

UNIVERSIDAD AUTÓNOMA CHAPINGO

POSGRADO EN INGENIERÍA AGRÍCOLA Y USO INTEGRAL DEL AGUA

RECONOCIMIENTO AUTOMÁTICO DEL TALLO EN CULTIVO DE TOMATE EN INVERNADERO MEDIANTE VISIÓN POR COMPUTADORA

TESIS

QUE COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:

MAESTRO EN INGENIERÍA AGRÍCOLA
Y USO INTEGRAL DEL AGUA

PRESENTA:

ABNER ORLANDO CHÁVEZ SÁNCHEZ



DIRECCIÓN GENERAL ACADÉMICA
DEPTO. DE SERVICIOS ESCOLARES
OFICINA DE EXÁMENES PROFESIONALES

Chapingo, Estado de México, Diciembre de 2016



**RECONOCIMIENTO AUTOMÁTICO DEL TALLO EN CULTIVO DE TOMATE
EN INVERNADERO MEDIANTE VISIÓN POR COMPUTADORA**

Tesis realizada por **Abner Orlando Chávez Sánchez** bajo la dirección del Comité Asesor indicado, aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

**MAESTRO EN INGENIERÍA AGRÍCOLA
Y USO INTEGRAL DEL AGUA**

DIRECTOR:



DR. PEDRO CRUZ MEZA

ASESOR:



DR. NOÉ VELÁZQUEZ LÓPEZ

ASESOR:



DR. AGUSTÍN RUIZ GARCÍA

AGRADECIMIENTOS

Al consejo Nacional de Ciencia y Tecnología (CONACYT) por el financiamiento otorgado para realizar mis estudios de maestría.

A la Universidad Autónoma Chapingo (UACH), al posgrado en Ingeniería Agrícola y Uso Integral del Agua (IAUIA) por brindarme la oportunidad de realizar mis estudios.

Al Dr. Pedro Cruz Meza por aceptar dirigir este trabajo, su confianza, paciencia y pero sobre todo por sus valiosos consejos.

Al Dr. Noé Velázquez López, por sus enseñanzas y valiosas sugerencias en el proceso de revisión de la tesis y artículos científicos.

Al Dr. Agustín Ruiz García, por sus valiosas sugerencias en la realización de este trabajo, pero sobre todo por su apoyo incondicional para el proceso de revisión de los artículos científicos.

Agradezco infinitamente a la Dra. Luz Abril Torres Méndez, quien me brindó la oportunidad de realizar una estancia de investigación en el CINVESTAV Unidad Saltillo, por facilitarme su vasto conocimiento en visión por computadora y por su apoyo incondicional en el desarrollo de la investigación. Además por su gran calidad humana.

A todos los profesores que han contribuido a mi formación académica y humana.

DEDICATORIA

A mis padres Joaquín Chávez López y Lourdes Sánchez López quienes me dieron la vida y motivan constantemente para alcanzar mis anhelos.

A mis hermanos

Erick Omar, Edwin Gabriel, Diego Joaquín, Brenda Florida, Irvin Darío y Yael Vidal por su amistad.

A Ana Cristina Sánchez Martínez por su apoyo incondicional

A mis amigos y compañeros

DATOS BIOGRÁFICOS

Abner Orlando Chávez Sánchez nació el 10 de Mayo de 1992 en San Pedro Mártir, Oaxaca, México. Curso su educación básica en escuelas públicas de su lugar de origen. Posteriormente en el 2007 ingresó a la Preparatoria Agrícola de la Universidad Autónoma Chapingo (UACH), terminando en el 2010.

En ese mismo año ingreso a la licenciatura en Ingeniería Mecánica Agrícola de la misma Universidad, concluyendo en el 2014 y obteniendo el título de Ingeniero mecánico agrícola. En agosto del 2014 ingresa a la Maestría en Ingeniería Agrícola y Uso Integral del Agua en la orientación de Mecanización agrícola

CONTENIDO

ÍNDICE DE CUADROS	III
ÍNDICE DE FIGURAS.....	IV
RESUMEN.....	V
ABSTRACT	V
CAPITULO 1. INTRODUCCIÓN GENERAL	1
1.1 Contexto de la visión artificial en la robótica agrícola	1
1.2 Descripción del trabajo	4
1.3 Hipótesis.....	5
1.4 Objetivos	5
1.5 Literatura citada	6
CAPITULO 2. REVISIÓN DE LITERATURA.....	10
2.1 Visión por computadora en la agricultura	11
2.2 Visión por computadora en invernadero	12
2.3 Visión basada en control para robots agrícolas.....	19
2.4 Esquemas de visión para robots agrícolas.....	19
2.5 Reconocimiento de órganos de la planta.....	23
2.6 Literatura citada	30
CAPITULO 3. SEGMENTACIÓN AUTOMÁTICA DE CULTIVOS EN INVERNADERO UTILIZANDO EL ESPACIO DE COLOR CIE L*A*B*	42
3.3 Resumen.....	42
3.2 Introducción	44

3.3	Materiales y métodos	47
3.4	Resultados y discusión	53
3.5	Conclusiones	59
3.6	Literatura citada	60
CAPITULO 4. RECONOCIMIENTO AUTOMÁTICO DEL TALLO DE PLANTAS DE TOMATE EN INVERNADERO MEDIANTE VISIÓN POR COMPUTADORA .		66
4.1	Resumen.....	66
4.2	Introducción	67
4.3	Materiales y métodos	70
4.4	Resultados y discusión	86
4.5	Conclusiones	91
4.6	Literatura citada	92
CAPITULO 5. CONCLUSIONES Y RECOMENDACIONES		97
5.1	Conclusiones	97
5.2	Recomendaciones.....	99
Apéndice 1. Código para segmentar imágenes del cultivo en invernadero.....		100

ÍNDICE DE CUADROS

Cuadro 1. Especificaciones técnicas de la webcam Logitech modelo C920	47
Cuadro 2. Evaluación de la media de calidad de segmentación por el método ATRWG para el enfoque propuesto y cuatro más.	55
Cuadro 3. Resultados de la calidad de segmentación obtenida con el nuevo método de evaluación.	58
Cuadro 4. Calidad de segmentación con diferente resolución de imágenes evaluado mediante el método propuesto por ATRWG.	58
Cuadro 1. Especificaciones técnicas de las webcam Logitech C920 utilizadas	71

ÍNDICE DE FIGURAS

Figura 1. Algunos ejemplos de algoritmos de visión por computadora y sus aplicaciones.....	11
Figura 2. Esquemas de visión y algoritmos de reconocimiento para diferentes frutos....	20
Figura 1. Equipo para la adquisición de imágenes.....	48
Figura 2. a) Distribución del color del cultivo en el canal a. b) Distribución del color del cultivo en el canal b.....	53
Figura 3. Imágenes del cultivo de jitomate y su segmentación manual.....	54
Figura 4. Ejemplo de resultados de la segmentación..	55
Figura 5. Comparación de la calidad de segmentación del método propuesto con cuatro enfoques, utilizando el índice propuesto por ATRWG.....	56
Figura 6. Resultados de la segmentación bajo invernaderos estructuralmente diferentes	57
Figura 7. Resultado de segmentación de imágenes de tamaño 2112 x 2816 píxeles.....	59
Figura 1. Equipo para la adquisición de imágenes.....	71
Figura 2. Diagrama de flujo para localizar el tallo del cultivo de jitomate a partir de imágenes RGB.	73
Figura 3. Diagrama de flujo para la extracción de hilos de tutoreo	81
Figura 4. Ejemplo del resultado de la segmentación en imágenes rectificadas.	86
Figura 5. Resultado del algoritmo para extracción de tallos potenciales en el 96a par de imágenes.....	88
Figura 6. Etapas para la extracción de puntos pertenecientes al tallo en el 36a par de imágenes.....	89
Figura 7. Ejemplos de los resultados de la localización del tallo bajo diferentes condiciones estructurales e iluminación.....	91

RECONOCIMIENTO AUTOMÁTICO DEL TALLO DE PLANTAS DE TOMATE EN INVERNADERO MEDIANTE VISIÓN POR COMPUTADORA

AUTOMATIC RECOGNITION OF GREENHOUSE TOMATO STEM THROUGH COMPUTER VISION

Abner Orlando **Chávez Sánchez**¹ y Pedro **Cruz Meza**²

RESUMEN

Las labores de cultivo y específicamente la poda del tomate en invernadero demanda una alta cantidad de mano de obra, lo que se refleja en altos costos de producción. La automatización de dicha labor podría ser la solución. Sin embargo un robot para realizar esta actividad debe conocer en todo momento la ubicación del tallo de la planta como una referencia de la posición de las yemas axilares y hojas. Por lo tanto, en este trabajo se propone un nuevo método para el reconocimiento del tallo. Consecuentemente, se desarrolló un algoritmo para segmentar el cultivo en la imagen usando su modelo de distribución de color en el espacio CIE $L^*a^*b^*$. Un par de webcam logitech modelo c920 fueron montadas en una configuración de stereo vision con una línea base de 50 mm. El algoritmo de reconocimiento se basa en el uso de información 3D para segmentar objetos a una distancia específica. Además incluye componentes novedosos como la segmentación del cultivo, el uso de hilos de tutorio como indicador visual, la implementación de interpolación mediante trazadores cúbicos y el uso de componentes principales para determinar la orientación de segmentos. El experimento muestra que el rendimiento del método propuesto fue de 95.0% con una desviación estándar de 9.9%. Además la calidad de segmentación por el método propuesto fue de 93.6% con una desviación estándar de 2.4%. En general, el algoritmo de reconocimiento del tallo en conjunto con el de segmentación puede ser implementado como primer paso para el reconocimiento de yemas axilares y hojas en cultivos que crecen a lo largo de un hilo de tutorio.

Palabras clave: Características invariantes, segmentación, estereo visión, iluminación natural.

ABSTRACT

The cultivation tasks and specifically pruning of the greenhouse tomato requires extensive human labor, which is reflected in high production costs. The automation of this work could be the solution. However, a robot to perform this activity must know at all times the location of the plant stem, as a reference of the position of the axillary buds and leaves. Therefore, a new method for stem recognition is proposed in this paper. Consequently, an algorithm was developed to segment the crop into the image, using its color distribution model in the CIE $L^*a^*b^*$ space. Two Logitech model c920 webcams were mounted in a stereo vision configuration with a baseline of 50 mm. The recognition algorithm is based on the use of 3D information for segmenting objects to a specific distance. It also includes novel components such as the crop segmentation from images, use of support wire as a visual cue, the implementation of interpolation using cubic splines and the use of principal components to determine segment orientation. Experiment shows that the performance of the proposed method was 95.0%, with a standard deviation of 9.9%. In addition the segmentation quality of the proposed method was 93.6%, with a standard deviation of 2.4%. Overall, the proposed stem recognition algorithm in conjunction with the crop segmentation can be used as a first step for the recognition of axillary buds and leaves in crops that grow along a support wire.

Keywords: Invariant features, segmentation, stereo vision, natural lighting.

¹Thesis author

²Thesis advisor

¹Autor de la tesis

²Director de tesis

CAPITULO 1

INTRODUCCIÓN GENERAL

1.1 Contexto de la visión artificial en la robótica agrícola

En las últimas cuatro décadas la población mundial se ha duplicado por lo que la demanda de alimentos crece en respuesta (Boserup, 2005). En consecuencia a esto y debido a que las tierras para producción de cultivos son limitadas se genera la necesidad de aumentar el rendimiento de los cultivos al mismo tiempo de reducir los costos de producción. El manejo de hortalizas en invernadero puede brindar una solución mediante la prolongación de su temporada y la intensificación de la producción agrícola (Li *et al.*, 2016). La intensificación no solo es en el sentido de rendimiento y en la producción durante todo el año, sino también en el de los costos de producción.

Valenciano *et al.*, (2003) en un estudio realizado para la medición de costos en la producción intensiva de tomate en suelo y sustrato, reportan que la mano de obra es el costo más importante representando casi el 35% del costo total para el caso de producción en suelo y 32% para la producción en sustrato. El concepto más relevante para el caso de la producción en suelo son las labores de cultivo (podas, tutorio), estas representan el 46% del total, la cosecha el 32% y el mantenimiento y la limpieza el 14%. Para el caso de la producción en sustrato los dos conceptos más importantes son la recolección y las labores de cultivo con un porcentaje conjunto que alcanza casi el 90%. Por lo tanto, mantener el ritmo de producción con respecto al aumento de la población, así como, reducir los altos costos debido a la mano de obra empleada en labores de cultivo, cosecha y monitoreo se está convirtiendo en un desafío para la agricultura.

Una alternativa prometedora es el uso de robots agrícolas, éstos tienen la capacidad de mantener la calidad del producto o trabajo (Van Henten *et al.*, 2003). Los robots agrícolas están enfocados principalmente a tareas de cosecha. Así pues, existen numerosas investigaciones en las que plantean metodologías para esta actividad. Entre ellas se encuentran, cosecha de pepino (van Henten *et al.*, 2002; Yuan *et al.*, 2011), fresa (Leonard *et al.*, 2013; Qingchun *et al.*, 2012), chile pimiento (Barth *et al.*, 2016), jitomate (Zhao *et al.*, 2016b). A pesar de que la cosecha robotizada tiene un gran potencial para el futuro, en la actualidad el rendimiento global del robot es a menudo insuficiente para competir con la operación manual (Grift *et al.*, 2008). Por otro lado, a pesar de que las labores de cultivo consumen el mayor porcentaje de mano de obra, no se han reportado trabajos relacionados con la aplicación de robots para realizar de dichas labores.

El cuello de botella para promover la aplicación del robot en la agricultura se encuentra en el rendimiento del control basado en visión (Zhao *et al.*, 2016a). Este es el principio mediante el cual el robot lleva a cabo la tarea que le es designada bajo la guía de información visual. Este se construye con dos objetivos: 1) el reconocimiento del objeto de trabajo, 2) la coordinación ojo-mano (Hashimoto, 2003). El reconocimiento automático del objeto consiste en su identificación y localización en una escena compleja natural. La coordinación ojo-mano de robot de cosecha tiene que ver con la interacción entre el robot, la percepción visual del espacio de trabajo y el control de actuadores (Goncalves y Torres, 2010).

Por lo tanto el algoritmo de reconocimiento es un factor clave en el rendimiento de un sistema de visión. Se han reportado para el reconocimiento de diversos frutos: tomate (Arefi y Motlagh, 2013; Wei *et al.*, 2014; Yamamoto *et al.*, 2014), chile pimiento (Song *et al.*, 2014), manzanas (Zhao *et al.*, 2005; Kelman y Linker, 2014), fresa (Xie *et al.*, 2011), cítricos (Kurtulmus *et al.*, 2011; Lu *et al.*, 2014) y kiwi (Zhan *et al.*, 2013). Para la aplicación de robots en labores de cultivo son necesarios algoritmos para el reconocimiento de los distintos órganos de la planta, tales como; tallos, hojas, yemas axilares, raquis de hojas y pedúnculos de los raquis centrales de las hojas.

En estos casos la dificultad se incrementa debido a que no es posible el uso de estructuras geométricas, ya que órganos como yemas axilares y tallos presentan características físicas y color similar. La detección y localización del tallo se ha estudiado principalmente en condiciones de iluminación controlada (Paprocki *et al.*, 2012). Sin embargo, en condiciones de iluminación natural se han reportado pocos trabajos. Van Henten *et al.* (2006) describen el modelo de un robot para la poda de hojas, en el que implementan un sistema de visión para detectar y localizar tallos y hojas basado en las propiedades espectrales (NIR) e imágenes del espectro visible RGB. Noble y Li (2012) examinaron características para la segmentación de frutas, hojas y tallos utilizando también imágenes espectrales.

Más tarde, Bac *et al.* (2013) desarrollan un algoritmo con distintos objetivos, uno de ellos fue clasificar la vegetación en obstáculos blandos (parte superior de una hoja, parte inferior de la hoja y pecíolo) y obstáculos duros (tallo y fruta). Fernández *et al.* (2013) propone un algoritmo que combina imágenes RGB y multiespectrales para la discriminación de órganos de la vid en ambientes naturales no estructurados. Luo *et al.* (2016) proponen un método para la adquisición de información espacial a partir de racimos de uvas. En estos trabajos ya sea a base de imágenes multiespectrales o imágenes RGB se extraen características del cultivo por lo que el primer objetivo es segmentarlo.

En esta sentido se ha venido investigando la segmentación de cultivos en ambientes no controlados. Bai *et al.* (2013) describe un algoritmo de segmentación donde establece el modelo de color del cultivo en el espacio de color CIE $L^*a^*b^*$. Así mismo se han reportado diversas metodologías para la segmentación de cultivos (Kataoka *et al.*, 2003; Neto, 2004; Hague *et al.*, 2006; Zheng *et al.*, 2009; Yu *et al.*, 2013), sin embargo todos estos trabajos se realizan a campo abierto o utilizando cámaras multiespectrales, lo que resulta en un alto costo de implementación, al menos de nuestro conocimiento no se ha reportado algún trabajo implementando un sistema de bajo costo en invernaderos bajo iluminación natural y sin la necesidad de iluminación artificial. Por lo tanto, es necesario desarrollar algoritmos de segmentación robustos a la iluminación natural dentro de un invernadero, que permitan extraer características de los órganos de la planta para su reconocimiento.

La segmentación es de igual manera esencial para el monitoreo del cultivo, a pesar de que el tratamiento ante plagas y enfermedades y la aplicación de nutrientes consumen solo el 7 % de la mano de obra (Valenciano *et al.*, 2003), es de vital importancia para el manejo del cultivo. Por lo anterior en este trabajo se desarrollan algoritmos para la extracción de información visual de imágenes del cultivo de tomate.

1.2 Descripción del trabajo

El objetivo de este trabajo se aborda en tres capítulos. En el capítulo 2 se presenta una descripción del estado del arte sobre la aplicación de la visión por computadora en la agricultura. En él se presentan los principales algoritmos reportados para solucionar los problemas comunes en invernadero (monitorizar requerimientos nutricionales, control de plagas y enfermedades, estimación del requerimiento hídrico), además, se presentan los esquemas de visión y algoritmos utilizados por robots agrícolas para automatizar la cosecha de frutos.

En el tercer capítulo se desarrolla un algoritmo para segmentar imágenes del cultivo de tomate en invernadero, el rendimiento obtenido se compara con métodos reportados en la literatura. Se evalúa la relación entre la dimensión de las imágenes y la calidad de segmentación obtenida, así mismo, se determina el tiempo computacional para cada caso.

En el cuarto capítulo se desarrolla y evalúa un algoritmo para el reconocimiento del tallo del cultivo de tomate en imágenes. Estas fueron adquiridas en 4 invernaderos con características diferentes. El algoritmo se compone de 7 etapas incluyendo la rectificación y estéreo-correspondencia, un apartado de segmentación, extracción de tallos potenciales y segmentos de hilo de tutoreo, la determinación del tallo verdadero, interpolación y la extracción de coordenadas (u, v) del tallo.

Finalmente en el quinto capítulo se realizan las conclusiones y recomendaciones generales del trabajo de investigación.

1.3 Hipótesis

Las imágenes de tomate en invernadero tienen la capacidad de proporcionar información visual para segmentar el cultivo y llevar a cabo el reconocimiento del tallo. Esto como primera etapa para el realizar el reconocimiento de los diferentes órganos de la planta y su posterior implementación en un robot agrícola para la poda de hojas y yemas axilares.

1.4 Objetivos

1.4.1 Objetivo general

Desarrollar un algoritmo para el reconocimiento automático del tallo de la planta de tomate en imágenes adquiridas en invernadero bajo condiciones de iluminación natural mediante visión por computadora.

1.4.2 Objetivo específico

- 1) Adquirir una base de datos de imágenes del cultivo de tomate en invernadero en condiciones de iluminación natural.
- 2) Determinar el modelo de distribución de color del cultivo en el espacio de CIE $L^*a^*b^*$.
- 3) Desarrollar un algoritmo para segmentar imágenes de tomate adquiridas en invernaderos bajo condiciones de iluminación natural (Apéndice 1).
- 4) Evaluar el rendimiento del algoritmo de segmentación del cultivo de tomate en condiciones de iluminación natural.
- 5) Evaluar el rendimiento del algoritmo de reconocimiento del tallo en imágenes del cultivo de tomate en condiciones de iluminación natural.
- 6) Implementar el algoritmo de segmentación y de reconocimiento del tallo de la planta de tomate en el lenguaje de programación C++ utilizando las librerías de OpenCV en su versión 3.1.0.
- 7) Proponer una metodología para la adquisición de imágenes.

1.5 Literatura citada

- Arefi, A., & Motlagh, A. M. (2013). Development of an expert system based on wavelet transform and artificial neural networks for the ripe tomato harvesting robot. *Australian Journal of Crop Science*, 7(5), 699.
- Bac, C. W., Hemming, J., & van Henten, E. J. (2013). Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper. *Computers and Electronics in Agriculture*, 96, 148-162. doi: <http://dx.doi.org/10.1016/j.compag.2013.05.004>
- Bai, X. D., Cao, Z. G., Wang, Y., Yu, Z. H., Zhang, X. F., & Li, C. N. (2013). Crop segmentation from images by morphology modeling in the CIE L*a*b* color space. *Computers and Electronics in Agriculture*, 99, 21-34. doi: <http://dx.doi.org/10.1016/j.compag.2013.08.022>
- Barth, R., Hemming, J., & van Henten, E. J. (2016). Design of an eye-in-hand sensing and servo control framework for harvesting robotics in dense vegetation. *Biosystems Engineering*, 146, 71-84. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2015.12.001>
- Boserup, E. (2005). *The conditions of agricultural growth: The economics of agrarian change under population pressure*: Transaction Publishers.
- Fernández, R., Montes, H., Salinas, C., Sarria, J., & Armada, M. (2013). Combination of RGB and multispectral imagery for discrimination of cabernet sauvignon grapevine elements. *Sensors*, 13(6), 7838-7859. doi: 10.3390/s130607838
- Goncalves, P., & Torres, P. (2010). *Learning Approaches to Visual Control of Robotic Manipulators*: Citeseer.
- Grift, T. E., Zhang, Q., Kondo, N., & Ting, K. C. (2008). Review of Automation and Robotics for the BioIndustry. *Journal of Biomechatronics Engineering*, 1(1), 37-54.
- Hashimoto, K. (2003). A review on vision-based control of robot manipulators. *Advanced Robotics*, 17(10), 969-991.
- Hague, T., Tillett, N., & Wheeler, H. (2006). Automated crop and weed monitoring in widely spaced cereals. *Precision Agriculture*, 7(1), 21-32.
- Kataoka, T., Kaneko, T., Okamoto, H., & Hata, S. (2003, 20-24 July 2003). *Crop growth estimation system using machine vision*. Paper presented at the Advanced

- Intelligent Mechatronics, 2003. AIM 2003. Proceedings. 2003 IEEE/ASME International Conference on.
- Kelman, E., & Linker, R. (2014). Vision-based localisation of mature apples in tree images using convexity. *Biosystems Engineering*, 118, 174-185. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2013.11.007>
- Kurtulmus, F., Lee, W. S., & Vardar, A. (2011). Green citrus detection using ‘eigenfruit’, color and circular Gabor texture features under natural outdoor conditions. *Computers and Electronics in Agriculture*, 78(2), 140-149. doi: <http://dx.doi.org/10.1016/j.compag.2011.07.001>
- Leonard, K., Strawbridge, R., Lindsay, D., Barata, R., Dawson, M., & Averion, L. (2013, 24-27 June 2013). *Minimal Geometric Representation and Strawberry Stem Detection*. Paper presented at the Computational Science and Its Applications (ICCSA), 2013 13th International Conference on.
- Li, C., Chen, J., Wang, J., Han, P., Luan, Y., Ma, X., & Lu, A. (2016). Phthalate esters in soil, plastic film, and vegetable from greenhouse vegetable production bases in Beijing, China: Concentrations, sources, and risk assessment. *Science of The Total Environment*, 568, 1037-1043. doi: <http://dx.doi.org/10.1016/j.scitotenv.2016.06.077>
- Lu, J., Sang, N., Hu, Y., & Fu, H. (2014). Detecting citrus fruits with highlight on tree based on fusion of multi-map. *Optik - International Journal for Light and Electron Optics*, 125(8), 1903-1907. doi: <http://dx.doi.org/10.1016/j.ijleo.2013.04.135>
- Luo, L., Tang, Y., Zou, X., Ye, M., Feng, W., & Li, G. (2016). Vision-based extraction of spatial information in grape clusters for harvesting robots. *Biosystems Engineering*, 151, 90-104. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2016.08.026>
- Neto, J. (2004). A combined statistical-soft computing approach for classification and mapping weed species in minimum-tillage systems.
- Noble, S. D., & Li, D. (2012). *Segmentation of greenhouse cucumber plants in multi-spectral imagery*. Paper presented at the International Conference of Agricultural Engineering, CIGR-Ageng, Valencia, Spain.

- Paproki, A., Sirault, X., Berry, S., Furbank, R., & Fripp, J. (2012). A novel mesh processing based technique for 3D plant analysis. *BMC Plant Biology*, 12(1), 63.
- Qingchun, F., Wengang, Z., Quan, Q., Kai, J., & Rui, G. (2012, 25-27 May 2012). *Study on strawberry robotic harvesting system*. Paper presented at the Computer Science and Automation Engineering (CSAE), 2012 IEEE International Conference on.
- Song, Y., Glasbey, C. A., Horgan, G. W., Polder, G., Dieleman, J. A., & van der Heijden, G. W. A. M. (2014). Automatic fruit recognition and counting from multiple images. *Biosystems Engineering*, 118, 203-215. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2013.12.008>
- Valenciano, J., Moreno, M. C. E., & Mesa, J. C. P. (2003). *Costes de producción y utilización de la mano de obra en tomate: un estudio empírico para el cultivo bajo plástico en Almería*. Paper presented at the Anales de economía aplicada 2003, Asociación Española de Economía Aplicada, ASEPELT.
- van Henten, E. J., Hemming, J., Van Tuijl, B., Kornet, J., Meuleman, J., Bontsema, J., & Van Os, E. (2002). An autonomous robot for harvesting cucumbers in greenhouses. *Autonomous Robots*, 13(3), 241-258.
- Van Henten, E. J., Van Tuijl, B. A. J., Hemming, J., Kornet, J. G., Bontsema, J., & Van Os, E. A. (2003). Field Test of an Autonomous Cucumber Picking Robot. *Biosystems Engineering*, 86(3), 305-313. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2003.08.002>
- Van Henten, E. J., Van Tuijl, B. A. J., Hoogakker, G. J., Van Der Weerd, M. J., Hemming, J., Kornet, J. G., & Bontsema, J. (2006). An Autonomous Robot for De-leafing Cucumber Plants grown in a High-wire Cultivation System. *Biosystems Engineering*, 94(3), 317-323. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2006.03.005>
- Wei, X., Jia, K., Lan, J., Li, Y., Zeng, Y., & Wang, C. (2014). Automatic method of fruit object extraction under complex agricultural background for vision system of fruit picking robot. *Optik - International Journal for Light and Electron Optics*, 125(19), 5684-5689. doi: <http://dx.doi.org/10.1016/j.ijleo.2014.07.001>
- Xie, Z., Ji, C., Guo, X., & Zhu, S. (2011). Detection and location algorithm for overlapped fruits based on concave spots searching. [Article]. *Nongye Jixie*

- Xuebao/Transactions of the Chinese Society of Agricultural Machinery*, 42(12), 191-196.
- Yamamoto, K., Guo, W., Yoshioka, Y., & Ninomiya, S. (2014). On Plant Detection of Intact Tomato Fruits Using Image Analysis and Machine Learning Methods. *Sensors*, 14(7), 12191.
- Yu, Z., Cao, Z., Wu, X., Bai, X., Qin, Y., Zhuo, W., . . . Xue, H. (2013). Automatic image-based detection technology for two critical growth stages of maize: Emergence and three-leaf stage. *Agricultural and Forest Meteorology*, 174, 65-84.
- Yuan, T., Ji, C., Chen, Y., Li, W., & Zhang, J. (2011). Greenhouse cucumber recognition based on spectral imaging technology. [Article]. *Nongye Jixie Xuebao/Transactions of the Chinese Society of Agricultural Machinery*, 42(SUPPL.), 172-176.
- Zhan, W., He, D., & Shi, S. (2013). Recognition of kiwifruit in field based on Adaboost algorithm. [Article]. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*, 29(23), 140-146. doi: 10.3969/j.issn.1002-6819.2013.23.019
- Zhao, J., Tow, J., & Katupitiya, J. (2005, 2-6 Aug. 2005). *On-tree fruit recognition using texture properties and color data*. Paper presented at the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems.
- Zhao, Y., Gong, L., Huang, Y., & Liu, C. (2016a). A review of key techniques of vision-based control for harvesting robot. *Computers and Electronics in Agriculture*, 127, 311-323. doi: <http://dx.doi.org/10.1016/j.compag.2016.06.022>
- Zhao, Y., Gong, L., Zhou, B., Huang, Y., & Liu, C. (2016b). Detecting tomatoes in greenhouse scenes by combining AdaBoost classifier and colour analysis. *Biosystems Engineering*, 148, 127-137. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2016.05.001>
- Zheng, L., Zhang, J., & Wang, Q. (2009). Mean-shift-based color segmentation of images containing green vegetation. *Computers and Electronics in Agriculture*, 65(1), 93-98.

CAPÍTULO 2.

REVISIÓN DE LITERATURA

Como seres humanos percibimos la estructura tridimensional del mundo que nos rodea con aparente facilidad. Sin embargo, describirlo a través de una o más imágenes y a partir de estas reconstruir propiedades, tales como la forma y su distribución de color es muy complejo (Szeliski, 2010). La visión por computadora tiene como finalidad la extracción de información del mundo físico a partir de imágenes, utilizando para ello una computadora. Por lo tanto se desarrolla sobre la representación de una realidad que le proporciona información sobre brillo, colores, formas, textura, etcétera. Estas representaciones suelen estar en forma de imágenes estáticas, escenas tridimensionales o imágenes en movimiento (Vélez *et al.*, 2003).

Los investigadores de visión por computadora han estado desarrollando técnicas matemáticas para la recuperación de la forma tridimensional y la apariencia de los objetos a partir de imágenes. Se han desarrollado técnicas fiables para calcular con precisión el modelo 3D parcial de un entorno (Figura 1 a). Teniendo en cuenta un amplio conjunto de puntos de diferentes vistas del objeto, podemos crear modelos de superficie 3D densos (Figura 1 b). Podemos realizar el seguimiento de personas en movimiento en fondos complejos (Figura 1 c y d). Sin embargo, a pesar de todos estos avances, el sueño de tener una computadora que interprete una imagen al mismo nivel que un niño de tres años sigue siendo difícil de alcanzar (Batchelor, 2012).

2.1 Visión por computadora en la agricultura

La visión por computadora se originó en la década de 1950, basándose en el reconocimiento estadístico de patrones. El trabajo principal se centró en el procesamiento, análisis y comprensión de imágenes en dos dimensiones. Debido al nivel de desarrollo de las computadoras, no pudo cumplir con el requisito en tiempo real. Esta tecnología se aplicó en un margen muy estrecho a la agricultura (Zhang y Fu, 2014). A finales de 1970, con la mejora continua de hardware y software, así como el desarrollo de disciplinas relacionadas como el procesamiento de imágenes, reconocimiento de patrones y visión por computadora se introdujo en la investigación del desarrollo agrícola. Donde se han desarrollado diversos trabajos brindando una solución a tareas básicas. En la actualidad, la investigación en el campo de la ingeniería agrícola es un tema muy activo.

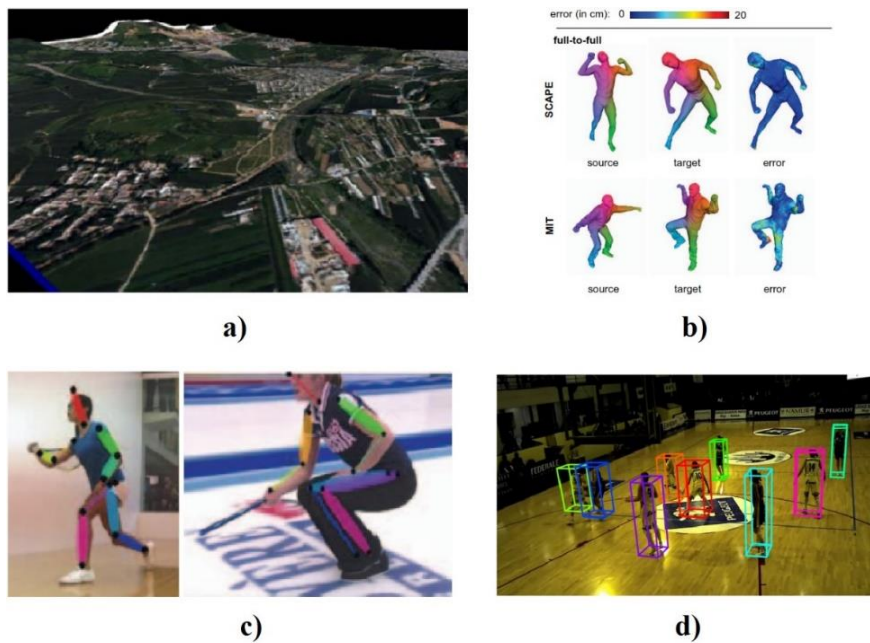


Figura 1. Algunos ejemplos de algoritmos de visión por computadora y sus aplicaciones. a) Generación de ortofotos y modelos tridimensionales de superficie (Wang et al., 2014). b) Algoritmo para obtener mapas 3D densos de personas (Wei et al., 2015). c) diseño sistemático de redes convolucionales para el aprendizaje de características (Wei et al., 2016). d) Seguimiento de múltiples objetos en movimiento y reconstrucción en tres dimensiones (Sekii, 2016).

2.2 Visión por computadora en invernadero

El desarrollo de nuevas tecnologías para la agricultura como son los invernaderos, considerados como sistemas biofísicos con entradas, salidas y bucles de control de procesos (Ehret *et al.*, 2001). Ayudan mediante la automatización de éstos a mejorar las condiciones de producción y por lo tanto obtener mayores rendimientos. Dichas condiciones están determinadas por la observación del cultivo. Sin embargo, la extracción de información en base a muestreos destructivos o invasivos no son prácticos para aplicaciones de monitoreo y control en tiempo real. La visión por computadora tiene el potencial de identificar tensiones emergentes y determinar el factor de estrés (Story *et al.*, 2010). Por esta razón se ha venido desarrollando algoritmos para brindar una solución rápida y oportuna en diferentes aspectos de la producción del cultivo, en las siguientes sub-secciones se muestran algunos ejemplos.

2.2.1 Monitorizar requerimientos nutricionales

Hetzroni *et al.* (1994) utilizaron una red neural y métodos estadísticos para determinar deficiencias de nutrientes (hierro, zinc y nitrógeno) en plantas de lechuga. Así mismo Ling *et al.* (1996) usaron características espectrales y morfológicas de las hojas de lechuga para determinar deficiencia de nutrientes. También Story *et al.* (2010) desarrollan un algoritmo para detectar deficiencias de calcio en el cultivo de lechuga cultivadas en condiciones de invernadero usando cambios de color y morfológicos de la planta. La metodología desarrollada es capaz de identificar deficiencias de calcio en plantas de lechuga 1 día antes de que se perciba por la visión humana. Los marcadores más prometedores para la detección oportuna de la deficiencia de calcio son: el área del cultivo segmentada en la imagen (obtenida a partir de una vista superior), la energía, la entropía, y la homogeneidad.

Posteriormente Mao *et al.* (2015) presentan un método para estimar el contenido de nitrógeno (N) en lechuga, usando una combinación de espectroscopia y visión por computadora. Para la espectroscopia, se usaron 73 variables correspondientes a longitudes de onda.

Para la visión por computadora, 11 características se extrajeron incluyendo el área segmentada a partir de imágenes adquiridas como una proyección de la parte superior del cultivo, la entropía, la energía, y la homogeneidad como característica de textura. Obtuvieron un coeficiente de correlación del 0.8864.

2.2.2 Identificar plagas

Otro problema muy común en la agricultura es el manejo de plagas, estas son unos de los factores críticos causantes de pérdidas económicas en cultivos bajo invernadero (Xia *et al.*, 2015). Se ha tenido una atención especial, principalmente a plagas como la mosca blanca (*Bemisia tabaci* Genn), pulgones (*Aphis gossypii* Glover) y trips (*Thrips tabaci* L.). Martin y Thonnat (2007) presentan un enfoque de visión cognitiva que ajusta los parámetros óptimos para la segmentación de la mosquita blanca en hojas, basados en técnicas de aprendizaje adaptativo. Boissard *et al.* (2008) también propusieron un enfoque multidisciplinario de visión cognitiva aplicable a mosquita blanca pero en rosas. Solis-Sánchez *et al.* (2009) utilizaron características geométricas (excentricidad, tamaño) de la mosquita blanca para su segmentación a partir de imágenes. Xia *et al.* (2012) también desarrollaron un algoritmo robusto para reconocer mosquita blanca, basado en el análisis multifractal.

Li *et al.* (2015) adoptaron el análisis multifractal para la segmentación de imágenes de mosca blanca en base a la selección de mínimos regionales, la precisión obtenida fue de 88,6%. Xia *et al.* (2015) proponen un método automático de identificación de plagas comunes (mosca blanca, pulgón y trips) adecuado para gran escala. Obtienen una R^2 de 0,934 para la mosca blanca, 0,925 para trips y 0,945 para áfidos.

2.2.3 Identificar enfermedades

El control de enfermedades es una prioridad para los productores, y la detección temprana de estas es un componente esencial de un programa de manejo integrado. Se han reportado una amplia gama de técnicas de procesamiento de imagen y de aprendizaje aplicados a la detección de enfermedades.

Técnicas basadas en cámaras hiper-espectrales muestran resultados prometedores, sin embargo tienen limitaciones para su implementación (Mahlein *et al.*, 2013; Zhao *et al.*, 2012). Además, requieren de una calibración frecuente y son sensibles a los cambios de iluminación (Kim, 2011).

Los sistemas basados en cámaras RGB para detectar enfermedades y plagas en campo han logrado tasas de detección favorables (Velázquez-López *et al.*, 2011; Zhang y Meng, 2011; Zheng *et al.*, 2009), sin embargo, introducen limitaciones en las imágenes de entrada que los hacen inadecuados para la detección autónoma. Más tarde Wspanialy y Moussa (2016) presentan un algoritmo para la detección temprana de enfermedades en invernaderos, el sistema utiliza iluminación artificial para eliminar los efectos causados por la iluminación natural y el método Hough forest para detectar cenicilla, obteniendo una tasa de detección del 85%.

2.2.4 Determinar requerimientos hídricos

El consumo de agua es actualmente uno de los principales factores que aumentan la productividad agrícola. El principio fundamental para la gestión del riego es el de satisfacer las necesidades de agua de los cultivos, al mismo tiempo de mantener los beneficios agronómicos con el menor consumo posible. El porcentaje de cobertura del cultivo (PGC) obtenido mediante la aplicación de técnicas de segmentación en imágenes es un parámetro útil para determinar requerimientos hídricos, ya que está directamente relacionada con la evapotranspiración (ET) (Escarabajal *et al.*, 2015a),

Se han logrado resultados satisfactorios utilizando el PGC como estimador del coeficiente de cultivo en tomate (Hanson y May, 2006), frijoles (Medeiros *et al.*, 2001) y cebolla (López-Urrea *et al.*, 2009). En lechuga, Fernández-Pacheco *et al.* (2014) desarrollaron una metodología para calcular el coeficiente de cultivo a partir de imágenes de la cubierta vegetal. Escarabajal *et al.* (2015b) desarrollan una metodología para estimar la profundidad de las raíces durante el crecimiento del cultivo de lechuga lo que permite la optimización del riego basándose en el contenido de agua en el suelo.

Escarabajal *et al.* (2015a) desarrollan una metodología para la programación del riego basado en el balance hídrico del cultivo de lechuga, utilizando imágenes para calcular las necesidades de agua en base al coeficiente de cultivo, lograron un aumento del 6,93% en la producción y una reducción del 17,80% en el consumo de agua en comparación con la parcela de referencia.

2.2.5 Obtención de parámetros de crecimiento del cultivo

Los parámetros de crecimiento del cultivo son una herramienta importante para la toma de decisión en sistemas de control inteligente en invernaderos. Los métodos tradicionales para la medición de la variable de crecimiento no se pueden integrar en los sistemas de control en tiempo real. Además, el crecimiento y desarrollo de cultivos es un proceso muy complicado que se ve afectado por diferentes factores entre ellos; genéticos y ambientales (Sun *et al.*, 2016). Actualmente existen varios métodos no destructivos para la estimación de estos parámetros. Li *et al.* (2003) desarrollaron un algoritmo para monitorear el crecimiento de plántulas de pepino en tiempo real, utilizan el área del cultivo proyectada en la imagen para estimar la altura. Obtienen una $R^2 = 0.927$. Escarabajal *et al.* (2015a) presentan una metodología para la obtención de modelos que estimen el crecimiento de las raíces del cultivo de lechuga a partir del porcentaje de cobertura del suelo. Obtienen una $R^2 = 0.981$. Además se han usado enfoques basados en modelos de estructura morfológica (Weilong *et al.*, 2008), fusión de imágenes con sistemas láser para la reconstrucción de parámetros espaciales (Feng *et al.*, 2013), teledetección (Liang *et al.*, 2013), y espectroscopía (Lin *et al.*, 2013a).

Los trabajos reportados para la identificación de malezas, la nutrición, el requerimiento hídrico, etc., se han centrado en algoritmos basados en la segmentación de la vegetación para la estimación de parámetros.

2.2.6 Segmentación de imágenes de cultivos

Sin embargo, la segmentación de imágenes de cultivos se complica por el hecho de que el color en ella varía con respecto a la iluminación en condiciones naturales (Bai *et al.*, 2013; Kim *et al.*, 2008), por ejemplo, la temperatura de color de la luz del día abarca un amplio rango y varía con la acumulación de nubes, el clima y la posición del sol.

Para una fuente de luz diurna, la temperatura de color por lo general oscila entre 3400 K (puesta del sol) a 9500 K (alba). Incluso al mediodía la temperatura de color puede variar de 6000 K (soleado) a 9300 K (cielo nublado) (Tian y Slaughter, 1998).

En general, se sabe que se consigue robustez en la segmentación si un espacio de color separa eficientemente la crominancia y luminancia de una imagen, y además se utiliza un modelo plausible de distribución de crominancia (Terrillon *et al.*, 1998). Los investigadores han considerado otras características del cultivo para compensar la variación de iluminación. Guijarro *et al.* (2011) proponen un enfoque automático para la segmentación de plantas, suelo y cielo, basado en la textura de estas tres clases y en un enfoque fuzzy clustering supervisado para identificar sub-texturas automáticamente.

Persson y Astrand (2008) investigaron modelos de forma activa para la extracción y clasificación del cultivo de remolacha azucarera. Compararon los resultados de la clasificación automática de plantas con la realizada manualmente obteniendo un rendimiento del 81-87%. Sin embargo, características de forma para muchos cultivos no son muy estables y significativas (Bai *et al.*, 2013). Los enfoques de segmentación basados en color tienen ventajas significativas, ya que es una señal visual eficaz y esencial para distinguir un objeto de entre otros (Kim *et al.*, 2008). Además, el color es casi invariante frente a cambios de tamaño, orientación y oclusión parcial del objeto bajo condiciones de iluminación constante (Phung *et al.*, 2005). Por lo que en esta revisión, nos concentramos en la segmentación de cultivos basados en el color.

Recientemente, se han propuesto diversos métodos para solucionar el problema de segmentación de cultivos en ambientes naturales, Woebbecke *et al.* (1995) proponen el uso de las coordenadas cromáticas rojo (r), verde (g) y azul (b) definidas como:

$$r = \frac{R}{R+G+b}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B} \quad (1)$$

Donde R, G y B son las coordenadas normalizadas del espacio de color RGB en un rango de 0 a 1, calculados como:

$$R = \frac{R}{R_{max}}, \quad G = \frac{G}{G_{max}}, \quad B = \frac{B}{B_{max}} \quad (2)$$

Donde $R_{max} = G_{max} = B_{max} = 255$. Así mismo proponen una modificación al canal H del espacio de color HSI para generar índices de color que permitan segmentar plantas bajo diversas condiciones de suelo, residuos de cosechas anteriores e iluminación. Los índices reportados son la substracción de la coordenada cromática verde de la roja ($r - g$), la substracción de la coordenada cromática azul de la verde ($g - b$), la división del primer y segundo índice $\left(\frac{g-b}{|r-g|}\right)$ y el índice de exceso de verde (Ec 3).

$$ExG = 2g - r - b \quad (3)$$

Tian y Slaughter (1998) desarrollan un algoritmo de segmentación adaptativo a las condiciones ambientales (EASA), para la segmentación de plantas a campo abierto. Kataoka *et al.* (2003) desarrollan el índice de vegetación CIVE (índice de color para la extracción de la vegetación) (Ec. 4).

$$CIVE = 0.441r - 0.811g + 0.385b + 18.78745 \quad (4)$$

El valor del umbral se determina mediante Otsu (1975). Este algoritmo fue desarrollado como una primera etapa para determinar la relación entre el área de cobertura de las plantas y las dimensiones reales de esta, usando una curva de crecimiento (Gompertz) y una función exponencial. Neto (2004) reporto un índice (Ec. 5) ExGR (índice de exceso de verde menos exceso de rojo) para segmentar el cultivo en condiciones naturales en función de tres clases, residuos de cultivos anteriores, el suelo y el cultivo.

$$ExGR = ExG - ExR \quad (5)$$

Donde $ExR = 1.4r - g$. Hague *et al.* (2006) describen un método para la segmentación automática de cultivos de cereales en imágenes, proponen un índice vegetativo (VEG), que es invariante a la iluminación natural y se calcula a partir de los canales rojo, verde y azul (Ec. 6).

$$VEG = \frac{g}{r^a b^{1-a}} \quad (6)$$

Donde $a = 0.667$. Perez *et al.* (2000) desarrollan una metodología para detectar malezas de hoja ancha en cultivos de cereales, utilizando información de color para discriminar entre la vegetación y el fondo. Para reducir los problemas de iluminación es usado un índice de diferencia normalizada (NDI), que se define como:

$$NDI = \frac{G-R}{G+R} \quad (7)$$

Donde G es el valor del pixel en el canal verde y R es el valor del pixel en el canal rojo. Este provee valores en un rango de -1 a 1. Esto genera un histograma bimodal por lo que el umbral para la segmentación es obtenido mediante (Otsu, 1975). Sin embargo, como lo menciona Pérez *et al.* (2000), estos métodos tienden a generar umbrales erróneos en presencia de cambios de iluminación.

Otro enfoque son los algoritmos basados en aprendizaje, Zheng *et al.* (2009) propusieron un algoritmo de segmentación mediante una red neuronal BPNN (Backpropagation Neural Network), en la que se supone que la imagen está compuesta de vegetación y suelo. Bai *et al.* (2013) utilizan un método para establecer el modelo de color del cultivo en el espacio de color CIE $L^*a^*b^*$ y a partir de esto realizar la segmentación de imagen. La calidad de segmentación que obtienen es del 87,2% evaluado mediante ATRWG (Automatic Target Recognition Working Group) y el 96,0% para otro método de evaluación. Yu *et al.*, (2013) desarrolló un nuevo método para la segmentación de cultivos que toma en cuenta la variación de iluminación en ambientes complejos. El método se basa en suponer que bajo una intensidad específica, la forma del histograma de tonalidad se comporta de manera similar a una distribución de Gauss. El proceso de calificación requiere de la media y la desviación estándar de cada nivel de intensidad. Así como de un valor para el umbral k adecuado ya que afecta en gran medida la precisión de la segmentación.

Lu *et al.* (2016) proponen un método formulado como un problema de segmentación semántica que se compone de dos etapas: 1) la generación de propuestas de región y 2) la predicción del modelo de color. Para la primera etapa se realiza la segmentación y posteriormente se agrupan las regiones mediante una iteración lineal simple.

En la segunda se usa un conjunto de redes neuronales, con el fin de agrupar categorías en regiones que contienen una región semántica de esa categoría. Este método hace una fuerte suposición que consiste en que la parte de la planta a segmentar se encuentra en una posición específica dentro de la escena. Así pues el color es una característica clave que se ha aplicado ampliamente en el reconocimiento de plantas, frutas, verduras, etc. En este sentido, varios estudios han demostrado la importancia de seleccionar el espacio de color óptimo para cada aplicación (Łuszczkiewicz-Piątek, 2014).

García-Mateos *et al.* (2015) en un estudio comparativo sobre la segmentación de plantas y suelo en cultivos de lechuga, encontraron que el espacio de color CIE L*a*b* es el más adecuado para la segmentación de este cultivo. Sin embargo, aunque las mismas conclusiones fueron apoyadas por otros investigadores (Łuszczkiewicz-Piątek, 2014; SHIH y LIU, 2005), el espacio de color óptimo depende de las características de cada aplicación.

2.3 Visión basada en control para robots agrícolas

A pesar del desarrollo de tecnologías robóticas agrícolas, la falta de algoritmos robustos para el reconocimiento de frutos y en general órganos de la planta ha limitado su aplicación. El uso de la información visual para el control de un manipulador robótico se conoce como control basado en visión (Shirai y Inoue, 1973). Esta técnica se componen de dos etapas; el reconocimiento del fruto y la coordinación ojo-mano (Hashimoto, 2003), el reconocimiento automático del fruto consiste en la identificación y localización de la fruta en una escena natural compleja. Estas dos tareas son la base de la operación mediante el cual el robot manipula el objeto de trabajo. La coordinación ojo-mano de robot consiste en la interacción entre la percepción visual del espacio de trabajo y los actuadores (Goncalves y Torres, 2010).

2.4 Esquemas de visión para robots agrícolas

Las tareas de reconocimiento de órganos de la planta se llevan a cabo por diversos sensores visuales. De acuerdo con el esquema de adquisición de imágenes, estos se clasifican en dos dimensiones (2D) y en tres dimensiones (3D). Las imágenes 2D adquiridas pueden indicar características morfológicas del objetivo como color, forma y textura.

Los sensores visuales tridimensionales proporcionan mapas de coordenadas 3D de la escena que pueden inferir sobre la forma y la ubicación espacial del objeto. El sistema de visión también tiene relación con el proceso de reconocimiento. Como se muestra en la Figura. 2, para identificar diferentes objetos, es necesario seleccionar sensores visuales que propician la aplicación de un determinado algoritmo de reconocimiento.

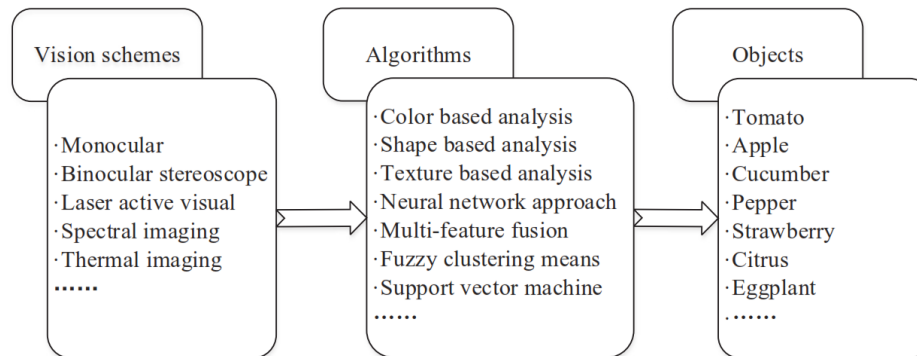


Figura 2. Esquemas de visión y algoritmos de reconocimiento para diferentes frutos (Zhao *et al.*, 2016a)

2.4.1 Esquema monocular

El esquema monocular es un sistema de visión artificial que se compone de una sola cámara, con sensores de dispositivo de carga acoplada (CCD) o sensores semiconductores complementarios de óxido metálico (CMOS). Este esquema es ampliamente usado para la implementación de sistemas autónomos de cosecha en árboles frutales (Jiménez *et al.*, 2000a). Jiménez *et al.* (2000b) describen un sistema de visión artificial a base de un láser para el reconocimiento automático naranja a través del robot AGRIBOT. Integran tanto la distancia como la información de reflectancia para generar cuatro estimadores característicos que den evidencia de la existencia de objetos esféricos.

Bulanon *et al.* (2004) describen un sistema de visión artificial en tiempo real para guiar un manipulador robótico para la cosecha de manzanas. Este se compone de una cámara de vídeo CCD de color y es capaz de reconocer la fruta bajo diferentes condiciones de iluminación. Zhao *et al.* (2005) utilizaron una cámara de color para identificar las manzanas en base a las características de color (verde y rojo) y textura e informó una precisión del 90%.

Baeten *et al.* (2008) desarrollaron un sistema de visión monocular mediante la combinación de una cámara y una fuente de luz de alta frecuencia para detectar las manzanas en el entorno al aire libre. Edan *et al.* (2000) desarrollaron un sistema visión compuesto por múltiples cámaras monoculares para reconocer y localizar el melón, informa que esta configuración podría mejorar la precisión del reconocimiento.

2.4.2 Esquema de estereo-visión

Por otro lado el sistema de visión estereoscópica está compuesto por dos cámaras separadas en una cierta distancia, de tal forma que capturan la misma escena en dos imágenes. Yang *et al.* (2007) presentan un método para detectar y reconocer las agrupaciones de tomate maduros en la planta en un invernadero. Utilizan un sistema de visión estereo compuesto por cámaras de color. El método que proponen realiza una reconstrucción 3D, introduciendo el algoritmo de capa de color creciente (CLG). Xiang *et al.* (2014) también presentan un método basado en visión estereoscópica para el reconocimiento de tomates agrupados. La tasa de precisión fue de 87,9%, esto cuando la oclusión debido a las hojas y ramas fue menor del 25%. La distancia de adquisición y el tiempo medio de ejecución de este método fueron de 300-500 mm y aproximadamente 0,5 s, respectivamente.

Si *et al.* (2015) describen un algoritmo de reconocimiento de manzanas basado en la diferencia de color del canal R-G y la diferencia de color de la relación $(R - G) / (G - B)$. Donde si un píxel cumple con la relación $R - G > 0$ y $(R - G)/(G - B) > 1$ el píxel se clasifica como una manzana, de lo contrario el píxel era definido como fondo. Reportan que el 89,5% de los frutos fueron reconocidos con éxito. La principal desventaja del esquema de visión estereoscópica es su complejidad y tiempo de cálculo debido al procesamiento para la correspondencia del par de imágenes.

2.4.3 Esquema visual activo mediante laser

Existen varias técnicas para obtener la información de profundidad, sin embargo para garantizar una mayor precisión un láser a través de un sistema visual activo es la mejor opción (Zhao *et al.*, 2016a).

Tankgaki *et al.* (2008) desarrollaron un robot para la cosecha de fresa, sus principales partes son un manipulador de 4 grados de libertad (DOF), un sensor de visión en 3D, un efector final y un ordenador. El sensor de visión 3D esta equipado con diodos láser rojo e infrarrojo, ambos rayos láser escanean el objeto al mismo tiempo. Yin *et al.* (2009) desarrollan un algoritmo para el reconocimiento de tomate maduro. Realiza una segmentación mediante K-means clustering usando el espacio de color CIE L*a*b*. Finalmente, hacen uso de un sistema laser para controlar el trayecto del brazo robótico al fruto.

Zhang *et al.* (2015) proponen un nuevo método para el reconocimiento del tallo de la manzana usando un sistema de visión por computadora combinado con un infrarrojo cercano y técnicas de reconstrucción 3D, obtuvieron una precisión del 97,5%. A pesar de que la exactitud del sistema visual activo a base de laser es prometedor su complejidad a menudo limita su aplicación práctica.

2.4.4 Esquema mediante imágenes térmicas

Las imagen térmicas también llamadas termógrafos infrarrojos son una visualización de la radiación infrarroja (Li *et al.*, 2014). Debido a su estructura física y características, las hojas acumulan menos calor que los frutos, por lo que la distribución de temperatura en la planta puede ser utilizada para el reconocimiento de frutos (Vadivambal y Jayas, 2011). Xu y Ying (2004) presentan un método para reconocer cítricos por medio de la termografía infrarroja. Encontraron una diferencia en la distribución térmica del fruto, hojas y ramas de aproximadamente 1 ° C.

Bulanon *et al.* (2004) desarrollan también un enfoque para el reconocimiento de cítricos en ambientes naturales mediante un estudio de la variación temporal térmica del fruto. Stajnko *et al.* (2004) reportan un método para estimar el número de manzanas a campo abierto, para lo cual hacen uso de una cámara térmica y obtienen un coeficiente de correlación $R^2 = 0.83$. El principal problema reportado mediante el uso de cámaras térmicas es que la precisión del método se ve afectado por la sombra de los árboles.

2.4.5 Esquema de imágenes espectrales

El desarrollo de imágenes espectrales ha permitido que sean utilizadas para generar algoritmos de reconocimiento de frutas. Kane and Lee (2007) presentan un método para la creación de mapas de rendimiento de cítricos, para lo cual utilizan una cámara en la longitud del infrarrojo cercano, equipada con filtros (1064, 1150 y 1572 nm) intercambiables. Obtuvieron un rendimiento del 84,5%.

Safren *et al.* (2007) proponen un método basado en visión artificial para la estimación del rendimiento de manzanas en árboles con diferentes etapas de crecimiento, para ello usaron imágenes hiperespectrales donde un algoritmo de múltiples etapas fue desarrollado. El rendimiento del método se evaluó en 88,1%, con una tasa de error global del 14,1%. Okamoto and Lee (2009) desarrollaron un método de procesamiento de imagen para detectar cítricos verdes en árboles individuales. Emplearon una cámara hiperespectral de 369 a 1042 nm para adquirir imágenes de frutos verdes de tres diferentes variedades, mostraron que las tasas de éxito de detección fueron 70 a 85%.

Con la finalidad de realizar el reconocimiento y la cosecha de frutos que presentan colores similares Yuan *et al.* (2011) desarrollan un método para seleccionar un espectro (350-1200nm), en el que se analizan y verifican bandas sensibles para el reconocimiento del pepino mediante análisis de varianza y el uso de una red neuronal. Posteriormente comparan la reflectancia espectral de los órganos (frutos, hojas y flores) de la plantas. La red entrenada llega a una tasa de identificación del 95% para el reconocimiento del pepino.

2.5 Reconocimiento de órganos de la planta

Un robot para la cosecha de frutos en el interior de un invernadero se compone de un manipulador y un efector final. El mismo principio puede ser utilizado para desarrollar robots que automaticen labores culturales como; la poda de yemas axilares y hojas. Estos robots deben tener la capacidad de planificar movimientos para alcanzar su objetivo sin producir colisiones, por lo que es necesario conocer la posición de obstáculos y órganos de la planta que manipulara. Sin embargo, es necesario que exista una etapa de reconocimiento. Por consiguiente, la detección de órganos de la planta es un tema en desarrollo.

2.5.1 Reconocimiento del fruto

Como se mencionó anteriormente el algoritmo de reconocimiento es un factor clave que afecta directamente el rendimiento de un sistema robótico. Así pues numerosas investigaciones informaron sobre algoritmos de reconocimiento para la cosecha robotizada de frutos, sin embargo para órganos de la planta como yemas axilares y hojas estas son limitadas. Los algoritmos de reconocimiento para frutos pueden clasificarse en enfoques de análisis de características únicas, enfoques de análisis de la fusión de múltiples características y enfoques de reconocimiento de patrones.

Análisis de características individuales

El color es una de las características más prominentes usadas para segmentar el fruto de un fondo natural. Sin embargo, esta metodología es sensible a las condiciones de iluminación. Para compensarlo, se utilizan diversos espacios de color tales como; RGB, rgb, XYZ, L*a*b*, L*u*v*, HSV, HLS (García-Mateos *et al.*, 2015; Yin *et al.*, 2009).

Arefi *et al.* (2011) desarrollan un algoritmo de segmentación de tomates maduros para su cosecha mediante un brazo robótico. El algoritmo es adaptable a las condiciones de iluminación en el interior del invernadero, este se desarrolla en dos pasos: (1) extracción del tomate maduro mediante la combinación de los espacios RGB, HSI, y YIQ. (2) se localiza el tomate maduro mediante las características morfológicas de la imagen. Obtienen una precisión total del 96,36%. Mao *et al.* (2009) desarrollan un sistema de visión artificial binocular para la segmentación de manzanas, utilizaron el índice de color Drg-Drb. Posteriormente las manzanas fueron segmentados usando el algoritmo de exploración bidireccional. Obtuvieron una tasa de clasificación correcta del 90%.

Wei *et al.* (2014) también presentan un método para la extracción automática de frutos en condiciones agrícolas. El método se basa en la extracción de características en el espacio de color OHTA para posteriormente utilizarlas como entrada en la determinación del umbral por Otsu (1975), muestran que el método propuesto puede extraer el fruto con una exactitud del 95%.

Sin embargo cuando los frutos no presentan diferencia significativa en color que el resto de los órganos de la planta no es posible hacer uso de este enfoque (Reis *et al.*, 2012), por lo que los investigadores se han dado a la tarea de desarrollar nuevas metodologías para el reconocimiento del fruto. Una de estas metodologías alternas son los algoritmos de reconocimiento basados en la extracción de características geométricas, principalmente para frutos esféricos como tomates, manzanas y cítricos (Liu *et al.*, 2007; Xie *et al.*, 2010). La principal característica de este enfoque es que no se ve afectado por la iluminación.

Xie *et al.* (2007) presentan un método para el reconocimiento de fresa basado en la transformada de Hough. Se compone de 3 pasos, 1) etiquetado de la región en la imagen segmentada, 2) obtención del rectángulo mínimo que encierra la región (MER) y 3) efectuar el reconocimiento basado en la transformada de Hough el cual extrae el contorno de la fresa en el MER. Los resultados experimentales muestran que el método es eficaz tanto para situaciones donde se presenta una sola fresa y situación complejas donde la pérdida del contorno de la fresa es menos de 1/2. El error de reconocimiento en promedio es del 4,8%. Xie *et al.* (2011) también propone un método basado en la transformación circular de Hough, donde la tasa de detección de melocotones es del 90%.

Kelman and Linker (2014) presentan una metodología para el reconocimiento de manzanas en condiciones naturales de iluminación utilizando un análisis de forma. La metodología consiste en una prueba de convexidad que identifica bordes que pudieran corresponder a los objetos convexos tridimensionales. Después, se realiza un segundo análisis para eliminar falsos positivos y eliminar múltiples detecciones de manzanas. Obtuvieron una detección del 94% para manzanas visibles en las imágenes. Con el fin de aumentar la fiabilidad del reconocimiento para entornos no controlados, frutos parcialmente ocluidos y características de fondo similar, investigadores desarrollan algoritmos basados en múltiples características (color, geometría y textura) para reconocer los frutos (Zhao *et al.*, 2016a).

Análisis mediante múltiples características

Hannan *et al.* (2009) desarrollan un algoritmo reconocimiento naranjas en ambientes naturales. Consta de una etapa de segmentación mediante el uso del coeficiente de cromacidad del color rojo, etiquetado de regiones, filtrado por tamaño, extracción y detección basada en perímetro de donde obtuvieron una tasa de falsos positivos del 4%.

Zhao *et al.* (2005) presentan un algoritmo para localizar manzanas en una sola imagen, para ello combinan las características de enrojecimiento combinado con la detección de bordes y textura. Mediante este método obtuvieron una precisión del 90% en el reconocimiento.

Lu *et al.* (2014) desarrollan un método basado en el mapa de aberración cromática (CAM) y el mapa de luminancia (LM) para segmentar cítricos en ambientes naturales. La precisión alcanzada por el método es del 86.81%. Rakun *et al.* (2011) describen un método para la detección del número de manzanas y su diámetro, basado en la combinación de propiedades de color, textura y forma 3D del objeto. Diversos métodos para el reconocimiento de frutos basados en la fusión multi-sensor han sido utilizados (Bulanon y Kataoka, 2010; Bulanon *et al.*, 2009; Feng *et al.*, 2014).

Para hacerle frente al problema de reconocimiento de frutos en la copa de los árboles Bulanon *et al.* (2009) y Wachs *et al.* (2010) utilizaron la fusión de imágenes infrarrojas (IR) y visibles para mejorar su reconocimiento. Los autores argumentaron que la precisión del reconocimiento mediante el enfoque de fusión (74%) se incrementó en comparación con el convencional utilizando solamente imágenes a color (66%) o IR (52%).

Reconocimiento de patrones

Los enfoques de reconocimiento de patrones se han investigado durante mucho tiempo para su aplicación en agricultura protegida. A principios de la década de 1970, Parrish and Goksel (1977) sugirieron que los enfoques de reconocimiento de patrones podrían ser utilizados para el reconocimiento de frutos.

Slaughter y Harrell (1989) desarrollaron un modelo que hace uso del clasificador bayesiano para segmentar naranjas en ambientes naturales de producción utilizando sólo la información de color. El clasificador bayesiano segmentó correctamente más del 75% de los píxeles del fruto

Con el fin de resolver el problema de superposición en el reconocimiento de plántulas, Pastrana y Rath (2013) presentan un enfoque que consta de dos pasos principales: 1) simplificación de la forma de la hoja mediante el uso de una aproximación a la elipse y 2) agrupamiento de las hojas (elipses) encontradas en las plántulas utilizando modelos de forma activa. Kurtulmus *et al.* (2011) desarrollaron un algoritmo para detectar y contar cítricos verdes inmaduros en condiciones naturales utilizando imágenes a color. Donde un análisis de textura Gabor y un enfoque 'eigenfruit' (inspirada en el reconocimiento de rostros 'Eigenface') fueron utilizados para la detección de cítricos verdes. Obtuvieron una precisión del 75,3% para verdaderos positivos. Zhan *et al.* (2013) desarrollaron un método basado en el algoritmo Adaboost para la segmentación del kiwi. Así mismo Zhao *et al.* (2016b) desarrollaron un algoritmo para el reconocimiento automático de tomates maduros en invernadero, el método consiste en una fusión de características de color con el clasificador AdaBoost, sostuvieron que la combinación de estas dos metodologías tiene la capacidad de detectar correctamente más del 96% de tomates maduros.

Actualmente, se está prestando cada vez más atención a los métodos inteligentes de reconocimiento de patrones tales como las redes neuronales artificiales (ANN), máquinas de vectores de soporte (SVM) y el reconocimiento de patrones difusos. ANN y SVM son algoritmos de aprendizaje supervisados que tienen la capacidad de aprender de los datos a través de un proceso de entrenamiento iterativo y mejorar su rendimiento después de cada iteración (Zhao *et al.*, 2016a). Gatica *et al.* (2013) presentan un método para el reconocimiento de aceituna. Este comprende dos etapas, la primera se centró en decidir si es o no el candidato identificado en la imagen un fruto de aceituna, y la segunda en analizar si las aceitunas se solapan en las imágenes. Los análisis requeridos en estas dos etapas se llevan a cabo mediante la aplicación redes neuronales.

Arefi y Motlagh (2013) también desarrollaron un algoritmo para el reconocimiento de tomate maduro, esto se llevó a cabo en dos etapas en la primera se elimina el fondo de la imagen, en la segunda se clasifican tomates maduros de tomates verdes. Esta última se llevó a cabo mediante la integración de procesamiento de imágenes y una red neural artificial. El procesamiento de imágenes se basó en transformación wavelet, la red neuronal desarrollada fue feedforward con dos capas ocultas. Sostienen que la precisión de reconocimiento de tomate maduro es de 95%

Para mejorar la identificación de frutos, Song *et al.* (2014) utilizaron un modelo basado en diccionarios (BoW) para localizar el chile pimienta. El modelo BoW representa cada imagen mediante una distribución de frecuencia de su vocabulario visual. Wang *et al.* (2015) presentaron un algoritmo Fuzzy Clustering Means (FCM) para reconocer los tomates agrupados. La precisión de este algoritmo se verificó de acuerdo con una comparación con K-means y un umbral dado por Otsu (1975). Con el fin de acelerar el cálculo de FCM, Xiong *et al.* (2013) presentaron un algoritmo FCM mejorado basado en la fusión del algoritmo de interpolación cúbica y el algoritmo FCM. Este se encuentra disponible para el reconocimiento de lichis, dentro de sus mejoras se encuentran el que permitía que el sistema de reconocimiento funcionara en tiempo real.

2.5.2 Reconocimiento de tallos, hojas y peciolo

Un robot agrícola en invernadero debe evitar colisiones con la estructura del mismo, así como con los órganos de la planta, para ello es necesario construir un mapa de obstáculos que le permita planificar el movimiento libre de colisiones. Así mismo, en tareas de cosecha y labores de cultivo conocer la posición de los órganos de la planta es indispensable. Por todo esto, el reconocimiento de los órganos de la planta es un tema que necesita ser investigado. Actualmente se ha venido investigando en diversos cultivos.

Noordam *et al.* (2004) desarrollaron un algoritmo para la cosecha de rosas, en el cual estiman la posición de corte del tallo. Evalúan técnicas basadas en estéreo-visión, triangulación laser e intersección volumétrica inversa, sostienen que esta última es la más prometedora para localizar el tallo. Van Henten *et al.* (2006) presentan un robot autónomo para la eliminación de hojas de plantas de pepino cultivadas en invernadero.

En el sistema de visión los tallos son identificados por la reflexión de objetos densos con un alto contenido de agua en la longitud de onda de 850 nm. Por otra parte, afirman que las hojas muestran aproximadamente la misma reflectancia a 850 y 970 nm. Los tallos son localizados al segmentar estructuras delgadas y largas utilizando la erosión y la dilatación.

Lu *et al.* (2011) desarrollan una metodología para reconocer órganos de la planta que podrían fungir como obstáculos para el manipulador. La metodología es desarrollada a base de características espectrales en la longitud de onda visible e infrarroja. Seleccionaron cinco filtros de banda estrecha en la región Vis-NIR para capturar cinco imágenes. Estas fueron filtradas usando el algoritmo de fracción de ruido mínimo (MNF). Posteriormente las imágenes tridimensional fueron construidas para reconocer ramas, utilizando el mapa espectral de ángulos (SAM). Obtienen que imágenes multiespectrales en combinación con MNF-SAM se puede ser eficaz para reconocer órganos de la planta bajo condiciones de iluminación no controladas. Jidong *et al.* (2012) también desarrollan un algoritmo para reconocer ramas y troncos del árbol de manzana, reportan una tasa de detección del 95%.

Noble y Li (2012) realizan un análisis de características espectrales de frutos y hojas para aplicaciones de segmentación. Sostienen que se existen diferencias entre las características espectrales observadas en bandas alrededor de 550, 790, 820, 850, 940 y 970 nm.

Bac *et al.* (2013) proponen un algoritmo también haciendo uso de características espectrales para la segmentación de los órganos de la planta, basado en un clasificador robusto a las variaciones de iluminación. Este clasifica la vegetación principalmente en partes blandas (parte superior de una hoja, parte inferior de la hoja y pecíolo) y obstáculos duros (tallo y fruta). Hacen uso de un sistema multiespectral con iluminación artificial para atenuar las perturbaciones causadas por la iluminación natural. El fondo es segmentado en la longitud del infrarrojo cercano (>900 nm) y objetos como cintas de riego, estacas, elementos de construcción y alambres de soporte, se retiraron usando un umbral en la longitud de onda azul (447 nm).

Para clasificar la vegetación, utilizaron árboles de regresión (CART). Sostienen que obtuvieron una tasa de detección de verdaderos positivos para obstáculos duros de 91,5%, para obstáculos blandos de 40,0%, para tallos de 78,7%, para la parte superior de una hoja de 68,5%, para la parte inferior de una hoja de 54,5%, para las frutas y peciolo de 49,5%. Como lo menciona el autor estos resultados son insuficientes para construir un mapa preciso de obstáculos.

Fernández *et al.* (2013) proponen un algoritmo de enmascaramiento secuencial basado en el método K-means que combina imágenes RGB y multiespectrales para discriminar elementos de la vid. Los píxeles de la imagen se clasifican en cinco grupos que corresponden a las hojas, tallos, ramas, frutos y el fondo. Utilizan una plataforma sensorial que integra una cámara CCD y filtros servo-controlado.

Finalmente Bac *et al.* (2014) desarrollan un método para reconocer el tallo del cultivo de chile pimienta, hacen uso del hilo de tutorio para estimar la posición de este. Una cámara RGB es montada en una corredera neumática para obtener un par de imágenes a una línea base de 1 cm, adicionalmente hacen uso de iluminación artificial para atenuar los efectos causados por la iluminación natural. El algoritmo consta de cinco pasos, obtuvieron que la tasa de detección del alambre (verdaderos positivos) fue más favorable en virtud de una irradiancia moderada (94%) que para una irradiación fuerte (74%). El error de reconocimiento del tallo en el plano horizontal se midió por la distancia euclidiana. Donde el error fue menor para los segmentos interpolados (0,8 cm), que para los segmentos extrapolados (1,5 cm).

2.6 Literatura citada

- Arefi, a., & motlagh, a. M. (2013). Development of an expert system based on wavelet transform and artificial neural networks for the ripe tomato harvesting robot. *Australian journal of crop science*, 7(5), 699.
- Arefi, a., motlagh, a. M., mollazade, k., & teimourlou, r. F. (2011). Recognition and localization of ripen tomato based on machine vision. [article]. *Australian journal of crop science*, 5(10), 1144-1149.

- Bac, c. W., hemming, j., & van henten, e. J. (2013). Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper. *Computers and electronics in agriculture*, 96, 148-162. Doi: <http://dx.doi.org/10.1016/j.compag.2013.05.004>
- Bac, c. W., hemming, j., & van henten, e. J. (2014). Stem localization of sweet-pepper plants using the support wire as a visual cue. *Computers and electronics in agriculture*, 105, 111-120. Doi: <http://dx.doi.org/10.1016/j.compag.2014.04.011>
- Baeten, j., donné, k., boedrij, s., beckers, w., & claesens, e. (2008). *Autonomous fruit picking machine: a robotic apple harvester*. Paper presented at the field and service robotics.
- Bai, x. D., cao, z. G., wang, y., yu, z. H., zhang, x. F., & li, c. N. (2013). Crop segmentation from images by morphology modeling in the cie l*a*b* color space. *Computers and electronics in agriculture*, 99, 21-34. Doi: <http://dx.doi.org/10.1016/j.compag.2013.08.022>
- Batchelor, b. G. (2012). *Machine vision handbook*: springer london.
- Boissard, p., martin, v., & moisan, s. (2008). A cognitive vision approach to early pest detection in greenhouse crops. *Computers and electronics in agriculture*, 62(2), 81-93.
- Medeiros, g. A., arruda, f. B., sakai, e. L., & fujiwara, m. (2001). The influence of crop canopy on evapotranspiration and crop coefficient of beans (*phaseolus vulgaris* l.). *Agricultural water management*, 49(3), 211-224. Doi: [http://dx.doi.org/10.1016/s0378-3774\(00\)00150-5](http://dx.doi.org/10.1016/s0378-3774(00)00150-5)
- Edan, y., rogozin, d., flash, t., & miles, g. E. (2000). Robotic melon harvesting. *Ieee transactions on robotics and automation*, 16(6), 831-835. Doi: 10.1109/70.897793
- Ehret, d. L., lau, a., bittman, s., lin, w., & shelford, t. (2001). Automated monitoring of greenhouse crops. *Agronomie*, 21(4), 403-414.
- Escarabajal, d., molina, j. M., fernández, d. G., cavas, f., & garcía-mateos, g. (2015a). Digital photography applied to irrigation management of little gem lettuce. *Agricultural water management*, 151, 148-157. Doi: <http://dx.doi.org/10.1016/j.agwat.2014.08.009>

- Escarabajal, d., molina, j. M., fernández, d. G., & garcía-mateos, g. (2015b). Methodology for obtaining prediction models of the root depth of lettuce for its application in irrigation automation. *Agricultural water management*, 151, 167-173. Doi: <http://dx.doi.org/10.1016/j.agwat.2014.10.012>
- Feng, j., liu, g., si, y., wang, s., & zhou, w. (2013). Construction of laser vision system for apple harvesting robot. *Transactions of the chinese society of agricultural engineering*, 29(1), 32-37.
- Feng, j., zeng, l., liu, g., & si, y. (2014). Fruit recognition algorithm based on multi-source images fusion. *Trans. Csam*, 45, 73-80.
- Fernández-pacheco, d. G., escarabajal-henarejos, d., ruiz-canales, a., conesa, j., & molina-martínez, j. M. (2014). A digital image-processing-based method for determining the crop coefficient of lettuce crops in the southeast of spain. *Biosystems engineering*, 117, 23-34.
- Fernández, r., montes, h., salinas, c., sarria, j., & armada, m. (2013). Combination of rgb and multispectral imagery for discrimination of cabernet sauvignon grapevine elements. *Sensors*, 13(6), 7838-7859. Doi: 10.3390/s130607838
- García-mateos, g., hernández-hernández, j. L., escarabajal-henarejos, d., jaén-terrones, s., & molina-martínez, j. M. (2015). Study and comparison of color models for automatic image analysis in irrigation management applications. *Agricultural water management*, 151, 158-166. Doi: <http://dx.doi.org/10.1016/j.agwat.2014.08.010>
- Gatica, g., best, s., ceroni, j., & lefranc, g. (2013). Olive fruits recognition using neural networks. *Procedia computer science*, 17, 412-419. Doi: <http://dx.doi.org/10.1016/j.procs.2013.05.053>
- Goncalves, p., & torres, p. (2010). Learning approaches to visual control of robotic manipulators: citeseer.
- Guijarro, m., pajares, g., riomoros, i., herrera, p. J., burgos-artizzu, x. P., & ribeiro, a. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and electronics in agriculture*, 75(1), 75-83. Doi: <http://dx.doi.org/10.1016/j.compag.2010.09.013>

- Hague, t., tillett, n., & wheeler, h. (2006). Automated crop and weed monitoring in widely spaced cereals. *Precision agriculture*, 7(1), 21-32.
- Hannan, m., burks, t., & bulanon, d. M. (2009). A machine vision algorithm combining adaptive segmentation and shape analysis for orange fruit detection. *Agricultural engineering international: cigr journal*.
- Hanson, b. R., & may, d. M. (2006). Crop coefficients for drip-irrigated processing tomato. *Agricultural water management*, 81(3), 381-399. Doi: <http://dx.doi.org/10.1016/j.agwat.2005.04.007>
- Hashimoto, k. (2003). A review on vision-based control of robot manipulators. *Advanced robotics*, 17(10), 969-991.
- Hetzroni, a., miles, g., engel, b., hammer, p., & latin, r. (1994). Machine vision monitoring of plant health. *Advances in space research*, 14(11), 203-212.
- Jidong, l., de-an, z., wei, j., yu, c., & ying, z. (2012, 18-20 may 2012). *Research on trunk and branch recognition method of apple harvesting robot*. Paper presented at the measurement, information and control (mic), 2012 international conference on.
- Jiménez, a. R., ceres, r., & pons, j. (2000a). A survey of computer vision methods for locating fruit on trees. *Transactions of the asae*, 43(6), 1911.
- Jiménez, a. R., ceres, r., & pons, j. (2000b). A vision system based on a laser range-finder applied to robotic fruit harvesting. *Machine vision and applications*, 11(6), 321-329.
- Kane, k. E., & lee, w. S. (2007). *Multispectral imaging for in-field green citrus identification*. Paper presented at the 2007 asae annual meeting.
- Kataoka, t., kaneko, t., okamoto, h., & hata, s. (2003, 20-24 july 2003). *Crop growth estimation system using machine vision*. Paper presented at the advanced intelligent mechatronics, 2003. Aim 2003. Proceedings. 2003 ieee/asme international conference on.
- Kelman, e., & linker, r. (2014). Vision-based localisation of mature apples in tree images using convexity. *Biosystems engineering*, 118, 174-185. Doi: <http://dx.doi.org/10.1016/j.biosystemseng.2013.11.007>

- Kim, c., you, b.-j., jeong, m.-h., & kim, h. (2008). Color segmentation robust to brightness variations by using b-spline curve modeling. *Pattern recognition*, 41(1), 22-37. Doi: <http://dx.doi.org/10.1016/j.patcog.2007.04.011>
- Kim, y.-h. (2011). Quantization-aware sensor selection for source localization in sensor networks. *Journal of information and communication convergence engineering*, 9(2), 155-160.
- Kurtulmus, f., lee, w. S., & vardar, a. (2011). Green citrus detection using ‘eigenfruit’, color and circular gabor texture features under natural outdoor conditions. *Computers and electronics in agriculture*, 78(2), 140-149. Doi: <http://dx.doi.org/10.1016/j.compag.2011.07.001>
- Li, c., teng, g., zhao, c., qiao, x., & wu, c. (2003). Development of non-contact measurement on plant growth in greenhouse using computer vision. *Transactions of the csae*, 19(3), 140-143.
- Li, l., zhang, q., & huang, d. (2014). A review of imaging techniques for plant phenotyping. *Sensors*, 14(11), 20078.
- Li, y., xia, c., & lee, j. (2015). Detection of small-sized insect pest in greenhouses based on multifractal analysis. *Optik-international journal for light and electron optics*, 126(19), 2138-2143.
- Liang, d., guan, q., huang, w., huang, l., & yang, g. (2013). Remote sensing inversion of leaf area index based on support vector machine regression in winter wheat. *Transactions of the chinese society of agricultural engineering*, 29(7), 117-123.
- Lin, h., liang, l., zhang, l., & du, p. (2013a). Wheat leaf area index inversion with hyperspectral remote sensing based on support vector regression algorithm. *Transactions of the chinese society of agricultural engineering*, 29(11), 139-146.
- Ling, p., giacomelli, g. A., & russell, t. (1996). Monitoring of plant development in controlled environment with machine vision. *Advances in space research*, 18(4), 101-112.
- Liu, j., li, p., & li, z. (2007, 18-21 aug. 2007). *A multi-sensory end-effector for spherical fruit harvesting robot*. Paper presented at the 2007 ieee international conference on automation and logistics.

- López-urrea, r., martín de santa olalla, f., montoro, a., & lópez-fuster, p. (2009). Single and dual crop coefficients and water requirements for onion (*allium cepa* l.) Under semiarid conditions. *Agricultural water management*, 96(6), 1031-1036. Doi: <http://dx.doi.org/10.1016/j.agwat.2009.02.004>
- Lu, h., cao, z., xiao, y., li, y., & zhu, y. (2016). Region-based colour modelling for joint crop and maize tassel segmentation. *Biosystems engineering*, 147, 139-150.
- Lu, j., sang, n., hu, y., & fu, h. (2014). Detecting citrus fruits with highlight on tree based on fusion of multi-map. *Optik - international journal for light and electron optics*, 125(8), 1903-1907. Doi: <http://dx.doi.org/10.1016/j.ijleo.2013.04.135>
- Łuszczkiewicz-piątek, m. (2014). Which color space should be chosen for robust color image retrieval based on mixture modeling *image processing and communications challenges* 5 (pp. 55-64): springer.
- Mahlein, a. K., rumpf, t., welke, p., dehne, h. W., plümer, l., steiner, u., & oerke, e. C. (2013). Development of spectral indices for detecting and identifying plant diseases. *Remote sensing of environment*, 128, 21-30. Doi: <http://dx.doi.org/10.1016/j.rse.2012.09.019>
- Mao, h., gao, h., zhang, x., & kumi, f. (2015). Nondestructive measurement of total nitrogen in lettuce by integrating spectroscopy and computer vision. *Scientia horticulturae*, 184, 1-7.
- Mao, w., ji, b., zhan, j., zhang, x., & hu, x. (2009, 17-19 oct. 2009). *Apple location method for the apple harvesting robot*. Paper presented at the image and signal processing, 2009. Cisp '09. 2nd international congress on.
- Martin, v., y thonnat, m. (2007, 29-31 oct. 2007). *A cognitive vision approach for image segmentation thresholding images of historical documents with back-to-front interference*. Paper presented at the 19th ieee international conference on tools with artificial intelligence(ictai 2007).
- Neto, j. (2004). A combined statistical-soft computing approach for classification and mapping weed species in minimum-tillage systems.
- Noble, s. D., & li, d. (2012). *Segmentation of greenhouse cucumber plants in multi-spectral imagery*. Paper presented at the international conference of agricultural engineering, cigr-ageng, valencia, spain.

- Noordam, j., hemming, j., van heerde, c., golbach, f., van soest, r., & wekking, e. (2004). *Automated rose cutting in greenhouses with 3d vision and robotics: analysis of 3d vision techniques for stem detection*. Paper presented at the international conference on sustainable greenhouse systems-greensys2004 691.
- Okamoto, h., & lee, w. S. (2009). Green citrus detection using hyperspectral imaging. *Computers and electronics in agriculture*, 66(2), 201-208. Doi: <http://dx.doi.org/10.1016/j.compag.2009.02.004>
- Otsu, n. (1975). A threshold selection method from gray-level histograms. *Automatica*, 11(285-296), 23-27.
- Parrish, e., & goksel, a. (1977). Pictorial pattern recognition applied to fruit harvesting. *Transactions of the asae*, 20(5), 822-0827.
- Pastrana, j. C., & rath, t. (2013). Novel image processing approach for solving the overlapping problem in agriculture. *Biosystems engineering*, 115(1), 106-115. Doi: <http://dx.doi.org/10.1016/j.biosystemseng.2012.12.006>
- Perez, a., lopez, f., benlloch, j., & christensen, s. (2000). Colour and shape analysis techniques for weed detection in cereal fields. *Computers and electronics in agriculture*, 25(3), 197-212.
- Persson, m., & åstrand, b. (2008). Classification of crops and weeds extracted by active shape models. *Biosystems engineering*, 100(4), 484-497. Doi: <http://dx.doi.org/10.1016/j.biosystemseng.2008.05.003>
- Phung, s. L., bouzerdoun, a., & chai, d. (2005). Skin segmentation using color pixel classification: analysis and comparison. *Ieee transactions on pattern analysis and machine intelligence*, 27(1), 148-154. Doi: 10.1109/tpami.2005.17
- Rakun, j., stajanko, d., & zazula, d. (2011). Detecting fruits in natural scenes by using spatial-frequency based texture analysis and multiview geometry. *Computers and electronics in agriculture*, 76(1), 80-88. Doi: <http://dx.doi.org/10.1016/j.compag.2011.01.007>
- Reis, m. J. C. S., morais, r., peres, e., pereira, c., contente, o., soares, s., . . . Bulas cruz, j. (2012). Automatic detection of bunches of grapes in natural environment from color images. *Journal of applied logic*, 10(4), 285-290. Doi: <http://dx.doi.org/10.1016/j.jal.2012.07.004>

- Safren, o., alchanatis, v., ostrovsky, v., & levi, o. (2007). Detection of green apples in hyperspectral images of apple-tree foliage using machine vision. *Transactions of the asabe*, 50(6), 2303-2313.
- Sekii, t. (2016). *Robust, real-time 3d tracking of multiple objects with similar appearances*. Paper presented at the proceedings of the iee conference on computer vision and pattern recognition.
- Shih, p., & liu, c. (2005). Comparative assessment of content-based face image retrieval in different color spaces. *International journal of pattern recognition and artificial intelligence*, 19(07), 873-893. Doi: doi:10.1142/s0218001405004381
- Shirai, y., & inoue, h. (1973). Guiding a robot by visual feedback in assembling tasks. *Pattern recognition*, 5(2), 99-108. Doi: [http://dx.doi.org/10.1016/0031-3203\(73\)90015-0](http://dx.doi.org/10.1016/0031-3203(73)90015-0)
- Si, y., liu, g., & feng, j. (2015). Location of apples in trees using stereoscopic vision. *Computers and electronics in agriculture*, 112, 68-74. Doi: <http://dx.doi.org/10.1016/j.compag.2015.01.010>
- Slaughter, d. C., & harrell, r. C. (1989). Discriminating fruit for robotic harvest using color in natural outdoor scenes. *Transactions of the asae*, 32(2), 757-0763.
- Solis-sánchez, l., garcía-escalante, j., castañeda-miranda, r., torres-pacheco, i., & guevara-gonzález, r. (2009). Machine vision algorithm for whiteflies (*bemisia tabaci* genn.) Scouting under greenhouse environment. *Journal of applied entomology*, 133(7), 546-552.
- Song, y., glasbey, c. A., horgan, g. W., polder, g., dieleman, j. A., & van der heijden, g. W. A. M. (2014). Automatic fruit recognition and counting from multiple images. *Biosystems engineering*, 118, 203-215. Doi: <http://dx.doi.org/10.1016/j.biosystemseng.2013.12.008>
- Stajanko, d., lakota, m., & hočevan, m. (2004). Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging. *Computers and electronics in agriculture*, 42(1), 31-42. Doi: [http://dx.doi.org/10.1016/s0168-1699\(03\)00086-3](http://dx.doi.org/10.1016/s0168-1699(03)00086-3)

- Story, d., kacira, m., kubota, c., akoglu, a., & an, l. (2010). Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments. *Computers and electronics in agriculture*, 74(2), 238-243.
- Sun, g., li, y., zhang, y., wang, x., chen, m., li, x., & yan, t. (2016). Nondestructive measurement method for greenhouse cucumber parameters based on machine vision. *Engineering in agriculture, environment and food*, 9(1), 70-78.
- Szeliski, r. (2010). *Computer vision: algorithms and applications*: springer science & business media.
- Terrillon, j.-c., david, m., y akamatsu, s. (1998). *Detection of human faces in complex scene images by use of a skin color model and of invariant fourier-mellin moments*. Paper presented at the pattern recognition, 1998. Proceedings. Fourteenth international conference on.
- Tian, l. F., y slaughter, d. C. (1998). Environmentally adaptive segmentation algorithm for outdoor image segmentation. *Computers and electronics in agriculture*, 21(3), 153-168. Doi: [http://dx.doi.org/10.1016/s0168-1699\(98\)00037-4](http://dx.doi.org/10.1016/s0168-1699(98)00037-4)
- Vadivambal, r., y jayas, d. S. (2011). Applications of thermal imaging in agriculture and food industry—a review. [journal article]. *Food and bioprocess technology*, 4(2), 186-199. Doi: 10.1007/s11947-010-0333-5
- Van henten, e. J., hemming, j., van tuijl, b., kornet, j., meuleman, j., bontsema, j., & van os, e. (2002). An autonomous robot for harvesting cucumbers in greenhouses. *Autonomous robots*, 13(3), 241-258.
- Van henten, e. J., van tuijl, b. A. J., hoogakker, g. J., van der weerd, m. J., hemming, j., kornet, j. G., & bontsema, j. (2006). An autonomous robot for de-leafing cucumber plants grown in a high-wire cultivation system. *Biosystems engineering*, 94(3), 317-323. Doi: <http://dx.doi.org/10.1016/j.biosystemseng.2006.03.005>
- Velázquez-lópez, n., sasaki, y., nakano, k., mejía-muñoz, j. M., & romanchik kriuchkova, e. (2011). Detección de cenicilla en rosa usando procesamiento de imágenes por computadora. *Revista chapingo. Serie horticultura*, 17, 151-160.
- Vélez, j., moreno, f., sánchez, á., & esteban, j. L. (2003). Visión por computador. *Dykinson, madrid*.

- Wang, h., li, j., wang, l., guan, h., & geng, z. (2014, 13-18 july 2014). *Automated mosaicking of uav images based on sfm method*. Paper presented at the 2014 ieee geoscience and remote sensing symposium.
- Wei, l., huang, q., ceylan, d., vouga, e., & li, h. (2015). Dense human body correspondences using convolutional networks. *Arxiv preprint arxiv:1511.05904*.
- Wei, s.-e., ramakrishna, v., kanade, t., & sheikh, y. (2016). Convolutional pose machines. *Arxiv preprint arxiv:1602.00134*.
- Wei, x., jia, k., lan, j., li, y., zeng, y., & wang, c. (2014). Automatic method of fruit object extraction under complex agricultural background for vision system of fruit picking robot. *Optik - international journal for light and electron optics*, 125(19), 5684-5689. Doi: <http://dx.doi.org/10.1016/j.ijleo.2014.07.001>
- Weilong, d., peiliang, m., & zhijun, c. (2008). Visual modeling and simulation of plant growth based on plant functional-structural mutual feedback mechanism. *Transactions of the chinese society of agricultural engineering*, 2008(11).
- Woebbecke, d., meyer, g., von bargaen, k., & mortensen, d. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the asae*, 38(1), 259-269.
- Wspanialy, p., & moussa, m. (2016). Early powdery mildew detection system for application in greenhouse automation. *Computers and electronics in agriculture*, 127, 487-494. Doi: <http://dx.doi.org/10.1016/j.compag.2016.06.027>
- Xia, c., chon, t.-s., ren, z., & lee, j.-m. (2015). Automatic identification and counting of small size pests in greenhouse conditions with low computational cost. *Ecological informatics*, 29, 139-146.
- Xia, c., lee, j.-m., li, y., chung, b.-k., & chon, t.-s. (2012). In situ detection of small-size insect pests sampled on traps using multifractal analysis. *Optical engineering*, 51(2), 027001-027001-027001-027012.
- Xiang, r., jiang, h., & ying, y. (2014). Recognition of clustered tomatoes based on binocular stereo vision. *Computers and electronics in agriculture*, 106, 75-90. Doi: <http://dx.doi.org/10.1016/j.compag.2014.05.006>

- Xie, z., ji, c., guo, x., & ren, s. (2010). An object detection method for quasi-circular fruits based on improved hough transform. *Transactions of the chinese society of agricultural engineering*, 26(7), 157-162.
- Xie, z., ji, c., guo, x., & zhu, s. (2011). Detection and location algorithm for overlapped fruits based on concave spots searching. [article]. *Nongye jixie xuebao/transactions of the chinese society of agricultural machinery*, 42(12), 191-196.
- Xie, z., zhang, t., & zhao, j. (2007). Ripened strawberry recognition based on hough transform. *Transactions of the chinese society for agricultural machinery*, 38(3), 106-109.
- Xiong, j., zou, x., wang, h., peng, h., zhu, m., & lin, g. (2013). Recognition of ripe litchi in different illumination conditions based on retinex image enhancement. *Transactions of the chinese society of agricultural engineering*, 29(12), 170-178.
- Xu, h., & ying, y. (2004). *Detecting citrus in a tree canopy using infrared thermal imaging*. Paper presented at the optical technologies for industrial, environmental, and biological sensing.
- Yang, l., dickinson, j., wu, q. M. J., & lang, s. (2007, 4-6 dec. 2007). *A fruit recognition method for automatic harvesting*. Paper presented at the mechatronics and machine vision in practice, 2007. M2vip 2007. 14th international conference on.
- Yin, h., chai, y., yang, s. X., & mittal, g. S. (2009, 4-7 dec. 2009). *Ripe tomato recognition and localization for a tomato harvesting robotic system*. Paper presented at the soft computing and pattern recognition, 2009. Socpar '09. International conference of.
- Yuan, t., ji, c., chen, y., li, w., & zhang, j. (2011). Greenhouse cucumber recognition based on spectral imaging technology. [article]. *Nongye jixie xuebao/transactions of the chinese society of agricultural machinery*, 42(suppl.), 172-176.
- Zhan, w., he, d., & shi, s. (2013). Recognition of kiwifruit in field based on adaboost algorithm. [article]. *Nongye gongcheng xuebao/transactions of the chinese society of agricultural engineering*, 29(23), 140-146. Doi: 10.3969/j.issn.1002-6819.2013.23.019

- Zhang, b., huang, w., wang, c., gong, l., zhao, c., liu, c., & huang, d. (2015). Computer vision recognition of stem and calyx in apples using near-infrared linear-array structured light and 3d reconstruction. *Biosystems engineering*, 139, 25-34. Doi: <http://dx.doi.org/10.1016/j.biosystemseng.2015.07.011>
- Zhang, f., y fu, l. S. (2014). *Application of computer vision technology in agricultural field*. Paper presented at the applied mechanics and materials.
- Zhang, m., & meng, q. (2011). Automatic citrus canker detection from leaf images captured in field. *Pattern recognition letters*, 32(15), 2036-2046. Doi: <http://dx.doi.org/10.1016/j.patrec.2011.08.003>
- Zhao, j., tow, j., & katupitiya, j. (2005, 2-6 aug. 2005). *On-tree fruit recognition using texture properties and color data*. Paper presented at the 2005 ieee/rsj international conference on intelligent robots and systems.
- Zhao, y., gong, l., huang, y., & liu, c. (2016a). A review of key techniques of vision-based control for harvesting robot. *Computers and electronics in agriculture*, 127, 311-323. Doi: <http://dx.doi.org/10.1016/j.compag.2016.06.022>
- Zhao, y., gong, l., zhou, b., huang, y., & liu, c. (2016b). Detecting tomatoes in greenhouse scenes by combining adaboost classifier and colour analysis. *Biosystems engineering*, 148, 127-137. Doi: <http://dx.doi.org/10.1016/j.biosystemseng.2016.05.001>
- Zhao, y., he, y., & xu, x. (2012). A novel algorithm for damage recognition on pest-infested oilseed rape leaves. *Computers and electronics in agriculture*, 89, 41-50. Doi: <http://dx.doi.org/10.1016/j.compag.2012.07.014>
- Zheng, l., zhang, j., & wang, q. (2009). Mean-shift-based color segmentation of images containing green vegetation. *Computers and electronics in agriculture*, 65(1), 93-98.

CAPÍTULO 3

SEGMENTACIÓN AUTOMÁTICA DE CULTIVOS EN INVERNADERO UTILIZANDO EL ESPACIO DE COLOR CIE L*A*B*

Abner Orlando Chávez-Sánchez^a, Luz Abril Torres-Mendez^b, Noé Velázquez-López^{c*},
Agustín Ruiz-García^c, Pedro Cruz-Meza^d

^aUniversidad Autónoma Chapingo, Posgrado en Ingeniería Agrícola y Uso Integral del Agua. Carretera
México-Texcoco km. 38.5, Chapingo, Estado de México. C.P. 56230, México.

^bCINVESTAV - Unidad Saltillo Ramos Arizpe, Coah., 25900

^cUniversidad Autónoma Chapingo, Departamento de Irrigación. Carretera México-Texcoco km. 38.5,
Chapingo, Estado de México. C.P. 56230, México.

^dUniversidad Autónoma Chapingo, Departamento de Ingeniería Mecánica Agrícola. Carretera México-
Texcoco km. 38.5, Chapingo, Estado de México. C.P. 56230, México.

3.3 Resumen

Los sistemas de visión por computadora son cada vez más usados en la agricultura de precisión. Dentro de estos, la segmentación de imágenes es un paso esencial en aplicaciones de cosecha, monitoreo y manejo. Existen diferentes técnicas para la segmentación de imágenes de cultivos en ambientes naturales a campo abierto, pero son pocos o nulos en condiciones bajo invernadero, en los cuales varía tanto la iluminación como la estructura y cubierta del invernadero. Además la incidencia de la radiación solar sobre la lente de la cámara también representa un problema. Por lo anterior, el objetivo de este trabajo fue desarrollar una técnica de segmentación automática de imágenes de tomate cultivado en invernadero mediante el espacio de color CIE L*a*b y adaptable a las condiciones de iluminación natural.

* Corresponding author. Tel.: (52)595-952-1551.

E-mail address: nvelazquez@taurus.chapingo.mx.

Se utilizaron 50 imágenes de tamaño 640×480 píxeles, las cuáles fueron tomadas en distintas etapas de crecimiento y de iluminación. Como primer paso, se analizaron los histogramas de los canales a y b para encontrar la distribución del color del cultivo y a partir de estas distribuciones se definieron los umbrales para la segmentación. El método de segmentación propuesto se evaluó haciendo uso del método de evaluación definido por Automatic Target Recognition Working Group (ATRWG) y se comparó con otros 4 enfoques. La calidad media de segmentación del método propuesto fue de 89.3%, superando por 4% o más a los otros métodos. La desviación estándar para el método propuesto fue de 4.0%, que es inferior a los obtenidos por los otros métodos. Por otra parte, se encontró que al incrementar la resolución de la imagen mejoró la calidad de segmentación en 3% y la desviación estándar disminuyó en 0.35%. Sin embargo, el tiempo computacional se incrementó ligeramente en función del tamaño de la imagen. Finalmente, se comprobó que el método propuesto es robusto a la variación de la iluminación en el interior de diferentes invernaderos y tiene una calidad mejor que los otros cuatro enfoques definidos para campo abierto. En general, el método de segmentación de cultivos propuesto se puede utilizar para la segmentación de cultivos en agricultura protegida de manera eficaz y eventualmente se podría aplicar para diversas tareas en las cuales es necesario extraer el cultivo tales como: aplicación de agroquímicos, detección de frutos, podas y otras.

Palabras clave: Visión por computadora; tomate; umbralización; iluminación natural.

3.2 Introducción

Los sistemas de visión por computadora son cada vez más usados en la agricultura de precisión con la finalidad de monitorizar características específicas de los cultivos, tales como: el crecimiento (Sakamoto *et al.*, 2012; Zhu *et al.*, 2016), la estimación del rendimiento (Gongal *et al.*, 2016; Payne *et al.*, 2013), reconocimiento de surcos para orientación de la maquinaria (Vidović *et al.*, 2016; Xue *et al.*, 2012), detección de enfermedades (Chung *et al.*, 2016; Pires *et al.*, 2016; Pourreza *et al.*, 2015), identificación de malezas para aplicación localizada de herbicidas (Arribas *et al.*, 2011; Burgos-Artizzu *et al.*, 2011; Tellaeche *et al.*, 2011), requerimiento nutricional del cultivo (Romualdo *et al.*, 2014), así como el reconocimiento de partes específicas de la planta que tras su manejo ayuden a mejorar la calidad del fruto (Kurtulmuş y Kavdir, 2014).

Actualmente, los invernaderos considerados como sistemas biofísicos con entradas, salidas y bucles de control de procesos, mejoran las condiciones de producción (Ehret *et al.*, 2001). Sin embargo, el monitoreo de factores bióticos y requerimientos nutricionales está restringido a pruebas de laboratorio que resultan muy costosas y tardadas o en su defecto dependen de la experiencia del ser humano para detectarlos.

La visión por computadora tiene la capacidad de solventar este problema, así pues se ha venido investigando en diferentes áreas en el interior de invernaderos: monitorizar el requerimiento nutricional (Mao *et al.*, 2015; Story *et al.*, 2010), identificar plagas (Li *et al.*, 2015; Xia *et al.*, 2015), detectar enfermedades (Ma *et al.*, 2015; Velázquez-López *et al.*, 2011), obtener parámetros de crecimiento del cultivo (Andújar *et al.*, 2016; Sun *et al.*, 2016), reconocer frutos para la cosecha mecanizada (Bac *et al.*, 2014; Rajendra *et al.*, 2009; Van Henten *et al.*, 2003; Xiang *et al.*, 2014), determinar requerimientos hídricos (Farkas, 2010; Mangus *et al.*, 2016) y evadir obstáculos (Nissimov *et al.*, 2015).

La segmentación de imágenes es una etapa importante y esencial para las tareas anteriores (Bai *et al.*, 2013), sin embargo, no existen métodos que tomen en cuenta factores como la variación de iluminación en el interior de invernaderos. Lo que relaciona directamente su rendimiento con la calidad de la imagen capturada. Ésta última es afectada por la variabilidad, la aleatoriedad y la complejidad de la iluminación natural (Ye *et al.*, 2015).

Así mismo, la complejidad de esta actividad se incrementa debido a la variación del color de la planta causado principalmente por las condiciones de iluminación y la etapa fenológica del cultivo. Incluso en una determinada etapa fenológica y bajo una misma condición de iluminación este varía entre sus órganos.

Se han desarrollado diversas técnicas para la segmentación de imágenes de cultivos en ambientes naturales a campo abierto con la finalidad de dar solución al problema de la iluminación, pero son pocos o nulos en condiciones bajo invernadero. Éstas técnicas hacen uso de las características de color de las plantas, por ejemplo: el índice vegetativo (VEG) diseñado para funcionar en condiciones de iluminación natural (Hague *et al.*, 2006), el índice de exceso de verde (ExG) (Woebbecke *et al.*, 1995), el índice de exceso de verde menos el exceso de rojo (ExGR) (Camargo, 2004), así como, el índice de diferencia normalizada (NDI) (Pérez *et al.*, 2000). Los algoritmos anteriores se basan en la hipótesis de que el color verde es el predominante entre los canales de la imagen en escenas agrícolas y en la combinación de estos directamente del espacio RGB (red, Green y blue). La combinación de canales provee una imagen a escala de grises y el problema se resume en encontrar un umbral para la segmentación, usando generalmente el método de Otsu (1975). Más tarde Yu *et al.* (2013) desarrollaron un método para la segmentación de cultivos que toma en cuenta la variación de la iluminación en ambientes complejos. El método se basa en suponer que en una intensidad específica la forma del histograma de tonalidad se comporta de manera similar a una distribución de Gauss. Éste método requiere de un umbral adecuado ya que afecta en gran medida la precisión de la segmentación.

Otro enfoque son los algoritmos basados en el aprendizaje. Zheng *et al.* (2009) propusieron un algoritmo de segmentación mediante una red neuronal BPNN (Back propagation Neural Network), en la que suponen que la imagen está compuesta de vegetación y no vegetación. Bai *et al.* (2013) desarrollaron un método de segmentación basado en el espacio de color CIE $L^*a^*b^*$ utilizando el método de modelación morfológica con la finalidad de establecer el modelo de color del cultivo.

Más tarde Lu *et al.* (2016) proponen un método de modelado de color formulado como un problema de segmentación semántica que se compone de dos etapas: 1) la generación de propuestas de la región, 2) La predicción del modelo de color. Este método hace una fuerte suposición que consiste en que la parte de la planta a segmentar se encuentra en una posición específica dentro de la escena. Además, Guo *et al.* (2013) propusieron un método para segmentar la vegetación en condiciones de iluminación natural que se basa en un proceso de aprendizaje de máquina mediante árboles de decisión (CART).

Todos los métodos de segmentación se basan en el uso de algún espacio de color por lo que Hernández-Hernández *et al.* (2016) desarrollaron un método de entrenamiento que selecciona automáticamente el espacio de color óptimo. Adicionalmente, realiza la combinación de canales de diferentes espacios de color para segmentar cultivos en ambientes naturales. Reporta un error de 0.3% con respecto a la segmentación manual. Los métodos desarrollados hasta ahora tienen la capacidad de adaptarse a cambios de iluminación con cierta precisión. El principal inconveniente es que necesitan una gran cantidad de muestras en la etapa de entrenamiento, además, el rendimiento del segmentado depende de que en el aprendizaje se puedan presentar todas las posibles condiciones de iluminación. Por lo tanto, el objetivo de este trabajo fue desarrollar una técnica de segmentación de imágenes de tomate cultivado en invernadero mediante el espacio de color CIE L^*a^*b y adaptable a las condiciones de iluminación natural. El método propuesto eventualmente se podría aplicar para diversas tareas en las cuales es necesario extraer el cultivo tales como: aplicación de agroquímicos, detección de frutos, podas y otras.

3.3 Materiales y métodos

3.3.1 Descripción del experimento y adquisición de imágenes

El estudio se llevó a cabo haciendo uso del cultivo de jitomate (*Solanum lycopersicum* L.) sembrado en 3 invernaderos, ubicados en los siguientes estados del país: Universidad Autónoma Chapingo, Estado de México, México (19°29'34.0"N 98°53'01.5"W), San Pedro Mártir, Oaxaca, México (16° 44' 24"N, 96° 43' 30"W) y Universidad Autónoma Agraria Antonio Narro, Coahuila, México (25°21'5"N 101°1'47"W). La segunda y tercera ubicación utilizan plástico (polietileno de baja densidad) como material de cubierta, mientras que la primera vidrio. Así mismo, el medio de producción para el primer y tercer invernadero es hidroponía, sin embargo, para el segundo es suelo.

Las imágenes fueron adquiridas utilizando una cámara web marca Logitech modelo C920 de un sistema de estereo visión (Cuadro 1). El sistema se fijó en un tripie orientando el eje óptico (ángulo pitch) de las cámaras web con una pequeña inclinación hacia el suelo y en contra del sol con la finalidad de evitar que la radiación incida directamente sobre la lente. Las cámaras web se colocaron a una distancia entre 50 a 70 cm del objetivo (Figura 1). El autofocus fue desactivado, mientras que la auto exposición se mantuvo activada.

Cuadro 1. Especificaciones técnicas de la webcam Logitech modelo C920

Especificaciones	HD pro webcam C920
Definición de videos	Full HD 1080p con H.264
Calidad de la foto	15 MP
Autofoco	Autofoco con 20 pasos
Corrección automática de iluminación	Premium



Figura 1. Equipo para la adquisición de imágenes.

Se adquirieron 216 imágenes en formato tif de tamaño 640x480 pixeles, tomadas bajo diferentes condiciones de iluminación (soleado, nublado, lluvioso). Las imágenes fueron tomadas de 9 am a 2 pm a partir del 16 de Noviembre del 2015 al 18 de Abril del 2016 en diferentes etapas del crecimiento del cultivo. Para el análisis se tomó una muestra al azar de 50 imágenes.

3.3.2 Método de segmentación propuesto

El algoritmo se desarrolló en una Laptop marca Toshiba modelo Satellite P745 con las siguientes características: procesador Intel Core i5 con 2.4 GHz y 6GB de memoria RAM. Se utilizaron las librerías de OpenCV versión 3.1.0 en el sistema operativo Ubuntu 2014 LTS. El método propuesto hace uso del espacio de color CIE $L^*a^*b^*$. En éste, L representa la luminosidad, a el canal que agrupa los colores correspondientes del verde al rojo en un rango de 0 a 255, respectivamente y b el canal que contiene los colores del amarillo al azul en la escala de 0 a 255 respectivamente. La primera etapa del estudio consistió en analizar los histogramas de los canales a y b de las imágenes de la base de datos. El histograma de una imagen a escala de grises está definido según la Ec. (1) (Cuevas *et al.*, 2010).

$$f(x) = \text{card}^2\{I(u, v) \mid I(u, v) = x\} \quad (1)$$

I es la matriz de la imagen a escala de grises, u el número de la filas de ésta, y v el número de columnas. El método de segmentación propuesto consiste en determinar el umbral para los canales a y b , que corresponden al cultivo. En el Anexo 1 se muestra el algoritmo del método propuesto.

3.3.3 Determinación del umbral para binarizar el canal a

El primer paso es transformar la imagen fuente del espacio de color RGB al espacio de color $CIE\ L^*a^*b^*$. Posteriormente se extrae el histograma del canal a y de este las frecuencias relativas diferentes de cero. Estos valores con sus correspondientes niveles de intensidad se almacenan en un vector de puntos $\mathbf{w}$. Como resultado del análisis de los histogramas del canal a se encontró que la distribución es normal del lado izquierdo (Figura 3 a); entonces la tarea principal recae en delimitar un rango en el cual el umbral óptimo para la segmentación de la planta se encuentra. El rango se delimita de la esperanza matemática (μ_x) a un valor c , que es el nivel de intensidad en el cual se presenta el argumento máximo de la primera derivada de la función definida en la ecuación (1). Por otro lado, μ_x se calcula primero obteniendo una aproximación en un rango de 0 a 127 con la ecuación (2), donde 127 es el nivel de intensidad máximo del color verde. Después a través de una técnica de refinamiento iterativo, la cual consiste en minimizar el error relativo definido por la ecuación (3).

$$\mu_x = \sum_x x f_x(x), \quad 0 \leq x \leq 127 \quad (2)$$

$$\varepsilon_t = \frac{\sum(f_i \text{ real} - f_i \text{ estimada})}{\sum f_i \text{ real}} \quad (3)$$

Para cada iteración se toma un nuevo rango del nivel de intensidades del vector $\mathbf{w}$. Dicho rango es delimitado en el extremo superior por los mínimos locales de las frecuencias del vector $\mathbf{w}$ menores o iguales a 127 y mayores que la primera aproximación de μ_x . Los niveles de intensidad en los que se presentan los mínimos locales en el rango de μ_x a 127 (Ec. 5) se almacenan en el vector $\mathbf{q}$.

$$f(x): [\mu_x, 127] \subseteq \mathbb{R} \rightarrow \mathbb{R}, \exists x_0 \in [\mu_x, 127], P = (x_0, f(x_0)) \quad (4)$$

$$P(x_0, f(x_0)) \text{ mínimo local de } f \Leftrightarrow E'(x_0) | \forall x \neq x_0 \in E'(x_0): f(x) > f(x_0) \quad (5)$$

Posteriormente, en las siguientes iteraciones para el cálculo de μ_x los rangos son asignados en dos pasos: primero se fija el valor máximo, que corresponde a cada elemento del vector $\mathbf{q}$ y para cada componente de este se asignan como valor inferior del rango cada elemento que corresponde al nivel de intensidad del vector $\mathbf{w}$, siempre y cuando éste sea al menos dos unidades menor al valor de la primera aproximación de μ_x . Posteriormente c es calculado (Ec. 6), este valor representa la separación más evidente entre el color verde y rojo del canal.

$$f'(x): [\mu_x, 255] \subseteq \mathbb{R} \rightarrow \mathbb{R}, \exists! c | \forall x \in [\mu_x, 255]: f'(c) > f'(x), x \neq c \quad (6)$$

Donde $f'(x)$ es la primera derivada de las frecuencias de intensidad del vector $\mathbf{w}$ calculada mediante diferencias finitas divididas hacia atrás (Chapra *et al.*, 2007) y c es el nivel de intensidad en el cual se presenta el argumento máximo de la primera derivada. En la siguiente etapa se obtienen los niveles de intensidad que corresponden a los mínimos locales de las frecuencias del vector $\mathbf{w}$ en el rango definido por μ_x y c , estos son almacenados en el vector $\mathbf{v}$. Donde el umbral deseado (*tha*) es el argumento máximo del vector $\mathbf{v}$, que puede escribirse como:

$$tha = \begin{cases} \operatorname{argmax} \mathbf{v} & ; \mu_x < v_i < c \\ c, & \text{de otra forma} \end{cases} \quad (7)$$

3.3.4 Determinación del umbral para binarizar el canal $\mathbf{b}$

De la imagen transformada al espacio de color $CIE L^*a^*b^*$, se extrae el histograma del canal $\mathbf{b}$ (Ec. 1) y de este las frecuencias relativas diferentes de cero. Estos valores con sus correspondientes niveles de intensidad se almacenan en un vector de puntos $\mathbf{wb}$. Como resultado del análisis de los histogramas del canal $\mathbf{b}$ se encontró que la distribución es normal del lado derecho (Figura 3 a); entonces al igual que en el caso del canal $\mathbf{a}$ la tarea principal recae en delimitar un rango en el cual el umbral óptimo se encuentra, éste se delimita del valor cb (Ec. 11) a la esperanza matemática de la distribución normal (μ_x). El valor cb es el nivel de intensidades que corresponde al argumento mínimo de la primera derivada de las frecuencias del vector $\mathbf{wb}$, mientras que μ_x es calculada primero obteniendo una aproximación en un rango de 129 a 255 con la Ec. (8). Después a través de una técnica de refinamiento iterativo, que consiste en minimizar el error relativo definido en la Ec. (3).

Para cada iteración se toma un nuevo rango del nivel de intensidades del vector **wb**. Dicho rango es delimitado en el extremo inferior de la distribución por los mínimos locales de las frecuencias del vector **wb** mayores que 128 y menores que la primera aproximación de μ_x . Los niveles de intensidad en los que se presentan los mínimos locales en el rango de 128 a μ_x (Ec. 10) son almacenados en un vector **qb**.

$$\mu_x = \sum_x x f_x(x), 129 < x < 255 \quad (8)$$

$$f(x): [128, \mu_x] \subseteq \mathbb{R} \rightarrow \mathbb{R}, \exists x_0 \in [128, \mu_x], P = (x_0, f(x_0)) \quad (9)$$

$$P(x_0, f(x_0)) \text{ mínimo local de } f \Leftrightarrow E'(x_0) | \forall x \neq x_0 \in E'(x_0): f(x) > f(x_0) \quad (10)$$

En las iteraciones posteriores para el cálculo de μ_x los rangos son asignados en dos pasos, el primero de ellos consiste en fijar el valor mínimo para ello se toma cada uno de los elementos del vector **qb** como valor inferior. Y para cada elemento del vector **qb** se asignan como valor superior del rango cada uno de los elementos correspondientes al nivel de intensidad del vector **w**, siempre y cuando éste sea al menos cuatro unidades mayor al valor de la primera aproximación de μ_x . Posteriormente, el nivel de intensidades **cb** es calculado (Ec. 11), este valor representa la separación más evidente entre el color azul y amarillo del canal.

$$f'(x): [0, \mu_x] \subseteq \mathbb{R} \rightarrow \mathbb{R}, \exists! cb | \forall x \in [0, \mu_x]: f'(cb) > f'(x), x \neq c \quad (11)$$

Donde $f'(x)$ es la primera derivada de las frecuencias de intensidades del vector **wb** calculada también por diferencias finitas divididas hacia atrás y **cb** es el nivel de intensidades que corresponde al argumento mínimo de la primera derivada de las frecuencias del vector **wb**. En la siguiente etapa se obtienen los niveles de intensidad que corresponden a los mínimos locales de las frecuencias del vector **wb** en el rango **cb** a μ_x , estos son almacenados en un vector **vb**. Donde el umbral deseado (**thb**) es el argumento mínimo del vector **vb**, que puede escribirse como:

$$thb = \begin{cases} \text{argmin } vb & ; \text{ } cb < vb_i < \mu_x \\ cb; & \text{de otra forma} \end{cases} \quad (12)$$

3.3.5 Generación de la máscara de segmentación.

Los umbrales obtenidos son usados para binarizar los canales correspondientes de la imagen. Hasta este punto se cuenta con dos imágenes binarias, una del canal a y otra del b . Mediante la conjunción de píxel por píxel de estas dos imágenes se obtiene la máscara de segmentación. Finalmente, esta es usada para segmentar la imagen original.

3.3.6 Evaluación de la calidad de segmentación

La calidad de segmentación se evaluó mediante dos enfoques, uno de ellos es el propuesto por Automatic Target Recognition Working Group (ATRWG) definido como:

$$Calidad = \left| \frac{A \cap B}{A \cup B} \right| \quad (13)$$

Donde A es el conjunto de píxeles que corresponden al cultivo obtenido mediante la segmentación manual y B es el conjunto de píxeles correspondientes al cultivo como resultado de la segmentación automática por los diferentes métodos incluyendo el propuesto. $||$ denota la cardinalidad del conjunto, la calidad igual a uno representa una segmentación perfecta. El segundo enfoque es el propuesto por Xiao *et al.* (2011), este puede escribirse como:

$$Calidad = \frac{|F_0 \cap F_R| + |B_0 \cap B_R|}{|F_0| + |B_0|} \quad (14)$$

Donde F_0 y B_0 representan el cultivo y el fondo respectivamente, segmentados manualmente. F_R y B_R son el resultado de la segmentación automática por los diferentes métodos. $|\cdot|$ Denota la cardinalidad del conjunto

3.3.7 Evaluación del método propuesto en comparación con otros enfoques y con imágenes de diferente resolución.

Para validar que el método de segmentación de imágenes de cultivos propuesto sea robusto a las variaciones de iluminación dentro del invernadero, se aplicó a la base de datos de imágenes recolectadas. Posteriormente el método propuesto se comparó con cuatro enfoques de segmentación de cultivos: ExG, ExGR, NDI y el índice de extracción de color de la vegetación (CIVE) (Kataoka *et al.*, 2003).

Todos estos enfoques tienen que fijar un umbral para la segmentación final, es decir, para discriminar entre las plantas y el suelo, para ello utilizan el método de Otsu (1975). Finalmente, el rendimiento del método propuesto se comparó con imágenes de diferente resolución. Incluyendo imágenes de tamaño 640x480 y 2112 x 2816 píxeles.

3.4 Resultados y discusión

3.4.1 Análisis de los histogramas

Experimentalmente se observó que el color del cultivo en el histograma de los canales *a* y *b* tienen una distribución normal. Ésta se distribuye del lado izquierdo para el primer caso (*Figura , a*) y del lado derecho para el segundo (*Figura , b*). Una hipótesis similar plantean en Yu *et al.* (2013). Este es el punto de partida para la determinación del umbral óptimo.

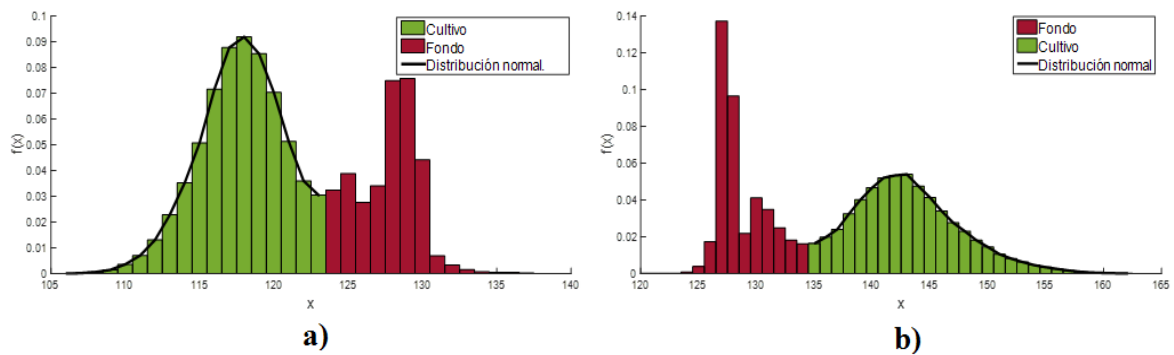


Figura 2. a) Distribución del color del cultivo en el canal a. b) Distribución del color del cultivo en el canal b.

3.4.2 Evaluación del método propuesto en comparación con otros enfoques

El resultado de la segmentación manual de la base de datos se muestra en la *Figura* .

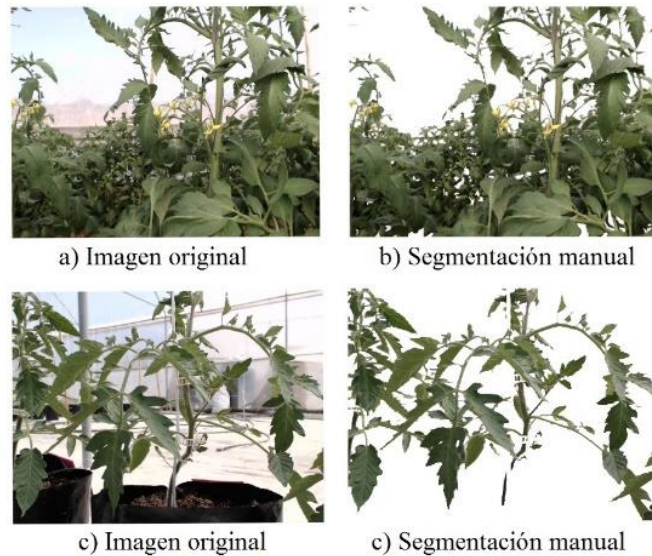


Figura 3. Imágenes del cultivo de jitomate y su segmentación manual.

Así mismo en la *Figura* se presenta un ejemplo del resultado obtenido de la segmentación por el método propuesto. Como se puede observar los métodos basados en ExG (*Figura* c), ExGR (*Figura* d) y NDI (*Figura* f) pierden píxeles correspondientes al cultivo. El índice ExG (*Figura* c) además de perder píxeles correspondientes al cultivo clasifica parte del fondo como cultivo (tonos azules principalmente), mientras que el CIVE (*Figura* e) a pesar de segmentar correctamente el cultivo es susceptible a tomar parte del fondo. El resultado por el método propuesto (*Figura* g) es el más cercano a la segmentación realizada manualmente. La *Figura* h) es el resultado de la segmentación por el método propuesto en donde se puede apreciar lo declarado anteriormente.

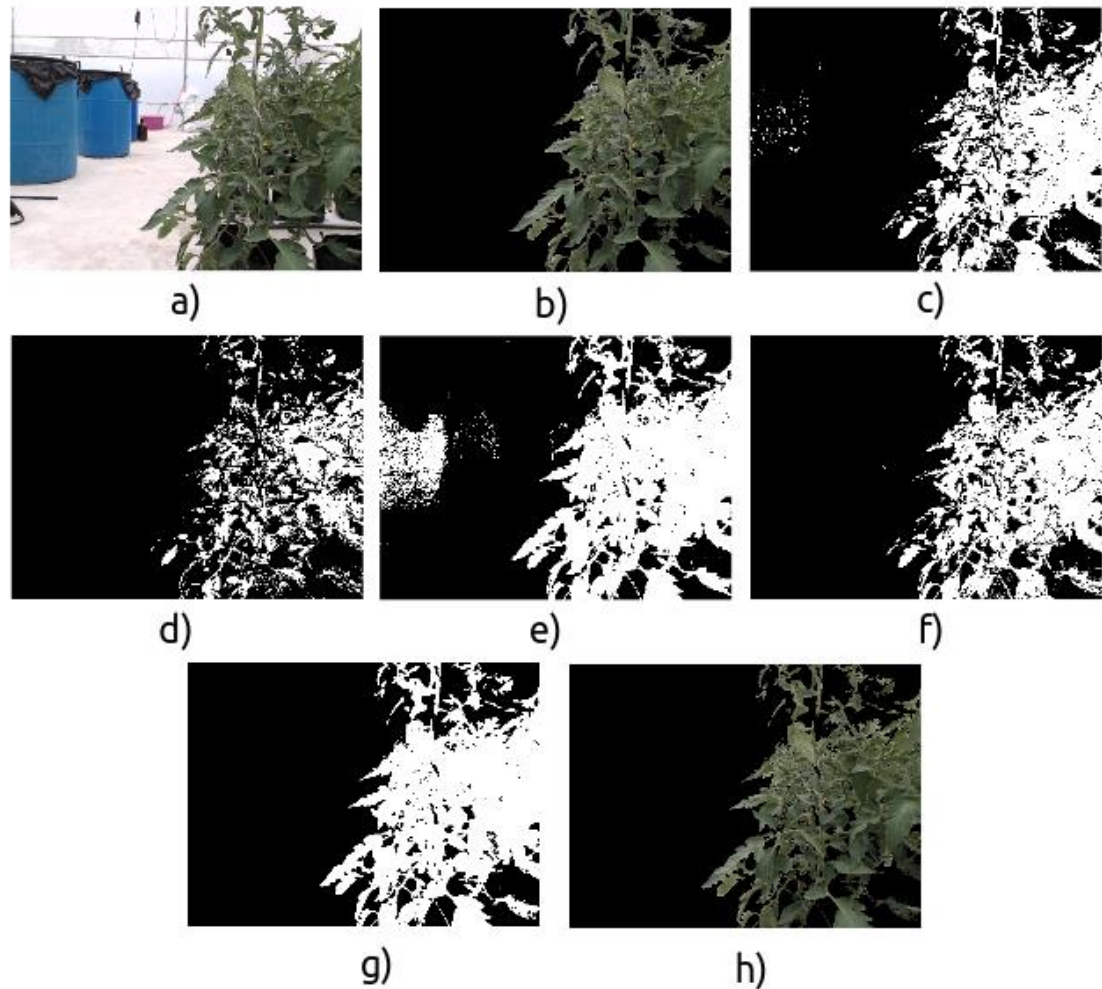


Figura 4. Ejemplo de resultados de la segmentación. (a) Imagen original; (b) Resultado de la segmentación manual; (c-f) Resultados de la segmentación de ExG, ExGR, CIVE, NDI, respectivamente; (g) Resultado del método propuesto; (h) Mascara de segmentación dado por g).

La media de la calidad de segmentación del método propuesto y los otros cuatro enfoques obtenida por el método de evaluación propuesto por ATRWG se presentan en el *Cuadro 2*. Posteriormente se grafican y se presentan en la Figura 6.

Cuadro 2. Evaluación de la media de calidad de segmentación por el método ATRWG para el enfoque propuesto y cuatro más.

	Método				
	propuesto	ExG	ExGR	CIVE	NDI
$\mu(\%)$	89.341	58.569	39.131	85.195	54.652
$\sigma(\%)$	4.044	12.502	13.522	7.586	15.583

Los datos del *Cuadro 2* muestran que la calidad media de segmentación del método propuesto supera a los otros métodos cuando menos por el 4%, además la desviación estándar es menor situándose por debajo del índice CIVE por más del 3%. Esto puede explicarse principalmente porque los 4 enfoques de comparación fueron desarrollados para segmentar imágenes a campo abierto, sin embargo, al no haber algoritmos desarrollados para invernadero, éstos se analizaron para verificar su precisión en condiciones bajo invernadero.

De acuerdo con los experimentos los enfoques ExG, ExGR y NDI suelen proporcionar una calidad de segmentación insuficiente en ambientes naturales bajo invernadero, ya que son más susceptibles a cambios de iluminación. Estos mostraron una diferencia en la media de calidad con respecto al método propuesto de 30.8%, 50.1%, 4.0% y 34.7% respectivamente. En cuanto al CIVE, aunque este difiere por menos del 5% en la media de calidad de segmentación, como se muestra en la *Figura (e)* es susceptible a tomar píxeles del tono azul como parte del cultivo. Por lo tanto, con una base de datos de 50 imágenes tomadas bajo diferentes condiciones de iluminación en 3 invernaderos con diferentes condiciones estructurales, se muestra que el método propuesto es robusto a los cambios de iluminación.

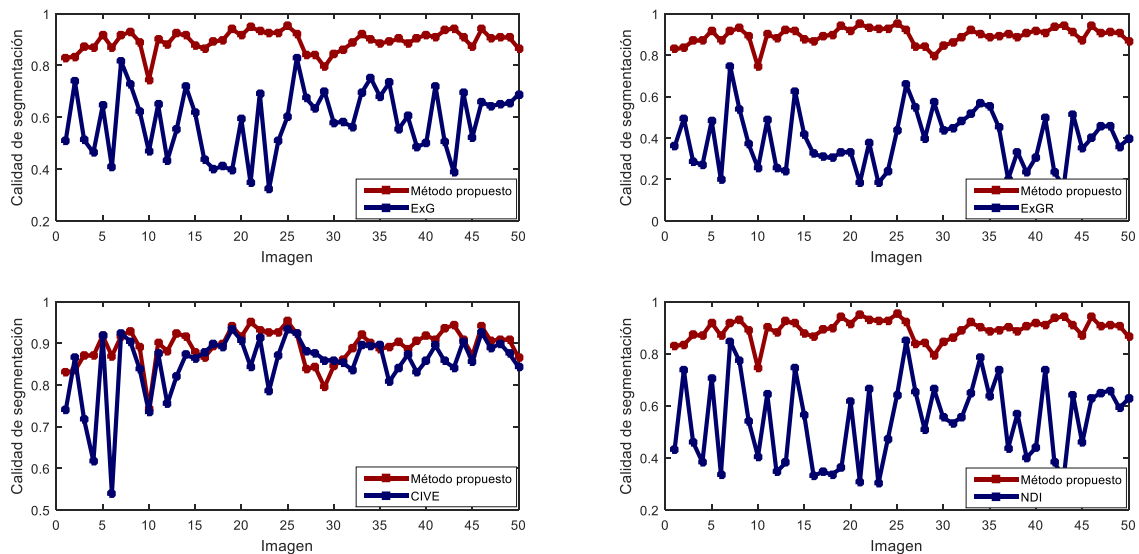


Figura 5. Comparación de la calidad de segmentación del método propuesto con cuatro enfoques, utilizando el índice propuesto por ATRWG.

Los buenos resultados del método propuesto se deben a que el modelo de color del cultivo obtenido del análisis de los histogramas es invariante ante los cambios de iluminación. El método se enfoca principalmente en la distribución de colores en el canal *a* y *b* para extraer tonos que corresponden a las plantas, esto permite aplicarlo en invernaderos con diferentes condiciones (cubierta, medio de producción e iluminación). Así mismo, el modelo de color del cultivo en los canales *a* y *b* es apoyado en gran medida por la metodología de adquisición de imágenes.

Con la finalidad de presentar los resultados de la calidad de segmentación del método propuesto bajo diferentes condiciones de iluminación se eligieron imágenes representativas de tres invernaderos estructuralmente diferentes, los resultados se presentan en la *Figura* . La *Figura* (a) corresponde a un invernadero de cubierta plástica con la superficie cubierta por una malla blanca y un sistema de producción agrícola en hidroponía. La *Figura* (b) corresponde a un invernadero de cubierta de vidrio, la superficie es de concreto y el sistema de producción es hidroponía. La *Figura* (c) fue tomada en un invernadero de cubierta plástica donde la superficie del suelo se encuentra descubierta y la producción es en suelo. Como se muestra en la *Figura* el método propuesto de segmentación de imágenes de cultivos en ambientes naturales bajo invernadero es robusto a las variaciones de iluminación y a las condiciones de producción.

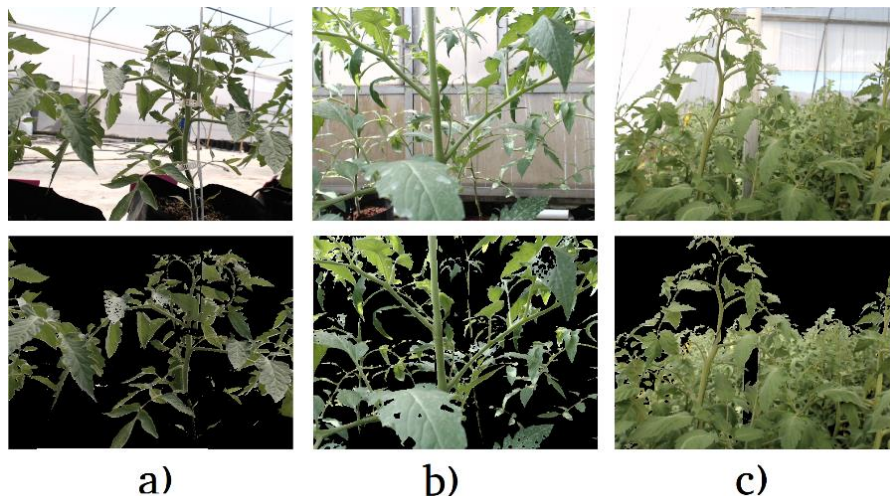


Figura 6. Resultados de la segmentación bajo invernaderos estructuralmente diferentes

Por otra parte, se compararon los resultados del *Cuadro 2* con los obtenidos al evaluar el índice de calidad descrito en la Ec. (14) (*Cuadro 3*).

Cuadro 3. Resultados de la calidad de segmentación obtenida con el nuevo método de evaluación.

	Método propuesto	ExG	ExGR	CIVE	NDI
$\mu(\%)$	93.592	72.998	60.630	90.460	70.304
$\sigma(\%)$	2.380	13.079	16.386	6.319	15.010

Del *Cuadro 3* se puede apreciar que el método propuesto tiene una calidad de segmentación de más del 4% que el presentado en el *Cuadro 2*. Así mismo es cuando menos 3% mayor que los otros métodos usando el mismo índice de evaluación (Ec. 14), la desviación estándar muestra estar por debajo del 3%. Los índices de segmentación ExG, ExGR y NDI mostraron una diferencia en la media de calidad con respecto al método propuesto de 20.6%, 33.0%, 3.1% y 23.3% respectivamente.

3.4.3 Rendimiento del método propuesto comparado con imágenes de diferente resolución.

En la *Figura* se presenta el resultado de aplicar el método propuesto a imágenes de tamaño 2112 x 2816 píxeles (base de datos proporcionada por el grupo de visión activa del CINVESTAV unidad Saltillo), la evaluación de este en comparación con la base de datos propia se muestra en el *Cuadro 4*.

Cuadro 4. Calidad de segmentación con diferente resolución de imágenes evaluado mediante el método propuesto por ATRWG.

	Base de datos propia (640 x 480 píxeles)	Base de datos CINVESTAV (2112 x 2816)
$\mu(\%)$	89.341	92.467
$\sigma(\%)$	4.044	3.687

Del *Cuadro 4*, la media de la calidad de la segmentación con imágenes de mayor resolución (2112 x 2816 píxeles) es mayor en 3% y la desviación estándar menor en 0.35% que la media y la desviación estándar obtenida con imágenes de menor resolución (640 x 480 píxeles). Esto se debe principalmente a que hay una mejor diferenciación entre los píxeles que corresponden al cultivo y al fondo. Así mismo, el tiempo computacional se incrementa proporcionalmente con el tamaño de la imagen, para las imágenes de tamaño 640 x 480 píxeles se tuvo un tiempo promedio de procesamiento de 9.171 ms, mientras que para las imágenes de tamaño 2112 x 2816 píxeles el tiempo promedio de procesamiento fue de 146.938 ms. Esto representa un incremento de aproximadamente 16 veces el tiempo necesario para procesar una imagen de dimensión 640 x 480 píxeles. De los resultados anteriores se demuestra que el uso de imágenes de mayor tamaño incrementa ligeramente el rendimiento de la segmentación, sin embargo incrementa significativamente el tiempo de procesamiento.



Figura 7. Resultado de segmentación de imágenes de tamaño 2112 x 2816 píxeles.

3.5 Conclusiones

En el trabajo se desarrolló un nuevo método de segmentación de imágenes para el cultivo de tomate en invernadero basado en el espacio de color CIE $L^*a^*b^*$ en condiciones de iluminación natural. Este método emplea los histogramas de los canales a y b para modelar la distribución del color del cultivo y a partir de estas distribuciones se definen los umbrales para la segmentación. Se analizaron 50 imágenes de las cuales la media de la calidad de segmentación del método propuesto fue de 89.3% utilizando la metodología de evaluación de ATRWG, superando por 4% o más a los métodos de segmentación ExG, ExGR, CIVE y NDI.

La desviación estándar para el método propuesto fue de 4.0%, que es inferior a los obtenidos por los otros métodos. Además, la calidad de segmentación de los métodos se evaluaron con el índice propuesto por Xiao *et al.* (2011), obteniéndose mejores resultados con el nuevo método. Por otra parte, se encontró que al incrementar la resolución de la imagen de 640 x 480 píxeles a 2112 x 2816 píxeles mejoró la calidad de segmentación en 3% y la desviación estándar disminuyó en 0.35%. Sin embargo, el tiempo computacional se incrementa en función del tamaño de la imagen pasando de 9.2 ms a 146.9 ms en promedio, respectivamente. En general, los resultados del experimento indican que el método propuesto es robusto a las variaciones de iluminación natural y a las condiciones estructurales del invernadero en el que se trabaje. El método propuesto puede extenderse para la segmentación de otros cultivos en invernadero aplicando el mismo procedimiento de adquisición de las imágenes, como una primera etapa para obtener diversos parámetros de la planta.

3.6 Literatura citada

- Andújar, D., Ribeiro, A., Fernández-Quintanilla, C., & Dorado, J. (2016). Using depth cameras to extract structural parameters to assess the growth state and yield of cauliflower crops. *Computers and Electronics in Agriculture*, 122, 67-73.
- Arribas, J. I., Sánchez-Ferrero, G. V., Ruiz-Ruiz, G., & Gómez-Gil, J. (2011). Leaf classification in sunflower crops by computer vision and neural networks. *Computers and Electronics in Agriculture*, 78(1), 9-18.
- Bac, C. W., Hemming, J., & van Henten, E. J. (2014). Stem localization of sweet-pepper plants using the support wire as a visual cue. *Computers and Electronics in Agriculture*, 105, 111-120. doi: <http://dx.doi.org/10.1016/j.compag.2014.04.011>
- Bai, X. D., Cao, Z. G., Wang, Y., Yu, Z. H., Zhang, X. F., & Li, C. N. (2013). Crop segmentation from images by morphology modeling in the CIE L*a*b* color space. *Computers and Electronics in Agriculture*, 99, 21-34. doi: <http://dx.doi.org/10.1016/j.compag.2013.08.022>

- Burgos-Artizzu, X. P., Ribeiro, A., Guijarro, M., & Pajares, G. (2011). Real-time image processing for crop/weed discrimination in maize fields. *Computers and Electronics in Agriculture*, 75(2), 337-346.
- Cuevas, E., Zaldívar, D., & Pérez, M. (2010). Procesamiento digital de imágenes con MATLAB y Simulink. *Alfaomega Ra-Ma (México)*.
- Chapra, S. C. C., Chapra, R. P. S. C., Canale, R. P., Chapra, S. C. C., Chapra, R. P. C., Canale, R. P., . . . Raymond, P. (2007). *Métodos numéricos para ingenieros*: McGraw-Hill.
- Chung, C.-L., Huang, K.-J., Chen, S.-Y., Lai, M.-H., Chen, Y.-C., & Kuo, Y.-F. (2016). Detecting Bakanae disease in rice seedlings by machine vision. *Computers and Electronics in Agriculture*, 121, 404-411.
- Ehret, D., Lau, A., Bittman, S., Lin, W., & Shelford, T. (2001). Automated monitoring of greenhouse crops. *Agronomie*, 21(4), 403-414.
- Farkas, I. (2010). Use of Computer Vision Methods for Irrigation Control. *IFAC Proceedings Volumes*, 43(26), 289-293.
- Gongal, A., Silwal, A., Amatya, S., Karkee, M., Zhang, Q., & Lewis, K. (2016). Apple crop-load estimation with over-the-row machine vision system. *Computers and Electronics in Agriculture*, 120, 26-35.
- Guo, W., Rage, U. K., & Ninomiya, S. (2013). Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model. *Comput. Electron. Agric.*, 96, 58-66. doi: 10.1016/j.compag.2013.04.010
- Hague, T., Tillett, N., & Wheeler, H. (2006). Automated crop and weed monitoring in widely spaced cereals. *Precision Agriculture*, 7(1), 21-32.
- Hernández-Hernández, J. L., García-Mateos, G., González-Esquiva, J. M., Escarabajal-Henarejos, D., Ruiz-Canales, A., & Molina-Martínez, J. M. (2016). Optimal color space selection method for plant/soil segmentation in agriculture. *Computers and Electronics in Agriculture*, 122, 124-132. doi: <http://dx.doi.org/10.1016/j.compag.2016.01.020>

- Kataoka, T., Kaneko, T., Okamoto, H., & Hata, S. (2003, 20-24 July 2003). *Crop growth estimation system using machine vision*. Paper presented at the Advanced Intelligent Mechatronics, 2003. AIM 2003. Proceedings. 2003 IEEE/ASME International Conference on.
- Kurtulmuş, F., & Kavdir, İ. (2014). Detecting corn tassels using computer vision and support vector machines. *Expert Systems with Applications*, 41(16), 7390-7397.
- Li, Y., Xia, C., & Lee, J. (2015). Detection of small-sized insect pest in greenhouses based on multifractal analysis. *Optik-International Journal for Light and Electron Optics*, 126(19), 2138-2143.
- Lu, H., Cao, Z., Xiao, Y., Li, Y., & Zhu, Y. (2016). Region-based colour modelling for joint crop and maize tassel segmentation. *Biosystems Engineering*, 147, 139-150.
- Ma, J., Li, X., Wen, H., Fu, Z., & Zhang, L. (2015). A key frame extraction method for processing greenhouse vegetables production monitoring video. *Computers and Electronics in Agriculture*, 111, 92-102.
- Mangus, D. L., Sharda, A., & Zhang, N. (2016). Development and evaluation of thermal infrared imaging system for high spatial and temporal resolution crop water stress monitoring of corn within a greenhouse. *Computers and Electronics in Agriculture*, 121, 149-159.
- Mao, H., Gao, H., Zhang, X., & Kumi, F. (2015). Nondestructive measurement of total nitrogen in lettuce by integrating spectroscopy and computer vision. *Scientia Horticulturae*, 184, 1-7.
- Neto, J. (2004). A combined statistical-soft computing approach for classification and mapping weed species in minimum-tillage systems.
- Nissimov, S., Goldberger, J., & Alchanatis, V. (2015). Obstacle detection in a greenhouse environment using the Kinect sensor. *Computers and Electronics in Agriculture*, 113, 104-115.
- Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica*, 11(285-296), 23-27.
- Payne, A. B., Walsh, K. B., Subedi, P., & Jarvis, D. (2013). Estimation of mango crop yield using image analysis–segmentation method. *Computers and Electronics in Agriculture*, 91, 57-64.

- Perez, A., Lopez, F., Benlloch, J., & Christensen, S. (2000). Colour and shape analysis techniques for weed detection in cereal fields. *Computers and electronics in agriculture*, 25(3), 197-212.
- Pires, R. D. L., Gonçalves, D. N., Oruê, J. P. M., Kanashiro, W. E. S., Rodrigues, J. F., Machado, B. B., & Gonçalves, W. N. (2016). Local descriptors for soybean disease recognition. *Computers and Electronics in Agriculture*, 125, 48-55.
- Pourreza, A., Lee, W. S., Ehsani, R., Schueller, J. K., & Raveh, E. (2015). An optimum method for real-time in-field detection of Huanglongbing disease using a vision sensor. *Computers and Electronics in Agriculture*, 110, 221-232.
- Rajendra, P., Kondo, N., Ninomiya, K., Kamata, J., Kurita, M., Shiigi, T., . . . Kohno, Y. (2009). Machine vision algorithm for robots to harvest strawberries in tabletop culture greenhouses. *Engineering in Agriculture, Environment and Food*, 2(1), 24-30.
- Romualdo, L., Luz, P. H. d. C., Devechio, F., Marin, M., Zúñiga, A., Bruno, O. M., & Herling, V. R. (2014). Use of artificial vision techniques for diagnostic of nitrogen nutritional status in maize plants. *Computers and Electronics in Agriculture*, 104, 63-70.
- Sakamoto, T., Gitelson, A. A., Nguy-Robertson, A. L., Arkebauer, T. J., Wardlow, B. D., Suyker, A. E., . . . Shibayama, M. (2012). An alternative method using digital cameras for continuous monitoring of crop status. *Agricultural and Forest Meteorology*, 154, 113-126.
- Story, D., Kacira, M., Kubota, C., Akoglu, A., & An, L. (2010). Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments. *Computers and Electronics in Agriculture*, 74(2), 238-243.
- Sun, G., Li, Y., Zhang, Y., Wang, X., Chen, M., Li, X., & Yan, T. (2016). Nondestructive measurement method for greenhouse cucumber parameters based on machine vision. *Engineering in Agriculture, Environment and Food*, 9(1), 70-78.
- Tellaeche, A., Pajares, G., Burgos-Artizzu, X. P., & Ribeiro, A. (2011). A computer vision approach for weeds identification through Support Vector Machines. *Applied Soft Computing*, 11(1), 908-915.

- Van Henten, E. J., Van Tuijl, B. A. J., Hemming, J., Kornet, J. G., Bontsema, J., & Van Os, E. A. (2003). Field Test of an Autonomous Cucumber Picking Robot. *Biosystems Engineering*, 86(3), 305-313. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2003.08.002>
- Velázquez-López, N., Sasaki, Y., Nakano, K., Mejía-Muñoz, J. M., & Romanchik Kriuchkova, E. (2011). Detección de cenicilla en rosa usando procesamiento de imágenes por computadora. *Revista Chapingo. Serie horticultura*, 17, 151-160.
- Vidović, I., Cupec, R., & Hocenski, Ž. (2016). Crop row detection by global energy minimization. *Pattern Recognition*, 55, 68-86.
- Woebbecke, D., Meyer, G., Von Bargen, K., & Mortensen, D. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE*, 38(1), 259-269.
- Xia, C., Chon, T.-S., Ren, Z., & Lee, J.-M. (2015). Automatic identification and counting of small size pests in greenhouse conditions with low computational cost. *Ecological Informatics*, 29, 139-146.
- Xiang, R., Jiang, H., & Ying, Y. (2014). Recognition of clustered tomatoes based on binocular stereo vision. *Computers and Electronics in Agriculture*, 106, 75-90. doi: <http://dx.doi.org/10.1016/j.compag.2014.05.006>
- Xiao, Y., Cao, Z., & Zhuo, W. (2011). Type-2 fuzzy thresholding using GLSC histogram of human visual nonlinearity characteristics. *Optics Express*, 19(11), 10656-10672. doi: 10.1364/oe.19.010656
- Xue, J., Zhang, L., & Grift, T. E. (2012). Variable field-of-view machine vision based row guidance of an agricultural robot. *Computers and Electronics in Agriculture*, 84, 85-91.
- Ye, M., Cao, Z., Yu, Z., & Bai, X. (2015). Crop feature extraction from images with probabilistic superpixel Markov random field. *Computers and Electronics in Agriculture*, 114, 247-260.
- Yu, Z., Cao, Z., Wu, X., Bai, X., Qin, Y., Zhuo, W., . . . Xue, H. (2013). Automatic image-based detection technology for two critical growth stages of maize: Emergence and three-leaf stage. *Agricultural and Forest Meteorology*, 174, 65-84.

- Zheng, L., Zhang, J., & Wang, Q. (2009). Mean-shift-based color segmentation of images containing green vegetation. *Computers and Electronics in Agriculture*, 65(1), 93-98.
- Zhu, Y., Cao, Z., Lu, H., Li, Y., & Xiao, Y. (2016). In-field automatic observation of wheat heading stage using computer vision. *Biosystems Engineering*, 143, 28-41.

CAPÍTULO 4

RECONOCIMIENTO AUTOMÁTICO DEL TALLO DE PLANTAS DE TOMATE EN INVERNADERO MEDIANTE VISIÓN POR COMPUTADORA

Abner Orlando Chávez-Sánchez^{a*}, Luz Abril Torres-Mendez^b, Noé Velázquez-López^c,
Agustín Ruiz-García^c, Pedro Cruz-Meza^d

^a*Universidad Autónoma Chapingo, Posgrado en Ingeniería Agrícola y Uso Integral del Agua. Carretera México-Texcoco km. 38.5, Chapingo, Estado de México. C.P. 56230, México.*

^b*CINVESTAV - Unidad Saltillo Ramos Arizpe, Coah., 25900*

^c*Universidad Autónoma Chapingo, Departamento de Irrigación. Carretera México-Texcoco km. 38.5, Chapingo, Estado de México. C.P. 56230, México.*

^d*Universidad Autónoma Chapingo, Departamento de Ingeniería Mecánica Agrícola. Carretera México-Texcoco km. 38.5, Chapingo, Estado de México. C.P. 56230, México.*

4.1 Resumen

Las labores de cultivo y cosecha en plantas de tomate en invernadero demandan una alta cantidad de mano de obra, lo que se ve reflejado en los costos de producción. Una solución es implementar brazos robóticos para realizar esta actividad, sin embargo, para ello es necesario conocer la posición de la planta y específicamente la del tallo. Este es una referencia visual de la ubicación de los órganos de la planta, además, delimita el corte de las hojas y yemas axilares en el caso de la poda. Este es un tema que solo se ha estudiado haciendo uso de sistemas multiespectrales e iluminación artificial. Por lo tanto, este trabajo tiene como objetivo desarrollar un algoritmo para el reconocimiento del tallo de la planta de tomate en invernadero bajo condiciones de iluminación natural, empleando un sistema estereoscópico de bajo costo.

El sistema de estereo visión está compuesto por dos cámaras web marca Logitech modelo C920 con una línea base de 50 mm. La adquisición de imágenes se efectuó inclinando ligeramente los ejes ópticos de las cámaras web hacia el suelo, es decir, se ubicaron de tal forma que la radiación no incidiera directamente en las lentes. El autofocus fue desactivado, mientras la auto exposición se mantuvo activada. El algoritmo propuesto se compone de las siguientes etapas: obtención del mapa de disparidad, segmentación del cultivo, clasificación de segmentos del tallo y del hilo de tutoreo, extracción de segmentos del tallo de interés. Finalmente, se interpola mediante trazadores cúbicos para estimar la posición del tallo en áreas donde no se cuenta con información. Los resultados muestran que la calidad media de reconocimiento del tallo mediante el método propuesto es del 95.0%. Además se muestra que el método propuesto es robusto a la variación de iluminación y a la etapa de crecimiento del cultivo en el interior de invernaderos estructuralmente diferentes. En general, es posible hacer uso del algoritmo desarrollado para reconocer el tallo de plantas de tomate de manera eficaz. Este trabajo es la primera etapa para el desarrollo de algoritmos de reconocimiento de otros órganos de la planta, que permitan automatizar labores de cultivo y cosecha.

Palabras clave: Estereo visión, segmentación, PCA, iluminación natural.

4.2 Introducción

El desarrollo de nuevas tecnologías para la agricultura como son los invernaderos, considerados como sistemas biofísicos con entradas, salidas y bucles de control de procesos mejoran las condiciones de producción con la finalidad de obtener mayores rendimientos (Ehret *et al.*, 2001). Las actividades fundamentales dentro de un invernadero son las labores de cultivo (poda, tutoreo y limpieza) y la cosecha, éstas conjuntamente representan el 85% de la mano de obra empleada (de Pablo *et al.*, 2003).

de Pablo *et al.* (2003) realizaron un estudio para estimar los costos de producción del tomate cultivado en suelo y sustrato, donde para el primer caso reportaron que la mano de obra representa el 35% del costo total. Del total de este porcentaje las labores de cultivo (podas, tutoreo, etc.) representan el 46 %, la cosecha el 32% y, el mantenimiento y la limpieza el 14%.

Por lo tanto, el aumento de costos debido a la alta cantidad de mano de obra empleada en labores de cultivo y cosecha, así como, la agricultura sostenible que demanda urgentemente el aumento de la eficiencia de los recursos en la producción de cultivos sin dañar el medio ambiente se está convirtiendo en un desafío para la agricultura. La automatización y la robótica agrícola pueden brindar una solución (Bergerman *et al.*, 2013). Estas tecnologías han reducido los costos de producción, elevado la calidad de los productos agrícolas y mejorado el control del medio ambiente (Edan *et al.*, 2009).

Un enfoque en el cual la robótica agrícola se basa es el control por visión. Este es un sistema mediante el cual el robot lleva a cabo tareas de recolección de frutos guiado mediante información visual. Se compone de dos etapas: 1) El reconocimiento del fruto, 2) La coordinación ojo-mano (Hashimoto, 2003). Los objetos de trabajo de robots de cosecha presentan un alto grado de variedad en tamaño, forma, color, textura y dureza como resultado de diferencias ambientales y genéticas. Además, el espacio de trabajo es complejo y poco estructurado con grandes variaciones de iluminación y alto grado de oclusión. Por lo tanto, el cuello de botella para promover la aplicación de estos sistemas en labores de cultivo y cosecha se encuentra en la falta de algoritmos robustos para el reconocimiento de frutos y órganos de las plantas.

La visión por computadora en invernaderos tiene la capacidad de solventar este problema, así pues existen numerosas investigaciones que mencionan algoritmos para el reconocimiento de frutos, capaces de proveer la posición del objeto de trabajo y que han sido empleados para la cosecha de frutos forma robotizada. Estos enfoques hacen uso de técnicas basadas en el tipo de cámaras (RGB, térmicas y multiespectrales) y en la configuración de estas (monocular, visión estereoscópica, visión activa mediante rayos láser) (Zhao *et al.*, 2016a).

Yang *et al.* (2007) presentaron un algoritmo para el reconocimiento de tomates agrupados en invernadero, el método realiza una reconstrucción 3D del entorno basado en visión estereoscópica. Zhao *et al.* (2016b) y Zhao *et al.* (2016c) desarrollaron un algoritmo que combina el clasificador AdaBoost y el análisis de color para la detección automática de tomates maduros en invernadero. Reportaron una detección de más del 96% de tomates maduros.

Así pues se han desarrollado algoritmos para el reconocimiento de diferentes frutos: tomates (Arefi *et al.*, 2011; Xiang *et al.*, 2014), manzanas (Kelman y Linker, 2014; Si *et al.*, 2015), fresas (Xie *et al.*, 2011), pimiento dulce (Barth *et al.*, 2016; Vitzrabin y Edan, 2016), melocotón (Kurtulmus *et al.*, 2014), pepino (Yuan *et al.*, 2011) y kiwi (Zhan *et al.*, 2013).

Los robots de cosecha que aplican estos algoritmos se desplazan por el invernadero manejando una planta a la vez (Kohan *et al.*, 2011), se plantea el mismo principio para realizar labores de cultivo como la poda de yemas axilares y hojas. Para la cosecha de frutos, se propone reducir el tiempo computacional realizando el reconocimiento del tallo como indicador visual de la ubicación y tendencia de crecimiento de la planta, lo que provee una estimación indirecta de la ubicación de los frutos (Hemming *et al.*, 2014; Van Henten *et al.*, 2003). Para el caso de la poda de yemas axilares y hojas es indispensable contar con el tallo como referencia visual para delimitar el corte de yemas axilares y hojas, En los siguientes apartados se presentan los trabajos reportados para realizar esta actividad.

Bac *et al.* (2013) desarrollaron un algoritmo para identificar órganos de la planta, entre ellas el tallo del pimiento dulce basado en un sistema multiespectral con iluminación artificial para atenuar las perturbaciones causadas por las condiciones naturales. Realizó una clasificación basada en píxeles de donde obtuvo una tasa de identificación del tallo del 40.0 % con una desviación estándar del 12.4 %.

Fernández *et al.* (2013) proponen un algoritmo para discriminar órganos de la vid (hojas, tallos, ramas, frutos y el fondo), obteniendo una precisión de clasificación del tallo del 57.2%. Bac *et al.* (2014) desarrollaron un algoritmo capaz de localizar el tallo del cultivo de pimiento dulce, usando la detección del hilo de tutorio como referencia visual. La tasa de detección del hilo fue más favorable en condiciones de radiación moderada (94%) llegando a ser de hasta 74% con radiación fuerte. Este método tiene el inconveniente que necesita contar con una fuente de iluminación artificial, además de que el rendimiento varía en dependencia de las condiciones de iluminación. Zhang *et al.* (2015) desarrollaron un algoritmo para el reconocimiento del tallo de la manzana mediante la reconstrucción 3D.

Adiciona el uso de imágenes multiespectrales y un sistema visual activo con rayos láser, obteniendo una precisión del 97.5%. Sin embargo a pesar de la exactitud de estos métodos, la complejidad así como el costo elevado debido al sistema multiespectral e iluminación artificial a menudo limitan su aplicación práctica.

El objetivo de esta investigación fue desarrollar un algoritmo para el reconocimiento automático del tallo de la planta de tomate de crecimiento indeterminado en invernaderos bajo iluminación natural. El sistema diseñado está basado en la visión estereoscópica y en el análisis de características invariantes del tallo. La localización espacial del tallo es la primera etapa para la automatización de las tareas agrícolas, dentro de ellas la poda de cultivos en invernadero mediante un robot.

4.3 Materiales y métodos

4.3.1 Cultivo y adquisición de imágenes

El estudio se llevó a cabo haciendo uso del cultivo de jitomate (*Solanum lycopersicum L.*) sembrado en 4 invernaderos, dos de éstos están ubicados en la Universidad Autónoma Chapingo, Estado de México, México (19°29'34.0"N 98°53'01.5"W), el tercero en San Pedro Martír, Oaxaca, México (16° 44' 24"N, 96° 43' 30"W) y el cuarto en la Universidad Autónoma Agraria Antonio Narro, Coahuila, México (25°21'5"N 101°1'47"W).

Tres de los invernaderos utilizan plástico (polietileno de baja densidad) como material de cubierta mientras que uno vidrio. Así mismo el medio de producción para 3 de los invernadero es hidroponía, sin embargo, para uno es suelo.

Las imágenes fueron adquiridas utilizando un sistema de estereo visión compuesto por dos cámaras web marca Logitech modelo C920 con una distancia base de 50 mm (

Cuadro). El sistema de estereo vision se fijó en un tripie con eje el óptico (ángulo pitch) inclinado ligeramente hacia el suelo. Las cámaras web se ubicaron de tal forma que la radiación directa no incidiera en las lentes, el autofocus de mantuvo desactivado, mientras que la auto exposición se mantuvo activada. Las cámaras se colocaron a una distancia entre 50-70 cm del objetivo (*Figura*).

Cuadro 1. Especificaciones técnicas de las webcam Logitech C920 utilizadas

Especificaciones	HD pro webcam C920
Definición de videos	Full HD 1080p con H.264
Calidad de la foto	15 MP
Autofoco	Autofoco con 20 pasos
Corrección automática de iluminación	Premium



Figura 1. Equipo para la adquisición de imágenes.

Se adquirieron 134 pares de imágenes en formato tif de tamaño 640 x 480 pixeles, tomadas bajo diferentes condiciones de iluminación. Las imágenes fueron tomadas de 9 am a 2 pm a partir del 16 de Noviembre del 2015 al 18 de Abril del 2016 en diferentes etapas de crecimiento del cultivo. Por cada imagen izquierda se registró 1 tallo correspondiente a la planta de interés.

4.3.2 Desarrollo del algoritmo

El algoritmo (*Figura*) se desarrolló haciendo uso de las bibliotecas de procesamiento de imágenes de Opencv versión 3.1.0 en el lenguaje de programación C++. Se utilizó una Laptop marca Toshiba modelo Satellite P745 con procesador Intel Core i5 y 2.4 GHz, 6GB de RAM, y el sistema operativo Ubuntu 2014 LTS. El algoritmo inicia con la rectificación y estéreo-correspondencia para obtener información 3D de la escena, esto permite segmentar la hilera de plantas más cercana a la cámara. En la siguiente etapa se realiza una segmentación del cultivo en la imagen izquierda. A continuación se extraen segmentos del tallo del cultivo, estos pueden pertenecer a diferentes plantas que se encuentra en la escena. Después, los hilos de tuteo son localizados. Posteriormente con el resultado de los dos pasos anteriores se hace una correspondencia con la finalidad de garantizar una correcta localización del tallo

En la siguiente etapa se realiza una interpolación con los segmentos del tallo para determinar su trayectoria en lugares en los que no se cuenta con información. Finalmente se obtienen las coordenadas en la imagen que representan el tallo que nos interesa monitorear.

Rectificación y estéreo-correspondencia

Las imágenes fueron rectificadas utilizando los parámetros intrínsecos y extrínsecos de las cámaras. Estos se determinaron mediante el procedimiento descrito por Zhang (2000). Dados los parámetros intrínsecos y extrínsecos es posible satisfacer la restricción epipolar (Brown et al., 2003).

Posteriormente, se implementó el algoritmo Semiglobal Matching modificado (Hirschmuller, 2008) para la estimación de correspondencias, el algoritmo busca coincidencias entre bloques a lo largo de la línea epipolar y para ello se utilizan las imágenes RGB originales. Los parámetros del algoritmo se establecieron como: la ventana (SADWindowSize) de amplitud 5x5 píxeles, se fijaron disparidades mínimas y máximas correspondientes a un rango de profundidad de 0.3 a 0.8 m debido a que es el rango en el que se encuentran los tallos de las plantas, el tamaño máximo de las regiones lisas de disparidad (speckleWindowSize) se establecieron en 200, el margen en porcentaje por el que el valor mínimo de la función de costos calculada debe superar al segundo mejor valor para asumir una coincidencia correcta (uniquenessRatio) se fijó en 50 y la variación máxima de disparidad dentro de cada componente conectado se fijó en 16. La salida de esta primera etapa es un mapa de disparidad en el que el tallo de las plantas es el componente de mayor interés.

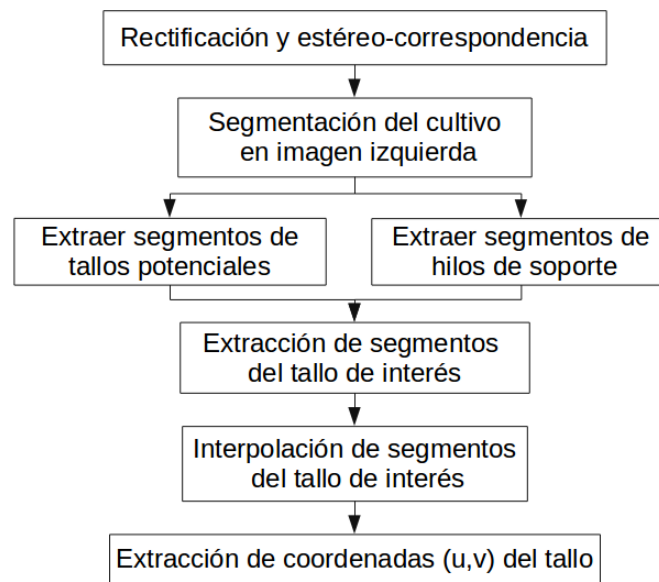


Figura 2. Diagrama de flujo para localizar el tallo del cultivo de jitomate a partir de imágenes RGB.

Segmentación del cultivo

En la etapa de segmentación se hace uso del espacio de color *CIE L*a*b** de donde se analizan los canales *a* y *b* de cada una de las imágenes de la base de datos con la finalidad de definir umbrales óptimos en sus histogramas. Experimentalmente se observó que el color del cultivo en el histograma de los canal *a* y *b* tienen una distribución normal y esta se distribuye del lado izquierdo para el canal *a* y del lado derecho para el ultimo. El histograma de una imagen a escala de grises está definido según la ecuación 1 (Cuevas et al., 2010).

$$f(x) = \text{card}^2\{I(u, v) \mid I(u, v) = x\} \quad (1)$$

Donde, *I* es la imagen a escala de grises, *u* el número de la filas de ésta, y *v* el número de columnas. El método de segmentación propuesto consiste en determinar y aplica umbrales para los canales *a* y *b*.

Determinación del umbral para binarizar el canal a

El primer paso es transformar la imagen fuente del espacio de color *RGB* al espacio de color *CIE L*a*b**. Posteriormente se extrae el histograma del canal *a* y de este los valores de la ecuación (1) diferentes de cero.

Estos valores con sus correspondientes niveles de intensidad se almacenan en un vector de puntos **w**. Como resultado del análisis de los histogramas del canal *a* se encontró que la distribución es normal del lado izquierdo (Figura 3 a); entonces la tarea principal recae en delimitar un rango en el cual el umbral óptimo para la segmentación de la planta se encuentra. El rango se delimita de la esperanza matemática (μ_x) a un valor *c*, que es el nivel de intensidad en el cual se presenta el argumento máximo de la primera derivada de la función definida en la ecuación (1). Por otro lado, μ_x se calcula primero obteniendo una aproximación en un rango de 0 a 127 con la ecuación (2), donde 127 es el nivel de intensidad máximo del color verde. Después a través de una técnica de refinamiento iterativo, la cual consiste en minimizar el error relativo definido por la ecuación (3).

$$\mu_x = \sum_x x f_X(x), 0 \leq x \leq 127 \quad (2)$$

$$\varepsilon_t = \frac{\sum(f_i \text{ real} - f_i \text{ estimada})}{\sum f_i \text{ real}} \quad (3)$$

Para cada iteración se toma un nuevo rango del nivel de intensidades del vector $\mathbf{w}$. Dicho rango es delimitado en el extremo superior por los mínimos locales de las frecuencias del vector $\mathbf{w}$ menores o iguales a 127 y mayores que la primera aproximación de μ_x . Los niveles de intensidad en los que se presentan los mínimos locales en el rango de μ_x a 127 (Ec. 5) se almacenan en el vector $\mathbf{q}$.

$$f(x): [\mu_x, 127] \subseteq \mathbb{R} \rightarrow \mathbb{R}, \exists x_0 \in [\mu_x, 127], P = (x_0, f(x_0)) \quad (4)$$

$$P(x_0, f(x_0)) \text{ mínimo local de } f \Leftrightarrow E'(x_0) | \forall x \neq x_0 \in E'(x_0): f(x) > f(x_0) \quad (5)$$

Posteriormente, en las siguientes iteraciones para el cálculo de μ_x los rangos son asignados en dos pasos: primero se fija el valor máximo, que corresponde a cada elemento del vector $\mathbf{q}$ y para cada componente de este se asignan como valor inferior del rango cada elemento que corresponde al nivel de intensidad del vector $\mathbf{w}$, siempre y cuando éste sea al menos dos unidades menor al valor de la primera aproximación de μ_x . Posteriormente c es calculado (Ec. 6), este valor representa la separación más evidente entre el color verde y rojo del canal.

$$f'(x): [\mu_x, 255] \subseteq \mathbb{R} \rightarrow \mathbb{R}, \exists! c | \forall x \in [\mu_x, 255]: f'(c) > f'(x), x \neq c \quad (6)$$

Donde $f'(x)$ es la primera derivada de las frecuencias de intensidad del vector $\mathbf{w}$ calculada mediante diferencias finitas divididas hacia atrás (Chapra *et al.*, 2007) y c es el nivel de intensidad en el cual se presenta el argumento máximo de la primera derivada. En la siguiente etapa se obtienen los niveles de intensidad que corresponden a los mínimos locales de las frecuencias del vector $\mathbf{w}$ en el rango definido por μ_x y c , estos son almacenados en el vector $\mathbf{v}$. Donde el umbral deseado (*tha*) es el argumento máximo del vector $\mathbf{v}$, que puede escribirse como:

$$tha = \begin{cases} \text{argmax } \mathbf{v} & ; \mu_x < v_i < c \\ c, & \text{de otra forma} \end{cases} \quad (7)$$

Determinación del umbral para binarizar el canal b.

De la imagen transformada al espacio de color *CIE L*a*b**, se extrae el histograma del canal *b* (Ec. 1) y de este las frecuencias relativas diferentes de cero. Estos valores con sus correspondientes niveles de intensidad se almacenan en un vector de puntos **wb**. Como resultado del análisis de los histogramas del canal *b* se encontró que la distribución es normal del lado derecho (Figura 3 a); entonces al igual que en el caso del canal *a* la tarea principal recae en delimitar un rango en el cual el umbral óptimo se encuentra, éste se delimita del valor *cb* (Ec. 11) a la esperanza matemática de la distribución normal (μ_x). El valor *cb* es el nivel de intensidades que corresponde al argumento mínimo de la primera derivada de las frecuencias del vector **wb**, mientras que μ_x es calculada primero obteniendo una aproximación en un rango de 129 a 255 con la Ec. (8). Después a través de una técnica de refinamiento iterativo, que consiste en minimizar el error relativo definido en la Ec. (3). Para cada iteración se toma un nuevo rango del nivel de intensidades del vector **wb**. Dicho rango es delimitado en el extremo inferior de la distribución por los mínimos locales de las frecuencias del vector **wb** mayores que 128 y menores que la primera aproximación de μ_x . Los niveles de intensidad en los que se presentan los mínimos locales en el rango de 128 a μ_x (Ec. 10) son almacenados en un vector **qb**.

$$\mu_x = \sum_x x f_x(x), 129 < x < 255 \quad (8)$$

$$f(x): [128, \mu_x] \subseteq \mathbb{R} \rightarrow \mathbb{R}, \exists x_0 \in [128, \mu_x], P = (x_0, f(x_0)) \quad (9)$$

$$P(x_0, f(x_0)) \text{ mínimo local de } f \Leftrightarrow E'(x_0) | \forall x \neq x_0 \in E'(x_0): f(x) > f(x_0) \quad (10)$$

En las iteraciones posteriores para el cálculo de μ_x los rangos son asignados en dos pasos, el primero de ellos consiste en fijar el valor mínimo para ello se toma cada uno de los elementos del vector **qb** como valor inferior. Y para cada elemento del vector **qb** se asignan como valor superior del rango cada uno de los elementos correspondientes al nivel de intensidad del vector **w**, siempre y cuando éste sea al menos cuatro unidades mayor al valor de la primera aproximación de μ_x . Posteriormente, el nivel de intensidades *cb* es calculado (Ec. 11), este valor representa la separación más evidente entre el color azul y amarillo del canal.

$$f'(x): [0, \mu_x] \subseteq \mathbb{R} \rightarrow \mathbb{R}, \exists! cb | \forall x \in [0, \mu_x]: f'(cb) > f'(x), x \neq c \quad (11)$$

Donde $f'(x)$ es la primera derivada de las frecuencias de intensidades del vector **wb** calculada también por diferencias finitas divididas hacia atrás y cb es el nivel de intensidades que corresponde al argumento mínimo de la primera derivada de las frecuencias del vector **wb**. En la siguiente etapa se obtienen los niveles de intensidad que corresponden a los mínimos locales de las frecuencias del vector **wb** en el rango cb a μ_x , estos son almacenados en un vector **vb**. Donde el umbral deseado (thb) es el argumento mínimo del vector **vb**, que puede escribirse como:

$$thb = \begin{cases} \text{argmin } vb & ; \quad cb < vb_i < \mu_x \\ cb; & \text{de otra forma} \end{cases} \quad (12)$$

Generación de la máscara de segmentación.

Los umbrales obtenidos son usados para binarizar los canales correspondientes de la imagen. Hasta este punto se cuenta con dos imágenes binarias, una del canal a y otra del b . Mediante la conjunción de píxel por píxel de estas dos imágenes se obtiene la máscara de segmentación. Finalmente, esta es usada para segmentar la imagen original

Extracción de tallos potenciales.

La extracción de tallos potenciales se basa en el análisis de las características morfológicas de los objetos del mapa de disparidad. Esta etapa se desarrolla en dos fases principales, la primera de ellas es un proceso de filtrado a través del cual se obtienen segmentos potenciales a ser parte del tallo. Se definen como segmentos a una agrupación de píxeles que representan un objeto o parte de él en el mapa de disparidades. La segunda fase es un proceso iterativo mediante el cual se agrupan objetos que comparten características específicas como la relación entre ellos, su posición y orientación del componente principal. A continuación se describen cada una de estas fases.

Extracción de tallos potenciales.

Filtrado del mapa de disparidad por profundidad y por aplicación de la máscara de segmentación. En esta primera sub-etapa se obtienen los objetos del mapa de disparidad en un rango de 100 a 255 en la escala de niveles de intensidad. Lo que equivale a extraer objetos a una distancia de 80 cm de las cámaras debido a que los tallos siempre se encontraron en este rango. Con la finalidad de reducir la información y obtener regiones pertenecientes únicamente al cultivo, la imagen binaria obtenida se filtra con la máscara de segmentación. Posteriormente al resultado obtenido se le aplica la transformación morfológica opening con una elipse de dimensión 3 x 3 como elemento estructurante, el objetivo es eliminar pequeños objetos. La salida es una imagen binaria con regiones pertenecientes al cultivo y a una distancia máxima de 80 cm.

Filtro de objetos por tamaño, orientación y elongación. En la máscara obtenida en el paso anterior se asume que el tallo del cultivo que nos interesa monitorear se encontrara en algún momento en un rango de 240 píxeles a lo ancho de la imagen. A este rango se le aplica un algoritmo de etiquetado de objetos con lo cual identificamos cada una de las regiones presentes, estas pueden pertenecer al tallo o a segmentos de hojas por lo que en los siguientes pasos se extraen solo aquellas que pertenecen al tallo del cultivo.

Se aplica un filtro por tamaño, donde las regiones con un área menor a 150 píxeles son eliminados. El tallo de la planta es un órgano alargado y se encuentran generalmente en una posición vertical o cercana a ésta, por lo que centramos toda la atención en la búsqueda de segmentos con una elongación menor a 0.2 y orientación entre $\pi/4$ y $3\pi/4$ radianes. Esto se logra haciendo uso del algoritmo PCA (Principal component analysis) (Hotelling, 1933) usando el método de covarianza, el cual nos devuelve la matriz $\mathbf{V}$ de eigenvectores de dimensión 2×2 , donde cada fila representa el eigenvector de un componentes principal. Los eigenvalores se obtienen en una matriz diagonal $\mathbf{D}$ también de dimensión 2×2 , donde $\mathbf{D}_{ij}$ con $i = j$ es el j -th eigenvalor correspondiente al j -th eigenvector. Mediante estos se calcula la elongación (Ec. 13) y la orientación (Ec. 14). Estas regiones obtenidas son denominadas regiones base y son utilizadas para generar un historial de su ubicación.

$$elongación = \frac{D_{i,j}}{D_{i+1,j+1}}, j = i \quad (13)$$

$$\alpha = \begin{cases} \arctan\left(\frac{-V_{i,j}}{V_{i,j+1}}\right) + \pi & \text{si } V_{i,j} < 0 \\ \arctan\left(\frac{-V_{i,j}}{V_{i,j+1}}\right) & \text{si } V_{i,j} > 0 \end{cases} \quad (14)$$

Extracción de líneas verticales mediante operaciones morfológicas. Se extraen mediante una operación de erosión seguida por una dilatación con un elemento estructurante rectangular de dimensión 2 píxeles de ancho y de largo el número de columnas de la imagen / 30. Esta sub-etapa fracciona las regiones filtradas en los pasos anteriores, se identifican y se filtran por tamaño eliminando aquellas compuestas por menos de 150 píxeles. Posteriormente se realiza una segunda aplicación del PCA mediante el cual obtenemos el eigenvalor del componente principal, el centroide y la orientación de cada región. Mediante este eigenvalor es posible identificar la región más alargada (tiene mayor probabilidad de ser un segmento del tallo), la salida es un vector en el cual se ordenan las regiones en dependencia de su elongación, así como de sus características (orientación, centroide). Para asegurar una detección correcta del tallo se restringe la dimensión del vector a más de 3 segmentos, de lo contrario se continúa con el procesamiento del siguiente par de imágenes.

Clasificación de segmentos potenciales.

Esta fase consiste en un proceso iterativo a través del cual se clasifican los segmentos tomando en cuenta características como; su posición, orientación del componente principal así como la relación entre ellos. El proceso consiste en tomar cada uno de los segmentos iniciando por el de mayor elongación y a partir de este formar un posible tallo. Para cada uno de ellos se realiza el siguiente proceso iterativo, se evalúa a que región base pertenece posteriormente se extraen todas las regiones actuales que pertenecen a esta, se eliminan segmentos que contienen su centroide en la misma fila que algún elemento de mayor elongación y finalmente se determinan los segmentos que se encuentran en la posición superior e inferior.

Definimos como el segmento superior aquel que se encuentra más cercano a la fila 0 e inferior aquel que se encuentra más cercano a la fila mayor. Finalmente, se realiza una búsqueda de segmentos para formar un posible tallo, esto a partir del segmento superior hacia la fila 0 y a partir del segmento inferior hacia la fila mayor.

Para el caso de la búsqueda de regiones hacia la fila 0 se realiza nuevamente un proceso iterativo en el que la entrada del ciclo es el segmento superior obtenido en el paso anterior y a partir de este se genera un nuevo elemento para alimentar el proceso. Este finaliza cuando no es posible incluir otro segmento, es decir cuando se llega a la fila 0 de la imagen. En este proceso el primer paso es obtener la orientación de la parte inicial del segmento, es decir, de una sección lo suficientemente grande para garantizar que la orientación del componente principal se encuentre entre $\pi/4$ y $3/4 \pi$ radianes. Esto define la tendencia real del crecimiento del tallo. Posteriormente se define un rango de ± 20 pixeles en dirección transversal y a lo largo de la orientación de la parte inicial del segmento, a través de este se realiza una búsqueda de algún elemento y en caso de encontrarse se procede a identificar a que región base pertenece. Se extraen las regiones actuales que pertenecen a ella, se filtran por coincidencia de su centroide con algún elemento de alguna región de mayor elongación de la región base. Estas regiones son tomadas como parte del posible tallo y se define la región superior que alimenta nuevamente al ciclo.

Para el caso de la búsqueda de regiones hacia la fila mayor, se repite el proceso para la búsqueda de posibles tallos hacia la fila 0, con la diferencia que en este caso la búsqueda se realiza en dirección contraria por lo que la entrada al ciclo es el elemento inferior y por cada uno de ellos se genera un nuevo elemento que alimente al ciclo. Este finaliza cuando no es posible incluir un nuevo elemento, es decir cuando se llega a la fila mayor de la imagen sin encontrar una nueva región. La salida de esta etapa es un vector con los centroides de cada uno de los segmentos que conforman el posible tallo.

Extracción de segmentos de hilos de tutoreo

Los hilos de tutoreo son una referencia indirecta de la posición del tallo de la planta de interés. El algoritmo descrito a continuación es desarrollado para segmentar hilos de tutoreo de color blanco y negro, se compone de 9 sub-etapas (*Figura 3*).

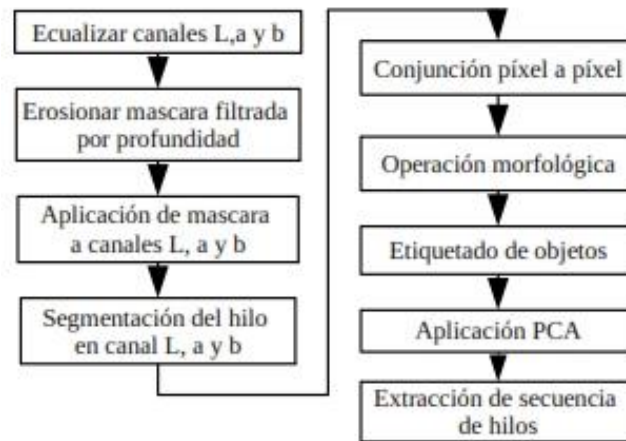


Figura 3. Diagrama de flujo para la extracción de hilos de tutoreo

Para hilos de tutoreo blancos, se busca que en los canales L , a y b se aprecie con intensidades mayores a 130 en la escala 0 a 255. Para ello se lleva a cabo la ecualización del canal L y a , en el caso del b se invierten cada uno de los píxeles de la imagen y posteriormente se ecualiza. En hilos de tutoreo negros se busca que este se aprecie en el canal L con una intensidad menor a 175, en el a y b con una intensidad menor a 80. Para ello se ecualiza directamente el canal a . L y el b se invierte para posteriormente ecualizarlo.

En la segunda sub-etapa, a la máscara resultante del mapa de disparidad tras la aplicación del filtro por profundidad se le aplica una operación morfológica de erosión con lo cual se evita que el contorno de las regiones sean detectadas como hilo en pasos posteriores. En la siguiente sub-etapa la máscara resultante es aplicada a cada uno de los canales obtenidos de la sub-etapa 1.

En la cuarta sub-etapa se realiza la segmentación por umbral, para hilos de tutoreo blancos este se fija en 130 para cada canal. Para hilos de tutoreo negros el umbral para el canal L se fija en 175 y para a y b en 80 en la escala 0 a 255. En la quinta sub-etapa se realiza una conjunción píxel a píxel de las tres mascarar binarias. Posteriormente en la sexta sub-etapa se aplica la operación morfológica de erosión con un elemento estructurante elíptico de dimensión 3×9 píxeles seguida por una dilatación con un elemento estructurante también elíptico pero con una dimensión de 3×3 píxeles, esto con la finalidad de eliminar pequeños objetos y mantener aquellos que presentan una estructura más alargada en posición vertical. En la séptima sub-etapa se delimita el área de búsqueda a 240 columnas situados en el centro de la imagen donde se aplica un algoritmo de etiquetado de objetos. Esto nos proporciona la ubicación y el número de píxeles que conforman cada segmento de hilo. De esta forma es posible filtrar dichas regiones por tamaño, se determinó que regiones menores a 30 píxeles causan ruido, por lo que se eliminan.

Los hilos utilizados para el tutoreo del jitomate se visualizan en la imagen también como segmentos alargados, muy angostos y con una orientación muy cercana a la vertical. Estas características son utilizadas para su identificación. En la octava sub-etapa se utiliza el PCA para reducir dimensionalmente las regiones de tal forma que nos permita obtener características de dos componentes principales, el eigenvector nos proporciona una referencia de la orientación de la componente, mientras que el eigenvalor cuantifica la dimensión del eigenvector por lo que es una medida indirecta de la dimensión del objeto en la dirección del eigenvector.

Un filtro para la orientación del componente principal es usado, la cual está determinada a partir de la componente x y y del eigenvector como se describe en la ecuación 14. El filtro clasifica los segmentos como partes del hilo de tutoreo si el ángulo de su componente principal se encuentra entre $\pi/4$ y $3\pi/4$ radianes.

Así mismo se diseña un filtro basado en un coeficiente de elongación, es decir, la relación entre la longitud y anchura del objeto. Se calcula como el cociente entre el máximo y mínimo valor de su eigenvalor (Ec. 13). Este parámetro produce una medida entre $[0, 1]$. Un objeto redondo produce una elongación de 1, mientras que un objeto alargado tiene un valor de 0. El filtro clasifica los objetos como segmentos de hilo de tutoreo si el coeficiente de elongación es menor a 0.1.

Los filtros anteriores tienen la capacidad de discriminar entre la orientación del componente principal de los objetos y la forma de estos, sin embargo, no consideran el tamaño de los objetos en dirección del segundo componente principal (anchura). Debido a que los segmentos de hilo se visualizan como objetos angostos, es decir, el eigenvalor correspondiente al segundo componente principal es pequeño. Los objetos se filtran de tal forma que se clasifican como segmentos de hilo de tutoreo aquellos que su eigenvalor de su segunda componente principal sea menor a 13 píxeles.

En la última sub-etapa se extrae una secuencia de regiones del hilo en dirección al crecimiento del tallo, para ello se sigue un proceso iterativo en el que la entrada es el centroide de cada una de estas. i se define como el i -ésimo centroide y N el número total de elementos. Para el i -ésimo elemento se establece un rango de ± 20 columnas (píxeles), donde si alguno de los $N-i$ centroides se encuentran se incluye en la secuencia. Finalmente se extrae la secuencia de segmentos de hilo con mayor número de elementos, esta secuencia de hilos nos proporcionan información acerca de la posición del tallo que nos interesa monitorear. Se considera que existe suficiente información para estimar la posición del tallo de interés, si más de dos objetos son identificados como segmentos de hilo.

Extracción de segmentos del tallo de interés.

De los apartados anteriores tenemos secuencias de segmentos que conforman posibles tallos e hilos de tutoreo. El objetivo de esta sección es identificar de estos cuál es el verdadero, para ello realizamos una concatenación entre cada una de las secuencias de posibles tallos e de hilo de tutoreo.

Las coordenadas de los centroides de los segmentos que conforman cada uno de los posibles tallos son promediados, de tal forma que para cada uno de ellos tenemos una coordenada u,v (fila, columna) que define una línea vertical.

Una estrategia para identificar el tallo verdadero es elegir el que minimice la discrepancia entre las líneas generadas y los centroides de los segmentos de hilos de tutorio. De tal forma que el objetivo es minimizar la suma de los cuadrados de los residuos (Ec. 15) entre la coordenada del centroide de los segmentos del hilo de tutorio y la recta generada por el promedio de los centroides de las regiones de cada posible tallo.

$$Sr = \sum_{i=1}^n e_i^2 \quad (15)$$

Donde Sr es la suma de los cuadrados de los residuos (píxeles), e es la diferencia entre la coordenada v de la recta obtenida y la coordenada v del centroide de cada segmento del hilo de tutorio (píxeles) y n es el número total de segmentos del hilo de tutorio. Al finalizar esta etapa tenemos los segmentos del tallo de interés, así como los segmentos de hilo que sirven de tutorio a este tallo.

Extracción de coordenadas (u, v) del tallo de interés

En esta etapa los segmentos del tallo de interés obtenidos previamente son conectados para representar un tallo. La estimación de valores intermedios entre los segmentos son obtenidos mediante una interpolación con trazadores cúbicos siguiendo la metodología planteada por Cheney y Kincaid (1985). El objetivo es obtener un polinomio de tercer grado (Ec. 16) para cada intervalo entre puntos.

$$\begin{aligned} f_i(x) = & \frac{f_i''(x_{i-1})}{6(x_i - x_{i-1})} (x_i - x)^3 + \frac{f_i''(x_i)}{6(x_i - x_{i-1})} (x - x_{i-1})^3 \\ & + \left[\frac{f(x_{i-1})}{x_i - x_{i-1}} - \frac{f''(x_{i-1})(x_i - x_{i-1})}{6} \right] (x_i - x) \\ & + \left[\frac{f(x_i)}{x_i - x_{i-1}} - \frac{f''(x_i)(x_i - x_{i-1})}{6} \right] (x - x_{i-1}) \end{aligned} \quad (16)$$

Para esto, por cada segmento de tallo se extraen tres puntos para generar la interpolación, uno corresponde a su centroide y los otros dos son ubicados a 2/6 de la longitud mayor a partir del centroide en dirección a esta longitud. Los puntos obtenidos y los resultantes de la aplicación de los trazadores cúbicos son las coordenadas (u, v) del tallo en la imagen.

4.3.3 Evaluación del rendimiento para el reconocimiento de hilos de tutoreo

El rendimiento para la detección de hilos de tutoreo en imágenes adquiridas en ambientes naturales se evaluó mediante la cuantificación de segmentos clasificados como verdaderos positivos (VP) (Ec. 17).

$$\frac{VP}{VP+FP} \quad (17)$$

Donde FP son segmentos clasificados como de falsos positivos.

4.3.4 Evaluación del rendimiento para el reconocimiento del tallo.

La precisión del reconocimiento de tallos en imágenes adquiridas en invernadero, se determinó mediante la cuantificación de coincidencias entre píxeles segmentados manualmente y los obtenidos mediante el algoritmo propuesto (Ec. 18).

$$\frac{|TA \cap TM|}{|TA|} \quad (18)$$

Donde TM representa las coordenadas (píxeles) del tallo segmentado manualmente en la imagen (tallo verdadero), TA es la coordenada obtenida como resultado del método propuesto. $|\cdot|$ Denota la cardinalidad del conjunto, valores igual a uno representan un reconocimiento perfecto y valores cercanos a cero representan una localización mal realizada. Se calculan la media y la desviación estándar del reconocimiento del tallo.

4.4 Resultados y discusión

En la Figura se presentan algunas imágenes rectificadas con su correspondiente resultado obtenido de la etapa de segmentación. Se obtuvo una calidad media del 89.3% con una desviación estándar de 4.0% evaluado por el método definido por Automatic Target Recognition Working Group (ATRWG). A pesar del buen resultado obtenido se plantea que es posible mejorar la calidad de segmentación a través del uso del canal L.

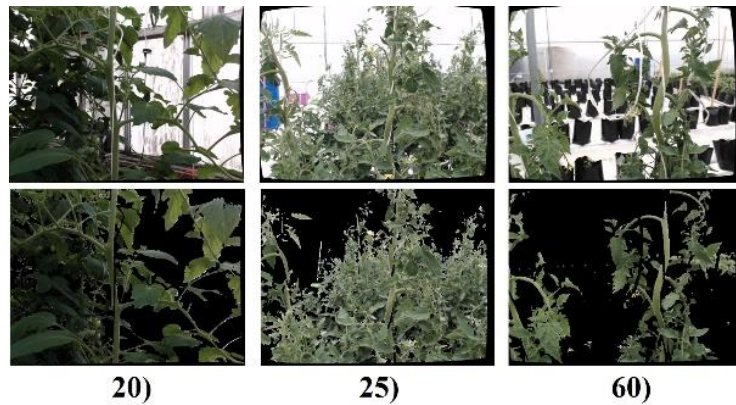


Figura 4. Ejemplo del resultado de la segmentación en imágenes rectificadas.

La salida de cada sub-etapa para la extracción de un posible tallo se presenta en la Figura . En ella se aprecia que el método propuesto muestra buenos resultados, la principal causa de errores se le atribuye al principio para extraer posibles tallos, este se basa en la clasificación de regiones por lo que no considera características de color o textura de la correspondiente región en la imagen RGB, por lo que esto puede ser optimizado en trabajos futuros.

La Figura a) ilustra el mapa de disparidad sobrepuesto en la imagen izquierda, en la Figura b) se presenta el resultado de la sub-etapa para el filtrado por profundidad y la conjunción con la máscara de segmentación. La Figura d) muestra el resultado de la sub-etapa para la extracción de segmentos verticales.

En la etapa de clasificación la Figura e) presenta como a partir de un elemento se extraen todos aquellos que pertenecen a su respectivo segmento base para posteriormente determinar en la Figura f) el elemento superior e inferior con su correspondiente tendencia. Finalmente en la Figura g) se ilustra el resultado después de realizar la búsqueda de segmentos hacia la fila 0 y hacia la fila mayor. En la Figura 7 b) se presenta el resultado de la extracción de segmentos del hilo de tutoreo (segmentos rojos). Se obtuvo una tasa de verdaderos positivos del 60.9% para hilos de tutoreo de color negro. Para los hilos de tutoreo de color blanco adquiridas bajo las mismas condiciones de iluminación se obtuvo una tasa de verdaderos positivos del 89.0%.

Esto representa una mejora de 30% en la detección que es atribuido principalmente a que el hilo de color blanco tiene una reflexión mayor, por lo que en la imagen se aprecia como un objeto brillante. Además la dimensión transversal es mayor por lo que facilita su detección. Bac *et al.* (2014) obtuvo una tasa de verdaderos positivos para la detección de hilos de tutoreo de color blanco del 94% evaluado en chile pimiento y bajo irradiación

moderada, sin embargo este resultado depende del uso de iluminación artificial para reducir el efecto de la iluminación natural. Así mismo, para mejorar la calidad de detección de hilos de tutoreo de color blanco es necesario disminuir el tiempo de exposición de la cámara, sin embargo este es un tema que será tratado en trabajos futuros.

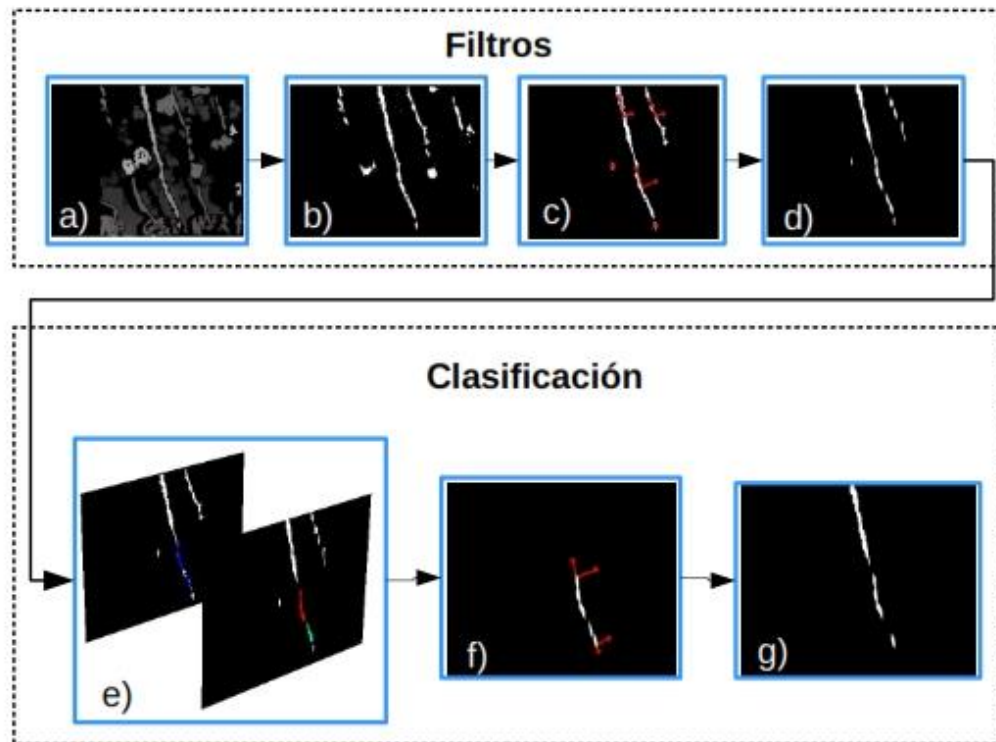


Figura 5. Resultado del algoritmo para extracción de tallos potenciales en el 96a par de imágenes.

En la Figura c) se muestran los segmentos (azul) que componen el tallo potencial que minimiza la diferencia de cuadrados de los residuos entre la posición de este y los segmentos del hilo de tutoreo.

La salida de la etapa para la extracción de puntos que antecede a la aplicación de la interpolación mediante trazadores cúbicos se presenta en la Figura d). Finalmente, en la Figura se muestra la salida del algoritmo aplicado a una serie de imágenes de la base de datos. Donde la tasa de localización del tallo es del 95.0% con una desviación estándar del 9.9%.

Algunas de las cualidades del algoritmo propuesto son: su capacidad para adaptarse a variaciones de iluminación en un entorno natural como lo es el invernadero, basado solamente en un sistema estereoscópico de bajo costo.

Bac *et al.* (2013) propusieron un algoritmo que dentro de sus funciones tiene la tarea de localizar el tallo del pimiento dulce, sin embargo hace uso de un sistema multiespectral con iluminación artificial lo que incrementa el costo de implementación.

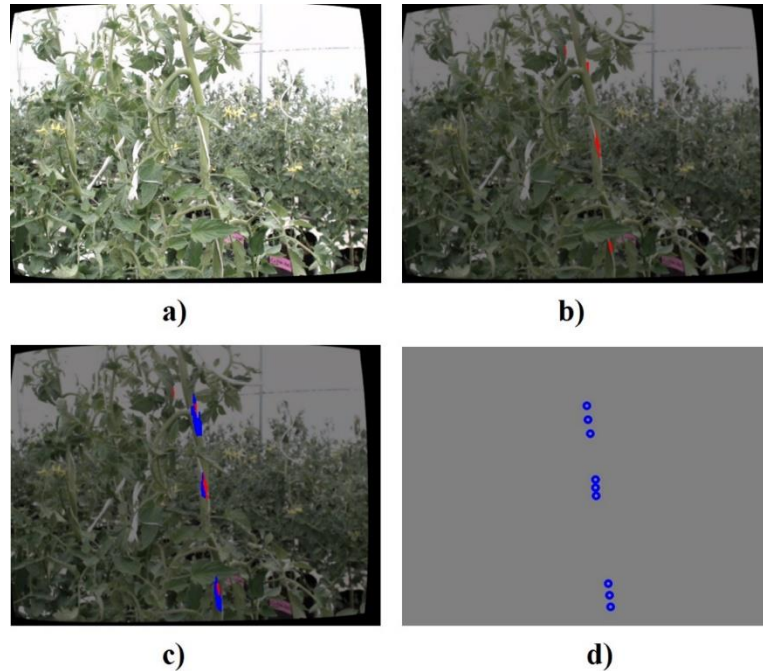


Figura 6. Etapas para la extracción de puntos pertenecientes al tallo en el 36a par de imágenes.

Fernández *et al.* (2013) proponen un algoritmo para discriminar órganos de la vid entre ellos el tallo, sin embargo también hacen uso de imágenes multiespectrales. Bac *et al.* (2014) en su estudio para la localización del tallo del pimiento dulce reportan un rendimiento del 94%. A pesar de esto, dicho rendimiento está en función de la radiación presente en el momento llegando a ser del 74% con una irradiación fuerte, aunado a esto es necesario hacer uso de iluminación artificial. Sin embargo, no fue posible comparar el método propuesto con los anteriores bajo las mismas condiciones de iluminación, debido a que hacen uso de software que requiere del pago de licencias o utilizan metodologías de evaluación basado en regiones y no en píxeles.

En este trabajo se incluye un algoritmo para la extracción de posibles tallos a partir de segmentos del mismo, lo que incrementa los puntos de referencia que se ve reflejado en una mejor detección (Figura). Otra de las cualidades del algoritmo propuesto es la capacidad de localizar el tallo sin la necesidad de realizar algún tratamiento a la planta como la eliminación de hojas. Como se muestra en la Figura (2-60) en casos en los que el tallo se encuentra totalmente expuesto el algoritmo lo detecta en su totalidad. Una desventaja del método propuesto es el tiempo de procesamiento requerido, siendo este en promedio de 260.1 milisegundos para el procesamiento de un par de imágenes estereoscópicas de dimensión 640 x 480 píxeles.

El método propuesto a partir de puntos de referencia pertenecientes al tallo implementa una interpolación mediante trazadores cúbicos, las coordenadas de los píxeles pertenecientes a los puntos de referencia y a las generadas por la interpolación son evaluadas con la segmentación del tallo realizada manualmente. Para disminuir el error es necesario obtener una mayor cantidad de dichos puntos, para ello se tienen que tener en cuenta las siguientes consideraciones; generar un mapa de disparidad más denso. Debido a las características físicas de las cámaras utilizadas para el sistema estereo no fue posible ajustar la distancia base a menos de 5 cm con la finalidad de disminuir las oclusiones generadas por el mismo cultivo. Bac *et al.* (2014) recomienda una línea base de 1 cm para localizar el tallo del pimiento dulce, otra consideración a tener en cuenta es el ajuste del tiempo de exposición de las cámaras, es posible incluir en trabajos futuros un sensor de radiación para tener un parámetro que permita adaptar el tiempo de exposición.

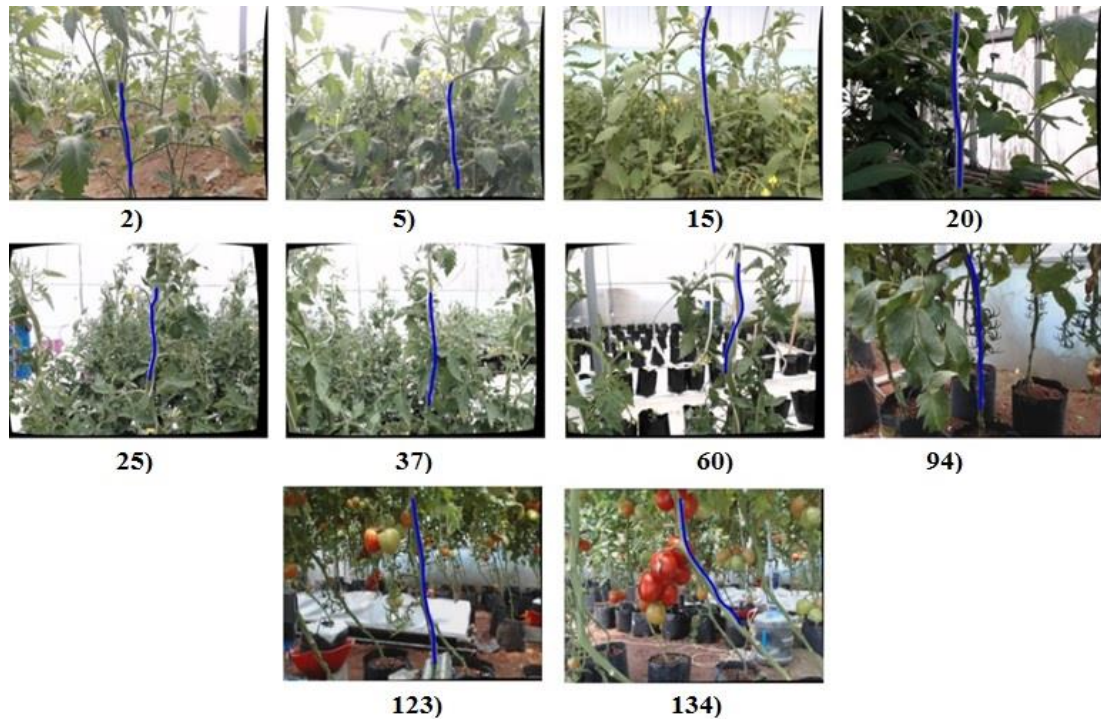


Figura 7. Ejemplos de los resultados de la localización del tallo bajo diferentes condiciones estructurales e iluminación.

4.5 Conclusiones

El algoritmo desarrollado es capaz de localizar el tallo del cultivo usando visión estereoscópica con una línea base de 50 mm, los segmentos del mismo e hilos de tutoreo. Además, incluye características novedosas como lo son la segmentación del cultivo, la extracción de segmentos del tallo, el uso de características invariantes del tallo, la aplicación del PCA para determinar la orientación del tallo y un algoritmo de interpolación mediante trazadores cúbicos para estimar la posición del tallo en lugares donde no se cuenta con información. Se compara el método propuesto con la extracción del tallo realizada manualmente haciendo uso de una base de datos de 134 pares de imágenes, donde la media del rendimiento obtenido para la localización del tallo por el método propuesto es de 95.0%, con una desviación estándar del 9.9%. Es posible mejorar este rendimiento mediante la adaptación del tiempo de exposición y reduciendo la línea base con la finalidad de reducir la oclusión generada por el cultivo.

En general, los resultados del experimento indican que el método propuesto es robusto a las variaciones de iluminación, así como a las condiciones estructurales del invernadero en el que se trabaje. Por lo tanto, es posible aplicar este método en tareas de cosecha, para estimar la posición de los frutos y como primera etapa para desarrollar algoritmos que permitan automatizar las labores de cultivo como es la poda de yemas axilares y hojas en cultivos de crecimiento indeterminado bajo invernadero.

4.6 Literatura citada

- Arefi, A., Motlagh, A. M., Mollazade, K., & Teimourlou, R. F. (2011). Recognition and localization of ripen tomato based on machine vision. [Article]. *Australian Journal of Crop Science*, 5(10), 1144-1149.
- Bac, C. W., Hemming, J., & van Henten, E. J. (2013). Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper. *Computers and Electronics in Agriculture*, 96, 148-162. doi: <http://dx.doi.org/10.1016/j.compag.2013.05.004>
- Bac, C. W., Hemming, J., & van Henten, E. J. (2014). Stem localization of sweet-pepper plants using the support wire as a visual cue. *Computers and Electronics in Agriculture*, 105, 111-120. doi: <http://dx.doi.org/10.1016/j.compag.2014.04.011>
- Barth, R., Hemming, J., & van Henten, E. J. (2016). Design of an eye-in-hand sensing and servo control framework for harvesting robotics in dense vegetation. *Biosystems Engineering*, 146, 71-84. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2015.12.001>
- Bergerman, M., Henten, E. v., Billingsley, J., Reid, J., & Mingcong, D. (2013). IEEE Robotics and Automation Society Technical Committee on Agricultural Robotics and Automation [TC Spotlight]. *IEEE Robotics & Automation Magazine*, 20(2), 20-125. doi: 10.1109/mra.2013.2255513

- Brown, M. Z., Burschka, D., & Hager, G. D. (2003). Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8), 993-1008. doi: 10.1109/tpami.2003.1217603
- Cuevas, E., Zaldívar, D., & Pérez, M. (2010). Procesamiento digital de imágenes con MATLAB y Simulink. *Alfaomega Ra-Ma (México)*.
- Chapra, S. C. C., Chapra, R. P. S. C., Canale, R. P., Chapra, S. C. C., Chapra, R. P. C., Canale, R. P., . . . Raymond, P. (2007). *Métodos numéricos para ingenieros*: McGraw-Hill.
- Cheney, W., & Kincaid, D. (1985). Numerical methods and computing. *Pacific Grove: California*.
- de Pablo, V. J., Moreno, M. C. E., & Mesa, J. C. P. (2003). *Costes de producción y utilización de la mano de obra en tomate: un estudio empírico para el cultivo bajo plástico en Almería*. Paper presented at the Anales de economía aplicada 2003.
- Edan, Y., Han, S., & Kondo, N. (2009). Automation in Agriculture. In Y. S. Nof (Ed.), *Springer Handbook of Automation* (pp. 1095-1128). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Ehret, D., Lau, A., Bittman, S., Lin, W., & Shelford, T. (2001). Automated monitoring of greenhouse crops. *Agronomie*, 21(4), 403-414.
- Fernández, R., Montes, H., Salinas, C., Sarria, J., & Armada, M. (2013). Combination of RGB and multispectral imagery for discrimination of cabernet sauvignon grapevine elements. *Sensors*, 13(6), 7838-7859. doi: 10.3390/s130607838
- Hashimoto, K. (2003). A review on vision-based control of robot manipulators. *Advanced Robotics*, 17(10), 969-991.
- Hemming, J., Bac, C., van Tuijl, B., Barth, R., Bontsema, J., Pekkeriet, E., & Van Henten, E. (2014). A robot for harvesting sweet-pepper in greenhouses.

- Hirschmuller, H. (2008). Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2), 328-341. doi: 10.1109/tpami.2007.1166
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417-441. doi: 10.1037/h0071325
- Kelman, E., & Linker, R. (2014). Vision-based localisation of mature apples in tree images using convexity. *Biosystems Engineering*, 118, 174-185. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2013.11.007>
- Kohan, A., Borghaee, A., Yazdi, M., Minaeid, S., & Sheykhdavudi, M. (2011). Robotic harvesting of rosa damascena using stereoscopic machine vision. *World Appl. Sci. J*, 12(2), 231-237.
- Kurtulmus, F., Lee, W. S., & Vardar, A. (2014). Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. *Precision Agriculture*, 15(1), 57-79. doi: 10.1007/s11119-013-9323-8
- Si, Y., Liu, G., & Feng, J. (2015). Location of apples in trees using stereoscopic vision. *Computers and Electronics in Agriculture*, 112, 68-74. doi: <http://dx.doi.org/10.1016/j.compag.2015.01.010>
- Van Henten, E. J., Van Tuijl, B. A. J., Hemming, J., Kornet, J. G., Bontsema, J., & Van Os, E. A. (2003). Field Test of an Autonomous Cucumber Picking Robot. *Biosystems Engineering*, 86(3), 305-313. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2003.08.002>
- Vitzrabin, E., & Edan, Y. (2016). Adaptive thresholding with fusion using a RGBD sensor for red sweet-pepper detection. *Biosystems Engineering*, 146, 45-56. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2015.12.002>

- Xiang, R., Jiang, H., & Ying, Y. (2014). Recognition of clustered tomatoes based on binocular stereo vision. *Computers and Electronics in Agriculture*, 106, 75-90. doi: <http://dx.doi.org/10.1016/j.compag.2014.05.006>
- Xie, Z., Ji, C., Guo, X., & Zhu, S. (2011). Detection and location algorithm for overlapped fruits based on concave spots searching. [Article]. *Nongye Jixie Xuebao/Transactions of the Chinese Society of Agricultural Machinery*, 42(12), 191-196.
- Yang, L., Dickinson, J., Wu, Q. M. J., & Lang, S. (2007, 4-6 Dec. 2007). *A fruit recognition method for automatic harvesting*. Paper presented at the Mechatronics and Machine Vision in Practice, 2007. M2VIP 2007. 14th International Conference on.
- Yuan, T., Ji, C., Chen, Y., Li, W., & Zhang, J. (2011). Greenhouse cucumber recognition based on spectral imaging technology. [Article]. *Nongye Jixie Xuebao/Transactions of the Chinese Society of Agricultural Machinery*, 42(SUPPL.), 172-176.
- Zhan, W., He, D., & Shi, S. (2013). Recognition of kiwifruit in field based on Adaboost algorithm. [Article]. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*, 29(23), 140-146. doi: 10.3969/j.issn.1002-6819.2013.23.019
- Zhang, B., Huang, W., Wang, C., Gong, L., Zhao, C., Liu, C., & Huang, D. (2015). Computer vision recognition of stem and calyx in apples using near-infrared linear-array structured light and 3D reconstruction. *Biosystems Engineering*, 139, 25-34. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2015.07.011>
- Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), 1330-1334. doi: 10.1109/34.888718

- Zhao, Y., Gong, L., Huang, Y., & Liu, C. (2016a). A review of key techniques of vision-based control for harvesting robot. *Computers and Electronics in Agriculture*, 127, 311-323. doi: <http://dx.doi.org/10.1016/j.compag.2016.06.022>
- Zhao, Y., Gong, L., Huang, Y., & Liu, C. (2016b). Robust Tomato Recognition for Robotic Harvesting Using Feature Images Fusion. *Sensors*, 16(2), 173.
- Zhao, Y., Gong, L., Zhou, B., Huang, Y., & Liu, C. (2016c). Detecting tomatoes in greenhouse scenes by combining AdaBoost classifier and colour analysis. *Biosystems Engineering*, 148, 127-137. doi: <http://dx.doi.org/10.1016/j.biosystemseng.2016.05.001>

CAPÍTULO 5

CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones

El color del cultivo en la imagen adquirida en el interior de un invernadero es muy sensible a la variación de iluminación, también se ve afectado por la configuración de la cámara. Como primera aportación para el desarrollo de un algoritmo que reconozca y localice en el espacio 3D los órganos de la planta de tomate, así como, la implementación de un robot agrícola para la poda, se desarrolla un algoritmo para el reconocimiento del tallo en imágenes adquiridas en el interior un invernadero con iluminación natural.

Se hace uso de un sistema de estéreo visión basado en dos cámaras logitech modelo c920 con una línea base de 50 mm, el cual permite discriminar entre hileras de plantas que se encuentran a diferente distancia de la cámara, este método utiliza componentes novedosos como la implementación de análisis de componentes principales para filtrar regiones y extraer la secuencia de tallos potenciales. Además el uso de trazadores cúbicos para interpolar secciones del tallo en donde no se cuenta con información, y finalmente el uso de los segmentos del hilo de tutoreo para inferir la posición del tallo.

Se evaluó el rendimiento de la detección de hilos de tutoreo para hilos de color blanco y negro, para el primer caso se obtuvo una tasa de verdaderos positivos del 89% y para el segundo del 60.9%. Los experimentos muestran que el hilo de color blanco es más apropiado para la implementación del algoritmo, esto se debe a que tiene una sección transversal mayor.

La calidad media del reconocimiento del tallo se evaluó en 95.0% con una desviación estándar del 9.9%, el tiempo de procesamiento por cada par de imágenes se cuantificó en 260.1 milisegundos. El uso de sistemas multispectrales resulta en altos costos de inversión por lo que no se contemplan en este estudio

Con la finalidad de reducir la información en la imagen RGB capturada, se propone un nuevo método de segmentación de imágenes del cultivo en invernadero, el cual a partir del modelo de distribución de color en los canales a y b del espacio CIE $L^*a^*b^*$ determina umbrales óptimos. Para validar el método propuesto se comparó la media de la calidad de segmentación con 4 enfoques reportados para iluminación natural a campo abierto.

La media de la calidad de segmentación para el método propuesto fue de 93.6%, superando por 20.6%, 32.7%, 3.1% y 23.3% a los índices ExG, ExGR, CIVE y NDI respectivamente. En cuanto a la desviación estándar, esta fue de 2.4% siendo inferior por lo menos en 3% que los otros métodos.

Así mismo, se evaluó la calidad de segmentación en imágenes de mayor resolución (2112 x 2816 píxeles), se obtuvo una media de la calidad de 97.5% con una desviación estándar de 0.97%, aunque la calidad incremento en mas del 3%, el tiempo computacional se incrementó proporcionalmente con el tamaño de la imagen, para las imágenes de tamaño 640 x 480 píxeles se tuvo un tiempo promedio de procesamiento de 9.2 ms, mientras que para las imágenes de tamaño 2112 x 2816 píxeles el tiempo promedio de procesamiento fue de 146.8 ms, esto representa un incremento de aproximadamente 16 veces el tiempo necesario para procesar una imagen de dimensión 640 x 480 píxeles.

Los experimentos muestran también que para la adquisición de imágenes es necesario que el eje óptico de la cámara se ubique con una pequeña inclinación en su ángulo pitch hacia el suelo, esto con la finalidad de reducir la incidencia de la radiación directa en la lente de la cámara.

En general, los resultados del experimento indican que el método propuesto para reconocer el tallo de la planta en la imagen mediante la segmentación del cultivo es robusto a las variaciones de iluminación y se puede aplicar en algoritmos para el reconocimiento de órganos de la planta de cultivos de crecimiento indeterminado en el interior de invernaderos.

5.2 Recomendaciones

La adquisición de imágenes debe hacerse ubicando el eje óptico de la cámara con una pequeña inclinación en su ángulo pitch hacia el suelo esto con la finalidad de reducir la incidencia de la radiación directa en la lente de la cámara.

Las webcam Logitech modelo c920 permiten modificar sus parámetros (tiempo de exposición, distancia focal, balance de blancos y ganancia) directamente en el código principal, mediante la modificación de estos es posible hacer frente a diversos problemas.

El efecto de la iluminación en la imagen puede ser resuelto mediante el ajuste del tiempo de exposición, apoyándose en un sensor de radiación solar. Así mismo, este ajuste tiene la capacidad de acentuar la diferencia entre tonos de color de hojas y del tallo.

Es necesario en trabajos futuros para el reconocimiento del tallo, incluir en el análisis de regiones sus componentes de color. En el algoritmo de segmentación se plantea el análisis para la determinación de umbrales óptimos mediante la fusión de los histogramas del canal a y b del espacio CIE L a*b*.

Apéndice 1. Código para segmentar imágenes del cultivo en invernadero

```
/* **** */
/* SEGMENTACIÓN DEL CULTIVO DE JITOMATE BAJO INVERNADERO */
/* EN CONDICIONES DE ILIMINACIÓN NATURAL */
/* */
/* by Chávez Sánchez Abner Orlando, 2016 */
/* email: os.abner92@gmail.com */
/* */
/* Última modificación: 02/09/2016 */
/* **** */

#include "opencv2/highgui/highgui.hpp"
#include "opencv2/imgproc/imgproc.hpp"
#include <iostream>
#include <math.h>

#define _USE_MATH_DEFINES

using namespace cv;
using namespace std;

void extraccion_normal(vector<Point> &hist_clean, int rango[],
float &et, float &esperanza, float &var);
void integracion(float &integ, float &b, float &n);
void error_estimado_int(float &arg, float &error_est);
void diferenciacion(vector<float> &fun, vector<float> &derivada,
float &h);

/*
 * Dibuja el hisograma de los canales a y b.
 */
void histogramacielab(Mat& cha, Mat& chb){

    int histSize = 256;

    float range[] = { 0, 256 } ;
    const float* histRange = { range };

    bool uniform = true; bool accumulate = false;

    Mat a_hist, b_hist;

    calcHist( &cha, 1, 0, Mat(), a_hist, 1, &histSize, &histRange,
uniform, accumulate );
    calcHist( &chb, 1, 0, Mat(), b_hist, 1, &histSize, &histRange,
uniform, accumulate );

    int hist_w = 512; int hist_h = 400;
```

```

    int bin_w = cvRound( (double) hist_w/histSize );

    Mat histImagecha( hist_h, hist_w, CV_8UC1, Scalar( 0) );
    Mat histImagechb( hist_h, hist_w, CV_8UC1, Scalar( 0) );

    normalize(a_hist, a_hist, 0, histImagecha.rows, NORM_MINMAX, -
1, Mat() );
    normalize(b_hist, b_hist, 0, histImagechb.rows, NORM_MINMAX, -
1, Mat() );

    for( int i = 1; i < histSize; i++ )
    {
        line( histImagecha, Point( bin_w*(i-1), hist_h -
cvRound(a_hist.at<float>(i-1)) ) ,
                Point( bin_w*(i), hist_h -
cvRound(a_hist.at<float>(i)) ) ,
                Scalar(255), 2, 4, 0 );
        line( histImagechb, Point( bin_w*(i-1), hist_h -
cvRound(b_hist.at<float>(i-1)) ) ,
                Point( bin_w*(i), hist_h -
cvRound(b_hist.at<float>(i)) ) ,
                Scalar( 255), 2, 4, 0 );
    }

    namedWindow("Canal a", CV_WINDOW_AUTOSIZE );
    imshow("Canal a", histImagecha );
    namedWindow("Canal b", CV_WINDOW_AUTOSIZE );
    imshow("Canal b", histImagechb );
}

/*
 * Lee la lista de imágenes
 */
static bool readStringList(const string& filename, vector<string>&
l)
{
    l.resize(0);
    FileStorage fs(filename, FileStorage::READ);
    if (!fs.isOpened())
        return false;
    FileNode n = fs.getFirstTopLevelNode();
    if (n.type() != FileNode::SEQ)
        return false;
    FileNodeIterator it = n.begin(), it_end = n.end();
    for (; it != it_end; ++it)
        l.push_back((string)*it);
    return true;
}

```

```

/*
 * Obtiene el Threshold del canal b
 */
void threshold_channel_b(Mat& I, int& posmin){
    int histSize = 256;
    int i, j;
    float range[] = { 0, 256 };
    float* q;
    const float* histRange = { range };
    bool uniform = true; bool accumulate = false;

    vector<Point> hist_clean; //guarda valores de histograma si
ceros

    Mat b_hist;

    calcHist(&I, 1, 0, Mat(), b_hist, 1, &histSize, &histRange,
uniform, accumulate);

    // Extrae histograma sin ceros
    for (i = 0; i < 1; ++i){
        q = b_hist.ptr<float>(i);
        for (j = 0; j < 256; ++j){
            if (q[j] != 0){
                Point coor_hist = Point(static_cast<int>(j),
static_cast<int>(q[j]));
                hist_clean.push_back(coor_hist);
            }
        }
    }
    int posinf = 0;
    for (i = 0; i < (int)hist_clean.size(); i++){
        if (hist_clean[i].x < 129){
            posinf++;
        }
    }
    int rango[2] = {129,255};
    float et;
    float esperanza; // E[x]
    float var; // varianza
    extraccion_normal(hist_clean, rango, et, esperanza, var); //
extraccion de condiciones iniciales
    // cout << "error relativo: " << et << endl;
    // cout << "E[x]: " << esperanza << endl;
    // cout << "Var[x]: " << var << endl;

/*
 * Calculamos mínimos en el rango [128, E[x]]

```

```

    */
    vector<int> minimos;
    for (i = 1; i < (int)hist_clean.size()-1; i++){
        if ((hist_clean[i].x < esperanza) &&(hist_clean[i].x >
128)){
            if (( hist_clean[i].y < hist_clean[i-1].y) &&
(hist_clean[i].y < hist_clean[i+1].y)){
                minimos.push_back(hist_clean[i].x);
            }
        }
    }

    /*contamos numero de niveles a la derecha de la E[x]*/
    int count_bar = 0;
    for (i = 0; i < (int)hist_clean.size(); i++){
        if (hist_clean[i].x > esperanza){
            count_bar++;
        }
    }
    /* para cada minimo encontrado se corre un ajuste de la
distribucion normal
    * la funcion normal que mejor se ajuste entre xi donde xi <
(E[x]- 4 posiciones en hist_clean)
    * y minimo_i sera la seleccionada para establecer un
intervalo de confianza del 99%, este valor
    * sera el threshold buscado*/
    float error_min = 100000;
    float esp_real = 0;
    float var_real = 0;
    if (count_bar > 2){
        if (minimos.size() > 0){
            for (i = 0; i < (int)minimos.size(); i++){
                rango[0] = minimos[i];
                for (j = hist_clean[hist_clean.size()-1].x; j >
hist_clean[hist_clean.size()-1].x - (count_bar - 2); j--){
                    rango[1] = j;
                    extraccion_normal(hist_clean, rango, et,
esperanza, var);
                    if (et < error_min){
                        esp_real = esperanza;
                        var_real = var;
                        error_min = et;
                    }
                }
            }
        }
        else{
            esp_real = esperanza;
            var_real = var;
        }
    }
}

```

```

    }else {
        esp_real = esperanza;
        var_real = var;
    }
    int Ex = round(esp_real);
    int Varx = round(var_real);
    float dsv = sqrt(var_real);
    // cout << "E[x]: " << Ex << endl;
    // cout << "Var[x]: " << Varx << endl;

    /*Calculamos la derivada con diferencia hacia atras en el
    intervalo compuesto por 127 a 255
    * filtro espacial [-1 1]
    */
    float der1;
    vector<Point> der_prim;
    for (i = 1; i < (int)hist_clean.size(); i++){
        if (hist_clean[i].x < Ex){
            der1 = hist_clean[i].y-hist_clean[i-1].y;
            Point der = Point(static_cast<int>(hist_clean[i].x),
                             static_cast<float>(der1));
            der_prim.push_back(der);
        }
    }
    /*ordenar argumentos 1ra derivada*/
    int Temp, Tempx;
    for(int i=0;i<(int)der_prim.size();i++){
        for(int j=0;j<(int)der_prim.size()-1;j++){
            if(der_prim[j].y > der_prim[j+1].y){// cambia
                Temp=der_prim[j].y;
                Tempx=der_prim[j].x;
                der_prim[j].y=der_prim[j+1].y;
                der_prim[j].x=der_prim[j+1].x;
                der_prim[j+1].y=Temp;
                der_prim[j+1].x=Tempx;
            }
        }
    }
    /*Extraemos argmin f'(x) */
    int argmin = 0;
    for (i = der_prim.size()-1; i > -1; i--){
        if (der_prim[i].x < 130){
            argmin = der_prim[i].x;
            break;
        }
    }

```

```

    }
}
if (argmin == 0){
    argmin = der_prim[der_prim.size()-1].x;
}
// cout << "argmin: " << argmin << endl;

/*extraccion de minimos en el rango E[x], argmin */
vector<int> mini_reales;
if (minimos.size() > 0){
    for (i = 0; i < minimos.size(); i++){
        if ((minimos[i] < Ex) && (minimos[i] > argmin)){
            mini_reales.push_back(minimos[i]);
        }
    }
}
/*definicion de threshold*/
if (mini_reales.size() > 0){
    posmin = mini_reales[0];
}else{
    posmin = argmin;
}

} // fin Threshold_channel_b

/*
 * Función para la segmentacion del cultivo
 */
void threshold_channel_a(Mat& I, int& posmin){

    int histSize = 256;
    int i, j;
    float range[] = { 0, 256 };
    float* q;
    const float* histRange = { range };
    bool uniform = true; bool accumulate = false;

    vector<Point> hist_clean; //guarda valores de histograma si
ceros

    Mat a_hist;

    calcHist(&I, 1, 0, Mat(), a_hist, 1, &histSize, &histRange,
uniform, accumulate);

    // Extrae histograma sin ceros
    for (i = 0; i < 1; ++i){
        q = a_hist.ptr<float>(i);

```

```

        for (j = 0; j < 256; ++j){
            if (q[j] != 0){
                Point coor_hist = Point(static_cast<int>(j),
                                         static_cast<int>(q[j]));
                hist_clean.push_back(coor_hist);
            }
        }
    }
    int rango[2] = {0,127};
    float et;
    float esperanza; // E[x]
    float var;       // varianza
    extraccion_normal(hist_clean, rango, et, esperanza, var); //
    extraccion de condiciones iniciales
    // cout << "error relativo: " << et << endl;
    // cout << "E[x]: " << esperanza << endl;
    // cout << "Var[x]: " << var << endl;

    /*
    * Calculamos mínimos en el rango [E[x], 128]
    */
    vector<int> minimos1;
    for (i = 1; i < (int)hist_clean.size()-1; i++){
        if ((hist_clean[i].x > esperanza) &&(hist_clean[i].x <=
128)){
            if (( hist_clean[i].y < hist_clean[i-1].y) &&
(hist_clean[i].y < hist_clean[i+1].y)){
                minimos1.push_back(hist_clean[i].x);
            }
        }
    }

    int count_bar = 0;
    for (i = 0; i < (int)hist_clean.size(); i++){
        if (hist_clean[i].x < esperanza){
            count_bar++;
        }
    }

    /* para cada minimo encontrado se corre un ajuste de la
distribucion normal
    * la funcion normal que mejor se ajuste entre xi donde xi <
(E[x]- 4 posiciones en hist_clear)
    * y minimo_i sera la seleccionada para establecer un
intervalo de confianza del 99%, este valor
    * sera el threshold buscado*/
    float error_min = 100000;
    float esp_real = 0;

```

```

float var_real = 0;
if (count_bar > 2){
    if (minimos1.size() > 0){
        for (i = 0; i < (int)minimos1.size(); i++){
            rango[1] = minimos1[i];
            for (j = 0; j < count_bar - 2; j++){
                rango[0] = j;
                extraccion_normal(hist_clean, rango, et,
esperanza, var);
                if (et < error_min){
                    esp_real = esperanza;
                    var_real = var;
                    error_min = et;
                }
            }
        }
    }else{
        esp_real = esperanza;
        var_real = var;
    }
}
}
}
else {
    esp_real = esperanza;
    var_real = var;
}
}
int Ex = round(esp_real);
int Varx = round(var_real);
//    cout << "E[x]: " << Ex << endl;
//    cout << "Var[x]: " << Varx << endl;

```

```

/*Calculamos 1er derivada con diferecncia hacia atras en el
intervalo compuesto por 127 a 255
* filtro espacial [-1 1]
*/
float der1;
vector<Point> der_prim;
for (i = 1; i < (int)hist_clean.size(); i++){
    if (hist_clean[i].x > Ex){
        der1 = hist_clean[i].y-hist_clean[i-1].y;
        Point der = Point(static_cast<int>(hist_clean[i].x),
                        static_cast<float>(der1));
        der_prim.push_back(der);
    }
}
/*ordemar argumentos 1ra derivada*/
int Temp, Tempx;
for(int i=0;i< (int)der_prim.size();i++){

```

```

        for(int j=0;j<(int)der_prim.size()-1;j++){
            if(der_prim[j].y < der_prim[j+1].y){// cambia
" <" a ">" para cambiar la manera de ordenar
                Temp=der_prim[j].y;
                Tempx=der_prim[j].x;
                der_prim[j].y=der_prim[j+1].y;
                der_prim[j].x=der_prim[j+1].x;
                der_prim[j+1].y=Temp;
                der_prim[j+1].x=Tempx;
            }
        }
    }
    /*Extraemos argmax f'(x) */
    int argmax = 0;
    for (i = 0; i < 5; i++){
        if (der_prim[i].x > 126){
            argmax = der_prim[i].x;
            break;
        }
    }
    if (argmax == 0){
        argmax = der_prim[0].x;
    }
    // cout << "argmax: " << argmax << endl;

    /*
    * Calculamos mínimos en el rango [E[x], argmax f'(x)]
    */
    vector<int> minimos;
    for (i = 1; i < (int)hist_clean.size()-1; i++){
        if ((hist_clean[i].x > Ex) &&(hist_clean[i].x <= argmax)){
            if ((hist_clean[i].y < hist_clean[i-1].y) &&
(hist_clean[i].y < hist_clean[i+1].y)){
                minimos.push_back(hist_clean[i].x);
            }
        }
    }

    /*definicion de threshold*/
    if (minimos.size() > 0){
        posmin = minimos[minimos.size()-1];
    }else{
        posmin = argmax;
    }

} // FIN threshold_channel_a

```

```

void extraccion_normal(vector<Point> &hist_clean, int rango[],
float &et, float &esperanza, float &var){
    esperanza = 0;
    var = 0;
    et = 0;
    /*****cálculo de estadísticas*****/
    float sum_fre_abs = 0; // suma de datos en y del histograma en
rango normal
    int i;
    vector<Point> datosnormal; //vector de datos en el rango de
la distribucion normal
    /*extare vector de datos en el rango de la distribucion normal
    * y calcula suma total de datos en y */
    for (i = 0; i < (int)hist_clean.size();i++){
        if ((hist_clean[i].x >= rango[0])&&(hist_clean[i].x <=
rango[1])){
            datosnormal.push_back(hist_clean[i]);
            sum_fre_abs += hist_clean[i].y;
        }
    }
    float esperanza2 = 0; //E[X^2]
    /*calcula E[x], E[X^2] y suma de cuadrados totales (sst) */
    for (i = 0; i < (int)datosnormal.size(); i++){
        esperanza += (float)datosnormal[i].x *
((float)datosnormal[i].y / sum_fre_abs);
        esperanza2 += pow((float)datosnormal[i].x,2) *
((float)datosnormal[i].y / sum_fre_abs);
    }

    var = esperanza2- pow(esperanza,2); // calcula la varianza de
los datos del hist en rango p/distribucion
    float dsv = sqrt(var); // calcula desviación estandar de los
datos del hist en rango

    vector<Point> dat_estimados;
    /*Estima valores de frecuencia relativa a partir de la
distribucion con E[x] y Var calculadas*/
    for (int i = 0; i < (int)datosnormal.size(); i++){
        float exponente = -1*(1/(2*var)) * pow((datosnormal[i].x -
esperanza),2);
        float fxX = ((1/(dsv*sqrt(2*M_PI))) * pow(M_E,exponente)) *
sum_fre_abs;
        Point est = Point(static_cast<int>(datosnormal[i].x),
static_cast<float>(fxX));
        dat_estimados.push_back(est);
    }
    float ssr = 0; //suma de cuadrados de los residuos
    for (int i = 0; i < (int)dat_estimados.size(); i++){
        ssr += abs(datosnormal[i].y - dat_estimados[i].y);
    }
}

```

```

    }
    et = ssr/sum_fre_abs;
}

/*
 * Segmentación del cultivo
 */
void Segmentacion(Mat& I , Mat& cha, Mat& chb, Mat& seg, Mat&
maskseg) {

    int threshold_val_indice = 0, threshold_valb = 0;

    Mat cha_binarizado, chb_binarizado;

    threshold_channel_a(cha,threshold_val_indice);
    threshold_channel_b(chb,threshold_valb);
    cout << "cha: " << threshold_val_indice << endl <<
        "chb: " << threshold_valb << endl;

    inRange(cha, Scalar(threshold_val_indice),
Scalar(255),cha_binarizado);
    inRange(chb, Scalar(threshold_valb), Scalar(255),
chb_binarizado);

    bitwise_and(~cha_binarizado,chb_binarizado, maskseg);
    I.copyTo(seg, maskseg);
}

/*
 * Función principal
 */
int main()
{
    string imagelistfn = "name_list_images.yaml"; //Lista de
imagenes a leer

    vector<string> imagelist;

    bool ok = readStringList(imagelistfn, imagelist);

    if (!ok || imagelist.empty())
    {
        cout << "can not open " << imagelistfn << " or the string
list is empty" << endl;
    }

    Mat img1;
    int i, j;
    int nimages = (int)imagelist.size();

```

```

for (i = j = 0; i < nimages; i++){
    char filename_mask[80];
    sprintf(filename_mask, "Image_mask%u.tif", i+1);
    char filename_a[80];
    sprintf(filename_a, "Image_a%u.tif", i+1);
    const string& filename = imagelist[i];
    img1 = imread(filename);
    Mat clone_i = img1.clone();
    Mat imgLab, b, a;
    vector<Mat> channels;

    imshow("original", img1);
    GaussianBlur(img1, img1, Size(3, 3), 0, 0); // Aplicar
filtro gaussiano

    cvtColor(img1, imgLab, CV_BGR2Lab);

    split(imgLab, channels);
    a = channels[1];
    b = channels[2];

    //histogramacielab(a,b);    // calcula y muestra el
histograma de los canales a y b

    Mat segmenta(img1.size(), CV_8UC3, Scalar(0,0,0));
    Mat mascaraseg(img1.size(), CV_8UC1, Scalar(0));

    Segmentacion(clone_i, a, b, segmenta, mascaraseg);
    namedWindow("segmenta", CV_WINDOW_NORMAL);
    imshow("segmenta", segmenta);
    j++;
    char c = (char)waitKey(0);
    if (c == 27 || c == 'q' || c == 'Q')
        break;
}
cout << j << " Número de imágenes detectadas.\n";
nimages = j;
return 0;
}

```