

Type: PhD summary

Section: Image Processing and Imaging Techniques

Participant in the 5th Ed. Justiniano Casas Prize (2016)

## Application of imaging techniques for the evaluation of phenolic maturity of grape seeds

### Aplicación de técnicas de imagen para la evaluación de la madurez fenólica de semillas de uva

Francisco J. Rodríguez-Pulido<sup>S</sup>, M. Lourdes González-Miret<sup>S</sup>, Francisco J. Heredia<sup>S\*</sup>

*Laboratorio Color y Calidad de Alimentos. Universidad de Sevilla.*

*Facultad de Farmacia, 41012-Sevilla*

(\*) E-mail: [heredia@us.es](mailto:heredia@us.es)

S: SEDOPTICA member

DOI: 10.7149/OPA.50.1.49503

#### ABSTRACT:

Quality of wine depends fundamentally on the condition of the components of grapes at the moment of harvest. In this respect, their seeds directly affect to structure, astringency, stability, and indirectly to colour of wine. In this work, the grape seeds have been evaluated by techniques of image analysis. In all cases, the methodology was developed trying to entail the minimum sample preparation and maximum quickness of the method, and as far as possible, substitute the chemical analysis that usually involves long and tedious processes that hamper monitoring of ripening in a rapid way. The proposed methodology for measuring and evaluating colour and near infrared spectrum has yielded satisfactory results for estimating the ripening of grape seeds, and it has been useful for establishing stages of ripeness from their optical features.

**Key words:** Colorimetry, Image Analysis, Grape seeds, Hyperspectral Imaging.

#### RESUMEN:

La calidad del vino depende en gran medida del estado en que se encuentren los diferentes componentes de la uva en el momento de la recolección y, en este sentido, las semillas afectan directamente a su astringencia, así como a la estructura y estabilidad del color del vino. En el presente trabajo se han evaluado semillas de uva de diferentes variedades aplicando distintas técnicas de análisis de imagen. En todos los casos se han desarrollado metodologías que implican la mínima preparación de muestra y la máxima rapidez del método, con el objeto de poder sustituir, en la medida de la posible, el análisis químico, que suele suponer largos y tediosos procesos que impiden un seguimiento rápido de la maduración. La metodología propuesta para la medición y evaluación del color y el espectro infrarrojo cercano ha dado resultados satisfactorios en la estimación del estado de madurez de las semillas de uva, y ha sido útil para establecer etapas de maduración a partir de sus propiedades ópticas.

**Palabras clave:** Colorimetría, Análisis de Imagen, Semillas de uva, Análisis de imágenes hiperespectrales.

#### REFERENCIAS Y ENLACES

- [1] H. R. Schultz and G. V. Jones, "Climate Induced Historic and Future Changes in Viticulture," *J Wine Res* **21**, 137–145 (2010).  
<https://doi.org/10.1080/09571264.2010.530098>
- [2] B. W. Zoecklein, *Análisis y producción de vino*. Zaragoza: Acribia, (2001).
- [3] B. Gordillo, F. J. Rodríguez-Pulido, M. L. Escudero-Gilete, M. L. González-Miret, and F. J. Heredia, "Comprehensive Colorimetric Study of Anthocyanic Copigmentation in Model Solutions. Effects of pH

- and Molar Ratio," *J Agric Food Chem* **60**, 2896–2905 (2012).  
<https://doi.org/10.1021/jf2046202>
- [4] R. Ristic and P. G. Iland, "Relationships between seed and berry development of *Vitis Vinifera* L. cv Shiraz: Developmental changes in seed morphology and phenolic composition," *Aust J Grape Wine Res* **11**, 43–58 (2005).  
<https://doi.org/10.1111/j.1755-0238.2005.tb00278.x>
- [5] F. Zamora-Marín, "El cambio climático, una amenaza para nuestra viticultura," *Enólogos* **29**, 28–31 (2006).
- [6] M. Gómez-Míguez, M. L. González-Miret, and F. J. Heredia, "Evolution of colour and anthocyanin composition of Syrah wines elaborated with pre-fermentative cold maceration," *J Food Eng* **79**, 271–278 (2007).  
<https://doi.org/10.1016/j.jfoodeng.2006.01.054>
- [7] A. González Marcos, F. J. Martínez de Pisón Ascacibar, A. V. Pernía Espinoza, F. Alba Elías, M. Castejón Limas, J. Ordieres Meré, and E. Vergara González, *Técnicas y algoritmos básicos de visión artificial*. Logroño, Spain: Universidad de La Rioja. Servicio de Publicaciones, (2006).
- [8] P. Jackman, D.-W. Sun, C.-J. Du, and P. Allen, "Prediction of beef eating qualities from colour, marbling and wavelet surface texture features using homogenous carcass treatment," *Pattern Recognit* **42**, 751–763 (2009).  
<https://doi.org/10.1016/j.patcog.2008.09.009>
- [9] C.-J. Du and D.-W. Sun, "Multi-classification of pizza using computer vision and support vector machine," *J Food Eng* **86**, 234–242 (2008).  
<https://doi.org/10.1016/j.jfoodeng.2007.10.001>
- [10] S. Pacifici, "Análisis de densidades mamográficas en espacio RGB," *Imagen Diagnóstica* **3**, 45–49 (2012).  
<https://doi.org/10.1016/j.imadi.2012.10.003>
- [11] International Organization for Standardization, "Graphic technology and photography – Colour characterisation of digital still cameras (DSCs) – Part 1: Stimuli, metrology and test procedures," ISO 17321-1:2012, (1999).
- [12] J. Hutchings, "14 Calibrated colour imaging analysis of food," in *Colour in food: improving quality*, D. B. MacDougall, Ed. Cambridge: CRC Press, (2002).
- [13] K. León, D. Mery, F. Pedreschi, and J. León, "Color measurement in L\*a\*b\* units from RGB digital images," *Food Res Int* **39**, 1084–1091 (2006).  
<https://doi.org/10.1016/j.foodres.2006.03.006>
- [14] T. Brosnan and D.-W. Sun, "Improving quality inspection of food products by computer vision—a review," *J Food Eng* **61**, 3–16 (2004).  
[https://doi.org/10.1016/S0260-8774\(03\)00183-3](https://doi.org/10.1016/S0260-8774(03)00183-3)
- [15] R. Ferrer-Gallego, J. M. Hernández-Hierro, J. C. Rivas-Gonzalo, and M. T. Escribano-Bailón, "Feasibility study on the use of near infrared spectroscopy to determine flavanols in grape seeds," *Talanta* **82**, 1778–1783 (2010).  
<https://doi.org/10.1016/j.talanta.2010.07.063>
- [16] M. R. Luo, G. H. Cui, and Li.C., "British Patent entitled Apparatus and method for measuring colour (DigiEye System), Derby University Enterprises Limited," (2001).
- [17] F. J. Rodríguez-Pulido, R. Ferrer-Gallego, M. Lourdes González-Miret, J. C. Rivas-Gonzalo, M. T. Escribano-Bailón, and F. J. Heredia, "Preliminary study to determine the phenolic maturity stage of grape seeds by computer vision," *Anal Chim Acta* **732**, 78–82 (2012).  
<https://doi.org/10.1016/j.aca.2012.01.005>
- [18] F. J. Rodríguez-Pulido, J. M. Hernández-Hierro, J. Nogales-Bueno, B. Gordillo, M. L. González-Miret, and F. J. Heredia, "A novel method for evaluating flavanols in grape seeds by near infrared hyperspectral imaging," *Talanta* **122**, 145–150 (2014).  
<https://doi.org/10.1016/j.talanta.2014.01.044>
- [19] F. J. Rodríguez-Pulido, L. Gómez-Robledo, M. Melgosa, B. Gordillo, M. L. González-Miret, and F. J. Heredia, "Ripeness estimation of grape berries and seeds by image analysis," *Comput Electron Agric* **82**, 128–133 (2012).  
<https://doi.org/10.1016/j.compag.2012.01.004>
- [20] F. J. Rodríguez-Pulido, B. Gordillo, M. Lourdes González-Miret, and F. J. Heredia, "Analysis of food appearance properties by computer vision applying ellipsoids to colour data," *Comput Electron Agric*

99, 108–115 (2013).

<https://doi.org/10.1016/j.compag.2013.08.027>

- [21] F. J. Rodríguez-Pulido, F. J. Heredia, J. M. Zaldívar-Cruz, and M. L. González-Miret, "Application of Image Analysis to the Color-Phenolic Composition Relationships of Grape Seeds," in *Color in Food*, CRC Press, 113–120 (2012).  
<https://doi.org/10.1201/b11878-15>
- [22] F. J. Rodríguez-Pulido, D. F. Barbin, D.-W. Sun, B. Gordillo, M. L. González-Miret, and F. J. Heredia, "Grape seed characterization by NIR hyperspectral imaging," *Postharvest Biol Technol* **76**, 74–82 (2013).  
<https://doi.org/10.1016/j.postharvbio.2012.09.007>
- [23] N. Vivas, Y. Glories, L. Lagune, C. Saucier, and M. Augustin, "Estimation du degré de polymérisation des procyanidines du raisin et du vin par la méthode au p-diméthylaminocinnamaldéhyde," *J Int Sci Vigne Vin* **28**, 319–336 (1994).
- [24] F. J. Rodríguez-Pulido. Tesis Doctoral: "Aplicación de técnicas de imagen para la evaluación de la madurez fenólica de semillas de uva" (2013).  
<https://www.educacion.gob.es/teseo/mostrarRef.do?ref=1059201>

## 1. Introducción

### 1.a. Importancia de las semillas de uva en la producción del vino

El clima es uno de los factores más importantes en el cultivo de la vid, ya que se trata de una planta sensible a las heladas en invierno y primavera y exigente de calor, no solo para su desarrollo vegetativo, sino también para la maduración de sus frutos, que precisan iluminación y temperaturas adecuadamente altas. Por estas circunstancias, la vid se ha cultivado tradicionalmente en climas no extremos de tipo mediterráneo, y su área general de cultivo corresponde a las dos zonas terrestres comprendidas entre los paralelos 30° y 50° latitud norte, y los 30° y 40° latitud sur [1]. La importancia socioeconómica que supone el cultivo de la vid ha hecho que los diferentes componentes de la uva sean objeto de numerosos estudios por su repercusión en la calidad final del vino. Una familia de éstos son los compuestos fenólicos, responsables de la estructura, cuerpo, aroma y directamente implicados en el color del vino. No solamente influyen en la calidad, también son objeto de estudio por el interés farmacológico que tienen sus propiedades antioxidantes. Los fenoles se encuentran en las partes sólidas de las bayas de *Vitis vinifera*, y pasan al vino entre un 40 y un 60% durante el proceso de vinificación [2].

En el desarrollo de la uva podemos diferenciar dos tipos de maduración. Una de ellas sucede en la pulpa del fruto y se denomina maduración industrial o glucoacídica. Durante ésta, la concentración de ácidos disminuye y el contenido en azúcar aumenta. Este contenido en azúcar es proporcional a la cantidad de alcohol que tendrá el vino resultante y, por eso, la maduración industrial ha sido el criterio tradicional para decidir la fecha de vendimia. El otro tipo, la maduración fenólica, se caracteriza por la biosíntesis de compuestos polifenólicos en las partes sólidas del fruto. Entre los fenoles sintetizados en la piel de la uva tinta destacan los antocianos, que son los pigmentos responsables del color característico del vino tinto. En las semillas también se biosintetizan compuestos fenólicos, que actúan como antioxidantes naturales y afectan a la estructura, astringencia y estabilidad del vino. Los fenoles de las semillas no son coloreados, pero son determinantes en la estabilidad del color del vino por el fenómeno químico de la copigmentación [3]. A lo largo de la maduración, las semillas cambian de apariencia debido a procesos de deshidratación y pardeamiento enzimático, y existen estudios que demuestran una fuerte correlación entre el color y la cantidad de fenoles extraíbles. Es por ello que la apariencia de la semilla se considera tradicionalmente uno de los criterios más utilizados para evaluar la madurez de la semilla de uva [4].

En climas cálidos, el cultivo de la vid y la elaboración de vinos de calidad se enfrentan a problemas relacionados con altas temperaturas ambientales. Con primaveras secas y veranos calurosos, el periodo comprendido entre el cambio de color de la uva (envero) y su madurez industrial óptima se reduce provocando un importante desfase entre la madurez de la pulpa y la de las partes sólidas (hollejos y semillas). Esto dificulta que las uvas alcancen la correcta madurez aromática y fenólica, produciendo colores pobres e irregulares [5]. Estos problemas son bien conocidos por los enólogos, de manera que las "caídas de color" en vinos producidos en climas cálidos han llevado en muchas regiones a, incluso, desestimar las elaboraciones de tintos, o a considerar el desarrollo de técnicas más o menos avanzadas para promover la síntesis de pigmentos o su extracción durante la maceración prefermentativa en frío [6]. El estado de maduración de la uva, y en concreto la madurez fenólica de las semillas, resultan de extraordinaria

importancia en la producción de vinos tintos de calidad en las zonas más afectadas por los efectos del cambio climático.

### **1.b. Análisis de imagen en el visible e infrarrojo cercano**

La ciencia informática avanza hacia la creación de ordenadores cada vez más rápidos, expertos y autónomos. Desde la creación del primer ordenador, el ser humano ha soñado con la creación de máquinas capaces de relacionarse con su entorno como lo hace él mismo, interactuando con lo que nos rodea como lo hacen los sentidos. Entre ellos, los avances más significativos se han alcanzado en dotar a una máquina con la capacidad de ver. Existen muchas analogías entre la visión de un ordenador y la visión de un ser humano pues, en ambos casos hay que adquirir la imagen e interpretarla. Para un ordenador, la cámara cumple las funciones fisiológicas del ojo, y es su sensor el que hace las veces de retina. Un ordenador procesando las imágenes realiza la función de cerebro, que debe reconocer, interpretar las formas y discernir qué información es útil en la imagen recibida. El proceso de análisis de imágenes abarca varias etapas entre las que se pueden destacar la adquisición, preprocesado, segmentación (identificación de objetos), extracción de información e interpretación de los resultados [7].

#### ***Colorimetría y análisis de imagen***

La medida del color se realiza normalmente con instrumentos que miden sobre una superficie concreta, ya sea obteniendo directamente los valores triestímulo o analizando el espectro visible y calculando los valores triestímulo a partir de él mediante las directrices de la CIE. En ocasiones es necesaria la evaluación de objetos cuyo color no es homogéneo, o puede suceder que el área del objeto a evaluar sea demasiado pequeña comparada con el área de medida los equipos colorimétricos. La mejora en la calidad de imagen ha hecho que el uso de cámaras digitales para la medida del color se haya desarrollado y extendido rápidamente. Gracias a una cámara digital, la medida del color va más allá, y permite evaluar aspectos relacionados con la apariencia y la textura, atributos hasta hace poco relacionados únicamente con la percepción humana [8]. Sin embargo, es necesario conocer las limitaciones del uso de cámaras digitales para la evaluación del color. Si bien la CIE trata de recomendar espacios de color que promuevan su comprensión independientemente del instrumento utilizado para mejorar la reproducibilidad de las medidas, el espacio de color RGB utilizado por cámaras digitales depende del dispositivo con el que se adquiere o donde se muestre, y por ello no se debe usar como tal para medidas colorimétricas absolutas. Otra particularidad del uso de cámaras digitales como colorímetros es la falta de linealidad entre valores RGB y los valores triestímulo XYZ. La imposibilidad de alternar de un espacio a otro con una matriz de cambio de espacio vectorial se subsana mediante ajustes no lineales. A pesar de estas limitaciones, algunos autores siguen utilizando las coordenadas RGB para la caracterización colorimétrica manteniendo siempre las mismas condiciones de medida [9, 10].

Como se ha visto anteriormente, los valores RGB de una cámara no coinciden con los valores CIE XYZ, ya que por lo general, los sensores de una cámara no tienen la misma sensibilidad espectral que los observadores estándares. El proceso por el cual se correlacionan las señales RGB con los valores CIE XYZ se denomina "caracterización de la cámara", y para ello se necesitan materiales de referencia. Existen dos tipos de caracterización: espectral y colorimétrica. Para la caracterización espectral, se mide directamente la respuesta de la cámara a una fuente monocromática en intervalos de 10 nanómetros. Este método permite medir la sensibilidad espectral de la cámara y se puede relacionar con las funciones de igualación CIE [11]. La caracterización colorimétrica se realiza mediante patrones de color, cuyos valores triestímulo XYZ son conocidos, se registran con la cámara digital a caracterizar para medir sus valores RGB bajo condiciones de iluminación fijas. La conversión entre valores RGB y XYZ se realiza mediante un ajuste por mínimos cuadrados utilizando transformaciones lineales y no lineales [12, 13]. El error entre los valores observados y predichos en este tipo de caracterizaciones se expresa mediante fórmulas de diferencias de color.

#### ***Análisis de imágenes hiperespectrales***

Los sistemas de visión artificial equipados con cámaras convencionales permiten evaluar características de apariencia tales como el color, forma, tamaño, textura, etc. Sin embargo, no permiten el análisis de otras propiedades más relacionadas con la composición química. Estas características pueden medirse mediante otras técnicas espectroscópicas más allá del intervalo visible del espectro. Las técnicas de imagen hiperespectral añaden las ventajas de las técnicas espectroscópicas y la visión artificial tradicional. Por un lado, permiten la medida de la huella espectral de una muestra, relacionada directamente con la composición química. Y por otro lado, la técnica permite la obtención de esa huella espectral en cada píxel de una imagen, lo que implica la medida de un parámetro químico y también su distribución espacial.

El concepto de imagen hiperespectral está alejado de la vida cotidiana, pero sus aplicaciones resultan muy beneficiosas en numerosos estudios. Una imagen común en color tiene tres bandas, roja, verde y azul; y la combinación de éstas produce la formación del resto de colores. Las imágenes hiperespectrales tienen entre decenas y cientos de bandas, y cada banda se corresponde con una longitud de onda diferente, independientemente de la región del espectro electromagnético. Con ellas se logra llegar a donde el ojo humano o una cámara convencional no pueden. Así, una imagen hiperespectral o hipercubo está compuesta por una batería de imágenes de una misma escena, donde cada imagen representa en escala de grises la reflectancia de la muestra a una única longitud de onda. Cada píxel de la imagen contiene un espectro completo de esa posición, que queda archivado como una huella dactilar. En el hipercubo, dos de las dimensiones corresponden a las coordenadas espaciales  $x$  e  $y$ , y una tercera dimensión corresponde a la longitud de onda ( $\lambda$ ). Aunque se desarrollaron originariamente para la detección remota desde satélite, esta técnica de imagen ha desarrollado todo su potencial como herramienta analítica en campos como la agricultura, la astronomía, la industria farmacéutica o la medicina. En ciencia de los alimentos se utiliza para detectar enfermedades en frutos y cultivos, así como para conocer la composición química y el grado de maduración [14].

### 1.c. Color en las semillas de uva

El estudio del cambio de color de las semillas de uva durante la maduración y sus implicaciones químicas ya ha sido estudiado con anterioridad [4], aunque esta evaluación del color se ha hecho generalmente de manera visual. Otros estudios han ahondado en sus propiedades espectrales utilizando técnicas espectrorradiométricas convencionales pero sin contar con técnicas de imagen [15], aunque en el caso de las bayas sí existen estudios que evalúan su composición mediante técnicas de imagen. El objetivo general de la tesis fue profundizar en el conocimiento del color y la apariencia de las semillas de uva en relación con su composición química y, en concreto, con el estado de maduración fenólica. Hasta ese momento, no existían prácticamente estudios que hubieran evaluado sus particularidades colorimétricas y espectrales en el infrarrojo cercano mediante técnicas de imagen.

## 2. Materiales y Métodos

### 2.a. Muestras

Se utilizaron uvas (*Vitis vinifera* L.) de viñedos localizados principalmente en la Denominación de Origen Condado de Huelva en las cuatro campañas comprendidas entre 2009 y 2012. Se muestrearon cuatro viñedos con diferencias en variedad de uva y tipo de suelo. De variedades tintas se consideraron Tempranillo y Syrah, las cuales están bien extendidas en toda España y se caracterizan por ser bastante resistentes a enfermedades y a diferentes condiciones agroclimáticas. Se consideró también la variedad de uva blanca Zalema, autóctona del Condado de Huelva. Esta variedad se muestreó en dos viñedos que diferían en el tipo de suelo (arena y barro). Para uno de los ensayos se utilizaron semillas de uva de la variedad Graciano pertenecientes a la Denominación de Origen Calificada Rioja.

### 2.b. Análisis de imagen

La adquisición de imágenes en el visible se realizó con el sistema DigiEye® (VeriVide, Reino Unido) [16]. El equipo consta de una cabina de iluminación equipada con dos emuladores D65, que iluminan la cúpula superior y es esta la que proporciona luz difusa y homogénea a la bandeja de muestra (Fig.1). Dispone en la parte superior de una cámara digital Nikon® D80 que adquiere las imágenes y las envía a un ordenador que convierte la señal RGB a unidades CIELAB.

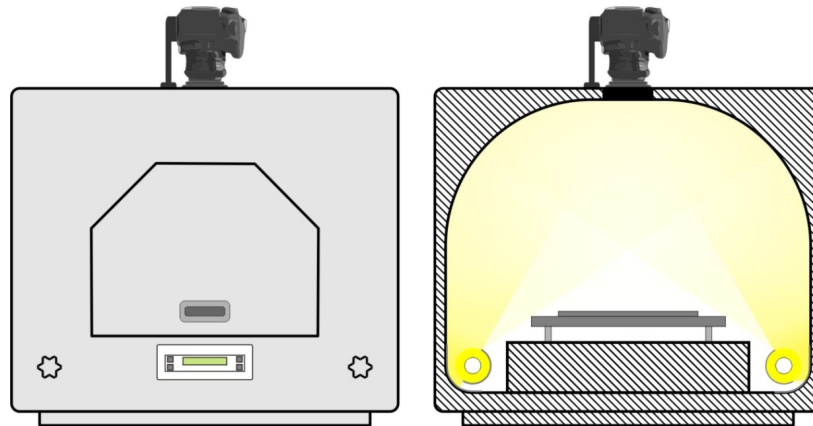


Fig.1. Esquema del exterior e interior de la cabina del sistema DigiEye.

Las imágenes hiperespectrales se adquirieron con una cámara suministrada por Infaimon® (Barcelona, España) que consta de los siguientes componentes (Fig. 2): Cámara Xeva 1.7-320 (Xenics®) útil en el intervalo 900-1700 nm; espectrógrafo N17E (Specim®); MirrorScanner (Specim®) como sistema de barrido basado en un espejo móvil; y dos lámparas halógenas de 70 W (Prilux®) separadas 50 cm y orientadas a 45° respecto la superficie de muestra. Como bandeja de muestra se utilizó una superficie blanca de polietileno, cuyo espectro característico facilitó los procesos de segmentación de los hipercubos.

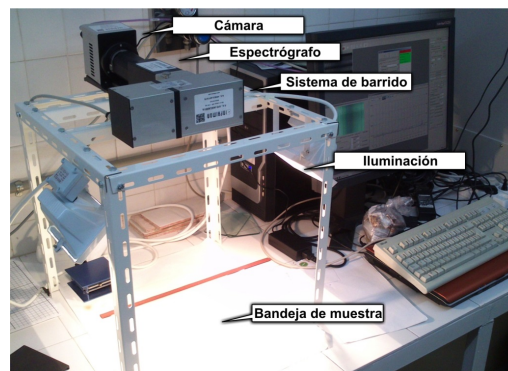


Fig.2. Cámara hiperespectral.

El procesado de imágenes, tanto en el visible como en el infrarrojo cercano, se realizó mediante software desarrollado por nuestro grupo bajo programación MATLAB R2009b (The Mathworks®), utilizando los paquetes de herramientas *Image Processing*, *Fuzzy Logic*, *Partial Differential Equation* y *Statistics*.

### 2.c. Análisis químico y quimiométrico

Los análisis químicos de referencia se realizaron siguiendo los métodos descritos por Rodríguez-Pulido et al. [17, 18]. La estadística descriptiva, análisis de la varianza (ANOVA) y el análisis general discriminante por pasos (FS-GDA) se realizó con Statistica 8.0 (Statsoft®). Los análisis de componentes principales (PCA) y las regresiones múltiples por mínimos cuadrados parciales (PLSR) se realizaron con MATLAB y Unscrambler 9.7 (CAMO®).

## 3. Resultados y discusión

### 3.a. Análisis de imagen en el visible.

El primer hito en este trabajo fue la puesta a punto de la metodología para la evaluación de las características de color y apariencia de semillas de uva, y estudiar cómo estas características evolucionan durante la maduración. Se consideraron semillas de uva de las cosechas de 2009 y 2010. El algoritmo programado permitió la medida del color en el espacio CIELAB, y en cuanto a apariencia, se midió en cada semilla el largo, ancho, relación de aspecto, grado de redondez, perímetro, área y heterogeneidad, definiéndose ésta como la fracción de píxeles cuya intensidad difiere más de un 10 % de la media, como se describe en Rodríguez-Pulido et al. [19].

En este estudio no se observaron cambios en la morfología de las semillas a lo largo del periodo de tiempo considerado, así que esta morfología no parecía un buen indicativo para medir la madurez. Sin embargo, sí se observaron diferencias significativas ( $p < 0.05$ ) entre las variedades Syrah, Tempranillo y Zalema, aunque de esta última variedad no se pudo diferenciar por el tipo de suelo. Se aplicó un análisis discriminante por pasos, que incluye iterativamente variables hasta la máxima optimización del modelo cuyos resultados están resumidos en la Tabla 1. De entre las variables utilizadas para la clasificación, el análisis discriminante incluyó secuencialmente *relación de aspecto*, *grado de redondez* y *ancho de semilla* en las ecuaciones como las de mayor poder de clasificación.

TABLA 1. Resumen de los resultados del análisis discriminante.  
Clasificación (a) según semillas de uva blanca o tinta, (b) según variedad, y (c) según variedad incluyendo tipo de suelo.

		Clasificación		
		(a)	(b)	(c)
tinta	Syrah	100%	100%	100%
	Tempranillo		87.7%	87.7%
blanca	Zalema arena	100%	100%	57.2%
	Zalema barro			57.2%

Respecto a la evolución del color durante la maduración, tanto la claridad, como el croma y el tono experimentaron un descenso que, conforme avanzaba en el tiempo, se hacía menos pronunciado. Otra observación colorimétrica fue que el pardeamiento de la semilla no ocurría de manera homogénea, sino que aparecía una zona oscura en la parte más ancha de la semilla y ésta se extendía hacia el pico durante la maduración. Para evaluar el fenómeno se observaron los histogramas de los píxeles de semillas antes y después de suceder el pardeamiento, y se observó que ambos histogramas presentaban un valle en el valor  $L^* = 50$ . Así, se ha propuesto un *índice de pardeamiento* definido como la fracción de píxeles cuyo valor de  $L^*$  es menor a 50 unidades CIELAB.

#### Propuesta de un índice de maduración de semillas

Otro aspecto que se estudió fue la evaluación dentro del espacio CIELAB de la dispersión de los puntos que sufrían las semillas de uva a lo largo de la maduración. Esta dispersión daba una idea de la heterogeneidad cromática, pero presentaba problemas a la hora de ser cuantificada, ya que las medidas existentes de heterogeneidad cromática en el espacio CIELAB no tienen en cuenta la orientación o la forma de esas nubes de puntos. Para afrontar este problema se propuso una metodología que ajustaba las nubes de puntos en el espacio de color a un elipsoide cuyo tamaño, forma y orientación podían cuantificarse y ser útiles con fines predictivos no solo en semilla de uva sino también en otros alimentos [20].

Esta metodología se utilizó para la creación de un índice de maduración de semillas de uva. Para ese fin, se consideró un número importante de semillas ( $n = 7635$ ) para que quedara bien representada toda la casuística posible de apariencia. De cada una de ellas se midió tanto el color, como las nuevas variables propuestas a partir del ajuste de sus nubes de puntos a elipsoides. El nuevo espacio multidimensional consideró hasta treinta y una variables por semilla. Para pre-establecer niveles de maduración que sirvieran como base en el modelo, se seleccionaron manualmente 1114 semillas en los cuatro estados ya descritos en bibliografía [4] (fig.3).

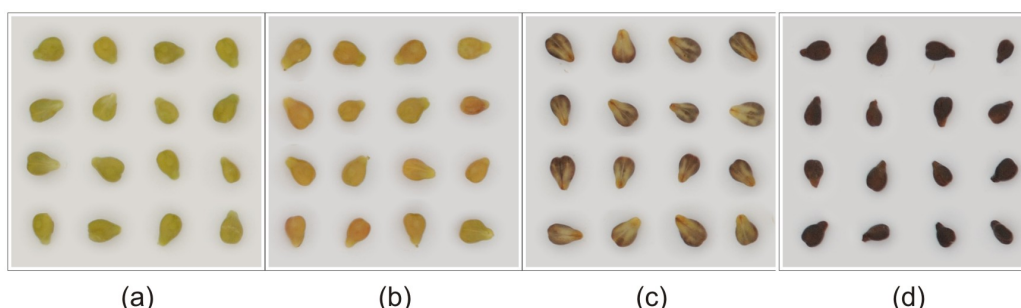


Fig.3. Apariencia de algunas semillas elegidas inicialmente para el modelo. (a) verde, (b) amarilla, (c) intermedia, (d) parda.

Fijados los valores de cada variable en cada uno de los cuatro niveles preliminares, se asignó a cada fecha de muestreo el nivel cuya distancia en el espacio multidimensional fuera mínima. Un modelo con solo cuatro

niveles, de los cuales dos de ellos corresponden a estadios muy tempranos, es claramente insuficiente desde un punto de vista enológico. Por ello, se programaron procedimientos lineales y no lineales para el diseño de un modelo con un mayor número de etapas. Así se consiguieron diferentes alternativas, desde las más sencillas con centroides equidistantes, tanto temporal como colorimétricamente, hasta las más complejas, basadas en modelos no lineales con promedios no equidistantes. La figura 4 muestra el resultado de aplicar uno de los algoritmos lineales probados.

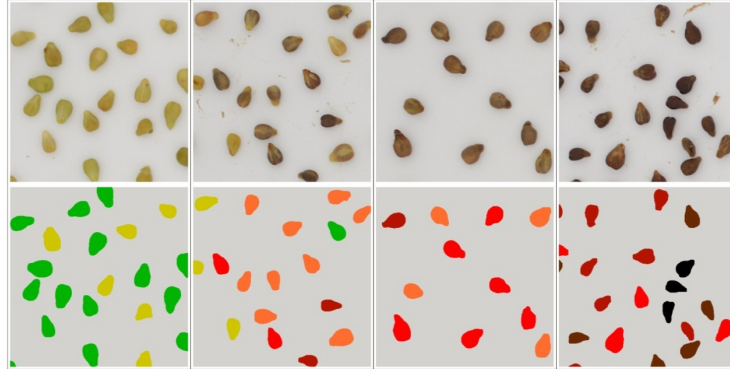


Fig.4. Aplicación de la asignación de semillas en niveles de maduración. verde (I), amarillo (II), naranja (III), rojo (IV), rojo oscuro (V), marrón (VI), negro (VII)

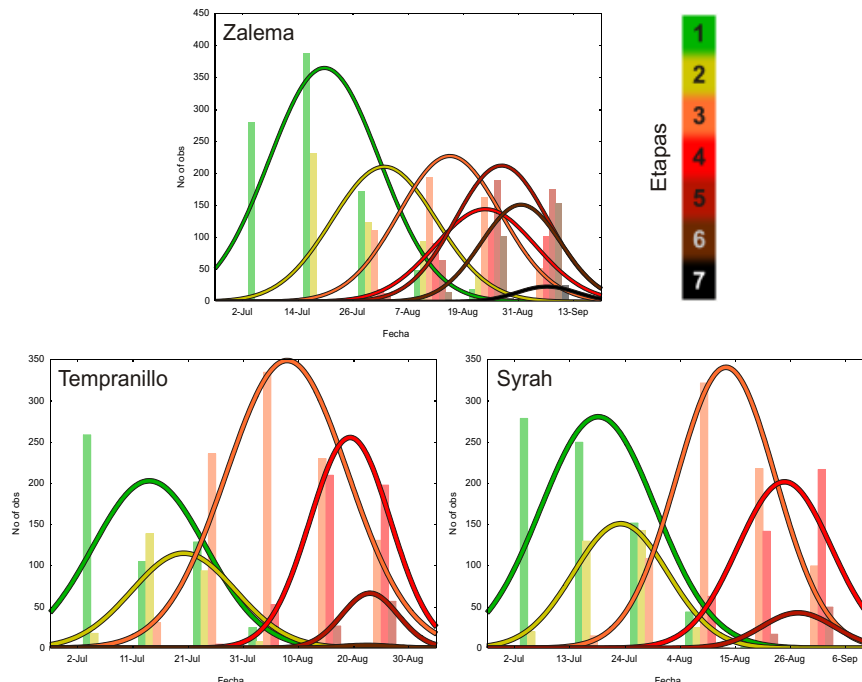


Fig.5. Distribución de las semillas según su nivel de maduración y variedad.

Con este modelo se pudieron establecer siete niveles que resultaron de la inter y extrapolación de los cuatro preliminares dentro del espacio multivariante considerado. En este caso, cada nivel equidistaba temporalmente de otro aproximadamente diez días, lo que resultó más coherente con criterios enológicos. Como era previsible, en cada muestreo se encontraron semillas en diferentes estados de maduración. La presencia relativa de cada nivel dependía de la variedad y fecha de muestreo. Para la variedad Zalema, los niveles 1 y 2 predominaron en los estadios iniciales y, a partir de finales de julio, las semillas comenzaron a pardearse de manera heterogénea; de hecho, las muestras de las dos últimas semanas de agosto contenían semillas pertenecientes a todos los niveles del modelo. Aplicando el método desarrollado para el establecimiento del nivel de maduración, se encontró que, en el momento de la vendimia, dos tercios de las semillas estaban en los estados 5 y 6, y solo un 10% de las semillas había alcanzado la etapa 7. Los perfiles de las variedades tintas fueron muy similares entre sí, y diferentes al de la variedad blanca (Fig. 5). En ellas, los dos niveles iniciales se encontraron más solapados que en las uvas blancas y, a mitad de verano, ya no se encontraron semillas en ninguno de esos niveles presentando una evolución más homogénea. De hecho,



al desaparecer las semillas de los niveles iniciales, el nivel 3 (principio del pardeamiento) fue el estado dominante para Syrah y Tempranillo, a diferencias de Zalema donde sus semillas coexistieron en diferentes niveles al mismo tiempo. Probablemente debido a la temprana vendimia de las variedades tintas, muy pocas alcanzaron los niveles 6 y 7. Merece la pena destacar que, a pesar de la similitud de los perfiles de distribución de las dos variedades tintas, existió un desfase en el tiempo entre ambas, concretamente de una semana, y coincidió a su vez con el desfase que hubo entre ambas vendimias.

### ***Relación entre el color y la composición fenólica de semillas de uva***

Esta relación se estudió a través de dos ensayos individuales. El primero de ellos estudió las relaciones univariantes entre cada coordenada colorimétrica de las semillas y su composición química analizada por técnicas cromatográficas. De los compuestos analizados, el ácido protocatéquico y la catequina mostraron una evolución más pronunciada. En la matriz de correlaciones se observó que prácticamente todas las relaciones fueron significativas ( $p < 0.05$ ). Además, no solo correlacionaron las variables colorimétricas cuantitativas, sino también el tono, lo que pone de manifiesto cambios en la composición de la semilla desde un punto de vista cualitativo.

En el otro ensayo se evaluó la relación entre color y composición mediante técnicas multivariantes. A diferencia del resto de estudios, para este ensayo se utilizaron semillas de uva tinta de la variedad Graciano de la Denominación de Origen Calificada Rioja durante la campaña de 2010, y el análisis químico se realizó mediante cromatografía líquida de alta eficacia con espectrometría de masas como sistema de detección (HPLC-MS). Se identificaron un total de veintinueve compuestos fenólicos, que fueron agrupados en las siguientes familias: ácidos hidroxibenzoicos, monómeros de flavanoles, dímeros, trímeros y flavanoles galoilados. Los resultados estuvieron en concordancia con otros estudios publicados anteriormente. Con la premisa de que unas variables podrían influir más que otras en la relación apariencia-composición, se aplicó una regresión múltiple por pasos (forward stepwise multiple regression) donde se intentó predecir el momento de la maduración en la que se encuentra la muestra, de composición química definida, a partir de las variables obtenidas por el análisis de imagen. El modelo iterativo de regresión incluyó secuencialmente la claridad, el croma, la longitud de la semilla, el grado de redondez y la relación de aspecto, como variables independientes en la predicción. De ellas, fueron estadísticamente significativas ( $p < 0.05$ ) la claridad y la longitud de la semilla. Al evaluar la relación entre los valores predichos y los observados, se alcanzó un buen coeficiente de determinación ( $R^2 = 0.97$ ,  $p = 0.0089$ ). Este hecho, unido a la estrecha banda de regresión obtenida, indicó una alta predictibilidad de la etapa de madurez a partir de los datos de análisis de imagen [17, 21].

### **3.b. Análisis de imagen en el infrarrojo cercano.**

#### ***Puesta a punto de la metodología y caracterización varietal***

Para este ensayo se consideraron de nuevo las variedades tintas Tempranillo y Syrah, y la variedad blanca Zalema cultivada en dos tipos de suelo. El primer paso fue optimizar la adquisición, calibrar las imágenes y programar los algoritmos de segmentación y extracción de los espectros de reflectancia.

Se realizó un análisis de componentes principales a partir de la información espectral y se observó que la fecha de muestreo influía fuertemente en el primer componente principal independientemente de la variedad. Esto se debía a que la pérdida de agua que sufre la semilla durante la maduración es aún más fuerte que las diferencias en la composición fenólica. Atendiendo a las dos siguientes componentes principales, donde se alcanzaba el 99% de la varianza, se observó una clara separación según la variedad. Incluso para la variedad Zalema, la información espectral era diferente según el tipo de suelo. A partir de los pesos de las variables se realizó una selección de las seis longitudes con más influencia con el fin de poder extraer la misma información a partir de menos bandas. Estas longitudes de onda fueron 928, 940, 1148, 1325, 1620 y 1656 nm, y los resultados del análisis de componentes principales fue prácticamente el mismo que a partir del espectro completo (Fig.6).

Para verificar el potencial de las bandas seleccionadas se aplicó un análisis discriminante utilizando como variables independientes el espectro completo y las seis bandas individuales, y siendo la variedad la variable categórica dependiente. Cuando se utilizó el espectro completo todas las muestras de semilla se clasificaron correctamente, distinguiendo incluso el tipo de suelo en la variedad Zalema. El porcentaje disminuyó al utilizar solamente las seis longitudes de onda, pero fue aún satisfactorio, ya que la clasificación fue correcta para más de un 96 % de los casos cuando la información espectral se había reducido de 240 a solamente 6 bandas [22].

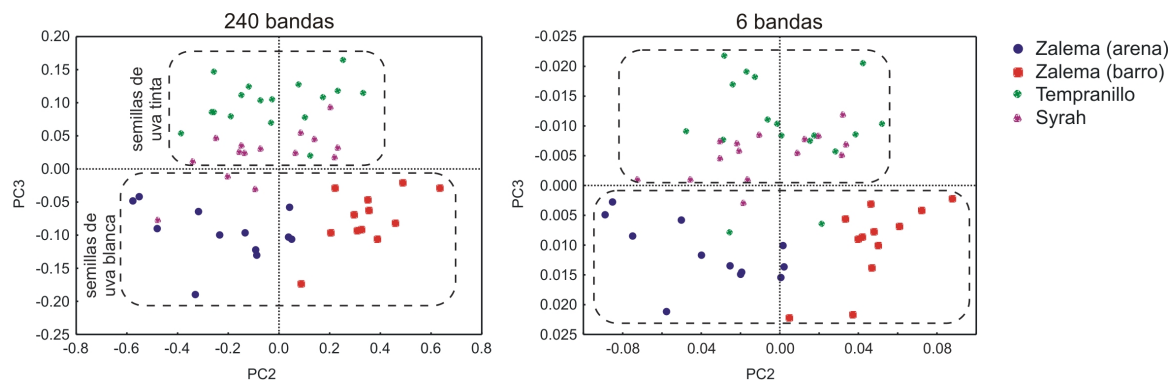


Fig.6. Puntuaciones de PC2 y PC3 de las muestras de semillas de uva según el número de bandas espectrales utilizadas

**Relación entre la composición y el perfil hiperespectral en semillas de uva**

Para este último ensayo se consideraron semillas de uva tinta Tempranillo y de uva blanca Zalema muestreadas desde principios de julio hasta la fecha de vendimia dos veces por semana. El número de muestras se dividió en tres sets. Dos de ellos se utilizaron como set de calibración y un tercio como set de predicción en los modelos. Se adquirieron las imágenes hiperespectrales de las semillas mediante la metodología puesta a punto anteriormente y se extrajo el espectro de cada muestra. Se analizó a continuación el contenido en flavanoles [23], como principales responsables de la composición fenólica en las semillas. La extracción de compuestos se realizó de dos formas diferentes, de tal manera que cada resultado representara el contenido flavanólico total y la fracción de flavanoles que las semillas ceden durante el proceso de vinificación (extracción con vino sintético).

La Tabla 2 resume los resultados de los modelos de predicción mediante regresión múltiple por mínimos cuadrados parciales. En cuanto a las variedades, se observan mejores resultados cuando son consideradas individualmente que cuando se consideran las dos juntas. Este resultado es obvio, pues la diferencia varietal influye en la varianza de la información espectral y afecta a los modelos. Por otro lado, se observaron coeficientes de determinación R<sup>2</sup> más altos y errores medios de calibración y predicción (RMSEC y RMSEP) más bajos cuando se quería estimar la cantidad de flavanoles presente en las extracciones con vino sintético, lo que resulta más útil desde un punto de vista enológico.

TABLA 2. Resultados de calibración y predicción de los modelos de regresión obtenidos a partir de la información espectral. RMSEC y RMSEP están expresados en mg de catequina por gramo de semilla.

	N	Tipo de extracción	Factores PLS	R <sup>2</sup> <sub>c</sub>	RMSEC	R <sup>2</sup> <sub>p</sub>	RMSEP
Todas las muestras	95	vino sintético	3	0.82	0.92	0.85	0.88
		total	3	0.73	4.01	0.75	3.86
Zalema	50	vino sintético	2	0.83	0.98	0.85	0.92
		total	1	0.82	2.90	0.82	2.93
Tempranillo	45	vino sintético	2	0.88	0.67	0.88	0.69
		total	6	0.94	2.09	0.88	2.89

**4. Conclusiones**

La tesis que este trabajo resume [24] ha supuesto el primer acercamiento de las técnicas de imagen a la caracterización fenólica de las semillas de uva, las cuales repercuten de manera activa en la calidad final del vino. Las particularidades colorimétricas y de apariencia de estas semillas han requerido el desarrollo de nuevas metodologías para evaluar la heterogeneidad cromática, basadas en la creación de nuevas variables derivadas del color, para medir el grado de maduración. A pesar de que los compuestos fenólicos estudiados no son coloreados, se han podido establecer relaciones estadísticas uni y multivariantes para la predicción de la composición química a partir de los datos obtenidos por análisis de imagen en el visible.

Más allá del color y la apariencia, se ha desarrollado la metodología basada en imágenes hiperespectrales en el infrarrojo cercano para la evaluación de la huella espectral de las semillas de uva. Esta huella ha

resultado ser útil para la caracterización varietal, distinguiendo incluso para una misma variedad el tipo de suelo en el que las uvas fueron cultivadas. Esta diferenciación varietal en el espectro infrarrojo seguía siendo efectiva tras realizar una drástica reducción de variables. Las longitudes de onda seleccionadas tras la reducción pertenecían a aquellas transiciones vibracionales de los enlaces típicos de los compuestos fenólicos presentes en semillas de uva. Las técnicas hiperespectrales resultaron también útiles para la predicción de su composición fenólica. Así, se evaluó con precisión el contenido en flavanoles, principales fenoles presentes en las semillas. Esta capacidad predictiva mejoró cuando, en lugar de evaluar el contenido total, se consideraron solo la cantidad de flavanoles que se transfieren al mosto durante el proceso de vinificación.

En resumen, las técnicas de análisis de imagen desarrolladas, tanto en la región visible del espectro como en el infrarrojo cercano se presentan como una alternativa para la estimación de características químicas de las semillas, debido a su simplicidad, versatilidad, rapidez y bajo coste. No necesitan tratamiento previo de las muestras ni uso de reactivos, lo cual las hace potencialmente útiles para la estimación rápida del estado de maduración de las semillas en la industria vitivinícola. Desde la publicación de los artículos resultantes de la tesis que este trabajo resume hasta la actualidad han surgido más estudios con otras variedades y condiciones agronómicas, lo que corrobora la idoneidad de estas técnicas con fines analíticos en el campo de la enología.

## **Agradecimientos**

Este trabajo ha sido realizado con el apoyo institucional del Ministerio de Economía y Competitividad del Gobierno de España con la concesión de los proyectos AGL2008-05569-C02-02, AGL2011-30254-C02-02 y AGL2014-58486-C2-2-R.