

**DESARROLLO DE TÉCNICAS ÓPTICAS DE ANÁLISIS RÁPIDO PARA EL  
CONTROL DE CALIDAD EN EL PROCESO DE ELABORACIÓN DE  
ACEITE DE OLIVA VIRGEN**

El Director

Tutora UCO

Fdo. José Antonio García Mesa  
Investigador Titular  
  
IFAPA Venta del Llano

Fdo. María Dolores Pérez Marín  
Profesora Departamento  
Producción Animal  
Universidad de Córdoba

Trabajo presentado para aspirar al Grado de Doctor

LA DOCTORANDA,

Fdo. Elena Guzmán Jiménez

Licenciada en Biología por la Universidad de Jaén

TITULO: *DESARROLLO DE TÉCNICAS ÓPTICAS DE ANÁLISIS RÁPIDO  
PARA EL CONTROL DE CALIDAD EN EL PROCESO DE  
ELABORACION DE ACEITE DE OLIVA VIRGEN*

AUTOR: *ELENA GUZMAN JIMENEZ*

---

© Edita: Servicio de Publicaciones de la Universidad de Córdoba.  
Campus de Rabanales  
Ctra. Nacional IV, Km. 396 A  
14071 Córdoba

[www.uco.es/publicaciones](http://www.uco.es/publicaciones)  
[publicaciones@uco.es](mailto:publicaciones@uco.es)

---

José Antonio García Mesa, Investigador titular del Centro IFAPA Venta del Llano (Instituto de Investigación Agraria y Pesquera de Andalucía), en calidad de Director de la Tesis Doctoral presentada por la Licenciada Elena Guzmán Jiménez

**CERTIFICA:** Que la Tesis Doctoral “**DESARROLLO DE TÉCNICAS ÓPTICAS DE ANÁLISIS RÁPIDO PARA EL CONTROL DE CALIDAD EN EL PROCESO DE ELABORACIÓN DE ACEITE DE OLIVA VIRGEN**” ha sido desarrollada totalmente en los laboratorios del Centro Venta del Llano del IFAPA (Jaén), y que a mi juicio, reúne todos los requisitos exigidos a este tipo de trabajo.

Y para que conste y surta los efectos pertinentes, expiden el presente certificado en la ciudad de Jaén, a 1 Marzo de 2013.

JOSÉ ANTONIO GARCÍA MESA

Mediante la defensa de esta Memoria, se pretende optar a la obtención del Doctorado Internacional, habida cuenta que la doctoranda reúne los requisitos exigidos para tal mención:

1. Se cuenta con los informes favorables de dos doctores pertenecientes a Instituciones de Enseñanza o Investigación de países europeos distintos del nuestro.
2. En el Tribunal que ha de evaluar la Tesis Doctoral, existe un miembro de un Centro de Investigación de otro país europeo.
3. Parte de la defensa de la Tesis Doctoral se realiza en la lengua oficial de otro país europeo.
4. La doctoranda ha realizado una parte del trabajo experimental en el Centro Wallon de Investigaciones Agronómicas de Gembloux (CRAW) en Bélgica, gracias a la concesión de una ayuda para estancias en el extranjero asociada a la beca de formación de personal investigador del Instituto Nacional de Investigaciones Agroalimentarias (INIA) y la Consejería de Agricultura, Pesca y Medio Ambiente.

Quiero agradecer al Instituto Nacional de Investigación Agraria y Alimentaria (INIA) la concesión de una beca para la formación del personal investigador del subprograma FPI-INIA, en régimen de cofinanciación entre el IFAPA y el INIA, que está cofinanciada por el Fondo Social Europeo, ya que ha posibilitado mi dedicación a este trabajo durante estos años.



Instituto de Investigación y Formación Agraria y Pesquera  
CONSEJERÍA DE AGRICULTURA, PESCA Y MEDIO AMBIENTE



Unión Europea

Fondo Social Europeo

A mi director, José Antonio García Mesa por sus orientaciones y su ayuda en la realización de este trabajo y por la oportunidad de comenzar mi carrera investigadora.

Muchas gracias en especial, a Juan y Vincent, por su apoyo, por su ayuda incondicional en cualquier momento, su paciencia, por su tiempo y su trabajo y porque gracias a ellos esta Tesis es como es.

A todo el equipo del Laboratorio, Maite, M<sup>a</sup> Carmen y Toñi por su enseñanza y ayuda en las determinaciones analíticas y en particular a Gloria por su gran ayuda y por ser muchos días la “percha de mis penas” que me ha escuchado y ha sido capaz de tranquilizarme en mis días de agobio

A todos mis compañeros,

Yosra, Maika, Gema, Elena, Avir, Abraham, Aymen, también a los últimos en llegar, Angélica, Leli y Estrella. A Lola, Contreras, Conchi, Noelia, por nuestros grandes momentos de “descompresión”, a todos por tratarme de la mejor forma y porque de esa forma da gusto ir a trabajar. A Miguel y Alfonso por su gran ayuda, por hacer posibles todos nuestros “inventos” y en resumen por hacerme el trabajo mucho más fácil.

A Ginés y todos los compañeros del campo por su trabajo en la recogida de muestras y en todo lo que los he necesitado.

Finalmente, a mis amigos por su confianza y su apoyo en todo momento. A mi familia y a “Fer”, por vuestro cariño y por entenderme, sobre todo en estos últimos meses bastante difíciles para mí. A mi abuelito, que aunque siga sin entender mi trabajo, sólo por estar ahí, ya es una gran ayuda. Y a vosotras, que aunque no lo podáis ver, sé que estáis conmigo.

A todos los que antes o después han pasado por mi vida ya que de una manera u otra han colaborado en la realización de este trabajo.

A TODOS, GRACIAS

## LISTA DE NOMENCLATURAS

<b>Acrónimos</b>	<b>Definición (castellano)</b>
UV	Ultravioleta-Visible
meq	mili-equivalentes
nm	nanometros
K232	coeficiente de extinción a la longitud de onda de 232 nm
K270	coeficiente de extinción a la longitud de onda de 270 nm
Hz	Hercios
IR	Región del Infrarrojo
FT	Transformada de Fourier
CCD	Dispositivo de cargas interconectadas
A/D	Convertor de imagen análogo/digital
mW	Mili vatio
RA	Ángulo recto
Bit	Dígito binario
PC	componentes principales
LOOCV	Validación cruzada dejando uno fuera
RMSEC	Error típico residual para el colectivo de calibración
RMSECV	Raíz cuadrada de la media del error en validación cruzada
RMSEP	Raíz cuadrada de la media del error de predicción
R2	Coficiente de determinación para el proceso de calibración

r <sup>2</sup>	Coeficiente de determinación para el proceso de validación cruzada
DT	Desviación típica
RPD	Desviación residual predictiva



	<u>Página</u>
<b><i>Lista de nomenclaturas</i></b>	
<b><i>Objeto</i></b>	3
<b><i>Aim</i></b>	6
<b><u>CAPITULO I. INTRODUCCION</u></b>	
<b><i>I.1 Generalidades: El Aceite de oliva</i></b>	11
<b><i>I.2 Calidad del Aceite de oliva</i></b>	12
<b><i>I.3 Técnicas ópticas para el análisis del aceite de oliva:</i></b>	
<b><i>Fundamentos y Antecedente Bibliográficos</i></b>	19
I.3.1 Espectroscopia Raman	22
I.3.2 Espectroscopia de Fluorescencia	30
I.3.3 Visión Artificial	37
<b><i>I.3.4 Instrumentación</i></b>	40
<b>Referencias</b>	
<b><u>CAPÍTULO II. QUIMIOMETRIA</u></b>	
<b><i>II.1 Generalidades</i></b>	57
<b><i>II.2 Tratamiento de datos</i></b>	58
II.2.1 Herramientas Quimiométricas	58
II.2.2 Pre-procesamiento de datos	59
II.2.3 Tratamiento Quimiométrico de datos espectroscópicos	64
○ Análisis exploratorio de datos	66
○ Técnicas Predictivas	70
❖ Selección de muestras de calibración y validación	72

❖	Técnicas de detección de datos aberrantes	74
❖	Selección del modelo de calibración	76
❖	Validación del modelo	77
○	Técnicas de Clasificación	79
	II.2.4 Tratamiento digital de imágenes	83

## Referencias

### **CAPITULO III. PLAN DE TRABAJO**

	<b><i>III.1 Tareas iniciales</i></b>	94
	<b><i>III.2 Ejecución de los trabajos experimentales</i></b>	95

### **CAPITULO IV. METODOLOGÍA**

	<b><i>IV.1 Introducción</i></b>	99
	<b><i>IV.2 Metodología desarrollada para sistemas espectroscópicos</i></b>	100
	IV.2.1 Descripción y preparación de las muestras	100
	IV.2.2 Instrumentación y métodos de análisis	101
	IV.2.2.1 Espectroscopia Raman	101
	III.2.2.2 Espectroscopia de Fluorescencia	103
	IV.2.3 Métodos de referencia	104
	IV.2.4 Construcción y validación de modelos cuantitativos y cualitativos	109

---

<b>IV.3 Metodología desarrollada para sistemas de visión artificial</b>	113
IV.3.1 Descripción y preparación de las muestras	113
IV.3.2 Instrumentación y métodos de análisis	113
IV.3.3 Métodos de referencia	114
IV.3.4 Construcción de los modelos cuantitativos y cualitativos	114

## **Referencias**

### **CAPITULO V. RESULTADOS**

<b>V.1 Introducción</b>	125
<b>V.2 Sistemas espectroscópicos</b>	126
V.2.1. Aplicación de Espectroscopia Raman de baja resolución para el análisis de aceites de oliva oxidados	126
V.2.2. Evaluación de la calidad global del aceite de oliva mediante Espectroscopia de Fluorescencia	133
V.2.3 Sensor portátil Raman para la discriminación rápida de aceitunas según la calidad del fruto	144
<b>V.3 Sistemas de Visión Artificial</b>	
V.3.1. Determinación del índice de madurez de frutos intactos de aceituna mediante análisis de imagen	151

V.3.2. Método automático para la detección en línea de suciedad en el proceso de producción del aceite de oliva	160
V.3.3 Sistema de visión infrarroja para la detección automática de la calidad en frutos de aceituna	167
V.3.4 Sistema visible de visión para la determinación on-line de parámetros de calidad de frutos de aceituna	172
<b>V.4 Discusión de los resultados</b>	176

***Bibliografía***

**CAPITULO VI. CONCLUSIONES/ CONCLUSIONS**

<b><i>Conclusiones y Líneas Futuras</i></b>	184
<b><i>Conclusions and Future Lines</i></b>	188

**ANEXO 1**

***Publicaciones y Manuscritos en revisión***

<i>Application of low-resolution Raman spectroscopy for the analysis of oxidised olive oil</i>	193
--	-----

<i>Evaluation of the overall quality of olive oil using fluorescence</i>	
<i>Spectroscopy</i>	199
<i>A portable sensor for the rapid discrimination of olives according to</i>	
<i>fruit quality</i>	218
<i>Determination of Olive Maturity Index of intact fruits using image</i>	
<i>Analysis</i>	224
<i>Automatic and on-line method for detection of dirt on olive oil production</i>	
<i>process</i>	247
<i>Infrared machine vision system for the automatic detection of olive fruit</i>	
<i>quality</i>	264
<i>Using a visible vision system for on-line determination of quality parameters</i>	
<i>of olive fruits</i>	277

# Objeto/Aim



UNIVERSIDAD  
E  
CÓRDOBA

Actualmente en el mercado existe gran oferta de productos alimenticios diversos, lo que exige avances adecuados en las técnicas de procesamiento e industrialización con el fin de garantizar su calidad. La calidad alimentaria comprende todas las características de un producto, incluidos los factores externos (apariencia, tamaño, color, textura, sabor...) e internos (físico-químicos), factores que lo hacen aceptable para los consumidores. La calidad, en este caso del aceite de oliva, se define como un líquido aceitoso obtenido a partir de frutas recién cosechadas y saludables en condiciones óptimas de madurez, evitando cualquier manipulación que cambie su naturaleza química durante el proceso de extracción o durante el almacenamiento. La calidad del aceite de oliva puede verse afectada por varios factores, tales como factores agronómicos, incluidas las condiciones estacionales o de cultivo y sistemas de maduración, u otros factores que se relacionan principalmente con los parámetros técnicos, tales como el método y la duración del almacenamiento o la tecnología del proceso.

Los análisis de control de calidad se realizan continuamente en la industria del aceite de oliva mediante métodos formalmente establecidos para asegurar el cumplimiento de los principales parámetros de calidad, tanto en aceitunas intactas como en el aceite de oliva. La mayoría de estos análisis comprenden procedimientos lentos y costosos que son altamente empíricos y cuya precisión depende de factores experimentales. Por lo tanto se hace necesario el desarrollo de nuevos métodos analíticos alternativos para la correcta supervisión del control de calidad del aceite de oliva y los frutos de aceituna de un modo más rápido y más barato, no sólo a nivel de laboratorio, sino también en la industria.

## Objetivo

---

El objetivo general de esta tesis es proponer alternativas a los métodos utilizados actualmente para evaluar los distintos parámetros de calidad de control de aceite de oliva y de las aceitunas intactas.

En este trabajo, esto se hace mediante el desarrollo de diversos métodos ópticos no destructivos que pueden ser fácilmente transferibles a la industria. Se han desarrollado métodos basados en la espectroscopia vibracional (Raman e infrarrojo), espectroscopia de fluorescencia y sistemas de visión artificial en combinación con herramientas quimiométricas de análisis e imagen, ayuda a asegurar una calidad correcta y completa del aceite de oliva virgen. Un factor importante al determinar la calidad del aceite de oliva es la determinación de su estado oxidativo. En esta tesis se hace uso de la técnica de espectroscopia Raman no-destructiva, con un instrumento portátil y de bajo coste (**Capítulo 3**), que es un método rápido, económico y respetuoso del medio ambiente que podría acoplarse a las fibras ópticas e insertadas directamente en el proceso. Esto permite un control completo de la calidad del aceite, antes, durante y después de la transformación de las aceitunas. Otros parámetros de calidad importantes del aceite de oliva es la acidez, que se ha determinado, junto con el estado de oxidación, mediante el uso de la espectroscopia de fluorescencia.

El objetivo de este estudio fue determinar si la espectroscopia de fluorescencia contiene suficiente información sobre los parámetros de calidad para el control de rutina y se puede utilizar para una caracterización global de los aceites de oliva (**Capítulo 4**). Con el fin de asegurar un aceite de oliva de buena calidad, se debe hacer un estudio completo de las aceitunas antes de llegar a la cadena de producción. Por esta razón, en esta tesis, se ha desarrollado un sistema Raman portátil para evaluar la calidad de la pasta de aceituna. Esto se aplicó para la discriminación de las aceitunas de acuerdo a su origen "suelo" / "árbol" destinada a garantizar la calidad de las aceitunas antes de entrar en el proceso de producción del aceite (**Capítulo 5**). Por otra



parte, se ha desarrollado un sistema de visión artificial para evaluar la calidad del fruto de aceituna en la fábrica antes de su procesamiento, usando la captura de imagen digital, tanto en el espectro visible como en el rango infrarrojo.

Este sistema permite establecer una metodología de trabajo para determinar la madurez de las aceitunas at-line y on-line (**Capítulo 6** y **Capítulo 9**), la cuantificación de la posible suciedad (barro, piedras, ramas, hojas) presentes en las aceitunas (**Capítulo 7**) y la cuantificación de aceitunas sanas y / o dañadas por cualquier causa (enfermedad, accidente cerebrovascular, tierra, hielo, etc) dando lugar a la pérdida de la calidad potencial (**Capítulo 8**).

En resumen, el objetivo de esta tesis está orientado reducir el número de análisis realizados diariamente en aceitunas y aceite de oliva con el fin de obtener información sobre sus parámetros de calidad principales. Los métodos de análisis propuestos permiten una evaluación rápida y barata de las aceitunas y aceite de oliva, para desarrollar nuevas aplicaciones de los datos espectroscópicos y ópticos, que pueden ser utilizados para una completa trazabilidad y con fines de control, para caracterizar los productos y al mismo tiempo proporcionar a los consumidores una mayor confianza y calidad.

Currently on the market there is large panoply of diverse food products, which demands adequate advances in processing techniques and industrialization to ensure their correct quality. Food quality comprises all characteristics of a product, including external factors (appearance, size, colour, texture, flavour...) and internal (physicochemical) factors that made it acceptable to consumers. Quality in the case of olive oil, is defined as an oily liquid obtained from freshly and healthy fruits harvested in optimal conditions of maturity, avoiding any manipulation that changes their chemical nature during the extraction process or during storage. The quality of olive oil can be affected by several factors, such as inherent agronomical factors, including seasonal conditions or different cultivation and ripening systems. Other factors relate mainly to technical parameters, such as storage method and duration or the technology of the process.

Quality control analyses are continuously conducted in the olive oil industry using formally established methods to ensure compliance with the main quality parameters in both intact olives and in the olive oil. Most of these analyses comprise slow and costly procedures that are highly empirical and accuracy depends on experimental factors. Therefore it becomes necessary to develop alternative new analytical methods for correctly monitoring the quality control of olive oil and fruits in a faster and cheaper way, not only at laboratory level but also at the mill industry directly.

The general objective of this thesis is to propose alternatives to the currently used methodologies to evaluate the different quality control parameters of olive oil and intact olives.

In this work, this is done by developing various non-destructive optical methods that could be easily transferable to industry. Methods based on vibrational spectroscopy (Raman and Infrared), fluorescence spectroscopy and artificial vision systems have been developed in combination with chemometric and image analysis tools helping to ensure a correct and

complete quality of virgin olive oil. One important factor when determining the quality of olive oil is the determination of its oxidative status. In this thesis this is done using the non-destructive. Raman spectroscopy technique with a portable and low cost instrument (**Chapter 3**), which is a rapid, inexpensive and environmentally friendly method that could be coupled to optical fibers directly inserted into the process. This allows a complete quality control of the oil before, during and after the processing of the olives. Other important quality parameters include the acidity of olive oil, which has been determined, together with the oxidation state, by using fluorescence spectroscopy.

The aim of this study was to determine whether fluorescence spectroscopy contains enough information quality parameters for routine control and can be used for a global characterization of olive oils (**Chapter 4**). In order to assure a good quality olive oil, a complete survey of the olives should be done before reaching the production chain. For this reason, in this thesis, a portable Raman system for evaluating the quality of the olives from the olive paste has been developed. This was applied for the discrimination of olives according to their 'ground'/'sound' origin aimed to ensure the quality of olives before they entered the oil production process (**Chapter 5**). Moreover, an artificial vision system for evaluating the quality of the olive fruit in the factory before olive oil processing has been also developed using digital image capture in both the visible spectrum and in the infrared range.

This system allowed to establish a working methodology to determine at-line and on-line the ripeness in olives (**Chapter 6 and Chapter 9**), the quantification of possible dirt (mud, stones, branches, leaves) present in olives (**Chapter 7**) and quantification of healthy and / or damaged olives by any cause (disease, stroke, soil, ice, etc.) giving rise a loss of potential quality (**Chapter 8**).

In summary, this thesis aimed to reduce the number of analyses daily performed in a large numbers of olives and olive oil samples in order to obtain information about their main quality parameters. The analytical methods proposed here allowed a rapid and inexpensive assessment of olives and olive oil samples and to develop novel specifications from spectroscopic and optical data, which can be used for a complete traceability and control purposes to characterize those products and at the same time providing consumers with added confidence in their authenticity.

# Capítulo I. *Introducción*



UNIVERSIDAD  
DE  
CÓRDOBA

## **I.1 Generalidades: El aceite de oliva**

El aceite de oliva se define como el zumo oleoso del fruto de las aceitunas (*Olea Europaea*), obtenido únicamente por procedimientos mecánicos y otros procesos físicos en condiciones, generalmente térmicas, que no ocasionen la alteración del aceite y que no hayan recibido tratamiento distinto al de lavado, decantación, centrifugado y la filtración [1].

La calidad del aceite de oliva virgen emana de ser un producto natural que puede consumirse directamente, lo que no puede hacerse con ningún otro aceite vegetal.

El Convenio Internacional del aceite de oliva de 1986, modificado en 1993 y prorrogado en último lugar en 2003 reserva la denominación de “Aceite de oliva” al aceite procedente únicamente del fruto del olivo, con exclusión de los obtenidos por disolventes, por procedimientos de re esterificación y de mezcla con aceites de otra naturaleza. Según esta normativa dentro de los aceites de oliva se establecen diferentes categorías (virgen extra, virgen, lampante, refinado o de oliva) según sus características físico-químicas y organolépticas como muestra la Figura 1.

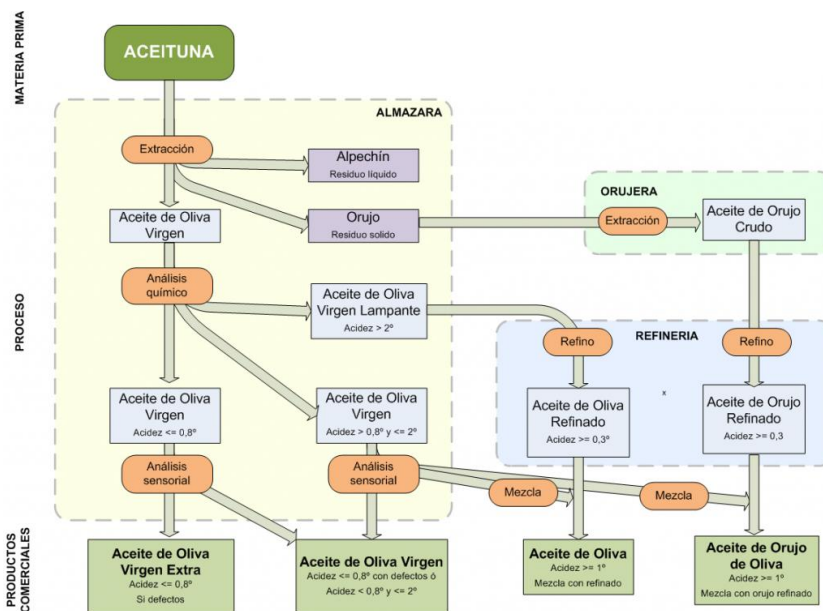


Figura 1. Tipos de aceite de oliva y de orujo

## I.2 Calidad del aceite de oliva

Genéricamente, la calidad puede definirse como “la propiedad o conjunto de propiedades inherentes a una cosa, que permiten apreciarla, mejor o peor que las restantes de su especie”.

La organización Internacional de la Normalización define el concepto de calidad en su norma ISO: 9000:2000 como “El conjunto de características de un producto, proceso o servicio que le confiere la aptitud para satisfacer necesidades expresadas o implícitas”.

Aplicada al aceite de oliva, Burón y García Teresa (1979) [2] la definen como “El conjunto de propiedades o atributos que él posee y que determina el grado de aceptación del consumidor respecto a un determinado uso”.

La calidad del aceite de oliva virgen tal y como se define en la Unión Europea, deriva de ser un fruto, deriva de ser un zumo de fruta recogida en optimas condiciones de madurez y salud y por procedimientos mecánicos y/o físicos. Cuando la aceituna se deteriora o el proceso de extracción no es el adecuado, el aceite pierde sus cualidades potenciales adquiriendo defectos que disminuyen su calidad. Los métodos analíticos ponen de manifiesto el deterioro del aceite y proporcionan información del estado del mismo.

El control de calidad de los aceites está claramente establecida en el Reglamento (CE) nº 2568/91 modificado por el (CE) nº 1989/03 de 6 de Noviembre de 2003 [3,4], que define las características físico-químicas y organolépticas de los aceites de oliva y de orujo de oliva, así como de los métodos de valoración de esas características. Recientemente, el Reglamento (UE) nº 61/2011 de la comisión por el que se modifica el Reglamento (CEE) nº 2568/91 relativo a las características de los aceites de oliva y de los aceites de orujo de oliva y sobre sus métodos de análisis, ha incluido la determinación del contenido en esteres alquílicos como nuevo parámetro a considerar en la categoría de aceite virgen extra [5].

Los parámetros de calidad reglamentada del aceite permiten evaluar sus características físico-químicas son:

**El grado de acidez:** determina la cantidad de ácidos grasos libres presentes en el aceite expresados en ácido oleico (%). La presencia de ácidos grasos libres es una anomalía resultante, entre otros factores, del mal estado de los frutos, mal tratamiento, o mala conservación de los mismos. La normativa actual, reglamento (CE) Nº 1989/2003, considera el límite para los aceites extras en 0,8°.



**Índice de peróxidos:** Determina el estado de oxidación primaria de un aceite antes de que se aprecie el olor y sabor a rancio.

Las grasas se oxidan al entrar en contacto con el oxígeno del aire. Cuando una grasa comienza a oxidarse se forman diversos compuestos; entre ellos, se encuentran los peróxidos, que se consideran los primeros productos de la oxidación.

Este índice también indica el deterioro que pueden haber sufrido ciertos componentes de interés nutricional como es la vitamina E, y se mide en mili equivalentes (meq) de oxígeno activo por kg. Su límite para el consumo es de 20.

**Absorbancia en el ultravioleta:** Esta analítica se fundamenta en la medida espectrofotométrica ultravioleta del coeficiente de extinción molar a distintas longitudes de onda: 232 y 270 nm.

K232. Indica, al igual que el índice de peróxidos, la oxidación inicial de un aceite, cuantificando su absorción de luz en la región UV a la longitud de onda de 232 nm. Su límite máximo para el virgen extra es de 2,5 y para el virgen, 2,6.

K270. Mide la absorbancia de un aceite a la longitud de onda de 270 nm. Detecta un estado oxidativo más avanzado. A medida que el proceso oxidativo avanza, los peróxidos se van modificando obteniéndose otro tipo de componentes: alfa-dicetonas o cetonas  $\alpha$ -insaturadas que absorben luz UV a distinta longitud de onda (270 nm) que los hidroperóxidos. En un aceite obtenido de una aceituna sana, que no haya sido sometido a ningún tratamiento diferente de las operaciones físicas propias de la extracción, su valor es generalmente inferior a los límites establecidos (0,22 para el virgen extra y 0,25 para el virgen). La alta absorbancia a 270 nm está relacionada

con la oxidación del aceite de oliva virgen, con el proceso de refinación o con ambas cosas a la vez.

**Contenido en ceras y en ésteres metílicos y etílicos:** el contenido en ceras y en ésteres metílicos (FAMEs) y etílicos (FAEEs) de los ácidos grasos de los aceites de oliva se ha incluido recientemente como un parámetro útil de calidad de los aceites de oliva virgen extra. La determinación de las ceras y los ésteres de alquilo se separan en función del número de átomos de carbono. El método se recomienda como herramienta para distinguir entre el aceite de oliva y el aceite de orujo de oliva, y como parámetro cualitativo para los aceites de oliva vírgenes extra, ya que permite detectar las mezclas fraudulentas de aceites de oliva vírgenes extra con aceites de menor calidad, ya sean vírgenes, lampantes o desodorizados.

La grasa se fracciona mediante cromatografía con columna de gel de sílice hidratado. La fracción eluida en primer lugar en las condiciones de se recupera y se analiza directamente mediante cromatografía de gases con columna capilar.

El límite para el contenido en ceras en aceites de oliva virgen extra que recoge el reglamento es un máximo de 250 mg/ kg, y la relación entre ésteres metílicos y ésteres etílicos no mayor a 1.5 o el sumatorio de ambos no superior a 75 mg/kg.

**Calidad sensorial del aceite:** el método oficial de análisis empleado hasta la fecha es el llamado test del panel de expertos catadores, la (UE) y el Consejo Oleícola Internacional (COI) han establecido parámetros y normativas de calidad definidas en el reglamento CE N° 2568/91, modificado por el CE N° 1989/03 del 6 de noviembre del 2003, las cuales clasifican a los aceites de oliva en categorías: Extra-virgen, virgen fino, virgen corriente y lampante, de acuerdo a su calidad y en función de sus características fisicoquímicas y

sensoriales, estableciendo así los procedimientos oficiales a seguir para cada determinación. El objetivo de este análisis es evaluar y cuantificar los atributos positivos y negativos del aceite de oliva, para la obtención de una puntuación organoléptica por parte de los catadores. El fundamento de este análisis se basa en el empleo de los sentidos olfato-gustativos y retro nasales, y su interpretación expresada por el grado de aceptabilidad de los catadores.

Estas características organolépticas se valoran según un método de cálculo estadístico de la mediana según el Reglamento nº 640/2008. Desde el punto de vista de clasificación comercial del aceite, las diferentes categorías se establecen según los límites expuestos en la Tabla 1 para la mediana del defecto y del frutado, teniendo que cumplir para la clasificación ambos límites.

Categoría	Mediana (Me) del defecto de intensidad	Mediana del frutado
Virgen Extra	0	>0
Virgen	$0 < Me \leq 3.5$	>0
Lampante	>3.5	=0

**Tabla1.** Clasificación de los aceites de oliva vírgenes según los resultados del análisis sensorial (Reglamento nº 640/2008)

La obtención de estos aceites de oliva de calidad depende de una gran cantidad de factores como los agronómicos como pueden ser la variedad, el medio agrológico, la recolección o el transporte, o factores de carácter industrial como todos los referentes al proceso de extracción.

Uno de los primeros factores que influyen sobre la calidad y seguridad del aceite de oliva virgen es la producción del fruto del que se obtiene, es decir, de la aceituna. Es de gran importancia el mantener un control integral de la calidad higiénica desde el origen hasta el producto final, y de aquellos aspectos de la producción primaria que pueden afectar al aceite. Para obtener aceites de calidad se deben cuidar todos los procesos previos, tanto del olivo como del sistema de recolección y el transporte a la industria, siendo éstos determinantes para la obtención de aceites de calidad.

No se puede desligar por tanto, los factores de obtención del fruto, del proceso de extracción de aceite, su almacén, envasado y distribución.

Estos factores que afectan a la calidad del aceite se pueden clasificar en agronómicos e industriales.

Los factores agronómicos se pueden dividir en intrínsecos (no se pueden modificar) y extrínsecos (pueden ser controlados). Entre los factores agronómicos intrínsecos que más afectan a la calidad encontramos el medio agrológico, éste tiene incidencia en la composición acídica de los aceites de un mismo cultivar [6] y presenta una clara influencia sobre la fracción insaponificable lo que se traduce en aceites de diferentes caracteres sensoriales. Entre los factores agronómicos extrínsecos se recogen las prácticas culturales, la recolección y el transporte del fruto a la almazara. De ellas las que más influencia tienen sobre la calidad son, los tratamientos fitosanitarios y la recolección del fruto. Tres son los aspectos a considerar en la recolección del fruto bajo la óptica de la calidad: la época, la procedencia del fruto y la forma de realizarla.

Entre los factores de carácter industrial es importante la separación del fruto de diferente calidad, siendo esencial su separación en recepción, limpieza y lavado como en etapas posteriores. La molturación inmediata del fruto es muy importante también para la obtención de aceites de calidad, así como la molienda y el batido de la pasta han de hacerse controlando parámetros como son la temperatura, tiempo y separación de fases para no alterar la calidad de los aceites.

Una vez conseguidos los aceites, es necesario conservarlos en depósitos hasta su comercialización, teniendo en cuenta aspectos como temperatura, tamaño y material de los depósitos.

Como hemos visto hay numerosos factores que pueden influir en la calidad del producto final, por lo que es necesario mejorar los procedimientos de control del proceso productivo. Tanto los parámetros de análisis físico-químicos como de calidad sensorial utilizados como métodos de referencia para la valoración de la calidad del aceite de oliva, poseen desventajas como la de ser métodos altamente empíricos y su exactitud depende de factores experimentales, además son procedimientos lentos y costosos en tiempo y disolventes. Por ello se hace indiscutible la necesidad de estudiar y poner a punto nuevas tecnologías analíticas que mejoren el control de calidad en la industria de producción de aceite de oliva.

En este como otros sectores el principal impedimento para el avance científico y la implantación de nuevas tecnologías es el económico, de ahí surge la necesidad de estudiar técnicas analíticas rápidas, limpias y en las que el coste de los equipos no sea elevado como son las técnicas ópticas.

### I.3 Técnicas ópticas en el análisis del aceite de oliva

Actualmente las técnicas más prácticas y satisfactorias para la evaluación no destructiva de la calidad y clasificación de productos agrícolas son las técnicas electro-ópticas, basadas en las propiedades ópticas de los productos [7]. Se trata de métodos cuyo fundamento se centra en medir la interacción entre la materia y la energía en forma de luz.

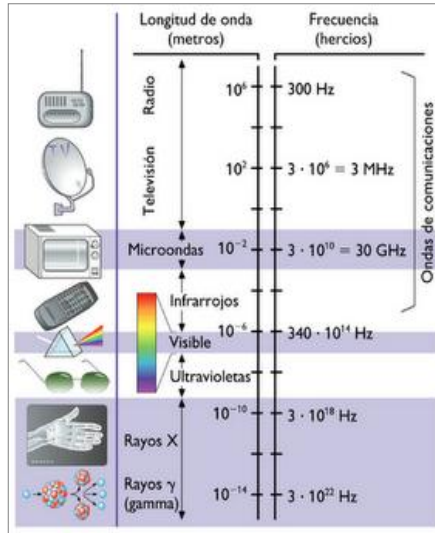
La información científica disponible señala que las técnicas de espectroscopia y otras técnicas ópticas están emergiendo con fuerza recientemente como técnicas analíticas en tecnología de los alimentos debido a la necesidad de reducir tanto el tiempo como el coste de los análisis unido a la creciente demanda de controles en línea y a las limitaciones de las técnicas tradicionales para resolver algunas cuestiones analíticas.

En sentido amplio, el concepto “luz” es un tipo de energía radiante (radiación electromagnética) que se transmite a través del espacio mediante un movimiento ondulatorio con una velocidad de propagación, una longitud de onda y una frecuencia determinadas.

Dicha longitud de onda permite clasificar la luz en varios rangos, abarcando el espectro electromagnético en todo su rango.

Estas ondas electromagnéticas tienen componentes eléctricos y magnéticos, alternantes y perpendiculares entre sí, y transportan energía de un lugar a otro.

Desde el siglo XIX, al conjunto de radiaciones que van aproximadamente desde 10<sup>-8</sup> m hasta 10<sup>-3</sup> m (ultravioleta, visible e infrarrojo) se denominó “espectro óptico”, y corresponde con la gama completa de la luz solar, que es sólo una parte del espectro electromagnético.



**Figura 2.** Escala que relaciona las diferentes ondas electromagnéticas con su frecuencia (Hz), longitud de onda (m).

En esta memoria de tesis los equipos ópticos ensayados se centran sólo en las zonas del espectro óptico correspondientes al visible e infrarrojo cercano:

Coloquialmente se entiende por luz sólo la parte visible (400-800 nm) del espectro electromagnético, es decir, el rango de longitudes de onda que es percibido por el ojo humano. El color de los objetos depende del espectro de la luz que incide y de la absorción del objeto, la cual determina qué ondas son reflejadas.

Los rayos infrarrojos o “radiación de calor” (800-15000 nm); están asociados al calor, pues a temperatura normal los objetos emiten espontáneamente radiaciones en el campo de los infrarrojos. Se considera dividido en tres regiones: infrarrojo cercano (800-1400 nm), infrarrojo medio (1400-3000 nm) e infrarrojo lejano (3000-10000 nm). Estas radiaciones son producidas por las vibraciones atómicas de los cuerpos calientes y absorbidas por ciertas moléculas [8].

La luz interacciona con la materia de varias maneras. Parte de la luz se modifica por factores geométricos, macroscópicos, y microscópicos, que causan distintas modificaciones en la radiación. Otra parte es absorbida por las sustancias que componen la materia, que queda excitada, y puede emitir de nuevo energía. Los efectos sobre la materia son muy diferentes dependiendo de la longitud de onda incidente sobre la superficie del objeto y de las características del propio producto. El análisis de la respuesta a esta interacción, proporciona valiosa información sobre la materia que se estudia, que van a influir en las características de los movimientos de las ondas; estos fenómenos ópticos clásicos se pueden agrupar en los siguientes tipos de interacción [9,10]

- Transmisión: refracción y dispersión.
- Reflexión: reflexión especular, dispersión y dispersión Raman.
- Absorción: absorción atómica, absorción molecular e inducida por un campo magnético.
- Emisión: no radiantes (tras una absorción), rayos X, rayos  $\gamma$ , fluorescencia y fosforescencia.

A continuación se describen con mayor detalle las tecnologías ópticas empleadas en la presente Memoria, así como los antecedentes bibliográficos descritos hasta la fecha se recogen de manera resumida las tecnologías utilizadas en esta Memoria para los diferentes trabajos con el objetivo de simplificar el análisis de los parámetros de calidad del aceite de oliva y del fruto de aceituna.

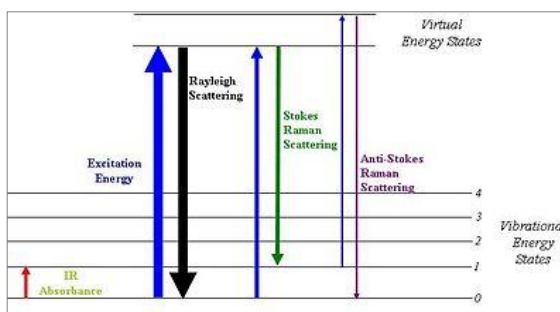


### I.3.1 Espectroscopia Raman

#### Fundamentos

La espectroscopia Raman es una técnica fotónica de alta resolución que proporciona en pocos segundos información química y estructural de cualquier material permitiendo su identificación. El análisis mediante espectroscopia Raman se basa en el examen de la luz dispersada por un material al incidir sobre él un haz de luz monocromático. Una pequeña parte de esta luz es dispersada inelásticamente experimentando ligeros cambios de frecuencia que son característicos del material analizado.

El efecto de dispersión Raman surge a partir de la interacción de la luz incidente con los electrones de una molécula iluminada. En la dispersión Raman la energía de la luz incidente no es suficiente para excitar la molécula a un nivel electrónico de mayor energía. Así el resultado de la dispersión Raman es cambiar el estado vibracional de la molécula.



**Figura 3.** Diagrama energético de una molécula mostrando el origen de la dispersión Raman (efecto Raman no resonante), y los diferentes mecanismos de los efectos Stokes y anti-Stokes.

Para que una molécula exhiba el efecto Raman, la luz incidente debe inducir un cambio en el momento dipolar o un cambio en la polarizabilidad molecular.

La luz dispersada contiene una pequeña porción de la luz debida a la dispersión Raman, además de la debida a la dispersión Raleigh normal. La dispersión Raman contiene líneas Stokes y anti-Stokes; sus frecuencias corresponden a la suma y diferencia de las frecuencias de la luz incidente y las frecuencias vibracionales moleculares permitidas. Cuando los fotones interactúan con una molécula, una parte de su energía se puede convertir en varios modos de vibración de la molécula. La luz dispersada pierde energía equivalente a la energía dada a las vibraciones moleculares (efecto Raman Stokes). Si la molécula se encontraba inicialmente en un estado rotacional o vibracional excitado, es posible que la luz dispersada tenga mayor energía que la incidente (efecto Raman anti-Stokes), siempre y cuando el sistema decaiga a un estado de menor energía que la del estado inicial. Como normalmente la población de los estados excitados es mucho menor que la del estado básico, la dispersión Raman Stokes es mucho más intensa que la dispersión anti-Stokes.

La espectroscopia Raman se basa en aprovechar la dispersión inelástica de la luz que se produce cuando un material es irradiado con luz monocromática, como es el caso del láser. La manera de recoger esta luz que contiene información y convertirla en datos útiles, es mediante la utilización de un equipo Raman y la posterior representación del espectro obtenido. Un espectro Raman no es más que la representación en un gráfico de la intensidad de las bandas Raman (eje de ordenadas) en relación con la longitud de onda (eje de abscisas). Normalmente, un espectro Raman se sitúa entre los  $100 \text{ cm}^{-1}$  y los  $4000 \text{ cm}^{-1}$ , con una anchura típica de las bandas que puede variar entre los  $5\text{-}20 \text{ cm}^{-1}$ .

La tecnología Raman es muy similar a la infrarroja, ambas son técnicas complementarias a la hora de estudiar los movimientos vibracionales. Mientras que la espectroscopia IR depende del cambio en la distribución de cargas de la molécula con la vibración, la espectroscopia Raman depende de

un cambio inducido en la distribución de cargas. Estas diferencias pueden llegar a ser muy importantes en la identificación de compuestos y por eso pueden ser utilizadas como técnicas complementarias.

La espectroscopia Raman ofrece varias ventajas como técnica de análisis respecto a otras tecnologías de análisis vibracional como FTIR y NIR, como son: Raman tiene gran interés para el estudio de sustancias en disolución acuosa, ya que el agua absorbe fuertemente en IR tapando otras bandas y en Raman no ocurre, con lo cual las interferencias del disolvente en disoluciones acuosas es escasa.

La medida mediante el equipo Raman se realiza directamente sobre el material a analizar, no requiere ningún tipo de preparación de la muestra (como el molido o la homogeneización), se pueden analizar materiales en estado sólido, líquido o gaseoso, y a través de contenedores. Por todo ello, los requerimientos de preparación de la muestra suelen ser bastante menores que con otras tecnologías se requieren la preparación de muestras muy delgadas y homogeneizadas. Los métodos para análisis son estables, no son afectados por variaciones de tamaño de partículas, humedad o temperatura ambiente.

La aparición en el mercado de equipos Raman portátiles especializados para los requisitos de la industria con tiempos de análisis de pocos minutos por muestra. Existen equipos Raman portátiles de muy pequeñas dimensiones. El bajo coste de este equipo es otra de las ventajas que presenta este trabajo ya que la tecnología infrarroja y otras técnicas utilizadas requieren unos equipos de análisis de valor igual o superior al necesario para llevar a cabo el análisis Raman.

En espectroscopia del infrarrojo se pierde potencia de procesos de reflexión y dispersión, para ello se suele hacer una medida de la corriente oscura y de un

blanco. Este paso no es necesario para esta tecnología Raman, lo que aumenta las ventajas de su aplicación.

Hay dos tecnologías comúnmente usadas en la instrumentación Raman. Ellas son FT Raman y Raman Dispersivo. La diferencia fundamental entre las dos tecnologías es la forma en que la señal Raman es analizada. En un sistema FT Raman el análisis espectral es logrado usando un interferómetro y transformadas de Fourier, mientras que en un sistema Raman dispersivo el análisis espectral es logrado mediante el uso de un monocromador.

La tecnología FT-Raman ha sido muy usada para definir las bandas asociadas a los distintos compuestos presentes en el aceite. La mayor parte de estas bandas se encontraba en la región por debajo de  $1800\text{ cm}^{-1}$  y la mayoría coinciden con las obtenidos mediante equipos Raman dispersivos, alrededor de  $1267\text{ cm}^{-1}$  ( $\delta$  cis [= CH]),  $1.302\text{ cm}^{-1}$  (en fase de giro del metileno) y  $1.442\text{ cm}^{-1}$  ( $\delta$  [CH<sub>2</sub>]),  $1.655\text{ cm}^{-1}$  (doble enlace cis  $\nu(\text{C} = \text{C})$ ) y  $1,747\text{ cm}^{-1}$  (éster  $\nu(\text{C} = \text{O})$ ) [10]. En el espectro obtenido del instrumento Raman portátil, existe una importante área por debajo de  $800\text{ cm}^{-1}$  que no coincide con el espectro de FT-Raman y todavía no ha sido descrito en la literatura.

#### *Aplicaciones al proceso de elaboración de Aceite de Oliva*

La caracterización de la aceituna por medio de tecnologías ópticas ha sido desarrollada principalmente por espectroscopia NIR y Raman.

A continuación en las siguientes tablas se muestra una revisión general de análisis aplicados al control de calidad en el proceso de elaboración de Aceite de Oliva mediante las tecnologías NIR y Raman.

Tecnología Infrarrojo cercano:

Título	Referencia
Medición de la adulteración de los aceites de oliva por espectroscopia de infrarrojo cercano	Wesley I. J. J. Am. Oil Chem. Soc. 72, pp. 289-292. 1995
Autenticación de aceite de oliva adulterado con aceites vegetales utilizando espectroscopia infrarroja con transformada de Fourier.	Tay, A. Technol. 35, 99-103. 1995
Análisis cuantitativo de adulterantes potenciales de aceite extra virgen de oliva mediante espectroscopia de infrarrojo	Lay, Y. Food Chem. 53, pp.95-98. 1995
Determination of oil and moisture in two-phases olive pomace by using near-infrared spectroscopy	García Mesa, J.A. Grasas y Aceites, Vol 47, No 5 (1996)
Uso de un sensor NIR on-line para la caracterización de aceites de oliva	Jiménez, A. Grasas y Aceites nº 51(5)311–315. 2000
Comparación de infrarrojo cercano, infrarrojo por transformada de Fourier y la transformada de Fourier-Raman para determinar la adulteración del aceite de orujo de oliva en aceite extra virgen de oliva	Yang H. & Irudayaraj J. J. Am. Oil Chem. Soc. 78(9), pp. 889-895. 2001
Análisis de aceituna intacta mediante NIRS	León, L. Grasas y Aceites 54, 41-47. 2003
Evaluación de la calidad mediante Espectroscopia de reflectancia NIR	Mailer, R.J. JAOCS 81, 823-827. 2004

Efecto del año de cosecha en aceitunas mediante Espectroscopia de reflectancia NIR	León, L. J. Agric. Food Chem. 52, 4957-4962. 2004
FT-NIR como herramienta para la clasificación y análisis cuantitativo de fruto de aceituna	Ayora-Cañada, MJ. Spectr Lett 38: 769-785. 2005
Determinación de parámetros en aceites comestibles mediante NIR	Armenta, S. Analytica Chimica Acta, 596(2),330-337. 2007
Evaluación preliminar de la aplicación de espectroscopia FT-IR para el control del origen geográfico y calidad de aceites de oliva vírgenes	Bendini, A. Journal of Food Quality 30 (4), 424-437. 2007
Determinación del origen geográfico y composición de aceites de oliva vírgenes mediante análisis quimiométrico de espectros NIR	Galtier, O. Anal Chim Acta. 9 595 (1-2): 136-44. 2007
Caracterización y clasificación de aceites de oliva italianos mediante infrarrojo medio y cercano	Sinelly, N. Journal of NearInfrared Spectroscopy, 16(3), 335-342. 2008
Estudio del envejecimiento oxidativo de los aceites comestibles mediante espectroscopia MCR-ALS métodos quimiométricos	Dreau, Y. Talanta, 77(5), 1748-1756. 2009
Predicción NIR de la humedad del fruto, acidez libre y contenido de aceite en aceitunas intactas	Cayuela, JA. Grasas y Aceites, Vol 60, nº 2. 2009
The Use of Near-Infrared Spectrometry in the Olive Oil Industry	Armenta, S.; Critical Reviews In Food Science And Nutrition

	Volume: 50 Issue: 6 Pages: 567-582. 2010
Crecimiento y maduración de la aceituna visto por espectroscopia vibracional	López-Sánchez, M. Journal of Agricultural and Food Chemistry, 58, 1, 82-87. 2010
Caracterización de aceitunas de mesa mediante espectroscopia NIR	Casale. Food Chem 122(4):5. 2010
Evaluación no destructiva de la maduración de la aceituna mediante espectroscopia en el infrarrojo cercano portátil	Gracia, A. Grasas y Aceites, Vol 62, nº 3. 2011
Caracterización no destructiva de aceitunas de mesa mediante espectroscopia NIR	Morales-Sillero, A. Journal of Food Engineering , Volume 107 (1). 2011

Tecnología Raman:

Título	Referencia
Detección de la adulteración del aceite de oliva virgen por espectroscopia Raman de transformada de Fourier	Baeten, V. Journal of Agricultural and Food Chemistry, 44(8), 225-230. 1996
Estudio del uso de la espectroscopia molecular para la autenticación de los aceites de oliva vírgenes parte	Davies A. N. Appl. Spectrosc. 54, pp. 1864-1867. 2000

Raman de transformada de Fourier espectroscopia	
Determinación del contenido de ácidos grasos libres en el aceite de oliva y aceitunas por espectrometría FT-Raman	Muik, B. <i>Analytica Chimica Acta</i> , 487, 2, 211-220. 2003
Espectrometría Raman de transformada de Fourier para el análisis cuantitativo del contenido de aceite y humedad en aceitunas	Muik, B. <i>Applied Spectroscopy</i> , 57, 2. 2003
Evaluación cuantitativa rápida de la adulteración de los aceites de oliva virgen con aceites de avellana utilizando espectroscopia Raman y Quimiometría	López-Díez EC. <i>J Agric Food Chem.</i> 51(21):6145-50. 2003
Sensor portátil Raman para la rápida identificación del aceite de oliva	Xiaohua, Q. <i>Technical Proceedings of the NSTI</i> , Vol. 2 pp. 97. 2008
Espectroscopia Raman visible para la discriminación de los aceites de oliva a partir de diferentes aceites vegetales y la detección de la adulteración	El-Abassy, R.M. <i>Journal of Raman Spectroscopy</i> , 40, 9, 1284-1289. 2009
Determinación Rápida de ácidos grasos libres en el aceite extra virgen de oliva por espectroscopia Raman y análisis multivariado	El-Abassy, R.M. <i>Journal of the American Oil Chemists Society</i> , 86, 6, 507-511. 2009
Autenticación rápida de la adulteración del aceite de Oliva mediante espectrometría de Raman	Zou, M.Q. <i>Journal of Agricultural and Food Chemistry</i> , 57, 14, 6001-6006. 2009



Evaluación de la degradación de calefacción convencional y microondas inducida en carotenoides en el aceite de oliva por espectroscopia Raman VIS y métodos clásicos	El-Abassy, R.M. Food Research International, 43, 3, 694-700. 2010
Predicción del origen geográfico de Aceite de Oliva Virgen por Quimiometría y tratamiento de espectros Raman	Dellaa, Y. XXII International Conference on Raman Spectroscopy, AIPConference Proceedings, 1267, 562-563 2010
Composición y autenticación de aceite de oliva virgen de las regiones francesas por tratamiento quimiométrico de espectros Raman	Korifi, R. Journal of Raman Spectroscopy 42, 7, 1540-1547. 2011
Autenticación rápida de aceite de oliva por espectroscopia Raman con el uso de componentes principales	Zhang, X.F. Analysis Analytical Letters, 44, 12, 2209-2220. 2011
Caracterización de biofenoles antioxidantes en aceite de oliva por métodos espectroscópicos	Paiva-Martins F. J Sci Food Agric. 30;91(2):309-14. 2011

### 1.3.2 Espectroscopia de Fluorescencia

#### *Fundamentos*

La espectroscopia de fluorescencia es una técnica analítica cuya teoría y metodología han sido ampliamente explotada para estudios de la estructura molecular por su accesibilidad y fácil manejo. Una indicación de esta popularidad es el creciente número de publicaciones de investigación sobre

fluorescencia, así como la introducción de nuevos instrumentos comercialmente disponibles para el análisis de fluorescencia. La fluorescencia visible inducida por luz ultravioleta se define como el fenómeno físico que se produce cuando los átomos que constituyen las moléculas de los materiales absorben la energía de la radiación ultravioleta, excitándose los electrones de sus orbitas interiores a niveles energéticos más altos. Esta excitación cesa rápidamente y estos electrones vuelven al nivel de origen, transformándose el exceso energético en otras formas de menor energía como vibración y radiación electromagnética de longitud de onda más larga que la incidente, la fluorescencia.

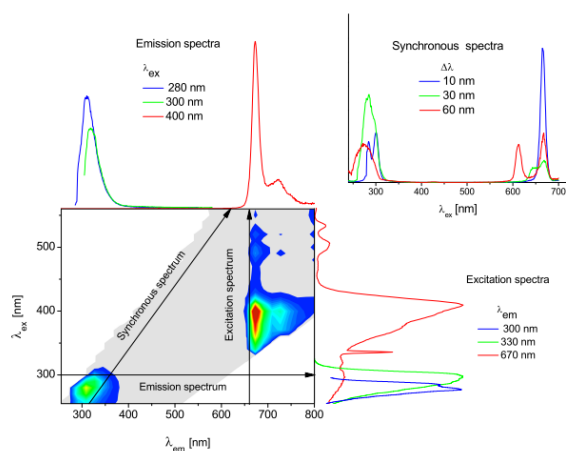
En la espectroscopia de fluorescencia, primero se excita la muestra mediante la absorción de un fotón de luz, desde su estado electrónico basal a uno de los distintos estados vibracionales del estado electrónico excitado. Las colisiones con otras moléculas causan que la molécula excitada pierda energía vibracional hasta que alcanza el estado vibracional más bajo del estado electrónico excitado.

La molécula desciende luego a uno de los distintos niveles de vibración del estado electrónico basal, emitiendo un fotón en el proceso. Como las moléculas pueden caer a cualquiera de los diferentes niveles de vibración en el estado basal, los fotones emitidos tendrán diferentes energías y, por lo tanto, frecuencias. Así pues, mediante el análisis de las diferentes frecuencias de luz emitida por espectrometría de fluorescencia, junto con sus intensidades relativas, se puede determinar la estructura de los diferentes niveles de vibración.

En un proceso de fluorescencia convencional, se miden las diferentes frecuencias de luz fluorescente emitida por una muestra, manteniendo la luz de excitación a una longitud de onda constante. A esto se le llama espectro

de emisión. Un espectro de excitación se mide mediante el registro de una serie de espectros de emisión utilizando luz de diferentes longitudes de onda.

Otro tipo de medición de este proceso, es la llamada fluorescencia sincrónica. En ella se elige una variación de longitud de onda constante entre las longitudes de onda de excitación y emisión de la especie a determinar, y se mueven simultáneamente a la misma velocidad los dos monocromadores. El espectro de fluorescencia sincrónica nos aporta una selectividad mayor de la convencional.



**Figura 4.** Diferentes mediciones de la Fluorescencia. Fluorescencia de una muestra de aceite de oliva virgen extra diluido (1% v/v en hexano)

Las ventajas de esta técnica son su velocidad de análisis, la no necesidad de disolventes y reactivos para el análisis y el requisito de sólo pequeñas cantidades de muestra, además de ser una técnica no invasiva. Estas ventajas han permitido utilizar esta tecnología para múltiples trabajos relacionados con la tecnología de los alimentos.

Aplicaciones al proceso de elaboración de Aceite de Oliva

Se ha utilizado la espectroscopia de fluorescencia en múltiples trabajos, a continuación en la tabla se recogen los trabajos realizados hasta la fecha.

Espectroscopia de Fluorescencia:

Título	Referencia
Fluorescencia de aceites de olive bajo luz ultravioleta	Glantz A. Industrial & Engineering Chemistry Analytical Edition, Vol.2 No.3, pp. 256-258. 1930
Evaluación exploratoria del deterioro del aceite de fritura mediante espectrometría de Fluorescencia	Engelsen, S. A. Journal of the American Oil Chemistry Society, 74, 1495–1508 1997
Medida del espectro de Fluorescencia de aceites de oliva y otros aceites vegetales	Kyriakidis N. B. Journal of AOAC International, Vol.83, No.24, pp. 1435-1439. 2000
Detección de aceite de avellana en aceite de oliva virgen por un método espectrofluorimétrico.	Sayago A. European Food Research and Technology, Vol.218, No.5, pp. 480-483. 2004
Análisis de agrupaciones aplicado al análisis exploratorio de los Aceites de Oliva Comerciales de España mediante espectroscopia de fluorescencia de excitación- emisión	Guimet, F. Journal of Agricultural and Food Chemistry, Vol.52, No.22, pp. 6673-6679. 2004

<p>Aplicación del Análisis de Componentes Principales y análisis factorial paralelo al análisis exploratorio Aceites de Oliva por medio de la matriz de espectroscopia fluorescencia de excitación-emisión</p>	<p>Guimet, F. <i>Analytica Chimica Acta</i>, Vol.515, No.1, pp. 75-85. 2004</p>
<p>Uso espectroscopia de fluorescencia de excitación-emisión y de tres vías de análisis</p>	<p>Guimet, F. <i>Analytica Chimica Acta</i>, Vol.544, No.1-2, pp. 143-152. 2005</p>
<p>Aplicación de la matriz de fluorescencia excitación-emisión combinado con el análisis discriminante lineal de Fisher para la clasificación del aceite de oliva</p>	<p>Guimet, F. <i>Chemometrics and Intelligent Laboratory Systems</i>, Vol.81, No.4, pp. 94-106, ISSN 0169-7439. 2006</p>
<p>Espectroscopia de fluorescencia de excitación - de emisión en combinación con métodos de tres vías de análisis como una técnica complementaria caracterización del aceite de oliva</p>	<p>Guimet, F. <i>Journal agricultural and Food Chemistry</i> pp. 9319-9328. 2005</p>
<p>Clasificación de los aceites comestibles utilizando espectroscopia de fluorescencia de barrido</p>	<p>Sikorska, E. <i>Food Chemistry</i>, 89, 217–225. 2005</p>
<p>Espectroscopia de fluorescencia para el seguimiento de deterioro de aceite extra virgen de oliva durante el calentamiento</p>	<p>Cheikhousman, R. <i>Analytical and Bioanalytical Chemistry</i>, 382(6), 1438–1443. 2005</p>
<p>Origen de la Denominación Aceite de Oliva Virgen francés de origen registrado predicho por el quimiométrico del espectro de fluorescencia</p>	<p>Dupuy, N. <i>Journal of Agricultural and Food Chemistry</i>, Vol.53, nº.24</p>

excitación-emisión	pp. 9361-9368. 2005
Clasificación de los aceites vírgenes comestibles de oliva lampante mediante espectroscopia de fluorescencia sincrónica y de luminiscencia total	Poulli K. Analytica Chimica Acta, Vol.542, No.2, pp. 151-156. 2005
Espectroscopia de fluorescencia sincrónica para la determinación cuantitativa de la adulteración del aceite de oliva virgen con aceite de girasol	Poulli,K.I.Analytical Bioanalytical Chemistry 386 (5) 1571-1575. 2006
Espectroscopia de fluorescencia de excitación -emisión combinada con métodos de tres vías de análisis como una técnica complementaria caracterización del aceite de oliva	Zandomeneghi, M. Journal of Agricultural and Chemistry 54 (14), pp. 5214-5215. 2006
Detección de la presencia de aceite de avellana en aceite de oliva refinado por Espectroscopia de Fluorescencia	Sayago, A. Journal of Agricultural and Food Chemistry, Vol.55, No.6, pp. 2068-2071. 2007
Espectroscopia de Fluorescencia para el control de aceite de oliva extra virgen durante el almacenamiento	Sikorska, E. Journal of Food Science and Technology, Vol.43, No.1-2, pp. 52-61. 2008
Aplicación de la espectroscopía de fluorescencia sincrónica para la determinación de la adulteración del aceite de oliva virgen extra	Dankowska, A. European Journal Science and Technology 111 (12), pp. 1233-1239. 2009
Determinación de polifenoles en el aceite de oliva	Gracia, A. Symposium on Food

por medio de análisis de fluorescencia	Processing, Monitoring Technology in Bioprocesses and Food Quality Management, pp. 1364-1368. 2009
Supervisión de la oxidación del aceite de oliva con el estrés térmico a través de la espectroscopia de fluorescencia sincrónica y ensayos d	Poulli, K.I. Food Chemistry 117, 499–503. 2009
Evaluación de deterioro térmico del aceite de oliva virgen por espectroscopia de fluorescencia	Tena, N. Journal of Agricultural Food Chemistry 57 (22), pp. 10505-10511. 2009
Cambios químicos del aceite de oliva virgen termo-oxidados determinada por espectroscopia fluorescencia de excitación-emisión (EEFS)	Tena, N. Research International 45 (1), pp. 103-108. 2012

Autores como Sikorska han confirmado la información disponible en los espectros de fluorescencia y demostraron que en el espectro de los aceites de oliva vírgenes de entre 400 y 700 nm medido a una longitud de onda de excitación de 365 nm presentan claras diferencias en comparación con los espectros de otros aceites vegetales. Los aceites de oliva vírgenes presentan dos picos a 445 y 475 nm, un pico intenso a 525 nm, y otro pico a 681 nm. Trabajos posteriores confirmaron que los picos a 445 y 475 nm estaban relacionados con los productos de oxidación de los ácidos grasos y el pico a 525 nm es derivado de la vitamina E. El pico a 681 nm se relaciona con los pigmentos del aceite más exactamente con las clorofilas. La baja intensidad

de los picos a 445 y 475 nm de aceites de oliva vírgenes se debe a su alto contenido de ácidos grasos monoinsaturados y antioxidantes fenólicos, que les proporcionan una mayor estabilidad frente a la oxidación, así los aceites refinados muestran sólo un pico intenso y amplio en torno a los 400-560 nm, que es debido a un estado de oxidación mayor de estos aceites como resultado de su gran contenido de ácidos grasos poli insaturados. Así esta técnica óptica nos permite diferenciar y clasificar los aceites según su composición y estado físico de una forma limpia, rápida y segura.

### I.3.3 Visión artificial

#### *Fundamentos*

La visión artificial es una técnica que utiliza señales digitales, obtenidas a través de un dispositivo de imagen o video, y permite la obtención, procesamiento y análisis de información de estas señales para reconocer una escena o características de una imagen.

La visión artificial la componen un conjunto de procesos destinados a realizar el análisis de imágenes. Estos procesos son: captación de imágenes, memorización de la información, procesado e interpretación de los resultados.

Esta técnica trata de simular el proceso del sentido de la visión de los seres vivos, según el cual el objeto es captado por los receptores de la retina y es transformado en impulsos nerviosos que son procesados por el cerebro. Este proceso lo lleva a cabo mediante la captación de imágenes en línea mediante cámaras basadas en matrices de sensores sensibles a la luz (CCD), el posterior tratamiento de las mismas mediante técnicas de procesamiento de imágenes y extracción de la información deseada del producto.



En esta tecnología los fotones son captados por una matriz de fotoreceptores que transforman la intensidad lumínica en carga eléctrica que va siendo transmitida a la fila adyacente para su lectura.

La carga analógica es digitalizada por un conversor A/D (8 bits=256 niveles de gris).

Los elementos de un sistema de visión son el sistema de iluminación (LED), la cámara de captura imagen (CCD) y el procesamiento de imagen (procesador y algoritmos de pretratamiento y filtrado de imagen, segmentación y reconocimiento de formas y clasificación).

#### *Aplicaciones al proceso de elaboración de Aceite de Oliva*

La visión artificial se utiliza para diferentes propósitos como automatizar procesos de control de calidad: detección de defectos y fallos en materias primas y productos.

Las principales aplicaciones de la visión artificial en el campo agroalimentario se pueden agrupar en:

Automatización de procesos para el control de calidad: detección de defectos o fallos en la materia prima, producto final o envases.

Automatización de procesos de clasificación en base a la forma, color y tamaño.

Automatizar procesos de producción: ajustar las condiciones del proceso en base a variables las medidas mediante análisis de imagen.

Automatizar procesos de detección de suciedad o cuerpos extraños.

En el área agroalimentaria, y más concretamente, el olivar y aceite de oliva se ha utilizado esta tecnología con varios propósitos, a continuación en la tabla se recogen estos trabajos

Título	Referencia
La aplicación de un algoritmo rápido para la clasificación de las aceitunas por visión artificial	Díaz, A. Food Research International 33, 30-309. 2000
Comparación de los tres algoritmos en la clasificación de las aceitunas de mesa por medio de visión artificial	Díaz, A. Journal of Food Engineering 61, 101-107. 2004
Clasificación de oliva de acuerdo a daños externos análisis de imagen	Riquelme M.T. Journal Food Engineering Volume 87, Issue 3, August 2008, Pages 371-379. 2007
Software de redes neuronales artificiales para estimación en tiempo real de parámetros cualitativos de aceite de oliva durante la extracción continua	Furferi R. Computers and Electronics in Agriculture Vol. 55, Issue 2, pp. 115-131. 2007
Clasificación de oliva de acuerdo a daños externos análisis de imagen	Riquelme, M.T. Journal of Food Engineering, 87, 371-379. 2008
Predicción automática del índice de madurez de aceitunas basada en el método ANN	Furferi, R. Journal of Food Engineering 101, 318-328. 2010

### I.3.4 Instrumentación

Los distintos procedimientos analíticos se han llevado a cabo con los siguientes instrumentos:

#### *Espectrómetro Raman*

El instrumento Raman utilizado fue un modelo de espectrómetro portátil RH-3000 (Ocean Optics). Este equipo está dotado de un láser de diodo a la longitud de onda de 785 nm conjuntamente con detectores de estado sólido tipo CCD, lo que permite lograr instrumentos de alta sensibilidad y gran solidez. Éste láser se relacionó con una fibra óptica para ser llevado a través de un recipiente de vidrio. Los espectros Raman se registraron 200 a 2700  $\text{cm}^{-1}$ , con una resolución espectral de 10  $\text{cm}^{-1}$ . El diodo láser del instrumento tenía un rango de intensidad de 0-290 mW.

El equipo Raman dispersivo utilizado cuenta con un láser de 785 nm, a estas longitudes de onda surge el fenómeno de la fluorescencia. Esta característica del equipo puede hacerlo quizá más interesante ya que este fenómeno puede poner de manifiesto compuestos que con infrarrojo o equipos de transformada de Fourier no se detectan debido a las longitudes de onda a la que excitan la muestra.

El equipo utilizado en los diferentes trabajos de esta Tesis es un equipo portátil, las dimensiones físicas del equipo son 20 x 18 cm y un peso de unos 4 kilos. La figura 5 muestra el equipo Raman portátil utilizado en los trabajos desarrollados en esta Tesis.



Figura. 5 Espectrómetro Raman utilizado

Aunque pueden darse pequeñas variaciones en función de las necesidades de análisis, en general, un equipo de espectroscopia Raman estará formado por los siguientes instrumentos:

<i>Instrumentos</i>	<i>Descripción</i>
Láser	Es la fuente de luz monocromática que va a excitar la muestra. El tipo de láser puede variar en función de la técnica utilizada, pero en general hablaremos de láseres de onda continua, con longitudes de onda variables entre el ultravioleta y el infrarrojo; o también podemos nombrar los láseres pulsados.
CCD	Una vez separada la señal en las distintas longitudes de onda que componen la luz, el detector CCD las digitaliza y las manda a un ordenador para poder trabajar con el espectro completo de la señal.
Equipo informático	Consta de un ordenador y un software capaz de tratar

	los datos digitales que recibimos del CCD.
Monocromador	Va a ser el encargado de separar espacialmente las distintas longitudes de onda de la luz.
Fibra óptica	Es el medio de transporte de la luz monocromática desde el láser hasta la muestra y su posterior reenvío hasta el monocromador.
Conjunto de lentes y filtros	El mecanismo de lentes y filtros del sistema se encargara de enfocar y concentrar la luz en el caso de las lentes, y eliminar componentes frecuenciales no deseadas (filtros).

**Figura. 6** Componentes de un sistema Raman dispersivo

### *Espectrofotómetro de Fluorescencia*

Las medidas espectroscópicas de fluorescencia se realizaron con un espectrofluorímetro modelo Cary Eclipse (Varian Ibérica, Madrid, España) (Figura 7). Este instrumento está equipado con una lámpara de xenón continua, monocromadores de excitación y emisión, y un fotomultiplicador (posibilidad de medir espectros sincrónicos o emisión/excitación). La celda de medida con cubeta presenta una medida de 10x10 mm (3 ml) se utiliza para medidas en "ángulo recto" (RA) y la anchura de rendija de excitación y emisión es variable de 5 a 50 nm. El rango del espectro de medida oscila entre desde 300-1200 nm. El tiempo de integración y el intervalo de aumento de la longitud de onda pueden modificarse según las necesidades y el uso de algoritmos de suavizado automático del espectro están incorporados en el software interno del equipo.



**Figura 7.** Espectrofotómetro de Fluorescencia utilizado

Existen dos tipos generales de instrumentos: Fluorímetro de filtro en el que el uso de estos filtros sirve para aislar la incidencia de luz y la fluorescencia de la luz, y los Espectrofotómetros, como el usado en este trabajo, que utilizan una red de difracción de monocromador para aislar la incidencia de la luz y la fluorescencia de la luz.

El funcionamiento básico de estos equipos es el siguiente: la luz procedente de una fuente de excitación pasa a través de un monocromador y golpea la muestra. Una porción de la luz incidente es absorbida por la muestra y algunas de las moléculas en la muestra fluorescente. La luz fluorescente es emitida en todas las direcciones. Algunas de estas luces fluorescentes pasan a través de un monocromador y alcanzan un detector, el cual es usualmente colocado a noventa grados de la incidencia del haz de luz para minimizar el riesgo de la transmisión o reflejo de la incidencia de la luz buscada en el detector. La Figura 8 muestra el diseño experimental de uno de estos equipos.

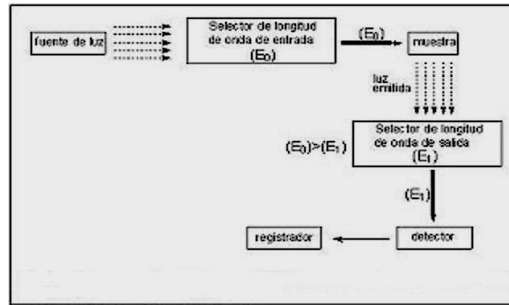


Figura 8. Diseño experimental de un equipo de fluorescencia

### *Cámara de Visión Artificial*

Los estudios de visión artificial se llevaron a cabo mediante una cámara digital modelo JAI AD-080CL dotada de dos sensores CCD adaptados a un prisma, uno color Bayer que puede generar imágenes RGB de 24 bits y uno infrarrojo que captura imágenes monocromo. Esta tecnología permite capturar dos imágenes exactas del mismo campo de visión, una de ellas en color (filtro Bayer) y la otra en infrarrojo. Este sistema de visión se basa en un prisma dicróico, que permite la separación precisa de las partes visibles (color) y del infrarrojo cercano del espectro en dos canales separados. El visible (color) canal se conoce como canal 1 y el canal infrarrojo cercano se conoce como Canal 2. Canal 1 y 2 puede ser configurado para funcionar por separado o de forma sincrónica. Al operar por separado cada canal puede ser activado y leer de forma independiente.

El prisma dicróico incorporado en la AD-080CL separa el visible (color) parte del espectro en una banda de longitud de onda de 400 nm a 650 nm (Canal 1) y la parte de zona de-IR en una banda que va desde 760 nm a 1000 nm (Canal 2).

La siguiente figura muestra el concepto de la separación en bandas visible e IR cercano.

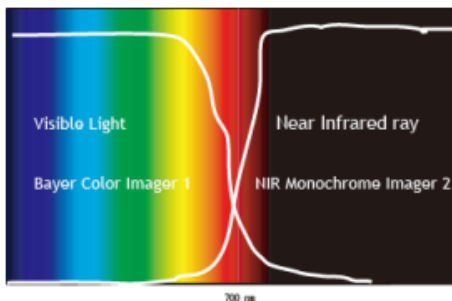
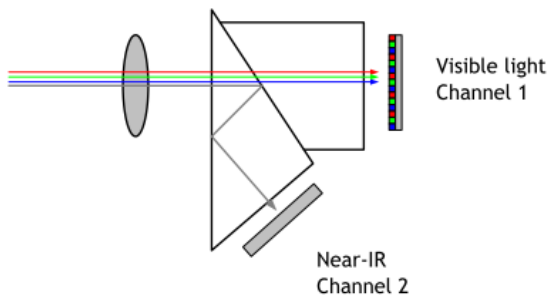


Figura 9. Diagrama conceptual para prisma óptico 2CCD

La cámara utiliza una interfaz estándar Camera Link, por lo que cada canal puede emitir imágenes con 8 o 10 bits (infrarrojo) y 24 bits (visible) de profundidad y una resolución de 1024x768 píxeles activos por canal.



Un sistema de visión consta de dos componentes principales:

Hardware: sistema de adquisición; ordenador con tarjeta digitalizadora, y sistema de visualización (cámara digital, iluminación).

Software: sistema de manipulación de imágenes, cuya función es adquirir, digitalizar, procesar y controlar la imagen.

En la figura 10 se muestra cada uno de los componentes que intervienen en un sistema de visión artificial.



**Figura 10.** Componentes de un sistema de visión artificial

En los sistemas de visión artificial, el éxito de las fases de pre-procesamiento y análisis depende altamente de la calidad de la información de las imágenes, esta calidad depende de factores como la iluminación, la óptica y los sensores utilizados para capturar las imágenes.

## Referencias

- [1] Sánchez Pineda. Calidad de los aceites (I) Alimentación, equipos y tecnología, 63-69 (2000)
- [2] Buron I., Gracia Teresa R. La calidad del aceite de oliva. Comunicaciones INIA. Madrid. (1979)
- [3] EC Commission Regulation No. 2568/91 (1991)
- [4] EC Commission Regulation No. 1989/03 (2003)
- [5] EC Commission Regulation No. 61/2011 (2011)
- [6] Uceda et al. Variability of fatty acid composition in olive (*Olea europaea* L.). Spanish Journal of Agricultural Research (2004) 2 (3), 353-359 (1994)
- [7] Chen, P. Quality evaluation technology for agricultural products. In: Proc. Int.Conf. on Agric. Machinery Engineering. Seoul. Korea (1996)
- [8] Moshenin, N.N. Physical properties of plant and animal materials. Gordon and Breach Science publishers, Inc (1970)
- [9] Ruiz-Altisent, M. Apuntes de clase de la asignatura de doctorado "Instrumentación de la calidad". Programa de Doctorado Energía, Maquinaria y Riego, del Dpto. de Ing. Rural, ETSI Agrónomos (UPM) (1998)
- [10] Valero, C. Aplicación de la espectroscopia de reflectancia difusa (ERDT) a la detección de parámetros de calidad interna de frutas. Tesis (2002)
- [11] Baeten, V. Oil and Fat Classification by FT-Raman Spectroscopy. J. Agric. Food Chem. 46, 2638-2646 (1998)
- [12] Gracia, A., Leon, L. Non-destructive assessment of olive fruit ripening by portable near infrared spectroscopy. Grasas y Aceites, 62, 3, 268-274 (2011)

- [13] Cayuela, J.A., Garcia, J.M., Caliani, N. NIR prediction of fruit moisture, free acidity and oil content in intact olives (2009)
- [14] Cayuela, J.A., Camino, M.D.P. Prediction of quality of intact olives by near infrared spectroscopy. *European Journal of Lipid Science and Technology*. 112, 11, 1209-1217 (2010)
- [15] Bendini, A., Cerretani, L., Di Virgilio, F., Belloni, P., Lercker, G., Toschi, T.G., In-process monitoring in industrial olive mill by means of FT-NIR. *European Journal of Lipid Science and Technology*. 109, 5, 498-504 (2007)
- [16] Ayora-Cañada, M.J., Muik, B., Garcia- Mesa, J.A., Ortega-Calderon, D., Molina-Diaz, A. Fourier-transform near-infrared spectroscopy as a tool for olive fruit classification and quantitative Analysis. *Spectroscopy Letters*. 2005, 38, 6, 769-785 by NIR spectroscopy. *Food Chemistry*. 122, 1261–1265 (2010)
- [17] Lopez-Sánchez, M., Ayora-Cañada, M.J., Molina-Diaz, A. Olive Fruit Growth and Ripening as Seen by Vibrational Spectroscopy. *Journal of Agricultural and Food Chemistry*, 58, 1, 82-87 (2010)
- [18] Muik, B., Lendl, B., Molina-Diaz, A., Ayora-Cañada, M.J. Fourier transform Raman Spectrometry for the quantitative analysis of oil content and humidity in olives. *Applied Spectroscopy*, 57, 2 (2003)
- [19] Korifi, R., Le Dreau, Y., Molinet, J., Artaud, J., Dupuy, N. Composition and authentication of virgin olive oil from French PDO regions by chemometric treatment of Raman spectra. *Journal of Raman Spectroscopy*, 42, 7, 1540-1547 (2011)
- [20] Dellaa, Y., Korifi, R., Le Dreau, Y., Artaud, J., Dupuy, N. Prediction of Geographical Origin of Virgin Olive Oil RDOs by Chemometric Treatment of Raman Spectra. XXII International Conference on Raman Spectroscopy, AIP Conference Proceedings, 1267, 562-563 (2010)
- [21] Zhang, X.F., Qi, X.H., Zou, M.Q., Liu, F. Rapid Authentication of Olive Oil by Raman Spectroscopy Using Principal Component. *Analysis Analytical Letters*, 44, 12, 2209-2220 (2011)

- [22] El-Abassy, R.M., Donfack, P., Materny, A. Visible Raman spectroscopy for the discrimination of olive oils from different vegetable oils and the detection of adulteration *Journal of Raman Spectroscopy*, 40, 9, 1284-1289 (2009)
- [23] Zou, M.Q., Zhang, X.F., Qi, X.H., Ma, H.L., Dong, Y. Liu, C.W., Guo, X. Wang, H. Rapid Authentication of Olive Oil Adulteration by Raman Spectrometry. *Journal of Agricultural and Food Chemistry*, 57, 14, 6001-6006 (2009)
- [24] El-Abassy, R.M., Donfack, P., Materny, A. Rapid Determination of Free Fatty Acid in Extra Virgin Olive Oil by Raman Spectroscopy and Multivariate Analysis. *Journal of the American Oil Chemists Society*, 86, 6, 507-511 (2009)
- [25] El-Abassy, R.M., Donfack, P., Materny, A. Assessment of conventional and microwave heating induced degradation of carotenoids in olive oil by VIS Raman Spectroscopy and classical methods *Food Research International*, 43, 3, 694-700 (2010)
- [26] Muik, B., Lendl, B., Molina-Diaz, A., Ayora-Canada, M.J. Direct, reagent-free determination of free fatty acid content in olive oil and olives by Fourier transform Raman spectrometry. *Analytica Chimica Acta*, 487, 2, 211-220 (2003)
- [27] Paiva-Martins, F., Rodriguez, V., Calheiros, R., Marques, M.P.M. Characterization of antioxidant olive oil biophenols by spectroscopic methods. *Journal of the Science of Food and Agriculture*, 91, 2, 309-314 (2011)
- [28] Guimet, F., Ferré, J., & Boqué, R. Rapid detection of olive pomace oil adulteration in extra virgin olive oils from the protected denomination of origin « Siurana » using excitation– emission fluorescence spectroscopy and three-way methods for analysis. *Analytica Chimica Acta*, 544, 143–152 (2005)
- [29] Poulli, K.I. Monitoring olive oil oxidation with the thermal stress through synchronous fluorescence spectroscopy and classical tests. *Food Chemistry* 117, 499–503 (2009)

- [30] Gracia, A. Determination of polyphenols in olive oil by fluorescence analysis. Symposium on Food Processing, Monitoring Technology in Bioprocesses and Food Quality Management, pp. 1364-1368 (2009)
- [31] Kyriakidis, N. B., & Skarkalis, P. Fluorescence spectra measurement of olive oil and other vegetable oils. *Journal of the American Oil Chemistry Society*, 83, 1435–1439 (2000)
- [32] Sikorska, E. Fluorescence Spectroscopy for monitoring extra virgin olive oil during storage. *Journal of Food Science and Technology*, Vol.43, No.1-2, pp. 52-61 (2008)
- [33] Wesley I. Measurement of the adulteration of olive oils by near infrared spectroscopy. *J. J. Am. Oil Chem. Soc.* 72, pp. 289-292 (1995)
- [34] Tay, A Authentication adulterated olive oil with vegetable oils infrared spectroscopy using Fourier transform. *Technol.* 35, 99-103. (1995)
- [35] Lay, Y. Quantitative analysis of potential adulterants of extra virgin olive oil by infrared spectroscopy. *Food Chem.* 53, pp.
- [36] Jiménez, A. Using an on-line NIR sensor for the characterization of olive oils. *Grasas y Aceites* nº 51(5)311–315 (2000)
- [37] Yang H. Comparison of near-infrared Fourier transform infrared and Fourier transform-Raman to determine the adulteration of olive pomace oil in extra virgin olive oil. *J. J. Am. Oil Chem. Soc.* 78(9), pp. 889-895 (2001)
- [38] León, L. Olive intact analysis using NIRS. *Grasas y Aceites* 54, 41-47. (2003)
- [39] Mailer, R.J. Quality assessment by NIR reflectance spectroscopy *JAOCs* 81, 823-827 (2004)
- [40] León, L. Effect of crop year in olives by NIR reflectance spectroscopy. *J.Agric. Food Chem.* 52, 4957-4962 (2004)
- [41] Ayora-Cañada, M.J. FT-NIR as a tool for classification and quantitative analysis of olive fruit. *Spectr Lett* 38: 769-785 (2005)

- [42] Armenta, S. Determination of parameters in edible oils by NIR *Analytica Chimica Acta*, 596(2),330-337 (2007)
- [43] Bendini, A. Preliminary evaluation of the application of FT-IR spectroscopy for monitoring the geographical origin and quality of virgin olive oil. *Journal of Food Quality*, 30(4), 424-437 (2007)
- [44] Galtier, O. Determination of the geographical origin and composition of virgin olive oils by chemometric analysis of NIR spectra. *Anal Chim Acta*. 9 595(1-2):136-44 (2007)
- [45] Sinelly, N. Characterisation and classification of Italian olive oils by near-and mid-infrared. *Journal of Near Infrared Spectroscopy*, 16(3), 335-342 (2008)
- [46] Dreau, Y. Oxidative aging study of edible oils by MCR-ALS spectroscopy chemometric methods. *Talanta*, 77(5), 1748e1756 (2009)
- [47] Casale. haracterization of table olives by NIR spectroscopy *Food Chem* 122(4):5 (2010)
- [48] Morales-Sillero, A. Non-destructive characterization of table olives by NIR spectroscopy. *Journal of Food Engineering* , Volume 107 (1) (2011)  
95-98 (1995)
- [49] Davies A. N. Study on the use of molecular spectroscopy for the authentication of virgin olive oils Part I: Fourier transform Raman spectroscopy *Appl. Spectrosc.* 54, pp. 1864-1867 (2000)
- [50] López-Díez EC. J Rapid quantitative assessment of the adulteration of virgin olive oils with hazelnut oils using Raman spectroscopy and chemometrics. *Agric Food Chem.* 51(21):6145-50 (2003)
- [51] Xiaohua, Q. Portable Raman Sensor for the rapid identification of olive oil *Technical Proceedings of the NSTI*, Vol. 2 pp. 97 (2008)
- [52] Glantz A. Olive oil fluorescence under ultraviolet light *Industrial & Engineering Chemistry Analytical Edition*, Vol.2, No.3, pp. 256-258 (1930)

- [53] Engelsen, S. A. Exploratory evaluation of frying oil deterioration fluorescence spectrometry *Journal of the American Oil Chemistry Society*, 74, 1495–1508 (1997)
- [54] Sayago A. Detection of hazelnut oil in virgin olive oil by a spectrofluorimetric method *European Food Research and Technology*, Vol.218, No.5, pp. 480-483 (2004)
- [55] Guimet, F. Cluster analysis applied to exploratory analysis of Commercial Olive Oils from Spain by fluorescence spectroscopy excitation-emission *Journal of Agricultural and Food Chemistry*, Vol.52, No.22, pp. 6673-6679. (2004)
- [56] Guimet, F. Application of Principal Component Analysis and Factor Analysis exploratory analysis parallel to olive oils by fluorescence spectroscopy matrix excitation-emission *Analytica Chimica Acta*, Vol.515, No.1, pp. 75-85 (2004)
- [57] Guimet, F. Application of Matrix fluorescence excitation-emission combined with the Fisher linear discriminant analysis for classification of olive oil *Chemometrics and Intelligent Laboratory Systems*, Vol.81, No.4, pp. 94-106, ISSN 0169-7439 (2006)
- [58] Guimet, F. Excitation fluorescence spectroscopy - emission methods in combination with analysis of three pathways as a complementary technique for the characterization of the olive oil *Journal of agricultural and Food Chemistry* 53 (24), pp. 9319-9328 (2005)
- [59] Sikorska, E. Classification of edible oils using scanning fluorescence spectroscopy. *Food Chemistry*, 89, 217–225 (2005)
- [60] Cheikhousman, R. Fluorescence spectroscopy for monitoring deterioration of extra virgin olive oil during heating *Analytical and Bioanalytical Chemistry*, 382(6), 1438–1443 (2005)
- [61] Dupuy, N. Origin of the name French virgin olive oil registered source predicted by chemometric analysis of the fluorescence spectrum of excitation-emission. *Journal of Agricultural and Food Chemistry*, Vol.53, nº.24, pp. 9361-9368 (2005)
- [62] Poulli K. Classification of edible virgin olive oils lampante by synchronous fluorescence spectroscopy and luminescence *Total. Analytica Chimica Acta*, Vol.542, No.2, pp. 151-156 (2005)

- [63] Poulli, K.I. Synchronous fluorescence spectroscopy for the quantitative determination of the adulteration of olive oil with sunflower oil. *Analytical and Bioanalytical Chemistry* 386 (5) pp. 1571-1575 (2006)
- [64] Zandomenighi, M. Fluorescence spectroscopy excitation-emission methods combined with three-way analysis as a complementary technique for the characterization of olive oil . *Journal of Agricultural and Food Chemistry* 54 (14), pp. 5214-5215 (2006)
- [65] Sayago, A. Detection of the presence of hazelnut oil in olive oil refined by Fluorescence Spectroscopy. *Journal of Agricultural and Food Chemistry*, Vol.55, No.6, pp. 2068-2071 (2007)
- [66] Dankowska, A. Application of synchronous fluorescence spectroscopy to determine adulteration of extra virgin olive oil. *European Journal of Lipid Science and Technology* 111 (12), pp. 1233-1239 (2009)
- [67] Gracia, A. Determination of polyphenols in olive oil by fluorescence analysis Symposium on Food Processing, Monitoring Technology in Bioprocesses and Food Quality Management, pp. 1364-1368 (2009)
- [68] Poulli, K.I. Monitoring olive oil oxidation with the thermal stress through synchronous fluorescence spectroscopy and classical tests *Food Chemistry* 117, 499–503 (2009)
- [69] Tena, N. Evaluation of thermal deterioration of virgin olive oil by fluorescence spectroscopy. *Journal of Agricultural and Food Chemistry* 57 (22), pp. 10505-10511 (2009)
- [70] Tena, N. Chemical changes virgin olive oil thermo-oxidized spectroscopy determined by fluorescence excitation-emission (EEFS). *Food Research International* 45 (1), pp. 103-108 (2012)
- [71] Diaz,A. Comparison of the three algorithms in the classification of table olives by means of artificial vision. *Journal of Food Engineering* 61,101–107. (2004)



[72] Diaz, A. The application of a fast artificial vision algorithm for classification of olives. *Food Research International* 33, 30-309 (2000)

[73] Riquelme M.T. Olive classification according to external damage using image analysis. *Journal of Food Engineering* Volume 87, Issue 3, August 2008, Pages 371–379 (2007)

[74] Furferi R. Artificial Neural Network for real time estimation of qualitative parameters of olive oil during continuous extraction. *Computers and Electronics in Agriculture* Vol. 55, Issue 2, pp. 115-131 (2007)

[75] Furferi, R. Automatic prediction olive maturity index based on ANN. *Journal of Food Engineering* 101,318–328 (2010)

# Capítulo II. *Quimiometria*



## II.1 Generalidades

El término Quimiometría fue introducido en 1972 por el sueco Svante Wold y Bruce R. Kowalski [1]. La quimiometría trata, específicamente, de todas aquellos procesos que transforman señales analíticas y datos más o menos complejos en información. La quimiometría utiliza métodos de origen matemático, estadístico y otros procedentes del campo de la lógica formal para conseguir sus fines. Por todo ello, la quimiometría se sitúa en un campo interdisciplinar. D.L Massart definió la quimiometria como la disciplina química que utiliza las matemáticas y estadísticas para diseñar o seleccionar procedimientos experimentales óptimos, para proporcionar la máxima información química relevante mediante el análisis de los datos químicos, y para obtener conocimiento acerca de los sistemas químicos [2]

Este tipo de información se obtiene aplicando una serie de técnicas quimiométricas que se han denominado genéricamente como pattern recognition. En español se ha traducido por técnicas de reconocimiento de patrones, de formas, de pautas, etc. Entre sus objetivos está la identificación de relaciones entre las muestras previamente analizadas a través de diversos métodos de análisis instrumental, que permitan su agrupamiento o clasificación según la similitud de sus espectros. De igual manera, debe permitir la clasificación de nuevas muestras como pertenecientes a un grupo determinado de una forma rápida y objetiva [3].

El desarrollo de la quimiometría fue tardío como resultado de la lenta evolución de la instrumentación científica. La progresiva automatización de los laboratorios y el de la instrumentación química, ha llevado consigo el incremento en el número de variables que son obtenidas con estos instrumentos y por tanto la necesidad de utilizar nuevos métodos avanzados de diseño de experimentos, calibración y análisis de esta gran cantidad de datos resultantes.

Con este fin se han diseñado varios paquetes estadísticos que facilitan el tratamiento de los datos químicos-analíticos como son Unscrambler, MatLab [4,5] u otros programas. Mediante éstos software se logra una mayor rapidez y comodidad a la hora de trabajar con los datos y obtener unos resultados.

Según el objetivo a alcanzar, se utilizan distintas técnicas estadísticas. En esta tesis la Quimiometría se ha utilizado para desarrollar modelos de calibración empíricos y su aplicación a muestras obtenidas en diferentes etapas del proceso de producción del aceite de oliva.

## **II.2 Tratamiento de datos**

### ***II.2.1 Herramientas Quimiométricas***

Todos los trabajos realizados en esta memoria tanto por métodos espectroscópicos como mediante visión artificial giran en torno a la misma unidad temática: aplicación de “propiedades ópticas” para analizar parámetros relacionados con la calidad de los frutos de aceituna y aceite de oliva. Estos estudios se centran en el estudio de las propiedades ópticas de la fruta y aunque aparentemente muestran una estructura diferente, presentan un esquema de trabajo común. Los datos originales (espectros o imágenes), proporcionados por los diferentes equipos detectores, son sometidos a un pre-tratamiento con el fin de acondicionarlos para un tratamiento matemático posterior; el modelo utilizado (regresión lineal, método discriminante, etc.) proporciona parámetros de caracterización de los productos que hace posible el desarrollo de modelos de calibrado o clasificación. Para cada estudio se han utilizado softwares específicos para realizar

todos los cálculos necesarios para conseguir los objetivos marcados en cada caso (desarrollo de modelos de calibración y predicción de parámetros de calidad, clasificación por categorías, según defectos, índice de madurez, etc.)

Todo el procesado de datos se realizó combinando algoritmos desarrollados con el programa matemático MATLAB (MathWorks, Inc.), que facilita el trabajo con matrices de datos y con cálculos de análisis estadístico, y el software Image Pro-Plus versión 6.0 MediaCybernetics, Inc) con el que se ha desarrollado algunos de los algoritmos de análisis de imagen.

### **II.2.2 Pre-procesamiento de datos**

El pre-procesado de los datos es un paso fundamental antes de aplicar cualquier técnica sobre los datos. El empleo de métodos de pre-procesado está ampliamente extendido en aplicaciones espectroscópicas como NIRS o Raman, para la corrección de las variaciones aleatorias y sistemáticas del espectro debidas a efectos físicos (dispersión de la luz, ruido, derivas, desplazamientos, etc.) y otras variaciones espectrales no relacionadas con las propiedades estudiadas. En muestras líquidas, como en nuestro caso el aceite de oliva, la temperatura o la turbidez son relevantes y deben ser consideradas antes de la etapa de procesado de los datos.

Los datos espectrales presentan diferentes ratios señal/ruido dependiendo del modo de medida, de la muestra a analizar y de las condiciones de medida. Las técnicas de corrección de señal eliminan el ruido presente, minimizando las contribuciones de variables físicas que incorporan información irrelevante en los espectros, logrando una mejor interpretación de los datos, construyendo de esta forma modelos más simples y robustos.

Algunos de los pre tratamientos más comunes y utilizados a lo largo de la presente tesis doctoral se describen brevemente a continuación.

### ***Centrado y Auto escalado***

Una técnica muy frecuente para trabajar con datos espectroscópicos es el centrado. En él, cada columna de datos es sustraída de todos los valores de la columna para obtener una matriz de datos donde la media de cada variable procesada es cero. Una vez centrados los datos la matriz ( $XTX$ ) resultante es la matriz de varianzas-covarianzas. La propiedad fundamental de los datos es que el valor medio de cada una de las variables es igual a cero. Este pre tratamiento no modifica la varianza de los datos [6].

El centrado no en todos los casos es conveniente, existen una serie de limitaciones en las cuales no debe aplicarse el centrado de los datos [7].

El Auto escalado (también denominado estandarización por columnas) se lleva a cabo después del centrado y divide los valores de cada columna entre la desviación típica de cada columna. El producto resultante es una matriz donde todas las columnas tienen de media cero y varianza unitaria lo que implica que toda la información restante está relacionada con la correlación entre variables. El escalado enfatiza las características con menor variabilidad por lo que resulta útil en el caso de determinaciones de componentes químicos minoritarios que puedan verse enmascarados. El inconveniente de ésta técnica se presenta en los espectros con un alto nivel de ruido ya que éste también se ve incrementado pudiendo llegar a enmascarar la información de interés [8].

### ***Corrección de la línea base (Baseline) y Normalización***

Es uno de los métodos habitualmente empleados para eliminar el efecto de nivel de la línea de base. Para ello se emplean uno a varios puntos del rango espectral para los que no existan variaciones significativas y se lleva a cabo la sustracción de ese valor a cada punto de la señal. La corrección simplemente mueve el espectro hacia arriba o abajo mediante la eliminación de la línea base. Las variaciones en la línea base de los espectros pueden producirse de forma no lineal provocando los llamados efectos de deriva.

Los métodos de normalización se emplean para la corrección de diferencias en la intensidad global debidas a factores sin interés en la estimación de la composición química como la dispersión de la luz, las variaciones en la penetración de la radiación o el tamaño de las muestras. El objetivo es asignar el mismo peso a cada señal mediante la corrección de esas diferencias de intensidad globales. La normalización básicamente lo que hace es convertir cada una de las filas de la matriz de datos en un vector normalizado de longitud unitaria dividiendo cada valor por la raíz cuadrada de la suma de los valores al cuadrado. La normalización por sí sola no es capaz de corregir las variaciones en la línea base por lo que debe ser aplicada conjuntamente con otro método para su corrección.

### ***Suavizado (Smoothing) y Derivadas primera y segunda***

El suavizado se basa en la reducción del ruido aleatorio existente en la señal instrumental. La técnica más utilizada es la propuesta por Savitzky-Golay [9]. Este método se basa en el principio de que, para un pequeño intervalo de longitudes de onda, se puede ajustar un polinomio del grado adecuado. Los nuevos valores tras el ajuste son una mejor estima que los valores medidos ya que se ha eliminado parte del ruido que los afectaba.

La diferenciación o cálculo de derivadas permite acentuar las diferencias existentes en los datos espectrales. Tanto la primera como la segunda derivada se utilizan a menudo para el tratamiento de los datos. La segunda derivada elimina el ruido de fondo lineal y constante. Los dos principales algoritmos de diferenciación son el de Savitzky-Golay y el de Norris. El primero, permite calcular derivadas de primer orden o mayor incluyendo un factor de suavizado que determina el número de variables adyacentes que se usarán en la estimación de la aproximación polinómica utilizada en la derivación. El algoritmo de Norris, a diferencia del anterior, solo permite el cálculo de derivadas de primer orden.

Una desventaja del uso de las derivadas es que disminuyen el valor de la relación señal-ruido, por esta razón, se recomienda realizar un suavizado de la señal antes de la diferenciación de los datos. Otra desventaja es que en ocasiones los modelos de calibración obtenidos mediante datos espectrales tratados con primera o segunda derivada, son menos robustos frente a cambios instrumentales, como por ejemplo derivas de la longitud de onda, que ocurren a lo largo del tiempo por lo que habría que revisar las calibraciones [10].

### ***Varianza Normal Estándar (SNV)***

Se trata de una transformación introducida por Banes y cols. [11] para la corrección de los efectos de la dispersión de la luz y los debidos al tamaño de partículas, aunque también permite reducir las diferencias en las intensidades globales de las señales.

En la transformación SNV cada espectro es centrado y escalado dividiendo por su desviación típica según se refleja en la Ecuación 2.



$$X_{ij}^{\text{SNV}} = \frac{X_{ij} - \bar{x}}{S_i}$$

**Ecuación 2.** Transformación de la Varianza Normal Estándar

Donde  $X_{ij}^{\text{SNV}}$  es el elemento del espectro transformado,  $X_{ij}$  corresponde al espectro original  $i$  a la longitud de onda  $m$ ,  $\bar{x}$  es la media del espectro  $i$  y  $S_i$  la desviación estándar de los valores de absorbancia de cada muestra. Cada espectro tratado de esta manera tiene media 0 y varianza igual a 1 y es, por tanto, independiente de los valores originales.

Al aplicar la transformación SNV se asume que los efectos multiplicativos de la señal son uniformes en todo el rango espectral por lo que en caso de no ser así se pueden introducir errores.

### ***Corrección de Dispersión Multiplicativa (MSC)***

La transformación de la dispersión multiplicativa es un tratamiento propuesto para datos espectrales que se centra en la eliminación de los efectos debidos a la dispersión de la radiación [12]. En él que se estiman los efectos de la dispersión de la radiación o los cambios en el paso óptico en relación a una muestra ideal que se introduce como referencia ( $m$ ). En principio, esta estimación se debería hacer en zonas del espectro que no contengan información química, sin embargo, en la práctica se corrige el espectro completo.

MSC se basa en dos premisas. La primera premisa es que un espectro correspondiente a una muestra es considerado como la suma de otros dos espectros, uno debido a la dispersión de la luz,  $d$ , y otro debido a la composición química de la muestra,  $c$ , como podemos ver en la Ecuación 3.

Teniendo en cuenta esto MSC trata de corregir el efecto del espectro de dispersión,  $d$ .

$$X_i = d_i + c_i$$

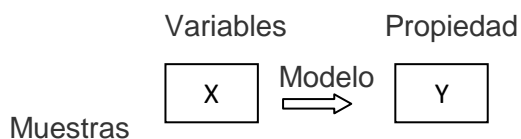
**Ecuación 3.** Descomposición de un espectro en su espectro debido a dispersión y su espectro debido a la descomposición química de la muestra

La segunda premisa asumida por MSC es que los coeficientes del espectro de dispersión,  $d$ , son los mismos para todas las muestras en todas las longitudes de onda y por lo tanto pueden ser estimados a partir de un rango de longitudes de onda empleando un espectro de referencia (normalmente el espectro medio del lote de calibración).

### II.2.3 Tratamiento quimiométrico de datos espectroscópicos

Para el cumplimiento de sus objetivos, el Reconocimiento de Patrones (pattern recognition) en Quimiometría utiliza el Análisis Multivariante, que es un conjunto de métodos que sirven para analizar el conjunto de datos multivariantes en el sentido de que hay varias variables medidas para cada individuo u objeto estudiado. Su razón de ser radica en un mejor entendimiento del fenómeno objeto de estudio obteniendo información que los métodos estadísticos univariantes y bivariantes son incapaces de conseguir.

El objetivo que se persigue es la relación entre las matrices  $X$  e  $Y$  representadas en la Figura 1.



**Figura 1.** Representación de la relación entre una matriz de datos X formada por variables descriptoras de cada uno de los objetos y una matriz Y que contiene otro tipo de variables

La relación entre matrices se establece mediante una ecuación o modelo matemático. La ecuación (1) representa un modelo en el que las variables y se relacionan con las variables X mediante la matriz de coeficientes B. Dado que cada valor medido de las variables de la matriz Y (variable respuesta) no resulta exactamente de la combinación lineal de las variables de la matriz X, existe una matriz que contiene los errores, E.

$$Y = X B + E \quad (1)$$

Podemos estar interesados en dos tipos de informaciones que proporciona la parte determinista de los modelos. Por un lado en el valor de los coeficientes, B. Estos coeficientes expresan la importancia de la relación entre las variables y, en muchos casos, explican el comportamiento de la variación existente en los datos.

Frecuentemente, sin embargo, estamos interesados en el carácter predictivo de los modelos. Una vez éstos han sido validados, nos interesa conocer el valor de la variable y, una vez que se han medido los valores de las variables x para el objeto, analizando las variables respuesta de la matriz Y son técnicas destructivas, difíciles, caras o tediosas de medir, así como perjudiciales para el medio ambiente (contenido en proteína, humedad, contenido graso, etc.), mientras que las variables de la matriz X se miden fácilmente (como por ejemplo un espectro Raman o de fluorescencia), de esta forma obtendremos una gran ventaja al poder predecir las primeras a partir de las segundas. El conjunto de técnicas que se utilizan para esta finalidad recibe el nombre de regresión multivariante.

Dicho conjunto de técnicas multivariantes pueden emplear diferentes métodos de trabajo los cuales podemos dividir en diferentes tipos: técnicas

exploratorias, predictivas y de clasificación, que se dividen a su vez en dos grupos, no supervisadas y supervisadas en función de si solo se usa la matriz espectral X o se dispone también de y (figura 2)

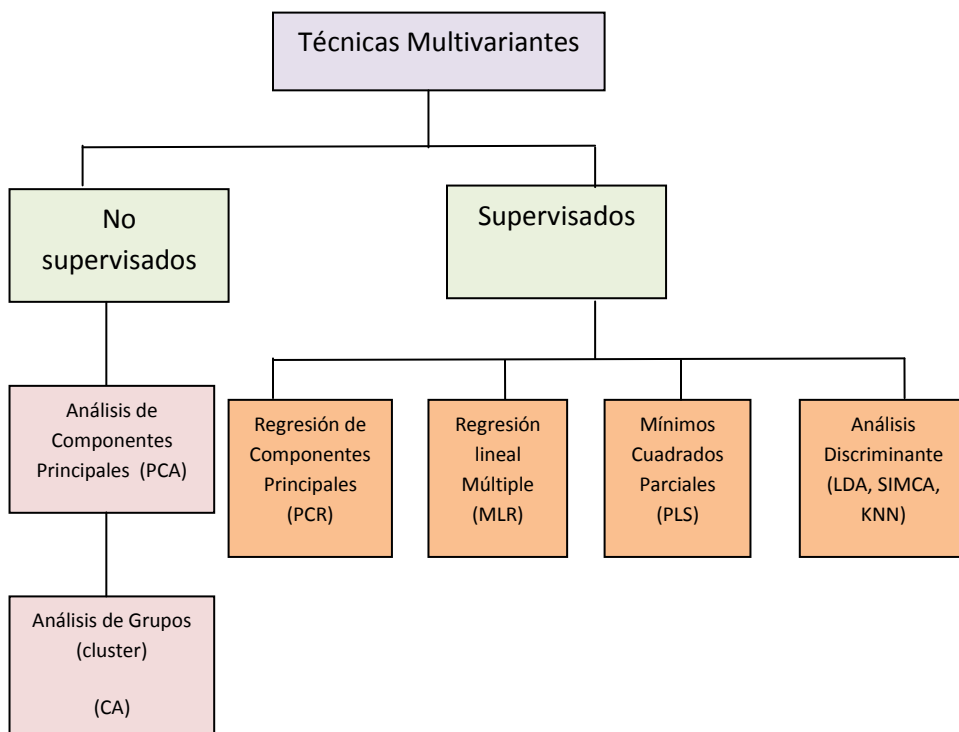


Figure 2. Clasificación de las técnicas multivariantes

### Análisis Exploratorio de Datos (Técnicas No supervisadas)

El análisis exploratorio de datos se utiliza para poner de manifiesto y resaltar la información contenida en una matriz de datos multidimensional.

Existen herramientas simples como son la observación directa de los datos o la representación de histogramas o gráficos de dispersión pero no son

suficientes para conocer por completo la estructura de los datos. Una de las técnicas de exploración más utilizadas y robusta es el análisis de componentes principales con el que se logra definir estructuras con espacios de más de 3 dimensiones, para identificar tendencias presentes en los datos y estudiar las causas o fuentes que provocan las variaciones.

### ***Análisis de Componentes Principales (PCA)***

Técnica que describe la variación en los datos multivariados a través de un pequeño número de variables no correlacionadas (es decir, los ejes principales o componentes principales, PC).

La característica principal del PCA es la combinación de las variables originales en una serie de nuevas variables en las que la máxima cantidad de información (variación espectral) presentada en una base de datos original se mantiene en las nuevas variables. PCA es una técnica que reduce la dimensionalidad preservando al mismo tiempo la máxima varianza.

Con esta técnica estadística y de proyección, se persigue maximizar la información de la varianza presente en un conjunto de datos y representarla en el menor número de dimensiones posible, utiliza como base la rotación propia de los ejes de los datos para conformar nuevos ejes que contienen la mayor información de la varianza posible, éstos nos permite buscar de forma secuencial las direcciones del espacio que ofrecen la mejor visión posible de la misma. En resumen, el objetivo principal es identificar tendencias ocultas presentes en los datos y estudiar las causas o fuentes de varianza a que obedecen, reduciendo las dimensiones del espacio de origen de modo que la información relevante contenida en la matriz multidimensional pueda quedar

reflejada, del mejor modo posible, sobre dos o tres dimensiones oblicuas obtenidas como combinaciones lineales de las variables originales [13].

El cálculo de las componentes principales comienza con la construcción de la matriz de covarianza. El estudio de esta matriz nos permite ver cómo están correlacionadas las variables entre sí. A partir de la matriz de covarianza se calculan los componentes principales que vienen dados según la siguiente expresión:

$$CP1 = a_{11}x_1 + a_{12}x_2 + \dots$$

$$CP2 = a_{21}x_1 + a_{22}x_2 + \dots$$

Donde  $x$  son las variables originales y los coeficientes  $a$  son los valores de los loadings, o contribución de cada variable a cada componente. Cuando en la ecuación de cada componente se sustituye el valor que cada variable toma para cada muestra, se obtienen las coordenadas de cada muestra en el espacio definido por las componentes principales, es decir, los scores.



Loadings: P

$$X = TPT + E - t_1p_1T + t_2p_2T + \dots + E$$

T: Scores o combinaciones lineales de las variables X originales.

P: Loadings o pesos de cada variable X en el nuevo espacio

Hay dos tipos de figuras para representar los resultados de los análisis de componentes principales. En el primer tipo se representan los scores, es

decir, los valores que toman las muestras en el nuevo espacio calculado siendo posible calcular relaciones entre ellas. En el segundo tipo se representan los loadings, indicando la importancia que tiene una cierta variable en cada componente. Estas representaciones permiten una interpretación eficaz del espacio de datos entero y su estudio es fundamental para la exploración de datos. Se les llama mapa de casos, cuando se representan los scores, y mapa de variables, cuando se representan los loadings.

A continuación se muestran estas representaciones para un conjunto de datos:

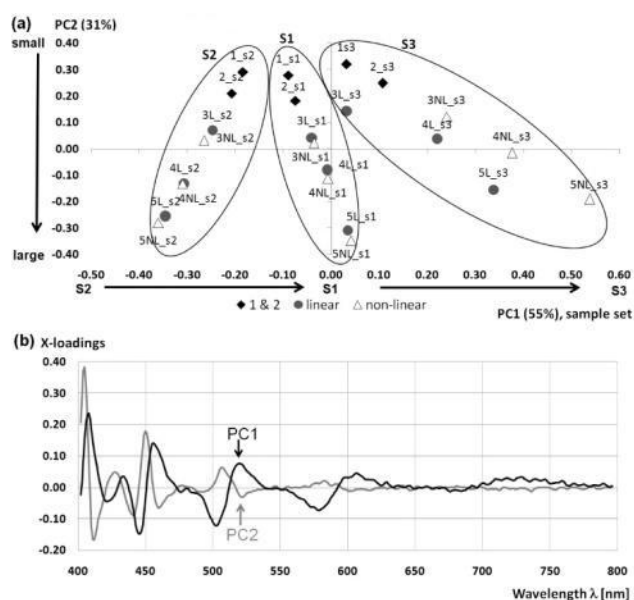


Figura 3 Gráfica de scores (a) y loadings (b) de un conjunto de datos espectroscópicos

El análisis de componentes principales da una visión general de los datos:

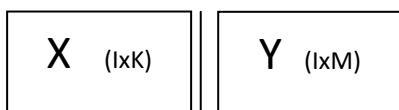
- Dispersión

- Tendencia de los datos
- Agrupaciones
- Datos aberrantes

## Técnicas Predictivas (Supervisadas)

### *Regresión por Mínimos Cuadrados Parciales (PLS)*

Es una técnica desarrollada en los años 60 por Svante Wold, que generaliza y combina características del análisis de componentes principales y de la regresión múltiple [14] y tiene un gran número de aplicaciones especialmente en espectroscopia. Es útil para predecir un conjunto de variables dependientes usando un extenso conjunto de variables independientes (predictores,  $y$ ). PLS produce la matriz  $W$  que explica la covarianza entre  $X$  e  $Y$ . La regresión consta de dos pasos fundamentales. Primero transforma a la matriz de predictores  $X$  de orden  $n \times p$ , con ayuda del vector de respuestas  $Y$  de orden  $n \times 1$ , en una matriz de componentes o variables no correlacionados entre sí pero que explican la mayor varianza de la matriz  $X$  y que están correlacionados con  $y$ ,  $T = (T_1, \dots, T_P)$  de orden  $n \times p$ , llamados componentes PLS. Segundo, calcula el modelo de regresión estimado usando el vector de respuestas original y como predictores, los componentes PLS. PLS determina los scores (variables latentes) teniendo en cuenta  $Y$  así como las variables predictoras ( $X_1, x_2, \dots, x_p$ ).  $T$ : scores



\_\_\_\_\_      \_\_\_\_\_  
 Loadings:  $p'(1 \times K)$     Loadings:  $q'(1 \times M)$



Los dos modelos desarrollados siguen estas ecuaciones:

$$\begin{array}{lcl}
 X = TP' + E & \longrightarrow & Y = XB + E \\
 Y = TQ' + F & & B = X + Y = (P') + Q = WQ'
 \end{array}$$

Donde:

B= matriz de regresión

W= matriz de pesos que refleja la covarianza entre X e y

P, Q= loadings de X e y respectivamente

T= scores PLS

E, F= matrices de residuales

El primer paso antes de aplicar el método de selección del mejor modelo (valores más elevados de correlación número de componentes óptimo), consiste en la construcción de dos conjuntos de muestras: el de calibración, que permitirá realizar los modelos PLS, y el conjunto de validación, que será utilizado, como su nombre indica, para validar o evaluar la capacidad de predicción del modelo elegido como óptimo, cuando es aplicado a muestras no utilizadas en el proceso de calibración.

El valor predicho (aproximación de la variable respuesta)  $\hat{Y}$ , B son los mínimos cuadrados estimados.

$$e_i = Y_i - \hat{Y}_i \text{ determinación del error o residuales}$$

El criterio de mínimos cuadrados es el que estima B y E de tal manera que los datos se ajustan a una línea recta que reduce al mínimo los errores de predicción

La línea de regresión se ajusta para minimizar la suma de los residuales cuadrados,  $\sum e_i^2$ .

$\sum e_i^2$  es la suma de los residuales cuadrados, que mide la varianza residual o inexplicada cuando los datos se ajustan a una línea recta.

### **Selección de muestras de calibración y validación**

A la hora de construir un modelo de calibración, los conjuntos de calibración y validación deben contener muestras que puedan aportar una variabilidad suficiente para todos los componentes presentes en las muestras que serán analizadas mediante el modelo, especialmente en cuanto a la variable de interés. El conjunto de calibración deberá comprender muestras uniformemente distribuidos en un rango mayor que el de las muestras que serán analizadas por el modelo. Para esto se necesitan buenos datos experimentales así como modelos estadísticos adecuados. Existen diferentes estrategias para la selección de conjuntos de datos adecuados para la calibración. Normalmente cuanto mayor es el número de muestras utilizado en la calibración, mejor es la capacidad predictiva del modelo y menor el error de predicción.

La situación real más simple que se puede encontrar es cuando es posible generar artificialmente las muestras en el laboratorio (forzando el deterioro de las muestras de aceituna obteniendo aceites con unos índices de calidad fuera de los valores normales encontrados en la planta de producción). Es muy recomendable incluir también muestras reales extraídas del proceso de producción ya que nos aseguramos de incluir en el proceso de estudio una gran variabilidad en las muestras y todas fuentes de variación inesperadas [15].

Existen diferentes métodos de selección de los conjuntos de calibración y validación: la división aleatoria, se da en el caso que no exista una base lógica para la división de datos, de tal forma que se asignan observaciones aleatoriamente para obtener el conjunto de datos de estimación y de predicción.

En este trabajo se ha utilizado un método de selección Dúplex [16]. El algoritmo Dúplex introducido por Snee en 1977, es un método de selección más idóneo ya que soluciona el problema encontrado con los otros métodos, generando subconjuntos de calibración y validación homogéneos e independientes entre sí.

El método se basa en encontrar los dos puntos que están más alejados (empleando la distancia Euclídea o la de Mahalanobis) y los selecciona para el conjunto de calibración. De los puntos restantes, los dos más objetos que se encuentren más alejados entre sí se seleccionan para el conjunto de validación. En un siguiente paso, se selecciona el punto que se encuentra más alejado de los puntos elegidos para el conjunto de calibración y se incluye en éste. Se repite este procedimiento seleccionando esta vez el punto que esté más alejado de los puntos del conjunto de validación. Así se van asignando los puntos alternativamente a los conjuntos de calibración y validación mediante el mismo procedimiento, obteniéndose conjuntos representativos y de igual tamaño.

Una vez revisada la distribución de los conjuntos de calibración y validación es también recomendable llevar a cabo un filtrado previo para la eliminación de potenciales muestras anómalas.

### **Métodos de detección de datos aberrantes (outliers)**

La presencia de observaciones atípicas en las muestras utilizadas para inferir parámetros de un modelo, producen malas influencias en los resultados, proporcionando modelos erróneos. Así, un primer problema que se puede plantear al decidir inferir un modelo a partir de una muestra, es que ésta puede contener observaciones no deseables, observaciones atípicas, que enturbian y proporcionan información errónea.

Como definición de observación atípica consideraremos la dada por Barnett y Lewis (1994) [17], donde éstas son observaciones que parecen ser discordantes o inconsistentes con respecto al resto de observaciones de la muestra en la que se encuentra.

Para proteger los resultados de las posibles observaciones atípicas, se definen unos procedimientos denominados métodos de detección de "outliers". Estos métodos se basan en ciertas medidas que nos permiten averiguar si, en una muestra dada para la estimación de unos determinados parámetros de un modelo, una observación es atípica; siendo candidata a ser estudiada a fondo para su posible eliminación de la muestra.

En la literatura existen una gran cantidad de métodos para detectar valores extremos en el conjunto de datos de predicción, autores como Fernández Pierna, J.A. (2002,2003) [18,19] hacen un compendio de las principales técnicas utilizadas para la detección de outliers, entre las que se destacan: método de la incertidumbre, método "convex full", distancia de Mahalanobis, X Residual, Potencial Functions, RHM (Resampling by the halfmeans method) y SVM (Smallest half-volume method).

En resumen podemos clasificar las técnicas actuales de detección de outliers en:

- Técnicas basadas en modelos. Se basan en el campo de la estadísticas, dada la premisa de conocer la distribución de los datos. Entre estas técnicas se resalta: Método de Incertidumbre y Método de "convex hull" .
- Técnicas basadas en proximidad. Esta técnica se fundamentan en el manejo de distancias entre objetos, a mayor distancia del objeto respecto a los demás datos, éste se considerado como un Outlier. Entre los principales métodos se encuentra: la distancia de Mahalanobis y X-residual.
- Técnicas basadas en densidad. Se hace uso de la estimación de densidad de los objetos, para ello, los objetos localizados en regiones de baja densidad, y que son relativamente distantes de sus vecinos se consideran anómalos. Entre los principales métodos se encuentra: el método de funciones potenciales (potencial function), método RHM (Resampling by the halfmeans method) o el método SHV (Smallest half-volume).

Entre las técnicas de detección de outliers basadas en la proximidad, se encuentra la distancia de Mahalanobis [20]. Esta distancia se define como la distancia de cada nueva observación es derivada de la distancia desde el centroide de la matriz de datos  $\bar{X}_c$  a la matriz de covarianza  $[S = \bar{X}_c, \bar{X}_c / I - 1]$  donde I es el numero de objetos del grupo de calibración.

La ecuación 1 muestra la distancia de Mahalanobis entre dos puntos definidos por los vectores x e y que está afectada por la matriz de covarianza S.

$$D(x,y)= \sqrt{(x-y)^T S^{-1} (x-y)}$$

**Ecuación 1.** Expresión general de la distancia de Mahalanobis

Los valores se comparan con un valor crítico de una distribución  $\chi^2$  (A grados de libertad, donde A es el complejidad del modelo de PLS y un  $\alpha= 0,95$ ) [21].

Otro método es el X-residual [22], en él se calculan el total de desviación estándar residual o diferencia entre la dispersión en el conjunto de calibración (se) y la desviación estándar residual (sei) del objeto de predicción. Entonces, estos valores se comparan, si sei es mucho mayor que si (tres veces más grande tal como se recomienda en la referencia [23]), esto es una indicación de un objeto anormal.

### **Selección del modelo de calibración**

Los estadísticos utilizados para la selección de las mejores ecuaciones de calibración fueron: el error típico residual para el colectivo de calibración (RMSEC), para el de validación cruzada (RMSECV) y el error de predicción (RMSEP). El coeficiente de determinación para el proceso de calibración ( $R^2$  o validación cruzada ( $r^2$ ), así como la desviación residual predictiva (RPD) que se define como el cociente entre la desviación típica y el error típico residual en la calibración cruzada (DT/RMSECV). Las mejores ecuaciones se seleccionaron atendiendo a los valores más bajos de error de calibración (RMSEC), error estándar de validación cruzada (RMSECV) y error estándar de validación (RMSEP). También se tuvieron en cuenta los valores más elevados de los coeficientes de determinación en calibración ( $R^2$ ) y en validación ( $r^2$ ) y los valores del índice RPD.

Para evaluar la precisión y exactitud de las predicciones de las ecuaciones se utiliza el Índice RPD. Por lo general, un buen modelo debe tener alta  $r^2$  cal,  $r^2$

valores y el valor de RPD, con baja RMSEC y los valores RMSEV (Wu et al., 2010) [24]. La relación de la desviación estándar de la variable de respuesta y el RMSEV (RPD) se calculó para evaluar el modelo de tasa de éxito y precisión. Los valores superiores a 3, valor mínimo recomendado por Williams y Sobering (1996) [25] confirman una elevada precisión de las calibraciones desarrolladas. Si el valor se encuentra entre 2,5 y 3 la calibración es buena y si se encuentra entre 2,5 y 2 la calibración es aceptable para análisis cuantitativos y cualitativos respectivamente.

### **Validación del modelo**

Para llevar a cabo una buena interpretación de un modelo de regresión debemos acompañar siempre nuestro estudio de la diagnosis y validación del modelo, con el fin de averiguar, por un lado, hasta qué punto los resultados obtenidos son extrapolables a la población de la que vienen los objetos seleccionados y, por el otro, por qué han aparecido dichos grupos. Esta validación se puede realizar de forma externa o interna.

#### *- Validación interna*

Se puede establecer utilizando procedimientos de validación cruzada. La validación cruzada o cross-validation es una técnica utilizada para evaluar los resultados de un análisis estadístico y garantizar que son independientes de la partición entre datos de calibración y validación. Existen diferentes tipos de validación cruzada: validación cruzada de K iteraciones, aleatoria o validación cruzada dejando uno fuera. En la presente tesis para el desarrollo de los modelos de calibración se ha utilizado el método de validación cruzada dejando uno fuera o leave-one-out cross-validation (LOOCV), que implica

separar los datos de forma que para cada iteración tengamos una sola muestra para los datos de prueba y todo el resto conformando los datos de entrenamiento. La evaluación viene dada por el error, de forma que para cada una de las N iteraciones se realiza un cálculo de error. El resultado final lo obtenemos realizando la media aritmética de los N valores de errores obtenidos, según la fórmula:

$$E = 1/N \sum_{i=1}^N E_i$$

Donde se realiza el sumatorio de los N valores de error y se divide entre el valor de N.

Para ello se dividen los datos en dos grupos y se aplica el algoritmo de clasificación a cada grupo comparando los resultados obtenidos en cada grupo. Con los valores predichos y los valores nominales se calcula el error a través de un estadístico llamado PRESS (la suma de cuadrados de los errores de predicción /prediction error sum of squares)

$$\text{PRESS: } \sum (e_a - e_{\text{pred}})^2$$

donde  $e_a$  es el valor nominal del componente en las muestras utilizadas para la calibración, y  $e_{\text{pred}}$  es el valor predicho por el modelo

Cuando se utiliza la validación cruzada para validar un modelo se puede definir la raíz cuadrada de la media del error en validación cruzada (RMSECV) como:

$$\text{RMSECV} = \sqrt{\text{PRESS}/n}$$

PRESS: corresponde a la suma de cuadrados de los errores en los valores predichos.



- *Validación externa*

Se puede realizar comparando los resultados obtenidos con un criterio externo (por ejemplo, predicciones obtenidas para muestras independientes no utilizadas en el proceso de calibración)

En la presente tesis la validación se realizó en todos los casos mediante un conjunto de muestras externo independiente a la calibración.

El modelo de regresión se calcula a partir de los datos del conjunto de calibración y, el modelo calculado, se valida con los datos del conjunto de validación que no han sido utilizados para calcular la ecuación de calibración del modelo. Se puede definir la raíz cuadrada de la media del error de predicción al cuadrado (RMSEP) para el conjunto de validación.

$$\text{RMSEP} = \sqrt{\sum(\hat{y}_i - y_i)^2 / n}$$

El método PLS puede corregir en determinados casos la falta de linealidad, aumentando el número de factores, pero esta opción debe desestimarse ya que se puede sobredimensionar el modelo (over-fitting). [26]

### **Técnicas de de Clasificación**

En un análisis clasificatorio se construyen modelos capaces de pronosticar la pertenencia de un objeto a una categoría sobre la base de las características del objeto.

*Métodos no supervisados*, se basan en descubrir agrupaciones naturales de variables en un espacio de N-dimensiones sin conocer a priori a qué clase pertenece cada muestra. Asimismo, permite detectar la presencia de outliers o muestras anómalas antes de incluirlas en el modelo predictivo. Entre los

métodos más comunes se incluye las Redes Neuronales Artificiales o el análisis de grupos (CA, cluster analysis) [27,28]

*Métodos supervisados*, en ellos la clasificación se basa en el conocimiento previo del sistema, con conjuntos de calibración o entrenamiento de clase conocida, que permiten definir cada clase. La calidad de los resultados viene influenciada por la calidad de los grupos de calibración. Es usual utilizar muestras adicionales pertenecientes a clases conocidas que no forman parte del conjunto de calibración original, para comprobar el modelo y controlar su precisión y robustez. Un ejemplo de estos métodos son el análisis discriminante lineal, método k-nn o el método SIMCA.

A continuación se comentan las distintas técnicas de clasificación supervisadas utilizadas en esta tesis:

#### *Métodos de Clasificación Supervisados*

Estos métodos de clasificación por clases por su parte, se basan en la creación de dominios en el espacio, cada uno de los cuales presentan límites para cada una de las clases, y prueba si los nuevos objetos caen dentro de ese dominio. En ellos se desarrolla un modelo matemático para cada clase dentro del conjunto de calibración. Por ello, una muestra puede clasificarse como perteneciente a una clase o a ninguna de ellas. Entre estos métodos se encuentran el análisis discriminante lineal, método K-nn y el modelado suave independiente de analogía de clases (SIMCA, Soft Independent Modelling of Class Analogy).

#### ***Análisis Discriminante Lineal (LDA)***

El análisis discriminante utiliza un algoritmo que busca combinaciones lineales de las variables originales que maximizan la varianza entre clases, a la vez que minimizan la varianza inter-clases [29]

Para construir un modelo es necesario asignar los objetos del conjunto de calibración a una clase dada, añadimos una variable categórica a la matriz de datos conteniendo tantas clases como sean necesarias.

Las funciones discriminantes se construyen de una en una, buscando las direcciones del espacio que hacen máxima la expresión:

$$\lambda = \text{SCD}/\text{SC1}$$

SCD: suma de los cuadrados de las distancias Euclídeas entre los objetos que pertenecen a distintas clases, en la dirección que indica la función discriminante

SC1: suma de los cuadrados de las distancias Euclídeas entre los objetos que pertenecen a la misma clase, también en la dirección de la función discriminante

Así es posible establecer fronteras explícitas que separan las regiones del espacio que pertenecen a cada categoría. La clasificación de los objetos se hace en función de sus probabilidades de pertenencia a una u otra clase. La región crítica que separa dos clases es una recta, o un plano o hiperplano, donde se igualan las probabilidades de pertenencia a una u otra clase, esta línea recta o plano es ortogonal a la función encontrada.

### ***Método K-NN***

Este método clasifica al objeto nuevo en la categoría a la que pertenecen sus k vecinos más próximos, siendo k un valor arbitrario, normalmente bajo e impar, 3 o 5 [30].

Para aplicar el método se hallan las distancias entre el objeto nuevo y todos los objetos del conjunto de calibración, en ocasiones se utiliza un sistema de puntos ponderados o “votos de distinta calidad”. Una forma frecuente de ponderar los puntos consiste en asignar k puntos a la categoría del vecino

más próximo, y así sucesivamente. Se suman los puntos obtenidos por cada clase, y el objeto nuevo se asigna a la clase que obtiene más puntos.

Las distancias pueden calcularse directamente sobre la matriz de objetos (variables sin escalar), y también antes o después de realizar una rotación propia, si se realiza la rotación propia se tiene la ventaja de trabajar en el espacio no correlacionado, donde las distancias Euclídeas tienen significado correcto.

El mayor inconveniente del método KNN respecto al análisis discriminante es que no calcula las probabilidades de pertenencia del objeto nuevo a cada categoría. Las ventajas de este método es que la calidad de las predicciones depende de la forma que tengan más clases. Por esta razón, el método puede clasificar correctamente objetos aunque las clases no sean linealmente ni cuadráticamente separables. El método KNN establece fronteras irregulares entre las clases.

### ***SIMCA (Soft Independent Modeling of Class Analogy)***

El método SIMCA de reconocimiento de patrones desde su introducción por Svante Wold en 1976 [31] ha sido uno de los métodos de clasificación más utilizados en el control de calidad de las industrias alimentarias.

SIMCA es un método supervisado de reconocimiento de patrones que se basa en el principio de analogía entre las muestras que pertenecen a una misma clase, y emplea para el cálculo de las distancias, los scores determinados mediante el análisis por componentes principales. (PCA). El método SIMCA calcula un modelo PCA para cada clase o categoría presente en el sistema objeto de estudio, posteriormente integra cada una de las clases y calcula sus límites o fronteras con una probabilidad dada, comúnmente del 95%.

En el caso en que no esté disponible la información externa, las categorías o clases se pueden definir a partir de una evaluación de las agrupaciones mediante el análisis exploratorio de datos (reconocimiento de patrones no supervisado) que proporcionan métodos como el análisis de componentes principal (PCA). Una vez definidas las clases o categorías, se construyen y mejora el modelo para el conjunto de muestras de calibración; y después éste se utiliza para predecir las clases de nuevas muestras desconocidas (conjunto de validación). Una muestra desconocida puede asignarse a una categoría, a más de una, o a ninguna; y se calcula la probabilidad de cada asignación.

#### **II.2.4 Tratamiento Digital de imágenes**

El objetivo general de un sistema de visión artificial es analizar imágenes para extraer superficies u objetos o para distinguir zonas de interés del resto de la superficie, es decir, extraer partes de una imagen que corresponden con las zonas de interés, este proceso se denomina segmentación.

El proceso de segmentación de imágenes se basa en las siguientes propiedades:

- Propiedades geométricas como la forma, tamaño, o localización.
- Propiedades de los materiales como el color, iluminación, textura o composición.

##### Propiedades geométricas

En los diferentes trabajos la separación de los objetos se llevo a cabo sobre las imágenes infrarrojas para identificar objetos, de modo que separan los diferentes objetos en la imagen para construir una máscara para aplicar a la imagen RGB a la clasificación de los colores y etiquetado. Los operadores

morfológicos se utilizan para mejorar problemas inherentes de la imagen original, como objetos que se tocan, o problemas del propio proceso de las imágenes, por ejemplo, pequeñas áreas mal iluminadas.

Con el fin de solucionar estos problemas se pueden aplicar distintos operadores morfológicos para: separar zonas, cribar áreas de determinados tamaños, erosionar o dilatar regiones, mirar conectividades, extraer bordes, etc.

Dos de las operaciones morfológicas más utilizadas en análisis de imágenes son la erosión y la dilatación. La erosión consiste en examinar cada píxel y cambiarlo de ON a OFF si alguno de sus vecinos está en OFF. Normalmente se utilizan como vecinos los ocho que rodean al píxel examinado, aunque para algunas aplicaciones se pueden utilizar conectividades de 4 vecinos (los 2 verticales y los 2 horizontales) e incluso conectividades de 2 vecinos (los verticales o los horizontales).

La dilatación es el proceso inverso, consiste en cambiar los píxeles de OFF a ON si alguno de sus vecinos esta ON.

Al resultado de una erosión más una dilatación se le denomina apertura ("opening"). El nombre proviene de la tendencia de esta secuencia de operaciones a separar (abrir) puentes de unión entre objetos próximos o a abrir cavidades próximas al borde.

El siguiente paso para identificar objetos en la imagen monocromática de IR, es la aplicación de filtros de imagen que sirven para la igualación de contraste y mejora espacial de la imagen. Las operaciones de filtrado espacial se aplican a una imagen para mejorar o atenuar detalle espacial con el fin de mejorar el rendimiento visual o para facilitar su posterior procesamiento, y constituyen una de las técnicas implicadas en la mejora de las imágenes. La

operación de filtrado espacial es un proceso local en el procesamiento de imágenes, en el sentido de cambiar el valor de cada píxel de acuerdo con los valores de los píxeles circundantes, se trata de la transformación de los niveles de gris originales, de manera que se asemejan o diferencian más que los de los píxeles vecinos. A continuación se enumeran los filtros más usados en los trabajos de esta tesis:

Filtro 'Flatten': reduce las variaciones de intensidad en los píxeles de fondo cuando encontramos condiciones de iluminación heterogéneas.

Filtro de Gauss: es un filtro muy usado antes de la detección de bordes. Se basa en un promedio de los píxeles de la imagen original, y se emplea para disminuir la ganancia de las altas frecuencias.

La tarea de identificación de objetos se llevó a cabo a continuación del filtrado utilizando dos algoritmos: Border-4 Neighbor para determinar los bordes de los objetos, teniendo en cuenta los valores de intensidad de píxeles en el histograma y el algoritmo de componentes conectados para identificar cada objeto como un valor para determinar los píxeles contenidos en ellos y enumerar cada conjunto de píxeles contenidos en un objeto [32]. A partir de los datos de la matriz de componentes similares pueden crear una máscara, en la que también se aplican varios filtros de tamaño de manera que se puede medir objetos completos y elimina automáticamente los objetos que se solapan con un área menor que o mayor que los límites definidos [33].

En base a estos objetos construir la imagen esqueleto. El esqueleto está destinado a representar la forma de un objeto con un número relativamente pequeño de píxeles. Así, todos los píxeles del esqueleto son estructuralmente necesarios. La posición, orientación y longitud de las líneas esqueleto corresponden a aquellos equivalentes de la imagen original. La tarea de

extracción de características de una imagen se simplifica para obtener su esqueleto en el que se encuentran perfectamente separados los objetos.

Este esqueleto se añade en la imagen RGB en la que se realiza la segmentación por color una vez los objetos están adecuadamente separadas.

### Propiedades de los materiales

El procedimiento que se realizó en este paso fue discriminar diferentes áreas de la imagen usando un algoritmo de clasificación del color. Dicho algoritmo convierte la imagen RGB al espacio  $L^*a^*b$ .

La segmentación basada en el color se puede realizar por dos métodos diferentes: por el método supervisado o no supervisado para la obtención y cuantificación del color predominante en cada uno de los objetos de la imagen.

En el método de supervisión es necesaria para determinar los colores que desee para cuantificar el color de su interés en la imagen. El color de interés en la imagen original se selecciona con una función de selección de polígonos, esta función toma una muestra para la región que contiene el color de interés, y utiliza un color promedio de cada región que se utiliza para identificar los colores y etiquetarlos (marcadores de colores). Esta clasificación se realiza por el método del vecino más cercano (KNN-k vecino más cercano) [34], donde cada píxel se clasifica en la misma clase que los datos de entrenamiento con una intensidad más alta. Los vecinos más cercanos  $k$  asigna un valor de 'a' y 'b' para cada marcador, así es posible clasificar cada píxel en la imagen para calcular la distancia euclídea entre los píxeles y los marcadores de color.

El método de segmentación del color no supervisado de utiliza técnicas de agrupación (clustering) [35,36]. Los algoritmos de agrupación realizan las



mismas funciones que los métodos de clasificación supervisados, pero sin el uso de datos de entrenamiento o referencia. En este algoritmo la idea básica es suponer que los píxeles de la imagen son puntos de un espacio tridimensional (RGB), de modo que los puntos de una región de color similar se agrupan juntos. Las técnicas de agrupamiento permiten la obtención de un punto representativo de cada grupo, con base en las medidas de similitud o distancia entre estos puntos. En este algoritmo un punto  $x$  se asigna a un grupo  $R$  cuyo centroide  $C$  esté más cerca de ese punto. Normalmente se usa como una medida de la distancia euclídea entre el punto y el centroide. De esta forma se procede a crear un vector que contiene las etiquetas para cada color, para luego inicializar la matriz que se usa para la clasificación, para posteriormente realizar la clasificación por color.

**Bibliografía**

- [1] Wold, S. Chemometrics: Theory and Application; American Chemical Society 243 (1998)
- [2] Massart, D.L., Vandeginste, L.M., Buydens, M.C. Chemometrics: A textbook. Elsevier Science Publisher, 23, 395-397 (1988).
- [3] Richard, G. Brereton. Chemometrics, Data Analysis for the Laboratory and Chemical Plant. John Wiley & Sons. Ltd (2002)
- [4] <http://www.camo.com/Software/Overview/unscrambler.htm>.
- [5] Matlab v7.0 (The Mathworks, Inc., Natick, MA).
- [6] Zeaiter, M., Rutledge, D. Preprocessing Methods. Comprehensive Chemometrics. Oxford, Elsevier 121-231 (2009)
- [7] Todeschini, R., Introduzione alla Chimimetria. EdiSES, Napoli (1998)
- [8] Forina, M., Introduzione alla Chimica Analitica con elementi di Chimimetria. ECIG (Edizioni Culturali Internazionali Genova), 1ª edizione (1993)
- [9] Savitzky, A., Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedure. Anal. Chem., 36: 1627-1639 (1964)
- [10] Bouveresse, E., Maintenance and Transfer of Multivariate Calibration Models Based on Near-Infrared Spectroscopy, doctoral thesis, Vrije Universiteit Brussel (1997)
- [11] Barnes RJ, Dhanoa MS, Lister SJ: Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra. Appl Spectrosc, 43(5):772-777 (1989)
- [12] Dhanoa MS, S. J. Lister RS, Barnes RJ: The Link between Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) Transformations of NIR spectra. Journal of Near Infrared Spectroscopy, 2:43 (1994)

- [13] Massart, D.L., Vandeginste, B.G.M., Deming, S.N., Michotte, Y., Kaufman, L., Data Handling in Science and Technology, volume 2. Chemometrics: a textbook. Elsevier Science Publishers, Amsterdam, The Netherlands (1988)
- [14] Wold, S. Pattern Recognition by means of disjoint principal components models. Pattern Recognition 8 (3), 127 (1978)
- [15] Snee, R.D. 1977 Validation of regression models: method and examples. Technometrics, 19: 415-428 (1977)
- [16] Kennard, R.W., Stone, L.A. 1969. Computer aided design of experiments. Technometrics, 11: 137-148
- [17] Barnett, V. y Lewis, T. (1994): Outliers in Statistical Data. 3st. Ed. Wiley and Sons.
- [18] J.A. Fernández Pierna, F. Wahl, O.E. de Noord, D.L. Massart Methods for outlier detection in prediction a Chemometrics and Intelligent Laboratory Systems 63, 27– 39 (2002)
- [19] J.A. Fernández Pierna, , L. Jina, M. Daszykowski, F. Wahl, D.L. Massart. A methodology to detect outliers/inliers in prediction with PLS Chemometrics and Intelligent Laboratory Systems 68, 17– 28 (2003)
- [20] S. De Vries, J.F. Cajo, C.J.F. ber Braak, Chemometr. Intell. Lab. Syst. 30 (1995) 239– 245.
- [21] Massart, S. de Jong, O.E. de Noord, C. Puel, B.G.M. Vandeginste, Internet J. Chem. 2 (1999) 19.
- [22] W.H. Beyer, Standard Mathematical Tables, 26th edn., CRC Press, Boca Raton, Florida, 1981
- [23] H. Martens, T. Naes, Multivariate Calibration, John Wiley & Sons Ltd., Chichester, 1989.

- [24] Wu D, He Y, Nie PC, Cao F, Bao YD. Hybrid variable selection in visible and near-infrared spectral analysis for non-invasive quality determination of grape juice. *Anal. Chim. Acta.*, 659: 229-237(2010)
- [25] Williams, P.C.; Sobering, D.C. How do we do it: a brief summary of the methods we use in developing near infrared calibrations. En: *Near Infrared Spectroscopy: The Future Waves*. Davies A.M.C., Williams P.C. (Eds.). NIR Publications, Chichester, UK, 185-188 (1996)
- [26] Shenk, J.S., Westerhaus, M.O. 1995b. Routine Operation, Calibration, Development and Network System Management Manual. NIRSystems, Inc., 12101 Tech Road,  
  
Silver Spring. MD 20904. USA
- [27] Wienke, D., Bydens, L. Adaptive resonance theory based neural Networks- the ART of real- time pattern recognition in chemical process monitoring. *Trends Anal.Chem.* 14, 398-406 (1995)
- [28] Massart, D.L. Kaufman, L. Interpretation of Analytical Chemical Data by the Use of Cluster Analysis. Wiley, New York (1983)
- [29] M. Bylesjö, M. Rantalainen, O. Cloarec, J.K. Nicholson, E. Holmes, J. Trygg. *Journal of Chemometrics.* 20 (2006) 8-10 341-351
- [30] T. Yu, The application of chemometrics derived pattern recognition methods to futures market analysis. *Chemistry Dissertations.* 11 (2009)
- [31] S.M. Wold. SIMCA: a method for analyzing chemical data in terms of similarity and analogy. In *Chemometrics: Theory and Application*; Kowalski, B. R., Ed.; ACS Symposium Series 52; American Chemical Society: Washington, DC, pp 243-282 (1977)
- [32] Chang, V.and Saavedra, J. Métodos alternativos para el mejoramiento automático del contraste de imágenes. Computer Engineering thesis, School of Informatic, University National of Trujillo (2001)

- [33] Russ, J.C. The image processing Handbook. CRC Press LLC 3rd Edition.
- [34] Richards, J. A. Remote Sensing Digital Image Analysis. Ed. Springer-Verlag (1999)
- [35] Thomas, S.W. Efficient inverse color map computation. In: Arvo, J. (Ed.), Graphics Gems II. Academic Press, Boston, pp. 116–125 (1991)
- [36] Floyd, R. W., & Steinberg, L. An Adaptive Algorithm for Spatial Gray Scale. In: International Symposium Digest of Technical Papers, Society for Information Displays, 36 (1975)

# Capítulo III.

## Plan de Trabajo



UNIVERSIDAD  
E  
CÓRDOBA

Para alcanzar los objetivos fijados en el desarrollo de este trabajo, ha sido necesario superar diversos niveles de formación en el conocimiento; sobre esta base, algunos objetivos se establecieron en el diseño experimental, mientras que otros surgieron durante el desarrollo de los trabajos experimentales y en el análisis de sus resultados, así como de nuevas ideas surgidas de actividades complementarias. Por tanto, los retos acometidos en este trabajo fueron los siguientes:

- Aplicación de técnicas emergentes en la caracterización de productos hortofrutícolas y aplicación en el análisis de frutos de aceituna. Utilización de equipos ópticos, para evaluar la calidad en frutos de aceituna y aceite de oliva.
- Desarrollo del trabajo en condiciones off-line y su posterior aplicación en condiciones on-line, con el fin de estudiar la viabilidad de los equipos y la posibilidad de su utilización para clasificar productos

Persiguiendo estos retos, los trabajos que han dado lugar a esta Tesis se desarrollaron durante cuatro años, periodo de duración de la beca de Formación de investigadores de tipo Predoctoral del Instituto Nacional de y Tecnología Agraria y Alimentaria (INIA), en el tema “Nuevas tecnologías analíticas para la mejora de la trazabilidad y calidad del aceite de oliva”.

### **III.1 Tareas iniciales**

Un primer paso anterior al desarrollo de los experimentos fue entender tanto la problemática del sector como las nuevas tecnologías aplicables en el ámbito de la agricultura. A este periodo corresponden las siguientes actividades:

- Inmersión en la actividad del laboratorio: conociendo y desarrollando los métodos de análisis actuales para el control de la calidad de aceitunas y aceite, con el fin de entender aún más la necesidad de unas tecnologías analíticas alternativas.
- Recopilación de información técnica sobre los principios de funcionamiento de diversos equipos ópticos, con el fin de conocer sus características y su funcionamiento, de cara a entender correctamente los equipos y sus modos de trabajo.
- Intensificación de la formación científica con la realización de los cursos de doctorado y numerosos cursos de formación sobre análisis de datos e imagen.
- Actualización de la revisión bibliográfica sobre técnicas no-destructivas, propiedades ópticas, métodos espectroscópicos, de visión artificial, etc.

### III.2 Ejecución de los trabajos experimentales

El proceso de estudio y formación, junto con la colaboración interna del grupo de trabajo, permitieron abordar los objetivos inicialmente propuestos, a través de la realización de los trabajos experimentales que se introducen a continuación, agrupados en dos grandes bloques ( datos espectroscópicos e imagen) :

#### Espectroscopia:

- *Aplicación de Espectroscopia Raman de baja resolución para el análisis de aceites de oliva oxidados*
- *Evaluación de la calidad global del aceite de oliva mediante Espectroscopia de Fluorescencia*



- *Sensor portátil Raman para la discriminación rápida de aceitunas según la calidad del fruto*

*Visión Artificial:*

- *Determinación del índice de madurez de frutos intactos de aceituna mediante análisis de imagen*
- *Método automático para la detección en línea de suciedad en el proceso de producción del aceite de oliva*
- *Sistema de visión infrarroja para la detección automática de la calidad en frutos de aceituna*
- *Sistema visible de visión para la determinación on-line de parámetros de calidad de frutos de aceituna*

# Capítulo IV. Metodología



UNIVERSIDAD  
E  
CÓRDOBA

## IV.1 Introducción

El objetivo de esta tesis es desarrollar nuevas tecnologías ópticas que permitan realizar análisis cuantitativos y cualitativos para ser usados en rutina por la industria.

La industria del aceite de oliva garantiza la calidad del producto mediante métodos de análisis que permiten evaluar sus características físico-químicas. El control de calidad de los aceites está claramente establecido en el Reglamento (CE) nº 2568/91 modificado por el (CE) nº 1989/03 de 6 de Noviembre de 2003 [1,2], que define las características físico-químicas y organolépticas de los aceites de oliva y de orujo de oliva, así como de los métodos de valoración de esas características. En este campo se han desarrollado diferentes metodologías espectroscópicas para cuantificar los principales parámetros de calidad del aceite de oliva como ya ha quedado recogido en el capítulo primero. Estas metodologías desarrolladas se han aplicado con el objetivo de sustituir los métodos de análisis habituales durante el proceso de extracción y elaboración del aceite de oliva virgen.

Las metodologías desarrolladas basadas en técnicas ópticas, pueden clasificarse en dos tipos; las basadas en técnicas espectroscópicas y las basadas en sistemas de visión artificial. La presentación conjunta del trabajo realizado según la tecnología utilizada nos mostrará las diferencias existentes entre los desarrollos y el tratamiento equivalente de los datos para lograr la adecuada clasificación de los frutos en cada uno de los equipos. Todos los trabajos realizados giran en torno a la misma unidad temática: aplicación de las “propiedades ópticas” para analizar frutos de aceituna y aceite de oliva mediante métodos no destructivos; de esta forma, tanto los trabajos realizados en el desarrollo de la investigación como sus correspondientes resultados se han agrupado en siete trabajos.

Todos los trabajos realizados se centran en el estudio de las propiedades ópticas de las aceitunas y aceite de oliva y aunque aparentemente muestran una estructura diferente, todos ellos siguen un esquema común. Los datos originales (espectros o imágenes), proporcionados por los diferentes equipos, son sometidos a un pre-tratamiento (eliminación de ruido o centrado entre otros) con el fin de acondicionarlos para el procesamiento matemático posterior (análisis discriminante, regresión lineal simple, clasificación, etc).

En los apartados que siguen a continuación, se describen las características más relevantes de las metodologías empleadas en el desarrollo de los nuevos métodos propuestos: el tipo de muestras analizadas, el registro de los espectros y/o toma de imágenes, la obtención de la propiedad analítica de interés mediante los métodos de referencia y el procesamiento de datos o imágenes.

## **IV.2 Metodología desarrollada para sistemas espectroscópicos**

### *IV.2.1 Descripción y preparación de las muestras*

La selección de un conjunto muestral adecuado es un paso muy importante del proceso metodológico, ya que de él depende en gran medida el éxito de los modelos construidos. Por ello es importante crear un grupo de muestras representativo que cubra el rango de valores esperables para el producto estudiado.

Las muestras utilizadas en los trabajos realizados son aceitunas cosechadas y aceite de oliva producido en las instalaciones del Centro IFAPA "Venta del Llano" de Mengíbar, Jaén. Las principales variedades de aceituna utilizadas son la variedad picual y arbequina. El conjunto de calibración deberá comprender muestras uniformemente distribuidos en un rango mayor que el de las muestras que serán analizadas por el modelo.

Normalmente cuanto mayor es el número de muestras utilizado en la calibración, mejor es la capacidad predictiva del modelo y menor el error de predicción.

Incorporar la variabilidad química al conjunto de calibración es una tarea que se reduce a preparar muestras en el laboratorio con el intervalo de valores deseado. El problema reside en cómo incorporar la variabilidad física del proceso al conjunto de calibración.

La situación real más simple que se puede encontrar es cuando es posible generar artificialmente las muestras en el laboratorio (forzando el deterioro de las muestras de aceituna obteniendo aceites con unos índices de calidad fuera de los valores normales encontrados en la planta de producción), así algunos grupos de muestras han sido sometidas a un proceso de oxidación o/y acidificación controlado con el objetivo de obtener muestras con valores extremos. Es muy recomendable incluir también muestras reales extraídas del proceso de producción ya que nos aseguramos de incluir en el proceso de estudio una gran variabilidad en las muestras y todas fuentes de variación inesperadas.

#### *IV.2.2 Instrumentación y métodos de análisis*

A continuación se describen los instrumentos espectroscópicos utilizados a lo largo de esta tesis y los diferentes métodos de análisis:

##### *IV.2.2.1 Espectroscopia Raman*

El instrumento Raman utilizado fue un modelo de espectrómetro portátil RH-3000 (Ocean Optics). Este equipo está dotado de un láser de diodo a la longitud de onda de 785 nm conjuntamente con detectores de estado sólido tipo CCD, lo que permite lograr instrumentos de alta sensibilidad y gran solidez. Éste láser se relacionó con una fibra óptica para ser llevado a través

de un recipiente de vidrio. Los espectros Raman se registraron 200 a 2700  $\text{cm}^{-1}$ , con una resolución espectral de 10  $\text{cm}^{-1}$ . El diodo láser del instrumento tenía un rango de intensidad de 0-290 mW.

En este trabajo el análisis sobre muestras de aceite de oliva se realizó directamente a través del contenedor de vidrio sin necesidad de realizar ningún tipo de procesado ni trasvase de la muestra.

Las medidas Raman en pasta de aceituna fueron tomadas sobre un instrumento prototipo que simula la salida de la pasta de aceituna del molino en la fábrica, en el que había una entrada y salida continua de aceituna. En este instrumento fue acoplada la sonda Raman, registrando automáticamente los espectros Raman directamente sobre la muestra a través de una ventana de vidrio.



Figura 1. Sistema diseñado para este trabajo con el fin de tomar las mediciones Raman de pasta de aceituna

Los espectros Raman de muestras de pasta de aceituna y aceite se registraron por triplicado, cada espectro analizado es el resultado de promediar estas tres medidas. La medida mediante el equipo Raman se realiza directamente sobre el material a analizar, no requiere ningún tipo de preparación de la muestra, se pueden analizar materiales en estado sólido, líquido o gaseoso, y a través de contenedores.

Diariamente, antes de comenzar el registro de los espectros de las muestras, se realiza la comprobación del correcto funcionamiento del equipo por comparación con medidas estándar.

### *III.2.2.2 Espectroscopia de Fluorescencia*

La fluorescencia del aceite de oliva fue medida mediante un espectrofotómetro de fluorescencia Cary Eclipse (Varian Ibérica, Madrid, España). Este instrumento está equipado con una lámpara continua de xenón, monocromadores de excitación y de emisión, y un fotomultiplicador. Una célula no fluorescente 10 x 10 mm (3 ml) se utilizó para tomar las medidas en "ángulo recto" (RA). En una primera parte del trabajo se obtuvo una matriz tridimensional (3D) para cada muestra, una matriz de fluorescencia de excitación-emisión (EEM). La principal ventaja de EEM es que se puede extraer más información sobre las especies fluorescentes porque hay una zona más amplia disponible para el análisis. Para este análisis de barrido se utilizó una longitud de onda de emisión de 300-800 nm y una longitud de onda de excitación de 300-400 nm. En un segundo análisis, el espectro de fluorescencia se midió a una longitud de onda específica (385 nm). El tiempo de integración fue 0,1 s, y la longitud de onda creciente mientras se escanea el espectro era 10 nm. Las condiciones experimentales fueron ensayadas para conseguir espectros de intensidad satisfactoria, buena resolución y relación señal-ruido.

Antes de cada medida los aceites se agitaron vigorosamente con el fin de homogeneizar las muestras. A continuación se vertieron en una cubeta de plástico, que se desechó en cada medida para garantizar la seguridad y fiabilidad de las mediciones y para mejorar la velocidad de método.

Cada muestra se analizó por duplicado, y la medición promedio (el valor de las dos exploraciones para cada muestra) fue usada, sin ningún tratamiento previo.

### *III.2.3 Métodos de referencia*

A continuación se describen las características más relevantes de los métodos usados como referencia:

Los principales parámetros que permiten la determinación de la calidad del aceite y aceituna y que han sido estudiados en estos trabajos se refieren a su estado de oxidación y grado de acidez.

#### *Determinación del grado de acidez*

Tiene por objeto la determinación de los ácidos grasos libres en el aceite de oliva. El aceite desde el punto de vista químico está compuesto por triglicéridos, es decir, ésteres de los ácidos grasos y la glicerina. La reacción de hidrólisis provoca la ruptura de estos, perdiendo ácidos grasos y dando diglicéridos y monoglicéridos. El porcentaje en que estos ácidos grasos libres se encuentran en el aceite, referidos a porcentaje de ácido oleico, es lo que se conoce como grado de acidez.

El método consiste en la disolución de la muestra en una mezcla de disolventes y valoración de los ácidos grasos libres mediante una disolución etanólica de hidróxido potásico, usando fenolftaleína como indicador.

Se disuelve la muestra de aceite filtrado en 50 ml de una mezcla de éter etílico y etanol 1:1, previamente neutralizada.

Valorar, agitando, con una disolución de hidróxido potásico 0,1 N hasta el viraje del indicador.

La acidez expresada en porcentaje de ácido oleico es igual a:



$$V * c * M/1000 * 100/P = V * c * M/ 10 * P$$

Siendo:

- V: volumen en ml de la disolución valorada de hidróxido potásico utilizada. Convenientemente corregidos teniendo en cuenta la prueba en blanco.
- c: concentración exacta, en moles litro, de la disolución de hidróxido potásico utilizada.
- M: peso molecular del ácido en que se expresa el resultado (ácido oleico=282)
- P: peso en gramos de la muestra utilizada.

Se utiliza la media aritmética de dos determinaciones.

#### Determinación de índice de peróxidos

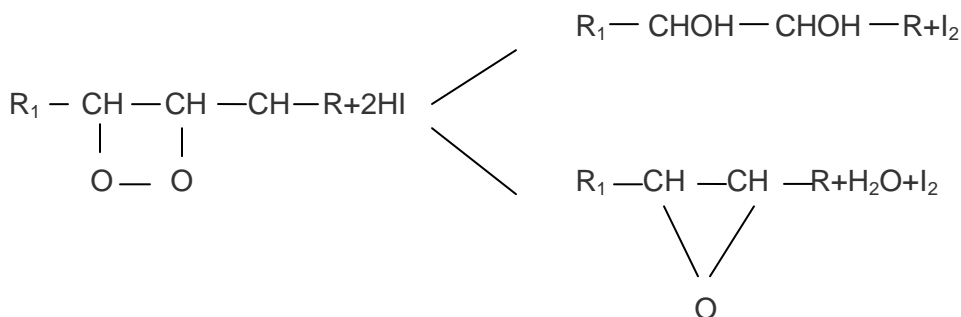
Los peróxidos son productos de oxidación existentes en una muestra, en un momento determinado. El índice de peróxidos mide el grado de oxidación primaria de un aceite, y nos indica el estado de conservación del mismo. Se expresa en miliequivalentes de oxígeno activo por kilo de grasa. Su determinación se realiza por volumetría. El índice de peróxidos es la cantidad de la cantidad de peróxidos en la muestra que provocan la oxidación del yoduro potásico en las condiciones de trabajo descritas. La muestra problema, disuelta en ácido acético y cloroformo, se trata con una solución de yoduro potásico. El yodo liberado se valora con solución valorada de tiosulfato sódico.

Los peróxidos se consideran como los primeros productos de oxidación de las grasas y su formación sigue, al menos durante las primeras etapas, una marcha paralela a la cantidad de oxígeno absorbido.

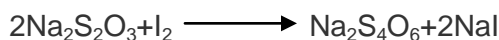
Existen varios factores que influyen sobre la velocidad de la oxidación de las grasas. Unos retardándola, como son ciertas sustancias denominadas antioxidantes y otras acelerándola. Dentro de éstos últimos están:

- Luz
- Calor
- Trazas metálicas
- Catalizadores orgánicos, etc.

El fundamento de la determinación volumétrica es la capacidad de los peróxidos para liberar yodo del IK según la siguiente reacción:



El yodo liberado se valoró con tiosulfato sódico según la reacción:



Es decir, un mol de oxígeno peroxidico deja en libertad un mol de yodo.

La muestra se tomará y almacenará resguardada de la luz, y se mantendrá refrigerada dentro de los envases de vidrio totalmente llenos y herméticamente cerrados con tapones de vidrio esmerilado y corcho.

El ensayo se realizará con luz natural difusa o con luz artificial. Pesar con precisión 0,001 g en una navecilla de vidrio.

Abrir un matraz e introducir la navecilla de vidrio que contenga la muestra problema. Añadir 10 ml de cloroformo. Disolver rápidamente la muestra problema mediante agitación. Añadir 15 ml de ácido acético y, a continuación, 1 ml de solución de yoduro potásico. Cerrar rápidamente el matraz, agitar durante 1 minuto y mantenerlo en la oscuridad durante 5 minutos exactamente, a una temperatura comprendida entre 15 y 25 °C.

Añadir 75 ml aproximadamente de agua destilada. Valorar (agitando al mismo tiempo vigorosamente) el iodo liberado con la solución de tiosulfato (solución 0,002 N si se presuponen valores bajos de oxidación y 0,1 si son altos), utilizando la solución de almidón como indicador.

Efectuar dos determinaciones por muestra.

Realizar simultáneamente el ensayo en blanco. Si el resultado sobrepasa 0,05 ml de la solución de tiosulfato sódico 0,01 N, sustituir los reactivos.

El índice de peróxidos (IP), expresado en miliequivalentes de oxígeno activo por kg de grasa se calcula mediante la fórmula siguiente:

$$IP = V \cdot N \cdot 1000 / P$$

siendo:

V : ml de solución valorada de tiosulfato sódico empleados en el ensayo, convenientemente corregidos para tener en cuenta el ensayo en blanco.

N : normalidad exacta de la solución de tiosulfato sódico empleada

P : peso en gramos de la muestra problema.

El resultado será la media aritmética de las dos determinaciones efectuadas

#### Prueba espectrofotométrica en el ultravioleta

La prueba espectrofotométrica en el ultravioleta puede proporcionar indicaciones sobre la calidad de una materia grasa, su estado de conservación y las modificaciones inducidas por los procesos tecnológicos.

Las absorciones en las longitudes de onda indicadas en el método se deben a la presencia de sistemas diénicos y triénicos conjugados. Los valores de estas absorciones se expresan en extinción específica  $E_{1\text{ cm } 1\%}$  (extinción de una solución de la materia grasa al 1% en el disolvente determinado, en un espesor de 1 cm) que se expresará convenientemente como K, también denominado coeficiente de extinción. El aceite de oliva virgen, de buena calidad y almacenado convenientemente, contiene muy pocos productos de oxidación y presenta una extinción específica en el entorno de 270 nm.

La materia grasa se disuelve en ciclohexano y se determina la extinción de la solución a las longitudes de ondas prescritas, respecto al ciclohexano puro. A partir de los valores espectrofotométricos se calculan las extinciones específicas.

La muestra debe ser perfectamente homogénea y estar exenta de impurezas en suspensión. Se pesan con precisión 0,25 g aproximadamente de la

muestra preparada y se colocan en un matraz aforado de 10 ml, se completa con ciclohexano y se homogeneiza. La solución resultante debe estar perfectamente clara. Se llena una cubeta con la solución obtenida y se miden las extinciones, usando como referencia el ciclohexano, a las longitudes de onda comprendidas entre 232 y 276 nm. Los valores de extinción deben estar comprendidos en el intervalo entre 0,1 y 0,8; en caso contrario es necesario repetir la medida utilizando soluciones más concentradas o más diluidas según el caso.

Se expresan las extinciones específicas o coeficientes de extinción a las diversas longitudes de onda, calculadas como sigue:

$$K_{\lambda} = E_{\lambda} / C \cdot e$$

siendo:

$K_{\lambda}$  : extinción específica a la longitud de onda " $\lambda$ "

$E_{\lambda}$  : extinción medida a la longitud de onda " $\lambda$ "

C : concentración de la disolución en g por 100 ml

e : espesor de la cubeta en cm.

Los resultados deben expresarse con dos cifras decimales.

#### *IV.2.4 Construcción y validación de modelos cuantitativos y cualitativos*

En este desarrollo se ha buscado el mejor método para procesar la información contenida en los espectros, para ello se han aplicado diferentes técnicas de análisis multivariante.

El primer paso es la selección del conjunto de muestras de calibración. Existen diferentes métodos de selección de los conjuntos de calibración y validación, en este trabajo se ha utilizado un método de selección de división aleatoria, denominado Dúplex, que asignan observaciones aleatoriamente generando subconjuntos de calibración y validación homogéneos e independientes entre sí.

La amplia cantidad de trabajos referentes a técnicas espectroscopias, pone de manifiesto la importancia de estas técnicas así como de las posibilidades de su aplicación; en la mayoría de los trabajos, agrupados en el capítulo 1, se puede observar que las técnicas de análisis multivariante más utilizadas en espectroscopia son los mínimos cuadrados parciales (PLS) y el análisis de componentes principales (PCA); de ahí que en el presente trabajo para el estudio de los espectros se haya optado por ambos métodos para el estudio en profundidad de los espectros procedentes de los equipos.

En todos los trabajos realizados el primer paso fue la exploración de los datos espectrales mediante el análisis de componentes principales (PCA) [3]. Este análisis se utilizó con el objeto de establecer posibles diferencias entre las características espectrales de las muestras. Se ensayaron diferentes pre-tratamientos, combinaciones y rangos de longitudes de onda, buscando la mejora de los datos espectrales, pero no se obtuvieron mejoras significantes, así como la aplicación de métodos de detección de datos aberrantes previo a la construcción del modelo de calibración.

El número de componentes principales utilizados siempre superó el 95% de la varianza explicada.

Esta técnica multivariante PCA permitió detectar diferencias entre muestras en base a sus correspondientes espectros, identificando un sub-espacio en el que las variaciones espectrales están más ligadas a la respuesta [4] lo que permitió posteriormente su análisis por el procedimiento del PLS.

Los métodos de análisis cuantitativos que han sido desarrollados en estos trabajos tenían como objetivo la determinación de propiedades analíticas de interés en las diferentes muestras. Estas propiedades fueron el índice de peróxidos, acidez y los coeficientes de extinción a 232 y 270 nm.

Estos métodos se basan en la utilización de modelos de calibración por mínimos cuadrados parciales (PLS)[5]. En la presente tesis para el desarrollo de los modelos de calibración se ha utilizado el método de validación cruzada dejando uno fuera o leave-one-out (LOOCV), que implica separar los datos de forma que para cada iteración tengamos una sola muestra para los datos de prueba y todo el resto conformando los datos de entrenamiento.

La selección del número de factores óptimo, para cualquiera de los modelos PLS, se hizo atendiendo a los valores más bajos de error de calibración (RMSEC), error estándar de validación cruzada (RMSECV) y error estándar de validación (RMSEP). También se tuvieron en cuenta los valores más elevados de los coeficientes de determinación en calibración ( $R^2$ ) y en validación ( $r^2$ ) y los valores del índice RPD.

Para llevar a cabo una buena interpretación de un modelo de regresión debemos acompañar siempre nuestro estudio de la diagnosis y validación del modelo, con el fin de averiguar, por un lado, hasta qué punto los resultados obtenidos son extrapolables a la población de la que vienen las muestras

seleccionadas. En la presente tesis la validación se realizó en todos los casos mediante un conjunto de muestras externo independiente a la calibración.

El modelo de regresión se calcula a partir de los datos del conjunto de muestras de calibración y, el modelo calculado, se valida con los datos del conjunto de muestras de validación que no han sido utilizados para calcular la ecuación de calibración del modelo. Para estos modelos se tuvo en cuenta el error de predicción al cuadrado (RMSEP) para el conjunto de validación.

En el trabajo referente a la discriminación de pasta de aceituna mediante su espectro Raman, las técnicas utilizadas fueron técnicas de clasificación que son modelos capaces de pronosticar la pertenencia de una muestra a una categoría sobre la base de las características espectrales del producto.

Los métodos utilizados en este trabajo, la clasificación se basa en el conocimiento previo del sistema, con conjuntos de calibración de clase conocida, que permiten definir cada clase.

En el caso en que no esté disponible la información externa, las categorías o clases se pueden definir a partir de una evaluación de las agrupaciones mediante el análisis exploratorio de datos (reconocimiento de patrones no supervisado) que proporcionan métodos como el análisis de componentes principal (PCA). Una vez definidas las clases o categorías, se construyen y mejora el modelo para el conjunto de muestras de calibración; y después éste se utiliza para predecir las clases de nuevas muestras desconocidas (conjunto de validación). Una muestra desconocida puede asignarse a una categoría, a más de una, o a ninguna; y se calcula la probabilidad de cada asignación.



En este trabajo se utilizaron los métodos de análisis discriminante lineal [6], método K-NN [7] y el modelado suave independiente de analogía de clases (SIMCA) [8]. Se realizó una comparación entre ellos para un conjunto de datos, obteniendo los mejores resultados con el método K-NN como se puede observar en el trabajo resultante correspondiente.

### **IV.3 Metodología desarrollada para sistemas de visión artificial**

#### *IV.3.1 Descripción y preparación de las muestras*

La aplicación del análisis de la imagen en la clasificación de aceitunas requiere una adecuada captación de las mismas, para poder realizar el correcto tratamiento de los datos que proporcionan. Las imágenes adquiridas de las aceitunas en las primeras etapas de este trabajo tenían como función esencial la caracterización de las diferentes categorías.

Se consideró que las imágenes disponibles podrían permitir realizar una adecuada clasificación de las aceitunas en función del color externos, suciedad, y superficie de defectos.

Se prepararon un número considerable de conjuntos de frutos, de la variedad picual principalmente, de diferentes zonas y momento de recolección, así como frutos del suelo, helados y enfermos, para que las imágenes registradas posean la máxima variabilidad.

#### *IV.3.2 Instrumentación y métodos de análisis*

En los trabajos realizados en esta tesis doctoral hemos utilizado diversas condiciones de trabajo y en consecuencia métodos de trabajo diferentes. Se han utilizado equipos para el análisis de imagen en laboratorio e instalaciones de trabajo en línea en almazara, mediante adquisición de

imágenes en el espectro visible (formato RGB) e infrarrojo. El equipo utilizado fue una cámara digital modelo JAI AD-080CL dotada de dos sensores CCD adaptados a un prisma, uno color Bayer que puede generar imágenes RGB de 24 bits y uno infrarrojo que captura imágenes monocromo. Esta tecnología permite capturar dos imágenes exactas del mismo campo de visión, una de ellas en color (filtro Bayer) y la otra en infrarrojo.

Uno de los factores más importantes es una adecuada captación de las imágenes, y para ello es fundamental una correcta iluminación de la muestra. La existencia de sombras o deficiencias en la iluminación hacen que el análisis sea más complicado, produciéndose por tanto una pérdida de información. En estos trabajos se realizaron pruebas con diferentes sistemas de iluminación, demostrándose que la luz halógena era la fuente óptima que permitía trabajar en ambos campos de visión (VIS e IR) simultáneamente.

En los diferentes trabajos se han utilizado frutos de características distintas (variedad, tamaño, color y condiciones sanitarias) el objetivo es determinar el índice de madurez, contenido en suciedad y calidad del fruto de aceituna que son parámetros que tienen interés comercialmente.

#### *IV.3.3 Métodos de referencia*

Actualmente la clasificación de los frutos de aceituna se realiza mediante métodos estrictamente visuales. Aunque la clasificación visual de las aceitunas se realiza correctamente por personal especializado, presenta serios problemas por coste de mano de obra, fatiga y fallos debidos al error humano. A partir de la clasificación efectuada por los sistemas de visión tanto en los instalados en las línea de proceso como los similares en el laboratorio, y aplicando un tratamiento de análisis de referencia (visual), realizados en laboratorio, nos ha permitido evaluar la eficacia de estos sistemas.

Los métodos de referencia utilizados para los diferentes trabajos de visión artificial han sido:

*Determinación del índice de madurez (IM)*

El IM fue determinado por un evaluador con experiencia cada día de muestreo, mediante una evaluación del color de la piel y la pulpa de las aceitunas [9]. El procedimiento consistió en la distribución de las aceitunas en ocho grupos de acuerdo con las características resumidas en la Tabla 1. El método oficial utiliza 100 aceitunas y calcula un valor global de índice de madurez para cada muestra o grupo de aceitunas.

El IM se calcula utilizando la siguiente ecuación:

$$IM = \sum i * ni/100$$

donde  $i$  es el número de grupo y  $Ni$  el número de aceitunas en ella.

Este método tiene que separar manualmente las aceitunas, cortar la pulpa para su examen y contar e identificar el grupo al que pertenecen.

Según este examen, las aceitunas se clasifican en uno de los grupos que se resumen a continuación:

Grupo Índice de Maduración	Descripción
0	Piel verde intenso
1	Piel verde-amarillento
2	Color piel con < mitad del fruto color rojizo, púrpura o negro
3	Color piel con > mitad del fruto color rojizo, púrpura o negro
4	Color piel toda púrpura o negra con la p blanca o verde
5	Color piel toda negra con < Mitad pulpa púrpura
6	Color piel toda negra con > Mitad pulpa púrpura
7	Color piel toda negra con toda la pulpa púrpura hasta el hueso

Tabla 1 Clases índice de madurez

### Cuantificación de la cantidad de suciedad

La cantidad de suciedad (hojas, piedras, barro) se pesó y se contaron manualmente el número de partículas en cada muestra con el fin de obtener valores de referencia de la cantidad de suciedad presente en cada imagen y para la validación del método de detección por visión artificial. Los resultados se obtuvieron como porcentaje del total de las hojas, piedras y aceitunas en cada muestra. Por lo tanto, podemos comparar los resultados obtenidos con

el método propuesto visión y el contenido real de la suciedad en cada muestra y se puede obtener una calibración que permite una estimación del peso o la cantidad de suciedad en las muestras externas.

#### Determinación del estado de calidad de las aceitunas

Para realizar la clasificación de las aceitunas en función de su grado de defecto o calidad, se definieron cinco clases o grupos. La tabla 2 define estas clases: aceitunas sanas, aceitunas con daños menores, aceitunas con daños moderados, aceitunas con daños severos, y las aceitunas defectuosas.

Categoría		
0	Aceituna Sana	100% area sana
1	Defectos Menores	>75% area sana
2	Defectos Moderados	>50 % area defecto
3	Defectos Severos	>75% area defecto
4	Aceituna Defectuosa	100 % area defecto

Tabla 2 Grupos de clasificación definidos para los defectos en aceitunas

#### IV.3.4 Construcción de los modelos cuantitativos y cualitativos

Para la creación de los modelos de clasificación de las aceitunas por su índice de madurez, cantidad de suciedad y calidad del fruto por análisis de imagen, el proceso fue simulado al completo en el propio laboratorio, para calibrar el equipo en condiciones más controladas, ya que las medidas en línea pueden verse afectadas por muchos más factores como la luz exterior, vibraciones propias de los frutos que pueden afectar a la imagen, etc. Una vez desarrollados en el laboratorio, se llevo a cabo la instalación de este equipo en línea, para probar su eficacia para medidas en línea.

La siguiente etapa sería preparar un software capaz de realizar todos los pasos que a continuación se presentan y que lleguen a clasificar con el menor error, cada fruto en la categoría que le corresponda.

El pre-tratamiento de las imágenes que pueden simplificarse en:

1. Identificación de objetos en la imagen IR
2. Extracción del color o defecto del área de cada aceituna
3. Cuantificación y etiquetado

El primer paso para identificar objetos en la imagen monocromática de IR, es la aplicación de filtros de imagen que sirven para la igualación de contraste y mejora espacial de la imagen. Las operaciones de filtrado espacial se aplican a una imagen para mejorar o atenuar detalle espacial con el fin de mejorar el rendimiento visual o para facilitar su posterior procesamiento, y constituyen una de las técnicas implicadas en la mejora de las imágenes.

La identificación de objetos se llevó a cabo a continuación del filtrado utilizando dos algoritmos: Border-4 Neighbor [9] para determinar los bordes de los objetos, teniendo en cuenta los valores de intensidad de píxeles en el histograma y el algoritmo de componentes conectados para identificar cada objeto como un valor para determinar los píxeles contenidos en ellos y enumerar cada conjunto de píxeles contenidos en un objeto. En base a estos objetos construir la imagen esqueleto. Una vez realizada la separación de los objetos en la imagen, se procede a realizar la segmentación necesaria para cada una de las metodologías. En los modelos de clasificación por color, como es el caso de la determinación del índice de madurez y suciedad en la aceituna, este esqueleto se añade en la imagen RGB en la que se realiza la segmentación por color una vez los objetos están adecuadamente separados. Y en el caso de la clasificación de los frutos según su calidad, se realiza una segmentación por niveles de grises directamente sobre los objetos.

El último paso sería la aplicación de algoritmos de agrupamiento y etiquetado para la cuantificación del parámetro analizado.

### Segmentación por color

Esta segmentación basada en el color se puede realizar por métodos supervisados o no supervisados, con el fin de obtener y cuantificar el color predominante en las aceitunas. Esto requiere la transformación de la imagen RGB utilizando diversas funciones para transformar el formato RGB en  $L^* a^* b^*$  espacio de color CIE  $L^* a^* b^*$ , donde  $L^*$  indica la luminosidad,  $a^*$  significa al eje de color de verde a rojo y  $b^*$  se refiere a la azul -amarillo.

Por un lado, en los métodos supervisados una región que contiene el color de interés (marcadores de color) se selecciona y se promedian. Esta clasificación se realiza mediante el método de los vecinos más próximos k-(KNN) [10], en la que cada pixel se clasifica en el mismo grupo que los marcadores de color con una intensidad similar. Los vecinos más próximos k asignan un valor de "a" y "b" para cada marcador y, por lo tanto, es posible clasificar cada píxel de la imagen con el fin de calcular la distancia euclídea entre los píxeles y marcadores de color.

Por otro lado, otras técnicas de segmentación de color utilizan algoritmos de clustering, que básicamente realizan las mismas funciones que los métodos de clasificación, pero sin utilizar datos de entrenamiento (marcadores) [11,12]. La idea básica es asumir que píxeles de la imagen son puntos en un espacio tridimensional (RGB). Por lo tanto, los puntos de una región de color similar se agrupan. Las técnicas de conglomerados nos permiten obtener un punto representativo de cada grupo, con base a las medidas de similitud o distancia entre estos puntos. En este algoritmo, un punto  $x_i$  se asigna a un grupo R, cuyo centroide C está más cerca de él. La distancia euclidiana se utiliza generalmente como una medida de la similitud entre el punto y el centroide. Los resultados de esta clasificación se expresa normalmente en términos del porcentaje total de píxeles de cada color en cada objeto. Gracias a estos

resultados, los frutos pueden ser clasificados de acuerdo con su nivel de madurez y cantidad de suciedad.

### Segmentación de niveles de grises

Usando las imágenes infrarrojas es posible distinguir las aceitunas aparentemente dañadas que parecen sanas en el espectro visible. La diferencia entre la piel sana y dañada puede verse claramente en las bandas NIR, que proporcionan un buen contraste entre la piel dañada y sana.

Los píxeles encontrados en un área defectuosa, pertenecen a un rango de altas intensidades de gris, y cada píxel de la fruta se agrupa ya sea defectuoso o no defectuoso, lo que indica que los píxeles con valores más altos de gris representan el área de defectos del fruto, y los valores más bajos representan las zonas sanas del fruto.

Esto nos permite medir el área del defecto en cada aceituna. Después de una evaluación de la zona de defecto, los resultados se recogen en porcentajes como la suma de todos los píxeles de las áreas ocupadas por los defectos y la suma de todos los píxeles de las áreas ocupadas por las áreas sanas de las aceitunas.



## Bibliografía

- [1] Richard, G. Brereton. Chemometrics, Data Analysis for the Laboratory and Chemical Plant. John Wiley & Sons. Ltd (2002)
- [2] <http://www.camo.com/Software/Overview/unscrambler.htm>.
- [3] Massart, D.L., Vandeginste, B.G.M., Deming, S.N., Michotte, Y., Kaufman, L., Data Handling in Science and Technology, volume 2. Chemometrics: a textbook. Elsevier Science Publishers, Amsterdam, The Netherlands (1988)
- [4] Roger, JM. 2005. Développements chimiométriques pour améliorer la robustesse des mesures spectrométriques appliquées aux agro-procédés. Mémoire d'habilitation. Academie de Montpellier, Université de Montpellier II.
- [5] Wold, S. Pattern Recognition by means of disjoint principal components models. Pattern Recognition 8 (3), 127 (1978)
- [6] M. Bylesjö, M. Rantalainen, O. Cloarec, J.K. Nicholson, E. Holmes, J. Trygg. Journal of Chemometrics. 20 (2006) 8-10 341-351
- [7] T. Yu, The application of chemometrics derived pattern recognition methods to futures market analysis. Chemistry Dissertations. 11 (2009)
- [8] S.M. Wold. SIMCA: a method for analyzing chemical data in terms of similarity and analogy. In Chemometrics: Theory and Application; Kowalski, B. R., Ed.; ACS Symposium Series 52; American Chemical Society: Washington, DC, pp 243-282 (1977)
- [9] Garcia JM, Seller S, Perez-Camino MC. J Agric Food Chem 44:3516–3520 (1996)
- [10] Richards, J. A.. Remote Sensing Digital Image Analysis. Ed. Springer-Verlag, 1999.
- [11] Thomas, S.W., 1991. Efficient inverse color map computation. In: Arvo, J. (Ed.),

Graphics Gems II. Academic Press, Boston, pp. 116–125.

[12] Floyd, R. W., & Steinberg, L. (1975). An Adaptive Algorithm for Spatial Grey Scale. In:

International Symposium Digest of Technical Papers, Society for Information Displays, 36

[12] Russ, J.C. The image processing Handbook. CRC Press LLC 3rd Edition.

# Capítulo V. Resultados



UNIVERSIDAD  
DE  
CÓRDOBA

## V.1 Introducción

Las técnicas de estudio de propiedades ópticas aplicadas a la evaluación de la calidad de frutos de aceituna y aceite de oliva ofrecen amplias posibilidades en su utilización tanto en los trabajos experimentales como en su empleo en los procesos de caracterización y clasificación durante el proceso de producción.

La metodología de aplicación conduce a resultados que revelan amplias posibilidades en el desarrollo de estas tecnologías avanzadas, sin embargo, su progreso implica la necesidad de actuar a diferentes niveles. En los trabajos realizados en esta tesis doctoral ha sido necesario recurrir a la utilización de equipamiento muy diverso y en consecuencia métodos de trabajo variados. Se han utilizado equipos espectroscópicos y de análisis de imagen en laboratorio e instalaciones de trabajo en la línea de procesado.

Por este motivo, este capítulo reúne los principales resultados obtenidos en los diferentes trabajos experimentales realizados para alcanzar los objetivos establecidos en esta Tesis, agrupados en siete trabajos monográficos clasificados en dos grandes grupos según el método óptico utilizado, sistemas espectroscópicos o visión artificial, y cuyo tema común y principal son “las propiedades ópticas”.

Los datos proporcionados en cada uno de los trabajos durante el desarrollo del trabajo experimental, constituyen la base de partida de los resultados y conclusiones; A continuación en el presente capítulo se muestran los principales resultados obtenidos para cada uno de los trabajos desarrollados:

## V.2 Sistemas espectroscópicos

### V.2.1. Aplicación de Espectroscopia Raman de baja resolución para el análisis de aceites de oliva oxidados

La tecnología FT-Raman ha sido usado para el estudio de la oxidación del aceite de oliva, mediante esta técnica se han obtenido bandas bien definidas que se pueden asociar a cambios producidos en el aceite durante la oxidación. Parte de estas bandas se encuentran en la región por debajo de 1800  $\text{cm}^{-1}$ , cerca de 1267  $\text{cm}^{-1}$  (doble enlace simétrico en roca  $\delta$  cis [= CH]), 1302  $\text{cm}^{-1}$  (en la fase de metileno giro), 1442  $\text{cm}^{-1}$  (modo de tijera de metileno  $\delta$  [CH<sub>2</sub>]), 1655  $\text{cm}^{-1}$  (doble enlace estiramiento cis  $\nu$  (C = C) y 1747  $\text{cm}^{-1}$  (estiramiento éster  $\nu$  (C = O)). [1]

En el espectro obtenido del instrumento Raman portátil hay un área importante por debajo de 800  $\text{cm}^{-1}$  que no se observa en los espectros FT-Raman y que hasta ahora no se ha descrito en la literatura disponible. La Figura 1 muestra un espectro típico de Raman de un aceite de oliva virgen obtenido usando el instrumento Raman portátil.

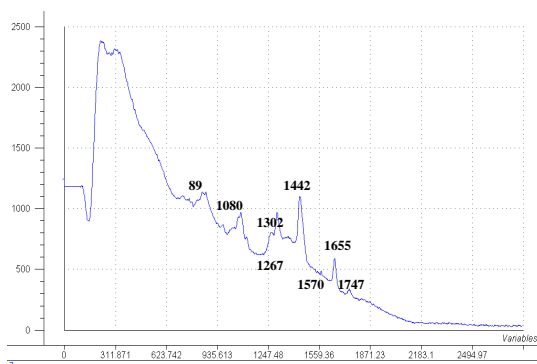


Figura.1 Espectro típico de Raman de un aceite de oliva virgen obtenido usando el instrumento Raman portátil.

Las muestras utilizadas en este estudio variaban ampliamente en su nivel de oxidación, como se muestra en la Tabla 1, que presenta los valores mínimo y

máximo, los valores medios y la desviación estándar para cada uno de los índices químicos estudiados (índice de peróxidos, determinación del coeficiente de extinción UV K<sub>232</sub> y K<sub>270</sub>).

	Max	Min	Mean	SD
IP (meq O <sub>2</sub> /kg)	74.59	3.68	22.22	10.57
K <sub>232</sub> (UV absorption)	6.52	1.30	3.18	0.74
K <sub>270</sub> (UV absorption)	1.02	0.09	0.48	0.20

Tabla 1. Desviación estándar, media y valores min-máx de los parámetros de referencia usados en el estudio.

El conjunto de datos fue inicialmente dividido de forma aleatoria en un conjunto de calibración y un conjunto de validación. El conjunto de calibración consistió en 100 espectros y se utilizó para construir el modelo, y el conjunto de validación contenía 26 espectros, que fueron utilizados para validar el modelo.

Con el fin de detectar la presencia de valores atípicos u outliers, se llevaron a cabo diferentes pruebas principalmente la aplicación de distintos tipos de pre-tratamientos y el uso de factores de PCA (análisis de componentes principales) y PLS (regresión por mínimos parciales). Como ejemplo, la Figura 2 muestra la primera variable latente vs la tercera después de centrar los datos, hay claramente algunas muestras fuera del rango percentil 95%. Luego, estas muestras (3 muestras para el conjunto de calibración y 4 muestras para el conjunto de validación) fueron eliminadas del conjunto de datos. Además, se aplicó el método de selección de variables (BVSPLS) [2] con el fin de seleccionar las regiones espectrales más importantes. Los

resultados (no mostrados) indicaron que ninguna región permitió mejorar los resultados en comparación con los modelos que utilizaban el espectro completo.

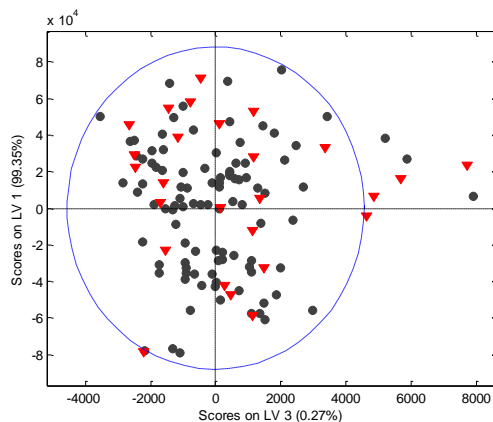


Figura 2. Primera variable latente vs la tercera del análisis PCA

Una vez que los valores anormales han sido eliminados, se construyeron los modelos de regresión para las tres propiedades estudiadas (valor de peróxidos, determinación del coeficiente de extinción UV K232 y K270) utilizando regresión por mínimos parciales (PLS). La Tabla 2 muestra un resumen de los resultados obtenidos utilizando PLS mediante el método de validación cruzada leave-one-out (LOOCV) para el conjunto de calibración (97 muestras) en términos de  $R^2$  y RMSECV, y los resultados de la predicción para el conjunto de validación (22 muestras) en términos de RMSEP y RPD.

	(LOOCV)			
	R2	RMSECV	RMSEP	RPD
IP	0.91	2.36	<b>2.57</b>	<b>4.11</b>
K <sub>232</sub>	0.88	0.36	<b>0.37</b>	<b>2.00</b>
K <sub>270</sub>	0.90	0.05	<b>0.08</b>	<b>2.50</b>

Tabla 2. Un resumen de los resultados obtenidos utilizando el método de validación cruzada leave-one-out (LOOCV)

El modelo de regresión PLS construido para la determinación del índice de peróxidos en el aceite de oliva, (Figura 3) fue construido usando siete variables latentes, obtuvo un RMSECV de 2,36 y un RMSEP de 2,57. El RPD fue 4,11, lo que indica que el modelo podría ser útil para el cribado inicial de las muestras.

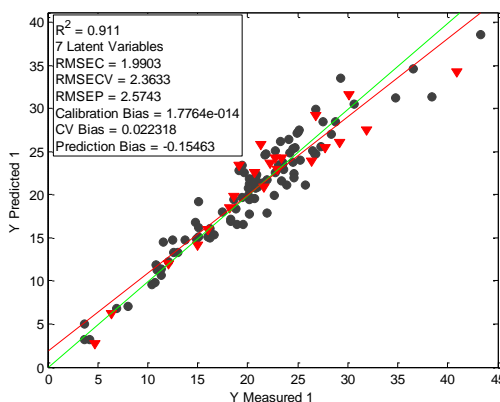


Figura 3. Diagrama de dispersión del valor real de peróxidos y el valor de peróxidos predicho por el modelo para el parámetro IP, conjunto de calibración (círculo) y conjunto de validación (triángulo).



Para la determinación del coeficiente de extinción UV K232, el mejor modelo PLS se construyó con espectros sin pre-tratamiento, utilizando seis factores PLS, y se obtuvo un valor de RMSECV de 0,36 y un RMSEP de 0,37. El modelo desarrollado obtuvo un valor de RPD de 2, por lo que podría ser utilizado para la selección inicial de muestras de aceite de oliva. La Figura 4 muestra los resultados para el conjunto de validación externa. Estos resultados indican que el modelo es consistente y da una buena predicción para un conjunto de muestras de validación.

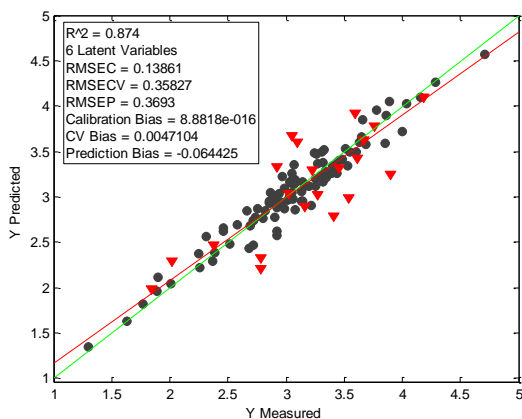


Figura 4. Diagrama de dispersión del valor real K232 y el valor K232 predicho por el modelo para el parámetro K232, conjunto de calibración (círculo) y conjunto de validación (triángulo).

Para la determinación del coeficiente de extinción UV K270, el modelo PLS se construyó utilizando siete factores, obtuvo un RMSECV de 0,07 y un RMSEP de 0,09 (Figura 5). Como en el caso de K232, la RPD de 2,50 sugerido que este modelo podría ser utilizado como un método de cribado. Estos resultados muestran que el modelo era consistente y dio la predicción correcta para un conjunto de muestras de validación. Es más estable que los

otros modelos, como se refleja en el RMSECV inferior, y tuvo una capacidad predictiva ligeramente mejor que los demás parámetros [3].

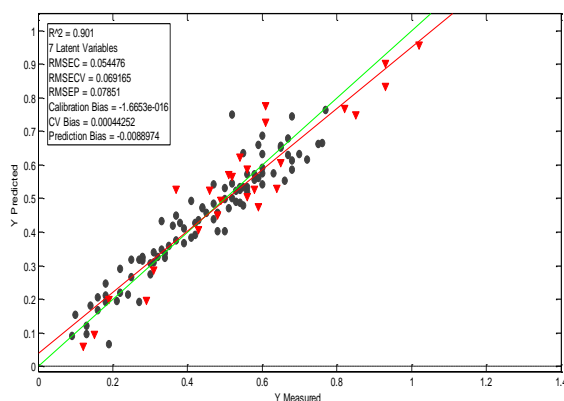


Figura 5. Diagrama de dispersión del valor real K270 y el valor K270 predicho por el modelo para el parámetro k270, conjunto de calibración (círculo) y conjunto de validación (triángulo).

Algunas muestras consistían en mezclas con diferentes concentraciones de las dos variedades de aceite picual y arbequina, por lo que se decidió aplicar un procedimiento diferente, el método de validación cruzada por clases (LMOCV). Cada mezcla fue considerada como una clase y se utilizó para realizar la validación cruzada. El procedimiento consistió en utilizar cada clase como un conjunto de validación externo. Un modelo de calibración con tres clases se desarrolló y se validó por la cuarta clase, este proceso se repitió cuatro veces, cada vez utilizando una clase para la validación y, finalmente, se obtuvo un modelo medio para el modelo de calibración.

Los resultados obtenidos (no se muestran) indicaron que los modelos utilizados para la validación con un grupo independiente de muestras dio valores de predicción y RSMECV similares a los obtenidos utilizando el

procedimiento LOOCV. Entonces, ya que este procedimiento es apto para modelos más sólidos, se decidió mantener estos modelos.

En los laboratorios, los análisis de grandes números de muestras de aceite se realizan generalmente a diario a fin de obtener información sobre los parámetros de calidad principales de aceite de oliva. Por tanto, es importante disponer de métodos analíticos para la evaluación rápida y económica de las muestras de aceite, que es esencial para la correcta clasificación de los aceites durante la preparación y en su posterior venta. Esto nos llevó a explorar la posibilidad de la tecnología de espectroscopia de dispersión Raman para analizar los diferentes parámetros de calidad de aceite de oliva virgen. En este estudio, se propuso un método eficaz para la determinación de estos tres parámetros de calidad (valor de peróxido, K232 absorción UV y la absorción UV K270) mediante la utilización de un espectrómetro Raman portátil y técnicas quimiométricas.

Este trabajo demostró la utilidad de la tecnología portátil, en este caso un espectroscopio Raman de baja resolución acoplado a un ordenador portátil para la determinación de la degradación oxidativa de los lípidos en aceite de oliva virgen.

Los resultados obtenidos con los modelos construidos PLS se relacionaron correctamente con los obtenidos en el laboratorio. Los mejores modelos fueron seleccionados sobre la base del error estándar de predicción y los valores de desviación residual predictiva (RPD). En todos los casos, los modelos obtuvieron un valor de RPD inferior a 5, que puede considerarse aceptable para el cribado y para generar datos de clasificación. A partir de estos resultados, sería posible utilizar este método para supervisar en línea el control de calidad en el proceso de elaboración del aceite de oliva. El principal inconveniente de esta técnica es que es indirecta, que requiere el uso del análisis de referencia y metodologías oficiales como un primer paso con el fin

de calibrar el instrumento. Sin embargo, hay varias ventajas claras de esta metodología: es más rápida que el método oficial, las muestras no necesitan ninguna preparación, puede ser usado en línea (en una fábrica o directamente en el campo), no hay necesidad de contacto con las muestras, y que hace que sea posible llevar a cabo un verdadero análisis a tiempo real.

En resumen, el uso de la espectroscopia Raman portátil, junto con algunas herramientas quimiométricas para analizar el estado de oxidación de aceite de oliva, es factible para el análisis en el momento de la preparación y como una técnica de cribado para el control de calidad en relación con la rastreabilidad y de almacenamiento.

### **V.2.2. Evaluación de la calidad global del aceite de oliva mediante Espectroscopia de Fluorescencia**

La caracterización del aceite de oliva se basa generalmente en los parámetros químicos (acidez; composición de ácidos grasos; índice de peróxidos (PV); absorbancia en el ultravioleta; trinoleína, y el contenido de esteroides [4-6], y el análisis sensorial [7]). La determinación de los parámetros químicos estudiados se realizó en el laboratorio, utilizando métodos analíticos oficiales [Reglamento (CEE) 2568/91] y que incluye cuatro parámetros de calidad (acidez, PV, K232 y K270). La Tabla 1 muestra los parámetros analizados, mostrando los valores máximo, mínimo y promedio obtenidos.

	Mean	Max	Min
PV	22.3	105.5	1.5
acidity	2.4	16.5	0.1
K232	2.4	3.18	1.4
K270	0.3	0.84	0.1

Tabla 1. Valores medios y mínimos-máximo de los valores de referencia utilizados en el estudio

La matriz de fluorescencia de excitación-emisión (EEM) se midió directamente a partir de muestras enteras, sin ningún tratamiento previo. Todas las muestras se midieron por duplicado, y se utilizó el valor promedio de cada muestra.

Durante la primera etapa del análisis de datos multivariantes, se realizó un primer pre-tratamiento de las señales medidas. El pre-tratamiento reduce o elimina las fuentes de variabilidad de naturaleza aleatoria en las señales, como el ruido. Diversos tratamientos fueron probados pero ninguno mejoró los resultados. La calibración fue dirigida a encontrar una relación lineal entre las variables de un modelo o un conjunto de datos experimentales.

Los estudios iniciales tuvieron como objetivo encontrar la relación entre las medidas de fluorescencia EEM y la acidez. El primer paso fue comprobar que no había ninguna correlación entre los parámetros de oxidación y acidez, ya que esto podría conducir a una medición errónea. La Figura 1 muestra la correlación entre el PV y la acidez, lo que confirma que no había relación entre estas variables.

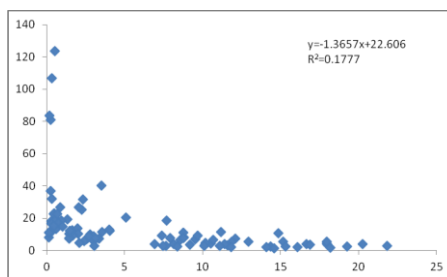


Figura. 1 Correlación entre el índice de peróxido (PV) y la acidez

La medida EEM de las muestras de aceite se utilizó para llevar a cabo una multi-regresión PLS (N-PLS) para correlacionar los valores de las variables proyectadas [8]. Este método, modela tanto las variables independientes (X) y la variable dependiente (Y) al mismo tiempo para encontrar las variables latentes en X que mejor predicen las variables latentes en Y. El número óptimo de factores se calculó utilizando el método de validación cruzada leave-one-out (LOOCV), siendo siete el número de factores utilizados siete, con un error para la validación cruzada (RMSCEV) de 1,77 y un coeficiente de correlación, obtenido en la etapa de validación, de 0,931 para el parámetro de acidez.

La etapa siguiente consistió en construir un modelo de calibración para estos parámetros utilizando una longitud de onda de excitación única con el fin de acelerar el análisis y evitar la necesidad de medir EEM a diferentes longitudes de onda. Un modelo de PLS [9] se aplicó a cada longitud de onda de excitación y los valores de error se compararon con el fin de seleccionar el valor mínimo. La longitud de onda se seleccionó, teniendo en cuenta los mejores valores de correlación con los diferentes parámetros y valores más bajos de error, obteniéndose estos valores para la longitud de onda de 380 nm. Los modelos PLS por lo tanto, se construyeron con las medidas de emisión de fluorescencia a 380 nm y una longitud de onda de excitación de

400-800 nm, a fin de verificar que la realización de una medición en esta longitud de onda obtiene la misma información que realizando una medida EEM completa.

El valor de RMSECV para todos los parámetros para cada longitud de onda de excitación se da en la figura 2, que muestra que a 380 nm se alcanza los valores más bajos para este error.

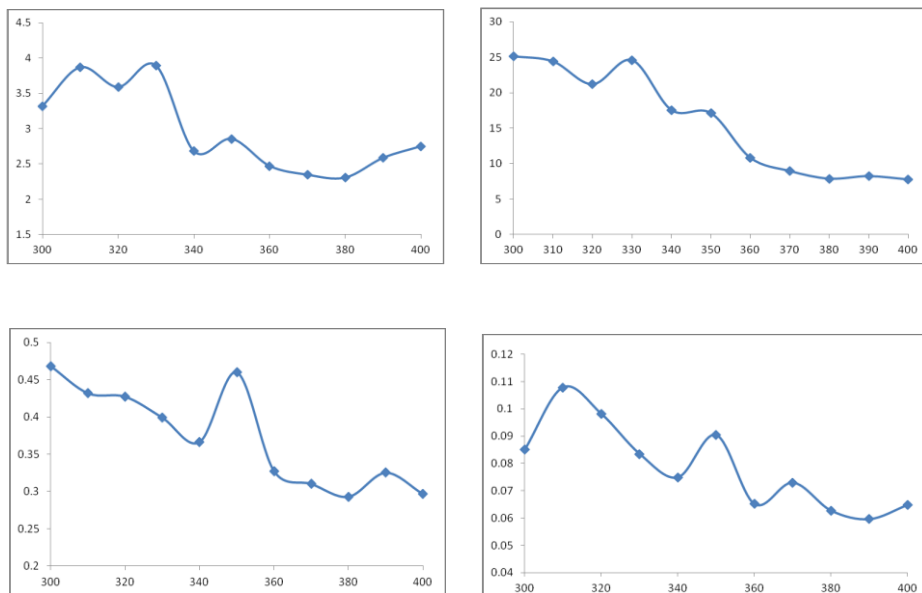


Figura 2. Error para la validación cruzada (RMSECV) para el índice de peróxidos (PV), acidez, K232 y K270, respectivamente, para cada longitud de onda de excitación

El número de factores usados en el método PLS se basa en el método de validación cruzada leave-one-out, el mismo procedimiento que para el método de N-PLS (siete); el RMSECV para la acidez fue 2,32 y el coeficiente de correlación obtenido en la etapa de validación fue 0,892. Estos resultados están de acuerdo con los obtenidos por Poulli et al. (2004) [10], quien encontró información en el espectro de fluorescencia en la región de 429-545 nm relacionada con el ácido oleico, pero no encontró un buen modelo predictivo en su caso.

El siguiente paso fue determinar la relación entre los espectros de fluorescencia y otros parámetros de calidad, tales como los relacionados con el estado de oxidación del aceite. El índice de peróxidos indican la presencia de productos de oxidación primarios (es decir, hidroperóxidos conjugados) y K232 y K270 indica la presencia de productos de oxidación secundaria (dienos-y-trienos conjugados).

Se estudió la relación entre los EEM de fluorescencia de los aceites y estos parámetros. Los altos valores de los parámetros indicaron que las muestras estaban degradadas, lo cual fue confirmado por sus altas intensidades de fluorescencia.

El trabajo previo ha demostrado que las muestras con menor PV (bajo nivel de oxidación primaria) muestran una fluorescencia suave alrededor de 415-600 nm. En contraste, la muestra con mayor PV (grandes cantidades de productos de oxidación primarios) tiene su fluorescencia máxima alrededor de esta zona.

Las muestras con el nivel más bajo de oxidación no mostraron fluorescencia alrededor de las longitudes de onda de 415-600 nm de emisión. La muestra con la oxidación alto, sin embargo, tuvo su máximo de fluorescencia alrededor de esta longitud de onda de emisión (Figura 3).



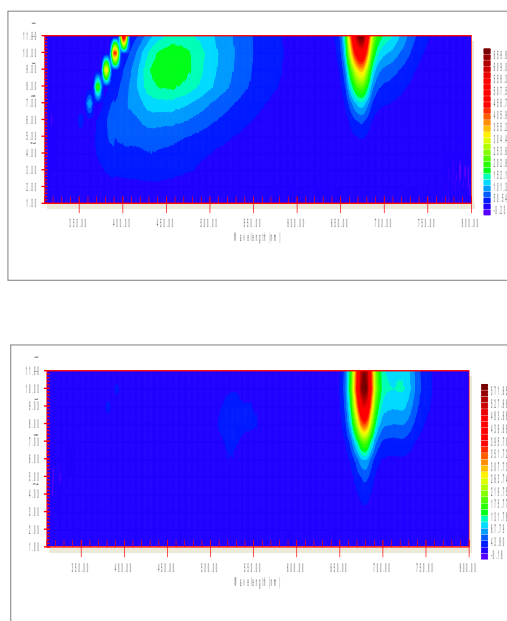


Figura 3. EEM entre 300-400 nm de excitación y 400-800 nm de emisión, de un aceite oxidado y un aceite fresco

El método de referencia para el análisis de los peróxidos en el laboratorio tiene un error en el análisis, ya que depende de muchos de varios factores externos. Además, el PV es muy variable debido a que los compuestos peroxídicos se descomponen rápidamente en otros compuestos de oxidación por el contacto con la luz y las altas temperaturas, y esto podría explicar los altos valores en el error de calibración para este parámetro.

Los resultados del análisis PLS para predecir los valores de PV calculado por medio del método de validación cruzada, obtuvo un RMSECV de 6,74 y un coeficiente de correlación en la validación de 0,967.

Un procedimiento similar se realiza para correlacionar las medidas de de fluorescencia con los parámetros K232 y K270. Durante la oxidación del aceite, altos niveles de peróxidos produce la conversión de los compuestos primarios de oxidación en productos de oxidación secundarios, que a menudo

dan mayor K232 y K270 y valores de los espectros de fluorescencia con picos en la región de 415-600 nm. Los valores altos de K270 en los aceites es mejor capturado por la fluorescencia. Los aceites con el valores altos de K270 tienen un pico ancho entre 415 y 600 nm, con un máximo de fluorescencia a 470 nm  $\lambda_{em}$  [11].

El modelo N-PLS para K232 y K270 obtuvo valores de RMSECV de 0,37 y 0,06, y coeficientes de correlación en la etapa de validación de 0,908 y 0,958, respectivamente.

Los resultados obtenidos usando el modelo de PLS para estos parámetros fueron valores de RMSECV de 8,17, 0,32 y 0,06, y los coeficientes de correlación en la etapa de validación de 0,965, 0,881 y 0,939, respectivamente.

Los resultados obtenidos con todos los métodos se resumen en la Tabla 2. Una comparación de los dos métodos muestra que los resultados son similares y que es posible utilizar la medida de la fluorescencia a una longitud de onda (380 nm) para obtener suficiente información para proporcionar una visión general de los parámetros de calidad del aceite de oliva.

	N-PLS		PLS		External Validation	
	R2	Error	R2	Error	R2	Error
Acidity	0.931	1.77	0.892	2.32	0.679	0.88
PV	0.967	6.74	0.965	8.17	0.902	5.81
K232	0.908	0.37	0.881	0.32	0.907	0.28
K270	0.978	0.06	0.939	0.06	0.924	0.08

Tabla 2. Resultados obtenidos utilizando los métodos N-PLS PLS y validación externa para cada uno de los parámetros estudiados

A fin de obtener resultados más robustos, el conjunto de datos se dividió aleatoriamente en un conjunto de calibración y un conjunto de validación. El conjunto de calibración, contiene 60 muestras, se utilizó para construir el modelo y el conjunto de validación, con 30 muestras, se utilizó para validarlo. El número de factores usados para el método PLS en el procedimiento de validación externa era de seis. Los resultados obtenidos de validación externa se dan en la Figura 4, donde los valores de error de predicción obtenidos fueron 0,88, 5,81, 0,28 y 0,08, y los coeficientes de correlación eran 0,679, 0,902, 0,907 y 0,924, respectivamente.

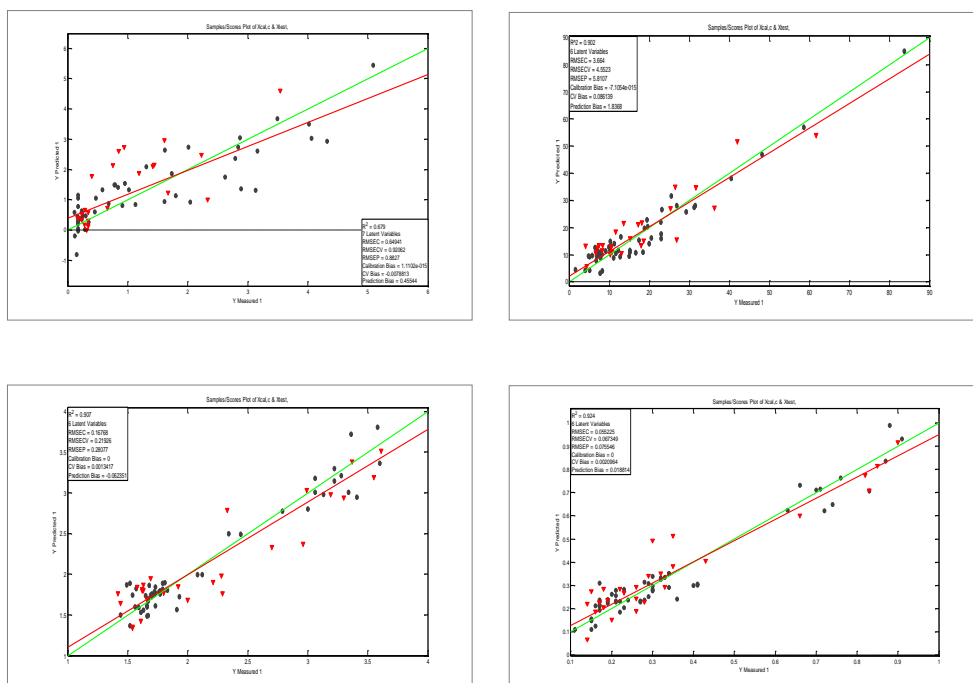


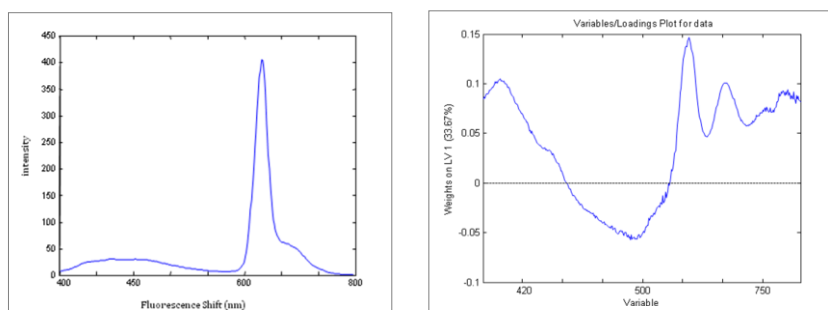
Figura. 4 Valores de referencia frente a los valores predichos en el conjunto de validación independiente. Acidez, índice de peróxidos, K232 y K270 fueron predichas por PLS validación externa

Los resultados obtenidos indican que en los modelos construidos para su validación, mediante un grupo independiente de muestras, se obtuvieron valores de predicción y RSMECV similares a los obtenidos mediante el procedimiento de validación cruzada, lo que indica que estos modelos podrían ser utilizados para llevar a cabo una evaluación global de la calidad del aceite.

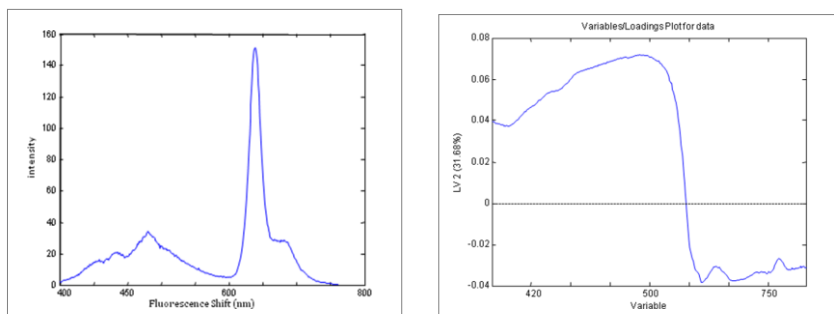
Con el fin de analizar el espectro de fluorescencia con más detalle, las diferentes áreas del espectro fueron estudiadas y relacionadas con cargas y puntuaciones de las variables (loadings and scores). Scores y loadings se utilizan para visualizar las relaciones entre los parámetros y la matriz X, respectivamente. Las relaciones entre los espectros y los parámetros fueron investigados mediante el uso de los scores y loadings.

La Figura 5 muestra el espectro típico de aceites de distinta calidad (diferentes niveles de oxidación y acidez) y su diagrama de cargas correspondiente, por lo que se pueden observar las áreas que tienen más peso en cada uno de los parámetros

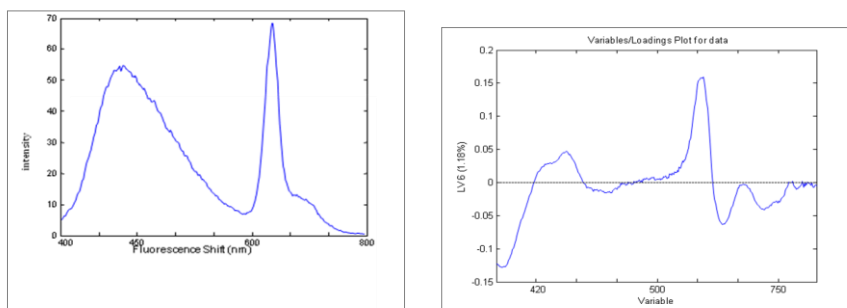
a)



b)



c)



d)

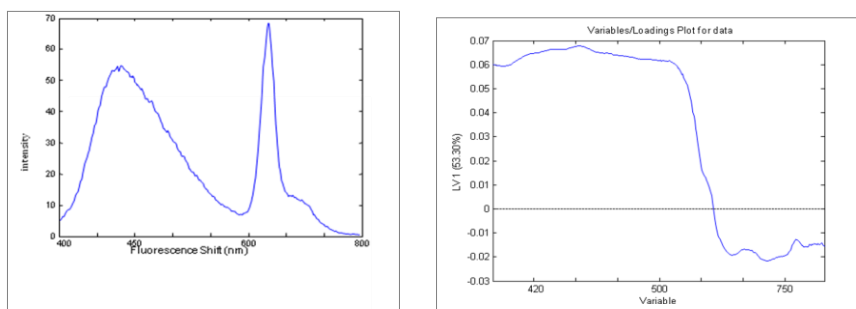


Figura 5. Espectro típico de aceites y diagramas de carga que muestran peróxidos, acidez, K232 y K270 valores: (a) alta PV y baja acidez, (b) bajo PV y acidez alta, y (c) y (d) altos K232 y K270 y baja acidez

Estos resultados están de acuerdo con los obtenidos por Guimet et al. (2005) [11] utilizando el EEM fluorescencia para construir un modelo de calibración para el parámetro índice de peróxidos con un RMSECV de 1,5 y un coeficiente de correlación en para la de validación de 0,84. Este trabajo mostró una correlación pobre, sin embargo, se obtuvieron mejores resultados entre la fluorescencia y el parámetro K232, lo que podría ser debido al mayor rango de valores y de las muestras utilizadas para desarrollar el modelo de calibración. Los valores obtenidos para el parámetro de K270 fueron un RMSECV de 0,08 y un coeficiente de correlación para la validación de 0,95, que eran muy similares a los valores obtenidos con validación externa. Este fue el mejor modelo de calibración para la predicción de este parámetro.

La ventaja de este método, por lo tanto, es que permite determinar estos parámetros con una medida simple a una longitud de onda de excitación única, sin la necesidad de escanear todo el espectro, lo que ahorra un tiempo considerable.

Estos resultados son prometedores debido a que proporcionan información útil que se puede utilizar para el control de la calidad del aceite de forma rápida y fácil y muestran el uso potencial de los instrumentos de fluorescencia para la evaluación global de la calidad del aceite de oliva y para la obtención de una visión general de la fluorescencia de diversas especies químicas desde un mismo análisis. Estos espectros se pueden medir fácil y rápidamente, sin necesidad de ningún tratamiento previo de la muestra de aceite, y posiblemente podría ser usado para identificar el aceite de oliva virgen. La característica de la fluorescencia de los aceites por lo tanto puede ser utilizado tanto para autenticar los aceites y para controlar su calidad, y podría proporcionar una alternativa interesante para otros métodos espectroscópicos, tales como espectroscopia NIR, utilizados ampliamente en

la industria del aceite, y este método podría sustituir al método de referencia actual.

### V.2.3 Sensor portátil Raman para la discriminación rápida de aceitunas según la calidad del fruto

El espectro Raman de los aceites de oliva fue estudiado con el fin de observar las bandas más características.

En los espectros de dispersión Raman, se pueden observar las mismas bandas que en el espectro FT-Raman. Las bandas más importantes del aceite de oliva virgen son en 1267 (en el plano  $\delta$  (= C-H) deformación en doble enlace cis no conjugada) 1302 (en fase movimiento de torsión de metileno), 1442 ( $\delta$  (CH<sub>2</sub>)) y la banda a 1655 ( $\nu$  (C = C)), dominan los espectros. La banda a 1747 cm<sup>-1</sup>, que se asigna a la  $\nu$  (C = O) vibración éster de, se puede encontrar tanto en el espectro del aceite de oliva como en el de las aceitunas [12, 13].



Figura 1. Espectro Raman de aceite de oliva mediante espectrometría Raman dispersiva

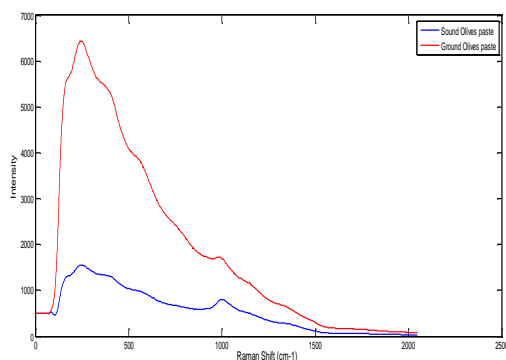


Figura 2. Espectro Raman de aceite y pasta de aceitunas de "vuelo" y de las aceitunas "árbol" mediante espectrometría Raman dispersiva

La principal diferencia en los espectros de las aceitunas y del aceite de oliva es el alto ruido de fondo de fluorescencia que caracteriza el espectro de la aceituna que se superpone con las bandas típicas del aceite. La diferencia entre los espectros de los diferentes tipos de aceitunas ("suelo" y "árbol") es también la fluorescencia, que es menor en las aceitunas de "árbol" y más alto en las de "suelo", aceitunas con algún tipo de daño (fermentación, congelación, etc). Esta fluorescencia es probablemente causada por la destrucción de las membranas celulares y las alteraciones en la materia vegetal, tales como la oxidación y reacciones enzimáticas. Las aceitunas se recogen a lo largo de toda la temporada (octubre a marzo) para que las aceitunas pertenezcan a todas las etapas posibles de su grado de maduración. De tal manera, que hay aceitunas sanas y dañadas desde el verde hasta negro, de esta forma, descartamos el color como el factor que diferencia estas clases.

Aparte de este fenómeno de fluorescencia, existen diversas variaciones de la intensidad relativa, pocos cambios en las posiciones de la banda, y pequeños cambios en los contornos espectrales entre las clases [14].



Para definir las categorías en la clasificación, se utilizó el método no supervisado PCA. Se hizo un pre-tratamiento de los datos compuesto sólo por un auto-escalado, que es un método de pre-procesamiento común que usa un centrado de la media, seguido por la división de cada columna (variable) por la desviación estándar de la columna. Este enfoque es válido para la corrección de distintas escalas y unidades de una variable, si la fuente predominante de la varianza en cada variable es la señal en lugar del ruido.

Previamente, el conjunto de datos inicialmente se dividió en un conjunto de calibración y un conjunto de validación. El conjunto de calibración se compuso de 142 espectros y se utilizó para construir el modelo, el grupo de validación consistió en 50 espectros y se utilizó para validar el modelo.

El modelo PCA se calculó con siete componentes principales, con una varianza explicada del 96%. La figura 4 muestra las puntuaciones de los dos primeros componentes principales del modelo, donde se encuentra una separación de muestras de aceitunas en diferentes grupos, el primer CP muestra la separación de muestras de acuerdo con "árbol" o "suelo", mientras que el segundo componente muestra la separación de las aceitunas de muestras de acuerdo con el período de cosecha a la que pertenecen. Estos resultados muestran que los frutos pueden ser diferenciadas en función del año de la cosecha, según la información señalada por Gurdeniz et al. [15].

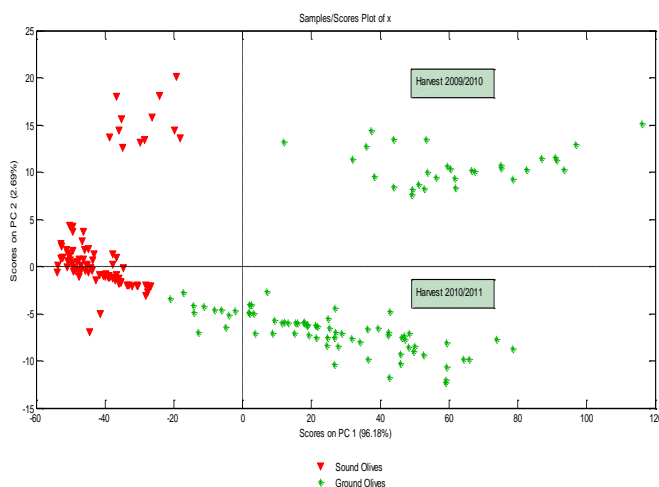


Figura 4. PCA1 vs PC2 para la calibración y validación establecidos

El análisis preliminar exploratorio de los datos mostró que la espectroscopia Raman puede ver la agrupación de las aceitunas en función de sus características espectrales y se puede utilizar para la clasificación supervisada con dos clases o categorías diferentes. Esto permitiría una separación de las aceitunas antes de su producción, en muestras de "árbol" separadas al entrar en el proceso de producción del aceite de las muestras "suelo".

El paso siguiente consistió en la clasificación multivariante, utilizando tres métodos supervisados: SIMCA, PLS-DA y KNN. Los modelos SIMCA y PLS-DA se construyeron con siete variables latentes principales para ambas clases. En el método KNN, el error de clasificación dependía del número de vecinos más cercanos elegido para el análisis ( $k$ ) [16]. El número óptimo de vecinos  $k$  fue seleccionado mediante un procedimiento de validación cruzada leave-one-out (LOOCV) [17], con valores de  $k$  entre 3, 5 y 7. Para  $k = 3$ , el reconocimiento y la capacidad de clasificación es del 100% para la

discriminación. Sin embargo, para  $K = 5$  y  $7$ , las capacidades de predicción proporcionadas con KNN fueron del 97% en ambos casos.

En la tabla 2 se presenta un resumen de los tres resultados del modelo para la calibración y validación de conjuntos en términos de capacidad de clasificación correcta (en porcentaje).

	SIMCA		PLS-DA		k-NN(3)	
	Cal Set	Val Set	Cal Set	Val Set	Cal Set	Val Set
Sound Olives	98 %	100 %	100 %	92 %	100 %	100%
Ground Olives	90 %	73 %	97 %	100 %	98 %	97 %

Tabla 2. Resultados de calibración y validación en términos de capacidad clasificación correcta en porcentaje para los métodos PLSDA, SIMCA y KNN

Los clasificadores lineales (SIMCA, PLS-DA) dieron buenos resultados para la discriminación entre las clases de aceitunas, con un bajo error de clasificación, pero KNN se encontró que fue la técnica de clasificación más eficaz.

El punto más importante es que KNN es un método no paramétrico, donde la calidad de las predicciones depende de la distribución de algunas clases. Así, el método pueden clasificar correctamente los objetos, aunque las clases no son linealmente separables o cuadráticas. El método KNN establece límites irregulares entre las clases, mientras que los métodos paramétricos como SIMCA y PLS-DA dependen de una distribución determinada por los límites estadísticos (teorema bayesiano) para establecer los límites entre las clases. En este estudio, las muestras se agruparon perfectamente y por lo tanto el

método KNN, que utiliza las distancias a los vecinos más cercanos para clasificar las muestras, fue el método más apropiado

Los resultados obtenidos con el método KNN clasificada en la validación externa correctamente todas las aceitunas de 'árbol', y sólo una muestra de aceitunas de "suelo" fue clasificada erróneamente, y los valores similares en la predicción indica que el modelo es bastante estable. Estos resultados mostraron que la técnica de clasificación más apropiada para la tarea de clasificación fue el método KNN, que tendía a producir resultados más robustos, a pesar de que se obtuvo una buena capacidad de clasificación con los modelos PLS-DA y SIMCA.

Finalmente, con los datos que teníamos de muestras de diferentes años de cosecha, y el uso de los métodos SIMCA y PLS-DA, tratamos de utilizar los datos de un año para calibrar las muestras de aceitunas, y los de otro año para hacer la validación de la muestra, pero los resultados no fueron satisfactorios (no se muestra). Las diferencias entre las distintas cosechas estaban directamente relacionadas con las fuentes de variación no incluidos en el PCA, hay grandes diferencias entre las muestras y las condiciones ambientales (temperatura, humedad, precipitación), el período de cosecha difieren para los dos años de cosechas consecutivos y la composición (perfil de ácidos grasos) también cambiaban. El origen geográfico y el año de la cosecha se sabe que tienen una influencia significativa en el perfil de ácidos grasos del aceite de oliva [18,19]. Un estudio anterior sobre Cornicabra aceite de oliva virgen reveló diferencias estadísticamente significativas en los índices de calidad, los principales ácidos grasos y esteroides composiciones en relación con los años de producción [20]. Los datos incluyen fuentes de variación que no fueron eliminados completamente por el pre-procesamiento, por lo que es necesario establecer un procedimiento para la transferencia de

modelos de calibración en los años posteriores a causa de estas variaciones entre las diferentes cosechas.

En este estudio se disponía sólo de datos de dos años de cosechas, por lo que era difícil la predicción. Serían necesarios datos de más años diferentes, a fin de obtener unas ecuaciones de predicción fiables y robustas [21, 22].

La combinación de espectroscopia Raman portátil de baja resolución y el análisis discriminante multivariado, puede ser una herramienta importante para el control de la calidad de las aceitunas antes de entrar en el proceso de producción del aceite. Esto es necesario para garantizar la buena calidad del aceite obtenido, mediante la creación de procesos de producción más controlados y ahorrando muchos gastos, evitando mezclas no deseadas.

Este método tiene muchas ventajas. Es rápido y barato, ya que no necesita preparación de la muestra. Además, no es destructivo y no utiliza productos químicos, y es por lo tanto respetuoso con el medio ambiente. Aparte de estas ventajas, esta técnica permite la posibilidad de acoplamiento de fibras ópticas, por lo que puede ser insertadas directamente en el proceso o conectadas a una célula de flujo a través del cual puede ser desviada la muestra de la línea de producción. Esto permitiría un control completo de la calidad del aceite, antes, durante y después de la transformación de las aceitunas.

### **V.3 Sistemas de Visión Artificial**

#### **V.3.1. Determinación del índice de madurez de frutos intactos de aceituna mediante análisis de imagen**

##### **V.3.1.1 Determinación del índice de madurez**

En este trabajo las muestras consistieron en conjuntos de aceitunas formados al azar y recogidos de diferentes zonas y fecha de recolección. Para cada de aceitunas, se calculó el índice de madurez (IM) de acuerdo con el procedimiento descrito en el capítulo anterior.

El procedimiento se realizó inicialmente en 40 imágenes de aceitunas separadas unas de otras, preparadas manualmente. Se aplicaron los algoritmos descritos con el fin de obtener los resultados del índice de maduración para cada fruto, obteniéndose los resultados como, el porcentaje existente de cada color en el objeto, pudiéndose clasificar así en las diferentes clases de maduración. La efectividad del método se demostró en la variedad Picual, las aceitunas de esta variedad utilizadas presentaban una gran variabilidad en el color y el tamaño.

Esta fase de prueba con imágenes de aceitunas separadas mostró que los valores proporcionados por el sistema de visión son muy similares a los proporcionados por el método tradicional basado en la determinación visual utilizado como referencia. La Tabla 2 muestra algunos resultados obtenidos por la aplicación del procedimiento propuesto en imágenes que contienen aceitunas con diferente índice de maduración. La figura 1 muestra una imagen original y la imagen después del análisis, los resultados muestran cada color del objeto con un color de píxel determinado.

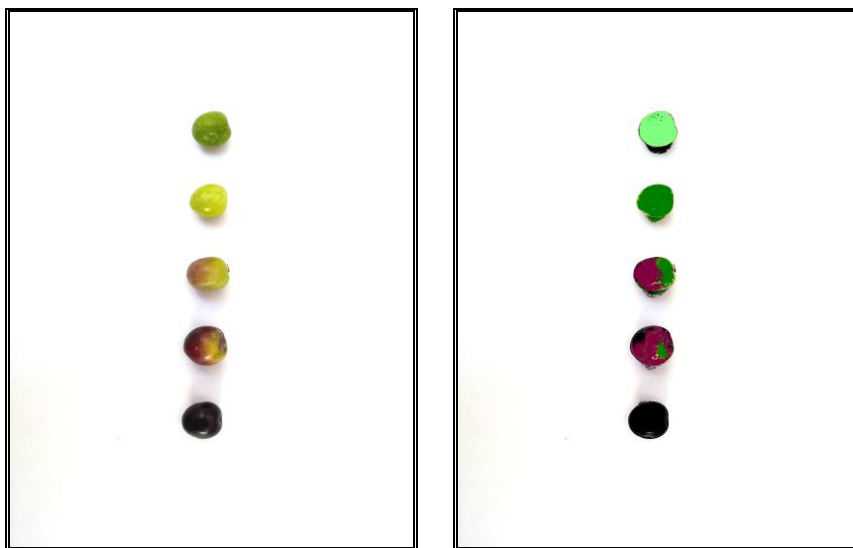


Figura 1. Imagen original y la imagen después del análisis, los resultados muestran cada color en el objeto con un color de píxel (negro =% negro; rojizo =% color marrón; verde oscuro= % verde, verde claro =% verde amarillento)

Objeto	IM				IM	IM Visual
	% Verde Intenso	% Verde-amarillo	% Rojo-marrón	% Negro	Estimado	Medido
1	100	0	0	0	0	0
2	57	43	0	0	0	1
3	33.3	66.7	0	0	1	1
4	0	100	0	0	1	1
5	1	99	0	0	1	1

6	0.6	99.4	0	0	1	1
7	2.8	97.2	0	0	1	1
8	0.7	99.3	0	0	1	1
9	0	40.4	59.6	0	3	3
10	0	67.4	32.6	0	2	2
11	0	33.6	66.4	0	3	3
12	0	51.9	48.1	0	2	2
13	0	60.6	39.4	0	2	3
14	0	49.9	50.1	0	3	3
15	0	0	100	0	3	3
16	0	31.3	68.7	0	3	3
17	0	0	0	100	4	4
18	0	0	0	100	4	4
19	0	0	0	100	4	4
20	0	0	0	100	4	4

Tabla 2. Resultados del índice de maduración estimado visualmente y calculado por el método de algunas muestras analizadas

La mayoría de los errores encontrados en el procedimiento de segmentación se deben a grupos aislados o en pequeños píxeles, principalmente ubicados en los límites de las regiones adyacentes (sombras).

Estos grupos de píxeles pueden ser fácilmente detectados y se descartan, ya que estos grupos no forman grandes regiones, sino más bien zonas



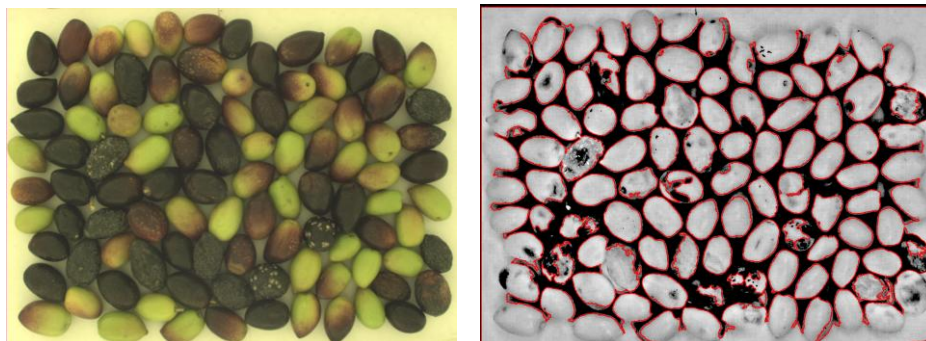
discontinuas, de modo que pueden descartarse debido a su pequeña área, utilizando el área central para aplicar la clasificación de los colores.

Los resultados de la predicción del índice de la maduración de los frutos por el método propuesto, muestra que los resultados obtenidos mediante análisis de imagen están fuertemente relacionados con IM visual medido.

Los errores de predicción son valores casi insignificantes o cero para otras muestras (resultados de la Tabla 2), estos valores son aceptables dada la subjetividad que caracteriza la evaluación del IM por el método tradicional.

Este protocolo se aplicó posteriormente a 15 imágenes complejas, con aceitunas que se superponen en diferentes capas. Alrededor de 100 aceitunas en cada imagen se analizaron y un promedio de 92% fueron detectados correctamente como objetos separados y analizados por el programa. La Figura 2 muestra un ejemplo de las imágenes analizadas.

Por lo tanto, este protocolo permite el análisis eficaz de muestras y puede mejorar sustancialmente los resultados a través de la mejora y normalización de las condiciones de iluminación en las muestras.



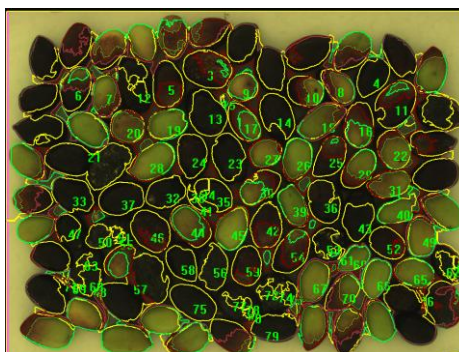


Figura 2. a) Representación de la muestra en el espacio RGB, b) la imagen después de la segmentación y c) la imagen obtenida después de aplicar el procedimiento completo

Los resultados del análisis de estas 15 imágenes se muestran en la Figura 3, donde es posible comparar los resultados obtenidos con ambos métodos.

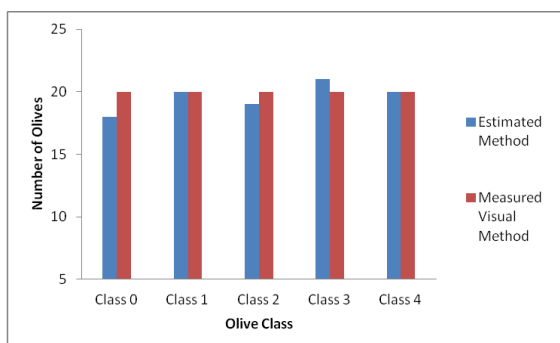


Figura 3. Representación de histograma de los resultados de una imagen analizada

Estos resultados confirman la alta correlación entre ambos métodos, la diferencia entre los valores proporcionados por el sistema de visión y el método visual clásico son muy bajos y aceptables dada la alta subjetividad del método visual utilizado como referencia.

### V.3.1.2. Estimación del tamaño y peso de las aceitunas

Esta parte del trabajo tiene por objeto dar una estimación del peso del fruto, calculando la zona ocupada por el objeto en píxeles de la imagen y

conociendo los pesos reales del fruto. Se analizaron 40 muestras de aceitunas, y se observó una correlación positiva entre el peso y el área ocupada por las aceitunas en píxeles, con un coeficiente de determinación de 0,91 (Fig. 4).

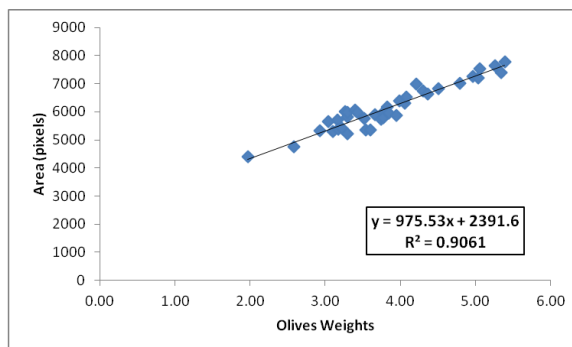


Figura 4. Correlación entre el área en píxeles ocupada y el peso real medido de las aceitunas

Estos resultados muestran la dependencia estadística entre el peso de los frutos y el área calculada en las imágenes. Los valores de error se pueden considerar mayores de lo deseable, debido a que en este análisis no se tiene en cuenta el índice de madurez de las muestras. El peso de las aceitunas puede depender del grado de maduración de los frutos, así como la cantidad de agua y aceite también aumenta hasta que alcanza su máximo en aceitunas más maduras. En el siguiente paso se construyó una recta de calibración para cada clase de maduración como se muestra en la Figura 5. La figura 5 muestra las líneas de calibración para diversas clases de maduración.

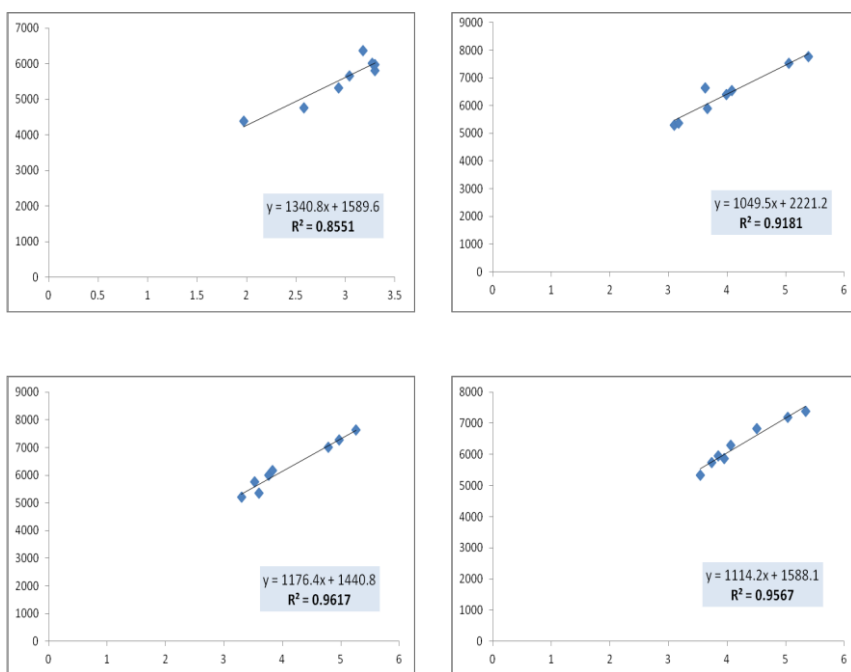


Figura 5. Correlación entre el área de las aceitunas en píxeles y se peso, para diferentes índices de madurez 1, 2, 3 y 4 respectivamente.

Con el fin de validar los datos de calibración, se midió el peso de 10 muestras externas de aceitunas. Los resultados de la validación mediante muestras muestran en la Tabla 3.

Objeto	Peso Real	Peso Estimado Calibración Global	Peso Estimado Calibración Parcial	Área Pixel	Error Relativo (CG)	Error Relativo (CP)
1	3.16	3.4	3.1	5717	0.08	0.02
2	3.4	3.8	3.3	6054	0.12	0.03
3	2	1.8	1.9	4111	0.10	0.05
4	3.43	3.7	3.6	5977	0.08	0.05
5	4.21	4.7	4.5	6986	0.12	0.07
6	3.22	3.1	3.4	5406	0.04	0.06
7	4.3	4.5	4.5	6740	0.05	0.05
8	3.77	3.5	3.8	5786	0.07	0.01
9	4.36	4.4	4.5	6658	0.01	0.03
10	3.3	3	3.3	5274	0.09	0.00

Tabla 3. Muestras de validación de datos y los valores de los errores relativos de ambas calibraciones

Por lo tanto, es necesario comprobar la influencia de la madurez para determinar el tamaño y el peso de los frutos y el uso de la recta de calibración apropiada para cada muestra de aceituna, este método puede ser una herramienta muy útil para evaluar la madurez del fruto de forma rápida y precisa, con un error relativo de no más de 0,07, por lo tanto, la determinación del tamaño fue resuelta correctamente. A partir de las características visuales y geométricas de los frutos se puede desarrollar un sistema de visión

automática para determinar el peso los frutos y mejorar el nivel de información a los agricultores.

Este método podría ser útil para la selección de la fecha de recogida, para rastrear cambios en el color de la fruta en el árbol hasta el valor seleccionado. Por lo tanto, el momento de la recolección de la aceituna no debe depender de una decisión subjetiva, sino de una medida uniforme, un método rápido y no destructivo.

La metodología propuesta tiene el potencial de ser utilizado para la clasificación del fruto en el control de calidad del proceso de extracción de aceite de oliva. De hecho, los resultados de este estudio sugieren la adquisición y análisis de imágenes en línea, éste será el objetivo de posteriores estudios.

El método descrito proporciona una estimación del valor del IM, este procedimiento requiere una cantidad insignificante de tiempo, es posible afirmar que sólo requiere la recogida de las muestras de aceituna de los árboles. Esto es importante para los expertos que trabajan en molinos de aceite, a la hora de realizar un seguimiento en tiempo real de los parámetros agronómicos y tecnológicos. En algunos casos, sin embargo, el propósito de los molinos de aceite es aumentar la calidad del aceite de oliva obtenido mediante la optimización de algunos parámetros tecnológicos.

El desarrollo de sistemas de clasificación no destructivos, aportan información rápida de los frutos y permiten asegurar que sus características cumplen las normas exigidas. En particular, las técnicas ópticas ofrecen posibilidades en la evaluación de características de la superficie y las cualidades internas o composición. Los sistemas utilizados para medir las propiedades ópticas, proporcionan una buena información sobre las características de cada fruto, útiles para la evaluación de los productos.

### **V.3.2. Método automático para la detección en línea de suciedad en el proceso de producción del aceite de oliva**

Con el fin de desarrollar este método, se analizaron una serie de imágenes reales de aceitunas. La cantidad de suciedad en cada imagen de aceitunas se calculó en peso y el número de partículas de acuerdo con el procedimiento descrito en la sección anterior.

El procedimiento se realizó en noventa imágenes de aceitunas, es decir, los algoritmos se han aplicado con el fin de obtener los resultados de la cantidad de suciedad, que muestra el porcentaje de hojas, piedras y aceitunas en cada imagen. La eficacia de este método se ha mostrado en variedad picual, cuyo color y tamaño presentan una gran variabilidad. Por lo tanto, este método podría ser potencialmente utilizado con otras variedades.

El contenido en hojas, piedras y aceitunas en cada imagen fue medido y, por lo tanto, este valor puede estar relacionado con el área en píxeles obtenidos por el método de visión artificial. Debido al hecho de que cada una de estas partículas de suciedad tienen una proporción área/ volumen diferente, es necesario construir un modelo de calibración para cada una de las partículas. En este sentido, se realizaron por separado análisis de imágenes con hojas, piedras y barro con el fin de obtener modelos predictivos más precisos. Todas las imágenes fueron capturadas por triplicado, con el fin de capturar la posible variabilidad dentro de cada conjunto de aceitunas.

En primer lugar, se analizaron 30 imágenes con diferentes cantidades de hojas a fin de obtener el modelo de calibración. La figura 1 muestra una imagen original RGB, así como la imagen después del análisis. El resultado indica el porcentaje total de píxeles ocupada por las hojas en toda la imagen.



Figura 1. Imagen original (a) e imagen después de la segmentación basada en el color, que muestra los resultados expresados como el área de píxeles ocupada por las hojas (b).

Los resultados de algunas imágenes analizadas y los valores reales de las hojas se midieron en términos de peso y el número de hojas en cada imagen. La Tabla 1 muestra el valor real en términos de peso y número de hojas, así como los valores obtenidos por el software después de aplicar segmentación y su desviación estándar.

Muestra	% Peso Hojas Medida Manual	% nº Hojas Medida Manual	% Medido por Método Visual
Imagen AV1	0.18	2.04	7.97±0.67
Imagen AV2	0.50	4.76	12.72±3.92
Imagen AV3	1.07	9.09	17.73±7.36
Imagen AV4	1.18	23.33	20.11±6.06
Imagen AV5	1.79	26.04	22.34±3.23
Imagen AV6	1.91	28.59	44.83±17.81
Imagen AV7	2.75	31.52	33.22±11.52
Imagen AV8	3.08	42.58	48.74±15.23
Imagen AV9	3.17	55.41	63.44±2.05

Tabla 1. Resultados de algunas imágenes obtenidas por el software después de la aplicación de algoritmos de segmentación para evaluar la cantidad de hojas en un conjunto de imágenes de aceitunas.



Gracias a estos resultados, se construyó una curva de calibración con el fin de estimar el peso o el número de hojas de acuerdo con el área en píxeles obtenidos por el software (Figura 2).

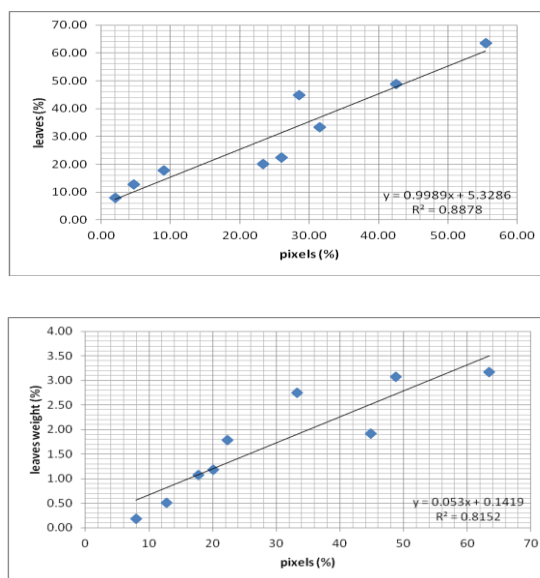


Fig. 2: Correlación entre los resultados obtenidos por ambos métodos en porcentaje de hojas (a) y peso de las hojas (b).

El mismo procedimiento se llevó a cabo para las piedras y barro, otras 30 imágenes con diferentes cantidades de piedras fueron analizadas con el fin de construir un modelo de calibración. La Figura 3 muestra una imagen original RGB, así como la imagen después del análisis. El resultado indica el porcentaje total de píxeles ocupada por las piedras en toda la imagen.



Fig. 3: Imagen original (a) e imagen después de la segmentación basada en el color, que muestra los resultados expresados en el área de píxeles ocupada por las piedras (b).

La Tabla 2 muestra el valor real en términos de peso y el número de piedras, así como los valores obtenidos por el software después de aplicar segmentación y su desviación estándar.

Sample	% Peso Piedras Medida Manual	% nº Piedras Medida Manual	% Medido por Método Visual
Imagen AV1	15.69	2.52	5.50±2.12
Imagen AV2	30.43	7.20	14.61±1.66
Imagen AV3	14.20	1.28	7.60±0.82
Imagen AV4	29.44	6.07	16.21±0.83
Imagen AV5	7.63	1.28	3.49±0.67
Imagen AV6	16.09	2.93	8.20±1.48
Imagen AV7	75.04	88.24	73.93±3.75
Imagen AV8	9.71	1.69	6.01±2.10
Imagen AV9	18.89	3.33	12.48±1.54

Tabla 2: resultados de algunas imágenes obtenidas por el software después de la aplicación de algoritmos de segmentación para evaluar la cantidad de piedras en un conjunto de imágenes de aceitunas.

Siguiendo el mismo procedimiento, se construyó una curva de calibración con el fin de estimar el peso o el número de piedras de acuerdo con el área en píxeles obtenidos por el software (Figura 4).

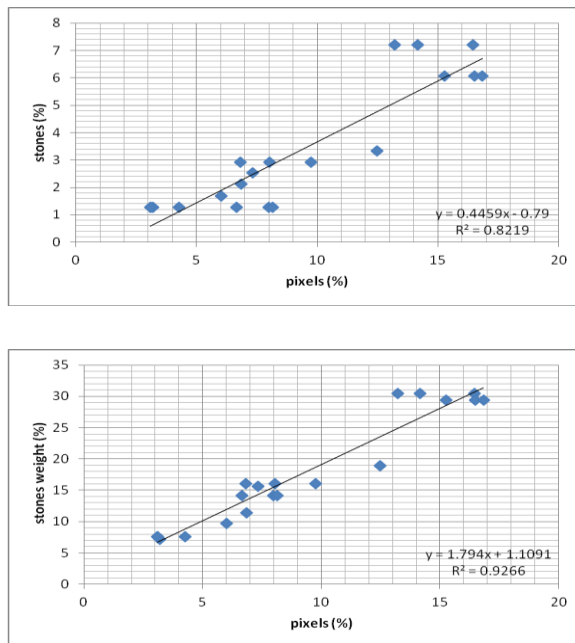


Figura 4. Correlación entre los resultados obtenidos por ambos métodos en porcentaje de piedras (a) y peso de piedras (b).

En esta parte del estudio se demostró que los valores proporcionados por el método de visión artificial son muy similares a los valores reales suministrados por el método manual utilizado como método de referencia. Los resultados visuales y numéricos indican que el método es muy eficaz para esta aplicación.

Además, estos resultados confirman la necesidad de utilizar curvas de calibración diferentes para diferentes tipos de partículas de suciedad con el fin de obtener resultados precisos. En situaciones reales, podría no ser necesario detectar hojas si el sistema de visión está implantado después del proceso de limpieza, que elimina las hojas. Sin embargo, sería muy útil utilizar este sistema de detección de suciedad, piedras y barro con el fin de detectar una gran cantidad de aceitunas que llegan a la fábrica en muy malas condiciones por haber sido recogidos del suelo.

En una segunda parte del estudio, se intentó determinar el nivel global de suciedad de las muestras, la obtención de los resultados como un porcentaje del área ocupada por la suciedad y el área ocupada por las aceitunas.

Estos algoritmos se han probado con otras 30 imágenes de aceitunas que presentan distribuciones diversas de suciedad (hojas, piedras y barro). La Tabla 4 muestra los resultados como el porcentaje de suciedad obtenido a partir del análisis de algunas imágenes.

La figura 5 muestra una imagen original, así como la imagen analizada después de aplicar el algoritmo de segmentación. Finalmente, la tabla 3 muestra los resultados del análisis de algunas imágenes



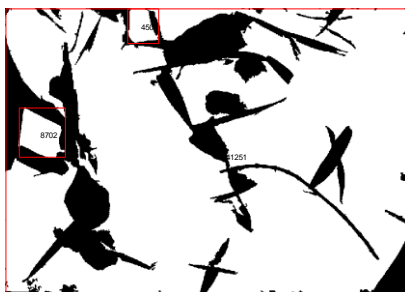


Figura 5. Imagen original (a) e imagen después de la segmentación basada en el color, que muestra los resultados expresados en el área de píxeles ocupada por la suciedad (b) y aceitunas (c).

<b>% medido por Método de Visión</b>	<b>% aceitunas</b>	<b>% suciedad</b>
<b>Img 1</b>	70	30
<b>Img 3</b>	74	26
<b>Img 2</b>	80.9	19.1
<b>Img 5</b>	87.3	12.7
<b>Img 4</b>	94.7	5.3

Tabla 3. Resultados de algunas imágenes obtenidas por el software después de la aplicación de algoritmos de segmentación para evaluar el índice de suciedad global.

Según estos resultados, este método proporcionaría una evaluación global de suciedad en las muestras al entrar en la fábrica. Como resultado, este método permite determinar el nivel de suciedad de forma rápida y automáticamente teniendo en cuenta el peso real de aceitunas que entran en la fábrica y, por lo tanto, nos permita evaluar el destino de las muestras en función de su estado sanitario.

Estos resultados muestran que a partir de las características visuales y geométricas, es posible realizar una predicción automática del porcentaje de suciedad presente en las muestras de aceitunas, el cual podría ser utilizado en el proceso de fabricación de aceite de oliva.

El presente trabajo describe un sistema basado en la segmentación de color y clasificación a través del método de vecino más cercano (KNN-k del vecino más cercano), que es capaz de estimar la cantidad de suciedad en lotes de oliva. Se puede utilizar como un control automático en un molino de aceite y puede ayudar a mejorar la información disponible acerca de las aceitunas al entrar en el proceso de elaboración.

Es necesaria una optimización y aumento de los modelos de calibración, el nuevo sistema de visión artificial puede ser la base para el desarrollo de un sistema de detección de suciedad en las aceitunas que podría ser incorporado para la detección en línea de suciedad en los lotes de aceitunas que entran en la fábrica.

En conclusión, el diseño off-line sistema de visión artificial puede ser utilizado como un modelo para el desarrollo de la detección de suciedad en línea en las fábricas reales.

### **V.3.3 Sistema de visión infrarroja para la detección automática de la calidad en frutos de aceituna**

Para poner a prueba este método, fueron tomadas 45 imágenes tanto de aceitunas sanas como de aceitunas defectuosas. La cantidad de ambas clases variaba entre las muestras, con el fin de obtener todas las combinaciones posibles de aceitunas sanas y defectuosas.

La Figura 2 muestra la imagen original RGB, la imagen IR y los resultados de imagen obtenidos después de aplicar los algoritmos para determinar los defectos en un conjunto de aceitunas. Como se puede ver en la imagen RGB no puede ser detectado visualmente el daño aparente en el fruto, mientras que en la imagen IR estos defectos se distinguen claramente.

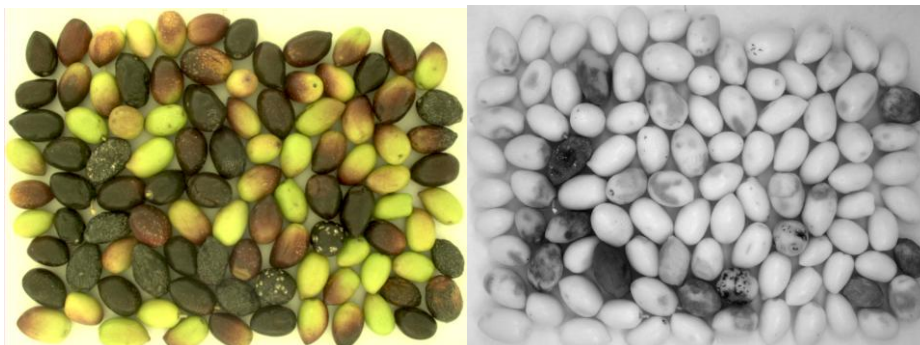


Figura 2. Imagen original RGB y la imagen IR

Los resultados de la detección de defectos (en porcentajes) para algunas de las imágenes analizadas se muestran en la Tabla 1. Los resultados se dan como porcentajes del área sana y defectuosa para cada aceituna.

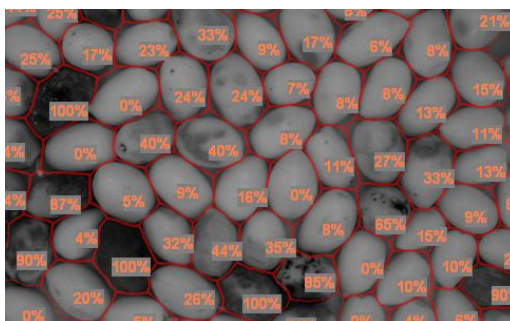


Figura 3. Imagen con resultados como los porcentajes defectuosos para cada aceituna

Objeto	% sano	% defecto
Aceituna1	78.82	21.18
Aceituna2	100.00	0.00
Aceituna3	62.56	37.44
Aceituna4	74.26	25.74
Aceituna5	45.60	54.40
Aceituna6	59.75	40.25
Aceituna7	27.03	72.97
Aceituna8	49.81	50.19
Aceituna9	18.85	81.15
Aceituna10	0.00	100.00

Tabla 1. Resultados de detección de defectos de algunas aceitunas en la imagen

Cada píxel del fruto fue clasificado como sano o defectuoso de acuerdo con los valores de sus componentes espectrales. Para cada aceituna, los resultados se expresaron como los porcentajes totales de la piel defectuosa y saludable.

Para la clasificación de las aceitunas en función de su grado de defecto, se agruparon en cinco clases. Tabla 2 define estas clases: aceitunas sanas, aceitunas con daños menores, aceitunas con daños moderados, aceitunas con daños severos, y las aceitunas defectuosas.

Categoría		
0	Aceituna Sana	100% area sana
1	Daños menores	>75% area sana
2	Daño Moderado	>50 % area defectuosa
3	Daño Severos	>75% area defectuosa
4	Aceituna Defectuosa	100 % area defectuosa

Tabla 2. Clasificación de oliva basada en el porcentaje de la piel sana y defectuosa



La Figura 3 muestra una imagen analizada, con los resultados de la clasificación. La Tabla 3 muestra los resultados de la clasificación de algunas imágenes de aceitunas, utilizando el método automático propuesto.

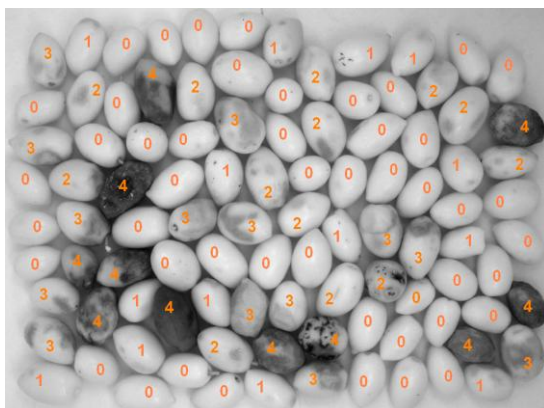


Figura 3. Resultados de clasificación para una imagen analizada

Imagen	% defectuosas	% sanas	Total nº. aceitunas	Class 0	Class 1	Class 2	Class 3	Class 4
Img 1	14.06	85.94	116	75	40	1	0	0
Img 2	34.95	65.05	80	7	35	29	0	9
Img 3	22.34	77.66	111	6	64	41	0	0
Img 4	15.88	84.12	110	14	78	18	0	0
Img 5	47.35	52.65	139	0	40	71	15	13
Img 6	68.56	31.44	107	0	3	51	30	23
Img 7	70.32	29.68	105	0	2	50	27	26
Img 8	68.42	31.58	122	0	0	80	22	20
Img 9	89.62	10.38	123	0	0	19	62	42
Img 10	70.98	29.02	115	0	0	62	34	19

Tabla 3. Resultados de la clasificación de algunas imágenes de aceitunas, utilizando el método automático propuesto

Estos resultados indicaron que las imágenes son una herramienta útil para la detección de daños en la superficie en la aceituna. Las imágenes espectrales proporcionan una herramienta valiosa para detectar defectos externos que no son claramente visibles para el ojo humano. Este sistema de visión en el espectro infrarrojo proporcionan un buen contraste entre el tejido defectuoso y sano, y son muy adecuadas para la detección de daños en los tejidos internos, tales como el daños de la piel y golpes. Este método basado un sistema de visión capaz de capturar radiación en todo el espectro infrarrojo, parece ser adecuado para la detección automática de los defectos en la aceituna. Estos resultados muestran que es posible extraer información a partir de las características visuales y el sistema de visión infrarroja y visible para desarrollar aplicaciones para la predicción automática del estado de calidad de las muestras de aceitunas en el inicio del proceso de producción de aceite de oliva. La metodología propuesta, se basa en algoritmos de detección de bordes y segmentaciones, es capaz de estimar el porcentaje de defectos en las aceitunas. Tiene potencial para la clasificación en línea de aceitunas, lo que ayudaría a mejorar el control de calidad del aceite de oliva en las fábricas.

Es necesaria una mayor optimización y calibración más precisa, el sistema de visión nuevo podría proporcionar la base para el desarrollo de un sistema que pueda ser incorporado para detección de defectos y estado de calidad en línea de las aceitunas.

Esta tecnología es una herramienta importante para la inspección automática y control de estos parámetros. El futuro de los sistemas espectrales aplicados a la inspección de alimentos es prometedor, con la industria y los consumidores cada vez más conscientes de la necesidad de garantizar la calidad y seguridad alimentaria.

### V.3.4 Sistema visible de visión para la determinación on-line de parámetros de calidad de frutos de aceituna

El objetivo de este trabajo fue desarrollar la aplicación en línea de las diferentes metodologías propuestas en los trabajos anteriores.

El procedimiento se realizó inicialmente en cincuenta imágenes capturadas directamente sobre la cinta de de entrada de las aceitunas en la fábrica. Estos algoritmos se aplican con el fin de determinar el índice de madurez y estado de calidad de los frutos. Los resultados se obtuvieron como el porcentaje de cada color y el porcentaje de defectos presentes en cada fruto. Según esta cuantificación, es posible clasificarlos de acuerdo a los diferentes grupos definidos. La Figura 2 muestra un ejemplo de imágenes analizadas en línea. Los resultados muestran la clasificación de las aceitunas para la determinación del índice de maduración basada en la segmentación por color.

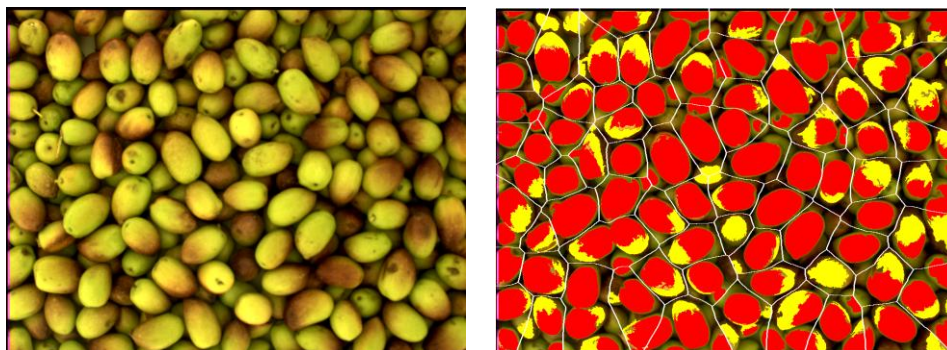


Figura 2. Imagen original tomada en línea e imagen tras el análisis (amarillo=% rojizo-marrón, rojo =% verde-amarillento)

Objec	% Verde Intenso	% Verde-amaril	% Marrón	% Negro	IM Estimada	IM Medida por Método Visual
1	0,00	29,98	36,47	33,55	3	3
2	0,00	0,00	59,66	37,86	3	4
3	0,00	79,60	0,00	0,00	1	1
4	0,00	0,00	20,23	77,91	4	4
5	0,00	0,00	56,78	42,13	3	3
6	0,00	0,00	0,00	89,61	4	4
7	0,00	0,00	0,00	94,69	4	4
8	0,00	0,00	54,26	45,73	3	3
9	0,00	39,13	23,68	37,19	2	2
10	25,31	65,80	0,00	0,00	1	1
11	0,00	0,00	70,14	24,47	1	1
12	0,00	58,45	0,00	41,55	3	3
13	0,00	21,78	41,37	36,84	2	2
14	0,00	0,00	71,88	0,00	1	1
15	23,68	42,92	0,00	56,97	4	3
16	24,05	39,13	37,19	0,00	2	2
17	0,00	67,12	0,00	0,00	1	1
18	0,00	0,00	59,66	37,86	3	4
19	0,00	0,00	35,96	64,04	4	4
20	66,40	21,54	0,00	0,00	0	0

Tabla 3 Resultados del índice de madurez estimado y medido por el método de visión de algunas muestras analizadas

En estudios anteriores se demostró que la predicción del índice de madurez del fruto con aceitunas individuales por el método de visión propuesto, está fuertemente relacionados con el IM medido mediante el método visual.

La determinación de detección de defectos se llevó a cabo mediante el mismo procedimiento que en el trabajo anterior, tal y como se describe en el capítulo de metodología. Para comprobar la eficacia del método en imágenes tomadas en línea, se analizaron un grupo de imágenes de este tipo, obteniéndose los resultados en forma de porcentaje de áreas sanas y área defectuosa en cada aceituna.

Cada píxel de la imagen en cada fruto, se clasifica en diferentes grupos: aceituna sana, o aceituna defectuosa, de acuerdo con los valores de sus componentes espectrales. Para cada aceituna, se obtuvo los resultados como el porcentaje total de defectos y el porcentaje total de área sana.

Para la clasificación de las aceitunas en función de su grado de defecto, fueron definidas cinco categorías diferentes. La tabla 2 define las categorías de la siguiente manera: aceitunas sanas, aceitunas con defectos menores, aceitunas con defectos moderados, aceitunas con defectos severos y las aceitunas defectuosas.

Categoría		
0	Aceituna Sana	100% area sana
1	Daños menores	>75% area sana
2	Daño Moderado	>50 % area defectuosa
3	Daño Severos	>75% area defectuosa
4	Aceituna Defectuosa	100 % area defectuosa

Tabla 2. Clasificación de oliva basada en el porcentaje de la piel sana y defectuosa

La Figura 3 muestra una imagen tomada en línea, con sus correspondientes resultados de clasificación. Los resultados de la clasificación de las aceitunas de varias imágenes de aceitunas analizadas se muestran en la tabla 6.



Figura 3 Resultados de la clasificación de una imagen tomada en línea

Imágenes	% Total defectuosas	% Total sanas	NºTotal Aceitunas detectadas	Clase 0	Clase 1	Clase 2	Clase 3	Clase 4
Img 1	10,00	90,00	139	90	10	0	0	0
Img 2	26,24	73,76	111	36	58	17	0	0
Img 3	28,08	71,92	125	43	55	27	0	0
Img 4	40,80	59,20	112	7	15	58	32	0
Img 5	57,31	42,69	102	21	25	22	28	6
Img 6	63,45	36,55	103	12	34	12	35	10
Img 7	64,73	35,27	105	21	12	36	24	12
Img 8	67,70	32,30	138	17	21	61	23	16
Img 9	69,33	30,67	134	7	15	68	32	12
Img 10	81,49	18,51	110	14	15	26	32	23

Tabla 6 Resultados de la clasificación de algunas imágenes de aceitunas por el método automático propuesto

De acuerdo con estos resultados, el método propuesto que consiste en la captura de una imagen en línea y su análisis, puede ser utilizado potencialmente para la clasificación de la fruta en el control de calidad en el

proceso de extracción en línea de aceite de oliva. Este proceso puede realizarse al mismo tiempo que la aceituna entra en la almazara, obteniendo la información a tiempo real.

Este sistema puede utilizarse potencialmente en línea en la clasificación de las aceitunas, lo que significa que ayudaría a mejorar el control de calidad del aceite de oliva en las fábricas.

#### **V.4 Discusión de los resultados**

La valoración conjunta de estos resultados pone de manifiesto las amplias posibilidades que estas tecnologías ópticas ofrecen en el sector del aceite de oliva y la necesidad de establecer metodologías de control y evaluación de su funcionamiento.

En el anexo 1 se incluyen todos estos trabajos completos, publicaciones originales y manuscritos enviados y en proceso de revisión, redactados íntegramente en inglés, de modo que se puede obtener una visión más global de cada uno de los trabajos. Cada trabajo incluye una introducción donde se muestra antecedentes y los objetivos específicos de cada trabajo, una sección de materiales y métodos, resultados y discusión, así como las conclusiones finales. En la introducción particular de cada uno de estos trabajos se incluye una revisión minuciosa de la literatura existente y los problemas metodológicos sobre las técnicas ópticas utilizadas, orientadas a cada caso en particular y enfocados todos ellos a: los frutos de aceituna y aceite de oliva.

Los resultados obtenidos en la realización de esta tesis doctoral han permitido avanzar en la metodología de aplicación de las propiedades ópticas al control no destructivo de estos productos y muestra las amplias posibilidades de

estos procedimientos de trabajo en el desarrollo de modelos y programas que pueden desarrollarse y aplicarse específicamente al producto en condiciones de producción y comercialización.

En las últimas décadas las grandes posibilidades de la creciente capacidad de cálculo de los equipos informáticos, unidos al avance en el desarrollo de métodos matemáticos, han enriquecido las posibilidades de la utilización de estas técnicas para la caracterización de aceitunas y aceite. Sin embargo, la diversidad de los frutos de distintas variedades y la propia heterogeneidad de cada variedad, hace necesario el estudio de sistemas concretos para cada proceso de selección.

Como resultado de los estudios realizados nos muestra que son amplias las posibilidades que se pueden plantear para progresar en estas líneas de investigación; entre estas podemos indicar:

- La variabilidad de los frutos de cada variedad, por lo que es necesario tratar de forma específica cada variedad, considerando minuciosamente sus características varietales. Para ello es necesario abordar estudios reiterados en que se consideren los factores que generan la variabilidad de las muestras. En cada uno de los trabajos se pone de manifiesto la necesidad de realizar estudios para contrastar la fiabilidad y repetibilidad, pues la eficacia de la selección puede estar afectada por numerosos factores, por lo que requiere ajustes en la configuración en función de las condiciones de trabajo, variedad, exigencias de separación, etc. Otra opción a desarrollar es el estudio de la fiabilidad de los equipos empleando muestras patrón con propiedades conocidas y estables.
  
- En el estudio de los espectros, puede ser necesario poner a punto sistemas de segregación que eliminen los casos anómalos, que contribuiría a reducir el tamaño de las muestras.



- Abordar el análisis de la calidad intrínseca de los modelos de calibración en los dispositivos, tanto para mejorar la rigurosidad de su trabajo, como para realizar su evaluación.
- Desarrollar un protocolo de ensayo orientado a las empresas que utilizan equipos ópticos similares, con el fin de que puedan comprobar el nivel de fiabilidad del software, así como profundizar en técnicas de transferencia de calibración que permitan mejorar la robustez de los modelos.

En definitiva, son propuestas que aunque deben abordarse de forma específica para cada situación, en conjunto aportaran una base para que se avance en la optimización del funcionamiento de los sistemas de propiedades ópticas.

## Bibliografía

- [31] Baeten, V. (1998) Oil and Fat Classification by FT-Raman Spectroscopy. *J. Agric. Food Chem.* 46, 2638-2646
- [2] Fernandez Pierna, J.A. A Backward Variable Selection method for PLS regression (BVSPLS) (2009). *Analytica chimica acta.* 642, n° 1-2, pp. 89-93
- [3] Sundaran, J. (2010) Determination of In-Shell Peanut Oil and Fatty Acid Composition Using Near-Infrared Reflectance Spectroscopy. *JAOCS, Journal of the American Oil Chemists&Society.* 39
- [4] Aparicio, R., Aparicio-Ruíz, R. (2000). Authentication of vegetable oils by chromatographic techniques. *J. Chromatogr. A* 881, 93-114
- [5] Armanino, C., Leardi, R., Lanteri, S., Modi, G. (1989). Chemometric analysis of Tuscan olive oils. *Chemom. Intell. Lab. Syst.* 5, 343-354
- [6] Marini, F., Balestrieri, F., Bucci, R., Magrí, A. D., A. L., Marini, D. (2004). Supervised pattern recognition to authenticate Italian extra virgin olive oil varieties. *Chemom. Intell. Lab. Syst.* 73, 85-93
- [7] Aparicio, R., Morales, M. T., Alonso, V. (1997). Authentication of European virgin olive oils by their chemical compounds, sensory attributes, and consumers' attitudes. *J. Agric. Food Chem.* 45, 1076-1083
- [8] Bro, R. Multi-way calibration. (1996). Multi-linear PLS. *J. Chemom.* 10, 47-62
- [9] Massart, D.L., Vandeginste, L.M., Buydens, M.C. (1988). *Chemometrics: A textbook.* Elsevier Science Publisher, 23, 395-397
- [10] Poulli, K., Mousdis, G.A., Georgiou, C.A. (2005). Classification of edible and lampante virgin olive oil based on synchronous fluorescence and total luminescence spectroscopy. *Analytica Chimica Acta* 542, 151–156
- [11] Guimet, F., Ferré, J., Boqué, R., Vidal, M., Garcia, J. (2005). Excitation - emission fluorescence spectroscopy combined with three-way methods of analysis as a

complementary technique for olive oil characterization. *Journal of Agricultural and Food Chemistry* 53 (24), pp. 9319-9328

[12] H. Sadeghi-Jorabchi, P.J Hendra, R.H. Wilson, P.S. Belton. Determination of the total unsaturation in oils and fats by Fourier transform Raman spectroscopy. *JAOCS* 67 (1990) 483-486

[13] V. Baeten, P. Dardenne, P. Hourant, M.T. Morales, R. Aparicio. Oil and Fat Classification by FT-Raman Spectroscopy. *J. Agric. Food Chem.* 46 (7) (1998) 2638-2646

[14] Muik, B., Lendl, B., Molina-Diaz, A., Ortega-Calderon, D., Ayora-Canada, M.J. Discrimination of olives according to fruit quality using Fourier transform Raman spectroscopy and pattern recognition techniques. *Journal of Agricultural and Food Chemistry*. 2004, 52, 20, 6055-6060

[15] G. Gurdeniz , B. Ozen , F. Tokatli. Classification of Turkish olive oils with respect to cultivar, geographic origin and harvest year, using fatty acid profile and mid-IR spectroscopy. *Eur Food Res Technol.* 227 (2008) 1275–128

[16] E. Sikorska, J. Chmielewski<sup>1</sup>, T. Górecki, I. Khmelinskii, M. Sikorski, D. De Keukeleire. Discrimination of Beer Flavours by Analysis of Volatiles Using the Mass Spectrometer as an Electronic Nose. *J. Inst. Brew.* 113(1), (2007)110–116

[17] K. Varmuza. Pattern Recognition in Chemistry. *Analytica Chimica Acta.*122 (1980) 3-15 227-240

[18] M. D`Imperio, G. Dugo, M. Alfa, L. Mannina, A.L Segre. Statistical Analysis on Sicilian olive oils. *Food Chem* .102 (2007) 956–965

- [19] E. Stefanoudaki, F. Kotsifaki, A. Koutsaftakis. Classification of the major Cretan cultivars based on their fatty composition. *J Am Oil Chem Soc* 76 (1999) 623–626
- [20] M.D. Salvador, F. Aranda, S. Gomez-Alonso, G. Fregapane. Influence of extraction system, production year and area on Cornicabra virgin olive oil: a study of five years. *Food Chem* 80 (2003) 359–366
- [21] Leon, L., Garrido-Varo, A. Downey, G. Parent and harvest year effects on near-infrared reflectance spectroscopic analysis of olive (*Olea europaea* L.) fruit traits. *Journal of Agricultural and Food Chemistry* . 2004, 52, 16, 4957-4962
- [22] R. Blanco Orús, F.J. García Ramos, B. Diezma Iglesias. Aplicabilidad de equipos portátiles de espectroscopia NIR para la estimación en parcela de azúcares en uva. *Tierras de Castilla y León: Agricultura*, 2011, pp 92-97

# Capítulo VI. Conclusiones



UNIVERSIDAD  
DE  
CÓRDOBA

Los resultados obtenidos en esta tesis han permitido mostrar las grandes ventajas de la aplicación de avanzadas técnicas ópticas y espectroscópicas para el análisis no destructivo de la aceituna y el aceite de oliva. Se muestra que el gran número de análisis realizados diariamente en las aceitunas y las muestras de aceite de oliva para obtener información sobre sus principales parámetros de calidad podría reducirse drásticamente por la utilización de los métodos propuestos en esta tesis. Estos métodos permiten la caracterización de las aceitunas y aceites de oliva mediante una información rápida para una completa trazabilidad y ofrece al consumidor mayor confianza en su autenticidad.

Con el fin de cumplir los objetivos de esta Tesis, los resultados presentados e incluidos en los trabajos publicados, permitieron extraer las siguientes conclusiones.

❖ Se ha propuesto el uso de un sensor óptico Raman portátil para determinar de forma rápida y directa el estado oxidativo de aceite de oliva virgen. Los parámetros de oxidación primaria y secundaria, como el índice de peróxidos, K232 y K270 fueron estudiados y determinados con éxito, en el momento de la preparación, mediante la aplicación de técnicas quimiométricas multivariantes.

❖ De manera similar, la espectroscopia de fluorescencia ha demostrado ser un método de elección para el análisis sencillo y rápido de los principales parámetros de calidad global de los aceites de oliva vírgenes (acidez, peróxidos, y la absorbancia UV). La espectroscopia de fluorescencia, en combinación con el análisis multivariante, se utilizó para caracterizar los aceites de oliva mediante el uso de diferentes modelos de regresión predictivos. Los resultados indicaron la posibilidad de diferenciar

y cuantificar de acuerdo con el estado de la calidad del aceite de oliva y el uso potencial de esta metodología como herramienta para un análisis global rápido y barato.

❖ Cuando se trata de la calidad de los frutos del olivo, el sistema portátil Raman ha demostrado su capacidad para clasificar las aceitunas en función de la calidad de su pasta. Los resultados proporcionaron información directa para evaluar la pasta de aceitunas en la entrada del proceso de producción de aceite de oliva. La combinación de la espectroscopia Raman portátil y el análisis discriminante multivariado proporciona información directa y puede ser una herramienta útil para el control de la calidad de las aceitunas antes de entrar en el proceso de producción del aceite.

En el ámbito de aplicación de los sistemas de visión artificial para el control y la clasificación de los frutos de aceituna de almazaras, se han desarrollado diferentes algoritmos que ofrece amplias posibilidades en la evaluación de las características de la superficie y / o de calidad. Se han creado algoritmos que permite la clasificación de las aceitunas en función de su índice de madurez mediante la detección del color de la piel del fruto en imágenes visibles e infrarrojas basados de procedimientos de segmentación. Estos algoritmos se han aplicado a las imágenes tomadas en condiciones de laboratorio y dentro de línea (off/in-line) comprobando la misma capacidad para el análisis en línea.

❖ Se ha desarrollado algoritmos que permiten una predicción rápida, automática y objetiva del índice de madurez las aceitunas. El método propuesto proporciona un cálculo más rápido del valor de índice de madurez (IM), en comparación con el método clásico utilizado por la industria de aceite de oliva que consiste en técnicas empíricas, lo que es

importante para los expertos que trabajan en molinos de aceite para realizar decisiones en tiempo real de los parámetros agronómicos y tecnológicos

❖ Desarrollo de algoritmos para la detección automática y la cuantificación de la suciedad y las impurezas en la recepción de la aceituna. Los resultados han mostrado que este algoritmo permite, a partir de las características visuales de los frutos, una predicción automática del porcentaje de suciedad presente en las aceitunas en la entrada de la fábrica antes del procesamiento de aceite de oliva. Así, es posible determinar el peso real de las aceitunas recepcionadas por diferencia de pesada.

❖ Finalmente, se han desarrollado también algoritmos para clasificar las frutas según su potencial para la obtención de aceites de buena calidad, permitiendo detectar la presencia de frutos que no reúnan las condiciones de calidad exigidas.

Los algoritmos propuestos se basan en la detección de posibles defectos en las frutas mediante el uso de un sistema de visión de infrarrojos. Los resultados han demostrado que el uso de los sistemas de visión visuales y geométricos combinados con iluminación infrarroja y visible, permite desarrollar aplicaciones para la predicción automática del estado de calidad de las aceitunas en la entrada del proceso de fabricación de aceite de oliva.

La realización de esta Tesis Doctoral ofrece una visión general para la modernización del análisis de los procesos actualmente utilizados para evaluar diferentes parámetros de calidad del aceite de oliva y de las aceitunas mediante el desarrollo de diferentes métodos no destructivos



ópticos para el análisis, y su transferencia a la industria para garantizar la calidad en el proceso de elaboración del aceite de oliva virgen.

Los estudios llevados a cabo durante esta tesis ofrecen nuevas líneas de trabajo futuro, entre ellos:

- La utilización de los frutos de variedades diferentes pueden generar problemas complejos, por lo que es necesario construir modelos de calibraciones específicos para cada tipo de fruto, considerando minuciosamente sus características varietales.
- Es necesaria la validación de los modelos de calibración utilizados en condiciones reales de trabajo, en este aspecto podría incluir la colaboración con los departamentos de I + D de las empresas.
- Trabajar en la profundización de las técnicas de transferencias entre equipos, que permiten generalizar el uso de tecnologías analíticas desarrolladas en la industria.

The results obtained in this thesis have permitted to show the great advantages of the implementation of advanced optical and spectroscopic techniques to non-destructive analysis of olive fruits and olive oil. It is shown that the large number of analyses daily performed in olives and olive oil samples to get information about their main quality parameters could be drastically reduced by the use of the methods proposed in this thesis. These methods allowed the development of novel specifications for a complete traceability and control purposes to characterize olives and olive oils and providing consumers with added confidence in their authenticity.

In order to fulfill the aims of this thesis, the results presented and included in the published papers, led to the following conclusions.

- ❖ The use of a portable Raman optical sensor has been proposed to determine the direct and rapid oxidative status of virgin olive oil. The parameters of primary and secondary oxidation such as peroxide value, K232 and K270 were studied and successfully determined, at the time of preparation, by applying multivariate chemometric techniques.
- ❖ Similarly, fluorescence spectroscopy has shown to be a method of choice for simple and rapid analysis of the main parameters of global quality (acidity, peroxide, and UV absorbance) for virgin olive oils. Fluorescence spectroscopy, in combination with multivariate analysis, was used to characterize olive oils, using different predictive regression models. The results indicated the possibility of differentiating and quantifying according to the state of the olive oil quality and then the potential use of the methodologies as tool for fast and inexpensive global analysis.
- ❖ When dealing with the quality of the olive fruits, the portable Raman system has been shown its ability to classify olives according to their paste quality. The results provided direct information to assess the olive paste at the

entrance of the process of production of olive oil. The combination of portable Raman spectroscopy and multivariate discriminant analysis provided direct information that could be important for the control of the quality of the olives prior to entering the oil production process.

In the area of application of machine vision systems for the monitoring and classification of olive mill fruits, different algorithms have been developed that offered broad possibilities in the evaluation of surface characteristics and /or quality. Algorithms have been created that allows the classification of olive fruits according to their ripeness index by detecting the skin color of the fruit in visible and infrared images based on segmentation procedures. These algorithms have been applied to images taken in at/on-line conditions checking the same capacity for on-line analysis.

❖ These algorithms allows for rapid prediction, automatic and objective olive maturity index. The proposed method provides a faster estimate of the value of maturity index (MI), compared with the classical method used by the olive oil industry consisting on empirical techniques, which is important for experts working in oil mills to perform real-time decisions.

❖ Moreover, alternative algorithms have been developed for the automatic detection and quantification of dirt and impurities in the olive fruit reception . The results have shown that this algorithm allows, from the visual characteristics of fruits, an automatic prediction of the percentage of dirt present in olives at the entrance of the factory before processing of olive oil. Thus, it is possible to determine the real weight of olive receptioned by difference of weighing.

❖ Finally, algorithms were also developed to classify fruits according to their potential for obtaining good quality oils, allowing the presence of fruit not satisfy the required quality conditions. The algorithms proposed is based on the detection of possible defects in fruits by using an infrared vision system. The results have shown that using visual and geometric vision systems

combined with infrared and visible lights, it is possible to develop applications for the automatic prediction of the health status of the olive samples at the entrance of the manufacturing process of olive oil.

The completion of this Doctoral Thesis gives an overview for the modernization of the analysis processes currently used to evaluate different quality parameters of olive oil and olive fruits by developing different non-destructive optical methods for the analysis, and their transferability to industry to ensuring quality in the process of elaboration of virgin olive oil.

The studies conducted during this thesis offered new lines of future work, among them:

- Use of fruits of different varieties can generate complex problems, so it is necessary construct calibrations models with each variety of olive fruit considering their varietal characteristics
- Is necessary the validation of calibration models used to in real working conditions, in this aspect could be include collaboration with I+D departments of companies.
- Working on deepening transfers techniques inter-equipments, that allow generalize the use of developed analytical technologies for industry

# ANEXO 1



UNIVERSIDAD  
E  
CÓRDOBA

**Application of low-resolution Raman  
spectroscopy for the analysis of oxidised olive  
oil**

**Food Control 22, 2036-2040 (2011)**



## Application of low-resolution Raman spectroscopy for the analysis of oxidized olive oil

Elena Guzmán<sup>a,\*</sup>, Vincent Baeten<sup>b</sup>, Juan Antonio Fernández Pierna<sup>b</sup>, José A. García-Mesa<sup>a</sup>

<sup>a</sup>Centro IFAPA 'Venta del Llano', Crta. Nacional Bailén-Motril Km 18.5, 23620 Mengíbar, Jaén, Spain

<sup>b</sup>Wallon Agricultural Research Centre, Valorisation of Agricultural Products Department, Food and Feed Quality Unit, Henseval Building, 24 Chaussée de Namur, 5030 Gembloux, Belgium

### ARTICLE INFO

#### Article history:

Received 28 December 2010

Received in revised form

24 May 2011

Accepted 26 May 2011

#### Keywords:

Raman spectroscopy

Olive oil

PLS

Oxidation process

Peroxide value

UV absorption values  $K_{232}$  and  $K_{270}$

### ABSTRACT

The aim of this study was to evaluate the potential of low-resolution Raman spectroscopy for monitoring the oxidation status of olive oil. Primary and secondary oxidation parameters such as peroxide value,  $K_{232}$  and  $K_{270}$  were studied. Low-resolution Raman spectra ranging from 200 to 2700  $\text{cm}^{-1}$  in a set of 126 oxidized and virgin olive oil samples were collected directly using a probe. Partial Least Squares was used to calibrate the Raman instrument for the different targeted parameters. The performance of the models was determined by using validation sets, and the best results obtained were:  $R^2 = 0.91$ , RMSEP = 2.57 for the peroxide value content;  $R^2 = 0.88$ , RMSEP = 0.37 for  $K_{232}$ ; and  $R^2 = 0.90$ , RMSEP = 0.08 for  $K_{270}$ . These results demonstrated that low-resolution Raman spectroscopy could be a relevant technique for evaluating the oxidation status of olive oils because the key oxidation parameters can be determined quickly and in a non-destructive and direct way.

© 2011 Elsevier Ltd. All rights reserved.

### 1. Introduction

The quality of olive oil can be affected by several factors, such as inherent agronomical factors, including seasonal conditions or different cultivation and ripening systems. Other factors relate mainly to technical parameters, such as storage method and duration or the technology of the process.

Oxidative rancidity, or the oxidation of oil, is the main parameter affecting its quality and takes place mainly during the production and storage of oil (Kiritsakis, Nanos, Polymenopoulos, Thomai, & Sfakiotakis, 1998). Some of the important parameters related to the oxidative deterioration of olive oil feature in the official regulations, including the peroxide value and the spectrophotometric UV absorption of olive oil at 232 nm ( $K_{232}$ ) during the initial stages of oxidation and absorption at 270 nm ( $K_{270}$ ) during secondary oxidation. Another important parameter is the anisidine index (Muik, Lendl, Molina-Díaz, Ortega-Calderón, & Ayora-Cañada, 2004), which is related to the measurement of secondary products for oxidation and produces rancidity in the oil. Peroxides are important indicators of oxidation and rancidity in foods; they are chemical compounds composed mainly of O–O bonds.

Peroxides negatively affect the nutritional value and global sensorial quality of foods (Vercellotti, St. Angelo Allen, Spanier, & Arthur, 1992). They are unstable and they are easily transformed, as a result of the secondary product accumulation of oxidation, into other products, such as conjugated aldehydes, ketones, conjugated dienes and trienes (double bonds C=C). The major secondary oxidation products formed are alcohols, aldehydes, saturated epoxy compounds and  $\alpha$ ,  $\beta$ -unsaturated epoxy compounds (Frankel, Selke, Neff, & Miyashita, 1992). These products reduce the sensory and nutritional quality of oil. A parameter such as  $K_{232}$  measures the oxidative state of oil compounds, such as conjugated dienes with an absorbance at 232 nm;  $K_{270}$  also measures conjugated dienes and trienes absorbing at 270 nm.

According to European Commission Regulation 1989/03 (EC Commission Regulation, 2003), extra virgin olive oil cannot have a peroxide value higher than 20 meq O<sub>2</sub>/kg, a UV absorption  $K_{232}$  value higher than 2.50 or a UV absorption  $K_{270}$  value no higher than 0.22.

Oils with peroxide content and UV absorption  $K_{232}$  and  $K_{270}$  values higher than these, indicate the presence of diene- and triene-conjugated systems and are then considered to be lampante oils; these oils need to be refined in order to make them acceptable for consumption. (Boskou, 1996; García, Seller, & Pérez-Camino, 1996).

The official method used for determining the peroxide value is based on the iodine index. This is obtained by titration with a sodium

\* Corresponding author.

E-mail address: [elena.guzman.jimenez@juntadeandalucia.es](mailto:elena.guzman.jimenez@juntadeandalucia.es) (E. Guzmán).

thiosulfate standard solution. The spectrophotometric absorptions at 232 nm ( $K_{232}$ ) and 270 nm ( $K_{270}$ ) are expressed as a specific extinction of a solution of oil dissolved in cyclohexane, respectively. These methods are highly empirical and their precision depends on several experimental factors, such as technician skill, light exposure and atmospheric oxygen. There are many factors that can affect the quality of olive oil, such as inherent agronomical factors or factors related to the various steps in olive oil processing, from extraction to bottling. Some of these factors could be changed by controlling the oil at the time of production, thus avoiding mixing oils of different qualities and thereby improving the quality of extra virgin oil. Oil analysis at the production stage is a laborious, costly and time-consuming procedure. There is not yet a method to directly determine the oxidative state of oil at the production stage and thus to improve oil production quality control.

Various techniques to solve this problem have been proposed in the literature. Techniques such as infrared spectroscopy have been studied and successfully used for characterizing some olive oil parameters (Armenta, Garrigues, & De la Guardia, 2007; Mailer, 2004; Sinelly, Tura, & Casiragui, 2008), for quality control and for determining the geographic origin of virgin olive oils (Bendini et al., 2007). Studies have also been done, using infrared spectroscopy, on the various parameters of oxidation of some edible oils: analyzing lipids by spectroscopy (Dobson, 1995), determining peroxide values in palm oil (Moh, 1999), monitoring the peroxide in soybean and corn oil (Yildiz, Randy, Wehling, & Susan, 2002), determining peroxide values in oxidized soybean oils (Yildiz, Randy, Wehling, & Susan, 2003), and determining the oxidation state in oxidized vegetal oils (Dreau, Dupuy, Artaud, Ollivier, & Kister, 2009). Raman spectroscopy is a promising tool in analytical chemistry because of its interesting characteristics, which include being non-destructive, fast and relatively inexpensive; it does not involve any chemical products and requires very little sample preparation. In addition, Raman spectra are the result of bands of fundamental vibrational transitions, which offer a large quantity of information about the molecular structure.

In the olive oil sector, the Raman spectroscopy technique has been used to determine unsaturation (Baeten, Hourant, Morales, & Aparicio, 1998), to detect adulteration in virgin olive oil and for authentication (Baeten & Meurens, 1996; Marigheto, Kemsley, Defernez, & Wilson, 1998). More recently, it has been used to determine the fat and moisture content in olives and for olive discrimination (Muik, Lendl, Molina-Díaz, Ortega-Calderón, & Ayora-Cañada, 2003). Work comparing Fourier Transform Infrared (FTIR) and Raman techniques for studying the oxidation of lipids in edible oils (Muik et al., 2007) and for monitoring lipid oxidation in edible oils by Raman spectroscopy (Muik, Lendl, Molina-Díaz, & Ayora-Cañada, 2005) has been also proposed.

Recent advances in Raman instrumentation have led to the development of portable Raman equipment. These portable systems, although they have a lower resolution than classic FT-Raman equipment, have contributed to increased industrial use of Raman spectroscopy, reflecting increased demand for online analysis (Xiaohua, Zhang, & Zou, 2008).

The aim of this study was to evaluate the potential of a portable Raman instrument with low resolution and chemometrics for determining quality parameters related to the oxidative state, such as the peroxide values,  $K_{232}$  and  $K_{270}$ , in virgin olive oil.

## 2. Material and methods

### 2.1. Samples and sample preparation

Four virgin olive oil samples freshly collected from the 2008–2009 harvest at the Instituto de Investigación y Formación

Agraria y Pesquera (IFAPA) center 'Venta Del Llano' (Mengíbar, Jaén) were used.

The samples consisted of two oil varieties, Picual (P) and Arbequina (A), in different proportions (P:A): 100:0, 60:40, 40:60 and 20:80. For each mixture, 32 glass containers of 30 ml of oil were produced, except for the last proportion where only 30 containers were produced. In total, 126 samples were available. These containers were subjected to a thermal oxidation process based on the Schaal Oven Storage Stability Test for olive oil (Frankel, July 1993; Malcolmson, 1992; Muik et al., 2007). This test involved placing the samples in an oven at a temperature of 100 °C for 189 h in order to extract typical oxidation compounds. Every 4 h, a container of each mixture was removed from the oven and left until it reached room temperature; the Raman spectra were then recorded.

### 2.2. Raman spectroscopy analysis

The instrument used was a portable Raman spectrometer model RH-3000 (Ocean Optics) with a low resolution; it had a 785 nm laser diode and was linked to an optical fiber that enabled measurements to be taken through a glass container. The Raman spectra were recorded from 200 to 2700  $\text{cm}^{-1}$ , with a spectral resolution of 10  $\text{cm}^{-1}$ . This measurement corresponds to the average value of the three scans performed for each sample. The laser diode of the instrument had an intensity range of 0–290 mW. Spectra were acquired with a laser intensity set at 190 mW, with an integration time of 10 s.

### 2.3. Reference chemical analysis

For all the samples, the references values for the peroxide (or index of peroxides [IP]) and the specific UV absorption at 232 nm ( $K_{232}$ ) and 270 nm ( $K_{270}$ ) were obtained following EU Regulations EEC/2568/91 and EEC/1429/92, which focus on the characteristics of olive oil and olive-residue oil and on relevant methods of analysis (EC Commission Regulation, 1991, 1992). The samples used in this study varied widely in oxidation level, as shown in Table 1, which presents the minimum and maximum values, the average values and the standard deviation for each of the chemical indices.

The peroxide value is the amount of peroxide in a sample (expressed in milliequivalents of active oxygen per kg fat) that causes the oxidation of potassium iodide. The sample, dissolved in acetic acid and chloroform, was treated with a potassium iodide solution. The liberated iodine was titrated with a sodium thiosulfate solution.

The  $K_{232}$  and  $K_{270}$  values were spectrophotometric measures for quantifying the UV absorption, at 232 nm ( $K_{232}$ ) and 270 nm ( $K_{270}$ ), of a solution of the oil in cyclohexane. Spectrophotometric analysis in UV provides information about the quality of the fat, the state of conservation of the oil and any deterioration occurring during the technological processes. High values of  $K_{232}$  and  $K_{270}$  indicate the presence of diene- and triene-conjugated systems.

### 2.4. Chemometrics

Two classic multivariate procedures were applied: the exploratory Principal Component Analysis (PCA) technique (Massart,

**Table 1**  
Standard deviation, mean and min-max values of the reference values used in the study.

	Max	Min	Mean	SD
IP (meq O <sub>2</sub> /kg)	74.59	3.68	22.22	10.57
$K_{232}$ (UV absorption)	6.52	1.30	3.18	0.74
$K_{270}$ (UV absorption)	1.02	0.09	0.48	0.20



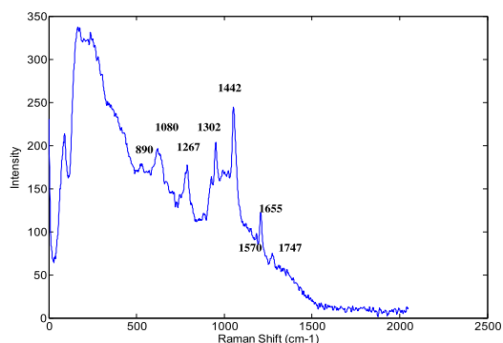


Fig. 1. Typical Raman spectrum of a virgin olive oil obtained with the portable Raman instrument.

Vandeginste, & Buydens, 1998); and the supervised Partial Least Squares (PLS) method (Wold, 1988).

PCA involves a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables, called principal components. This technique was used to make an initial exploratory analysis of the data showing evidence of the presence of some aberrant data or outliers. PLS reduces the spectral data to a series of key variables (latent variables) that take into account both the spectral values as reference values. It aims to identify the underlying factors, or linear combination of the X variables, that best model the Y-dependent variables. PLS can deal efficiently with datasets where there are many variables that are highly correlated and involve substantial random noise.

PLS models were validated by using the 'leave-one-out' cross validation (LOOCV) and 'leave-more-out' cross validation (LMOCV) approaches, and by using external validation. For LOOCV, each individual sample was used in turn for validation. For external validation, 97 samples were used to build the calibration model while another group of external samples (22) was used to validate the calibration model. For LMOCV, instead of using each sample for validation, a group of samples was used. Here, because of the different mixtures available, each mixture was taken out in order to use it for validating the model constructed with the other mixtures.

The quality of the models was assessed using various statistical parameters: the coefficient of determination ( $R_2$ ); the root mean squared error for calibration (RMSEC); the root mean squared error for cross-validation (RMSECV), which represents the model and the validation errors, respectively; and the residual predictive deviation (RPD). The model with the smallest RMSECV was assumed to have the best fit to the data, and the external validation set was then predicted. The RMSEP described the model's predictive ability based on the results of the external test set of 22 samples.

### 2.5. Software

All computations, chemometric analyses and graphics were performed using Matlab v7.4.0. (The Mathworks, Inc., Natick, MA, USA). For PCA and PLS, the PLS toolbox v. 4.11 (Eigenvector Research, Inc.) was used.

### 3. Results

Using FT-Raman, well-defined bands could be obtained that were associated with changes in oils during oxidation. Part of these

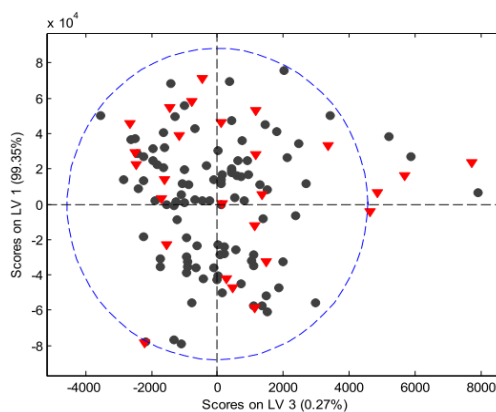


Fig. 2. LV1 vs. LV3 (calibration [circle] and validation [triangle] sets).

bands was in the region below  $1800\text{ cm}^{-1}$ , about  $1267\text{ cm}^{-1}$  (symmetric rock double bond in  $\text{cis } \delta [= \text{CH}]$ ),  $1302\text{ cm}^{-1}$  (in-phase twist methylene),  $1442\text{ cm}^{-1}$  (methylene scissoring mode of  $\delta [\text{CH}_2]$ ),  $1655\text{ cm}^{-1}$  (cis double bond stretching  $\nu (\text{C}=\text{C})$  and  $1747\text{ cm}^{-1}$  (ester stretching  $\nu (\text{C}=\text{O})$ ). (Baeten, 1998)).

In the spectrum obtained from the portable Raman instrument, there is an important area below  $800\text{ cm}^{-1}$  that does not match the FT-Raman spectra and has not yet been described in the literature.

Fig. 1 shows a typical Raman spectrum of a virgin olive oil obtained using the portable Raman instrument.

The dataset was initially split randomly into the calibration set and the test set. The calibration set consisted of 100 spectra and was used to build the model; the test set contained 26 spectra and was used to validate it.

In order to detect the presence of outliers, various tests were performed based mainly on applying different pre-treatments and using PCA and PLS factors. As an example, Fig. 2 shows the first latent variable vs the third one after centering the data; there are clearly a few samples outside the 95% percentile range. These samples (three samples for the calibration set, and four for the test set) were removed from the dataset. In addition, the Backward Variable Selection PLS (BVSPLS) method (Fernandez Pierna, 2009) was applied in order to select the most important spectral regions. The results (not shown) indicated that any particular region showed improved results compared with the models using the whole spectra.

Once the outliers had been removed, regression models for the three properties studied (peroxide value and UV extinction coefficients  $K_{232}$  and  $K_{270}$ ) were constructed using PLS. Table 2 shows a summary of the PLS results obtained using LOOCV for the calibration set (97 samples) in terms of  $R_2$  and RMSECV, and the

Table 2  
A summary of the results obtained using leave-one-out cross-validation (LOOCV) procedures.

	(LOOCV)			
	$R_2$	RMSECV	RMSEP	RPD
IP	0.91	2.36	2.57	4.11
$K_{232}$	0.88	0.36	0.37	2.00
$K_{270}$	0.90	0.05	0.08	2.50

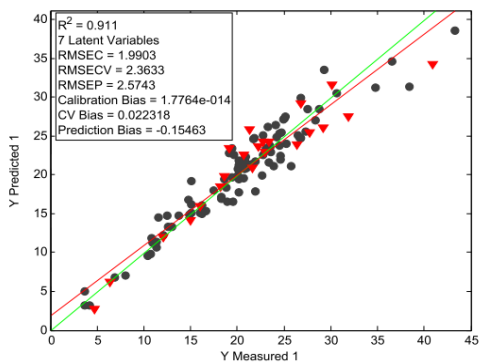


Fig. 3. Scatter plot of the actual Peroxide Value (IP) and predicted IP for calibration (circle) and validation (triangle) sets.

prediction results for the validation set (22 samples) in terms of RMSEP and RPD.

For determining the peroxide values shown, the final PLS regression model (Fig. 3) for determining the peroxide value in olive oil was built using seven latent variables; it had an RMSECV of 2.36 and an RMSEP of 2.57. The RPD was 4.11, indicating that the model could be useful for initial screening purposes.

For determining the UV extinction coefficient  $K_{232}$ , the best PLS model was built with spectra without pre-treatment, using six PLS factors; it had an RMSECV of 0.36 and an RMSEP of 0.37. The model developed gave an RPD value of 2, so it could be used for the initial screening of olive oil. Fig. 4 shows the results for the external validation set. These results indicate that the model is consistent and gives a good prediction for a validation sample set.

For determining the UV extinction coefficient  $K_{270}$ , the PLS model was built using seven factors; it had an RMSECV of 0.07 and an RMSEP of 0.09 (Fig. 5). As in the case of  $K_{232}$ , the RPD of 2.50 suggested that this model could be used as a screening method. These results show that the model was consistent and gave the correct prediction for a validation sample set. It was more stable

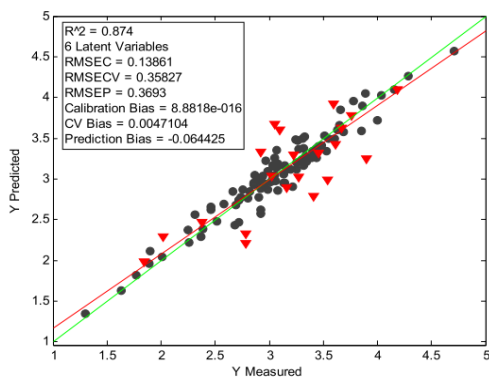


Fig. 4. Scatter plot of the actual  $K_{232}$  and predicted for parameter  $K_{232}$  for calibration (circle) and validation (triangle) sets.

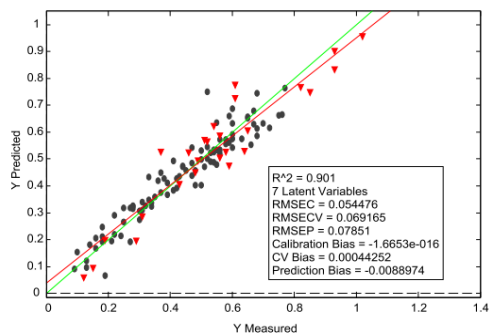


Fig. 5. Scatter plot of the actual  $K_{270}$  and predicted for parameter  $K_{270}$  for calibration (circle) and validation (triangle) sets.

than the other models, as reflected in the lower RMSECV, and it had a slightly better predictive ability than other parameters (Sundaran, 2010).

As some samples were mixtures with different concentrations of the two main varieties, it was decided to apply a different procedure, LMOCV. Each mixture was considered as a class and was used to perform CV. The LMOCV procedure consisted of using each class as an external validation set. A calibration model with three classes was developed and was validated by the fourth class; this process was repeated four times, each time using a class for the validation, and ultimately obtaining an average model for the whole calibration model.

The results obtained (not shown) indicated that the models used for the validation with an independent group of samples gave prediction and RMSECV values similar to those obtained using the LOOCV procedure. Then, because this procedure caters for more robust models, it was decided to keep these models.

#### 4. Discussion and conclusions

In laboratories, the analyses of large numbers of oil samples are usually performed daily in order to obtain information about the main quality parameters of olive oil. It is therefore important to have analytical methods available for the rapid and inexpensive assessment of oil samples, which is essential for correctly classifying oils during preparation and in their subsequent sale. This led us to explore the possibility of dispersive Raman spectroscopy technology for analyzing the various quality parameters of virgin olive oil. For this study, an effective method for determining these three quality parameters (peroxide value, UV absorption  $K_{232}$  and UV absorption  $K_{270}$ ) was proposed using a portable Raman spectrometer and some chemometric techniques.

This work demonstrated the usefulness of the portable technology, in this case a low-resolution Raman spectroscope linked to a laptop computer for determining the oxidative degradation of lipids in virgin olive oil. There is also an important area below  $800\text{ cm}^{-1}$  in the spectra obtained from the portable Raman instrument different from that obtained with the FT-Raman spectra. The information included in this area in the portable Raman has not been described yet in the literature and could be an interesting area for future research.

The results obtained with the constructed PLS models correlated well with those obtained in the laboratory. The best models were selected on the basis of the standard error of prediction and the

RPD values. In all cases, the models had an RPD value lower than 5, which can be considered acceptable for screening and for generate ranking data, especially for the IP parameter. From these results, and for performing calibration models with larger groups of samples, it would therefore be possible to use this method for monitoring online the quality control process for olive oil. The main drawback of this technique is that it is indirect; it requires using the reference and official methodologies as a first step in order to calibrate the instrument. There are several clear advantages of this methodology, however. It is faster than the official method, the samples do not need any preparation, it can be used online (in a factory or directly in the field), there is no need for contact with the samples, and it makes it possible to perform real-time analysis.

In summary, the use of portable Raman spectroscopy, combined with some chemometric tools for analyzing the oxidative state of olive oil, is feasible for analysis at the time of preparation and as a screening technique for quality control relating to traceability and storage.

## References

- Armenta, S., Garrigues, S., & De la Guardia, M. (2007). Determination of edible oil parameters by near infrared spectrometry. *Analytica Chimica Acta*, 596(2), 330–337.
- Baeten, V., & Meurens, M. (1996). Detection of virgin olive oil adulteration by Fourier transform Raman spectroscopy. *Journal of Agricultural and Food Chemistry*, 44(8), 2225–2230.
- Baeten, V., Hourant, P., Morales, T., & Aparicio, R. (1998). Oil and fat classification by FT-Raman spectroscopy. *Journal of Agricultural and Food Chemistry*, 46(7), 2638–2646.
- Baeten, V. (1998). Oil and fat classification by FT-Raman spectroscopy. *Journal of Agricultural and Food Chemistry*, 46, 2638–2646.
- Bendini, A., Cerretani, L., Di Virgilio, F., Belloni, P., Boloni-Carbognin, M., & Lercker, G. (2007). Preliminary evaluation of the application of the FTIR spectroscopy to control the geographic origin and quality of virgin olive oils. *Journal of Food Quality*, 30(4), 424437.
- Boskou, D. (1996). *Olive oil*. Chemistry and Technology. USA: AOCS Press.
- Dobson, G. (1995). Review: silver ion chromatography of lipids and fatty acids. *Journal of Chromatography Biomedical Applications*, 671(1–2), 197–222.
- Dreau, Y., Dupuy, N., Artaud, J., Ollivier, D., & Kister, J. (2009). Infrared study of aging of edible oils by oxidative spectroscopy index and MCR-ALS Chemometric method. *Talanta*, 77(5), 1748–1756.
- EC Commission Regulation No. 2568/91. (1991).
- EC Commission Regulation No. 1429/92. (1992).
- EC Commission Regulation No. 1989/03. (2003).
- Fernandez Pierna, J. A. (2009). A backward variable Selection method for PLS regression (BVSPLS). *Analytica Chimica Acta*, 642(1–2), 89–93.
- Frankel, E. N., Selke, E., Neff, E., & Miyashita, K. (1992). Autoxidation of polyunsaturated triacylglycerols. IV. Volatile decomposition products from triacylglycerols containing linoleate and linolenate. *Lipids*, 27(6), 442–446.
- Frankel, E. N. (July 1993). In search of better methods to evaluate natural antioxidants and oxidative stability in food lipids. *Trends in Food Science & Technology*, Vol 4(7), 220–225.
- García, J. M., Sella, S., & Pérez-Camino, M. C. (1996). Influence of fruit ripening on olive oil quality. *Journal of Agricultural and Food Chemistry*, 44, 3516–3520.
- Kiritsakis, A., Nanos, G. D., Polymenopoulos, Z., Thomai, T., & Sfakiotakis, E. M. (1998). Effect of fruit storage conditions on olive oil quality. *JAOCs, Journal of the American Oil Chemists' Society*, 75(6), 721–724.
- Maier, R. J. (2004). Rapid evaluation of olive oil quality by NIR reflectance spectroscopy. *Journal of the American Oil Chemists' Society*, 81(9), 823–827.
- Malcolmson, L. J. (1992). *Journal of the American Oil Chemists' Society*, Vol 71(N4), 435–440.
- Marigheto, A., Kemsley, K., Defernez, M., & Wilson, R. H. (1998). Comparison of mid-infrared and Raman spectroscopy for the authentication of edible oils. *Journal of the American Oil Chemists' Society*, 75(8), 987–992.
- Massart, D. L., Vandeginste, L. M., & Buydens, M. C. (1988). *Chemometrics: A textbook*, Vol. 23. Elsevier Science Publisher, 395–397.
- Moh, H. (1999). Determination of peroxide value in thermally oxidized crude palm oil by near infrared spectroscopy. *Journal of the American Oil Chemists' Society*, 76(1), 19–23.
- Muik, B., Lendl, B., Molina-Díaz, A., Ortega-Calderón, D., & Ayora-Cañada, M. J. (2003). Fourier transform Raman spectroscopy for the quantitative analysis of oil content and humidity in olives. *Analytica Chimica Acta*, 487, 211–220.
- Muik, B., Lendl, B., Molina-Díaz, A., Ortega-Calderón, D., & Ayora-Cañada, M. J. (2004). Discrimination of olives according to fruit quality using Fourier transform Raman spectroscopy and pattern recognition techniques. *Journal of Agricultural and Food Chemistry*, 52(20), 6055–6060.
- Muik, B., Lendl, B., Molina-Díaz, A. M., & Ayora-Cañada, M. J. (2005). Direct monitoring of lipid oxidation in edible oils by Fourier transform Raman spectroscopy. *Chemistry and Physics of Lipids*, 134, 173–182.
- Muik, B., Lendl, B., Molina-Díaz, A., Ortega-Calderón, D., Valcárcel, M., & Ayora-Cañada, M. J. (2007). Two-dimensional correlation spectroscopy and multivariate curve resolution for the study of lipid oxidation in edible oils monitored by FTIR and FT-Raman spectroscopy. *Analytica Chimica Acta*, 593, 54–67.
- Sinelly, N., Tura, D., & Casiragui, E. (2008). Characterisation and classification of Italian virgin oils by near- and mid-infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 16(3), 335–342.
- Sundaran, J. (2010). Determination of in-Shell Peanut oil and fatty acid Composition using near-infrared reflectance spectroscopy. *JAOCs, Journal of the American Oil Chemists & Society*, 39.
- Vercellotti, J. R., St. Angelo Allen, J., Spanier, & Arthur, M. (1992). *Method for determining oxidation of vegetable oils by near-infrared spectroscopy lipid oxidation in food* in ACS Symposium Series, Vol. 500, 1–11.
- Wold, S. (1988). *Chemometrics: Theory and Application*. American Chemical Society, 243.
- Xiaohua, Q., Zhang, X. F., & Zou, M. Q. (2008). A portable Raman sensor for the rapid identification of the olive oil. In *Technical Proceedings of the 2008 NSTI Nanotechnology Conference and Trade show, NSTI-Nanotech. Nanotechnology*, Vol. 2 (pp. 97).
- Yildiz, G., Randy, L., Wehling, & Susan, L. (2002). Monitoring of the peroxide in soybean and corn oil. *Journal of the American Oil Chemists' Society*, 79(11), 1085–1089.
- Yildiz, G., Randy, L., Wehling, & Susan, L. (2003). Comparison of four analytical methods for the determination of peroxide value in oxidized soybean oils. *Journal of the American Oil Chemists' Society*, 80(2), 103–107.

**Evaluation of the overall quality of olive oil  
using fluorescence spectroscopy**

**Food Chemistry (pendiente de aceptación)**

## Evaluation of the overall quality of olive oil using fluorescence spectroscopy

Elena Guzmán<sup>a</sup>, Vincent Baeten<sup>b</sup>, Juan Antonio Fernández Pierna<sup>b</sup>, José A. García-Mesa<sup>a</sup>

<sup>a</sup> Centro IFAPA 'Venta del Llano', Crta. Nacional Bailén-Motril Km 18.5, 23620 Mengíbar, Jaén, Spain

<sup>b</sup> Walloon Agricultural Research Centre, Valorization of Agricultural Products Department, Food and Feed Quality Unit, Henseval Building, 24 Chaussée de Namur, 5030 Gembloux, Belgium

Corresponding author E-mail address: elena.guzman.jimenez@juntadeandalucia.es(E. Guzmán).Phone: 34953366366

### Abstract

The fluorescence spectra of some olive oils were examined in their natural and oxidized state, with wavelength range emissions of 300-800 nm and 300-400 nm used as excitation radiation. The fluorescence emissions were measured and an assessment was made of the relationship between them and the main quality parameters of olive oils, such as peroxide value, K232, K270 and acidity. The quality control of these parameters involves sophisticated analytical techniques that usually require a large amount of reagents and time-consuming analysis; there is a need for rapid analytical techniques and a low-cost technology for olive oil quality control. The oxidized oils studied had a strong fluorescence band at 430-450 nm. Extra virgin olive oil gave a different but interesting fluorescence spectrum, composed of three bands: one low intensity doublet at 440 and 455 nm; one strong band at 525 nm; and one of medium intensity at 681 nm. The band at 681 nm was identified as the chlorophyll band. The band at 525 nm was derived, at least partially, from vitamin E.

The results presented demonstrate the ability of the fluorescence technique, combined with multivariate analysis, to characterize olive oils on the basis of all the quality parameters studied. Prediction models were obtained using various methods, such as partial least squares (PLS), N-way PLS (N-PLS) and external validation, in order to obtain an overall evaluation of oil quality.

The best results were obtained for predicting K270 with a root mean square (RMS) prediction error of 0.08 and a correlation coefficient obtained with the external

validation of 0.924. Fluorescence spectroscopy facilitates the detection of virgin olive oils obtained from defective or poorly maintained fruits (high acidity), fruits that are highly degraded in the early stages (with a high peroxide value) and oils in advanced stages of oxidation, with secondary oxidation compounds (high K232 and K270). The results indicate the potential of a spectrofluorimetric method combined with multivariate analysis to differentiate, and even quantify, the levels of oil quality. The proposed methodology could be used to accelerate analysis, is inexpensive and allows a comprehensive assessment to be made of olive oil quality.

**KEYWORDS:** Virgin olive oil; UV radiation; quality control

## 1. INTRODUCTION

The demand for safe, high-quality food requires a high level of quality control and processing, which in turn requires the appropriate analytical tools. Virgin olive oil is a high-quality product in terms of its sensory and nutritional properties, and is therefore governed by strict controls on verifying its quality and purity [1]. The comprehensive quality control of a virgin olive oil requires considerable analytical assessment with varying levels of complexity, ranging from simple analytical processes (e.g., UV absorbance) to more complex ones (e.g., sterols, aliphatic alcohols and other minor compounds). These procedures can be time consuming, and highlight the need to develop new analytical techniques that are simpler and faster.

Spectroscopic techniques are ideal for this purpose because they are simple, cost-effective, rapid and non-destructive. Among these techniques is fluorescence spectroscopy, which has widely been applied to olive oil analysis. In a recent study, Sikorska et al. reviewed these applications in detail [2]. Fluorescence spectroscopy enables valuable analytical information to be obtained for different applications: a) discrimination between quality grades; b) detection of adulteration; c) authentication of geographic origin; d) quantification of fluorescent components; e) monitoring of thermal and photo-oxidation; and f) assessment of quality changes of olive oil during storage.

Fluorescence spectroscopy has several advantages. For example, it is a spectroscopic quantitative analysis method that is more sensitive and selective in terms of organic and inorganic compounds, and in general it does not require using consumable reagents and sample pre-treatment [3].

The fluorescence emission spectra of olive oils contain information about their polyphenol and tocopherol content (300-390 nm) [4, 5]. Virgin olive oils present two smooth peaks at 445 and 475 nm, a peak intensity at 525 nm and another peak at 681 nm. Kyriakidis and Skarkalis [6] suggested that the peaks at 445 and 475 nm were related to the oxidation products of fatty acids and that the peak at 525 nm was derived from vitamin E. They also showed, however, that the addition of vitamin E acetate to virgin olive oil increases the fluorescence intensity not only at 525 nm, but also at 445 and 475 nm. They suggested that this was due to oxidized vitamin E that fluoresces at about this region. The peak at 681 nm is related to chlorophylls. The very low intensity of the peaks at 445 and 475 nm is attributable to the high content of monounsaturated fatty acids and phenolic antioxidants in virgin olive oils, which provide greater protection against oxidation.

Another important quality parameter is oil acidity (the free acids in olive oils in terms of oleic acid). Studies on this parameter have focused on the fluorescent components of the main triglycerides in oil, such as oleic, linoleic and palmitic acids. Acidity in olive oil results from hydrolytic rancidity, and the free fatty acids in olive oil derive from these acids. Oleic acid shows a band of fluorescence at 405 nm, and butyric acid (palmitic acid analogue) and linoleic acid show fluorescence bands at 273 and 325 nm. This has been reported in previous studies, which have confirmed that fluorescence intensity at 429-545 nm was due to oleic acid [7].

In recent years much research has been done on developing rapid methods that needs no preparation for assessing the quality of food. There has been an increase in the use of fluorescence spectroscopy, combined with multidimensional statistical techniques, to evaluate food quality. In most of the research conducted, the obtained fluorescence signal was assigned to specific fluorophores after fixing the excitation or emission wavelength.

The articles published on fluorescence in vegetable oils focus mainly on the characterization of these oils. In these studies, the variation in excitation and emission wavelengths allows for the simultaneous determination of compounds that are present in the oils as polyphenols [8, 9]. Guimet et al. (2004) [10] used fluorescence and PARAFAC to detect the adulteration of extra virgin olive oil with pomace oil; other authors have also used fluorescence to study the adulteration of olive oil with other vegetable oils [11,12,13,14].

There are several studies on the use of fluorescence spectroscopy for monitoring the deterioration of extra virgin olive oil in thermoxidized virgin olive oil [15, 16, 17, 18, 19]. There have also been studies on characterizing and classifying olive oils in which the fluorescence spectra are associated with quality parameters such as peroxide value (PV) and UV absorbance at 232 nm (K232) and 270 nm (K270); this work has attempted to determine these quality parameters, but has not produced accurate prediction models [20, 21, 22, 23].

In the light of these studies, and given that fluorophores in edible oils are chemical compounds with a very different nature, directly related to the mechanisms of oxidation, fluorescence spectroscopy is considered to be a suitable technique for evaluating the oxidative status of oils. The aim of this study, therefore, was to determine whether fluorescence spectroscopy provides enough information on quality parameters for routine control and could be used for a general characterization of olive oils.

## **2. MATERIAL AND METHODS**

### **2.1 Samples and sample preparation**

Ninety samples of fresh and heated virgin oils obtained from the 2010-2011 harvest at the Venta del Llano Instituto de Investigación y Formación Agraria y Pesquera (IFAPA) in Jaén, Spain were used.

In order to obtain a wide variety of oxidized oils and determine the relationship between PV), K232, K270, and fluorescence intensity, a rapid oxidation process was applied to 20 samples of olive oil. The oil was heated to 110°C for 96 h under an air



stream. The samples were taken at different times and then allowed to cool to room temperature, before the fluorescence spectra were recorded and the analysis of reference performed. The assessment of acidity involved storing 10 oil samples in anaerobic conditions in order to achieve high acidity values and obtain a wide range of acidity levels. Their fluorescence was then measured and analyzed.

## 2.2 Fluorescence spectroscopy analysis

A three-dimensional (3D) matrix for each sample was obtained, referred to as a fluorescence excitation-emission matrix (EEM). The main advantage of EEMs is that more information can be extracted from them about the fluorescent species because there is a wider area available for analysis. There are some examples in the literature of the use of EEMs for olive oils [24-25].

A spectrofluorimetric measurement of olive oil was conducted using a Cary Eclipse (Varian Ibérica, Madrid, Spain). This instrument was equipped with a continuous xenon lamp, excitation and emission monochromators, and a photomultiplier. A non-fluorescent cell 10 x 10 mm (3 ml) was used for 'right angle' (RA) fluorometry, and the excitation and emission slits (5 nm wide) were used for the emission spectra excited at  $\lambda_{em} = 300-800$  nm  $\lambda_{exc} = 300-400$  nm. In a second analysis, a fluorescence spectrum was measured at a specific wavelength (385 nm). The integration time was 0.1 s, and the increasing wavelength while scanning the spectrum was 10 nm. The experimental conditions were appropriate for providing satisfactory intensity spectra, resolution and signal-to-noise ratio.

The oils were shaken vigorously before measurement in order to homogenize the samples. They were then poured into a plastic cuvet, which was discarded at each stage to ensure the safety and reliability of the measurements and to improve method speed.

Each sample was analyzed in duplicate, and the average measurement (the value of the two scans for each sample) was measured using the original spectra, without any pre-treatment. All computations, chemometric analysis and graphics were performed using Matlab v7.4.0. (The MathWorks, Inc., Natick, MA, USA). For principal

component analysis (PCA) and PLS, the PLS Toolbox v 4.11 (Eigenvector Research, Inc.) was used.

### 2.3 Reference chemical analysis

For all the samples, the reference values for PV, acidity and specific UV absorption at 232 nm (K232) and 270 nm (K270) were obtained following EU Regulations EEC/2568/91 and EEC/1429/92, which focus on the characteristics of olive oil and olive-residue oil and on relevant methods of analysis [26,27]. The samples used in this study varied widely in oxidation level, as shown in Table 1, which presents the minimum and maximum values, the average values and the standard deviation for each of the chemical indices.

The PV is the amount of peroxide in a sample (expressed in milliequivalents of active oxygen per kg fat) that causes the oxidation of potassium iodide. The sample was dissolved in acetic acid and chloroform, and was then treated with a potassium iodide solution. The liberated iodine was titrated with a sodium thiosulfate solution.

Acidity is defined as the free acids in olive oils in terms of oleic acid. The free fatty acid content is expressed as acidity. This is calculated using the conventional method, which involves dissolving the sample in a solvent mixture and measuring the free fatty acids using an ethanolic solution of potassium hydroxide. A low level of acidity indicates that a virgin olive oil is derived from good quality olives and has been kept in peak condition throughout the production process. The K232 and K270 values were spectrophotometric measures for quantifying the UV absorption, at 232 nm (K232) and 270 nm (K270), of a solution of the oil in cyclohexane. Spectrophotometric analysis in UV provides information about the quality of the fat, the conservation status of the oil and any deterioration that might have occurred during the technological processes. High K232 and K270 values indicate the presence of diene- and triene-conjugated systems.

## 3. RESULTS

Olive oil authentication is usually based on chemical parameters (acidity; fatty acid composition; PV; UV absorbance; trisolein, acid and sterol content [28-30]; and

sensorial analysis [31]). Chemical analyses were performed by the laboratory, using official analytical methods [Regulation (EEC) 2568/91] and including four quality parameters (acidity, FV, K232 and K270). Table 1 shows the parameters analyzed, giving the maximum, minimum and average values.

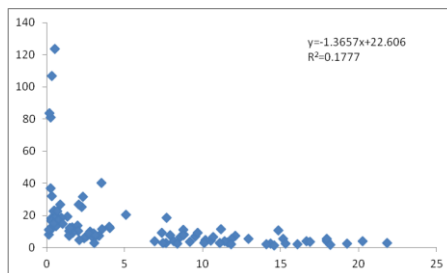
	Mean	Max	Min
PV	22.3	105.5	1.5
acidity	2.4	16.5	0.1
K232	2.4	3.18	1.4
K270	0.3	0.84	0.1

**Table 1.** Mean and minimum-maximum values of the reference values used in the study

The fluorescence EEM was measured directly from entire samples, without any pre-treatment. All the samples were measured in duplicate, and the average value of each sample was used.

During the first stage of the analysis of multivariate data, an initial pre-treatment of the measured signals was performed. This involved mathematical manipulation of the measuring signals prior to analysis. The pre-treatment reduced or eliminated sources of variability in signals of a random nature, such as noise. Various treatments were tested but none improved data results. The calibration of chemometrics was aimed at finding a linear relationship between the variables of a model or a set of experimental data.

The initial studies sought to find the relationship between the fluorescence EEM parameters and acidity. The first step was to check that there was no correlation between oxidation parameters and acidity, as this could lead to an erroneous measurement. Figure 1 shows the correlation between PV and acidity, confirming that there was no relationship between these variables.

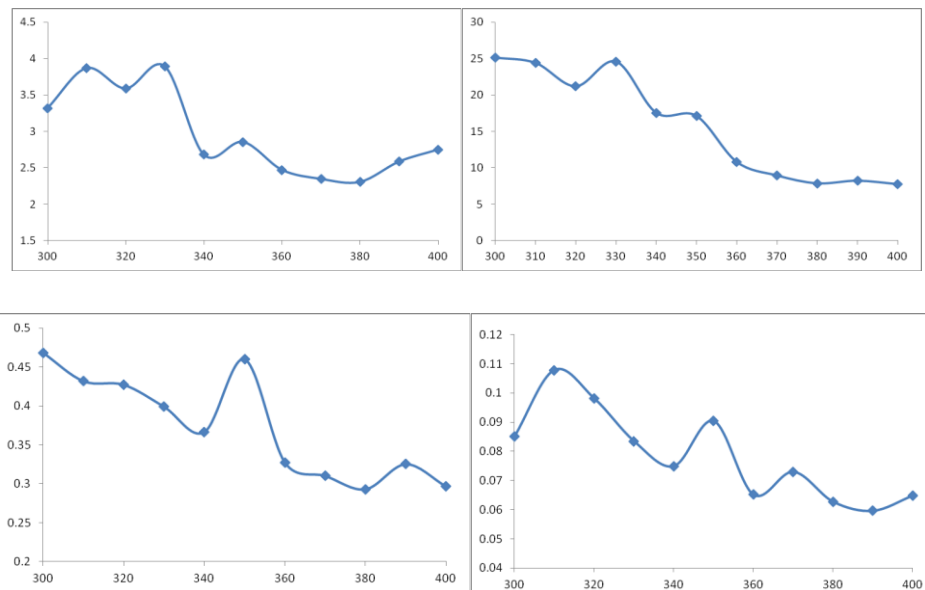


**Figure. 1** Correlation between peroxide value (PV) and acidity

The oil sample's EEM was used to conduct a multi-way PLS regression (N-PLS) to correlate the projected variable values [32]. This method models both the independent (X) and dependent (Y) variables simultaneously to find the latent variables in X that will best predict the latent variables in Y. The optimum number of factors calculated using the leave-one-out cross-validation procedure was seven, with an RMS error of cross-validation (RMSCEV) of 1.77 and a correlation coefficient, obtained in the validation step, of 0.931 for the acidity parameter.

The next stage involved constructing a calibration model for these parameters using a single excitation wavelength in order to accelerate the analysis and avoid the need to measure EEMs at different wavelengths. A PLS model [33] was applied to each wavelength of excitation and the error values were compared in order to select the minimum value. The wavelength was selected to give the best values of correlation with the different parameters and lower error values of 380 nm. The PLS models therefore had a fluorescence spectrum at 380 nm and an excitation wavelength of 400-800 nm, in order to verify that performing a measurement in this wavelength could be done using the same information as EEMs.

The RMSECV for all the parameters for each excitation wavelength is given in Figure 2, which shows that at 380 nm is reached the lowest values for error this error.



**Figure 2.** Root mean square error of cross-validation (RMSECV) for acidity, peroxide value (PV), K232 and K270, respectively, for each excitation wavelength

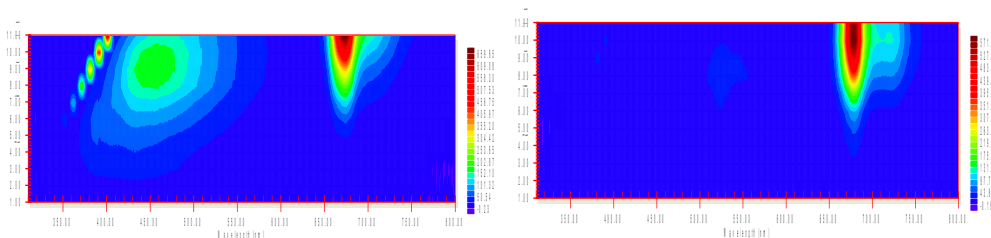
The number of factors used in the PLS method based on the leave-one-out cross-validation procedure was the same as that for the N-PLS method (seven); the RMSECV obtained for acidity was 2.32 and the correlation coefficient obtained in the validation step was 0.892. These results accord with those obtained Poulli et al. (2004) [7] who found information in the fluorescence spectrum in the 429-545 nm region related with oleic acid, but did not find a good predictive model in this case.

The next step was to determine the relationship between fluorescence spectra and other quality parameters, such as those relating to the oxidation state of the oil. The PVs indicate the presence of primary oxidation products (i.e., conjugated hydroperoxides) and K232 and K270 indicate the presence of secondary oxidation products (diene- and triene-conjugated systems).

The relationship between the fluorescence EEMs of oils and these parameters was studied. The high values of the parameters indicated that the sample was degraded, which was confirmed by their high fluorescence intensities.

Previous work has shown that samples with the lowest PV (low level of primary oxidation) display soft fluorescence around 415-600 nm. In contrast, the sample with the highest PV (large amounts of primary oxidation products) has its maximal fluorescence around this zone.

Samples with the lowest level of oxidation did not show fluorescence around the 415-600 nm emission wavelengths. The sample with high oxidation, however, had its maximum fluorescence around this emission wavelength (Figure 3).



**Figure 3.** EEMs between 300-400 nm excitation and 400-800 nm emission of oxidized oil and fresh oil samples

The reference method for analyzing peroxides in the laboratory has an error in the analysis, as it depends largely on external factors. In addition, the PV is very variable because peroxidic compounds decompose rapidly into other oxidation compounds for contact with light and high temperatures, and this could explain the high values in the calibration error for this parameter.

The results of PLS N-quantitative analysis for predicting the PV calculated by means of the method cross-validated with an RMSECV of 6.74 and a coefficient of correlation obtained in the validation of 0.967.

A similar procedure was performed to correlate the fluorescence parameters with K232 and K270. During oil oxidation, the conversion of primary oxidation compounds produces high levels of peroxides in secondary oxidation products, which often give higher K232 and K270 values and fluorescence spectra with peaks in the 415-600 nm region, although the variation of K270 values in oils is best captured by fluorescence.

Oils with the highest K270 value have a broad peak between 415 and 600 nm, with a maximum fluorescence at  $\lambda_{em}$  470 nm [20].

The N-PLS model for K232 and K270 gave RMSECVs of 0.37 and 0.06, and correlation coefficients in the validation stage of 0.908 and 0.958, respectively.

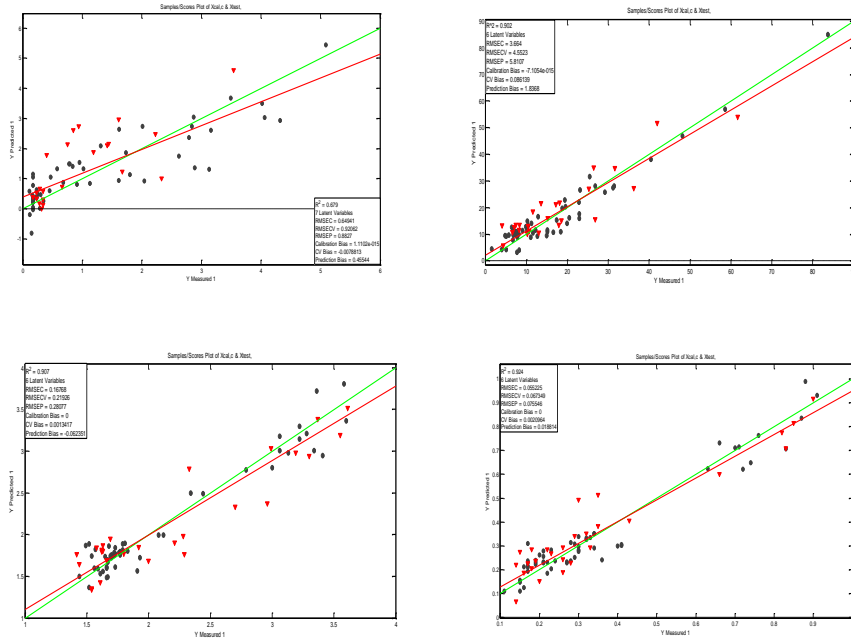
The results obtained using the PLS model for these parameters were RMSECVs of 8.17, 0.32 and 0.06, and correlation coefficients in the validation step of 0.965, 0.881 and 0.939, respectively.

The results obtained with all methods are summarized in Table 2. A comparison of two methods shows that the results are similar and that it is possible to use the fluorescence measure at a single wavelength (380 nm) to obtain enough information to provide an overview of the quality parameters of olive oil.

	N-PLS		PLS		External Validation	
	R2	Error	R2	Error	R2	Error
Acidity	0.931	1.77	0.892	2.32	0.679	0.88
PV	0.967	6.74	0.965	8.17	0.902	5.81
K232	0.908	0.37	0.881	0.32	0.907	0.28
K270	0.978	0.06	0.939	0.06	0.924	0.08

**Table 2.** Results obtained using the N-PLS and PLS methods and external validation for each of the parameters studied

In order to obtain more robust results, the dataset was randomly divided into a calibration set and a validation set. The calibration set, containing 60 samples, was used to construct the model and the validation set, with 30 samples, was used to validate it. The number of factors used for the PLS method in the external validation procedure was six. The results obtained from external validation are given in Figure 4, where the prediction error values obtained were 0.88, 5.81, 0.28 and 0.08, and correlation coefficients were 0.679, 0.902, 0.907 and 0.924, respectively.



**Figure. 4** Reference vs. predicted values in the independent validation set. Acidity, peroxide value, K232 and K270 were predicted by PLS external validation

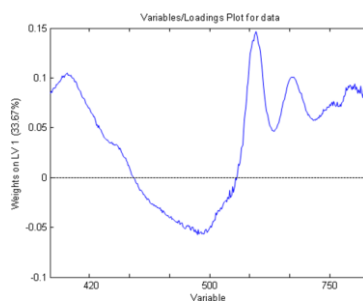
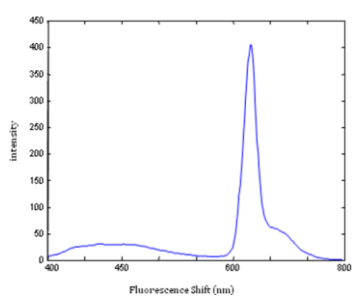
The results obtained indicated that the models constructed for validation, using an independent group of samples, gave prediction and RSMECV values similar to those obtained using the leave-one-out procedure, indicating that these models could be used to conduct an overall assessment of oil quality.

In order to analyze the fluorescence spectrum in more detail, different areas of the spectrum were studied and related to loadings and scores of the variables. Score vectors and loading vectors were used to visualize the relationships between the spectra and the parameters in matrix X, respectively. The relationships between the spectra and parameters were investigated by using the score plots and the loading plots.

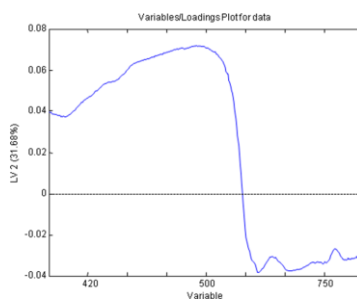
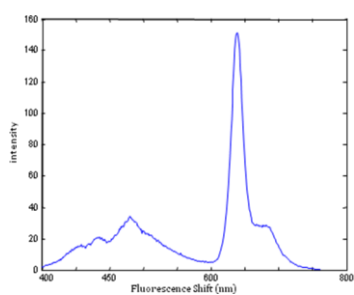
The following are the score and loading plots-plots of the parameters studied as well as the typical range of olive oils with different levels of oxidation and acidity, so that can observe the areas that have more weight on each parameters. Figure 5 shows the typical spectrum of oils of differing quality and their corresponding loading plot.



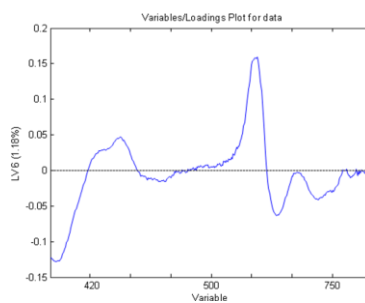
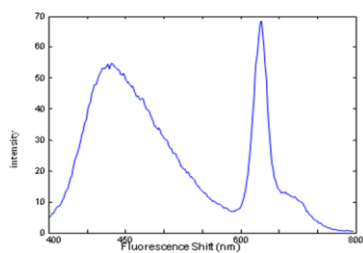
a)



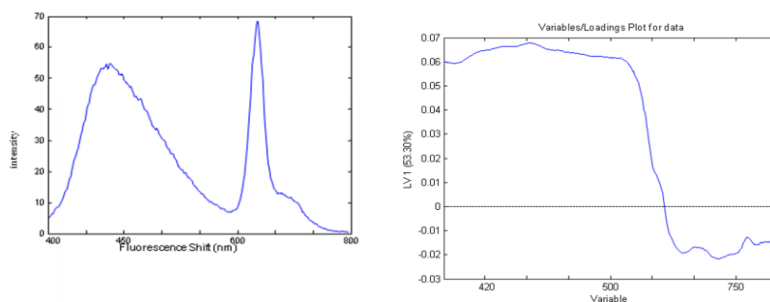
b)



c)



d)



**Figure 5.** Spectrum typical of oils and loading plots showing peroxide, acidity, K232 and K270 values: (a) high PV and low acidity, (b) low PV and high acidity, and (c) and (d) high K232 and K270 and low acidity

These results accord with those obtained by Guimet et al. (2005) [20] using the fluorescence EEM to construct a calibration model for the peroxide parameter with an RMSECV of 1.5 and a correlation coefficient in the validation step of 0.84. This work showed poor correlation, however, between fluorescence and the K232 parameter, which could be due to the larger range of values and samples used to develop the model calibration. The values obtained for the K270 parameter were an RMSECV of 0.08 and a correlation coefficient in the validation step of 0.95, which were very similar to the values obtained with external validation. This was the best calibration model for the prediction of this parameter.

The advantage of this method, therefore, is that it enables these parameters to be determined with a simple measure of a single excitation wavelength without the need to scan the full spectrum, thus saving considerable time.

These results are promising because they provide useful information that can be used for the rapid and easy monitoring of oil quality and show the potential use of fluorescence instruments for the overall evaluation of olive oil quality and for obtaining an overview of the fluorescence of various chemical species from the same analysis. These spectra can be measured quickly and easily, without the need for any pre-treatment of the oil sample, and could possibly be used to identify virgin olive oil. The fluorescence characteristic of oils could therefore be used both to authenticate oils

and to monitor their quality, and could provide an interesting alternative to other spectroscopic methods, such as the NIR spectroscopies used widely in the oil industry, and this method can substituted to actual reference method used.

### **Acknowledgment**

The authors thank the National Institute of Research and Technology Agriculture and Food (INIA) for funding of the scholarship of training FPI sub-INIA, ESF co-financed through the ESF Operational Programme for Andalusia 2007-2013 within axis 3 "Increasing and improving human capital" expenditure category "Developing human potential in the field of research and innovation."



### **References**

- [1] COI/T.15/NC No 3/REV. 6 – International Olive Council (IOC). Trade standard applying to olive oils and olive-pomace oils. Madrid, Spain
- [2] Sikorska, E., Khmelinskii, I., Sikorski, M. (2012) Analysis of Olive Oils by Fluorescence Spectroscopy: Methods and Applications. In Olive Oil - Constituents, Quality, Health Properties and Bioconversions. Edited by Boskou Dimitrios, ISBN 978-953-307-921-9. <http://www.intechopen.com>
- [3] Sikorska, E., Romaniuk, A., Khmelinskii, I. V., Herance, R., Bourdelande, J. L., Silorski, M., Koziol, J. (2004). Characterization of edible oils using total luminescence spectroscopy. J. Fluoresc. 14, 25–35
- [4] Zandomenighi, M., Carbonaro, L., Caffarata, C. (2005). Fluorescence of vegetable oils: olive oils. J. Agric. Food Chem. 53, 759–766
- [5] Giungato, P., Aveni, M., Rana, R., Notarnicola, L. (2004). Modifications induced by extra virgin olive oil frying processes. Ind. Aliment. 43, 369–375

- [6] Kyriakidis, N. B. and Skarkalis, P. (2000). Fluorescence spectra measurement of olive oil and other vegetable oils. *Journal of the American Oil Chemistry Society*, 83, 1435–1439
- [7] Poulli, K., Mousdis, G.A., Georgiou, C.A. (2005). Classification of edible and lampante virgin olive oil based on synchronous fluorescence and total luminescence spectroscopy. *Analytica Chimica Acta* 542, 151–156
- [8] Mújica-Ascencio, C., Moreno-García, E., Stolik Isakina, S., De La Rosa-Vázquez, J.M. (2010). Fluorescence spectroscopy and chemometric techniques to identify compounds in a mixture. *CONIELECOMP 2010 - 20th International Conference on Electronics Communications and Computers*, art. no. 5440781 , pp. 136-141
- [9] Gracia, A., Zude, M., León, L. (2009). Determination of polyphenols in olive oil by means of fluorescence analysis. *5th International Technical Symposium on Food Processing, Monitoring Technology in Bioprocesses and Food Quality Management*, pp. 1364-1368
- [10] Guimet, F., Ferré, J., Boqué, R. (2005). Rapid detection of olive pomace oil adulteration in extra virgin olive oils from the protected denomination of origin 'Siurana' using excitation – emission fluorescence spectroscopy and three-way methods for analysis. *Analytica Chimica Acta*, 544, 143–152
- [11] Poulli, K.I., Mousdis, G.A., Georgiou, C.A. (2006). Synchronous fluorescence spectroscopy for quantitative determination of virgin olive oil adulteration with sunflower oil. *Analytical and Bioanalytical Chemistry* 386 (5), pp. 1571-1575
- [12] Poulli, K.I., Mousdis, G.A., Georgiou, C.A. (2007). Rapid synchronous fluorescence method for virgin olive oil adulteration assessment. *Food Chemistry* 105 (1), pp. 369-375
- [13] Sayago, A., García-González, D.L., Morales, M.T., Aparicio, R. (2007). Detection of the presence of refined hazelnut oil in refined olive oil by fluorescence spectroscopy. *Journal of Agricultural and Food Chemistry* 55 (6), pp. 2068-2071
- [14] Dankowska, A., Małecka, M. (2009). Application of synchronous fluorescence spectroscopy for determination of extra virgin olive oil adulteration. *European Journal of Lipid Science and Technology* 111 (12), pp. 1233-1239
- [15] Cheikhousman, R., Zude, M., Bouveresse, D.J.-R., Léger, C.L., Rutledge, D.N., Birlouez-Aragon, I. (2005). Fluorescence spectroscopy for monitoring deterioration of extra virgin olive oil during heating. *Analytical and Bioanalytical Chemistry* 382 (6) , pp. 1438-1443
- [16] Poulli, K.I., Chantzou, N.V., Mousdis, G.A., Georgiou, C.A. (2009). Synchronous fluorescence spectroscopy: tool for monitoring thermally stressed edible oils. *Journal of Agricultural and Food Chemistry* 57 (18) , pp. 8194-8201

- [17] Poulli, K.I., Mousdis, G.A., Georgio, C.A.. (2009).Monitoring olive oil oxidation under thermal and UV stress through synchronous fluorescence spectroscopy and classical assays. *Food Chemistry* 117, 499–503
- [18] Tena, N., García-González, D.L., Aparicio, R. (2009). Evaluation of virgin olive oil thermal deterioration by fluorescence spectroscopy. *Journal of Agricultural and Food Chemistry* 57 (22), pp. 10505-10511
- [19] Tena, N., Aparicio, R., García-González, D.L. (2012). Chemical changes of thermoxidized virgin olive oil determined by excitation-emission fluorescence spectroscopy (EEFS). 2012. *Food Research International* 45 (1), pp. 103-108.
- [20] Guimet, F., Ferré, J., Boqué, R., Vidal, M., Garcia, J. (2005).Excitation - emission fluorescence spectroscopy combined with three-way methods of analysis as a complementary technique for olive oil characterization. *Journal of Agricultural and Food Chemistry* 53 (24), pp. 9319-9328
- [21] Zandomeneghi, M., Carbonaro, L., Zandomeneghi, G. (2006).Comment on excitation-emission fluorescence spectroscopy combined with three-way methods of analysis as a complementary technique for olive oil characterization. *Journal of Agricultural and Food Chemistry* 54 (14), pp. 5214-5215
- [22] Guimet, F., Boqué, R., Ferré, J. (2006).Application of non-negative matrix factorization combined with Fisher's linear discriminant analysis for classification of olive oil excitation-emission fluorescence spectra. *Chemometrics and Intelligent Laboratory Systems* 81 (1), pp. 94-106
- [23] Sikorska, E., Górecki, T., Khmelinskii, I. V., Sikorski, M., Koziol, J. (2005).Classification of edible oils using synchronous scanning fluorescence spectroscopy. *Food Chemistry*, 89, 217–225
- [24] Engelsen, S. A. (1997). Explorative spectrometric evaluations of frying oil deterioration. *Journal of the American Oil Chemistry Society*, 74, 1495–1508
- [25] Cheikhousman, R., Zude, M., Bouveresse, D. J. R., Leger, C. L., Rutledge, D. N., Birlouez-Aragon, I. (2005).Fluorescence spectroscopy for monitoring deterioration of extra virgin olive oil during heating. *Analytical and Bioanalytical Chemistry*, 382(6), 1438–1443
- [26] EC Commission Regulation No. 2568/91 (1991).
- [27] EC Commission Regulation No. 1429/92 (1992).
- [28] Aparicio, R., Aparicio-Ruíz, R. (2000).Authentication of vegetable oils by chromatographic techniques. *J. Chromatogr. A* 881, 93-114

[29] Armanino, C., Leardi, R., Lanteri, S., Modi, G. (1989). Chemometric analysis of Tuscan olive oils. *Chemom. Intell. Lab. Syst.* 5, 343-354

[30] Marini, F., Balestrieri, F., Bucci, R., Magrí, A. D., A. L., Marini, D. (2004). Supervised pattern recognition to authenticate Italian extra virgin olive oil varieties. *Chemom. Intell. Lab. Syst.* 73, 85-93

[31] Aparicio, R., Morales, M. T., Alonso, V. (1997). Authentication of European virgin olive oils by their chemical compounds, sensory attributes, and consumers' attitudes. *J. Agric. Food Chem.* 45, 1076-1083

[32] Bro, R. Multi-way calibration. (1996). Multi-linear PLS. *J. Chemom.* 10, 47-62

[33] Massart, D.L., Vandeginste, L.M., Buydens, M.C. (1988). *Chemometrics: A textbook*. Elsevier Science Publisher, 23, 395-397

**A portable sensor for the rapid discrimination  
of olives according to fruit quality.**

**Talanta 93, 94-98 (2012)**



## A portable Raman sensor for the rapid discrimination of olives according to fruit quality

Elena Guzmán<sup>a,\*</sup>, Vincent Baeten<sup>b</sup>, Juan Antonio Fernández Pierna<sup>b</sup>, José A. García-Mesa<sup>a</sup>

<sup>a</sup> IFAPA Centro Venta del Llano, Crta. Nacional Bailén-Motril Km 18.5, 23620 Mengíbar, Jaén, Spain

<sup>b</sup> Walloon Agricultural Research Centre, Valorisation of Agricultural Products Department, Food and Feed Quality Unit, Henseval Building, 24 Chaussée de Namur, 5030 Gembloux, Belgium

### ARTICLE INFO

#### Article history:

Received 20 October 2011

Received in revised form 19 January 2012

Accepted 29 January 2012

Available online 2 February 2012

#### Keywords:

Portable Raman spectroscopy

Pattern recognition

Olives

Olive oil quality

### ABSTRACT

In the real marketplace, providing high-quality olive oil is important from the perspective of both consumers and producers. Quality control should meet all requirements in the production process, from farm to packaging. The quality of olive oil can be affected by several factors, including agricultural techniques, seasonal conditions, farming systems, maturity, method and duration of storage, and process technology.

The quality of oil produced also depends largely on the quality of the olives. In an enterprise aimed at producing high-quality oils, olives with defects ('ground'; i.e., fallen to the ground) should be separated from healthy fruit ('sound'; i.e., collected directly from the tree), because a very small portion of low-quality fruit can ruin the whole batch.

The fruit falls partly because of its maturation process, but also because of pest and disease attack or weather conditions (strong wind). Fruit that has fallen to the ground can suffer a rapid deterioration in quality.

Currently, the separation of fruits is based mainly on visual inspection or information provided by the farmer. These are not very reliable procedures. Methods using analytical parameters to characterize the oil, such as acidity and peroxide value, can be applied, but they require a lot of time and materials. Alternative techniques are therefore needed for the rapid and inexpensive discrimination of olives as part of a quality control strategy.

The work described here aims to determine the potential of low-resolution Raman spectroscopy for the discrimination of olives before the oil processing stage in order to detect whether they have been collected directly from the tree (i.e., healthy fruit) or not. Low-resolution Raman spectroscopy was applied together with multivariate procedures to achieve this aim. PCA was used to find natural clusters in the data. Supervised classification methods were then applied: Soft Independent Modeling of Class Analogy (SIMCA), PLS Discriminate Analysis (PLS-DA) and *K*-nearest neighbors (KNN). The best results were obtained using the KNN method, with prediction abilities of 100% for 'sound' and 97% for 'ground' in an independent validation set.

These results demonstrated the potential of a portable Raman instrument for detecting good quality olives before the oil processing stage, by developing models that could be applied before this stage, thus contributing to an overall improvement in quality control.

© 2012 Elsevier B.V. All rights reserved.

### 1. Introduction

A series of important stages, from farm to packaging, are involved in producing good quality olive oil. The quality of the oil can be affected by several factors, including agricultural techniques used, seasonal conditions, farming systems, maturity, method and duration of storage, and process technology. The harvesting stage is a very important one in the production of good quality olive

oil. When harvesting olives, it is important to check the state of ripeness in order to ensure that, during the oil processing stage, the extraction yield and oil quality are optimal. Harvesting can have an important influence on the quality of the oil produced, as well as on the production costs and therefore on the olive farm economy. In terms of quality, there are three main issues to consider when harvesting olives: the collection time, the origin of the fruit, and the method of collection [1].

Under normal conditions, variety and agronomy have no clear influence on the quality regulated. Any variety and a range of agronomic conditions can provide oils that can be classified as extra virgin, if obtained from healthy olives collected at the optimum

\* Corresponding author.

E-mail address: [elena.guzman.jimenez@juntadeandalucia.es](mailto:elena.guzman.jimenez@juntadeandalucia.es) (E. Guzmán).



and appropriate time and put through the correct milling process [2,3].

A delay in harvesting olives can lead to the natural fall of the fruit (over-ripe olives), accelerated by wind; the degree to which this happens depends on the variety [4].

The fruit on the ground can suffer from a variety of defects caused by diseases, pests and ground frost, reducing the quality of the oil subsequently produced. In an earlier study, it was shown that some quality indexes for oil were not affected by frostbite, but the oil became less stable and there were some sensory changes.

When olives are on the ground for several days, this leads to a deterioration of the fruit pulp which, in turn, adversely affects oil quality. This loss of quality results in increased acidity, a higher peroxide index and a deterioration in the sensory characteristics of the oil, as demonstrated in several previous studies [5].

After harvesting, the most important factor affecting fruit quality is storage prior to production, which, if inadequate, can lead to higher acidity and peroxide values and promote enzymatic activities, leading to deterioration in sensory characteristics [6].

For all these reasons, it is very important to collect, transport and process healthy fruit separately from defective fruit, as small amounts of defective fruit can significantly alter the organoleptic characteristics of oils obtained from healthy fruits and thus reduce their quality. Fruit that has the potential to provide good quality oils (i.e., healthy olives picked from the tree) must follow a different route from the outset – from collection on the farm and factory reception and milling process to storage of the oil produced.

When olives arrive at the factory, there should be a control check whereby their quality is continuously evaluated. The oil production sector is making a major effort to improve the quality of virgin olive oil, but sometimes there is inadequate control and verification at the point when olives first enter the factory. In most cases, quality control at this stage is based on visual observation of the olives or information provided by the farmer. Thus, the method of olive harvest is an important checkpoint in the process of producing high quality olive oils, and alternative techniques are needed for the rapid and inexpensive discrimination of olives at this stage, as part of a quality control strategy. An important first step, therefore, lies in the quality control of the fruit at the factory reception stage and the classification of olives at the start of the production process, and the need to avoid mixing of olives of different qualities by using fast, safe and clean analytical methodologies.

The characterization of olive fruit by means of optical technologies has been developed mainly by NIR and Raman spectroscopies. With this purpose, NIR technology has been applied to the direct analysis of intact olives for the determination of fat content and moisture, including in some cases other parameters like free acidity, fatty acid composition, olive variety and fruit traits [7–13]. In other contributions, NIR spectra have been collected after grinding of olive fruits, in the resulting pulp [14–17]. On the other hand, Raman spectroscopy has been proposed for the determination of several compositional and quality parameters in olives after grinding. Thus, fat and moisture [18,19], free acidity [20], carotenoids and phenolic compounds [18] have been evaluated.

Concerning olive oil analysis, Raman spectroscopy has been mainly applied to authentication purposes, to classify PDO olive oils [21,22] and to detect adulteration with other low price oils [23–34]. This technique has been also applied to the determination of the oxidative degradation [35], free acidity [36,20] and the characterization of antioxidant olive oil biophenols [37].

To our knowledge, there are only two contributions dealing with the evaluation of fruit quality, from a point of view of origin of fruit (tree or soil) or status (sound or affected by some disease, frostbite or fermented). For this purpose, NIR [38] and Raman [39] spectroscopies have been applied, achieving a high classification capability.



Fig. 1. System designed for this work in order to take Raman measurements.

In this study, a portable low-resolution Raman spectroscopy method was applied for the discrimination of olives according to their 'ground'/'sound' origin. The work aimed to ensure the quality of olives before they entered the oil production process. In an earlier study, some applications of this technique for determining the quality parameters in olive oil oxidation [40] were reported. This technique was based on using an instrument that was small, inexpensive and easily installed for checking the quality of oil production process.

## 2. Materials and methods

### 2.1. Samples and sample preparation

One hundred ninety-two olives samples from 11 Spanish varieties, mainly Picual, were harvested and analyzed from October to March over 2 years (the 2009–2010 and 2010–2011 harvests) at the Instituto de Investigación y Formación Agraria y Pesquera (IFAPA) center 'Venta Del Llano' in Mengíbar, Jaén, Spain.

The samples included 86 samples of 'sound' olives picked from the tree and 106 olives collected from the ground, including frozen and fermented olives. Table 1 shows, for both harvests, the quality parameters for these samples, including humidity and oil content, as well as the quality parameters of the oil extracted in a laboratory mill [41]: free acidity, peroxide value, and spectrophotometric absorptions at 232 nm ( $K_{232}$ ) and 270 nm ( $K_{270}$ ).

Before the Raman measurements, the olives went through a system which was designed for this work in order to take measurements of several probes with different instruments simultaneously in the same experience (NIR, Raman and VIS measurements). This system is a roller mill that had a circular stainless steel component (4 cm in diameter) linked to the optical fiber probe Raman equipment.

All the olives entered the mill through a hopper and passed through the circuit where the Raman measurement was taken from the olive paste using a glass window (2 mm thick).

Fig. 1 shows the structure that is designed to take the Raman steps.

### 2.2. Raman spectroscopy analysis

The instrument used was a portable Raman spectrometer model RH-3000 (Ocean Optics); it had a 785 nm laser diode and was linked to an optical fiber that enabled measurements to be taken through a glass container. The Raman spectra were recorded from 200 to

**Table 1**  
Standard deviation, mean and min–max reference values of the olive samples used in the study.

	Harvest 2009/2010				Harvest 2010/2011			
	Max	Min	Mean	SD	Max	Min	Mean	SD
Humidity (%)	53.07	40.69	45.77	2.83	61.82	40.51	48.58	4.65
Oil content (%)	32.52	17.94	26.02	2.89	30.09	10.09	21.45	3.57
IP (meqO <sub>2</sub> /kg)	7.5	0.6	3.42	2.09	22.37	0.91	5.42	3.40
K <sub>232</sub> (UV abs.)	1.72	1.28	1.49	0.11	2.74	1.23	1.56	0.25
K <sub>270</sub> (UV abs.)	1.24	0.1	0.19	0.21	1.34	0.09	0.17	0.13
Acidity (%oleic ac.)	2.06	0.21	0.99	0.58	12.69	0.17	1.55	2.02

2700 cm<sup>-1</sup>, with a spectral resolution of 10 cm<sup>-1</sup>. The laser diode of the instrument had an intensity range of 0–290 mW. Spectra were acquired with a laser intensity set at 190 mW, and were the average of 10 scans, resulting in an acquisition time of 5 min.

The equipment used a laptop, with the physical dimensions of 20 (L) × 18 (A) and weighing about 4 kg, as a preliminary simulation of the measurements that could be taken by coupling this equipment in the factory mill, enabling it to be used during the production process.

### 2.3. Chemometrics

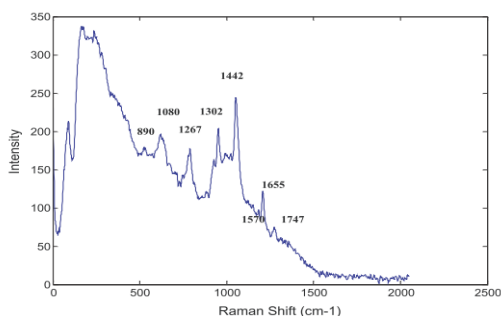
The first step involved classification, which was carried out using an exploratory analysis with Principal Components Analysis (PCA) in order to find trends that would group the samples based only on their spectral characteristics. PCA is an unsupervised technique that reduces the dimensionality of the original data matrix, retaining the maximum amount of variability. It is an important tool for in the good exploration of data [42].

In the second step, three supervised classification methods were used: Soft Independent Modeling of Class Analogy (SIMCA) [43], PLS Discriminant Analysis (PLS-DA) [44] and *K*-nearest neighbor (KNN) [45].

*SIMCA* classifies samples/spectra based on their similarities in a principal components space. A principal components model is built for each available class. For each modeled class, the mean orthogonal distance of training data samples from the line, plane or hyper-plane (calculated as the residual standard deviation) is used to determine a critical distance for classification. Based on these distances, the classification of a new set of samples is obtained by adjusting each of these samples to each model and deciding, at a 95% confidence level, if the sample belongs to the corresponding class or not. In this model, the distance of a point from a class was determined by the out-of-space distance; that is, by the Euclidean distance of the point from the subspace spanned by the *k* principal components used to model the class [46].

*PLS-DA* is performed in order to sharpen the separation between groups of observations, by rotating PCA components so that a maximum separation among classes is obtained, and to understand which variables carry class-separating information. PLS-DA consists of a classical PLS regression where the response variable is a categorical one (replaced by the set of dummy variables describing the categories) expressing the class membership of the statistical units [47].

*KNN* is a method for classifying objects based on the closest training examples in the feature space. By comparing the distances between the unknown object and the training samples, whose closeness is known in advance, KNN with shortest distances is selected, and the classification type represented by a simple majority of the KNN is assigned to the known object. The Euclidean distance is the most common algorithm used for this purpose [48].



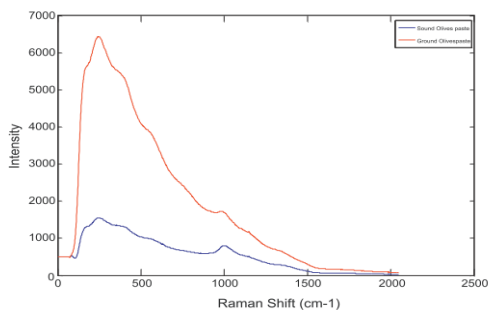
**Fig. 2.** Raman spectra of oil olive using dispersive Raman spectrometry.

### 2.4. Software

All computations, chemometric analyses and graphics were performed using Matlab v7.4.0 (The Mathworks, Inc., Natick, MA, USA). For PCA and supervised methods (*SIMCA*, *PLS-DA*, *KNN*), the *PLS* toolbox v. 4.11 (Eigenvector Research, Inc.) was used.

## 3. Results and discussion

Figs. 2 and 3 shows characteristic dispersive Raman spectra of olive oil, 'ground' olive paste and 'sound' olive paste. In the dispersive Raman spectra, the same bands as in the FT-Raman spectra can be observed. The major bands of virgin olive oil are at 1267 (in plane



**Fig. 3.** Raman spectra of sound and ground olive paste using dispersive Raman spectrometry.

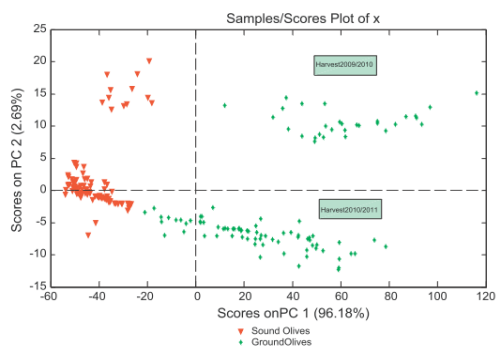


Fig. 4. PCA1 vs PC2 for calibration and validation set.

$\delta$  (=C–H) deformation in unconjugated *cis* double bond) 1302 (in-phase methylene twisting motion), 1442 ( $\delta$ (CH<sub>2</sub>)) and the band at 1655 ( $\nu$ (C=C)), dominate the spectra. The band at 1747 cm<sup>-1</sup>, which is assigned to the  $\nu$ (C=O) ester vibration, can be found in both oil and the olive spectra [49,50].

The main difference in the spectra of the olives and olive oil is the high fluorescence background that characterizes the olive spectrum that overlaps with the typical oil bands. The difference between the spectra of the different kinds of olives ('ground' and 'sound') is also the fluorescence that is lower in the 'sound' olives and higher in the 'ground' olives or those olives with some type of damage (fermented and frostbite). This fluorescence is probably caused by the destruction of cell membranes and alterations in vegetable matter, such as oxidation and enzyme reactions. The olives are collected throughout the whole season (October–March) so that case the olives belong to all possible ranges of maturity indexes. In such a way, there are healthy and ground olives from green to black, therefore the color is not the factor that differentiates these classes.

Apart from this fluorescence phenomenon, there are several variations of the relative intensity, little change in band positions, and small changes in the spectral contours between classes [39].

To define the categories in the classification, the unsupervised PCA method was used. Data pretreatment comprised only auto-scale, which is a common preprocessing method that uses mean-centering followed by the division of each column (variable) by the standard deviation of that column. This approach is valid for correcting different variable scaling and units if the predominant source of variance in each variable is a signal rather than noise. Under these conditions, each variable will be scaled so that its useful signal has an equal footing with the signal of other variables.

Previously, the dataset had initially been randomly split into a calibration set and a test set. The calibration set consisted of 142 spectra and was used to build the model; the test set contained 50 spectra and was used to validate it.

The PCA model was calculated for seven principal components, with an explained variance of 96%. Fig. 4 shows the scores for the first two principal components of the model, where a separation of olives samples into different groups can be observed; the first PC shows the separation of samples according to 'sound' or 'ground', whereas the second component shows the separation of olives samples according to the period harvest to which they belong. These results show that the fruits could be differentiated according to harvest year, as reported by Gurdeniz et al. [51].

The preliminary exploratory analysis of data showed that Raman spectroscopy can see the grouping of olives according to their spectral characteristics and can be used for supervised classification

Table 2  
Calibration and validation results in terms of classification ability given in percentage of PLS-DA, SIMCA and KNN classification.

	SIMCA		PLS-DA		k-NN(3)	
	Cal set	Val set	cal set	Val set	Cal set	Val set
Sound olives	98%	100%	100%	92%	100%	100%
Ground olives	90%	73%	97%	100%	98%	97%

with two different classes or categories. This would allow a separation of the olives before their production, with 'sound' samples prepared separately from 'ground' samples as they enter the oil production process.

The next step involved multivariate classification, using three supervised methods: SIMCA, PLS-DA and KNN. The SIMCA and PLS-DA models were constructed with seven main latent variables for both classes. In the KNN method, the classification error depended on the number of the nearest neighbors chosen for the analysis ( $k$ ) [52]. The best number of  $K$  neighbors was selected using a leave-one-out procedure [53], with  $K$  values between 3, 5 and 7. For  $K=3$ , the recognition and classification abilities were 100% for the discrimination. However, for  $K=5$  and 7, the prediction abilities provided with KNN were 97% in both cases.

Table 2 provides a summary of the three model results for the calibration and validation sets in terms of correct classification ability (in percentage).

The linear classifiers (SIMCA, PLS-DA) gave good results for discrimination between olive classes, with a low error of classification, but KNN was found to be the most effective classification technique.

The most important point is that KNN is a non-parametric method where the quality of predictions depends on the distribution of some classes. Thus, the method can correctly classify objects, although classes are not linearly or quadratic separable. The KNN method sets irregular boundaries between classes, while the parametric methods such as SIMCA and PLS-DA depend on a distribution determined by the statistical limits (Bayesian theorem) to establish the boundaries between classes. In this study, the samples were highly clustered and therefore the KNN method, which uses distances to the nearest neighbor to classify the samples, was the most appropriate method.

The results obtained using the KNN method correctly classified all 'sound' olives samples in external validation, and only one 'ground' olives sample was misclassified; the similar values in classification and prediction indicate that the model is fairly stable. These results showed that the most appropriate classification technique for the classification task was the KNN method, which tended to produce more robust results, although good classification ability was obtained with the PLS-DA and SIMCA models.

Finally, with the data we had from different years, and using SIMCA and PLS-DA methods, we tried to use the data from one year to calibrate olive samples and from another year to do the sample validation, but the results were not satisfactory (not shown). The differences between different harvests were directly related to sources of variation not included in the PCA; there were large differences between samples and environmental conditions (temperature, humidity and rainfall), the harvesting period differed for the two consecutive harvest years and the composition (fatty acid profile) also changed. Geographical origin and harvest year are known to have a significant influence on the fatty acid profile of olive oil [54,55]. An earlier study on Cornicabra virgin olive oil revealed significant statistical differences in quality indexes, major fatty acids and sterol compositions in relation to year of production [56]. The data included sources of variation that were not completely eliminated by preprocessing, so it is necessary to establish a process for the transfer of calibration models in subsequent years because of these variations among different harvests.

In this study there were only 2 years of data, so prediction was difficult. More data from different years are needed in order to obtain reliable and robust prediction equations [11,57].

The combination of portable low-resolution Raman spectroscopy and multivariate discriminant analysis in olives could be an important tool for checking the quality of olives before they enter the oil production process. This is necessary to ensure good quality oil by creating more controlled production processes and saving much expense by avoiding unwanted mixing.

This method has many advantages. It is rapid and inexpensive because it does not need sample preparation. Also, it is non-destructive and uses no chemicals, and is therefore environmentally friendly. Apart from these advantages, this technique allows the possibility of coupling optical fibers, which can be directly inserted into the process or connected to a flow-cell through which the sample can be diverted from the production line. This would allow complete control of the quality of the oil before, during and after the processing of the olives.

## References

- [1] D. Boskou, Olive Oil: Chemistry and Technology, AOCS Press, Champaign, IL, 1996, pp. 14–19.
- [2] IOOC, Trade Standard Applying to Olive Oil and Olive Pomace Oil; COI/T15 3, International Olive Oil Council, Madrid, Spain, 2003.
- [3] M. Uceda, M. Hermoso, R. Fernández-Escobar, L. Rallo, in: Barranco (Ed.), *El Cultivo del Olivo*, 2nd ed., Junta de Andalucía, Grupo Mundi-Prensa, Madrid, Spain, 1997.
- [4] L. Civantos, R. Contreras, R. Grana, Obtención del aceite de oliva virgen, Editorial Agrícola Española, S.A. Madrid, España, 1992.
- [5] J.R. Morello, M.J. Motilva, T. Ramo, M.P. Romero, Food Chemistry 81 (2003) 547–553.
- [6] A. Kiritisakis, G.D. Nanos, Z. Polymenopoulos, T. Thomai, E.M. Sfakiotakis, JAOCS 75 (1998) 6.
- [7] A. Gracia, L. Leon, Grasas y Aceites 62 (3) (2011) 268–274.
- [8] J.A. Cayuela, J.M. Garcia, N. Caliani, NIR prediction of fruit moisture, free acidity and oil content in intact olives, Grasas y Aceites 60 (2) (2009) 194–202.
- [9] J.A. Cayuela, M.D.P. Camino, European Journal of Lipid Science and Technology 112 (11) (2010) 1209–1217.
- [10] N. Dupuy, O. Galtier, Y. Le Dreau, C. Pinatel, J. Kister, J. Artaud, European Journal of Lipid Science and Technology 112 (4) (2010) 463–475.
- [11] L. Leon, A. Garrido-Varo, G. Downey, Journal of Agricultural and Food Chemistry 52 (16) (2004) 4957–4962.
- [12] L. Leon, L. Rallo, A. Garrido, Grasas y Aceites 1 (54) (2003) 41–47.
- [13] A. Morales-Sillero, V.M. Fernandez-Cabanas, L. Casanova, M.R. Jimenez, M.P. Suarez, P. Rallo, Journal of Food Engineering 107 (1) (2011) 99–106.
- [14] A. Bendini, L. Corretani, F. Di Virgilio, P. Belloni, G. Lercker, T.G. Toschi, European Journal of Lipid Science and Technology 109 (5) (2007) 498–504.
- [15] M.J. Ayora-Canada, B. Muik, J.A. Garcia, Mesa, D. Ortega-Calderon, A. Molina-Diaz, Spectroscopy Letters 38 (6) (2005) 769–785.
- [16] A. Jimenez, E. Izquierdo, F. Rodriguez, J.I. Duenas, C. Tortosa, Grasas y Aceites 51 (5) (2000) 311–315.
- [17] M. Casale, P. Zunin, M.E. Cosulich, E. Pitarino, P. Perego, Food Chemistry 122 (2010) 1261–1265.
- [18] M. Lopez-Sanchez, M.J. Ayora-Canada, A. Molina-Diaz, Journal of Agricultural and Food Chemistry 58 (1) (2010) 82–87.
- [19] B. Muik, B. Lendl, A. Molina-Diaz, M.J. Ayora-Cañada, Applied Spectroscopy 57 (2003) 2.
- [20] B. Muik, B. Lendl, A. Molina-Diaz, M.J. Ayora-Canada, Analytica Chimica Acta 487 (2) (2003) 211–220.
- [21] R. Korifi, Y. Le Dreau, J. Molinet, J. Artaud, N. Dupuy, Journal of Raman Spectroscopy 42 (7) (2011) 1540–1547.
- [22] Y. Dellaa, R. Korifi, Y. Le Dreau, J. Artaud, N. Dupuy, XXII International Conference on Raman Spectroscopy. AIP Conference Proceedings, 1267 (2010) 562–563.
- [23] X.F. Zhang, X.H. Qi, M.Q. Zou, F. Liu, Analysis Analytical Letters 44 (12) (2011) 2209–2220.
- [24] R.M. El-Abassy, P. Donfack, A. Materny, Journal of Raman Spectroscopy 40 (9) (2009) 1284–1289.
- [25] M.Q. Zou, X.F. Zhang, X.H. Qi, H.L. Ma, Y. Dong, C.W. Liu, X. Guo, H. Wang, Journal of Agricultural and Food Chemistry 57 (14) (2009) 6001–6005.
- [26] X.H. Qi, X.F. Zhang, M.Q. Zou, Y.Z. Chen, C. Wang, J.T. Ye, C.W. Liu, X. Guo, H. Wang, Clean Technology 2008: Bio Energy, Renewables, Green Building, Smart Grid, Storage, and Water, 2008, p. 371.
- [27] H.M. Heise, U. Damm, P. Lampen, A.N. Davies, P.S. McIntyre, Applied Spectroscopy 59 (10) (2005) 1286–1294.
- [28] V. Baeten, J.A.F. Pierna, P. Dardenne, M. Meurens, D.L. García-Gonzalez, R. Aparicio-Ruiz, Journal of Agricultural and Food Chemistry 53 (16) (2005) 6201–6206.
- [29] E.C. Lopez-Diez, G. Bianchi, R. Goodacre, Journal of Agricultural and Food Chemistry 51 (21) (2003) 6145–6150.
- [30] H. Yang, J. Irudayaraj, Journal of the American Oil Chemists Society 78 (9) (2001) 889–895.
- [31] A.N. Davies, P. McIntyre, E. Morgan, Applied Spectroscopy 12 (54) (2000) 1864–1867.
- [32] R. Aparicio, V. Baeten, Ocl-Oleagineux Corps Gras Lipides 5 (4) (1998) 293–295.
- [33] V. Baeten, P. Hourant, M.T. Morales, R. Aparicio, Journal of Agricultural and Food Chemistry 46 (7) (1998) 2638–2646.
- [34] V. Baeten, M. Meurens, M.T. Morales, R. Aparicio, Journal of Agricultural and Food Chemistry 44 (8) (1996) 2225–2230.
- [35] R.M. El-Abassy, P. Donfack, A. Materny, Food Research International 43 (3) (2010) 694–700.
- [36] R.M. El-Abassy, P. Donfack, A. Materny, Journal of the American Oil Chemists Society 86 (6) (2009) 507–511.
- [37] F. Paiva-Martins, V. Rodriguez, R. Calheiros, M.P.M. Marques, Journal of the Science of Food and Agriculture 91 (2) (2011) 309–314.
- [38] M.J. Ayora-Cañada, B. Muik, J.A. Garcia-Mesa, D. Ortega-Calderon, A. Molina-Diaz, Spectroscopy Letters 38 (6) (2005) 769–785.
- [39] B. Muik, B. Lendl, A. Molina-Diaz, D. Ortega-Calderon, M.J. Ayora-Canada, Journal of Agricultural and Food Chemistry 52 (20) (2004) 6055–6060.
- [40] E. Guzmán, V. Baeten, J.A.F. Pierna, J.A. Garcia-Mesa, Food Control 22 (12) (2011) 2036–2040.
- [41] J.M. Martínez Suárez, E. Muñoz Aranda, J. Alba Mendoza, A. Lanzón Rey, Grasas y Aceites 26 (6) (1975) 379–385.
- [42] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics: A Textbook, Elsevier Science Publisher, 1988, 23, pp. 395–397.
- [43] S.M. Wold, in: B.R. Kowalski (Ed.), Chemometrics: Theory and Application, ACS Symposium Series 52, American Chemical Society, Washington, DC, 1977, pp. 243–282.
- [44] S. Wold, Pattern Recognition 8 (1976) 127–139.
- [45] S. Wold, M. Sjostrom, in: B.R. Kowalski (Ed.), Chemometrics: Theory and Application, American Chemical Society, Washington, 1977, p. 243.
- [46] K. Vanden, M. Hubert, Chemometric and Intelligent Laboratory Systems 79 (1–2) (2005) 10–21.
- [47] M. Bylesjö, M. Rantalainen, O. Cloarec, J.K. Nicholson, E. Holmes, J. Trygg, Journal of Chemometrics 20 (8–10) (2006) 341–351.
- [48] T. Yu, Chemistry Dissertations, 11 (2009).
- [49] H. Sadeghi-Jorabchi, P.J. Hendra, R.H. Wilson, P.S. Belton, JAOCS 67 (1990) 483–486.
- [50] V. Baeten, P. Dardenne, P. Hourant, M.T. Morales, R. Aparicio, Journal of Agricultural and Food Chemistry 46 (7) (1998) 2638–2646.
- [51] G. Gunderiz, B. Ozon, F. Tokatli, European Food Research and Technology 227 (2008) 1275–1288.
- [52] E. Sikorska, J. Chmielewski, T. Görecki, I. Khmelinskii, M. Sikorski, D. De Keukeleire, Journal of the Institute of Brewing 113 (1) (2007) 110–116.
- [53] K. Varmuza, Analytica Chimica Acta 122 (3–15) (1980) 227–240.
- [54] M. D'Imperio, G. Dugo, M. Alfa, L. Mannina, A.L. Segre, Food Chemistry 102 (2007) 956–965.
- [55] E. Stefanoudaki, F. Kotsifaki, A. Koutsafakis, Journal of the American Oil Chemists Society 76 (1999) 623–626.
- [56] M.D. Salvador, F. Aranda, S. Gomez-Alonso, G. Fregapane, Food Chemistry 80 (2003) 359–366.
- [57] R. Blanco Ordás, F.J. García Ramos, B. Diezma Iglesias, Tierras de Castilla y León: Agricultura (2011) 92–97.

**Determination of Olive Maturity Index of intact  
fruits using image analysis.**

**Food Science and Technology (en revisión)**

## Determination of Olive Maturity Index of intact fruits using image analysis

Elena Guzmán<sup>a</sup>, Vincent Baeten<sup>b</sup>, Juan Antonio Fernández Pierna<sup>b</sup>, José A. García-Mesa<sup>a</sup>

<sup>a</sup> Centro IFAPA 'Venta del Llano', Crta. Nacional Bailén-Motril Km 18.5, 23620 Mengíbar, Jaén, Spain

<sup>b</sup> Walloon Agricultural Research Centre, Valorization of Agricultural Products Department, Food and Feed Quality Unit, Henseval Building, 24 Chaussée de Namur, 5030 Gembloux, Belgium

Corresponding author E-mail address: elena.guzman.jimenez@juntadeandalucia.es (E. Guzmán). Phone: 34953366366

### Abstract

In this work, the maturity index of different samples of olives was objectively assessed by image analysis obtained through machine vision, in which algorithms of color-based segmentation and operators to detect edges were used.

This method allows a fast, automatic and objective prediction of olive maturity index. This prediction value was compared to maturity index (MI), generally used by olive oil industry, based on the subjective visual determination of color of fruit skin and flesh. Machine vision was also applied to the automatic estimation of size and weight of olive fruits.

The proposed system was tested to obtain a good performance in the classification of the fruit in batches. When applied to several olive samples, the maturity index predicted by machine vision was in close agreement with the maturity index of fruits visually estimated, values that are currently used as standards. The evaluation of weight of fruit also provided good results ( $R^2 = 0.91$ ).

These results obtained by image analysis can be used as a useful method for the classification of olives at the reception in olive mill, allowing a better quality control of the production process.

**Key Words:** Algorithm, maturity index, olive fruit, image analysis

## Introduction

The quality of virgin olive oil depends on both agronomic and technological factors. Concerning the agronomic aspects, there are several main factors that can affect the quality of the olive oil: olive variety, irrigation rate, pest attacks, fruit diseases and harvest time (Uceda et al, 1975). The harvest moment is the first and crucial decision of the production process of virgin olive oil. This date has to be decided based on several factors: fruit retention force, oil content in the olive fruit, chemical composition, sensory attributes of olive oil and industrial yield. Because these factors evolve during fruit maturity, it is necessary to have techniques allowing the following up of the maturity process. The most common method to evaluate the maturity in olives is based on the visual determination of the maturity index (MI). This technique considers changes in the color of the skin and flesh during maturity of the fruit and gives them a classification (from 0 to 7) according to the color [2,3]. It is a very time consuming procedure and very subjective since it depends on the opinions and experience of the evaluator, as well as the environmental factors such as lighting, color appearance and condition of the fruits that can be highly variable. However, this method is subjective and depends on environmental conditions. Moreover, quite often the evaluators have to decide between colors that are difficult to distinguish. From a visual point of view and according to the stage of ripeness, the color of olives can change from intense green to black (Mínguez 1995; Roca, 2001) [2,3]. This is due to the fact that in the final stages, most of chlorophylls are degraded and replaced by anthocyanins, which allows the olive fruits to be more sensitive to external damage and infections [4, 5, 6].

This maturity index is useful for producers to identify the optimal time of harvesting the olives and also to increase the quantitative and qualitative characteristics of olive oil production. As example, for the Picual olive cultivar, a good balance between yield and quality of olive oil can be obtained with an MI value of between 3 and 4 regardless of the extraction process. [9] There are other different methodologies to determine the state of maturity of the olives such as the determination of fruit firmness [10,11], the measurement of the rate of respiration of the fruit [12], the use of non-destructive methods such as a densimeter [13,14,15] or by measuring the transmission of acoustic waves through the fruit [16]. Other authors have proposed a modification of

some classical chemical parameters as indicative of the maturity process. Thus, Mickelbart and James (2003) [17] proposed the use of the percentage of dry matter in the fruit and Cherubini et al. proposed the sugar content of olives as a technological maturity index. Recently, Near Infrared Spectroscopy (NIRS) has been also applied for the evaluation of the maturity of olive fruits. Gracia and León (2011) [18], studied the evolution of oil and moisture contents in intact olives during maturity. Cayuela et al. [19] carried out a similar study but including also the prediction of acidity and oil content referring to fresh weight, and to dry matter. In both works, the prediction of olive fruit and virgin olive oil of these quality parameters through the direct measuring of the fruit using NIRS was investigated. Bellicontro et al. have applied the NIR technology to the on-field measurement of individual concentration of main polyphenols (oleuropein, verbascoside, and 3, 4-DHPEA-EDA) and total polyphenols in olives, proposing these parameters as indicators of fruit maturity. [20] Morales-Sillero et al. (2011) studied the feasibility of NIR spectroscopy for the non-destructive characterization of table olive traits, including the maturity index. [21] Salguero-Chaparro et al (2012) studied the optimization of the acquisition parameters for on-line analysis of intact olive fruits by vis-NIR spectroscopy, which is a first step in order to implement the NIRS technology for process control in olive oil production [22].

Another important technique to evaluate the maturity status of olive fruits is computer vision. Many applications have been developed in the last years using this technique for classifying the fruit according to different criteria. The maturity level has been used to classify peaches during harvest by applying histograms of the red channel (R) and a combination of red and infrared images (red divided by the corresponding infrared image(R/IR). [23] Other studies have used a region-oriented segmentation algorithm for detecting the most common peel defects in citrus fruits. [24] This detection is mainly focused on the regions of interest consisting of the sound peel, the stem and the defects and allows successful segmentation of smaller defects, such as scale [25], quality control and defects in cherries [26, 27] and apples. [28, 29, 30] The detection can be also based on the shape, size and maturity of the fruits, as it was shown for cucumbers. [31] In this case, artificial classifiers such as linear discrimination analysis and multi-layer perceptron neural network were used driving to 100% of success rate in detecting cucumbers of three different shape categories: well-



formed, slightly deformed and seriously deformed. Other studies have measured the volume of watermelon by using water displacement, ellipsoid approximation and image processing methods. [32] Brosnan and Sun [33] presents a comprehensive review of image processing techniques for various food products as bakery products, meat and fish, vegetables or fruits. Moreover, previous studies have demonstrated the ability of sorting fruits according to the color. [34, 35, 36, 37] Significant correlations between shape and color during aging in apples have been found by evaluating the average fruit's diameters per tree with manual measurements at different developing stages of apple fruits [38] In the literature some works used the correlation between images and the degree of ripeness to characterize the ripening process and define appropriate threshold values for different cultivars in order to consider its possible applications in field and postharvest practice. [39, 40] As well, computer vision has been applied to table olive to detect external damage [41, 42, 43] and has become a widely used technology for the production of table olives at the industrial level.

Until now, all the studies developed using vision systems allowed obtaining a global and unique maturity index for the set of olives. The new vision method, based on the use of infrared and visible images, allows obtaining a maturity index for each individual olive in a set, which provides more quantity of information about the olives entering the factory and can help to take decisions on the process line. Therefore, for acceptance in the industry, the method for assessing the maturity index should be easy to use, inexpensive, and reliable. In order to reduce the subjectivity of the current method for determining the maturity index of the olives, this work aims to provide an automatic method that integrates a machine vision system for the classification of the olives. This is done by classifying the olives according to the maturity classes by an image segmentation algorithm based on color and edge detection. In order to improve the agreement of predicted values through computer vision, this work combines the information obtained by color analysis and MI prediction of individual olive fruits. The proposed methodology also provides more quality information such as the size and the average weight of the olives and could be useful in improving the performance of commercial systems allowing classifying and removing olives that do not meet the

required standards and then improving a better control of the extraction process of olive oil.

## **2. Materials and methods**

### **2.1. Samples**

Olive samples of Picual variety (2 kg) were collected in several olive mills in the province of Jaen (Spain) during November 2011. The colour of these olives ranged from dark green to completely black. These samples were used for the acquisition of images, synthetic samples were also carried out with the aim to incorporate into the analysis samples of olive with a full range of maturity index.

### **2.2. Visual determination of Maturity Index (MI)**

The RI was determined by an experienced evaluator every sampling day with an assessment of skin and flesh color of olives [7]. The procedure involved the distribution of olives into eight groups according to the characteristics summarized in Table 1. The official method uses 100 olives and calculates a global value of maturity index for each sample or group of olives.

The MI is calculated by using the following equation:

$$MI = \sum_i n_i / 100$$

where  $i$  is the group number and  $n_i$  the number of olives in it.

This method has to manually separate the olives, cut the pulp to examine them and count and identify the group to which they belong. This method also includes calibrations of other measures of fruits such as weight, size and diameter.

<b>Maturity index Group</b>	<b>Description</b>
<b>0</b>	Skin color deep green
<b>1</b>	Skin color yellow-green
<b>2</b>	Skin color with < half the fruit surface turning red, purple or black
<b>3</b>	Skin color with > half the fruit surface turning red, purple or black
<b>4</b>	Skin color all purple or black with all white or green flesh
<b>5</b>	Skin color all purple or black with < half the flesh turning purple
<b>6</b>	Color all purple or black with > half the flesh turning purple
<b>7</b>	Skin color all purple or black with all the flesh purple to the pit

**Table 1.** MI classification groups

## 2.2 Image System

Images were acquired with a JAI AD-080CL multi-spectral camera, combining a visible color channel (Bayer mosaic CCD) which can generate 24-bit RGB images and a NIR (Near Infrared) channel (monochrome CCD). The major advantage of this camera is the capability of capturing both channels simultaneously though the same optical path.

This camera uses a standard Camera Link interface, whereby each channel can output images with 8 or 10-bit and 24- bits, and a resolution of 1024x768 active pixels per channel. This camera was installed in an enclosed cabin equipped with a controlled lighting to achieve a consistent image in all acquisitions. The lighting was a halogen lamp with some filters to get a diffused light simulating the illuminant D65 (6,500 K), which is given as standard by the International Commission on Illumination, simulating the lighting conditions of a cloudy day at midday. Direct lighting on olive fruit was avoided by means of a reflecting surface placed between the scene and the illuminant. The images were taken at a distance of approximately 45 cm with a neutral white background.

### 2.3 Methodology

The algorithms used were developed with the image processing toolbox, version 7.4.0 of Matlab (The MathWorks, Inc., Natick, MA, USA) along with software Image Pro-Plus version 6.0 (MediaCybernetics, Inc).

The images were taken in duplicate for both channels, IR (monochromatic) and visible (RGB image), which are images whose pixels are specified by three values, one for each color component (red, green and blue).

#### 2.3.1 Estimation of Maturity Index

The technique used was segmentation, which is based on the identification of regions and edges using clusters of pixels selected according to different criteria (color, boundary, textures, etc.) [46, 47].

In order to develop a method for the automatic prediction of MI, the following steps have been carried out:

##### 2.3.1.1 Identification and separation of objects in the NIR image

The procedure of separation of objects consists on the use of the NIR monochrome image in order to segregate every single olive fruit. To identify objects in the NIR monochrome image, a first step performed a contrast equalization followed by a spatial filter enhancement 'Flatten', which reduces the intensity variations in the

background pixels. Spatial filtering operation was applied to the images to enhance or attenuate spatial detail in order to enhance visual performance or to facilitate further processing. The spatial filtering operation can be considered as a "local" procedure in image processing, in the sense of changing the value of each pixel according to the values of the surrounding pixels. From a practical point of view this steps consists mainly on classifying the original gray levels in the original NIR image according to the value in the neighboring pixels. This task was carried out using two algorithms: Border-4 Neighbor that determine the edges of the olive taking into account the intensity values of pixels in the histogram and considers the darker pixels as edges of objects ; b) Connected components algorithm that is used to identify each of olives with a value. This edge-based algorithm enumerates each set of pixels that make up an olive [48].

#### 2.3.1.2 Creating a mask and application of the mask to the color visible image

From the data in the array of connected components, can create a mask, on which also can apply various size filters so that we can measure complete objects or automatically removes the objects that overlap with one area less than or greater than those with an standard size olive [49]. Based on these objects is constructed the skeleton of image. The skeleton is intended to represent the shape of an object with a relatively small number of pixels. Thus, all pixels of the skeleton are structurally necessary. The position, orientation and length of the skeleton lines correspond to those equivalents of the original image, this skeleton simplifies the task of extracting features of an image.

Over this mask which reduces the image to a skeleton and eliminates protrusions, is added the RGB image on which the color-segmentation is performed over separated objects.

#### 2.3.1.3 Segmentation based on color and classification

This segmentation based on color can be performed by supervised or unsupervised methods in order to get and quantify the predominant color in the olives.

This step requires transforming the RGB image using various functions for transforms RGB format in  $L^*a^*b^*$  color space. The CIE  $L^*a^*b^*$  colour notation system was applied to determine the parameters  $L^*$ ,  $a^*$  and  $b^*$ ; where  $L^*$  indicates the lightness,  $a^*$  means the colour axis from green to red and  $b^*$  the blue-yellow one.

In the supervised methods a region containing the color of interest (color markers) is selected and then averaged. This classification is performed by using the method of the k-nearest neighbors (KNN) [50], where each pixel is classified in the same class as the color markers with a similar intensity. The k nearest neighbors assigned a value of 'a' and 'b' for each marker, then it is possible to classify each pixel in the image to calculate the Euclidean distance between pixels and color markers. Unsupervised methods for color segmentation techniques use clustering algorithms that essentially perform the same functions as classifiers methods, but without the use of training data. [51, 52] To compensate for the lack of training data, clustering methods iterate between image segmentation and characterize the properties of each class. In this sense, clustering methods are trained with the available data. The basic idea is to assume that image pixels are points in a three dimensional space (RGB), so that the points of a region of similar color are grouped together. Clustering techniques allow obtaining a representative point of each group, based on measures of similarity between these points. In this algorithm a point  $x_i$  is assigned to a group R whose centroid C is closer to it. Euclidean distance is usually used as a measure of similarity between the point and the centroid. The results of this classification are normally expressed in terms of the total percentage of pixels colored in each object. With these results fruits can be classified according to their level of maturity (Table 2).

Class	Olive skin Color
0	>50% Bright green
1	>50% Green-yellowish
2	%Green-yellowish with black and/or reddish %Reddish-brown
3	%Green-yellowish with black and/or reddish %Reddish-brown
4	100 % Blackish purple or black

**Table 2.** Olive classification based on olive skin

### 2.3.2 Estimation of size and weight

Another part of the study was to determine the size and weight of the fruit by image analysis. To do this, a calibration curve for different weights of the fruit and their respective areas in the image pixels has been performed using a group of training samples and then validated with a set of external validation samples.

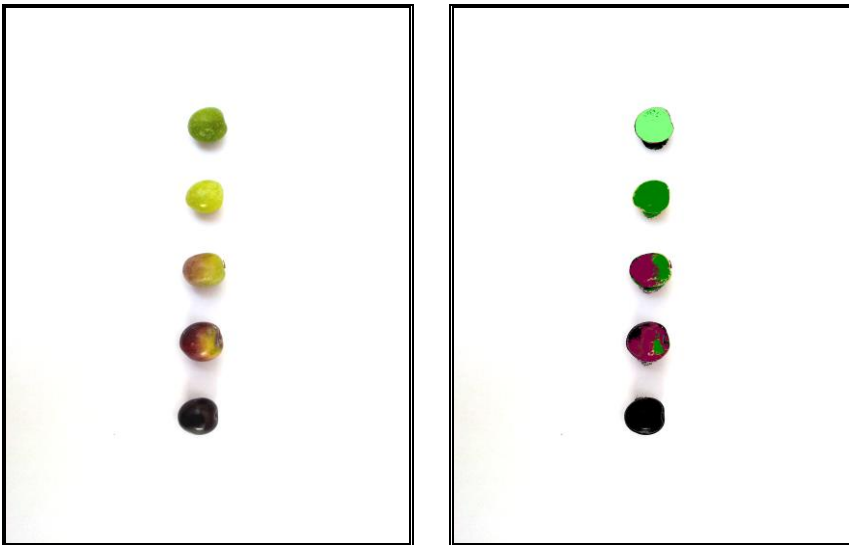
## 3. Results and Conclusion

### 3.1. Estimation of Maturity Index

All the available sets of olives were analyzed and for each batch of olives, the ripeness index was calculated according to the procedure described in the previous section.

The procedure was performed first in 40 synthetic images with separate olives, and applied these algorithms to obtain the results mentioned maturation index for each fruit, the results are obtained as the percentage of each color in the object and is classified the different classes. The effectiveness of the method was demonstrated in the Picual olive variety, olives of this variety used presents great variability in color and size. .

This test phase with simple images showed that the values provided by the vision system are very similar to those provided by the traditional method based on visual determination used as reference. Table 2 shows some results obtained by the application of the proposed procedure in images containing olives with different maturity index. And figure 1 shows an original image and the image after the analysis, results showing each object class with a pixel color



**Figure 1.** An original image and the image after the analysis results showing each object class with a pixel color (black=% black; brown=%Reddish-brown; green=% green)

Object	%	%	%	%	Estimated	Measured
	Bright-Green	Green-yellowish	Redish-Brown	Black	MI	Visual MI
1	100	0	0	0	0	0
2	57	43	0	0	0	1
3	33.3	66.7	0	0	1	1



4	0	100	0	0	1	1
5	1	99	0	0	1	1
6	0.6	99.4	0	0	1	1
7	2.8	97.2	0	0	1	1
8	0.7	99.3	0	0	1	1
9	0	40.4	59.6	0	3	3
10	0	67.4	32.6	0	2	2
11	0	33.6	66.4	0	3	3
12	0	51.9	48.1	0	2	2
13	0	60.6	39.4	0	2	3
14	0	49.9	50.1	0	3	3
15	0	0	100	0	3	3
16	0	31.3	68.7	0	3	3
17	0	0	0	100	4	4
18	0	0	0	100	4	4
19	0	0	0	100	4	4
20	0	0	0	100	4	4

**Table 2.** Results of the maturation index estimated and visually calculated of some samples analyzed

Most of the errors found in the pixel segmentation procedure are due to isolated or small clusters of pixels, mainly located at the boundaries of adjacent regions.

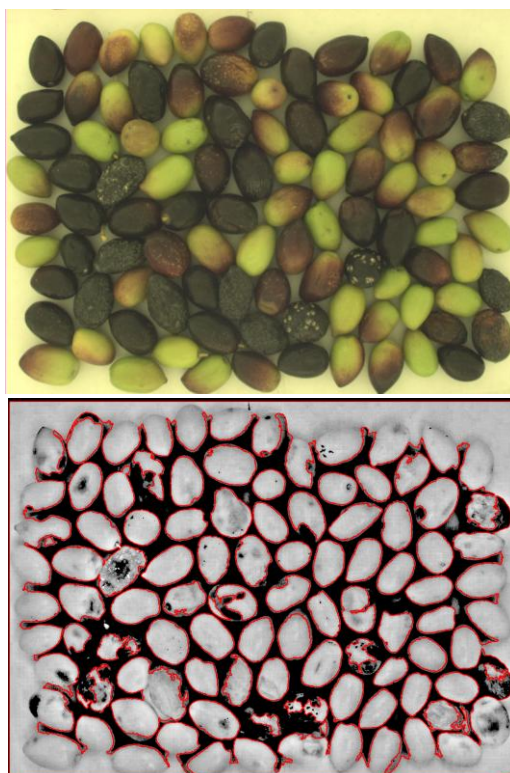
These groups of pixels may be easily detected and discarded, as these groups do not form broad regions but rather discontinuous are divided into segments, so that can be discarded due to their small area, using the object's center area where the color classification is applied.

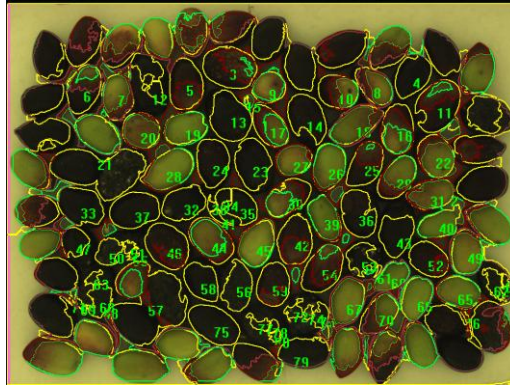
The results of prediction maturity index of fruit by the proposed vision method for images with individual olives are strongly related to visual RI measured.

The errors of prediction can be almost negligible or zero values for other samples (Results in Table 2), these values are acceptable given the subjectivity that characterizes the MI evaluation by the traditional method.

These protocols were applied to 15 complex images with overlapping olives. Around 100 olives in each image were analyzed and an average of 92% were correctly detected as separate objects and analyzed separately by the program. Figure 2 shows an example of analyzed images.

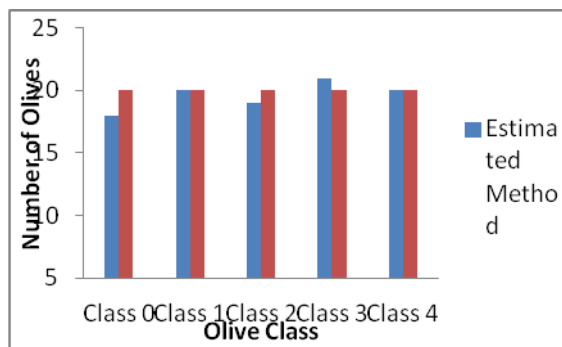
Therefore, this protocol allows effective analysis of samples and can substantially improve these results through the improvement and standardization of the lighting conditions in the samples.





**Figure 2.** a) Representation of sample in the RGB space, b) image after segmentation and c) the image obtained after applying the complete procedure

The results of the analysis of these 15 images are shown in Figure 3, where it is possible to compare the results obtained with both methods.

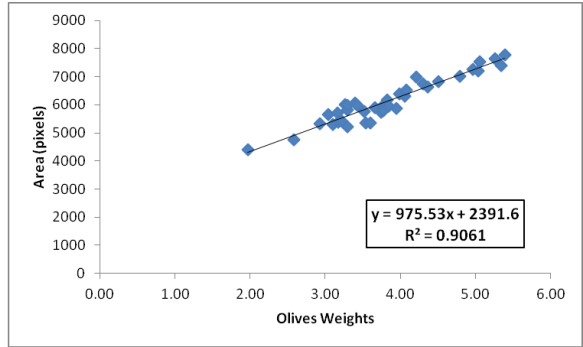


**Figure 3.** Histogram representation of the results of an analyzed image

These results confirm the high correlation between both methods as the difference values provided by the vision system and the classical visual method are very low and acceptable given the high subjectivity of the method used as reference.

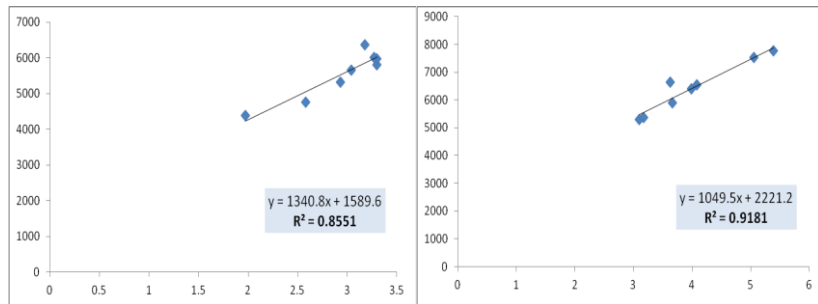
### 3.2. Estimation of olive size and weight

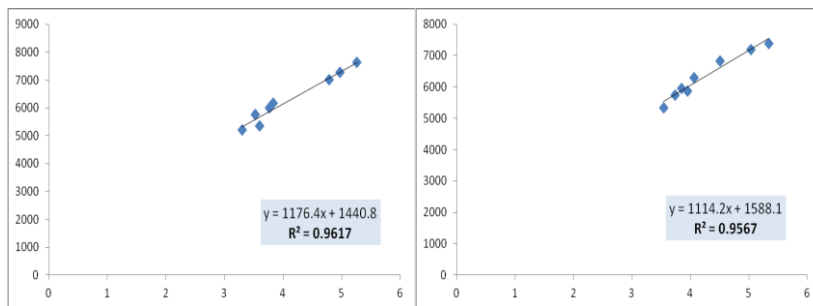
This part of the work aims giving an estimate of the fruit weight by the area calculated in pixels in the image processing. The real weights were previously measured in 40 samples of olives and a positive correlation with the area occupied by olives in pixels with a determination coefficient of 0.91 (Fig 4).



**Figure 4.** Correlation between olives areas in pixel and real olive weight measured

These results show the statistical dependence between the weight of fruit and visual images observed in the sample as the measured area. The error values can be considered higher than desirable because in this line is not taken into account the type of ripeness of samples. The weight of the olives can depend on the degree of maturity of the fruit, as well as the amount of water and oil also increases until it reaches its maximum in riper olives. In the next step was constructed a calibration line for each class of maturation as shown in Figure 5. Figure 5 shows the calibration lines for various classes of maturation.





**Figure 5.** Correlation between olives areas in pixel and olive weight for maturity index 1, 2, 3 and 4 respectively

In order to validate the calibration data using external samples, the weight of 10 samples of olives was measured. The results of validation using shown in Table 3.

Object	Real Weight	Estimate Weight	Estimate Weight	Area Pixel	Relative Error	Relative Error
		Global Calibration	Parcial Calibra		(GC)	(PC)
1	3.16	3.4	3.1	5717	0.08	0.02
2	3.4	3.8	3.3	6054	0.12	0.03
3	2	1.8	1.9	4111	0.10	0.05
4	3.43	3.7	3.6	5977	0.08	0.05
5	4.21	4.7	4.5	6986	0.12	0.07
6	3.22	3.1	3.4	5406	0.04	0.06
7	4.3	4.5	4.5	6740	0.05	0.05
8	3.77	3.5	3.8	5786	0.07	0.01
9	4.36	4.4	4.5	6658	0.01	0.03
10	3.3	3	3.3	5274	0.09	0.00

**Table 3.** Data validation samples and relative errors values for both calibrations

Therefore, it is important to check the influence of the maturity to determine the size and weight of the fruit. Using the appropriate calibration line for each sample of olives, this method can be a tool very useful to assess quickly and accurately fruit maturity with a relative error not more than 0.07, thus, determination of size was properly solved. From visual and geometric characteristics of the fruits can develop an automatic vision system of fruit weight to improve the level of information to farmers.

This method could be useful for the selection of the date of collection, to track changes in the color of the fruit on the tree until the selected value. Therefore, the time of harvesting the olives should not depend on a subjective decision, but a uniform measure, a method rapid and non destructive.

The proposed methodology has the potential to be used for the classification of fruit in the quality control of the extraction process of olive oil. Indeed, the results of this study suggests in its implementation and on-line consisting of image acquisition and analysis can occur at acceptable speeds, this shall be the future aim of studies.

The described method provides an estimate of the MI value, this procedure require a negligible amount of time, it is possible to state that the only time required is the one to collect the olive samples from the trees. This is important when the experts working in oil mills perform a real-time monitoring of the agronomical and technological parameters. In some cases, however, the purpose of oil mills is to increase the quality of the extracted olive oil by optimizing some technological parameters.

The development of classification systems for non-destructive methods supported make application to each fruit and ensure that their characteristics

meet the standards. In particular optical techniques offer possibilities in the assessment of surface characteristics and inner qualities or composition. The systems used to measure the optical properties, provide good information on the characteristics of each fruit, useful for assessment of products.

## Acknowledgment

The authors thank the National Institute of Research and Technology Agriculture and Food (INIA) for funding of the scholarship of training FPI sub-INIA, ESF co-financed through the ESF Operational Programme for Andalusia 2007-2013 within axis 3 "Increasing and improving human capital" expenditure category "Developing human potential in the field of research and innovation."



## References

- [1] Uceda, M, Frias L. Harvest dates. Evolution of the fruit oil content, oil composition and oil quality. In: Proceedings of the II Seminario Oleícola Internacional; International Olive Oil Council: Cordoba, Spain, pp 125–130 (1975)
- [2] Isabel Mínguez-Mosquera, Lourdes Gallardo-Guerrero. Disappearance of chlorophylls and carotenoids during the ripening of the olive. *Journal of the Science of Food and Agriculture* Volume 69, Issue 1, pages 1–6, September (1995)
- [3] Roca, M, Mínguez-Mosquera . Changes in chloroplast pigments of olive varieties during fruit ripening. *Journal of agricultural and food chemistry* 49 , 2, 832-839 (2001)
- [4] Ting SV, Rouseff RL. In citrus Fruits and their products. *Analytical Technol* 3,1–9 (1986)
- [5] Casas A, Mallent D. *Agrochim Technol Aliment* 28(2):184–201(1998)
- [6] Gnanasekharan V, Shewfelt RL, Chinnan MS. *J Food Sci* 57(1):149–154 (1992)
- [7] Garcia JM, Seller S, Perez-Camino MC. *J Agric Food Chem* 44:3516–3520 (1996)
- [8] Tovar MJ, Romero MP, Alegre S, Girona J, Motilva MJ. *J Sci Food Agric* 82:1755–1763 (2002)
- [9] Baccouri, B., Wissem, Z., Dhouha, K., Issam, N., Nabil, B. Influence of fruit ripening and crop yield on chemical properties of virgin olive oils from seven selected oleasters (*Olea europea L.*). *Journal of Agronomy* 6 (3), 388–396 (2007)
- [10] Taylor MA, Rabe E, Jacobs G, Dodd MC. *Post Biol Technol* 5:285–294 (1995)
- [11] Lehman-Salada L. *HortScience* 31(6):994–997 (1996)
- [12] Ranalli A, Tombesi A, Ferrante ML, de Mattia G. *RevIta delle Sostanze Grasse* 74:553–557 (1997)

- [13] Garc\_a JM, Guti\_rrez F, Barrera MJ, Albi MA. J Agric Food Chem 44:590–593 (1996)
- [14] Olmo M, Nadas A, Garc\_a JM. J Food Sci 65:365–369 (2000)
- [15] Garc\_a JM, Medina RJ, Ol\_as, JM. J Food Sci 63(6):1037–1041(1998)
- [16] Muramatsu N, Sakurai N, Yamamoto R, Nevins DJ. HortScienceS 31(7):1199–1202 (1996)
- [17] Mickelbart MV, James D. New Zealand J Crop Hort Sci 31:269–276 (2003)
- [18] Non-destructive assessment of olive fruit ripening by portable near infrared spectroscopy. Gracia, A. Leon, L.. Grasas y aceites, 62, 3, 268-274 (2011)
- [19] Cayuela, J.A., García, J.M, Caliani, N. NIR prediction of fruit moisture, free acidity and oil content in intact olives. Grasas y aceites, 60 (2), 194-202 (2009)
- [20] Bellincontro, A. Taticchi, A. Servili, M. Esposto, S. Farinelli, D. Mencarelli, F. Feasible Application of a Portable NIR-AOTF Tool for On-Field Prediction of Phenolic Compounds during the Ripening of Olives for Oil Production. Journal of agricultural and food chemistry, 60, 10, 2665-2673 (2012)
- [21] Ana Morales-Sillero, Víctor-Manuel Fernández-Cabanás, Laura Casanova, María-Rocío Jiménez, María-Paz Suárez, Pilar Rallo. Feasibility of NIR spectroscopy for non-destructive characterization of table olivetraits. Journal of Food Engineering ,107,1, 99-106 (2011)
- [22] Salguero-Chaparro, , Baeten. V., Abbas, O., Peña-Rodríguez, F. On-line analysis of intact olive fruits by vis–NIR spectroscopy: Optimisation of the acquisition parameters. Journal of Food Engineering, 112, 3, 152–157 (2012)
- [23] Cordero, S., Lleo, L., Barreiro, P., Ruiz-Altisent, M. Peach multispectral images related to firmness and maturity. In: World Congress: Agricultural Engineering for a Better World. Bonn, Germany (2006)
- [24] Blasco, J., Aleixos, N., Molto, E. Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm. Journal of Food Engineering 81, 535-543 (2007)
- [25] Kondo, N , Ahmad, U., Monta, M., Murasc, H. Machine vision based quality evaluation of Iyokan orange fruit using neural networks. Computers and Electronics in Agriculture 29, 135-147 (2000)
- [26] Rosenberger, C, Emile, B., Laurent, H. Calibration and quality



- control of cherries by artificial vision. *Journal of Electronical Imaging* 13 (3), 539-546 (2004)
- [27] Uthaisombut, P (1996). Detecting defects in cherries using machine vision. M.S. Thesis, Computer Science, Michigan State University.
- [28] Cheng, X., Tao, Y., Chen, Y.R., Luo, Y. NIR/MIR dual-sensor machine vision system for online apple stem-end/calyx recognition. *Transactions of ASAE* 46 (2), 551-558 (2003)
- [29] Kavdir, I., Guyer, D.E. Comparison of artificial neural networks and statistical classifiers in apple sorting using textural features. *Biosystem Engineering* 89 (3), 331-344 (2004)
- [30] Mehl, P.M., Chen, Y.R., Kim, M.S., Chan, D.E. Development of hyperspectral imaging technique for the detection of apple surface defects and contaminations. *Journal Food Engineering* 61, 67-81 (2004)
- [31] Abdullah, M.Z., Mohamad-Saleh, J., Fathinul-Syahir, A.S., Mohd-Azemi, B.M.N. Discrimination and classification of fresh-cut starfruits (*Averrhoa carambola* L.) using automated machine vision system. *Journal of Food Engineering* 76 (4), 506–523 (2006)
- [32] Koc, A.B. Determination of watermelon volume using ellipsoid approximation and image processing. *Postharvest Biology and Technology* 45 (1), 366–371 (2007)
- [33] Brosnan, T., Sun, D.W. Improving quality inspection of food products by computer vision-a review. *Journal of Food Engineering* 61, 3-16 (2004)
- [34] Pla, F., Sanchez, J.S., & Sanchiz, J.M. On-line machine vision system for fast fruit colour sorting using low-cost architecture. In: VIII SPIE Conference on Machine Vision Systems for Inspection and Metrology, Boston, Massachusetts, USA (1999)
- [35] Mendoza, F., Aguilera, J.M. Application of image analysis for classification of ripening bananas. *Journal of Food Science* 69 (9), 471–477 (2004)
- [36] Oberti, R., Guidetti, R., Mignani, I. Analisi multispettrale di immagini per l'individuazione precoce di danni meccanici su prodotti ortofrutticoli:

un'applicazione su pere decana. *Rivista di Ingegneria Agraria* 30 (3), 137–147 (1999)

[37] Peri, G., Romaniello, R., Amodio, M., & Colelli, G. The application of a fast colour quantization algorithm for sorting peaches for firmness. In: XXX CIOSTACGIR V CONFERENCE, Torino, Italy (2003)

[38] Stajanko, D., Cmelik, Z. Modelling of apple fruit growth by application of image analysis. *Agriculturae Conspectus Scientificus* 70 (1), 59–64 (2005)

[39] Bodria, L., Fiala, M., Guidetti, R., & Oberti, R. Optical techniques for assessing the fruit maturity stage. In: ASAE Annual International Meeting/CIGR XVth World Congress, Chicago, USA (2002)

[40] Choong, T.S.Y., Abbas, S., Shariff, A.R., Halim, R., Ismail, M.H.S., Yunus, R., Salmiaton, A., Fakhru-Razi, A., 2006. Digital image processing of palm oil fruits. *International Journal of Food Engineering* 2 (2), 7 (2006)

[41] Diaz, A., Faus, G., Blasco, M., Molto E. The application of a fast algorithm for the classification of olives by machine vision. *Food Research International* 33 , 30-309 (2000)

[42] R. Diaz L. Gil C. Serrano M. Blasco E. Moltó J. Blasco Comparison of three algorithms in the classification of table olives by means of computer vision *Journal of Food Engineering* 61 (2004) 101–107

[43] Riquelme, M.T., Barreiro, P., Ruiz-Altisent, M., Valero, C. Olive classification according to external damage using image analysis. *Journal of Food Engineering*, 87, 371–379 (2008)

[44] Ram, T, Wiesman, Z., Parmet, I., Edan Y. Olive oil content prediction models based on image processing. *Biosystems engineering* , 105, 221-232 (2010)

[45] Furferi, R., Governi, L. Volpe, Y. ANN-based method for olive Ripening Index automatic prediction. *Journal of Food Engineering* 101, 318–328 (2010)

[46] Medioni, G.; Lee, M-S; Tang, C.K. A Computational Framework for Segmentation and Grouping, Elsevier Science B.V., 2000. 0-444-50353-6 (2000)

[47] Plataniotis, K.N.; Venetsanopoulos, A.N. *Color Image Processing and Applications*, Springer Verlag. 3-540-66953-1 (2000)

[48] Chang, V. Saavedra, J. Metodos alternativos para el mejoramiento automatico del contraste de imágenes. Computer Engineering thesis, School of Informatic, University National of Trujillo (2001)

[49] Russ, J.C. *The image processing Handbook*. CRC Press LLC 3rd Edition.

[50] Richards, J. A.. *Remote Sensing Digital Image Analysis*. Ed. Springer-Verlag (1999)

[51] Thomas, S.W., 1991. Efficient inverse color map computation. In: Arvo, J. (Ed.), *Graphics Gems II*. Academic Press, Boston, pp. 116–125 (1991)

[52] Floyd, R. W., & Steinberg, L. An Adaptive Algorithm for Spatial Gray Scale. In: International Symposium Digest of Technical Papers, Society for Information Displays, 36 (1975)

**Automatic and on-line method for detection of  
dirt on olive oil production process.**

**Food Science and Technology (enviado)**

# **Automatic and on-line method for detection of dirt on olive oil production process**

Elena Guzmán<sup>a</sup>, Vincent Baeten<sup>b</sup>, Juan Antonio Fernández Pierna<sup>b</sup>, José A. García-Mesa<sup>a</sup>

a Centro IFAPA 'Venta del Llano', Crta. Nacional Bailén-Motril Km 18.5, 23620 Mengíbar, Jaén, Spain

b Walloon Agricultural Research Centre, Valorization of Agricultural Products Department, Food and Feed Quality Unit, Henseval Building, 24 Chaussée de Namur, 5030 Gembloux, Belgium

Corresponding author E-mail address: elena.guzman.jimenez@juntadeandalucia.es (E. Guzmán). Phone: 34953366366

## **Abstract**

Currently, there are no methods for automatic detection of dirt in food production factories; the dirt count is most commonly determined by manual methods and visual inspection.

This work is aimed at proposing the development of a new method to detect dirt on-line, as well as to measure the percentage of dirt on a sample of the oil in the manufacturing line.

Good results were obtained in the detection of dirt in the olive samples analysed, giving the results as a percentage of dirt with regards to the total olives in the sample. Visual and numerical results indicate that the vision system method designed is highly efficient in order to detect particles of dirt, showing that results are not significantly different from those calculated manually

The developed system is possibly a first step towards an evaluation on-line in the real dirty olive samples. This method allows automatic, quick and objective prediction of the percentage of dirt present in the olive samples which enter into the production process.

## 1. Introduction

In the process of olive oil production we can observe different factors influencing the final product quality. The harvest and transport stage is one of the most important since a great loss of oil quality depends on it. Olives harvested from the tree are of better quality than those fallen on the ground. During the harvest, both types of olives should be kept in separate containers and transported to the mill in different ways. Olives harvested from the ground, something which occurs easily in the season, carry a great amount of soil or stones and, therefore, must be thoroughly washed as soon as possible. At the time of entry, olives should be classified according to the first treatments to which they are subjected, i.e. cleaning and washing. There will be at least two distinct discharge points: one for dirty olives, i.e. being muddy with soil or stones and another one for clean olives. Then, they must pass through the cleaning and the washing process. Later, part of the impurities which accompany the olive harvest, mainly leaves and twigs, must be removed. They can represent between 5% and 10% by weight of fruit at the time entering the mill. High-performance systems have self-cleaning processes and elements for removal of dirt, such as the removal of leaves, twigs, dust, etc. If olives are heavily soiled with mud or dirt, they should be washed with potable water. This process is not necessary when olives have been harvested from the tree. Consequently, for this and other reasons, it is important to clean and classify olives. However, this is a step which increases the time of reception and in the case of clean olives, it is not advisable. Washing clean olives is of doubtful interest because the operations of extraction of residual water lead to a decrease in the increased extractability 0.3% of fat rich, a loss of stability and a lower intensity of the sensory attributes of oils, especially in the fruit [1]. Finally, after cleaning, the fruit is transferred to the weighing balance.

There is no evidence in the current literature of any on-line method for assessing the level of contamination in the processing plants.

For these reasons, it is necessary a classification method for olives according to their level of dirt, which would directly obtain the percentage of oil and dirt found in each sample. This would also facilitate their separation during the reception time, save time,

reduce the area required for the olive reception and basically simplify the step of olive reception.

In order to solve this problem, this paper provides an automatic method which integrates a machine vision system for the evaluation of dirt in olive samples by analysing the image segmentation algorithms based on colour detection, edges and morphological segmentation.

Artificial vision has found many applications in the food industry for food quality controls. This method allows non-destructive inspection to be fast, inexpensive and objective, which has been applied in different fields of industry. This technique is notable for its speed and accuracy, and because thanks to it, production is increased and quality requirements are ensured. Therefore, it represents a great help in the development of fully automated processes [2].

In recent years, methods based on instrumentation using image processing have been designed to perform real-time control and automated inspection, such as cameras and spectrometers which allow the detection of fruit characteristics. By interposing appropriate filters, we will achieve the analysis of non-visible areas which let us understand some characteristics of products which are not always detected visually by operators.

Among some of the applications developed using computer vision as a technique for classifying fruit, Brosnan and Sun [3] present a comprehensive review of image processing techniques for various food products.

Other studies have shown the ability to sort the fruit by its colour [4,5,6,7]. Moreover, there are some ranking algorithms based on fruit shape and size: in cucumber [8] and watermelon [9]. The literature presents many papers related to the correlation between the image and the degree of maturity of the different fruit [10, 11] in order to remove products which do not meet the required standards of classification.

Previous spectroscopy has been used as a rapid non-destructive method for determining the quality of olives in the process [12]. Nevertheless, this paper aims to improve the application of techniques for fast on-line analysis of olive oil and proposes

a direct method to determine on-line the level of dirt in the samples of olives. This non-destructive method offers many advantages in the quality control of the process, since it is an easy and quick way of obtaining data, which is difficult to obtain manually. The final methodology could be useful for evaluating the performance of commercial systems and better control of the extraction process of olive oil.

## **2. Materials and methods**

In order to develop a method to detect dirt automatically, the following tasks have been carried out:

- Data collection.
- Preliminary quantification of dirt by visual manual method.
- Image acquisition.
- Algorithms development and preliminary prediction of the amount of dirt by defining an Artificial Vision Method

### **2.1. Data collection and quantification of dirt by visual manual method**

Ninety real samples of picual variety olives were collected and analysed during the entire 2011-2012 olive season in the IFAPA Centre "Venta del Llano" (Mengibar, Jaén). Other samples collected in some factories in the province of Jaen were also used.

The amount of dirt (leaves, stones, mud) was weighed and counted in each sample in order to obtain reference values of the amount of dirt present in each image and to validate the method of detection by artificial vision. The results were obtained as a percentage of total leaves, stones and olives in each sample. Therefore, we can compare the results obtained with the vision method proposed and the real content of dirt in each sample and we can obtain a calibration which permits an estimation of weight or amount of dirt for external samples.



## 2.2. Image acquisition and algorithm development

Images were captured with a digital CCD (Charge-Coupled Device) Progressive Scan Camera JAI AD-080CL which has a visible colour channel (Bayer mosaic CCD). This system can generate 24-bits RGB images and use a standard Camera Link interface with a resolution of 1024x768 active pixels per channel. This camera was installed in an enclosed cabin with controlled lighting in order to achieve a consistent image in all captures. The lighting used was a halogen lamp with some filters to get a diffused light simulating the illuminant D65 (6,500 K). This is considered the standard by the International Commission on Illumination, which describes the ordinary lighting conditions at noon, thus avoiding glow effects on the fruit. The images were taken at a distance of approximately 45 cm in a neutral white background. For on-line application of this method, these devices were installed in the processing line, specifically, on the belt which transports the olives into the mill, where the lighting conditions are the same as described above. This high speed camera can take high quality pictures of the fruit, reaching motion captures up to 30 frames per second with full resolution.

The algorithms used were developed with the following software of image analysis: Image Pro-Plus version 6.0 (MediaCybernetics,Inc).

These images were taken in triplicate in RGB format; this means that they are images whose pixels are specified by three values, one for each colour component (red, green and blue) in each pixel.

The technique used was the segmentation, which allows us to group the pixels around the elements of the image pixels based on the techniques of identification of regions, edges, colour-based and morphological segmentations [13,14].

The segmentation based on colour can be performed by two different methods: either by supervised or unsupervised method in order to obtain and quantify the predominant colour in each of the element present in the image of olives. This requires transforming the RGB image using the functions 'makecform' and 'applycform' in L\*a\*b colour space.

On the one hand, in the supervised method, it is necessary to determine the colours of interest that you want to quantify in each image. The colour of interest in the original image is selected with the function 'roipoly'. This function takes a sample of the region containing the colour of interest, as well as it takes an average colour of each region through the 'median' function, used to identify the colours of the rating label (colour markers). This process is performed thanks to the method of nearest neighbour (KNN- k nearest neighbour) [15], in which each pixel is classified in the same group as the data training with a higher intensity. The k nearest neighbours assign a value of 'a' and 'b' for each marker and, therefore, it is possible to classify each pixel of the image so as to calculate the Euclidean distance between pixels and colour markers.

On the other hand, the method for unsupervised colour segmentation techniques uses grouping (clustering) [16,17]. Clustering algorithms essentially perform the same functions as the classifying methods, but without using data training. This is why they are unsupervised methods. In order to compensate for the lack of training data, clustering methods iterate between image segmentation and characterize the properties of each class. In this sense, clustering methods are trained with the available data. In this algorithm, the basic idea is to assume that image pixels are points of a three dimensional space (RGB). As a result, the points of a region of similar colour are grouped together. Clustering techniques allow us to obtain a representative point of each group, based on measures of similarity or distance between these points. In this algorithm, a point  $x_i$  is assigned to a group R, whose centroid C is closer to it. This is usually used as a measure for the Euclidean distance between the point and the centroid.

The results provide the percentage of pixels grouped by colour and, thus, we can obtain the total area of pixels occupied by dirt and the area occupied by olives. This way, we can compare the real amount of dirt and the visually quantified and measured amount of dirt in the sample estimated by the artificial vision method.

### 3. Results and Discussion

In order to develop and examined the approach proposed, a series of images of olives were analysed. The amount of dirt in each image of olives was calculated by weight and number of particles according to the procedure described in the previous section.

The procedure was performed in ninety images of olives, i.e. the algorithms were applied in order to get the results of the amount of dirt, showing the percentage of leaves, stones and olives in each image. The effectiveness of this method has been shown in the picual variety olive, whose colour and size present great variability. Consequently, this method could be potentially used with other varieties.

The content in leaves, stones and olives in each image is known and, therefore, this value can be related to the area in pixels obtained by the artificial vision method. Due to the fact that each of these dirt particles have a different ratio area / volume, it is desirable to construct a calibration model for each one. In this sense, images can be analysed separately with leaves, stones and mud in order to obtain more accurate predictive models. All images were captured in triplicate and used as the average value of the three replicates.

In the first place, 30 images were analysed with varying amounts of leaves in order to obtain the calibration model. Figure 1 shows an original RGB image, as well as the image after the analysis. The result indicates the total percentage of pixels occupied by leaves in the whole image.



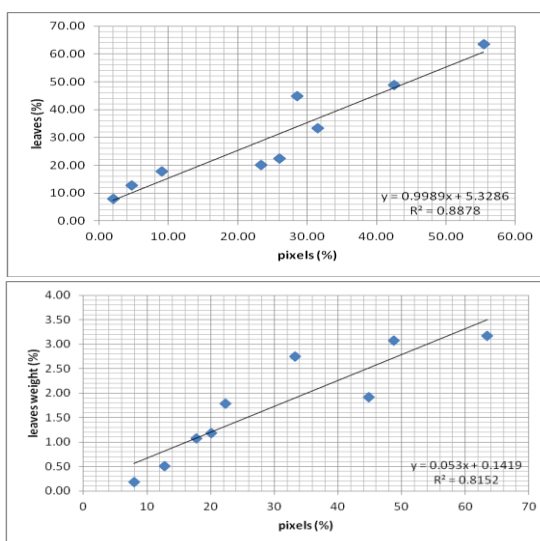
**Figure 1.** Original image (a) and image after the colour based segmentation, showing the results expressed in the pixel area occupied by leaves (b).

The results of some analysed images and real values of leaves were measured in terms of weight and number of leaves in each image. Table 1 shows the real value in terms of weight and number of leaves, as well as the values obtained by the software after applying segmentation and standard deviation.

Sample	% Weight leaves Manual Measure	% Leaves Manual Measure	% Measured Vision Method
Image AV	0.18	2.04	7.97±0.67
Image AV	0.50	4.76	12.72±3.92
Image AV	1.07	9.09	17.73±7.36
Image AV	1.18	23.33	20.11±6.06
Image AV	1.79	26.04	22.34±3.23
Image AV	1.91	28.59	44.83±17.81
Image AV	2.75	31.52	33.22±11.52
Image AV	3.08	42.58	48.74±15.23
Image AV	3.17	55.41	63.44±2.05

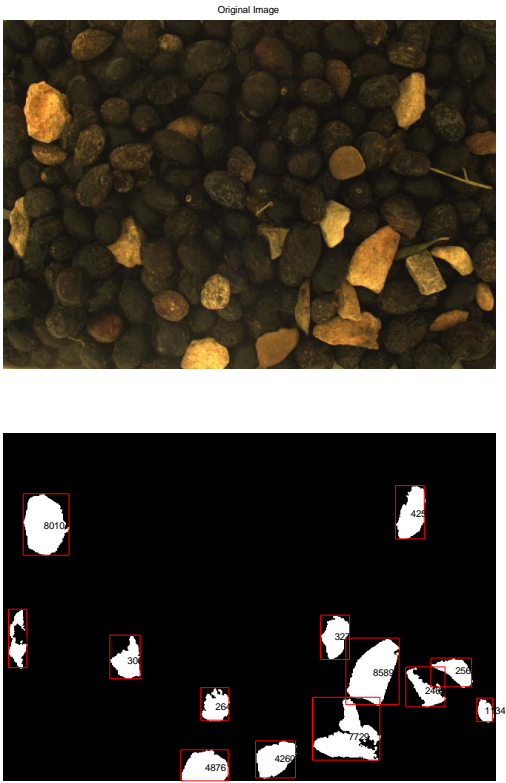
**Table 1.** Results of some images obtained by the software after applying segmentation algorithms to assess the amount of leaves in a set of images of olives.

Thanks to the analysis of these results, a calibration curve was built in order to estimate the weight or number of leaves according to the area in pixels obtained by the software (Figure 2).



**Fig 2:** correlation between results obtained by both methods in percentage of leaves (a) and weight of leaves (b).

The same procedure was conducted for stones and mud; other 30 images with different amounts of stones were analysed so as to build a calibration model. Figure 3 shows an original RGB image, as well as the image after the analysis. The result indicates the total percentage of pixels occupied by stones in the whole image.



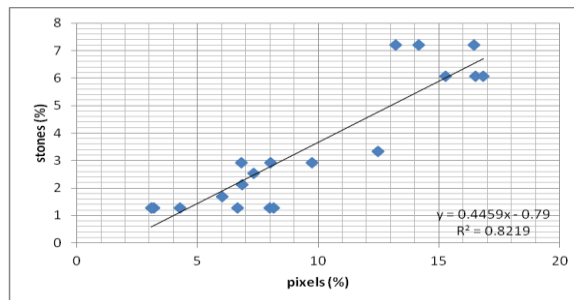
**Fig 3:** original image (a) and image after the colour based segmentation, showing the results expressed in the pixel area occupied by stones (b).

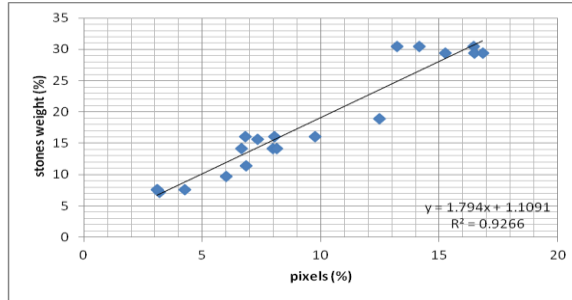
Table 2 shows the real value in terms of weight and number of stones, as well as the values obtained by the software after applying segmentation and standard deviation

Sample	% Weight stones Manual Measure	% Stones Manual Measure	% Measured Vision Method
Image A	15.69	2.52	5.50±2.12
Image A	30.43	7.20	14.61±1.66
Image A	14.20	1.28	7.60±0.82
Image A	29.44	6.07	16.21±0.83
Image A	7.63	1.28	3.49±0.67
Image A	16.09	2.93	8.20±1.48
Image A	75.04	88.24	73.93±3.75
Image A	9.71	1.69	6.01±2.10
Image A	18.89	3.33	12.48±1.54

**Table 2:** results of some images obtained by the software after applying segmentation algorithms to assess the amount of stones in a set of images of olives.

Following the same proceedings, thanks to the analysis of these results, a calibration curve was built in order to estimate the weight or number of stones according to the area in pixels obtained by the software (Figure 4).





**Fig 4.** Correlation between results obtained by both methods in percentage of stones (a) and weight of stones (b).

This phase of the test showed that the values provided by the artificial vision method are very similar to the real values provided by the manual method used as the reference method. Visual and numerical results indicate that the method is highly efficient for this application.

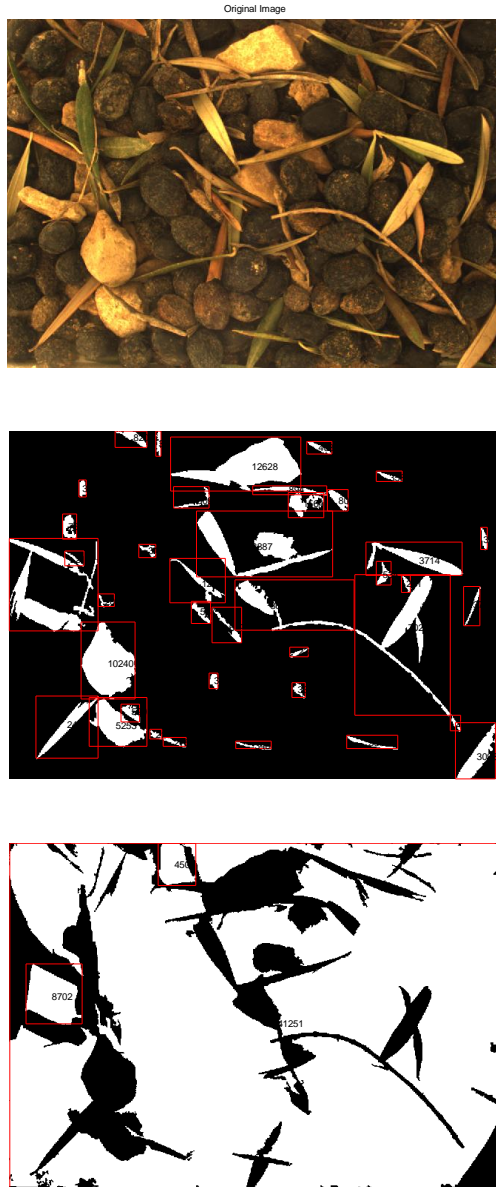
Furthermore, these results confirm the need to use different calibration curves for different types of ground particles in order to obtain accurate results. In real life, it might not be necessary to detect leaves if there is vision system implanted after the cleaning process, which removes the leaves. However, it would be very helpful to use this detection system of dirt, stones and mud in order to detect a lot of olives which arrive to the factory in very bad conditions for having been collected from the ground.

In addition, this method can be used to determine the global level of dirt of the samples, obtaining the results as a percentage of the area occupied by dirt and the area occupied by olives.

These algorithms were tested with other 30 images of olives presenting varied distributions of dirt (leaves, stones and mud). Table 4 shows the results as the percentage of dirt obtained from the analysis of some images.

Figure 5 shows an original image, as well as the analysed image after applying the segmentation algorithm. Finally, table 3 shows the results of the analysis of some images





**Fig 5.** Original image (a) and image after the colour based segmentation, showing the results expressed in the pixel area occupied by dirt (b) and olives (c).

<b>% measured by Vision Method</b>	<b>% olives</b>	<b>% dirt</b>
<b>Img 1</b>	70	30
<b>Img 3</b>	74	26
<b>Img 2</b>	80.9	19.1
<b>Img 5</b>	87.3	12.7
<b>Img 4</b>	94.7	5.3

**Table 3.** Results of some images obtained by the software after applying segmentation algorithms to assess the global dirt index.

According to the results, this method provides an overall evaluation of dirt in the samples when entering the factory. As a result, this method would allow us to determine the level of dirt quickly and automatically bearing in mind the real weight of olives entering the factory and, thus, would allow us to assess the destination of samples according to their sanitary condition.

#### **4. Conclusion**

These results show that from the visual and geometric characteristics, an automatic prediction of the percentage of dirt present in the samples of olives has been developed, which could be used in the manufacturing process of olive oil.

The present work describes a model system based on colour segmentation and classification through the method of nearest neighbour (KNN-k nearest neighbour), which is able to estimate the amount of dirt in olive batches. It can be used as an automatic control of an oil mill and it can help to improve the information available about olives when entering into the elaboration process.

After further optimization and increase of accurate calibrations, the new artificial vision system could be integrated into a grading machine for on-line detection of dirt in olives batches. Furthermore, this detection program can be the basis for the development of a detection system for dirt in olives.

In conclusion, the designed off-line computer vision system can be used as a blueprint for the development of on-line dirt detection in real factories.

## Acknowledgment

The authors thank the National Institute of Research and Technology Agriculture and Food (INIA) for funding of the scholarship of training FPI sub-INIA, ESF co-financed through the ESF Operational Programme for Andalusia 2007-2013 within axis 3 "Increasing and improving human capital" expenditure category "Developing human potential in the field of research and innovation."



## References

- [1] Hermoso, M., González, L., Uceda, M., García-Ortiz, A. Morales, F., Frias, L., Fernández, A. Elaboración de aceite de oliva de calidad (1991) Consejería de Agricultura y Pesca. Dirección general de investigación, tecnología y formación agroalimentaria y pesquera.
- [2] Sun, D.-W., 2000. Inspecting pizza topping percentage and distribution by a computer vision method. *Journal of Food Engineering* 44, 245\_/249.
- [3] Brosnan, T., Sun, D.W., 2004. Improving quality inspection of food products by computer vision-a review. *Journal of Food Engineering* 61, 3-16.
- [4] Pla, F., Sanchez, J.S., & Sanchiz, J.M. (1999). On-line machine vision system for fast fruit colour sorting using low-cost architecture. In: VIII SPIE Conference on Machine Vision Systems for Inspection and Metrology, Boston, Massachusetts, USA.
- [5] Mendoza, F., Aguilera, J.M., 2004. Application of image analysis for classification of ripening bananas. *Journal of Food Science* 69 (9), 471–477.
- [6] Oberti, R., Guidetti, R., Mignani, I., 1999. Analisi multispettrale di immagini per l'individuazione precoce di danni meccanici su prodotti ortofrutticoli: un'applicazione su pere decana. *Rivista di Ingegneria Agraria* 30 (3), 137–147.
- [7] Peri, G., Romaniello, R., Amodio, M., & Colelli, G. (2003). The application of a fast colour quantization algorithm for sorting peaches for firmness. In: XXX CIOSTACGIR V CONFERENCE, Torino, Italy.

- [8] Abdullah, M.Z., Mohamad-Saleh, J., Fathinul-Syahir, A.S., Mohd-Azemi, B.M.N., 2006. Discrimination and classification of fresh-cut starfruits (*Averrhoa carambola* L.) using automated machine vision system. *Journal of Food Engineering* 76 (4), 506–523.
- [9] Koc, A.B., 2007. Determination of watermelon volume using ellipsoid approximation and image processing. *Postharvest Biology and Technology* 45 (1), 366–371.
- [10] Bodria, L., Fiala, M., Guidetti, R., & Oberti, R. (2002). Optical techniques for assessing the fruit maturity stage. In: *ASAE Annual International Meeting/CIGR XVth World Congress*, Chicago, USA.
- [11] Choong, T.S.Y., Abbas, S., Shariff, A.R., Halim, R., Ismail, M.H.S., Yunus, R., Salmiaton, A., Fakhrol-Razi, A., 2006. Digital image processing of palm oil fruits. *International Journal of Food Engineering* 2 (2), 7.
- [12] E Guzmán, V Baeten, JAF Pierna A portable Raman sensor for the rapid discrimination of olives according to fruit quality ... - *Talanta*, 2012 – Elsevier
- [13] Medioni, G.; Lee, M-S; Tang, C\_K.; “A Computational Framework for Segmentation and Grouping”, Elsevier Science B.V., 2000. ISBN: 0-444-50353-6.
- [14] Plataniotis, K.N.; Venetsanopoulos, A.N.; “Color Image Processing and Applications”, Springer Verlag, 2000. ISBN: 3-540-66953-1.
- [15] Richards, J. A.. *Remote Sensing Digital Image Analysis*. Ed. Springer-Verlag, 1999.
- [16] Thomas, S.W., 1991. Efficient inverse color map computation. In: Arvo, J. (Ed.), *Graphics Gems II*. Academic Press, Boston, pp. 116–125.
- [17] Floyd, R. W., & Steinberg, L. (1975). An Adaptive Algorithm for Spatial Gray Scale. In: *International Symposium Digest of Technical Papers*, Society for Information Displays, 36

**Infrared machine vision system for the  
automatic detection of olive fruit quality.**

**Food Control (enviado)**

# **Infrared machine vision system for the automatic detection of olive fruit quality**

Elena Guzmán<sup>a</sup>, Vincent Baeten<sup>b</sup>, Juan Antonio Fernández Pierna<sup>b</sup>, José A. García-Mesa<sup>a</sup>

<sup>a</sup> Centro IFAPA 'Venta del Llano', Crta. Nacional Bailén-Motril Km 18.5, 23620 Mengíbar, Jaén, Spain

<sup>b</sup> Walloon Agricultural Research Centre, Valorization of Agricultural Products Department, Food and Feed Quality Unit, Henseval Building, 24 Chaussée de Namur, 5030 Gembloux, Belgium

Corresponding author E-mail address: elena.guzman.jimenez@juntadeandalucia.es (E. Guzmán). Phone: 34953366366

## **Abstract**

External quality is an important factor in the extraction of olive oil and the marketing of olive fruits. The appearance and presence of external damage are factors that influence the quality of the oil extracted and the perception of consumers, determining the level of acceptance prior to purchase in the case of table olives. The aim of this paper is to report on artificial vision techniques developed for the online estimation of olive quality and to assess the effectiveness of these techniques in evaluating quality based on detecting external defects. This method of classifying olives according to the presence of defects is based on an infrared (IR) vision system. Images of defects were acquired using a digital monochrome camera with band-pass filters near-infrared (NIR). The original images were processed using segmentation algorithms, edge detection and pixel value intensity to classify the whole fruit. The detection of the defect involved a pixel classification procedure based on nonparametric models of the healthy and defective areas of olives. Classification tests were performed on olives to assess the effectiveness of the proposed method.

This research showed that the IR vision system is a useful technology for the automatic assessment of olives that has the potential for use in offline inspection and for online sorting for defects and the presence of surface damage, easily distinguishing those that do not meet minimum quality requirements.

**Key Words:** Algorithm, surface damage, olive fruit, image analysis

## 1.Introduction

The presence of damage on the skin of the olives is the most decisive factor in determining the fruit's external quality. The classification of olives is important in order to determine whether they meet the requirements for marketing as table olives or for oil extraction. The control of olives is based on human supervision, and it is important to detect whether they have any disease or damage to the skin.

A new approach to the infrared (IR) system, which improves its effectiveness and can be optimized by using a set of IR monochrome images, makes it possible to discover and assess features that are difficult to see using traditional systems. The use of these techniques has increased in recent years because they provide important information on the natural attributes of the products studied.

Another important feature of these systems is that they enable these products to be studied in electromagnetic spectrum regions where the human eye cannot operate, such as the ultraviolet (UV), IR and near-infrared (NIR) regions. The high risk of human error in sorting processes is well known and is one of the most important obstacles that machine vision can help overcome.

For all these reasons, the application of machine vision to food analysis has increased considerably in recent years. Some research institutes and manufacturers of are seeking to develop automatic classification techniques. Moltó et al. (1998) studied image analysis techniques in order to measure some parameters that define the quality of various fruits and vegetables. Later, Moltó et al. (2000) attempted to determine citrus size, color, shape and defects. Heinemann et al. (1997) developed an algorithm to evaluate apple size, shape, color and surface defects using machine vision. These parameters have also been assessed by other authors using artificial vision systems [4,5,6,7,8].

There are some articles on the detection of damage and quality control in fruits. One of the first approaches to detecting bruising on apples was based on using interferential filters [9]. The techniques used to detect defects have varied from the use of wavelets [10] to the recognition of edges [11]. The methods for obtaining multiple images are based on using photographs [12], X-rays [13], and three-dimensional images [14].

The most recent techniques for defect detection combine IR and visible (RGB) data [15] or the use of hyperspectral imaging [16]. A multispectral vision system that was developed included four wavelength bands in the visible/NIR range for sorting apples [17]. The 750- and 850-nm bands, however, offered a good contrast between defective and healthy tissue and were well suited to detecting internal tissue damage. Bennedsen and Peterson (2005) [18] developed a multispectral machine vision system for detecting surface defects on apples, and other researchers [19] developed a multispectral inspection system for detecting and sorting citrus fruits according to 11 types of external defects by combining the information obtained from four image acquisition systems that are sensitive to NIR, visible, UV and fluorescent images. A further step was taken in the development of a prototype for the real-time inspection of citrus canker based on two wavelengths centered at 730 and 830 nm [20]. In order to discriminate between the stem end and the actual defects, they combined a further two bands. All these studies have found that the use of monochromatic images corresponding to wavelengths in the NIR contain a large amount of information on fruit surface defects and external damage that visible images do not provide.

Diaz et al. (2000) [21] used a vision system to link the quality parameters of table olives with their physicochemical characteristics. The system was based the development of a fast and efficient algorithm that would be incorporated into an automatic sorting process. Other studies on the use of artificial vision for olive fruits include work on classifying external damage contour edges by using visible images [22], predicting the ripening index [23] and determining the parameters of oil quality based on artificial vision and ANN [24].

In view of these studies have shown that the infrared region which provides better contrast in the detection of tissue damage is near infrared area of 700-900 nm.

This paper discusses recent advances in the use of this technology for assessing fruit quality and proposes a method for classifying olive fruits according to their quality by determining the percentage of external damage using an IR vision system with band ranging from 760 nm to 1000 nm. The lower cost of these systems allows them to now be used in laboratories for food quality, and they are an emerging and promising tool for food quality and safety control.



## **2. Materials and methods**

### **2.1. Image system**

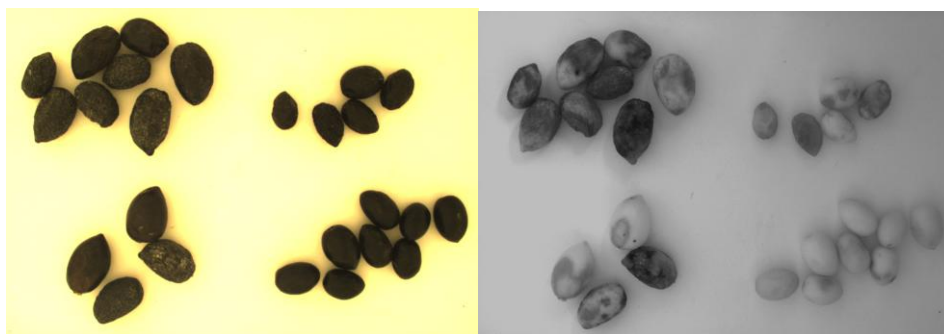
The images were acquired with a digital 2CCD (charge-coupled device) Progressive Scan Multi-Spectral Camera JAI AD-080CL, which combined a visible color channel (Bayer mosaic CCD) and an NIR channel (monochrome CCD) which provided IR information. The major advantage of this camera is that it captures both channels simultaneously through the same optical path. This camera is based on a dichroic prism, allowing precise separation of the visible (color) and Near-Infrared parts of the spectrum into two separate channels. The Visible (color) channel is referred to as Channel 1 and the Near-Infrared channel is referred to as Channel 2. Channel 1 and 2 can be configured to operate separately or synchronously. The dichroic prism incorporated in the AD-080CL separates the visible (color) part of the spectrum into a wavelength band from 400nm to 650nm (Channel 1) and the Near-IR part into a band ranging from 760 nm to 1000 nm (Channel 2). It can be used for inspection tasks, where images from the visible (color) and NIR spectra are required in combination. This system can generate images with 8- or 10-bit depth and 24-bits in RGB images, using a standard Camera Link interface with a resolution of 1024 x 768 active pixels per channel. The camera was installed in an enclosed cabin under controlled lighting in order to achieve a consistent image in all acquisitions. The lighting was provided by a halogen lamp with some filters, which created a diffused light simulating the illuminant D65 (6,500 K. This is the standard set by the International Commission on Illumination and simulates normal light conditions until noon, thus avoiding glow effects on the fruit. The images were taken from a distance of about 45 cm, with the olives on a neutral white background. For the online application of this method, the devices were installed in the processing line on a belt that transported olives into the mill, where the lighting conditions were the same as those described above. Being a high-speed camera, it can take high-quality pictures of the fruit, reaching a motion capture of up to 30 frames per second with full resolution.

### 2.3 Methodology

Freshly collected samples from the 2011-2012 olive crop at the Venta del Llano Instituto de Investigación y Formación Agraria y Pesquera (IFAPA) in Mengíbar in Jaén, Spain were used. Real samples were collected from a factory in Jaén province.

The algorithms used were developed with the image analysis software Image Pro-Plus version 6.0 (MediaCybernetics, Inc). Forty-five IR images of healthy olives and of olives with defects were acquired. The defects included damage with or without skin perforation, scar tissue, frost damage, visible flesh damage and recent bruises.

Using this technique it was possible to distinguish apparently damaged olives which looked healthy to the human eye. The difference between healthy and damaged skin could be seen clearly in the NIR bands, which provided a good contrast between the damaged and healthy skin and were well suited to detecting internal skin damage [25]. Figure 1 shows an RGB image of healthy and defective olives where the external damage is not clear and the corresponding IR image can where the damage can clearly be seen.



**Figure 1.** Visible (RGB) and infrared (IR) images of a batch of olives

In the wavelength bands of IR images, defect pixels covered a wide range of intensities, and each pixel of the fruit was grouped as either non-defective or defective. After selecting the images, areas of interest or objects in the image were selected using edge detection algorithms and labeling. [26,27]. Edge detection was performed to identify objects in the IR image, so that separate the different objects in the image to build a mask. This mask is applied to the original image to finally

segmentation of external defects inside each object. A routine was developed to calculate the position and size of each defect based on this segmentation.

### 2.3.1. Identification and separation of objects in the NIR image

The procedure to separate objects involved using an NIR monochrome image in order to segregate each olive fruit. In order to identify objects in the NIR monochrome image, a contrast equalization procedure was performed, followed by spatial filter enhancement ('Flatten'), which reduces the intensity variation in the background pixels. Spatial filtering was applied to the images in order to enhance or attenuate spatial detail, which would in turn enhance visual performance or facilitate further processing. Spatial filtering can be considered as a 'local' procedure in image processing in that it changes the value of each pixel according to the values of the surrounding pixels. From a practical point of view, this step consists mainly of classifying the original gray levels in the original NIR image according to the value of the neighboring pixels. This task was carried out using two algorithms: (a) Border-4 Neighbor, which determines the edges of the olive (taking into account the pixel intensity values in the histogram) and considers the darker pixels as edges of objects; and (b) the connected components algorithm, which is used to give each olive a value. The edge-based algorithm enumerates each set of pixels that make up an olive [28].

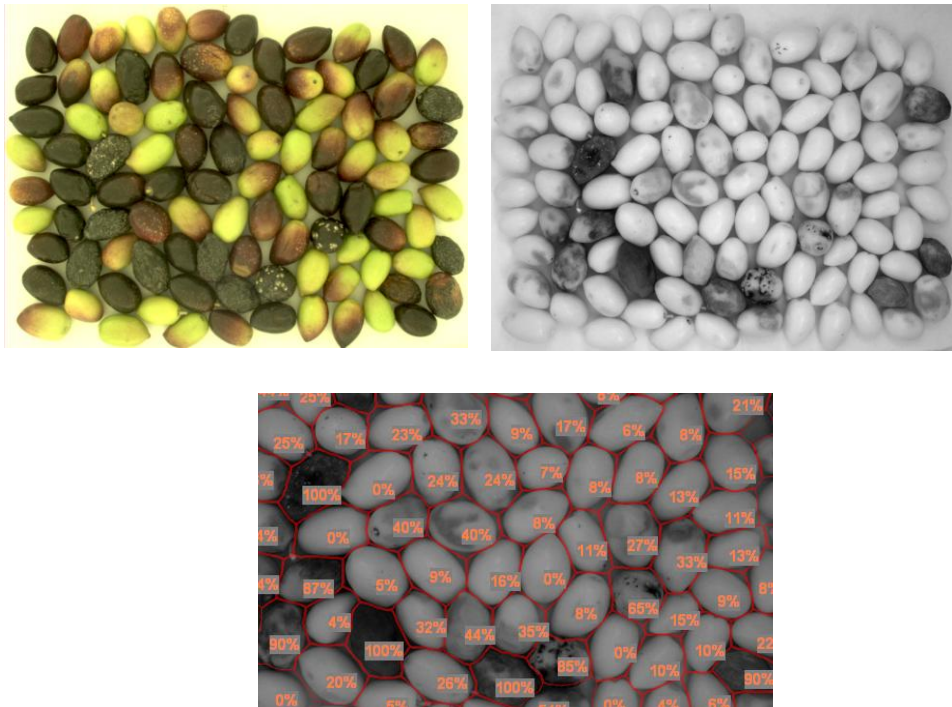
### 2.3.2. Creating a mask and segmentation of gray levels

Based on these objects, a skeleton image was constructed. It was intended to represent the shape of an object with a relatively small number of pixels. Thus, all pixels of the skeleton were structurally necessary. The position, orientation and length of the skeleton lines corresponded to those equivalents in the original image. The task of extracting the features of an image was simplified in order to obtain the skeleton. A filter mask 'pruning' was applied, which reduced the image to a skeleton and eliminated protrusions. The skeleton was added to the original IR image and marked the position of the defects in the original image. To that end, a gradient-based segmentation of gray levels, areas with the highest values of gray representing the area of pixel defects and areas with the lowest value of gray representing the healthy areas of the fruit. This made it possible to measure some geometric properties of the defects and, in particular, the defect area of each olive. After an evaluation of the

defect areas, the results were translated into percentages based on the sum of all pixels in the defective areas of the olives and the sum of all pixels in the healthy areas of the olives.

### 3. Results

In order to test this method, 45 sample images were taken of both healthy olives and defective olives. The number of olives in the images varied, in order to obtain all possible combinations of healthy and defective olives. Figure 2 shows the original RGB image, the IR image and the image results obtained after applying the algorithms to determine the defects in a set of olives. As can be seen in the RGB image, visually apparent damage on the fruit cannot be detected, whereas in the IR image these defects are clearly distinguished.



**Figure 2.** Original image in RGB, IR image and results in zoom images obtained

The defect detection results (in percentages) for some of the images analyzed are shown in Table 1. The results are given as the healthy and defective percentages for each olive.

Object	% healthy	% defective
Olive1	78.82	21.18
Olive 2	100.00	0.00
Olive 3	62.56	37.44
Olive 4	74.26	25.74
Olive 5	45.60	54.40
Olive 6	59.75	40.25
Olive 7	27.03	72.97
Olive 8	49.81	50.19
Olive 9	18.85	81.15
Olive 10	0.00	100.00

**Table 1.** Defect detection results for some olives in the image

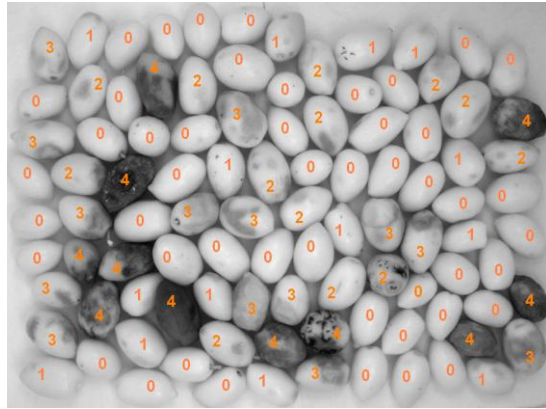
Each pixel of the fruit was classified as healthy or defective according to the values of its spectral components. For each olive, the results were given as the total percentages of defective and healthy skin.

In order to classify olives according to their grade defect, was defined five classes for grouped them into five classes. Table 2 defines these classes: healthy olives, olives with minor damage, olives with moderate damage, olives with severe damage, and defective olives.

Category		
0	Healthy Olive	100% healthy area
1	Minor defects	>75% healthy area
2	Moderate defects	>50 % defective area
3	Several defects	>75% defective area
4	Defective Olive	100 % defective area

**Table 2.** Olive classification based on percentage of healthy and defective skin

Figure 3 shows an analyzed image, with the classification results. Table 3 shows the results of the classification of some samples of olives, using the proposed automatic method.



**Figure 3.** Results of classification for an analyzed image

Images	% defective	% healthy	Total no. of olives	Class 0	Class 1	Class 2	Class 3	Class 4
Img 1	14.06	85.94	116	75	40	1	0	0
Img 2	34.95	65.05	80	7	35	29	0	9
Img 3	22.34	77.66	111	6	64	41	0	0
Img 4	15.88	84.12	110	14	78	18	0	0
Img 5	47.35	52.65	139	0	40	71	15	13
Img 6	68.56	31.44	107	0	3	51	30	23
Img 7	70.32	29.68	105	0	2	50	27	26
Img 8	68.42	31.58	122	0	0	80	22	20
Img 9	89.62	10.38	123	0	0	19	62	42
Img 10	70.98	29.02	115	0	0	62	34	19

**Table 3.** Results of the classification of some samples of olives, using the proposed automatic method

These results indicated that spectral components of IR images are a useful tool for detecting surface damage in olive fruits. Spectral imaging provides a valuable tool to detect external defects that are not clearly visible to the human eye. The near infrared bands provide a good contrast between defective and healthy tissue and are well suited for detecting damage to internal tissues, such as skin and shock damage. The method, based on selecting on a rank of spectral and applying the appropriate filter in

order to acquire spectral images, appears to be suitable for the automatic detection of defects in olive fruits. These results show that it is possible to extract information from the visual characteristics and infrared vision system and visible to develop applications for the automatic prediction of the health status of olive samples at the start of the process of producing olive oil. The proposed methodology, based on the edge-based and connected components algorithms, is able to estimate the percentage of defects in the olives. It has potential for the online classification of olives, which would help improve the quality control of olive oil in factories. After further optimization and more accurate calibration, the new vision system could be incorporated into a sorting machine for online screening of the health status of olives and thus provide the basis for a defect detection system for olives. The development of an offline machine vision system could be used as a model for defect detection in olives at the factory level. The future of spectral systems applied to food inspection is promising, with both industry and consumers becoming more aware of the need to ensure quality and food safety. This technology is an important tool for the automatic inspection and control of these parameters

After further optimization and more accurate calibration, the new vision system could be incorporated into a sorting machine for online screening of the health status of olives and thus provide the basis for a defect detection system for olives. The development of an offline machine vision system could be used as a model for defect detection in olives at the factory level. The future of spectral systems applied to food inspection is promising, with both industry and consumers becoming more aware of the need to ensure quality and food safety. This technology is an important tool for the automatic inspection and control of these parameters.

### **Acknowledgment**

The authors thank the National Institute of Research and Technology Agriculture and Food (INIA) for funding of the scholarship of training FPI sub-INIA, ESF co-financed through the ESF Operational Programme for Andalusia 2007-2013 within axis 3 "Increasing and improving human capital" expenditure category "Developing human potential in the field of research and innovation."

## References

- [1] Moltó, E., Ruiz, L.A., Aleixos, N., Vázquez, J., Juste, F. Machine vision for non-destructive evaluation of fruit quality. *Acta Horticulture* 421, pp. 85-90 (1998).
- [2] Moltó, E., Blasco, J., Aleixos, N. Multispectral, real-time inspection of citrus using a parallel architecture. In *AgEng 2000 – Agricultural Engineering into the Third Millennium*. Warwick, UK, 2-7 July (2000).
- [3] Heinemann, P. H., Hughes, R., Morrow, C.T., Crassweller, H.J., Sommer, H.J., Marrazzo, W.N., He, B. Machine vision based station for grading of Golden Delicious apples. *Proc. Sensors for Nondestructive Testing-97*, pp. 239-248 (1997).
- [4] Diaz, R., Gil, L., Serrano, C., Blasco, M., Molto, E., Blasco, J. Comparison of three algorithms in the classification of table olives by means of computer vision. *Journal of Food Engineering* 61 (1), 101-107 (2004).
- [5] Mateos, F., Patino, A., Perez, J., Tombs, J., Aguirre, M.A. The implementation in FPGA of a real-time processing system for the detection of defects on fruit. In: *V Jornadas de Computation Reconfigurable y Aplicaciones*, pp. 151-157. Granada, Spain (2005).
- [6] Tao, Y., Morrow, C. T., Heinemann, P. H., Sommer, J. H. Automated machine vision inspection of potatoes. *ASAE Paper No. 90-3531* (1990).
- [7] Varghese, Z., Morrow, C. T., Heinemann, P. H., Sommer, H. J., Tao, Y., Crassweller, R.M. Automated inspection of Golden Delicious apples using colour computer vision. *ASAE Paper No. 91-7002* (1991).
- [8] Brodie, J.R., Hansen, A. C., Reid, J. F. Size assessment of stacked logs via the Hough Transform. *Transactions of the ASAE*, 37(1), 303–310 (1994)
- [9] Leemans, V., Magein, H., Destain, M. F. Defects segmentation on 'Golden Delicious' apples using colour vision and Bayesian Method. *Computers and Electronics in Agriculture*, 23, 43–53 (1999).
- [10] Woodford, B. J., Kasabov, N. K., Wearing, C. H. 'Fruit Image Analysis Using Wavelets'. *Proceedings of the ICONIP/ANZIIS/ANNES'99 International Workshop*. University of Otago Press. pp. 88–91(1999).
- [11] Jimenez, A. Jain, A. Ceres, R. Pons, J. Automatic fruit recognition: A survey and new results using range/attenuation images. *Pattern Recognition*, 32 (10), pp. 1719-1736 (1999).



- [12] Blasco, J., Molto, E. Identification of defects in citrus using multispectral imaging. International Conference on Agricultural Engineering, AgEng 02, Budapest, Hungary. EurAgEng Paper No. 02-AE-031 (2002).
- [13] Krger, C., Bartle, C.M., West, J.G. Digital X-Ray imaging and image processing for object detection and closed containers. Proceedings (426) Computer Graphics and Imaging (2004).
- [14] Jain, A. and Flynn, P. Three-dimensional object recognition. ACM Computing Surveys (CSUR) – Annals of Discrete Mathematics, 24, Volume 17 Issue 1, March (1985).
- [15] Kleynen, O., Leemans, V., Destain, M. F. Development of a multi-spectral vision system for the detection of defects on apples. Journal of Food Engineering, 69, 41–49 (2005).
- [16] Peirs, A., Scheerlinck, N., De Baerdemaeker, J., Nicolai, B. M. Quality determination of apple fruits with a hyperspectral imaging system. AgEng 02. Budapest, Hungary. EurAgEng Paper No. 02-PH-028 (2002).
- [17] Rehkugler, G. E., and Throop, J. A. (1986). Apple sorting with machine vision. Transactions of the ASAE, 29(5), 1388–1397 (2002).
- [18] Bennedsen, B. S. and Peterson, D. L. Performance of a system for apple surface defect identification in near-infrared images. Biosystems Engineering, 90(4), 419–431(2005).
- [19] Blasco, J., Aleixos, N., Gómez, J., Moltó, E. Citrus sorting by identification of the most common defects using multispectral computer vision. Journal of Food Engineering, 83(3), 384–393.(2007)
- [20] Aleixos, N., Blasco, J., Navarrón, F., Moltó, E. Multispectral inspection of citrus in real-time using machine vision and digital signal processors. Computers and Electronics in Agriculture, 33(2), 121–137 (2002).
- [21] Diaz, R., Faus, G., Blasco, M., Blasco, J., Moltó, E. The application of a fast algorithm for the classification of olives by machine vision. Food Research International 33, 305-309 (2000).
- [22] Riquelme, M.T., Barreiro, P., Ruiz-Altisent, M., Valero, C. Olive classification according to external damage using image analysis. Journal of Food Engineering Volume 87, Issue 3, August 2008, Pages 371–379 (2007)
- [23] Furferi, R., Governì, L., Volpe, Y. ANN-based method for olive Ripening Index automatic prediction. Journal of Food Engineering 101 (2010) 318–328 (2010).
- [24] Furferi, R., Carfagni, M., Daou, M. Artificial neural network software for real-time estimation of olive oil qualitative parameters during continuous extraction. Computers and Electronics in Agriculture 55 (2007) 115–131 (2006).
- [25] Lu, R. and Y. R. Chen. Hyperspectral imaging for safety inspection of food and agricultural products. SPIE 3544:121–133 (1998).
- [26] Medioni, G., Lee, M-S., Tang, C.K. 'A Computational Framework for Segmentation and Grouping', Elsevier Science B.V. ISBN: 0-444-50353-6 (2000).

[27] Plataniotis, K.N. and Venetsanopoulos, A.N. 'Color Image Processing and Applications', Springer Verlag. ISBN: 3-540-66953-1 (2000).

[28] Chang, V. and Saavedra, J. Métodos alternativos para el mejoramiento automático del contraste de imágenes. Computer Engineering thesis, School of Informatic, University National of Trujillo (2001).

[29] Russ, J.C. The Image processing Handbook. CRC Press LLC, 3rd edition (2006)

**Using a visible vision system for on-line  
determination of quality parameters of olive  
fruits**

**Applied Sciences: Rapid Detection Systems  
(enviado)**

# Using a visible vision system for on-line determination of quality parameters of olive fruits

Elena Guzmán<sup>a</sup>, Vincent Baeten<sup>b</sup>, Juan Antonio Fernández Pierna<sup>b</sup>, José A. García-Mesa<sup>a</sup>

<sup>a</sup> Centro IFAPA 'Venta del Llano', Crta. Nacional Bailén-Motril Km 18.5, 23620 Mengíbar, Jaén, Spain

<sup>b</sup> Walloon Agricultural Research Centre, Valorization of Agricultural Products Department, Food and Feed Quality Unit, Henseval Building, 24 Chaussée de Namur, 5030 Gembloux, Belgium

Corresponding author E-mail address: elena.guzman.jimenez@juntadeandalucia.es (E. Guzmán). Phone: 34953366366

## Abstract

The increased expectations for food products of high quality and safety standards and the need for accurate, fast and objective quality determination of these characteristics in food products continue to grow. In this situation, new techniques are necessary to enable on-line control of quality parameters.

Computer vision provides one alternative for an automated, non-destructive and cost-effective technique in order to accomplish these requirements. The maturity index and sanitary conditions were objectively assessed on-line by image analysis obtained through machine vision, in which algorithms of colour-based segmentation, as well as the main operators to detect edges were used.

The proposed methodology is able to estimate the maturity index and the percentage of defects in olives. In addition, this system can be potentially used on-line in the classification of olives, which means that it could help to improve the quality control of olive oil in factories.

Key Words: on-line, algorithm, maturity index, olive fruit, image analysis

## 1. Introduction

There are many factors which can affect the quality of olive oil, such as inherent agronomical factors or factors related to the various steps in the olive oil processing, i.e. from extraction to bottling.

Sensory characteristics and chemical composition greatly influence the quality of a product. Traditionally, quality inspection of agricultural and food products has been performed by human graders. However, in most cases these manual inspections are

time-consuming and labour-intensive. Controlling the oil at the time of production, thus avoiding mixing oils of different qualities and thereby improving the quality of extra virgin oil is essential.

The appearance and colour of food are important factors which influence consumer preferences. During the ripeness time, olives follow different phases. Firstly, they are of a deep green colour with predominantly high levels of chlorophyll and give rise to bitter oil. In their final phase, they present a black colour paste, their chlorophyll is degraded and replaced by anthocyanin and they are more sensitive to external damage and infections [1,2,3].

The maturity index is useful for producers to identify the optimal time for harvesting the olives. Using this index may also help to increase the quantitative and qualitative characteristics of olive oil production. There are many systems used to determine the state of maturity of olives depending on various factors, such as fruit firmness [4,5], the calculation of dry matter [6], the calculation of the respiration rate of fruit [7] or non-destructive methods, such as the use of the hydrometer [8,9,10] or by measuring the transmission of acoustic waves through the fruit [11].

The most common method to evaluate this index in olives is called maturity index (MI), which considers changes in the colour of the skin and flesh during the ripening of fruit and gives them their colour according to a class (classes 1 to 7) [12, 13,14].

However, this method is subjective and depends on environmental conditions, which can affect the colour appearance of olives. At times, the evaluators have to decide between colours which are difficult to distinguish. Consequently, in order to provide a successful complement for the current industry, we need a method able to assess the maturity index accurately, which is, additionally, easy to be used, inexpensive, preferably non-destructive, and which can be coupled in the process line for the rapid determination of the maturity index.

Techniques such as infrared spectroscopy have been studied and successfully used for rapid characterization of some olives and some olive oil parameters [15,16,17,18] and, for quality control and for the determination of the geographic origin of virgin olive oils [19]. Studies using Raman spectroscopy have also been carried out for authentication purposes, for the classification of PDO olive oils [20,21] and for the detection of adulteration with other low price oils [22,23,24]. Moreover, this technique has been applied for the determination of the oxidative degradation [25], free acidity

[26,27], the characterization of antioxidant olive oil biophenols [28] and, finally, to provide some contributions dealing with the evaluation of fruit quality [29] .

At present, the olive oil sector is experiencing some changes in technology in the production lines allowing a continuous process control. The olive oil production focuses on producing high quality products. This new approach determines a transition from mass production of olive oil to lower production characterized by a higher quality of marketable products. Consequently, many scholars and scientists have focused on improving production processes as shown in the literature [30, 31]. This grade of oil production can be optimized with automatic control of the mechanical extraction lines. This would be capable of controlling the cycle of continuous extraction by measuring various technical parameters of olives and by determining the influence of these parameters on the quality of olive oil [32]. Some of the parameters which have been studied include, for instance, the oil temperature in the mixer, the mixing time and temperature of the oil centrifugal extractor. Some recent publications have explored the possibility of a specific software system, which is able to control the extraction process and the parameters involved. Artificial neural networks were used to estimate some parameters for the extraction of olive oil in real time [33,34,35], as well as the use of NIR sensors for the on-line characterization of olive oils [36]. Finally, other works have been carried out for the determination of parameters of oil and acid value (AV), the bitter taste (K225) and fatty acid composition (FAME) [37], for the monitoring of carotenoid and chlorophyll pigments in virgin olive oil [38], and the measurement of oil content and humidity in olive cakes [39].

In the food industry, artificial vision has found many applications for controlling product quality. This non-destructive method is fast, inexpensive, consistent and objective, and has expanded in many different industries. Its speed and accuracy satisfy ever-increasing production and quality requirements, hence aiding in the development of totally automated processes [40]. Some methods based on image processing have been designed over the last five years in order to conduct real-time control and automated inspection thanks to different instrumentation, such as cameras and spectrometers, which allow the detection of features in fruit. By interposing appropriate filters, we will achieve the analysis of non-visible areas which let us understand some characteristics of products, which are not always detected visually by the operators.

Many applications are developed using computer vision as a technique for classifying the fruit: peaches [41], citrus [42,43], cherries [44,45] and especially apples [46,47,48] Brosnan and Sun [49] present a comprehensive review of image processing techniques for various food products.

Over the previous years, studies have shown the ability to sort the fruit by its colour [50,51,52,53]. There are also some algorithm classification of the fruit on the basis of the shape and size of cucumbers [54] and watermelon [55]. Some significant correlations between data about the shape, colour and trend of aging, and the age of apples have been found [56].

In the literature there are works related to the correlation between the image and the degree of ripeness of different fruit [57,58] in order to remove products which do not meet the required standards of classification.

In previous works, we have used spectroscopy as a rapid non-destructive method to determine the quality of olives in the process [59]. This paper aims to improve the application of techniques for fast on-line analysis of olive oil and proposes a direct method to determine on-line the maturity of olives. This non-destructive method offers many advantages in the quality control of the process, since it is an easy and quick way of obtaining data, which is difficult to obtain manually. The final methodology could be useful for evaluating the performance of commercial systems and better control of the extraction process of olive oil.

## **2. Materials and methods**

### **2.1. Samples**

Olive samples of picual variety (2 kg) were collected in several olive mills in the province of Jaen (Spain) during November, 2011. The colour of these olives ranged from dark green to completely black. These samples were used for the acquisition of images; synthetic samples were also carried out so as to be included in the analysis of samples of olive with a full range of maturity index.

### **2.2. Visual determination of Maturity Index (MI)**

The MI was determined by an experienced evaluator; every sampling day an assessment of skin, flesh and colour of olives was performed [60]. The procedure involved the distribution of olives into eight groups according to the characteristics

summarized in Table 1. The official method uses 100 olives and calculates a global value of maturity index for each sample or group of olives.

The MI is calculated by using the following equation:

$$MI = \sum i * ni / 100$$

where i is the group number and ni the number of olives in it.

This method has to separate the olives manually, cut the pulp to examine them and count and identify the group to which they belong.

<b>Maturity index</b>	<b>Description</b>
<b>Group</b>	
<b>0</b>	Skin color deep green
<b>1</b>	Skin color yellow-green
<b>2</b>	Skin color with < half the fruit surface turning red, purple or black
<b>3</b>	Skin color with > half the fruit surface turning red, purple or black
<b>4</b>	Skin color all purple or black with all white or green flesh
<b>5</b>	Skin color all purple or black with < half the flesh turning purple
<b>6</b>	Color all purple or black with > half the flesh turning purple
<b>7</b>	Skin color all purple or black with all the flesh purple to the pit

**Table 1:** MI classification groups

## 2.2 Image System

Images were captured with a JAI AD-080CL multi-spectral camera, combining a visible colour channel (Bayer Mosaic CCD), which can generate 24-bit RGB images and a NIR (Near Infrared) channel (monochrome CCD). The major advantage of this camera is the capability of capturing both channels simultaneously through the same optical path.



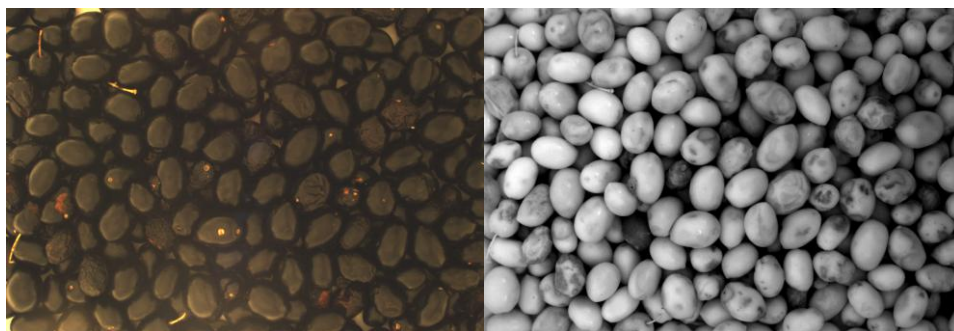
This camera uses a standard Camera Link interface, whereby each channel can output images with 8 or 10-bit and 24- bits, and a resolution of 1024x768 active pixels per channel. This camera was installed in an enclosed cabin equipped with a controlled lighting in order to achieve a consistent image in all captures. The lighting was a halogen lamp with some filters to get a diffused light simulating the illuminant D65 (6,500 K). This is considered the standard by the International Commission on Illumination, which describes the ordinary lighting conditions of a cloudy day at midday. Direct lighting on olive fruit was avoided by means of a reflecting surface placed between the scene and the illuminant. The images were taken at a distance of approximately 45 cm in a neutral white background.

### 2.3 Methodology

The algorithms used were developed with the image processing toolbox, version 7.4.0 of Matlab (The MathWorks, Inc., Natick, MA, USA), together with the software Image Pro-Plus version 6.0 (MediaCybernetics, Inc).

The images were taken in duplicate through both channels, namely, IR (monochromatic) and visible (RGB image). This means that they are images whose pixels are specified by three values, one for each colour component (red, green and blue), which are used for MI determination.

IR images are used to distinguish damaged olives which are healthy in appearance and which cannot be distinguished in the visible systems. The difference between damaged skins can be clearly seen at near-infrared bands, which offered good contrast between the defect and the sound skin and were well suited for detecting internal skin damage [8].



**Figure 1.** Sample representation in RGB and infrared space obtained in the on-line assay conditions

### 2.3.1 Estimation of Maturity Index and external defects

The technique used was the segmentation, which is based on the identification of regions and edges using clusters of pixels selected according to different criteria (colour, boundary, grey levels, etc.) [61, 62].

In order to develop a method for the on-line prediction of maturity index and sanitary conditions, the following steps have been carried out:

#### 2.3.1.1 Identification and separation of objects in the NIR image

The procedure of separation of objects consists in the use of the NIR monochrome image in order to segregate every single olive fruit. To identify objects in the NIR monochrome image, the first step performed is a contrast equalization followed by a spatial filter enhancement 'Flatten', which reduces the intensity variations in the background pixels. Spatial filtering operation was applied to the images to enhance or attenuate spatial detail in order to enhance visual performance or to facilitate further processing. The spatial filtering operation can be considered as a "local" procedure in image processing, in the sense of changing the value of each pixel according to the values of the surrounding pixels. From a practical point of view, this step mainly consists in classifying the original grey levels in the original NIR image according to the value in the neighbouring pixels. This task was carried out thanks to the use of two algorithms: on the one hand, Border-4 Neighbour, which determines the edges of the olive, taking into account the intensity values of pixels in the histogram and considering the darker pixels as edges of objects; on the other hand, a connected components algorithm which is used to identify each olive with a value. This edge-based algorithm enumerates each set of pixels which represent an olive [63].

#### 2.3.1.2 Creation of a mask and application of the mask to the original image

From the data in the array of connected components, a mask can be created. In addition, various size filters can be applied, so that we can measure complete objects or automatically remove the objects which overlap with an area smaller or bigger than

those with a standard size olive [64]. Based on these objects, the skeleton of image is constructed. The skeleton is intended to represent the shape of an object with a relatively small number of pixels. Thus, all pixels of the skeleton are structurally necessary. The position, orientation and length of the skeleton lines correspond to those equivalents to the original image; this skeleton simplifies the task of extracting features of an image.

This mask reduces the image to a skeleton and eliminates protrusions. The original image is added to the mask and the colour-segmentation and grey-segmentation are performed in separated objects.

### 2.3.1.3 Segmentation based on colour/grey and classification

This segmentation based on colour can be performed by supervised or unsupervised methods in order to get and quantify the predominant colour in the olives. This step requires transforming the RGB image using various functions to transform RGB format in  $L^*a^*b$  colour space. The CIE  $L^*a^*b^*$  colour notation system was applied to determine the parameters  $L^*$ ,  $a^*$  and  $b^*$ ; where  $L^*$  indicates the lightness,  $a^*$  means the colour axis from green to red and  $b^*$  refers to the blue-yellow one.

On the one hand, in the supervised methods a region containing the colour of interest (colour markers) is selected and then averaged. This classification is performed by using the method of the k-nearest neighbours (KNN) [65], where each pixel is classified in the same group as the colour markers with a similar intensity. The k nearest neighbours assign a value of 'a' and 'b' for each marker and, therefore, it is possible to classify each pixel of the image so as to calculate the Euclidean distance between pixels and colour markers.

On the other hand, unsupervised methods for colour segmentation techniques use clustering algorithms, which essentially perform the same functions as the classifying methods, but without using data training [66, 67]. In order to compensate for the lack of data training, clustering methods iterate between image segmentation and characterize the properties of each class. In this sense, clustering methods are trained with the available data. The basic idea is to assume that image pixels are points in a three dimensional space (RGB). Therefore, the points of a region of similar colour are grouped together. Clustering techniques allow us to obtain a representative point of

each group, based on measures of similarity or distance between these points. In this algorithm, a point  $x_i$  is assigned to a group  $R$ , whose centroid  $C$  is closer to it. The Euclidean distance is usually used as a measure for the similarity between the point and the centroid. The results of this classification are normally expressed in terms of the total percentage of pixels coloured in each object. Thanks to these results, fruits can be classified according to their level of maturity (Table 2).

Class	Olive skin Colour
0	>50% Bright green
1	>50% Green-yellowish
2	%Green-yellowish with black and/or Spots > %Reddish-brown
3	%Green-yellowish with black and/or Spots < %Reddish-brown
4	100 % Blackish purple or black

**Table 2** Olive classification based on olive skin

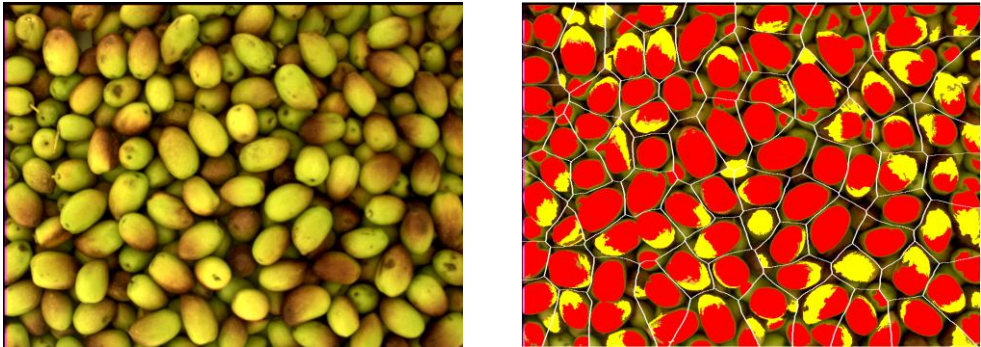
In the case of the determination of external defects, the skeleton is added to the original image of infrared and marks the position of the defects in the original image. To that end, we have a gradient-based segmentation of grey levels, which indicates that the values with the highest values of pixels represent the area of pixel defects and the lowest values represent the healthy areas of the fruit.

This allows us to measure some geometric properties of the defects, specifically, the defect area in each olive. After an evaluation of the area of defect, the results are collected in percentages as the sum of all pixels in areas occupied by the defects and the sum of all pixels in areas occupied by the sound areas of the olives.

### 3. Results and Discussion

Different sets of olives were analysed aiming at developing the proposed approach. The procedure described in the previous section was developed for each batch of olives.

The procedure was performed for the first time in fifty images captured directly on the olive production belt. These algorithms were applied to get the maturity index and quality state of each fruit. The results were obtained as the percentage of each colour and percentage of defects present in the object. These are classified according to different groups. Figure 2 shows an example of analysed on-line images. The results display the classification of olives for colour segmentation.



**Figure 2.** Original on-line image and image with analysed results with pixel index of different classes in objects (yellow=%Reddish-brown; red=% green-yellowish)

Object	% Bright-Green	% Green-yellowish	% Redish-Brown	% Black	Estimated RI	Measured Visual RI
1	0,00	29,98	36,47	33,55	3	3
2	0,00	0,00	59,66	37,86	3	4
3	0,00	79,60	0,00	0,00	1	1
4	0,00	0,00	20,23	77,91	4	4
5	0,00	0,00	56,78	42,13	3	3

6	0,00	0,00	0,00	89,61	4	4
7	0,00	0,00	0,00	94,69	4	4
8	0,00	0,00	54,26	45,73	3	3
9	0,00	39,13	23,68	37,19	2	2
10	25,31	65,80	0,00	0,00	1	1
11	0,00	0,00	70,14	24,47	1	1
12	0,00	58,45	0,00	41,55	3	3
13	0,00	21,78	41,37	36,84	2	2
14	0,00	0,00	71,88	0,00	1	1
15	23,68	42,92	0,00	56,97	4	3
16	24,05	39,13	37,19	0,00	2	2
17	0,00	67,12	0,00	0,00	1	1
18	0,00	0,00	59,66	37,86	3	4
19	0,00	0,00	35,96	64,04	4	4
20	66,40	21,54	0,00	0,00	0	0

**Table 3** Results of the maturity index estimated and visual calculated of some samples analysed

The results of prediction of the maturity index of fruit by the proposed vision method for images with individual olives are strongly related to visual RI measured.

The errors of prediction have been almost negligible or zero values for other samples, (see results in Table 2). These values are acceptable given the subjectivity which characterizes the MI on-line evaluation by the traditional method.

The determination of defects detection was carried out as described in the previous section. The results of some on-line images were analysed and they were obtained as a percentage of sound areas and defects in each olive.

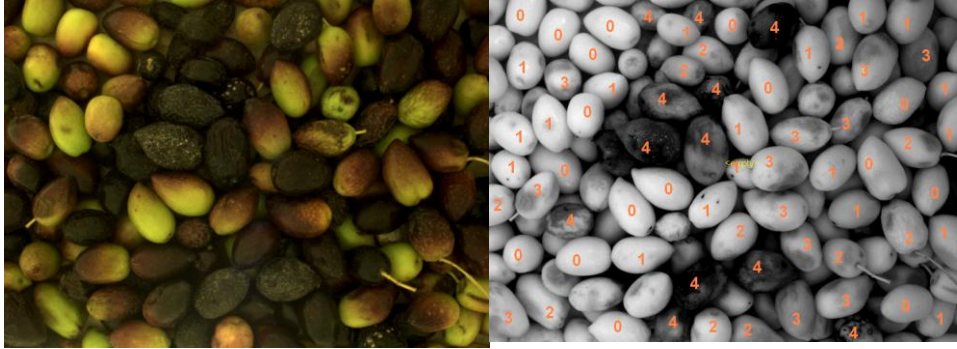
Each pixel of the fruit is classified into the different groups: sound olives or defective olives, according to the values of its spectral components. For each olive, results were obtained as the total percentage of defects and the total percentage of sound skin.

In order to classify olives according to their degree of defect, five different categories have been agreed. Table 2 defines those categories as follows: sound olives, olives with minor defects, olives with moderate defects, olives with several defects and defective olives.

Category		
0	Sound Olives	100% sound area
1	Minor defects	>75% sound area
2	Moderate defects	>50 % defective area
3	Several defects	>75% defective area
4	Defective Olives	100 % defective area

**Table 4** Olive classification based on percentage of sound and defective skin

Figure 3 shows an on-line image analysed with its corresponding results of classification. The results of classification of the olives to various images of olives analysed are shown in table 6.



**Figure 3** Results of classification for an on-line image analysed

Images	% Total defects	% Total sound	TotalOlives Detected	Class 0	Class 1	Class 2	Class 3	Class 4
Img 1	10,00	90,00	139	90	10	0	0	0
Img 2	26,24	73,76	111	36	58	17	0	0
Img 3	28,08	71,92	125	43	55	27	0	0
Img 4	40,80	59,20	112	7	15	58	32	0
Img 5	57,31	42,69	102	21	25	22	28	6
Img 6	63,45	36,55	103	12	34	12	35	10
Img 7	64,73	35,27	105	21	12	36	24	12
Img 8	67,70	32,30	138	17	21	61	23	16
Img 9	69,33	30,67	134	7	15	68	32	12
Img 10	81,49	18,51	110	14	15	26	32	23

**Table 6** Results of classification for some samples of olives by the proposed automatic method

According to these results, the proposed methodology can potentially be used for the classification of fruit in the quality control in the on-line extraction process of olive oil, which consists in the capture of an on-line image and its analysis. This process can occur at acceptable speeds and, therefore, this shall be the future aim of studies.

The described method provides an estimate of the MI value. This procedure requires a negligible amount of time; it is possible to state that the only time required is that needed to collect the olive samples from the trees. It is important to bear that in mind, given the fact that experts working in oil mills perform a real-time monitoring of the agronomical and technological parameters. In some cases, however, the purpose of



oil mills is to increase the quality of the extracted olive oil by optimizing some technological parameters.

The development of classification systems for non-destructive methods is applied to each fruit and ensures that their characteristics meet the standards. Concretely, optical techniques offer possibilities in the assessment of surface characteristics and inner qualities or composition. This system can be potentially used on-line in the classification of olives, which means that it would help to improve the quality control of olive oil in factories. Additionally, the systems used to measure optical properties provide significant information about the characteristics of each fruit, which is useful for the assessment of products.

The future of spectral systems applied to food inspection is promising, as both industry and consumers are increasingly aware of the need to ensure quality and food safety. Consequently, this technology is an essential tool for automatic inspection and control of these parameters.

### **Acknowledgment**

The authors thank the National Institute of Research and Technology Agriculture and Food (INIA) for funding of the scholarship of training FPI sub-INIA, ESF co-financed through the ESF Operational Programme for Andalusia 2007-2013 within axis 3 "Increasing and improving human capital" expenditure category "Developing human potential in the field of research and innovation."



### **References**

- [1] Ting SV, Rouseff RL (1986) *Anal Technol* 3:1–9
- [2] Casas A, Mallent D (1988) *Agrochim Technol Aliment* 28(2):184–201
- [3] Gnanasekharan V, Shewfelt RL, Chinnan MS (1992) *J Food Sci* 57(1):149–154
- [4] Taylor MA, Rabe E, Jacobs G, Dodd MC (1995) *Post Biol Technol* 5:285–294
- [5] Lehman-Salada L (1996) *HortScience* 31(6):994–997
- [6] Mickelbart MV, James D (2003) *New Zealand J Crop HortSci* 31:269–276
- [7] Ranalli A, Tombesi A, Ferrante ML, de Mattia G (1997) *Revista delle Sostanze Grasse* 74:553–557

- [8] Lu, R., and Y. R. Chen. Hyperspectral imaging for safety inspection of food and agricultural products. *SPIE* 3544:121–133 (1998)
- [9] Olmo M, Nadas A, Garc\_a JM (2000) *J Food Sci* 65:365–369
- [10] Garc\_a JM, Medina RJ, Ol\_as, JM (1998) *J Food Sci* 63(6):1037–1041
- [11] Muramatsu N, Sakurai N, Yamamoto R, Nevins DJ (1996b) *HortScienceS* 31(7):1199–1202
- [12] Garc\_a JM, Sellar S, P\_rez-Camino MC (1996a) *J Agric Food Chem* 44:3516–3520
- [13] Uceda, M, Frias L (1975) Harvest dates. Evolution of the fruit oil content, oil composition and oil quality. In: *Proceedings of the II Seminario Ole\_cola Internacional; International Olive Oil Council: Cordoba, Spain*, pp 125–130
- [14] Tovar MJ, Romero MP, Alegre S, Girona J, Motilva MJ (2002) *J Sci Food Agric* 82:1755–1763
- [15] Ayora-Cañada, M.J., Muik, B., García- Mesa, J.A., Ortega-Calderón, D., Molina-Díaz, A. Fourier-transform near-infrared spectroscopy as a tool for olive fruit classification and quantitative analysis. *Spectroscopy Letters*. 2005, 38, 6, 769-785
- [16] Cayuela, J.A., Garcia, J.M., Caliani, N. NIR prediction of fruit moisture, free acidity and oil content in intact olives
- [17] Cayuela, J.A., Camino, M.D.P. Prediction of quality of intact olives by near infrared spectroscopy. *European Journal of Lipid Science and Technology*. 2010, 112, 11, 1209-1217
- [18] Leon, L., Rallo, L., Garrido, A., Near-Infrared Spectroscopy (NIRS) analysis of intact olive fruit: an useful tool in olive breeding programs. *Grasas y Aceites*. 2003, 54, 1, 41-47
- [19] Bendini, A., Cerretani, L., Di Virgilio, F., Belloni, P., Boloni-Carbognin, M., LerckerG. (2007). Preliminary evaluation of the application of the ftir spectroscopy to control the geographic origin and quality of virgin olive oils. *Journal of Food Quality* 30 (4) 424437
- [20] Korifi, R., Le Dreau, Y., Molinet, J., Artaud, J., Dupuy, N. Composition and authentication of virgin olive oil from French PDO regions by chemometric treatment of Raman spectra. *Journal of Raman Spectroscopy*. 2011, 42, 7, 1540-1547
- [21] Dellaa, Y., Korifi, R., Le Dreau, Y., Artaud, J., Dupuy, N. Prediction of Geographical Origin of Virgin Olive Oil RDOs by Chemometric Treatment of Raman Spectra. *XXII International Conference on Raman Spectroscopy, AIP Conference Proceedings*.2010, 1267, 562-563
- [22] Zhang, X.F., Qi, X.H., Zou, M.Q., Liu, F. Rapid Authentication of Olive Oil by Raman Spectroscopy Using Principal Component. *Analysis Analytical Letters*. 2011, 44, 12, 2209-2220
- [23] El-Abassy, R.M., Donfack, P., Materny, A. Visible Raman spectroscopy for the discrimination of olive oils from different vegetable oils and the detection of adulteration *Journal of Raman Spectroscopy*. 2009, 40, 9, 1284-1289

- [24] Zou, M.Q., Zhang, X.F., Qi, X.H., Ma, H.L., Dong, Y., Liu, C.W., Guo, X., Wang, H. Rapid Authentication of Olive Oil Adulteration by Raman Spectrometry. *Journal of Agricultural and Food Chemistry*. 2009, 57, 14, 6001-6006
- [25] El-Abassy, R.M., Donfack, P., Materny, A. Assessment of conventional and microwave heating induced degradation of carotenoids in olive oil by VIS Raman Spectroscopy and classical methods *Food Research International*. 2010, 43, 3, 694-700
- [26] El-Abassy, R.M., Donfack, P., Materny, A. Rapid Determination of Free Fatty Acid in Extra Virgin Olive Oil by Raman Spectroscopy and Multivariate Analysis. *Journal of the American Oil Chemists Society*. 2009, 86, 6, 507-511
- [27] Muik, B., Lendl, B., Molina-Diaz, A., Ayora-Canada, M.J. Direct, reagent-free determination of free fatty acid content in olive oil and olives by Fourier transform Raman spectrometry. *Analytica Chimica Acta*. 2003, 487, 2, 211-220
- [28] Paiva-Martins, F., Rodriguez, V., Calheiros, R., Marques, M.P.M. Characterization of antioxidant olive oil biophenols by spectroscopic methods. *Journal of the Science of Food and Agriculture*. 2011, 91, 2, 309-314
- [29] Muik, B., Lendl, B., Molina-Diaz, A., Ortega-Calderon, D., Ayora-Canada, M.J. Discrimination of olives according to fruit quality using Fourier transform Raman spectroscopy and pattern recognition techniques. *Journal of Agricultural and Food Chemistry*. 2004, 52, 20, 6055-6060
- [30] Amirante P., Di Rienzo G. C., Di Giovacchino L., Bianchi B., Catalano P., Evoluzione tecnologica degli impianti di estrazione dell'olio di oliva, *OLIVÆ*, Vol.48, 1993, pp. 43-53.
- [31] Amirante P., Catalano P., Analisi teorica e sperimentale dell'estrazione dell'olio d'oliva per centrifugazione, *Rivista Italiana delle Sostanze Grasse*, Vol. LXX, 1993, pp. 329-335.
- [32] Catalano P., Cini E., Sarghini F., Applicazione di modelli nell'agroindustria per la ricerca e la gestione dei sistemi produttivi, *Proceedings of "Innovazione delle macchine e degli impianti nel settore agroalimentare per un'agricoltura multifunzionale nel rispetto dell'ambiente"*, Anacapri 5-6 Giugno 2006.
- [33] Furferi R., Carfagni M., Daou M., Artificial neural network software for real-time estimation of olive oil qualitative parameters during continuous extraction, *Computers and Electronics in Agriculture* Vol. 55, Issue 2, February 2007, pp. 115-131.
- [34] Cini E., Daou M., Furferi R., Recchia L., A modelling approach to extra virgin olive oil extraction, *J. of Ag. Eng. - Riv. di Ing. Agr.*, Vol. 4, 2007, pp. 1- 10.

- [35] Bordons C., Nunez-Reyes A., Model based predictive control of an olive oil mill, *Journal of Food Engineering* Vol. 84, Issue 1, January 2008, pp. 1-11. [7] Molina J.A., Pascual M.I., Using optical NIR sensor for on-line virgin olive oils characterization, *Sensors and Actuators B: Chemical* Vol. 107, 2005, pp. 64–68.
- [36] Molina J.A., Pascual M.I., Using optical NIR sensor for on-line virgin olive oils characterization, *Sensors and Actuators B: Chemical* Vol. 107, 2005, pp. 64–68.
- [37] AJ Marquez, AM Díaz Using optical sensor for on-line virgin olive oil characterization . *Sensors and Actuators B: Chemical*, 2005 - Elsevier
- [38] AJ Márquez Monitoring carotenoid and chlorophyll pigments in virgin olive oil by visible-near infrared transmittance spectroscopy. On-line application. - *Journal of near infrared spectroscopy*, 2003 –
- [39] Hermoso, M. Uceda Garcia-ortiz. Jimenez A. Beltran G. Preliminary results of « on-line» measure of oil content and humidity in olive cakes from the two phases decanted
- [40] Sun, D.-W., 2000. Inspecting pizza topping percentage and distribution by a computer vision method. *Journal of Food Engineering* 44, 245\_/249.
- [41] Cordero, S., Lleo, L., Barreiro, P., Ruiz-Altisent, M. 2006. Peach multispectral images related to firmness and maturity. In: *World Congress: Agricultural Engineering for a Better World*. Bonn, Germany
- [42] Blasco, J., Aleixos, N., Molto, E., 2007. Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm. *Journal of Food Engineering* 81, 535-543.
- [43] Kondo, N , Ahmad, U., Monta, M., Murasc, H , 2000. Machine vision based quality evaluation of lyokan orange fruit using neural networks. *Computers and Electronics in Agriculture* 29, 135-147.
- [44] Rosenberger, C, Emile, B., Laurent, H, 2004. Calibration and quality control of cherries by artificial vision. *Journal of Electronical Imaging* 13 (3), 539-546.
- [45] Uthaisombut, P. 1996. Detecting defects in cherries using machine vision. M.S. Thesis, Computer Science, Michigan State University.

- [46] Cheng, X., Tao, Y., Chen, Y.R., Luo, Y., 2003. NIR/MIR dual-sensor machine vision system for online apple stem-end/calyx recognition. *Transactions of ASAE* 46 (2), 551-558.
- [47] Kavdir, I., Guyer, D.E., 2004. Comparison of artificial neural networks and statistical classifiers in apple sorting using textural features. *Biosystem Engineering* 89 (3), 331-344.
- [48] Mehl, P.M., Chen, Y.R., Kim, M.S., Chan, D.E., 2004. Development of hyperspectral imaging technique for the detection of apple surface defects and contaminations. *Journal Food Engineering* 61, 67-81.
- [49] Brosnan, T., Sun, D.W., 2004. Improving quality inspection of food products by computer vision-a review. *Journal of Food Engineering* 61, 3-16.
- [50] Pla, F., Sanchez, J.S., & Sanchiz, J.M. (1999). On-line machine vision system for fast fruit colour sorting using low-cost architecture. In: VIII SPIE Conference on Machine Vision Systems for Inspection and Metrology, Boston, Massachusetts, USA.
- [51] Mendoza, F., Aguilera, J.M., 2004. Application of image analysis for classification of ripening bananas. *Journal of Food Science* 69 (9), 471–477.
- [52] Oberti, R., Guidetti, R., Mignani, I., 1999. Analisi multispettrale di immagini per l'individuazione precoce di danni meccanici su prodotti ortofrutticoli: un'applicazione su pere decana. *Rivista di Ingegneria Agraria* 30 (3), 137–147.
- [53] Peri, G., Romaniello, R., Amodio, M., & Colelli, G. (2003). The application of a fast colour quantization algorithm for sorting peaches for firmness. In: XXX CIOSTACGIR V CONFERENCE, Torino, Italy.
- [54] Abdullah, M.Z., Mohamad-Saleh, J., Fathinul-Syahir, A.S., Mohd-Azemi, B.M.N., 2006. Discrimination and classification of fresh-cut starfruits (*Averrhoa carambola* L.) using automated machine vision system. *Journal of Food Engineering* 76 (4), 506–523.
- [55] Koc, A.B., 2007. Determination of watermelon volume using ellipsoid approximation and image processing. *Postharvest Biology and Technology* 45 (1), 366–371.
- [56] Stajnko, D., Cmelik, Z., 2005. Modelling of apple fruit growth by application of image analysis. *Agriculturae Conspectus Scientificus* 70 (1), 59–64.
- [57] Bodria, L., Fiala, M., Guidetti, R., & Oberti, R. (2002). Optical techniques for assessing the fruit maturity stage. In: ASAE Annual International Meeting/CIGR XVth World Congress, Chicago, USA.
- [58] Choong, T.S.Y., Abbas, S., Shariff, A.R., Halim, R., Ismail, M.H.S., Yunus, R., Salmiaton, A., Fakhru-Razi, A., 2006. Digital image processing of palm oil fruits. *International Journal of Food Engineering* 2 (2), 7.
- [59] E Guzmán, V Baeten, JAF Pierna A portable Raman sensor for the rapid discrimination of olives according to fruit quality ... - *Talanta*, 2012 – Elsevier
- [60] Garc\_a JM, Guti\_rrez F, Barrera MJ, Albi MA (1996b) *J Agric Food Chem* 44:590–593

- [61] Medioni, G.; Lee, M-S; Tang, C\_K.; "A Computational Framework for Segmentation and Grouping", Elsevier Science B.V., 2000. ISBN: 0-444-50353-6.
- [62] Plataniotis, K.N.; Venetsanopoulos, A.N.; "Color Image Processing and Applications", Springer Verlag, 2000. ISBN: 3-540-66953-1.
- [63] Chang, V. and Saavedra, J. (2001) Métodos alternativos para el mejoramiento automático del contraste de imágenes. Computer Engineering thesis, School of Informatics, University National of Trujillo
- [64] Russ, J.C. The image processing Handbook. CRC Press LLC 3rd Edition.
- [65] Richards, J. A.. Remote Sensing Digital Image Analysis. Ed. Springer-Verlag, 1999.
- [66] Thomas, S.W., 1991. Efficient inverse color map computation. In: Arvo, J. (Ed.), Graphics Gems II. Academic Press, Boston, pp. 116–125.
- [67] Floyd, R. W., & Steinberg, L. (1975). An Adaptive Algorithm for Spatial Grey Scale. In: International Symposium Digest of Technical Papers, Society for Information Displays, 36

