

Visión artificial para detección automática de altura del cultivo

Fatima Martinez, José María Manzano, James Romaine, Pablo Millan
 Universidad Loyola Andalucía. {fbelen,jmanzano,jbromaine,pmillan}@uloyola.es

Resumen

En este trabajo se presenta la combinación de visión estéreo y un algoritmo basado en kernels para la identificación automática de la altura de un cultivo. La técnica desarrollada alivia la necesidad de realizar controles periódicos del cultivo en persona, reduciendo costes tanto por disminuir el tiempo del personal como por identificar rápidamente retrasos en el crecimiento del cultivo. La propuesta utiliza operaciones morfológicas y la transformada de Hough para el trazo de las rectas que delimitan la parte superior e inferior de la banda de la planta. El algoritmo es de bajo coste computacional, permitiendo su aplicación en dispositivos IoT con microprocesadores empotrados. Los resultados se aplican sobre imágenes reales, obteniendo un error en la estimación de altura del 3%.

Palabras clave: visión artificial, visión estereoscópica, kernels, agricultura inteligente

1. Introducción

1.1. Motivación

El estrés vegetal, causado por factores bióticos o abióticos, afecta negativamente al crecimiento de las plantas y reduce significativamente la productividad. En los huertos, la eficacia de las medidas correctoras depende de la detección e identificación a tiempo de la causa del estrés [13].

La detección temprana del estrés de las plantas es fundamental para minimizar la pérdida de productividad, tanto aguda como crónica. La gravedad de los daños depende de la duración entre el inicio y el momento de la corrección [6]. Una magnitud para la identificación del estrés es la medida de crecimiento del cultivo.

La variedad de formas, tamaños y colores de los cultivos agrícolas propicia la aplicación de sistemas de visión artificial como medio para sustituir la inspección visual realizada por humanos, optimizando el tiempo empleado por los agricultores [4]. La tecnología clave es el procesamiento eficiente de imágenes [16] y los algoritmos de

visión computacional, que permiten a los ordenadores procesar el contenido de una imagen, obteniendo características que son importantes para un objetivo determinado [10].

Para un agricultor, la calidad de su producto viene definida por su potencial de rendimiento y apetencia del mercado a la hora de agilizar la entrega de la cosecha. Para conocer el estado de la cosecha, los agricultores acuden al campo con una periodicidad alta. Si se contara con algoritmos capaces de estimar los parámetros requeridos de forma automática, en el terreno, se reduciría la necesidad de supervisión por parte de los productores.

Uno de estos parámetros de interés es la altura del cultivo. Conociendo sus distintos valores a lo largo del desarrollo de la cosecha, se puede estimar la productividad final. Además, si la altura no se desarrolla de la manera esperada es un indicio de la presencia de posibles factores dañinos, como plagas, falta de minerales o sequía, entre otros.

1.2. Estado del arte

Los algoritmos para detección automática de la altura del cultivo se han venido desarrollando en las últimas décadas. Inicialmente implicaban un alto coste computacional y de instrumentación además de una alta dependencia del entorno [12]. Un ejemplo de estos métodos es la aplicación del radar de apertura sintética (SAR) para medir la altura de los árboles mediante modelos interferométricos.

Otro método desarrollado anteriormente para la medición de la altura ha sido el sistema monocular basado en la visión artificial, que utiliza una sola cámara. En [2], se emplea el sistema monocular y algoritmos de visión artificial a bordo de un vehículo aéreo no tripulado para capturar imágenes de la vista superior de los árboles. En este trabajo, la información de profundidad se reconstruye mediante programación dinámica, con un algoritmo de modelado de oclusión por medio de un sistema LiDAR. La desventaja es que no se detectan todos los árboles y la técnica presenta un error registrado que oscila entre 1,1 y 1,8 m, estando el emisor láser a un rango de distancia con respecto al nivel del suelo de 230 a 280 metros.

Otros autores han abundado en el empleo del sistema LiDAR para la estimación de la altura [15], sin lograr mejoras muy notables en la detección.

Otra opción disponible para resolver este problema es la visión estéreo, que consiste en la utilización de dos sensores en paralelo, donde cada sensor toma una imagen al mismo tiempo para ser procesadas [9]. Esta técnica tiene como ventaja principal proporcionar la profundidad del objeto de estudio en relación a la posición de la cámara considerada como origen.

Inicialmente los estudios hechos con visión estéreo se realizaban generalmente para el análisis fenotípico, con cultivos individuales de pequeño tamaño [8]. Estos enfoques tienen limitaciones cuando se aplican automáticamente en cultivos reales, ya que son desarrollados en situaciones específicas, en un escenario ideal, que no involucran factores externos presentes en un montaje en campo.

En [7] se logró realizar la detección de la altura en cultivos por medio de la visión estéreo, a través del mapa de disparidad en distintos cultivos, donde se toma como objeto de interés, aquello que se encuentra a una distancia más corta del sistema óptico, descartando todo objeto que se pueda encontrar detrás de la planta.

La aplicación de equipos móviles empezó a ser más empleada en el ámbito de visión por ordenador. En [5] se acopla el sistema óptico a un tractor, para medir la altura de las plantas de algodón. Los resultados mostraron que el sistema podía medir la altura en condiciones de campo del mundo real. Se desarrolló un algoritmo para medir las alturas máximas y medias de las parcelas individuales de algodón, utilizando las imágenes de profundidad adquiridas por el sistema.

Estas técnicas se extendieron utilizando vehículos aéreos no tripulados. En [14] se realiza el cálculo de las características geométricas tridimensionales de los árboles individuales y de las filas de árboles. Estos se procesan para extraer las características de los cultivos por medio de la generación de modelos digitales de superficie con tecnología de vehículos aéreos no tripulados, y el uso de técnicas de análisis de imágenes basadas en objetos.

1.3. Contribución

En este artículo se propone el análisis de imágenes estereoscópicas, combinándose con un algoritmo basado en kernels para la determinación de altura de los cultivos. Una ventaja del método es la toma y procesamiento de la información de forma local, desde el mismo nodo situado en la cosecha, cosa que es posible debido al bajo costo computacional del algoritmo desarrollado. El método es aplicable

para la detección de altura en cultivos que crecen en bandas, donde la primera fila es representativa de toda la banda. Se referirá como *banda* al conjunto de plantas del mismo tipo cuyo límite se encuentre bien definido.

La detección se realiza con la utilización de un algoritmo basado en kernels para el trazo de la altura superior e inferior de la planta en píxeles, que puede traducirse en unidades de medida longitudinales a partir de la profundidad obtenida con la visión estereoscópica.

La técnica desarrollada emplea la transformada de Hough que, si bien se aplica en algunos algoritmos de visión por computador, es más habitual en sistemas de detección de rutas y no se han encontrado estudios que la empleen para la detección de alturas de cultivos.

El resto de este artículo está estructurado como sigue: en la Sección 2 se explica el método propuesto para detectar automáticamente la altura del cultivo. En la Sección 3 se muestra la aplicación del método a un caso real. El trabajo concluye con una breve discusión de los resultados en la Sección 4.

2. Método

En esta sección se detalla el algoritmo propuesto para la detección automática de la altura de un cultivo. Este algoritmo es capaz de extraer parámetros en una manera eficiente, pudiendo ser integrado en sistemas de bajo consumo, como microcontroladores ESP32 o dispositivos IoT.

La técnica propuesta combina dos procesos principales. El primer proceso conlleva el empleo de algoritmos basados en kernels para detección de los puntos de interés: aquellos que indican el principio y fin de la banda de plantas en una vista de alzado. El segundo proceso implica el uso de la visión estéreo, para la reconstrucción espacial de la escena. Fusionando ambos, se podrá obtener la altura de la banda de interés.

La estructura del método se representa en la Figura 1: los píxeles de la imagen capturada vienen denotados por los índices $i = 1, \dots, n$ y $j = 1, \dots, m$. Las coordenadas espaciales son X, Y, Z , medidas en metros. Como origen se toma el centro de la cámara izquierda.

2.1. Algoritmo basado en kernels

El objetivo de este primer proceso es proponer un algoritmo capaz de localizar el inicio y el fin de la planta en el eje vertical.

A grandes rasgos, el proceso propuesto en este artículo consistirá en identificar el tono de color

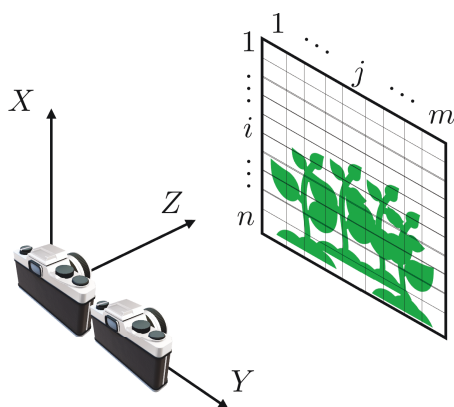


Figura 1: Esquema de la estructura considerada

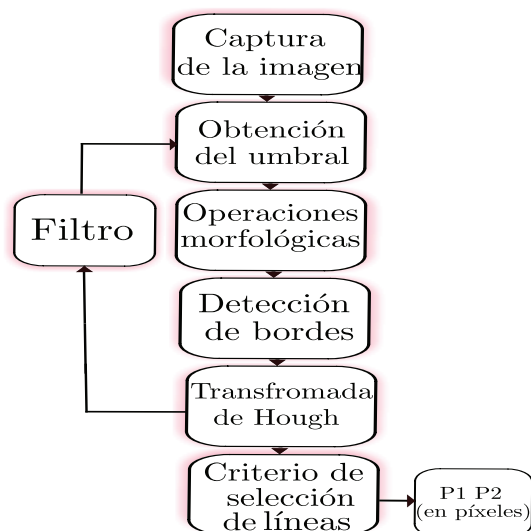


Figura 2: Algoritmo basado en kernels

predominante en la imagen correspondiente a la planta, para posteriormente aplicar operaciones morfológicas y transformaciones, con el fin de obtener las ecuaciones de las rectas que delimitan la banda, en términos de coordenadas en píxeles.

En la Figura 2 se representa el proceso del algoritmo basado en kernels, para lograr la detección de los píxeles de interés. Los píxeles de interés son los que delimitan la altura de la banda de la planta.

La primera tarea consiste en la captura de la imagen. Después, se realizarán dos procesos similares, para identificar por un lado la línea superior, y por otro la línea inferior. Ambos procesos se ejecutarán sobre la imagen original por separado.

En ambos casos, para lograr identificar los píxeles deseados, se procede a la obtención del rango de colores predominantes en la imagen, correspondientes a los tonos de verde presentes en ella.

Este rango se define mediante un umbral, para lo que se requiere el modelo matiz-saturación-valor

(HSV, por sus siglas en inglés) y el histograma de la imagen a procesar. Esto se hace con el fin de identificar los tonos correspondientes a la banda predominante en la imagen, para aplicar una máscara y aislar la zona de interés.

Una vez identificada el área de interés, se aplican operaciones morfológicas. En una operación morfológica, cada píxel de la imagen se ajusta en función del valor de otros píxeles de su entorno [11], mediante la operación de un kernel.

Las operaciones morfológicas consideradas son la erosión y la dilatación. En la erosión, el píxel i -ésimo de la imagen resultante toma el valor mínimo de entre aquellos que se encuentren en la región alrededor del píxel i -ésimo en la imagen original, donde la *región alrededor* viene definida por el tamaño del kernel. La dilatación es la operación opuesta: el píxel resultante toma el valor del píxel con mayor valor de luminosidad de entre los abarcados por el kernel.

Estas operaciones morfológicas se emplean para resaltar los cambios de tonos de verde, y luego pasan como entrada de un detector de bordes, para que, por medio de la transformada de Hough, sean detectadas las rectas presentes en la imagen [1]. Tras esto, se debe aplicar un criterio de selección de las líneas tomadas como correctas.

Por otro lado, la transformada de Hough es una técnica para la detección de figuras en imágenes digitales. Con esta transformada estándar es posible encontrar todo tipo de figuras que puedan ser expresadas matemáticamente. En este caso busca las líneas utilizando la ecuación de una recta. Para las bandas, las líneas determinadas son horizontales, ya que los cambios de tonos de verdes se dan horizontalmente en las bandas de las plantas.

Este proceso es iterativo, de forma que en el caso de no identificar ninguna línea, se aplica un filtro consistente en modificar el brillo de las imágenes. Para ello, se aumentan los valores de saturación (S) y matiz (H) del modelo HSV, multiplicándolos por un factor; siempre saturando en 255 los valores que resultasen superiores a 255. Entonces se repite el proceso descrito, como se indica en la Figura 2.

La transformada de Hough busca encontrar puntos alineados en la imagen, es decir, aquellos que satisfagan la ecuación de la recta. Puede detectar una forma determinada, incluso si están rotas o un poco distorsionadas.

La recta en coordenadas x, y se define mediante su pendiente m , y el valor de la ordenada del punto que corta al eje vertical en el plano b . Al aplicar un cambio de coordenadas de una recta representada en el sistema x, y al espacio de Hough, el cual es

un sistema de ejes ortogonales en un plano m, b , se tendrá un punto (m, b) en dicho espacio.

En la Figura 3 se observa que partiendo de unos puntos identificados en las coordenadas x, y , y llevándolos al espacio de Hough, se obtiene una recta para cada punto (x, y) . Dentro de las rectas trazadas en el espacio de Hough, se busca el punto de intersección entre estas rectas, (m_o, b_o) . Esta intersección no necesariamente se da de manera directa, es decir, no siempre todas las rectas se interceptan en el mismo píxel. Así, se procede a identificar cada píxel por los cuales atraviesan estas rectas y se incrementa el número de votos de estos, logrando un voto cada vez que una de las rectas trazadas atraviese el píxel. Finalmente, el píxel con el número de votos más elevado indica el punto (m_o, b_o) a ser tomado como válido. Con este punto (m_o, b_o) se traza una recta en las coordenadas x, y , la cual será la salida de la transformada de Hough.

Pueden existir singularidades en las rectas, como por ejemplo una pendiente vertical ($m = \infty$). Por tanto, se realiza una segunda transformación del plano de imágenes (coordenadas x, y) a polares, usando el plano θ, ρ . La ecuación aplicada es la de la recta en forma polar:

$$\rho = x \cos(\theta) + y \sin(\theta), \tag{1}$$

donde ρ es la distancia entre el origen y la recta trazada, y θ es el ángulo comprendido entre el eje de abscisa y ρ .

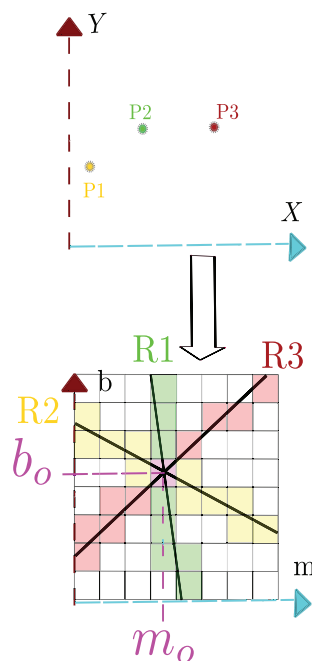


Figura 3: Transformada de Hough

La cantidad de líneas encontradas inicialmente por la transformada de Hough es dependiente de la sa-

lida de la operación morfológica, ya que las operaciones morfológicas resaltan los puntos característicos a ser analizados.

Nótese por tanto que dicha transformada de Hough devuelve no solo la línea superior e inferior que delimitan la banda, sino un conjunto de líneas horizontales posiblemente contenidas en la banda. Resta por tanto realizar un último proceso para identificar cuál de estas es la representativa.

Como criterio de selección para la línea superior que representará la copa de la banda, se denota a las líneas cercanas como un grupo. Un grupo de líneas se define como varias líneas cercanas entre sí y apartadas de otras. El umbral de proximidad seleccionado deberá ser dependiente de la resolución de la imagen capturada.

Por último, se calcula el promedio de cada grupo. Se seleccionan las líneas superior e inferior como aquellas de interés, en cada caso. Como se mostrará en la sección siguiente, las operaciones morfológicas son distintas entre el cálculo de la línea superior e inferior. Por ejemplo, en el caso inferior no es determinante el tono específico de verde, sino que es suficiente distinguir entre verde y el color del suelo.

Identificadas las líneas que delimitan la banda, se denota con los subíndices u a la superior y l a la inferior. Se escoge un píxel de cada una de estas dos rectas, denotados $p_u = (i_u, j_u)$ y $p_l = (i_l, j_l)$, teniendo en cuenta que estén alineados verticalmente; es decir, $j_u = j_l$. Serán empleados en el algoritmo de visión estereo para estimar finalmente la altura de la banda.

2.2. Visión estereoscópica

La visión estereoscópica es el proceso de extracción de información 3D, partiendo de un par de imágenes, mediante la estimación de la profundidad relativa de los puntos de la escena. En la literatura experta suelen encontrarse dos maneras de obtener las coordenadas espaciales de los píxeles: visión punto a punto y visión densa [3].

La visión estereo punto a punto es suficiente para el objetivo perseguido, pues gracias al proceso anterior se conocen los dos píxeles sobre los que se desea estimar sus coordenadas 3D. Sin embargo, gracias a la disponibilidad de algoritmos de visión estereo densa, resulta más sencillo operar sobre los puntos de interés una vez conocidas las coordenadas espaciales de toda la escena.

En la Figura 4 se muestra un esquema que explica el proceso que conlleva la reconstrucción 3D. La primera etapa consiste en conseguir los parámetros de las cámaras, tanto intrínsecos co-

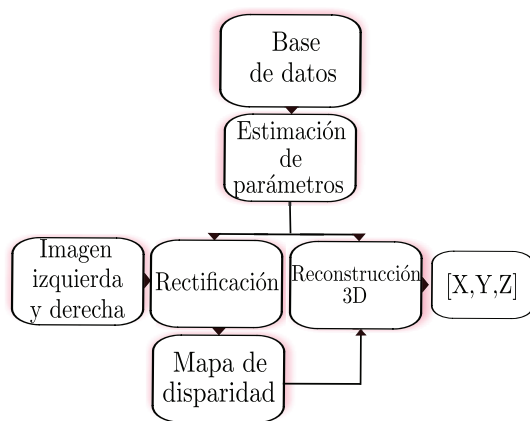


Figura 4: Esquema del proceso de visión estéreo

mo extrínsecos, y coeficientes de distorsión. Estos parámetros pueden o bien conocerse de fábrica, o bien estimarse en base a una serie de fotografías con ciertas características, como pueden ser aquellas que contengan un tablero de ajedrez.

La siguiente etapa es la rectificación, realizada con los parámetros estimados. Además de eliminar distorsiones introducidas por la cámara, tiene el objetivo principal de modificar las dos imágenes, independientemente de la manera en que hayan sido capturadas. Estas modificaciones permiten que las imágenes parecieran capturadas por cámaras perfectamente alineadas, haciendo coincidir un punto cualquiera en la horizontal de ambas imágenes.

La rectificación facilita la búsqueda de la correspondencia entre ambas imágenes, para dar lugar al mapa de disparidad.

El mapa de disparidad es la diferencia medida en píxeles sobre la coordenada horizontal de las dos proyecciones del mismo punto. Proporciona información en píxeles de las distancias reales, aunque agregadas entre las tres coordenadas espaciales.

La etapa final es la reconstrucción 3D, donde se obtienen las coordenadas X, Y, Z , en unidades de longitud, desde el origen a los objetos en las imágenes. Usando los píxeles p_u, p_l obtenidos previamente, se puede calcular la distancia \hat{h} entre ambos píxeles, medida en unidades de longitud, como

$$\hat{h} = \sqrt{(X_u - X_l)^2 + (Y_u - Y_l)^2 + (Z_u - Z_l)^2}, \quad (2)$$

donde los subíndices u y l hacen referencia a las líneas superior e inferior, respectivamente.

3. Resultados

En esta sección se detallan los resultados obtenidos por la metodología planteada previamente, en ensayos reales. Los algoritmos están desarrollados en un entorno de Python con la librería *OpenCV*.



Figura 5: Imagen izquierda rectificada de la banda de la planta

Para tomar las imágenes se utilizaron dos módulos de cámara USB, con chip OV2710, modelo HBV-1716WA. Para la base de datos usada para calibración, se capturaron 94 imágenes con las cámaras izquierda y derecha. En todas ellas aparece un tablero de ajedrez de 10×7 cuadros, donde cada cuadrado tiene un lado de 67 mm. Se obtienen los parámetros necesarios para la visión estéreo con la función `stereoCalibrate` de *OpenCV*.

Posteriormente, se toma una imagen que contiene una banda de plantas, sobre la que se aplicará el algoritmo propuesto. La imagen izquierda se muestra en la Figura 5. La cámara está ubicada a 2,49 m de distancia horizontal respecto a la banda. Esta imagen es la resultante tras haber sido sometida a la rectificación, con la cual se harán todos los procesos posteriores.

3.1. Línea superior

En la Figura 6a se muestra el resultado de haber aplicado una máscara a la imagen a procesar. Para ello se identifica el rango de valores aplicados a la máscara, cuyo matiz H va de 33 a 93, abarcando todos los tonos de verde.

Las operaciones morfológicas aplicadas son dos de erosión, y entre ellas una de dilatación. En la Figura 6b se muestra el resultado de haber aplicado erosión al resultado de la máscara, con un kernel K_1 de tamaño 10×3 píxeles. Tiene una dimensión reducida para detectar tonos bien dispersos.

En la Figura 6c se ve el resultado de haber aplicado dilatación. Para este proceso se utiliza el kernel K_2 , de 25×40 píxeles. Este kernel es de mayor dimensión en consideración al anterior, ya que tiene como objetivo unificar los tonos de verdes.

En la Figura 6d se muestra el resultado de haber aplicado la última operación morfológica, correspondiente a una erosión, con un kernel K_3 . Este kernel tiene el mismo tamaño que K_2 , con lo que se logra disminuir el área de análisis.

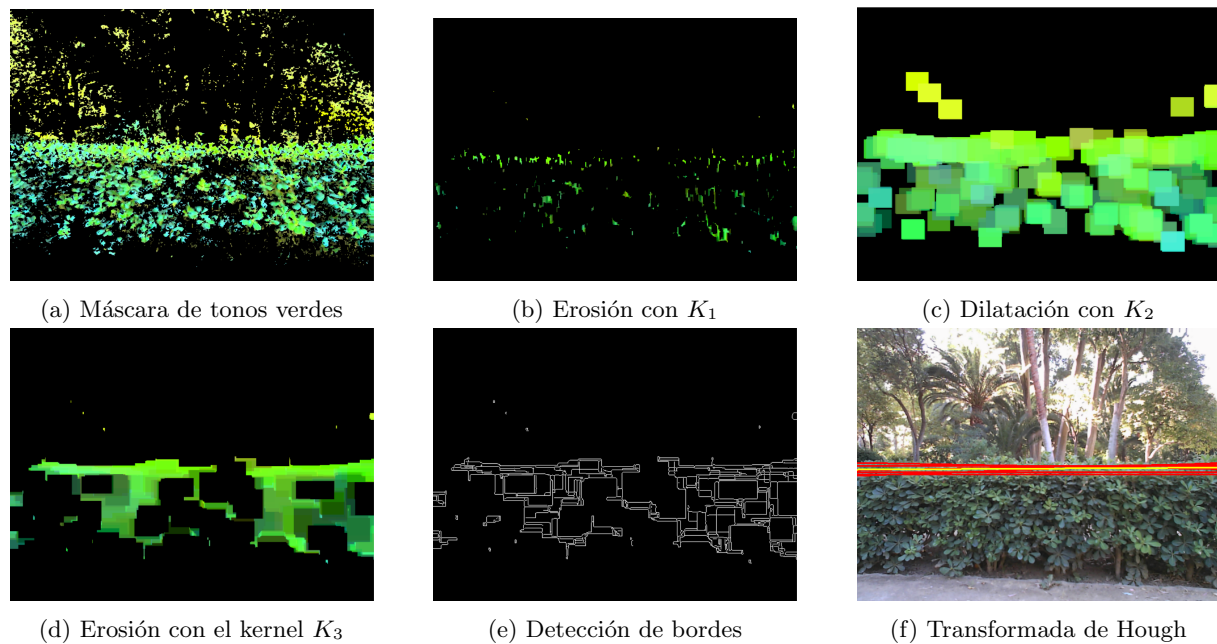


Figura 6: Proceso de identificación de la línea superior

Terminadas las operaciones morfológicas de erosión y dilatación, se aplica la detección de bordes, como se observa en la Figura 6e. Un detalle notable es que los trazos superiores forman líneas casi continuas. Son las rectas que representen a estas líneas las que deben ser identificadas por medio de la transformada de Hough.

En la Figura 6f se observan todas las líneas identificadas por la transformada de Hough en color rojo, a las cuales se les aplica el criterio de selección por cercanía.

En este estudio la imagen tiene una resolución de 640×480 píxeles, y se consideran cercanas a las líneas cuya separación sea menor a 90 píxeles con respecto a la primera línea identificada, formando así un grupo.

La línea amarilla representa el promedio de las líneas identificadas como grupo. Esta línea amarilla es la recta seleccionada para el posterior cálculo de la altura de la banda.

Cabe mencionar que el filtro diseñado para el caso en que la transformada de Hough no detecte ninguna línea ha tenido que ser aplicado cuatro veces. Este ha consistido en aumentar los valores de H y S en un 25 %, saturándolos en 255. Los resultados mostrados en la Figura 6 corresponden a la iteración final.

3.2. Límite inferior

Para la obtención de la línea inferior se procede de forma similar. En la Figura 7a se muestra el resultado de aplicar la máscara a la imagen rectificadas.

En este caso, solo se aplica una operación morfológica de dilatación, con un kernel K_4 de dimensión 12×70 píxeles. El resultado se muestra en la Figura 7b. Este kernel tiene como objetivo incrementar áreas verdes, y realizar un barrido para expandir horizontalmente el área verde sin centrarse en los detalles. Esto se debe a que la necesidad es identificar el color verde, sin tener en cuenta los detalles de su variación de tonos.

En la Figura 7c se ve el resultado de haber aplicado la detección de bordes, para ser analizado por la transformada de Hough.

En la Figura 7d se ven todos los trazos encontrados por la transformada de Hough, el cual es un número elevado de líneas. Esto se debe a que el kernel aplicado en la dilatación es un kernel que no toma como área de interés las zonas con cambios de tonos, sino que toma toda el área verde presente. La última línea trazada es la recta escogida para extraer el píxel de interés, que delimita el suelo de la base de la banda.

En la Figura 8 se observan las dos rectas que representan el inicio y final de la banda analizada.

La altura estimada es $\hat{h} = 0,87$ m, mientras la altura real es $h = 0,9$ m, lo que supone un error del 3.3 %.

3.3. Discusión

En esta sección se realiza un análisis de la influencia del tamaño del kernel en la detección de la altura de la banda.

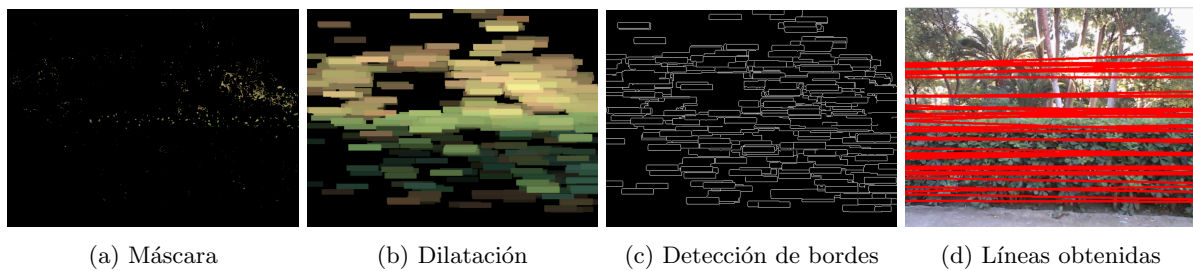


Figura 7: Proceso de identificación de la línea inferior



Figura 8: Rectas que delimitan la altura

Tabla 1: Prueba de diferentes tamaños del kernel K_1 en la identificación de la línea superior

Dimensión (píxeles)	\hat{h} (m)	Error (%)
10×1	0.776	14
8×3	0.848	6
10×3	0.87	3
17×3	0.822	9
20×3	0.795	12
10×4	0.648	28
20×7	Línea no encontrada	

En la Tabla 1 se muestran pruebas realizadas variando el tamaño de K_1 . Se encontraron las siguientes recomendaciones: un kernel de una gran dimensión, aplicado a la operación morfológica de erosión, hace poco probable la coincidencia plena con los tonos verdes de la imagen procesada. Al coincidir con el área de color negro de la imagen, tomará valor cero, por lo que se perderán los tonos de verde, y devolverá una imagen sin áreas de interés. Un kernel con una dimensión superior a 20×7 es considerado de gran dimensión para el caso de estudio, por lo que la transformada de Hough ya no es capaz de identificar ninguna línea.

Por el contrario, con un kernel muy pequeño como es el caso de la dimensión 10×1 , se tomarían demasiados puntos y no se filtraría únicamente el área de interés durante la operación morfológica de erosión.

En la Tabla 2 se muestran pruebas realizadas va-

Tabla 2: Prueba de diferentes tamaños del kernel K_4 en la identificación de la línea inferior

Dimensión (píxeles)	\hat{h} (m)	Error (%)
7×50	0.78	13
7×70	0.87	3
12×70	0.87	3
12×120	0.87	3
15×60	0.65	28
15×70	0.65	28
20×50	0.37	59
20×90	0.69	23

riando el tamaño de K_4 . Se resaltan los siguientes puntos: el kernel debe ser horizontal, ya que si es vertical la transformada de Hough identificará líneas verticales. Además, se estimarán como píxeles verdes a aquellos que no corresponden a la planta, lo que inducirá cálculos erróneos. El kernel debe tener un gran tamaño horizontal, ya que se desea hacer un barrido, para identificar el límite entre el inicio del área verde y el suelo. Siendo las dimensiones de K_4 igual a $n_4 \times m_4$, donde se recomienda que n_4 no sea mayor a 12 píxeles y m_4 se encuentre entre 70 y 120 píxeles.

4. Conclusiones

En este artículo se ha propuesto un algoritmo de visión por ordenador para una aplicación de agricultura inteligente: la detección de altura del cultivo. Mediante el uso de operaciones morfológicas, detección de bordes, y transformada de Hough, se detectan automáticamente los límites de la planta. Posteriormente, mediante el empleo de visión estéreo, se consigue estimar la altura de la banda. Los resultados han sido probados en imágenes reales, obteniendo un error de 3% en el cálculo de la altura.

Como trabajo futuro se prevé la implementación de los algoritmos desarrollados en dispositivos compactos, con capacidad de cómputo autónomo y de comunicación con la nube, para ayudar a los agricultores mediante monitoreo automático.

Agradecimientos

Los autores agradecen a la Junta de Andalucía (proyecto IRRIGATE-PY20 RE 017 LOYOLA) y a la Agencia Andaluza de Cooperación Internacional para el Desarrollo AACID (proyecto Agricultura 4.0) por financiar parcialmente este trabajo.

English summary

Computer vision for automatic detection of crop heights

Abstract

This article introduces a new algorithm based on stereoscopic imaging and kernels for the automatic and intelligent detection of crop heights. The kernel-based algorithm alleviates the need for humans in the loop cutting costs by prioritizing personnel time and quickly identifying stunted growth due to stress. The kernels algorithm consists of morphological operations and the Hough transform to identify the bottom and top of a specific crop by varying kernel sizes. The algorithm is computationally low cost allowing for IoT based application and produces as little as 3% error.

Keywords: Computer vision, Stereoscopic vision, Kernels, Intelligent agriculture

Referencias

- [1] Assidiq, A. A., Khalifa, O. O., Islam, M. R., Khan, S. (2008, May). Real time lane detection for autonomous vehicles. In 2008 International Conference on Computer and Communication Engineering (pp. 82-88). IEEE.
- [2] Cai, J., Walker, R. (2010). Height estimation from monocular image sequences using dynamic programming with explicit occlusions. *IET Computer Vision*, 4(3), 149-161.
- [3] Corke, P., & Khatib, O. (2011). *Robotics, vision and control: fundamental algorithms in MATLAB* (Vol. 73, p. 2). Berlin: Springer.
- [4] Gomes, J. F. S., & Leta, F. R. (2012). Applications of computer vision techniques in the agriculture and food industry: a review. *European Food Research and Technology*, 235(6), 989-1000
- [5] Jiang, Y., Li, C., Paterson, A. H. (2016). High throughput phenotyping of cotton plant height using depth images under field conditions. *Computers and Electronics in Agriculture*, 130, 57-68.
- [6] Kim, Y., Glenn, D. M., Park, J., Ngugi, H. K., Lehman, B. L. (2011). Hyperspectral image analysis for water stress detection of apple trees. *Computers and Electronics in Agriculture*, 77(2), 155-160.
- [7] Kim, W. S., Lee, D. H., Kim, Y. J., Kim, T., Lee, W. S., Choi, C. H. (2021). Stereo-vision-based crop height estimation for agricultural robots. *Computers and Electronics in Agriculture*, 181, 105937.
- [8] Malekabadi, A. J., Khojastehpour, M., & Emadi, B. (2019). Disparity map computation of tree using stereo vision system and effects of canopy shapes and foliage density. *Computers and electronics in agriculture*, 156, 627-644.
- [9] Rovira-Más, F., Zhang, Q., Reid, J. F. (2008). Stereo vision three-dimensional terrain maps for precision agriculture. *Computers and electronics in agriculture*, 60(2), 133-143.
- [10] Riquelme, M. T., Barreiro, P., Ruiz-Altisent, M., Valero, C. (2008). Olive classification according to external damage using image analysis. *Journal of Food Engineering*, 87(3), 371-379.
- [11] Remboski, T. B., de Souza, W. D., de Aguiar, M. S., Ferreira Jr, P. R. (2018, April). Identification of fruit fly in intelligent traps using techniques of digital image processing and machine learning. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing* (pp. 260-267).
- [12] Shimada, M., Muhtar, Q., Tadono, T., Wakabayashi, H. (2001, July). Tree height estimation using an airborne L-band polarimetric interferometric SAR. In *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium* (Cat. No. 01CH37217) (Vol. 3, pp. 1430-1432). IEEE.
- [13] Tian, H., Wang, T., Liu, Y., Qiao, X., Li, Y. (2020). Computer vision technology in agricultural automation—A review. *Information Processing in Agriculture*, 7(1), 1-19.
- [14] Torres-Sánchez, J., López-Granados, F., Serrano, N., Arquero, O., Peña, J. M. (2015). High-throughput 3-D monitoring of agricultural-tree plantations with unmanned aerial vehicle (UAV) technology. *PloS one*, 10(6), e0130479.
- [15] Yamamoto, K., Takahashi, T., Miyachi, Y., Kondo, N., Morita, S., Nakao, M.,... Murate, N. (2011). Estimation of mean tree height using small-footprint airborne LiDAR without a digital terrain model. *Journal of forest research*, 16(6), 425-431.
- [16] Zhang, F., Fu, L. S. (2014). Application of computer vision technology in agricultural field. In *Applied Mechanics and Materials* (Vol. 462, pp. 72-76).



© 2022 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution CC-BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>).