

# Reconstrucción 3D de racimos de uva basada en estéreo-visión

Carlos Alberto Molina Toscano  
Universitat Politècnica de València, carmotos@etsii.upv.es

Antonio José Sánchez Salmerón  
Universitat Politècnica de València, asanchez@isa.upv.es

Eugenio Ivorra Martínez  
Universitat Politècnica de València, euivmar@upvnet.upv.es

## Resumen

*El tamaño de las bayas de uva es uno de los parámetros fundamentales que determinan la calidad de la uva de vino. Recientemente en la literatura se han propuesto métodos de estimación de tamaño de las bayas utilizando sistemas monoculares junto con objetos de calibración. Sin embargo en este trabajo se describe un procedimiento automático para la reconstrucción 3D de racimos de uvas a partir de pares estéreo de imágenes. El método propuesto se basa en la correspondencia de los círculos detectados por medio del algoritmo Hough y la resolución del problema de correspondencias utilizando restricciones epipolares, de gradientes y de color.*

**Palabras Clave:** estereovisión, reconstrucción 3D, detección círculos.

## 1. Introducción

La calidad del vino depende principalmente de la calidad de los granos de uva con los que está hecho. Las prácticas actuales para estimar la calidad de la uva requieren costosos análisis destructivos o análisis subjetivos de expertos. Desafortunadamente, no solo la subjetividad de los análisis, sino también el corto periodo de tiempo disponible para llevar a cabo estos análisis durante la cosecha, la falta de herramientas de medida y su alto coste económico, hacen difícil medir la calidad de la uva.

El tamaño de las bayas de uva es uno de los parámetros fundamentales que determinan la calidad de la uva de vino, ya que dicho tamaño establece una relación entre la cantidad de piel y de pulpa del racimo. Recientemente se ha publicado un trabajo de investigación orientado a la estimación del tamaño de bayas de uva de vino [1], donde se detectan círculos sobre una imagen y se clasifican los granos de uva. En este trabajo se estima el tamaño de las bayas de uva mediante un sistema monocular y la inserción de un objeto de calibración de tamaño conocido en la

escena. Sin embargo sería interesante utilizar un sistema estereoscópico, evitando así la necesidad de tener que insertar un objeto de calibración conocido.

Un racimo de uva tiene una geometría compleja ya que está formado por un aglomerado de cuasi-esferas donde se producen muchos ocultamientos y recovecos. De ese modo, el control de calidad automático de los racimos mediante visión artificial resulta un problema complejo de resolver. Sin embargo existen muchos métodos que permiten extraer información 3D a partir de imágenes 2D. Estos métodos se pueden clasificar en activos y pasivos. Los métodos activos [2, 3] son aquellos que emiten un patrón de luz sobre la escena para analizar su deformación, mientras que los métodos pasivos, como la estereoscopia [4], solo utilizan la iluminación ambiente.

Aunque la estereoscopia es muy simple, dentro del ámbito de la inspección de alimentos, no se han realizado muchos trabajos utilizando esta técnica. Esto se debe principalmente a que el problema de correspondencia es difícil de resolver. Algunos trabajos interesantes que utilizan la estereoscopia son para, detectar malezas [5], medir el tamaño de granos de trigo [6], estimar la firmeza en salmón [7] y medir el volumen de rodajas de manzana durante el proceso de secado [8]. En todos estos casos se utilizaron los bordes como características para resolver el problema de correspondencias.

Con el objetivo de determinar automáticamente el tamaño de las bayas de uva, en este trabajo se presenta un procedimiento estereoscópico basado en la correspondencia de los círculos detectados en el par de imágenes por medio del algoritmo Hough [9] y la resolución del problema de correspondencias utilizando restricciones epipolares, de gradientes y de color.

El procedimiento se divide en 3 fases: procesado de imagen, búsqueda de correspondencias y reconstrucción 3D.

Para llevar a cabo cada una de las fases se utiliza la herramienta software MATLAB.

## 2. Procesado de imágenes

Se parte de un par de imágenes RGB capturadas con un sistema estereoscópico Bumblebee2 (Point Grey Research Inc, Richmond, BC, Canada) modelo BB2-08S2C-25, las cuales han sido obtenidas en condiciones de laboratorio con un fondo controlado. Inicialmente se corrige la distorsión de las imágenes capturadas y también se rectifican a partir de los datos de calibración de las cámaras [10], [11] y [12].

A continuación se detalla los pasos a seguir, a partir del par estéreo de imágenes RGB rectificadas de los racimos (Figura 1).

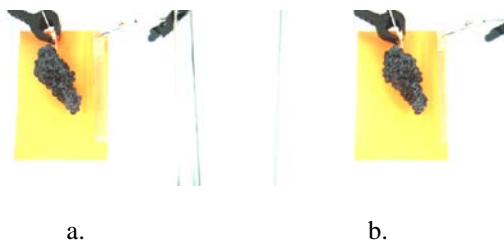


Figura 1: Imágenes RGB rectificadas. a. Izquierda. b. Derecha

### 2.1. Segmentación

Las imágenes han sido tomadas bajo condiciones de laboratorio, por lo tanto, se ha forzado a que haya un fondo de imagen que contraste, claramente, con el racimo de uvas. De esta manera el racimo se segmenta fácilmente del fondo, utilizando, en primer lugar, umbrales de color en R, G y B y posteriormente el algoritmo de Otsu [13] sobre las imágenes resultantes en escala de grises, obteniéndose así la máscara del racimo, Figura 2.a.

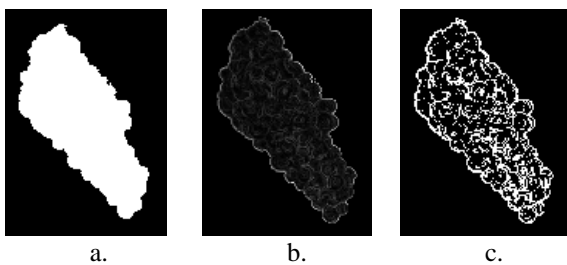


Figura 2: a. Máscara del racimo utilizando el método Otsu. b. Magnitud del gradiente. c. Bordes. Píxeles cuya magnitud de gradiente ha superado un umbral.

### 2.2. Estimación de círculos

El aspecto más importante a resolver para la estimación de los círculos, que representan las bayas, es la detección de bordes; ya que, dependiendo de la calidad de éstos (definición y cantidad de píxeles), la estimación puede ser más fiable o no.

#### 2.2.1. Bordes

Para la detección de bordes se ha utilizado un detector gradiente discreto ( $G$ ), empleado tanto en dirección horizontal ( $g_x$ ) como vertical ( $g_y$ ), aprovechando el valor de la magnitud.

$$|\nabla G| = \sqrt{\left(\frac{\Delta G}{\Delta x}\right)^2 + \left(\frac{\Delta G}{\Delta y}\right)^2} = \sqrt{g_x^2 + g_y^2} \quad (1)$$

El gradiente se aplica sobre la imagen de grises enmascarada, Figura 2.b. El gradiente se umbraliza para obtener aquellos bordes con magnitud más elevada, que implica que tienen más probabilidad de ser píxeles de borde, Figura 2.c.

#### 2.2.2. Detección de círculos

Se ha implementado el algoritmo Hough [9] para detectar los círculos en cada una de las imágenes, ver Figura 4. Este algoritmo se divide en 3 fases: acumulación en el espacio de parámetros, detección de picos y filtrado de resultados.

##### Acumulación

El espacio de parámetros define todos las posibles combinaciones que pueden tener los parámetros que definen un círculo: posición del centro ( $X$ ,  $Y$ ) y radio ( $R$ ). Se crea una matriz  $M_{m \times n \times p}$ , donde cada  $M(x, y, r)$  corresponde a la posición  $x$  e  $y$  del centro del círculo de radio  $r$ .

La matriz tiene sus celdas inicializadas a cero, cada vez que se encuentre un círculo de radio  $k$  y centro  $(i, j)$  el valor la celda  $M(i, j, k)$  se incrementa en una unidad, esto se hace con cada una de las opciones que se definen en el espacio de parámetros para cada característica de borde.

##### Detección de picos

Una vez que la matriz  $M$  está completada, se buscan aquellas celdas que tengan un mayor valor, lo que indica que hay más probabilidades de que sea un círculo. Se ordenan de mayor a menor probabilidad y se escoge un número fijo de círculos a detectar, que además, cumplan unos criterios de localización.

Se escoge un número lo suficientemente alto de círculos a detectar, se toman los 50 mejores, y como criterio de localización se ha determinado que no puede haber un solapamiento de círculos de más del 50%, es decir, la separación entre centros debe ser como mínimo del radio de la baya.

Cabe resaltar que debido a que normalmente los círculos de menor radio tienen mayor probabilidad de ser detectados que los de radio mayor, se ha optado por separar el proceso de detección por radios, es decir, se tomarán los 50 primeros círculos para un

determinado radio y el criterio de localización se cumplirá para estos. Por lo tanto, al unir todas estas detecciones, puede que haya varios círculos que estén solapados más allá del 50%.

### Filtrado

Para descartar, en la medida de lo posible, los falsos positivos de círculos detectados, se realiza un filtrado. El filtrado permite reducir, el coste computacional y el error en la búsqueda de correspondencias.

$$\theta = \tan^{-1} \left( \frac{g_y}{g_x} \right) \quad (2)$$

El filtrado se basa en la comparación de la orientación del gradiente (2) del círculo detectado sobre la imagen en escala de grises y un patrón ideal del mismo radio (Figura 3).

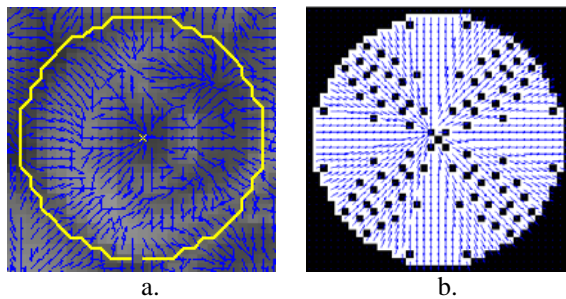


Figura 3. Comparación de la orientación del gradiente. a. Círculo detectado. b. Patrón

Si la discrepancia entre las orientaciones detectadas y las ideales supera un umbral, la detección de dicho círculo se descarta.

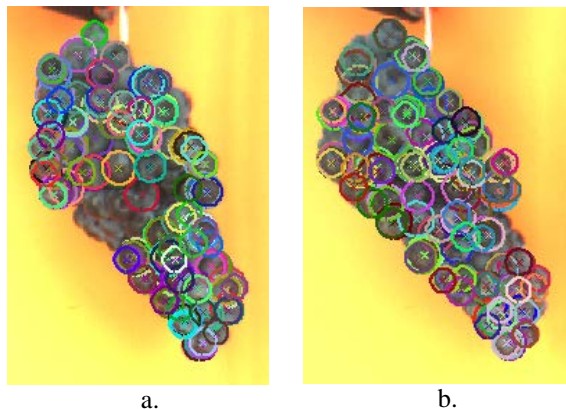


Figura 4: Círculos detectados con Hough. a. Izquierda. b. Derecha.

### 3. Correspondencias

La propuesta que se hace para resolver el problema de las correspondencias es definir dos índices, uno

asociado a la correspondencia que tienen cada par de círculos (coste de correspondencia ( $c_{ij}$ )), y otro asociado al ocultamiento del círculo (coste de ocultamiento ( $o_{ij}$ )).

Para establecer los índices, se mide la discrepancia que hay entre pares (izquierdo y derecho) de los círculos detectados. Los factores que determinan el valor (coste) son: características de color, disparidad y orientación del gradiente.

$$c_{ij} = c_{ijDISPARIDAD} + c_{ijGRADIENTE} + c_{ijCOLOR} \quad (3)$$

$i$ , círculo en la imagen izquierda.

$j$ , círculo en la imagen derecha.

$c_{ijDISPARIDAD}$ , disparidad entre centros  $i$  y  $j$ .

$c_{ijGRADIENTE}$ , error de comparación de la orientación del gradiente entre  $i$  y  $j$ .

$c_{ijCOLOR}$ , índice de comparación de color entre  $i$  y  $j$ .

$$o_{ij} = o_{ijDISPARIDAD} + o_{ijGRADIENTE} + o_{ijCOLOR} \quad (4)$$

Para reducir el número de posibles pares candidatos, se define un ancho de banda de búsqueda, donde, los círculos que caen dentro de ésta banda generan dos matrices  $m \times n$ , siendo  $m$  es el número de círculos detectados dentro de la banda en la imagen izquierda y  $n$  es el número de círculos detectados dentro de la banda en la imagen derecha.

En la primera de las matrices (matriz de coste correspondencia) cada celda alberga el valor del coste de correspondencia  $c_{ij}$ , ver Tabla 1. La segunda matriz (matriz de coste de ocultamiento) almacena en cada celda el valor del coste de ocultamiento  $o_{ij}$ , ver Tabla 2.

Tabla 1: Matriz de costes de correspondencia

	1	2	3	...	n
1	$c_{11}$	$c_{12}$	$c_{13}$		$c_{1n}$
2	$c_{21}$	$c_{22}$	$c_{23}$		$c_{2n}$
...					
m	$c_{m1}$	$c_{m2}$	$c_{m3}$		$c_{mn}$

Tabla 2: Matriz de costes de ocultamiento

	1	2	3	...	n
1	$o_{11}$	$o_{12}$	$o_{13}$		$o_{1n}$
2	$o_{21}$	$o_{22}$	$o_{23}$		$o_{2n}$
...					
m	$o_{m1}$	$o_{m2}$	$o_{m3}$		$o_{mn}$

El objetivo es encontrar un camino que vaya desde la posición (1, 1) hasta la (m, n) que minimice el coste total (suma de coste de las celdas de cada camino). El coste de cada celda es el mínimo entre su coste de correspondencia y su coste de ocultamiento. Para

hacer la búsqueda más eficiente se han ordenado los círculos de izquierda a derecha, y siguiendo un criterio de posicionamiento horizontal (una baya no puede estar arriba o debajo de otra), cuando se considera que un círculo encuentra su homólogo, se pasa al círculo de la siguiente posición, ya que no pueden existir dos círculos visibles con el mismo círculo homólogo.

Para resolver el problema de encontrar el camino de mínimo coste, se puede utilizar un algoritmo como el dijkstra [14], el resultado debería ser como el de la Tabla 3. En nuestro caso se utiliza programación dinámica.

Tabla 3: Matriz con camino de mínimo coste

	1	2	3	...	n
1	$o_{11}$	$c_{12}$			
2			$o_{23}$	$c_{2i}$	
...					$o_{jn}$
m					$c_{mn}$

## 4. Reconstrucción 3D

### 4.1. Triangularización

A partir de las parejas de círculos (proyecciones) resultantes de aplicar las correspondencias y, teniendo en cuenta las características intrínsecas (2) y extrínsecas (3) de la estéreo cámara se estima un modelo 3D de la posible esfera generada.

$$M = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$\alpha_x = \frac{f_x}{d_x}$ , distancia focal horizontal en píxeles.

$\alpha_y = \frac{f_y}{d_y}$ , distancia focal vertical en píxeles.

$f = f_x = f_y$ , distancia focal en mm.

$d_x, d_y$ , tamaño del píxel en x e y.

$x_0, y_0$ , punto principal expresado en píxeles.

$$RT = \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ O_{1 \times 3} & 1 \end{bmatrix} \quad (3)$$

$R_{3 \times 3}$ , matriz de rotación.

$T_{3 \times 1}$ , matriz de traslación.

En la generación del modelo 3D se tiene en cuenta la disparidad ( $d$ ) y la línea base ( $b$ ) de la cámara, con la que se estima la profundidad (5), del centro de la baya.

$$d = x_{izq} - x_{der} \quad (4)$$

$x_{izq}, x_{der}$ , posición del centro de la proyección izquierda y derecha en x.

$$Z = \frac{b \cdot f}{d} \quad (5)$$

$Z$ , profundidad del centro de la baya en coordenadas del mundo real.

El problema de reconstruir una esfera a partir de sus proyecciones tiene solución única, por lo que para calcular el resto de coordenadas ( $X, Y$ ) del centro de la baya, basta con calcular el factor de escalado  $\lambda$ .

$$\lambda = \frac{Z}{f} \quad (6)$$

Así,

$$X = \lambda \cdot x \quad Y = \lambda \cdot y \quad (7)$$

$x, y$ , centro de la baya en coordenadas de imagen.

En modelo 3D del racimo de uvas obtenido, Figura 8, se puede apreciar como existen bayas que son incoherentes, tienen una posición anómala absoluta y/o relativa con el resto de bayas, con el resultado que se esperaría. Por lo tanto se aplica una etapa final de refinamiento.

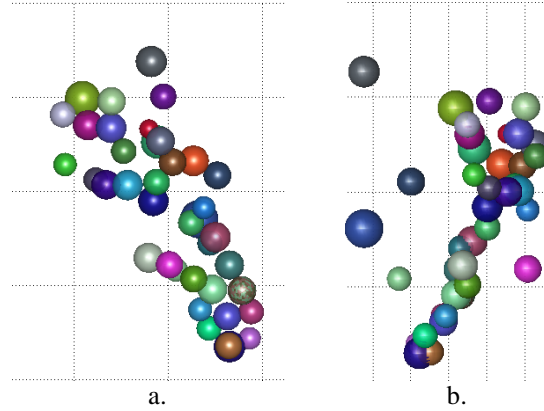


Figura 5: Racimo reconstruido. a. Alzado. b. Perfil

### 4.2. Refinamiento

El conocimiento, a priori, que se tiene de la morfología del racimo de uvas, puede ser aplicado para realizar un refinamiento final. Con el fin de obtener una reconstrucción de bayas que tenga más coherencia (el posicionamiento de las bayas sea acorde a lo que se esperaría de un racimo), es decir, la profundidad sea similar a la de las bayas aledañas, y estas a su vez deben formar una estructura similar a una colmena de abejas, como se observa claramente en la zona inferior de la Figura 4.a. Siguiendo estas premisas tal como se muestra en la Figura 4, la gran

mayoría de la bayas incoherentes han sido eliminadas del modelo.

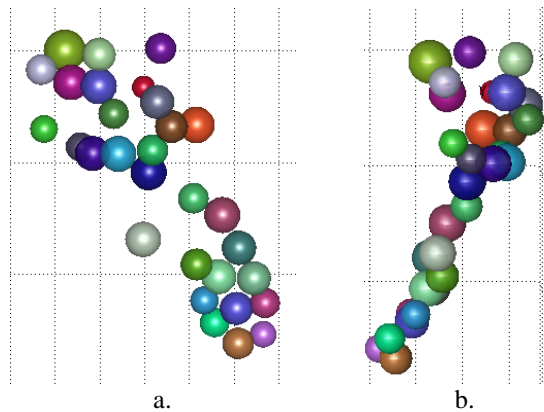


Figura 6: Racimo reconstruido refinado. a. Alzado. b. Perfil.

## 5. Resultados y conclusiones

Se han reconstruido 100 racimos de uva de 10 variedades distintas, obteniéndose un ratio medio de bayas correctamente reconstruidas del 25% (aproximadamente 15 bayas) con un error de precisión en la reconstrucción menor a 1 mm.

Teniendo en cuenta que en el proceso manual de caracterización del tamaño medio de bayas de un racimo se utiliza un muestreo de unas quince bayas. Se puede concluir que se ha reconstruido una cantidad suficiente de bayas que permiten estimar el tamaño de baya medio del racimo de forma automática.

Uno de los aspectos que mayor importancia tiene en la reconstrucción de un modelo 3D es la obtención de bordes, cuanto más definidos estén mejor será la detección de círculos y por tanto mejor será la estimación del modelo 3D.

Otros aspectos que tienen también importancia son la calibración de la cámara y la definición de los umbrales de detección.

Se han definido umbrales fijos, pero idealmente deberían ser dinámicos para que los bordes que no tengan una elevada magnitud pero que se hallen en zonas donde la magnitud media sea baja, no sean descartados.

### Agradecimientos

Este trabajo ha sido parcialmente financiado por el Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria de España (INIA) a través del proyecto RTA2012-00062-C04-02, soportado por

fondos europeos FEDER y el proyecto UPV-SP20120276.

### Referencias

[1] Roscher, R., Herzog, K., Kunkel, A., Kicherer, A., Töpfer, R., & Förstner, W. (2014). Automated image analysis framework for high-throughput determination of grapevine berry sizes using conditional random fields. *Computers and Electronics in Agriculture*, 100, 148–158.

[2] Udomkun, P., Nagle, M., Mahayothee, B., & Müller, J. (2014). Laser-based imaging system for non-invasive monitoring of quality changes of papaya during drying. *Food Control*, 42(0), 225–233. doi:10.1016/j.foodcont.2014.02.010

[3] Verdú, S., Ivorra, E., Sánchez, A. J., Girón, J., Barat, J. M., & Grau, R. (2013). Comparison of TOF and SL techniques for in-line measurement of food item volume using animal and vegetable tissues. *Food Control*, 33(1), 221–226. doi:10.1016/j.foodcont.2013.02.031

[4] Cyganek, B., & Siebert, J. P. (2011). An introduction to 3D computer vision techniques and algorithms. John Wiley & Sons.

[5] Saez, A. J., & Marchant, J. A. (2000). Fusing 3D information for crop/weeds classification. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on (Vol. 4, pp. 295–298 vol.4)*. doi:10.1109/ICPR.2000.902917

[6] Sun, C., Berman, M., Coward, D., & Osborne, B. (2007). Thickness measurement and crease detection of wheat grains using stereo vision. *Pattern Recognition Letters*, 28(12), 1501–1508. doi:10.1016/j.patrec.2007.03.008

[7] Quevedo, R., & Aguilera, J. M. (2010). Computer Vision and Stereoscopic for Estimating Firmness in the Salmon (Salmon salar) Fillets. *Food and Bioprocess Technology*, 3(4), 561–567. doi:10.1007/s11947-008-0097-3

[8] Sampson, D. J., Chang, Y. K., Rupasinghe, H. P. V., & Zaman, Q. U. (2014). A dual-view computer-vision system for volume and image texture analysis in multiple apple slices drying. *Journal of Food Engineering*, 127(0), 49–57. doi:10.1016/j.jfoodeng.2013.11.016

[9] D. Young, «Matlab Central. MathWorks Inc.» 2010. [Online]. Available: <http://www.mathworks.com/matlabcentral/fileexchange/26978-hough-transform-for-circles>. [Último acceso: 16 07 2014].

[10] Ricolfe-Viala, Carlos; Sanchez-Salmeron, Antonio-Jose; Camera calibration under optimal conditions, *Optics Express*, 19, 11, 10769-10775, 2011, Optical Society of America.

[11] Ricolfe-Viala, Carlos; Sanchez-Salmeron, Antonio-Jose; Martinez-Berti, Enrique; Accurate calibration with highly distorted images, *Applied optics*, 51, 1, 89-101, 2012, Optical Society of America.

[12] Sánchez, Antonio; Calatayud, Ricardo; ,Auto-calibración de un sistema binocular de visión activa, *Jornadas de Automática* 2001.

[13] Otsu, N., "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 9, No. 1, 1979, pp. 62-66.

[14] Dijkstra, E.W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik* 1, 269–271.