

## Detección de podredumbres en cítricos mediante análisis de imágenes *backscattering*

D. Lorente<sup>1</sup>, M. Zude<sup>2</sup>, J. Gómez-Sanchis<sup>3</sup>, C. Regen<sup>2</sup> y J. Blasco<sup>1</sup>

<sup>1</sup> Centro de Agroingeniería. Instituto Valenciano de Investigaciones Agrarias (IVIA). Cra. Moncada-Náquera km 5, 46113, Moncada (Valencia). E-mail: delia.lorente.g@gmail.com.

<sup>2</sup> Leibniz-Institute for Agricultural Engineering Potsdam-Bornim (ATB). Max-Eyth-Allee 100, 14469, Potsdam-Bornim, Germany.

<sup>3</sup> University of Valencia –Intelligent Data Analysis Laboratory. Av. Universitat S/N, 46100, Burjassot (Valencia).

### Resumen

La detección temprana y automática de infecciones en poscosecha de cítricos es un problema de especial interés comercial todavía sin resolver, ya que provocan graves pérdidas económicas para la industria. Actualmente, esta detección se realiza por personal entrenado iluminando la fruta con peligrosa luz ultravioleta. Este trabajo estudió el potencial de la técnica de imagen basada en *backscattering* con la finalidad de detectar podredumbres en cítricos. Las imágenes *backscattering* se tomaron de naranjas cv. 'Navelate' usando cinco diodos láser que emiten en longitudes de onda distintas (532, 660, 785, 830 and 1060 nm). Estas imágenes tienen simetría radial con respecto al punto de incidencia de luz, por lo cual se redujeron mediante promediado radial a un perfil monodimensional. Los perfiles radiales se describieron usando una función de distribución producto cruzado de Gaussiana y Lorentziana (GL) de cinco parámetros independientes. Los parámetros de la función GL para cada longitud de onda se usaron como vector de características para clasificar la fruta en dos clases distintas (sana y podrida) mediante un clasificador de tipo supervisado basado en análisis discriminante lineal (LDA). Con el objetivo de encontrar la combinación de longitudes de onda de láser más adecuada para detectar podredumbres en cítricos, las longitudes de onda se ordenaron y compararon en términos de su contribución en la detección. La menor tasa de acierto promedio del clasificador del 80.4% se obtiene usando una única longitud de onda. Sin embargo, al utilizar una combinación de dos longitudes de onda en el clasificador el incremento de la tasa de acierto fue del 10%, llegando a alcanzar un éxito del 90.2%. Esta mejora se debió, sobre todo, al aumento del éxito en la clasificación de naranjas con podredumbre. Finalmente, los mejores resultados se consiguieron utilizando cinco longitudes de onda con una tasa de acierto promedio del 96.1%, siendo el número de frutas bien clasificadas superior al 95.0% para ambas clases.

Palabras clave: Inspección de fruta; Cítricos; Imagen *backscattering*; Clasificador LDA; Función producto cruzado de Gaussiana-Lorentziana

### Detection of decay in citrus fruit by means of backscattering image analysis

#### Abstract

The early detection of fungal infections in citrus fruits still remains as one of the major problems in postharvest, since these cause great economic losses for the industry. Nowadays, this detection is carried out manually by trained workers by illuminating each fruit with ultraviolet light in dark chambers while they pass along a conveyor belt. In the present study, the potential of the backscattering imaging technique was evaluated for detecting apparent decay in citrus fruits after infection with *Penicillium digitatum*. Backscattering images of oranges cv. 'Navelate' with and without decay were obtained using laser light at different wavelengths (532, 660, 785, 830 and 1060 nm). The images of backscattered light have radial symmetry with respect to the incident point of the light, being reduced to a one-dimensional profile after radial averaging. The radial profiles were described using the Gaussian-Lorentzian cross product (GL) distribution function with five independent parameters. The GL parameters at each wavelength were used as input vector for classifying fruit samples into two classes (sound orange and orange with decay) using a supervised classifier based on linear discriminant

analysis (LDA). In order to figure out the most suitable combination of laser wavelengths to detect fruits with decay, the laser wavelengths were ranked in terms of their contribution to the detection of decay. The minimum classification average success rate of 80.4% was obtained using a single wavelength. However, by employing a two-wavelength combination in the classifier, the increase in the average success rate was 10%, achieving a success rate of 90.2%. This improvement was mainly due to the increment in the number of well-classified oranges with decay. Finally, the best results were achieved using the five laser wavelengths, increasing the average success rate up to 96.1%, with a percentage of well-classified fruit samples greater than 95.0% for both classes.

Keywords: Fruit inspection; Citrus fruits; Backscattering imaging; LDA classifier; Gaussian-Lorentzian cross product function

## Introducción

La detección temprana y automática de infecciones por hongos en poscosecha de cítricos es un problema de especial interés comercial todavía sin resolver, ya que provocan graves pérdidas económicas para la industria. Actualmente, esta detección se realiza por personal entrenado iluminando la fruta con luz ultravioleta. Sin embargo, este proceso no está exento de problemas ya que la exposición de las personas a este tipo de iluminación resulta dañina.

Una posible solución a este problema puede venir del uso de los sistemas automáticos de visión por computador, como los que ya se utilizan para estimar parámetros morfológicos como tamaño, forma, color o detectar la presencia de otros daños (Cubero et al., 2011). Sin embargo, algunos daños, como las podredumbres, son especialmente difíciles de detectar utilizando sistemas convencionales de visión por computador, como las cámaras RGB, ya que estos defectos son difícilmente visibles a simple vista por el ojo humano. Por lo tanto, se han incorporado otras tecnologías de visión por computador para la detección automática de podredumbres en cítricos, como el uso de fluorescencia inducida por luz ultravioleta (Kurita et al., 2009) o la introducción de sensores hiperespectrales (Lorente et al., 2012; Gómez-Sanchis et al., 2012).

Recientemente, los sistemas de visión basados en *backscattering* se han estudiado como una técnica alternativa de visión por computador para evaluar la calidad de la fruta. En concreto, muchos estudios se han centrado en usar esta técnica en manzanas y otras frutas (Lu, 2004; Qing, 2007), aunque no se ha usado todavía para detectar podredumbres en cítricos. Esta técnica emergente está basada en la retrodispersión de la luz (*backscattering*). Este fenómeno es un tipo de reflexión de la luz y sucede cuando la luz, en vez de reflejarse especular o difusamente, interacciona con los componentes internos de la fruta y luego vuelve de nuevo al exterior de la superficie, por lo cual, esta luz contiene información relacionada con la morfología y estructuras del tejido (Birth, 1976).

La técnica de imagen basada en *backscattering* podría ser una prometedora herramienta para detectar podredumbres en cítricos, puesto que el proceso de podredumbre también implica cambios estructurales en el tejido de la fruta. El principal objetivo de este trabajo fue estudiar el potencial de los sistemas de imagen basados en *backscattering* para la detección automática de podredumbres producidas por el hongo *Penicillium digitatum* en cítricos. Para ello, se tomaron imágenes *backscattering* de naranjas usando diodos láser que emiten en cinco longitudes de onda distintas y se utilizó una función de distribución producto cruzado de Gaussiana y Lorentziana (GL) de cinco parámetros independientes para caracterizar los resultantes perfiles *backscattering*. Además, con el objetivo de clasificar la fruta en naranjas sanas y podridas, se construyeron clasificadores basados en análisis discriminante lineal (LDA) usando los parámetros GL para cada longitud de onda como vector de características. El último objetivo fue comparar las longitudes de onda de láser en términos de su contribución a la detección de podredumbres, con la finalidad de encontrar la combinación de longitudes de onda más adecuada para detectar fruta podrida.

## Material y métodos

### Material vegetal

Para los experimentos se emplearon naranjas cv. Navelate (*Citrus sinensis* L. Osbeck). Se usaron un total de 100 frutos: 50 inoculados con *P. digitatum* y 50 inoculados con agua para fines de control. La inoculación de las esporas de los hongos se realizó empleando una disolución de esporas en suspensión con una concentración de  $10^6$  esporas/ml por ser una concentración habitualmente en ensayos para producir podredumbres (Palou et al., 2001). Los frutos se almacenaron durante cuatro días a 20°C y una humedad relativa del 65%. Tras este periodo, todas las frutas inoculadas con ambos hongos presentaban daños superficiales con un diámetro igual o superior a 30 mm.

### Sistema de adquisición de imágenes

En este trabajo se utilizó un sistema de imagen basado en *backscattering*. Este sistema se componía de una cámara CCD de alta eficiencia (JAI CV-A50 IR), cinco diodos láser que emiten en longitudes de onda distintas (532, 660, 785, 830 and 1060 nm) usados alternativamente como fuentes de luz y un ordenador para controlar la cámara. La Figura 1a muestra un ejemplo de imagen *backscattering*. Se configuró para capturar imágenes con un tamaño de  $720 \times 576$  píxeles y una resolución de 0.0730 mm/píxel. El ángulo de incidencia del haz de luz se fijó a  $7^\circ$  con respecto al eje vertical. En la Figura 1b se muestra el esquema del sistema de adquisición de imágenes. En total se adquirieron cinco imágenes para cada uno de los 100 frutos, lo que constituye una base de datos de 500 imágenes *backscattering*.

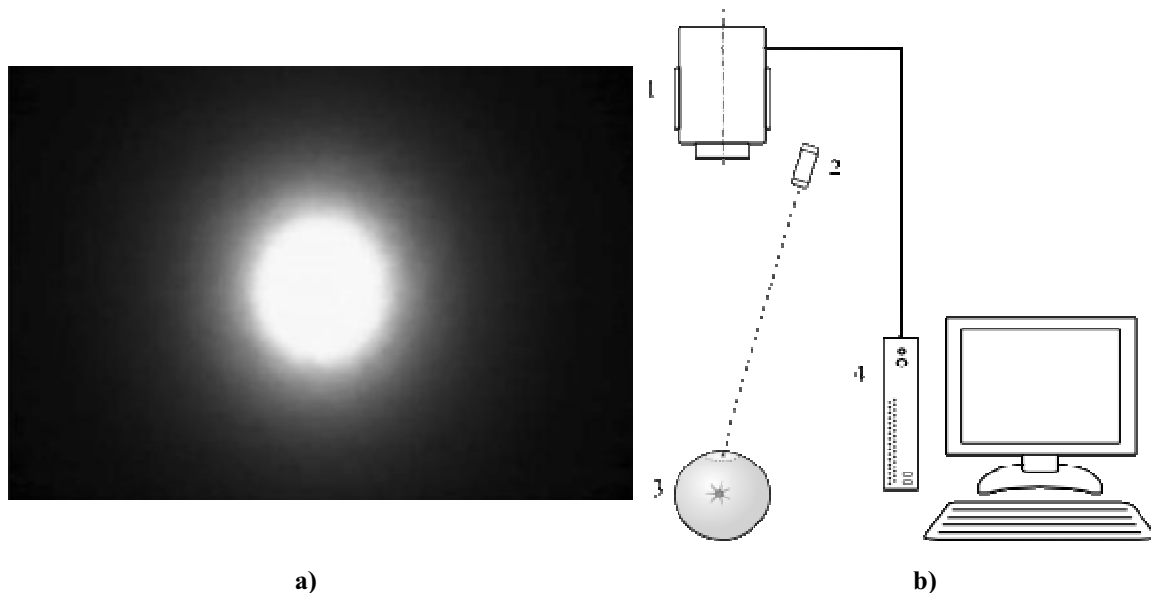


Figure 1. a) Ejemplo de imagen *backscattering*. b) Esquema del sistema de adquisición de imagen. 1: cámara CCD; 2: láser; 3: fruta; 4: ordenador

### Función para describir los perfiles *backscattering*

Las imágenes *backscattering* tienen simetría radial con respecto al punto de incidencia de luz (Figura 1), por lo cual, éstas se redujeron mediante promediado radial a un perfil monodimensional (Lu, 2004). La intensidad radial de los perfiles *backscattering* se obtuvo mediante el cálculo del valor medio de todos los píxeles dentro de cada anillo circular de un píxel de anchura. Los perfiles radiales se describieron usando una función de distribución producto cruzado de Gaussiana y Lorentziana

(GL), comúnmente usada en espectroscopia (Penache et al., 2002). La función GL se expresa matemáticamente con la Ecuación (1):

$$I(x) = a + \frac{b}{\left[1 + e\left(\frac{x-c}{d}\right)^2\right] \exp\left[\frac{1-e}{2}\left(\frac{x-c}{d}\right)^2\right]} \quad (1)$$

donde  $I$  es la intensidad luminosa de cada banda circular después del promediado radial;  $x$  es la distancia *scattering* (expresada en píxeles);  $a$  es el valor asintótico de la intensidad luminosa cuando  $x$  tiende a infinito;  $b$  es el valor máximo de intensidad luminosa en el centro;  $c$  es el centro;  $d$  es el ancho total correspondiente a la mitad del máximo valor de intensidad;  $e$  está relacionado con la forma, siendo de valor 0 con una función Lorentziana pura y 1 con una Gaussiana pura. La función GL se usó para ajustar los perfiles radiales para las cinco longitudes de onda para cada fruta. Se implementó un programa basado en regresión no lineal por mínimos cuadrados (Gelman y Hill, 2006) con el objetivo de estimar los cinco parámetros GL para cada longitud de onda. Todos los algoritmos de este trabajo se implementaron usando Matlab 7.9 (Mathworks, Inc.).

### Clasificador

En este trabajo se utilizó un clasificador basado en análisis discriminante lineal (LDA; Duda et al., 2001). LDA es una técnica de aprendizaje supervisado que se caracteriza por obtener una transformación lineal que maximiza el cociente de la varianza entre clases y la varianza entre los elementos de cada clase. Esta técnica estadística permite obtener un espacio vectorial transformado en el que las clases están máximamente separadas desde el punto de vista lineal. La principal diferencia con otro tipo de transformaciones como el análisis de componentes principales (PCA) es que LDA tiene en cuenta información supervisada del problema y no únicamente información sobre la distribución de los datos.

### Conjunto de datos etiquetado

Para construir el clasificador se etiquetó un conjunto de datos que constaba de 25 características asociadas a cada naranja. En concreto, cada patrón contenía los cinco parámetros de la función GL para cada longitud de onda y una clase asignada manualmente por un experto (sana y podrida). Estos datos se dividieron en dos subconjuntos: calibración (50% de las muestras) y validación (50%). El primer subconjunto se utilizó para construir el clasificador y el segundo para evaluar su rendimiento.

### Desarrollo y validación de los modelos de clasificación

El clasificador LDA y los parámetros GL para las cinco longitudes de onda de láser se usaron para clasificar las frutas. Con el objetivo de encontrar la combinación de longitudes de onda de láser más adecuada para detectar podredumbres en cítricos, las longitudes de onda se ordenaron y compararon en términos de su contribución en la detección. En primer lugar, el clasificador LDA se construyó y evaluó usando como vector de características los cinco parámetros GL asociados a cada longitud de onda individualmente. Se seleccionó la longitud de onda que tenía la mayor tasa de acierto promedio del clasificador. El siguiente paso fue la obtención de la mejor combinación de dos longitudes de onda. Cada una de las restantes longitudes de onda se añadió individualmente a la mejor longitud de onda, y se calcularon las tasas de acierto para todas las combinaciones de dos longitudes de onda. Este procedimiento se repitió para obtener la mejor combinación de tres longitudes de onda, y así sucesivamente hasta que todas las longitudes de onda se ordenaron.

## Resultados y discusión

La Tabla 1 muestra los resultados de clasificación correspondientes al conjunto de validación para las distintas combinaciones de longitudes de onda seleccionadas, incluyendo las tasas de acierto promedio del clasificador, así como de las matrices de confusión. Al comparar los resultados de clasificación, se observa que la menor tasa de acierto promedio del clasificador del 80.39% se obtuvo usando una única longitud de onda. Por otra parte, los mejores resultados se consiguieron utilizando cinco longitudes de onda con una tasa de acierto promedio del 96.08%. La matriz de confusión para el clasificador usando todas las longitudes de onda muestra que el número de frutas bien clasificadas era superior al 95.0% para ambas clases, a pesar de la gran similitud entre naranjas sanas y podridas.

Además, el incremento de la tasa de acierto fue del 10% al pasar de una única longitud de onda (80.39%) a la combinación de dos longitudes de onda (90.20%). A partir de las correspondientes matrices de confusión, se puede observar que la clasificación de naranjas podridas mejoró mucho para la combinación de dos longitudes de onda, incrementando de un 74.07 a un 92.59%, mientras que el número de naranjas sanas bien clasificadas permaneció constante (87.50%) en ambos casos. En la práctica, esta reducción del número de naranjas podridas mal clasificadas tiene mucha importancia para la implementación de sistemas de inspección, puesto que un número reducido de frutas infectadas puede contaminar una partida completa provocando grandes pérdidas económicas. Por otra parte, para el resto de casos, de una combinación de longitudes de onda a otra, el incremento en la tasa de acierto es sólo del 2% al añadir una longitud de onda más.

**Tabla 1.** Resultados de clasificación para las combinaciones de longitudes de onda seleccionadas

Número de longitudes de onda	Combinación de longitudes de onda (nm)	Tasa de acierto promedio (%)	Matriz de confusión		
				Sana (%)	Podrida (%)
1	532	80.39			
			Sana	87.50	25.93
			Podrida	12.50	74.07
2	532, 660	90.20			
			Sana	87.50	7.41
			Podrida	12.50	92.59
3	532, 660, 1060	92.16			
			Sana	91.67	7.41
			Podrida	8.33	92.59
4	532, 660, 1060, 830	94.12			
			Sana	95.83	7.41
			Podrida	4.17	92.59
5	532, 660, 1060, 830, 785	96.08			
			Sana	95.83	3.70
			Podrida	4.17	96.30

## Conclusiones

En este trabajo se estudió la viabilidad de los sistemas de imagen basados en *backscattering* para la detección de podredumbres en cítricos. Para esta detección se tomaron imágenes *backscattering* de naranjas para cinco longitudes de onda de láser. La función de distribución GL de cinco parámetros independientes se usó para describir los perfiles *backscattering*. Con el objetivo de clasificar la fruta

en naranjas sanas y podridas, se construyeron modelos de clasificación basados en LDA usando los parámetros GL para diferentes combinaciones de longitudes de onda como vector de características.

La menor tasa de acierto promedio del clasificador del 80.4% se obtuvo usando una única longitud de onda. Sin embargo, al utilizar una combinación de dos longitudes de onda en el clasificador el incremento de la tasa de acierto fue del 10%, llegando a alcanzar un éxito del 90.2%. Esta mejora se debió, sobre todo, al aumento del éxito en la clasificación de naranjas con podredumbre. Finalmente, los mejores resultados se consiguieron utilizando cinco longitudes de onda con una tasa de acierto promedio del 96.1%, siendo el número de frutas bien clasificadas superior al 95.0% para ambas clases. Por lo tanto, la técnica de imagen basada en *backscattering*, junto con la función GL, fue una herramienta adecuada para detectar podredumbres en cítricos.

## Agradecimientos

Este trabajo está parcialmente financiado por el Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) a través del proyecto RTA2009-00118-C02-01 y por el Ministerio de Ciencia e Innovación (MICINN) a través del proyecto DPI2010-19457, en ambos casos cofinanciados por fondos Europeos FEDER.

## Bibliografía

- Birth, G.S. (1976). How light interacts with foods. In: Gafney J.Jr.(Ed.), *Quality Detection in Foods*, ASAE, St. Joseph, USA, pp. 6-11.
- Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., and Blasco, J. (2011). Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food and Bioprocess Technology* **4**(4), 487-504.
- Duda, R.O., Hart, P.E., and Stork, D.G. (2001). *Pattern classification* (second edition). Wiley-Interscience.
- Gelman, A., and Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Gómez-Sanchis, J., Martín-Guerrero, J.D., Soria-Olivas, E., Martínez-Sober, M., Magdalena-Benedito, R., and Blasco, J. (2012). Detecting rotteness caused by *Penicillium* in citrus fruits using machine learning techniques. *Expert Systems with Applications* **39**(1), 780-785.
- Kurita, M., Kondo, N., Shimizu, H., Ling, P., Falzea, P.D., Shiigi, T., Ninomiya, K., Nishizu, T., and Yamamoto, K. (2009). A double image acquisition system with visible and UV LEDs for citrus fruit. *Journal of Robotics and Mechatronics* **21**(4), 533-540.
- Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O.L., and Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food and Bioprocess Technology* **5**(4), 1121-1142.
- Lu, R. (2004). Multispectral imaging for predicting firmness and soluble solids content of apple fruit. *Postharvest Biology and Technology* **31**, 147-157.
- Palou, L., Smilanick, J., Usall, J., and Viñas, I. (2001). Control postharvest blue and green molds of oranges by hot water, sodium carbonate, and sodium bicarbonate. *Plant disease* **85**, 371-376.
- Penache, C., Miclea, M., Bräuning-Demian, A., Hohn, O., Schössler, S., Jahnke, T., Niemax, K., and Schmidt-Böcking, H. (2002). Characterization of a high-pressure microdischarge using diode laser atomic absorption spectroscopy. *Plasma Sources Science and Technology* **11**(4), 476-483.
- Qing, Z., Ji, B., and Zude, M. (2007). Predicting soluble solid content and firmness in apple fruit by means of laser light backscattering image analysis. *Journal of Food Engineering* **82**, 58-67.