



EXTRACCIÓN DE CARACTERÍSTICAS DE LA CUBIERTA VEGETAL DEL VIÑEDO MEDIANTE IMÁGENES RGB Y RGIR OBTENIDAS DE FORMA DINÁMICA

Christian Correa Farias¹, Adolfo Moya-González¹, Eva Báguena Isiegas¹, Ana Herrero Langreo¹, María Paz Diago², Javier Baluja², Javier Tardaguila², Constantino Valero¹, Pilar Barreiro¹.

¹ E.T.S.I. Agrónomos. Departamento de Ingeniería Rural. Universidad Politécnica de Madrid. LPF-TAGRALIA. c.correaf@alumnos.upm.es, pilar.barreiro@upm.es, constantino.valero@upm.es

² Instituto de Ciencias de la Vid y del Vino (Universidad de La Rioja - CSIC - Gobierno de La Rioja) maria-paz.diago@unirioja.es, javier.tardaguila@unirioja.es

Resumen

Diversas investigaciones han intentado resolver el problema de identificación de frutos u hojas mediante imágenes digitales, pero sólo lo han logrado parcialmente. Por esto, el objetivo de este trabajo es explorar una metodología de identificación que permita estimar áreas de hojas y racimos en viñedos, empleando imágenes en el espectro visible (RGB) y en el infrarrojo cercano (RGIR). El problema de la identificación fue abordado por dos vías, forma y color. En el caso de la identificación por forma se empleó la transformada circular de Hough y en el de la identificación por color se emplearon las técnicas de clasificación no supervisada denominadas k-means y Fuzzy c-means. Se determinó que la clasificación mediante k-means sobre el espacio $L^*a^*b^*$, para imágenes RGB y sobre el índice SAVI en las imágenes RGIR, son las técnicas más adecuadas. En cuanto a la identificación por forma, ésta resultó aplicable sólo en condiciones muy particulares.

Palabras Clave: k-means, fuzzy c-means, SAVI.

Abstract

Several studies have attempted to solve the problem of identification of fruits or leaves with digital images, but have only partially achieved. Therefore, the aim of this paper is to explore a methodology to identify and estimate leaves and bunches areas in vineyards, using images in the visible (RGB) and near infrared (RGIR) spectrums. The identification problem was addressed in two ways, shape and color. In the case of the identification by form circular Hough transform was used and in identification by color unsupervised classification techniques called k-means and Fuzzy c-means were used. It was determined that the classification by k-means on the $L^*a^*b^*$ color space for RGB images and the SAVI index in RGIR images are the most appropriate techniques. As for the identification by shape, it was applicable only under very specific conditions.

Key Words: k-means, fuzzy c-means, SAVI

INTRODUCCIÓN

Dotar de visión a una máquina no resulta complejo, ni costoso debido a la masificación de las cámaras digitales y de los ordenadores. Sin embargo, el hacer que dicha máquina tenga la capacidad de identificar lo que está viendo, resulta más complejo de lo que aparenta.

Se han desarrollado diversas técnicas de identificación orientadas principalmente a resolver problemas del ámbito industrial, como por ejemplo, control de calidad en el tamaño o volumen de piezas mecánicas, o ya en el ámbito agrícola, identificación de defectos y tamaño de diversos productos hortofrutícolas (Bjurström *et al.* 2002, Fernández *et al.* 2008, Berenstein *et al.* 2010). Dichas técnicas se basan principalmente en la identificación por el color y la forma del producto en cuestión.

La mayor parte de las aplicaciones de estas técnicas se realizan bajo condiciones controladas (ambiente estructurado) de iluminación, velocidad y distancia al producto, lo que hace que los algoritmos empleados sean válidos sólo bajo esas condiciones. Así, las técnicas clásicas no son aplicables en ambientes no estructurados (Tian *et al.* 1998), por lo que en los últimos años se ha comenzado a emplear técnicas provenientes del área de la inteligencia artificial que permiten un mayor grado de generalización de los entornos y objetos a identificar (Ming *et al.* 2006).

Las técnicas de clasificación por color las podemos separar en supervisadas y no supervisadas (Yixin *et al.* 2009). Las supervisadas son aquellas en donde le señalamos al algoritmo cuántas clases tendremos y cómo será el prototipo de dicha clase. En las no supervisadas, no se conocen a priori los prototipos de las clases, se busca que en cada clase los elementos posean características afines y que sean más similares entre sí, que respecto a elementos pertenecientes a otras clases.

En entornos no estructurados las condiciones son variables, por lo que establecer a priori que características tendrán los elementos de una clase conduce a un sesgo que limita las posibles soluciones. Esto es debido a que las características que se impongan resultarán válidas sólo para situaciones particulares (Kotsiantis, 2007).

En el mundo real, no siempre se conoce la distribución de las clases e incluso, aún conociéndola, es difícil extraer objetos representativos de cada clase para preparar el conjunto de entrenamiento (González, 2010). Debido a esto, las técnicas de clasificación no supervisada resultan de interés para la clasificación en entornos no estructurados, ya que agrupan datos (píxeles en el caso de las imágenes) sin un criterio preestablecido, sólo minimizan la distancia entre píxeles dentro de cada grupo. Restando luego, sólo identificar qué datos han sido agrupados en cada clase.

Nos hemos propuesto emplear técnicas de clasificación no supervisadas para la identificación de uvas y hojas en viñedos mediante imágenes en los espectros visible e infrarrojo.

MATERIAL Y MÉTODOS

Para la adquisición de imágenes, se construyó de un vehículo remolcable con una estructura modular adecuada para la implementación de cámaras y sensores en campo. El remolque fue construido con unas dimensiones y peso adecuados para ser arrastrado mediante 'quads' o tractores viñeros sin pérdida de capacidad de maniobra. La estructura superior emplea perfiles de aluminio que permiten la fijación rápida en todas sus caras mediante. Esto permite la rápida adaptación en campo a las condiciones del cultivo y de los sensores a emplear (altura, distancia al cultivo). En la parte superior del remolque se dispone de una plataforma con elementos de sujeción para portar los PCs empleados para el ajuste, control y almacenamiento de datos procedentes de los distintos sensores. El remolque también dispone de un alojamiento para una batería de 12V la que se conecta un inversor de 12V DC/220V AC con protección por diodos, para la alimentación de los PCs.

Tabla 1: Velocidades reales ($m s^{-1}$) para cada marcha seleccionada en el avance del tractor y tiempos de integración establecidos para ambas cámaras (- combinación no usada)

Marcha seleccionada	Tiempo de integración (ms)					
	50	30	20	10	5	3
4HT	-	0,91	-	0,84	0,81	0,83
2HT	0,42	0,43	0,42	0,41	0,41	0,40
4HC	0,19	0,20	0,19	0,18	-	-
Paradas	0,00	0,00	-	0,00	-	-

Después de optimizar algunos parámetros como la frecuencia de grabación de imágenes o el campo de visión en función de la óptica disponible y las distancias al cultivo, se establecieron los parámetros variables a ensayar en las pruebas de campo sobre un viñedo (*V. vinifera L. cv. Tempranillo*) situado en Ayegui (lat: 42° 39'N; long 2° 03'W, 540 m) La Rioja (España) durante el mes de octubre de 2010, una semana antes la vendimia. La Tabla 1 muestra las velocidades reales (ms^{-1}) para los ensayos realizados a distintas marchas del tractor empleado para arrastrar el vehículo remolcable y para los distintos tiempos de integración seleccionados para las cámaras RGB y RGIR. Para la realización de los ensayos se implementaron en el vehículo remolcable los siguientes equipos de medida: cámara RGB: DuncanTech MS3100-RGB (VIS), cámara RGIR: DuncanTech MS3100-CER (VIS-IR), cámara de vídeo: Sony Handycam digital, GPS diferencial: Garmin 17xHVS y GPS RTK: Leica Zeno 10.

Espacios de color

Disponíamos de un conjunto de imágenes de un viñedo (200) tomadas tanto en el rango visible RGB, (Red, Green, Blue) como en el infrarrojo cercano RGIR (Red, Green, Infrared). Como primera etapa en el procesamiento de las imágenes RGB, estas fueron transformadas a los espacios de color HSV, HSI, CMYK, $L^*a^*b^*$, XYZ y Ohta, (Ohta *et al.*, 1980) para determinar así, cuál de ellos o su combinación es el más adecuado para el proceso de identificación de hojas, racimos, sarmientos y fondo. Esto bajo las premisas de que píxeles de un mismo color pertenecen al mismo tipo de objetos y de que algunos espacios de color diferencian mejor los objetos que otros.

Índices espectrales

Por otra parte las imágenes RGIR fueron transformadas a imágenes en las que cada píxel representa un índice espectral. Los índices empleados fueron: el Índice de Vegetación de Diferencia Normalizada NDVI (Ec.1) y el Índice de Vegetación Ajustado al Suelo (SAVI) (Ec.2; Chuiwieco *et al.*, 2002) ya que estos índices permiten separar vegetación del fondo o suelo.

$$NDVI = \frac{(IR - R)}{(IR + R)} \quad (1)$$

$$SAVI = \frac{(IR - R)}{(IR + R + L)} (1 - L) \quad (2)$$

Donde:

IR es el canal infrarrojo, R el canal rojo.

L corresponde a un factor de corrección para evitar el efecto del brillo del suelo.

Clasificación e Identificación

El problema de la identificación fue abordado por dos vías, forma y color. En el caso de la identificación por forma se empleó la transformada circular de Hough, como la describe Fernández *et al.* 2008 y en el caso de la identificación por color se emplearon las técnicas de clasificación no supervisada denominadas *k-means* y *Fuzzy C-Means* descritas por Ray *et al.*, 1999 y Chuang *et al.*, 2006. Estas metodologías básicamente agrupan píxeles de colores similares, en un número preestablecido de grupos.

Color

K-means Cluster. El algoritmo de k-means está basado en el análisis de las varianzas. Agrupa un conjunto de datos en un número predefinido de clases o grupos. Comienza con un conjunto aleatorio de centroides de cada una de las clases y continúa reasignando los datos del conjunto de píxeles a los distintos centroides, basándose en minimizar la distancia $J(V)$ (Ec.3) entre los píxeles y el centroide. El proceso de reasignación no se detiene hasta que se converge al criterio de parada (se alcanzó un número fijo de iteraciones o los centroides no cambian). Las distancias empleadas fueron Minkowski, Euclidiana, Manhattan y del Coseno.

$$J(V) = \sum_{j=1}^k \sum_{i=1}^n \|X_k - V_j\|^2 \quad (3)$$

Donde

X_k son los vectores de datos, píxeles en nuestro caso;

V_i son los centros de las clases;

$\|X_k - V_i\|$ es una medida de distancia

Fuzzy C-Means (FCM). Este algoritmo de clasificación es una variante del algoritmo k-means cluster. La diferencia entre ambos es que en k-means cada elemento pertenece a un único grupo, mientras que FCM permite la pertenencia parcial de un dato a más de un grupo. Otra diferencia radica en la forma en la que se actualizan los centroides (Ec.4, Ec.5 y Ec.6).

$$J_m(U, V) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \|X_k - V_i\|^2 \quad (4)$$

Donde:

X_k son los vectores de datos, píxeles en nuestro caso;

V_i son los centros de las clases;

$U = [u_{ik}]$ es una matriz de $c \times n$, donde u_{ik} es el i -ésimo valor de pertenencia del k -ésimo vector de datos, que cumple con la condición de

$$\sum_{i=1}^c u_{ik} = 1, \forall k = 1, 2, \dots, n;$$

$m=[1, \infty]$ es un factor de peso que controla el grado de borrosidad de la función de pertenencia.

El centro de cada clase se calcula según la Eq.5:

$$V_i = \frac{\sum_{k=1}^n u_{ik}^m X_k}{\sum_{k=1}^n u_{ik}^m} \quad (5)$$

Y el grado de pertenencia según Ec. 6:

$$u_{ik} = 1 / \sum_{j=1}^c \left(\frac{\|X_k - V_i\|}{\|X_k - V_j\|} \right)^{2/(m-1)} \quad (6)$$

Forma

Transformada circular de Hough. La transformada de Hough comienza obteniendo el mapa de bordes sobre una imagen en escala de grises, en nuestro caso el mapa fue generado empleando las técnicas de Sobel, Prewitt, Roberts y Canny (Heath *et al.* 1998), sobre los canales a^* y b^* del espacio de color $L^*a^*b^*$ y sobre el canal R del espacio RGB. La información obtenida se utiliza para deducir la ubicación de los centros (x, y) y los valores de los radios (r) de los posibles círculos, por último se lleva a cabo la detección de máximos. La desventaja de este método es la alta demanda de recursos computacionales, ya que debe realizar múltiples operaciones en forma recursiva y almacenar los resultados parciales de dichas operaciones (Fernández *et al.*, 2008.).

Estas técnicas fueron aplicadas sobre un conjunto de 200 imágenes RGB de 350x500 y otras 200 RGIR de la misma resolución. Luego procesadas en rutinas desarrolladas en Matlab 7.5, en un ordenador con procesador Pentium 4 ® de 1.6 MHz y 1Gb de RAM.

RESULTADOS Y DISCUSIÓN

De los espacios de color empleados los que mejores resultados muestran son el $L^*a^*b^*$ en particular los canales a^* b^* (magenta-verde y amarillo-azul) y el CMYK, en específico el canal Y (amarillo). Ya que dichos canales permiten separar de mejor modo vegetación, uvas, sarmientos y fondo.

Merecen especial atención los canales L (luminosidad) y K (negro) de los espacios antes mencionados, ya que estos guardan relación con el nivel de intensidad lumínica de los píxeles, lo que resulta útil a la hora de separar o unir conjuntos de píxeles de igual color pero distinta intensidad.

En cuanto a los índices de vegetación NDVI y SAVI, el que mejores resultados entregó fue SAVI con $L=0.5$, ya que permite eliminar el fondo de las imágenes y conservar sólo la vegetación. Mientras que NDVI tiende a confundir píxeles de vegetación con el fondo como se puede apreciar en la Figura 1c. Luego empleando SAVI con valores menores a 0 es posible crear dos grupos, vegetación y no vegetación lo que resulta útil al momento de corregir errores en la clasificación de las imágenes RGB.

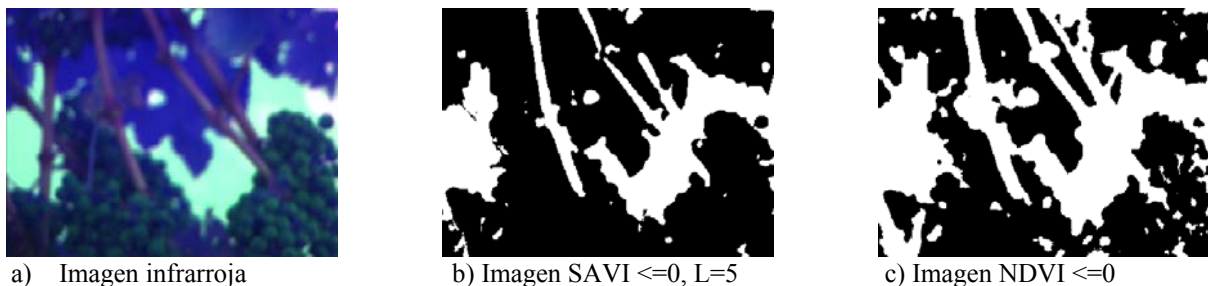


Figura 1. Índices espectrales umbralizados generados partir de la imagen infrarroja.

En lo que se refiere a los algoritmos de clasificación, aplicados sobre imágenes en el espectro visible, k-means, usando ocho clases en el espacio $L^*a^*b^*$ en particular sobre los canales a^*b^* , es capaz de separar adecuadamente las uvas del resto de los componentes de la imagen, así como las hojas (secas y verdes). De las distancias usadas para evaluar k-means, la que mejor desempeño mostró fue la denominada del coseno, ya que las otras tienden a incluir el fondo en la clase que contiene a la uva.

Sin embargo para determinar correctamente el área de hojas o de sarmientos, se debe aumentar la cantidad de clases, ya que como se observa en la Figura 2b, en la Clase 2 se

mezclan píxeles que corresponden a fondo, hojas y sarmientos. Este efecto se reduce al aumentar a 12 clases o al repetir la clasificación sobre la imagen generada por la Clase 2.

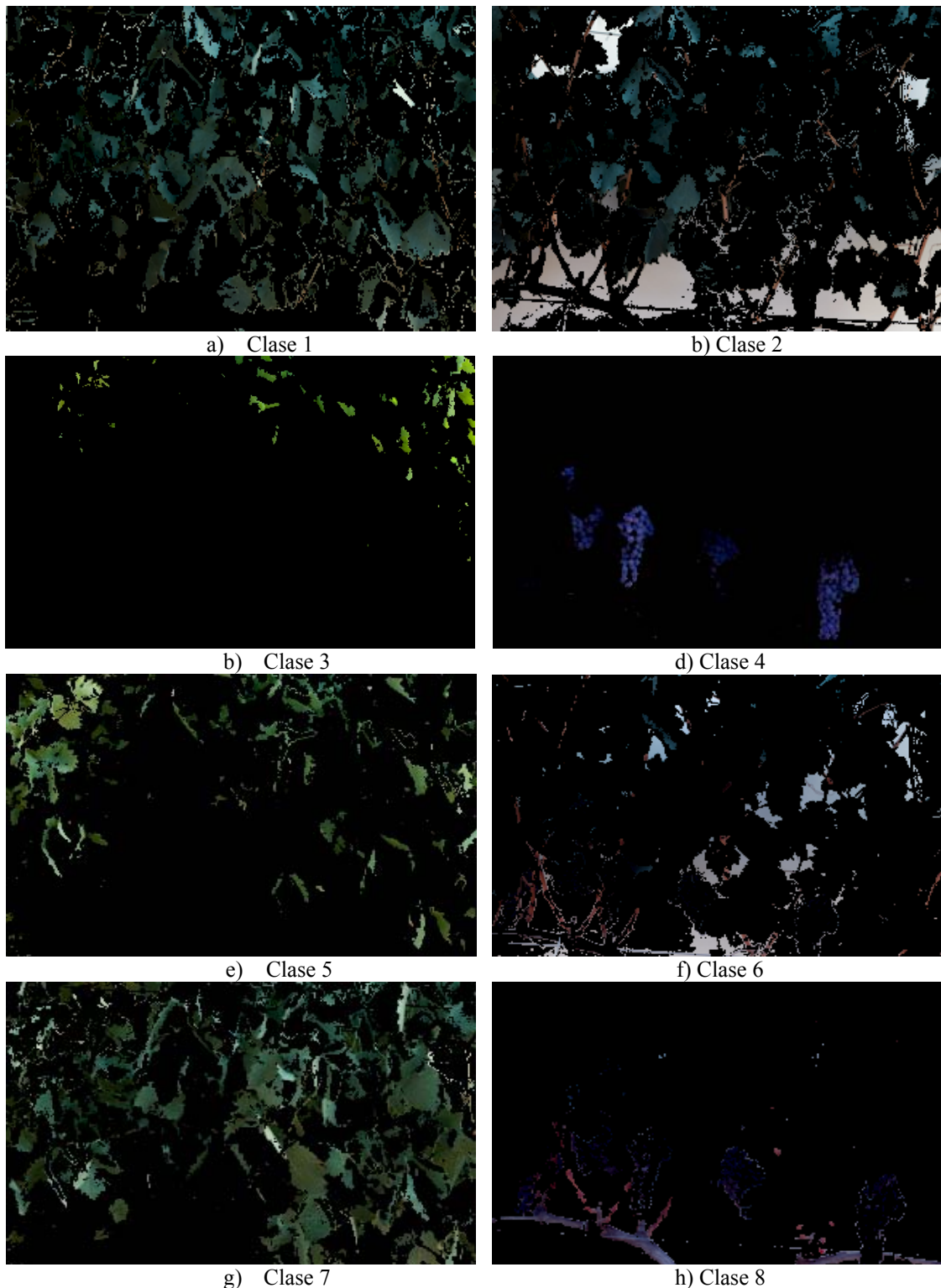
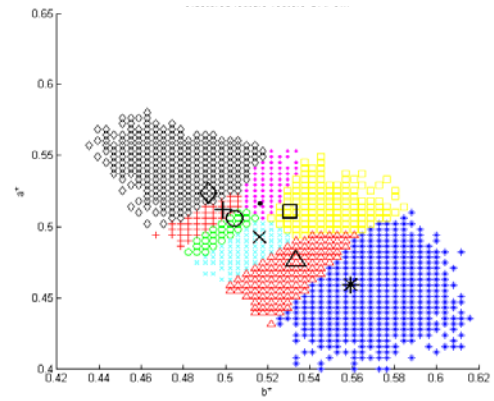


Figura 2. Clases generadas por k-means usando la Distancia del *Coseno* sobre canales a*b*

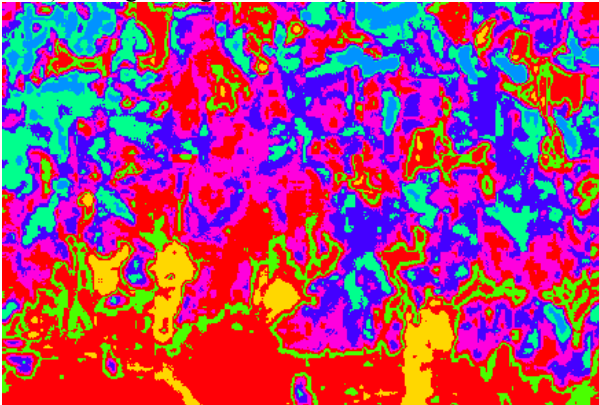
En cuanto a FCM, esta técnica de clasificación mostró mejores resultados que los de k-means, como se puede apreciar en la Figura 3c y 3d, en donde incluso se identifican racimos difíciles de visualizar en la imagen original. Sin embargo, requiere 55 segundos para clasificar una imagen, en comparación con los 9 segundos de k-means.



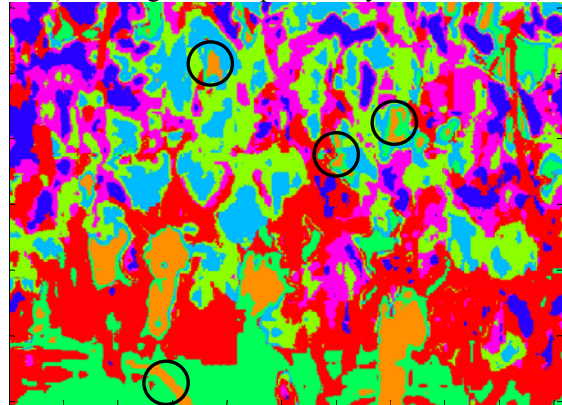
a) Imagen original en el espacio de color RGB



b) Clases generadas por FCM y sus centroides,



c) Clasificación por k-means, ocho clases.



d) Clasificación por FCM, ocho clases.

Figura 3. Resultados obtenidos con los diversos métodos de clasificación. Encerrado en los círculos se observan zonas mal clasificadas por FCM

Para evaluar estos métodos se empleó como referencia, el porcentaje de las áreas de los racimos clasificadas correctamente, Ec. 7, así como el porcentaje del número de racimos detectados correctamente, Ec. 8. En este último caso se eliminó todos aquellos grupos menores a 10 píxeles, ya que pueden ser considerados como ruido.

$$P_A = \frac{\text{Área Real}}{\text{Área detectada}} 100 \quad (7)$$

$$P_N = \frac{\text{Número de racimos Real}}{\text{Número de racimos detectados}} 100 \quad (8)$$

Los resultados pueden ser apreciados en la Tabla 2. Nótese que el porcentaje de racimos identificados correctamente en ambos métodos es igual, ya que una vez eliminados los grupos menores a 10 píxeles, sólo quedan los racimos y parte del tronco que ha sido mal clasificado.

Tabla 2: Desempeño de k-means y FCM como clasificadores.

		K-means	FCM
Porcentaje de área correctamente clasificada	P_A	85.1	90
Porcentaje de racimos identificados correctamente	P_N	87.5	87.5

Con respecto a la clasificación por forma, la transformada circular de Hough no mostró resultados que puedan ser empleados para la identificación, como se puede apreciar en la Figura 4, debido a que la frontera entre bayas es difusa porque comparte el color con sus

vecinas y a la sombra que proyectan entre sí. Estos resultados no mejoran significativamente al aplicar la transformada circular de Hough sobre la imagen de la Clase Uvas generada por k-means o FCM.



Figura 4. Transformada circular de Hough sobre el canal b^* . Busca círculos de radio entre 3 y 9 píxeles, sin tolerancia a incluir círculos concéntricos.

CONCLUSIONES

Si bien es cierto que los resultados mostrados por FCM resultan prometedores, esta técnica no puede ser aplicada en tiempo real (menos de 60ms) por la limitación del tiempo de procesado, por lo que la técnica de k-means resultaría más adecuada para dicho objetivo.

Ambas técnicas mostraron una excelente capacidad de generalización al ser aplicadas a imágenes con diversos grados de iluminación. Pero al mismo tiempo ambas técnicas presentan la desventaja de no ser consistentes en el tiempo, esto es, al aplicar la clasificación a una misma imagen, las etiquetas que le asigna a cada grupo son distintas, por lo que se debe trabajar en desarrollo de algoritmos que garanticen consistencia, o que caractericen cada grupo.

Ambos clasificadores sobreestiman la cantidad de racimos, esto ya que el tronco tiende a ser confundido con los racimos, debido a que poseen píxeles de igual color. Esto puede ser subsanado eliminando regiones que sean más anchas que largas, ya que los troncos generalmente tienen una posición horizontal, mientras que los racimos son por norma más largos que anchos.

Realizando esta modificación, el porcentaje de racimos detectados correctamente P_N sube a un 100%, pero no así el área clasificada correctamente P_A , la que sube apenas de un 85.1% a un 85.8% para k-means y de un 90% a un 90.7% para FCM.

En cuanto a las imágenes IR, estas separan muy bien uvas del resto y en particular del fondo o porosidad, al emplear un umbral adecuado para SAVI. Sin embargo, se debe emplear sobre la imagen SAVI alguna técnica de clasificación, como k-means, para distinguir entre hojas verdes secas y sarmientos.

De ser factible, se deberían emplear cuatro bandas para una clasificación robusta del color o de los índices de vegetación. Esto es, RGB convertido a L^*a^*b junto con el canal IR de la imágenes RGIR.

AGRADECIMIENTOS

Los autores agradecen la financiación del presente trabajo a la COMISIÓN EUROPEA a través del proyecto RHEA. “Robot Fleets for Highly Effective Agriculture and Forestry Management” del 7 Programa Marco (proyecto nº 245986). Se agradece la contribución de todos los participantes en el proyecto: Agencia Estatal Consejo Superior de Investigaciones Científicas - CSIC (Centro de Automática y Robótica, Instituto de Ciencias Agrarias, Instituto de Agricultura Sostenible), CogVis GmbH, Forschungszentrum Telekommunikation Wien Ltd., Cyberbotics Ltd, Università di Pisa, Universidad Complutense de Madrid, Tropical, Soluciones Agrícolas de Precisión S.L., Universidad Politécnica de Madrid - UPM (ETS Ingenieros Agrónomos, ETS Ingenieros Industriales), AirRobot GmbH & Co. KG, Università degli Studi di Firenze, Centre National du Machinisme Agricole, du Génie Rural, des Eaux et des Forêts -CEMAGREF, CNH Belgium NV, CNH France SA, Bluebotics S.A. and CM Srl.

BIBLIOGRAFIA

- Berenstein R., had B. Shahar, Amir Shapiro, and Yael Edan. 2010. *Grape clusters and foliage detection algorithms for autonomous selective vineyard sprayer*. *Intell. Serv. Robot.* 3, 4, pages 233-243.
- Bjurström, Håkan and Jon Svensson. 2002 *Assessment of Grapevine Vigour Using Image Processing*, Master Thesis, Linköping University, Sweden.
- Chuang K., H. Tzeng, S. Chen, J. Wu, T. Chen. 2006. *Fuzzy c-means clustering with spatial information for image segmentation*, *Computerized Medical Imaging and Graphics*, Pages 9-15.
- Chuvieco, E., Martín, M.P. y Palacios, A. 2002. "Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination" *Int. J. of Remote Sensing*, vol. 23, pp. 5103-5110.
- Fathy M., Mohsen Sharifi and M. Tayefeh Mahmoudi, “An Experimental Comparison of Commonly used Edge Detectors”, *The 2002 International Conference on Imaging Science, Systems, and Technology (CISST’02), The 2002 International Multi-conference on Computer Science*, CSREA Press, Las Vegas, USA, June 24-27, 2002.
- Fernández Leandro A.F., Manuel M. Oliveira. 2008. *Real-time line detection through an improved Hough transform voting scheme*. *Pattern Recognition*, Volume 41, Issue 1, Pages 299-314.
- González, D.P. 2010. *Algoritmos de agrupamiento basados en densidad y validación de clusters*. *Phd thesis. Universitat Jaume I*.
- Heath, M.; Sarkar, S.; Sanocki, T. & Bowyer, K. *Comparison of Edge Detectors: A Methodology and Initial Study* *Computer Vision and Image Understanding*, 1998, 69, 38 – 54.
- Kotsiantis, S. B. 2007. *Supervised Machine Learning: A Review of Classification Techniques*. *Informatics*, 31, 249-268.
- Tian L., David C. Slaughter. 1998. *Environmentally adaptive segmentation algorithm for outdoor image segmentation*, *Computers and Electronics in Agriculture*, Volume 21, Pages 153-168.
- Ming Gao; Tien-Fu Lu; 2006. "Image Processing and Analysis for Autonomous Grapevine Pruning," *Mechatronics and Automation*, Proceedings of the 2006 IEEE International Conference, vol., no., pp.922-927.

Ohta Yu-Ichi, Tokeo Kamode.1980. *Color information for region segmentation*. [J]. *Computer Graphics and Image Processing* 13.Pages 222-241.

Ray S. and R H Turi: 1999. *Determination of number of clusters in K-means clustering and application in color image segmentation*, *Proceedings of ICAPRDT'99*, pp 137-143.

Yixin Yan, Yongbin Shen, and Shengming Li. 2009. *Unsupervised Color-Texture Image Segmentation Based on A New Clustering Method*. In *Proceedings of the 2009 International Conference on New Trends in Information and Service Science (NISS '09)*. IEEE Computer Society, Washington, DC, USA, 784-787.