralelas y rectas en el campo, permanecen rectas a lo largo de la mayor parte del terreno explorado. En este mapa estabilizado hay una región en la que se produce un error de detección y seguimiento de los cultivos, tal y como se comentó en la sección 2.5. En dicha zona el algoritmo encuentra un rodal de mala hierba que produce una zona suficientemente cubierta por vegetación como para que, por posición y cantidad de verde detectado, sea confundido con la propia hilera de cultivo izquierda y utilizado en el cálculo de la matriz de homografía.

Aunque el rodal desaparece rápidamente, el sistema de limitar la búsqueda de puntos de interés alrededor de la última posición conocida provoca que el algoritmo no detecte la hilera de cultivo, que se está desplazando hacia la derecha en ese momento, hasta 43 fotogramas más tarde. La posición del punto de interés almacenada no se modifica al no detectarse vegetación en esos fotogramas y el error perdura en ese fragmento. En la Figura 22 del Anexo I se muestra una selección de los fotogramas correspondientes a ese fragmento del video.

En la Figura 16 se muestra un segundo mapa, generado utilizando otra secuencia de video de la colección, grabada, en este caso en un día soleado. La principal diferencia, en cuanto a la iluminación se refiere, es el cambio en la coloración dominante del suelo, que pasa de un tono rojizo a otro más arenoso y amarillo. Sin embargo, pese a dicho cambio de iluminación, el mapa generado muestra que los cultivos han sido detectados y seguidos con total éxito. Viendo el mapa previo a la estabilización (Figura 16a) ya se aprecia que el terreno de este campo de cultivo es más regular y suave que en el caso anterior, pero aún con esto, la versión estabilizada del mapa es significativamente mejor (Figura 16b), de hecho las hileras de cultivo quedan prácticamente rectas. Este campo presenta mucha menor concentración de rodales de mala hierba, lo que posiblemente mejora el funcionamiento del algoritmo.

Al margen del error de seguimiento mencionado, los mapas generados tras el proceso de estabilización muestran claramente y en sus posiciones correctas los cultivos y las malas hierbas presentes en el espacio entre hileras. Gracias a la estabilización las posiciones de las hileras de cultivo responden a la realidad. Se ve que los centros de las hileras de cultivo aparecen en color blanco, lo que indica la máxima certeza de que esa es su posición correcta (se detectan en esas zonas en muchos fotogramas) mientras que los rodales de mala hierba aparecen en distintas tonalidades de gris, en función del número de apariciones en esas posiciones concretas.

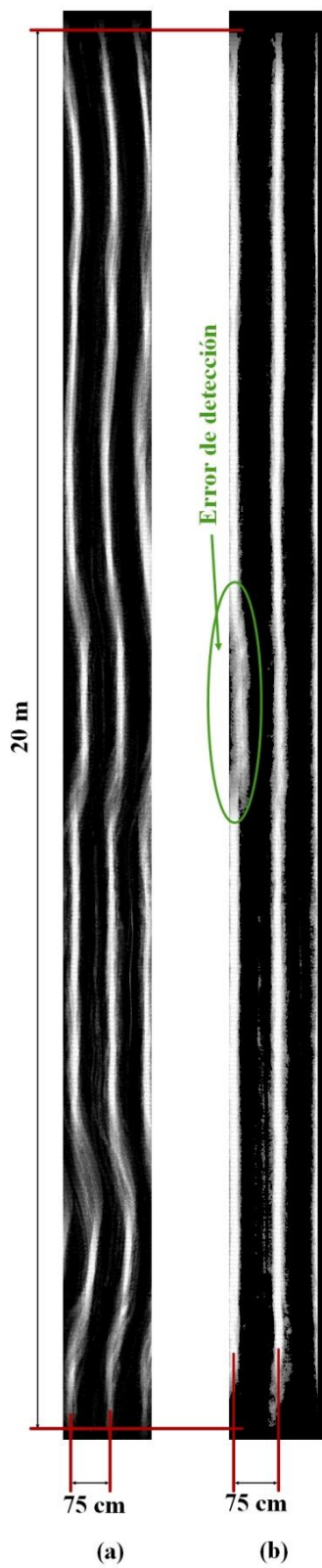


Figura 17. Comparación de los mapas generados antes (a) y después (b) de la estabilización para un video de 306 fotogramas grabado en un día parcialmente cubierto.

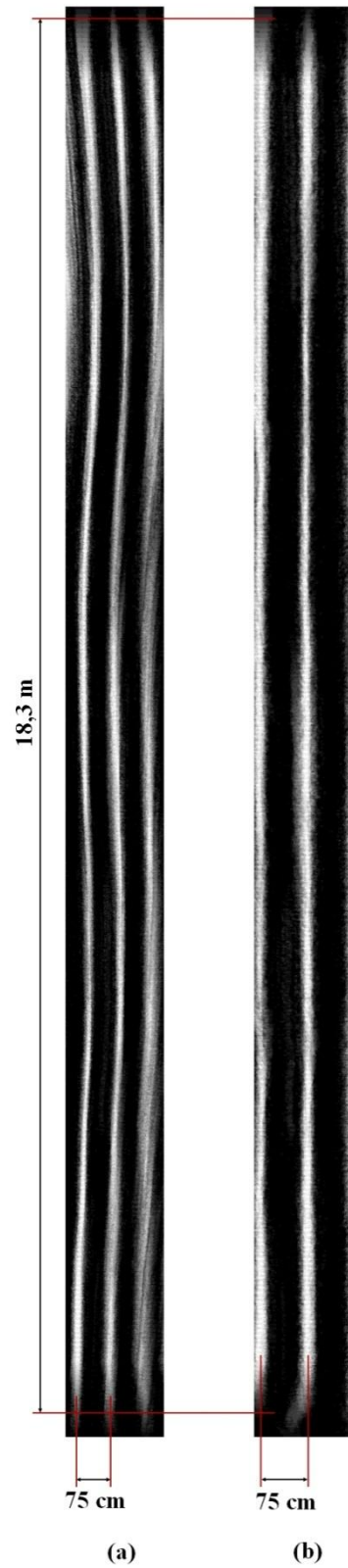
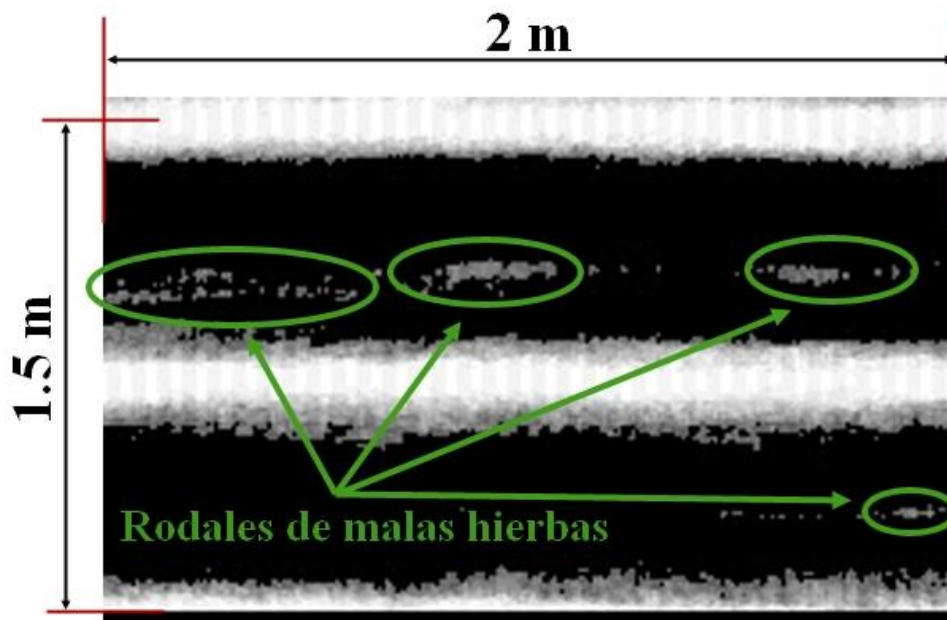


Figura 16. Comparación de mapas generados antes (a) y después (b) del proceso de estabilización para un video de 275 fotogramas grabado en un día soleado.

En la Figura 18 se presenta una ampliación de una zona del mapa de la Figura 17b. Ésta imagen muestra un fragmento del campo explorado de unos dos metros de longitud y se pueden apreciar los rodales de malas hierbas así como las hileras de cultivo, cada uno con sus distintas tonalidades de gris.



4. Conclusiones

La información precisa de los terrenos agrícolas es una parte crucial de la Agricultura de Precisión y los mapas una de las formas más útiles de organización de esta información. Esta información es necesaria para poder estudiar la evolución de los cultivos o la cobertura de malas hierbas y poder tomar decisiones al respecto de los tratamientos a aplicar u otro tipo de actuaciones sobre el terreno. El uso de vehículos autónomos para aplicar estos tratamientos es una de las posibilidades más interesantes a día de hoy, y esta opción requiere conocer con precisión las posiciones de los cultivos y los rodales de mala hierba de forma que se minimicen las cantidades de productos a aplicar. Las herramientas más útiles y económicas para recoger la información del cultivo son las cámaras de vídeo montadas sobre plataformas móviles, que permiten el muestreo continuo del terreno. Al emplear estos dispositivos en movimiento sobre el campo es inevitable un cierto nivel de ruido e inestabilidad en las secuencias grabadas que debe tenerse en cuenta antes de extraer información útil y precisa por lo que el proceso de estabilización juega un papel fundamental en todo este proceso de extracción de la información.

Las técnicas habituales de estabilización de secuencias de video utilizan elementos constantes que aparecen en todos o casi todos los fotogramas del video y cuyas posiciones se suponen estables para realizar las correcciones necesarias. En campos de cultivo la inclinación de las cámaras hacia el suelo (que elimina la línea del horizonte en la mayoría de fotogramas) y la ausencia de puntos fijos impide aplicar los métodos tradicionales de estabilización y convierte el problema en no trivial.

En este trabajo se ha presentado un método de estabilización de secuencias grabadas en terrenos agrícolas basado en la detección y el seguimiento de las líneas de cultivo en todos los fotogramas. Las hileras de cultivo se detectan y son seguidas por el algoritmo de forma correcta en una gran mayoría de los fotogramas tratados, y se eliminan las oscilaciones laterales y los giros de la cámara manteniendo una región de interés centrada en la imagen.

El algoritmo se ha probado en diferentes secuencias de video, grabadas en campos distintos bajo condiciones de iluminación natural variables y en operaciones de tratamiento real de los cultivos, lo que aumenta considerablemente las vibraciones generadas por el tractor al llevar acoplado un sistema de distribución de herbicida. Se han medido las distancias existentes en las imágenes entre ciertas características de referencia y se ha comprobado que éstas permanecen constantes una vez aplicado el proceso de estabilización. El método es capaz de corregir desviaciones laterales de los cultivos de su posición centrada de hasta un 66 % del espacio entre hileras de plantas.

El movimiento de la cámara sobre el tractor genera también movimientos no deseados en la dirección longitudinal (cabeceos de la cámara). Este tipo de movimientos no son detectables en las imágenes con que se ha trabajado debido a la simetría longitudinal de los cultivos y por tanto no han sido considerados en el método propuesto. El efecto de este tipo de inestabilidad en el mapa construido es la generación de rodales de malas hierbas con mayor extensión que la real en la dirección del movimiento del tractor (paralela a los cultivos). Pese a esto, el efecto sobre la imagen no es un problema serio ya que la precisión del mapa es superior a la mínima resolución del sistema GPS que guiaría un tratamiento automático con herbicida y por tanto los rodales, pese a su mayor longitud, seguirían quedando englobados en una celda unidad del mapa de tratamiento (aplicación o no aplicación de herbicida).

En general, los mapas generados son la prueba de la necesidad de un algoritmo de estabilización. Los mapas del terreno sin estabilización previa muestran unas hileras de cultivo serpenteantes que no recogen la geometría real del cultivo. El proceso de estabilización combinado con la eliminación de la perspectiva logra mantener la rectitud y el paralelismo de los surcos con éxito en las secuencias de video utilizada pese a irregularidad del terreno.

Se ha visto que se presentan algunos problemas en aquellas zonas del campo donde la infestación de mala hierba es alta (existen rodales compactos) y/o la densidad del cultivo es menor de lo esperable. En estas regiones el algoritmo tiene problemas para detectar de forma adecuada los cultivos y tiende a clasificar erróneamente los rodales de

malas hierbas como hileras de cultivo, especialmente si dichos rodales ocupan posiciones muy cercanas a donde se espera encontrar el cultivo. Otra situación que provoca errores es una alta cobertura de infestación en el espacio entre hileras de cultivo. Si esta situación se da, el algoritmo es incapaz de determinar la posición real del cultivo al estar basado en las diferencias de color entre las plantas y el resto. Sin embargo, esta situación es muy poco frecuente en la época en que se deben aplicar los tratamientos.

Se ha mostrado que el algoritmo es robusto frente a errores en la siembra (ausencia de plantas en áreas restringidas) y éstos tienen que ser muy significativos y venir acompañados de fuertes oscilaciones laterales de la cámara para que lleguen a afectar al rendimiento global del algoritmo.

Pese a los resultados obtenidos, hay una serie de mejoras previstas para el futuro. En primer lugar se pretende mejorar la robustez del algoritmo y su respuesta a errores de siembra. Se puede implementar un método de *memoria* para retener los cálculos hechos en los fotogramas fiables y utilizar esta información para mejorar la respuesta en caso de incertidumbre. Esto ya se hace de una forma muy sencilla al mantener los puntos del vector de centros de interés sin modificar siempre que no se detecta la línea de cultivo adecuada. Sin embargo, este mecanismo ha mostrado una respuesta inadecuada cuando la falsa detección de una línea de cultivo se produce junto con una fuerte oscilación de la cámara que conduce a un cambio significativo en la posición del cultivo. La ventana de búsqueda de nuevos centros del cultivo limita la búsqueda a una zona donde simplemente no hay más que suelo y por tanto la detección falla. Una forma, por ejemplo, de solventar este problema sería adaptar la búsqueda al desplazamiento medio de los cultivos.

Otra modificación interesante es la instalación de un dispositivo mecánico reductor de las vibraciones en el sistema de montaje de la cámara sobre el vehículo. Esto podría ayudar en la mejora de la calidad general de las imágenes y como consecuencia tendría un efecto directo sobre la calidad de la segmentación de la capa vegetal y en la precisión del mapa.

5. Agradecimientos

El Ministerio de Ciencia e Innovación y la Unión Europea han proporcionado un apoyo y financiación constantes durante todo el desarrollo de esta investigación a través de los proyectos PLAN NACIONAL AGL2008-04670- C03-02/AGR y el proyecto europeo RHEA (Robot Fleets for Highly Effective Agriculture and Forestry Management). Éste último, proyecto número NMP2-LA-2010-245986 del 7º Programa Marco.

Además me gustaría agradecer aquí a mis tutores Ángela Ribeiro y Gonzalo Pajares todo su apoyo y ayuda para la elaboración de este trabajo de investigación. A veces es necesario alguien que te reorienta cuando los pensamientos te llevan en la dirección errónea y que comparta la alegría cuando las cosas salen bien.

6. Referencias

1. Chartuni, E; de Assis de Carvalho, F.; Marçal, D.; Ruz, E. Nuevas herramientas para mejorar la gestión tecnológica en la empresa agropecuaria. *Revista Comunica*, 1, II Época. Instituto Interamericano de Cooperación para la Agricultura (IICA). **2007**, 24 – 31.
2. Marshall, E.J.P. Field-scale estimates of grass weed populations in arable land, *Weed Research*, 28 (3). **1988**. 191-198.
3. Johnson, G. A.; Mortensen, D. A.; Martin, A. R.. A simulation of herbicide use based on weed spatial distribution. *Weed Research*, 35 (3). **1995**. 197-205.
4. Tian, L.; Reid, J. F.; Hummel, J. W. Development of a Precision Sprayer for Site-Specific Weed Management. *Transaction of the American Society of Agricultural Engineers*, 42. **1999**. 893-900.
5. Medlin C. R.; Shaw, D. R. Economic comparison of broadcast and site-specific herbicide applications in nontransgenic and glyphosate-tolerant Glycine max. *Weed Science* 48 (5). **2000**. 653 – 661.
6. Timmermann, C.; Gerhards, R.; Kühbauch, W. The Economic Impact of Site-Specific Weed Control. *Precision Agriculture*, 4 (3). **2003**. 249-260.
7. Menges, R.M.; Nixon, P.R.; Richardson, A.J. Light reflectance and remote sensing of weeds in agronomic and horticultural crops. *Weed Science*, 33. **1985**. 569 – 581.
8. Richardson, A.J.; Menges, R.M.; Nixon, P.R. Distinguishing weed from crop plants using video remote sensing. *Photogrammetric Engineering and Remote Sensing*, 51. **1985**. 1785 – 1790.
9. López Granados, F.; Jurado-Expósito, M.; Atenciano Núez, S.; García-Ferrer, A.; Sánchez dela Orden, M.; G.-T. L. Spatial variability of agricultural soils in southern Spain, *Plant and Soil* 246, **2002**. 97–105
10. Rew L.J.; Cousens R. D. Spatial distribution of weeds in arable crops: are current sampling and analytical methods appropriate? *Weed Research* **2001**, 41, 1-18.
11. Colliver, C.T.; Maxwell B.D.; Tyler D.A.; Roberts D.W.; Long D.S. Georeferencing wild oat (*Avena fatua*) infestations in small grains (wheat and barley): accuracy and efficiency of three weed survey techniques. *Proceedings 3rd International Conference on Precision Agriculture* (eds P.C Roberts, R.H Rust and W.E Larson), Minneapolis, USA, **1996**. 453 – 463.
12. Murphy, D.P.L; Oestergaard, H; Schnug, E. Lokales Ressourcen Management- Ergebnisse und Ausblick (Local resources management – results and outlook). ATB/KTBL- Kolloquium Technik für Kleinräumige Bewirtschaftung (Potsdam-Borhim, Germany). **1994**. 90 – 101.
13. Schwartz, J.; Wartenberg, G. Wirtschaftlichkeit der teilflächenspezifischen Herbizidanwendung (Economic benefit of site-specific weed control). *Landtechnik* 54. **1999**. 334 – 335.
14. Billingsley, J., Schoenfisch, M. The successful development of a vision guidance system for agriculture. *Computers and Electronics in Agriculture*. **1997**. 16, 147–163.

15. Censi, A.; Fusiello, A.; Roberto, V. Image Stabilization by Features Tracking. *International Conference on Image Analysis and Processing* **1999**, 665-670.
16. Ertürk, S. Real-Time digital image stabilization using Kalman filters. *Real-Time Imaging*, **2002**, 8, 317-328.
17. Morimoto, C.; Chellappa, R. Fast Electronic Digital Image Stabilization for Off-Road Navigation. *Real-Time Imaging*, **1996**, 2, 285-296.
18. Hansen, P.; Anandan, P.; Dana, K.; VanDer Wal, G.; Burt, P. Real-time scene stabilization and mosaic construction. *Proceedings of the Second IEEE Workshop on Applications of Computer Vision*, **1994**, 54-62.
19. Irani, M.; Rousso, B.; Peleg, S. Recovery of ego-motion using image stabilization. *Proceedings of the 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. **1994**, 454-460.
20. Viéville, T.; Clergue, E.; Dos Santos Facao, P.E. Computation of ego motion using the vertical cue. *Machine Vision and Applications*, **1995**, 8, 41-56.
21. Vella, F.; Castorina, A.; Mancuso, M.; Messina, G. Digital image stabilization by adaptive Block Motion Vectors filtering. *IEEE Transactions on Consumer Electronics*, **2002**, 48 (3), 796-801.
22. Bradsky, G.; Kaehler, A. *Learning OpenCV*. 1st ed. O'Reilly Media Inc. 1005 Gravenstein Highway North, Sebastopol, CA. USA. 2008.
23. Gerhards, R.; Oebel, H. Practical experiences with a system for site-specific weed control in arable crops using real-time image analysis and GPS-controlled patch spraying. *Weed Research* 46, **2006**. 185-193.
24. Slaughter, D.C.; Giles, D.K.; Downey, D. Autonomous robotic weed control systems: a review. *Computers and Electronics in Agriculture*. **2008**, 61, 63-78.
25. Woebbecke, D.; Meyer, G.; Vonbargen, K.; Mortensen, D. Colour indices for weed identification under various soil, residue and lighting conditions. *Transactions of the American Society of Agricultural Engineers*, **1995**, 38 (1), 271-281.
26. Ribeiro, A.; Fernández-Quintanilla, C.; Barroso, J.; García-Alegre, M.C. Development of an image analysis system for estimation of weed. *Proceedings of the Fifth European Conference on Precision Agriculture*, **2005**, 169-174.
27. Burgos-Artizzu, X.P.; Ribeiro, A.; Tellaeche, A.; Pajares, G.; Fernandez-Quintanilla, C. Analysis of natural images processing for the extraction of agricultural elements. *Image and Vision Computing*, **2010**, 28, 138-149.
28. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems Man and Cybernetics*, 9 (1). **1979**. 62 - 66.
29. Burgos-Artizzu, X.P.; Ribeiro, A.; Guijarro, M.; Pajares, G. Real-time image processing for crop/weed discrimination in maize fields. *Computers and Electronics in Agriculture*. **2011**. 75. 337-346.
30. Sonka, M.; Glavac V.; Boyle, R. *Image processing, analysis, and machine vision*. 3rd Ed. Thomson Learning. 1120 Birchmount Road, Toronto, Ontario, Canada, 2008, 553-565.
31. Barnard, S.T. Interpreting perspective images. *Artificial Intelligence* **1983**, 21, 435-462.

32. Mallot, H. A.; Bühlhoff, H. H.; Little, J. J.; Bohrer, S. Inverse perspective mapping simplifies optical flow computation and obstacle detection. *Biological Cybernetics*, **1991**, 64, 177-185.
33. Bevilacqua, A.; Gherardi, A.; Carozza, L. Automatic perspective camera calibration based on an incomplete set of chessboard markers. *Sixth Indian conference on computer vision, graphics & image processing, ICVGIP '08*. **2008**. 126-133.

Anexo I: Secuencias de fotogramas

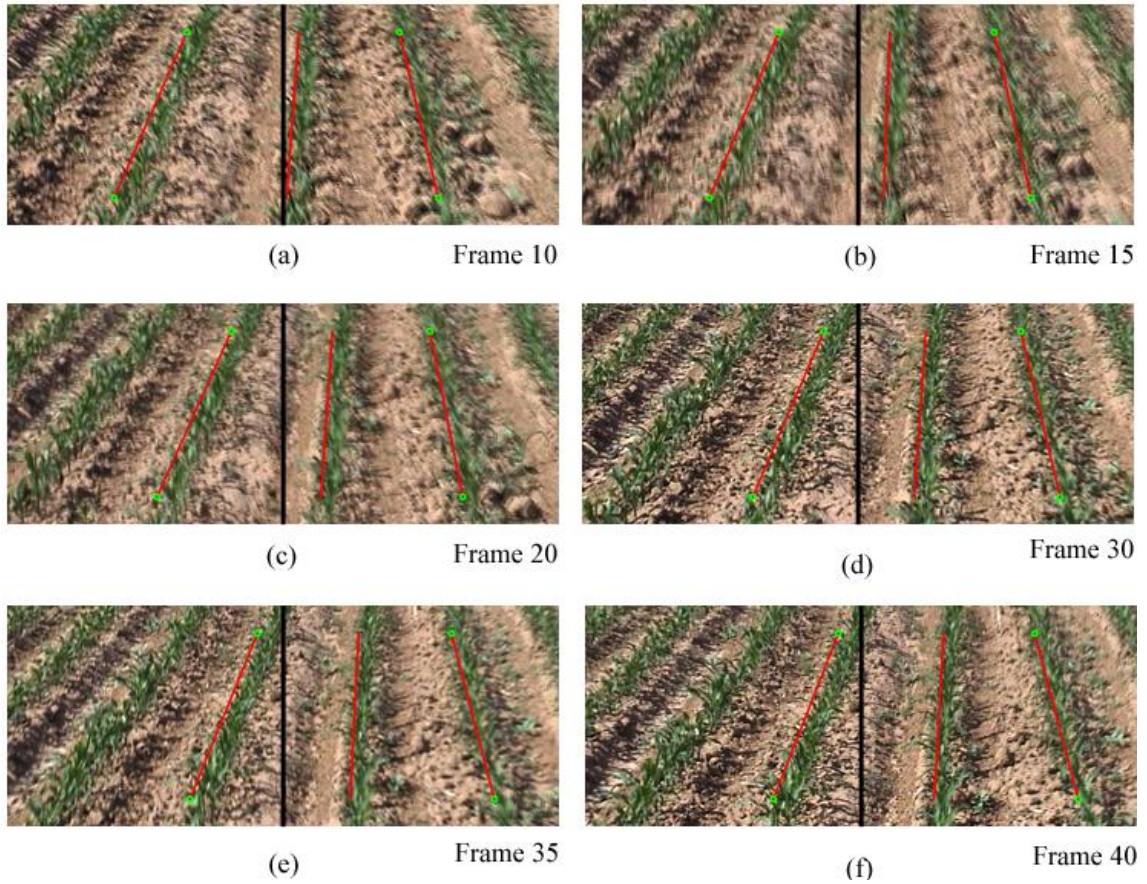


Figura 19. Secuencia de fotogramas pertenecientes a una grabación en un campo en día parcialmente cubierto. La irregularidad del terreno provoca significativas oscilaciones laterales en las posiciones de las hileras de cultivo que requieren estabilización.

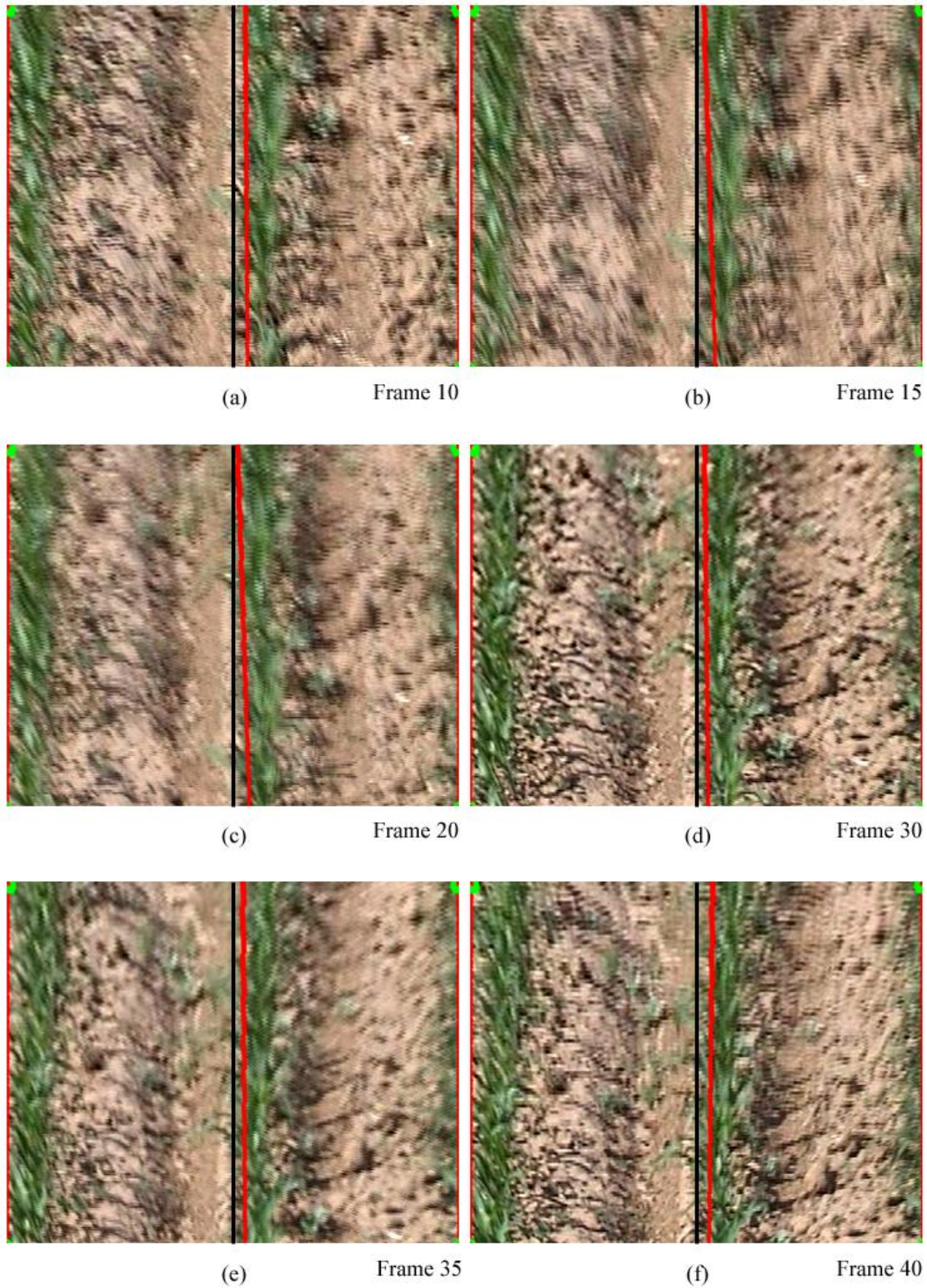
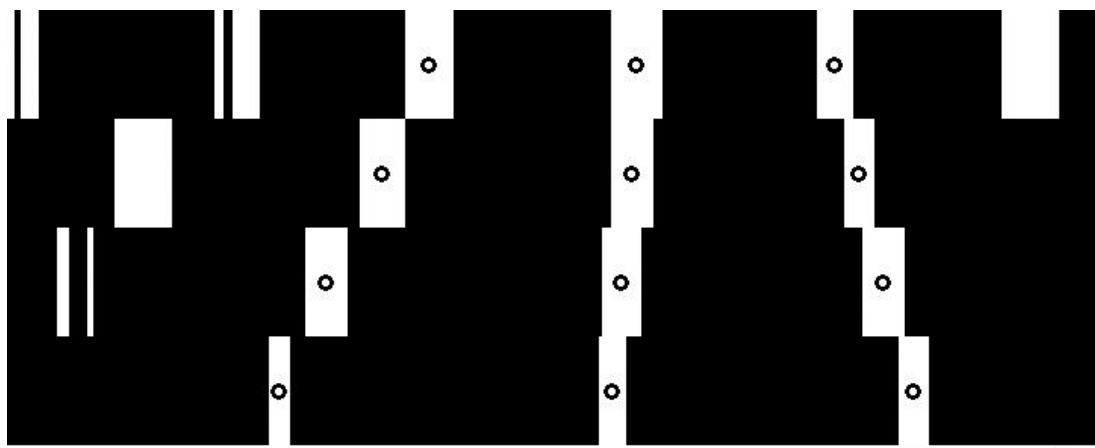
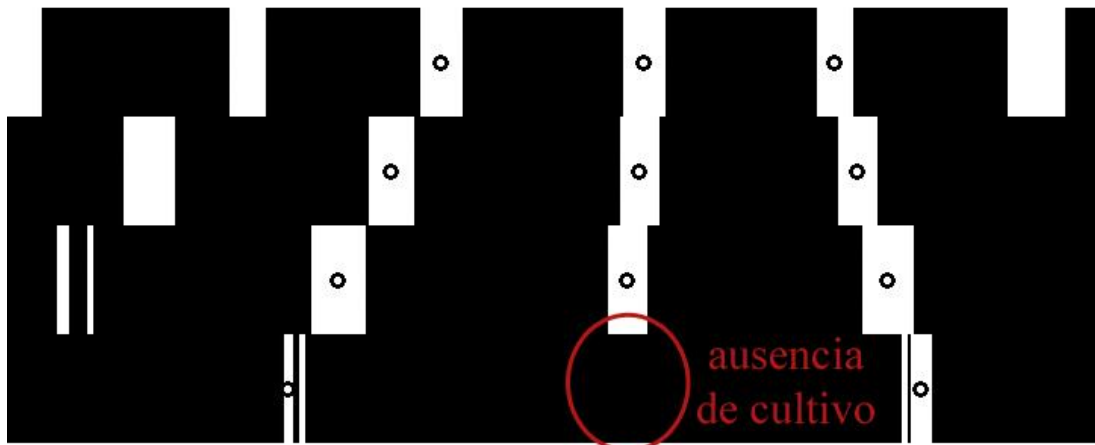


Figura 20. Secuencia de fotogramas en vista cenital y estabilizados correspondientes a los mostrados en la Figura 19. Ahora las hileras de cultivo permanecen en posiciones fijas y estables en todos los fotogramas.



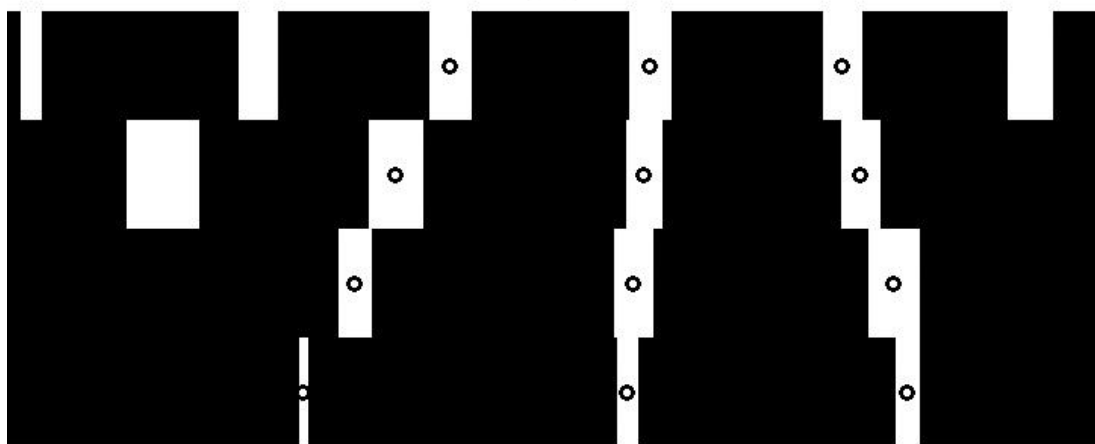
(a)

Fotograma 18



(b)

Fotograma 19



(c)

Fotograma 20

Figura 21. Secuencia de fotogramas en la que se ve como el algoritmo es capaz de detectar el cultivo pese a que éste desaparece en uno de los fotogramas. (a) Fotograma 18 con el cultivo presente y detectado. (b) Fotograma en el que el cultivo es inexistente en la parte inferior de la línea central. El algoritmo no puede por tanto detectarlo. (c) Fotograma en el que el cultivo reaparece y es detectado de nuevo por el algoritmo.

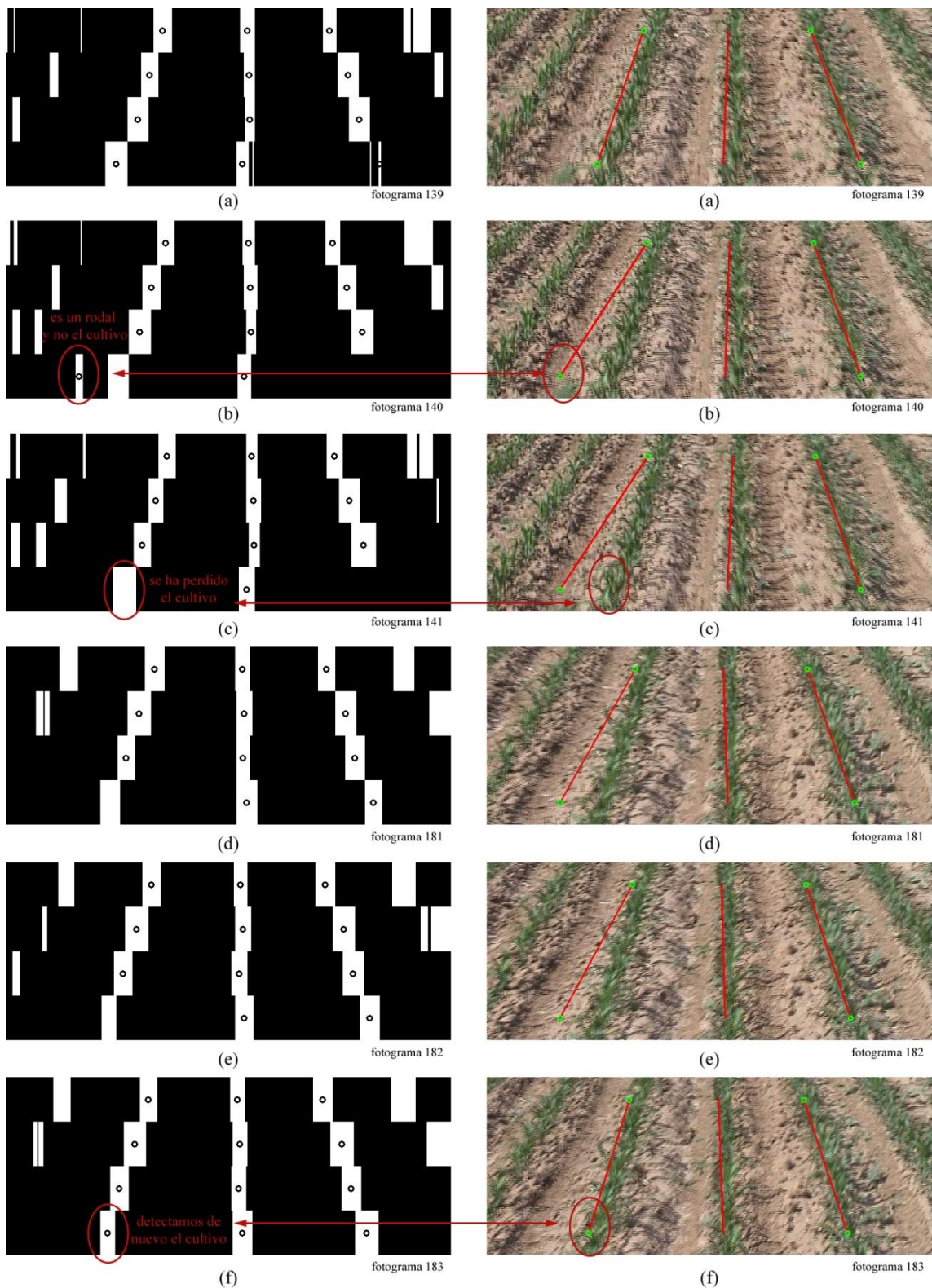


Figura 22. Secuencia de fotogramas en la que se aprecia un error en la detección de los cultivos debido a la presencia de un rodal de malas hierbas cerca de la posición donde debería estar el cultivo. El sistema falla en la detección en el fotograma 140 y se mantiene dicho error durante varios fotogramas, hasta que en el número 183 se recupera la normalidad al detectarse de nuevo el cultivo.