



Escola Tècnica Superior d'Enginyeria
de Telecomunicació de Barcelona

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Identificación automática de Espectros Raman de pigmentos mediante Análisis por Componentes Principales

Juan José González Vidal

Directora: Rosanna Pérez Pueyo

Proyecto Final de Carrera

Grup de Comunicacions Òptiques (GCO)

Departament de Teoria del Senyal i Comunicacions (TSC)

Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona (ETSETB)

Universitat Politècnica de Catalunya (UPC)

Agradecimientos:

Lejano parecía este día cuando inicié este viaje... Un viaje que emprendía con ilusión ante el reto que suponía todo lo que en Barcelona me aguardaba. Ha resultado ser una aventura llena de incertidumbres y de sacrificios, pero también de recompensas y de experiencias inolvidables. Y lo más importante, me ha brindado la oportunidad de conocer a una gente increíble, con la que sé que seguiré contando tras la etapa que concluye hoy con este proyecto.

Por ello, me gustaría agradecer a todos los amigos que me han acompañado en este trayecto, por ser las velas y el motor en este desafiante recorrido, con los que he compartido grandes momentos y que, poco a poco, se han convertido en parte de mi vida: Miquel, Xavi, Peter, Christian, Jus, Marta, Ester, Álvaro y Alberto. A mis compañeros de despacho, Gonzalo y Lorena, y al resto de compañeros, Maria, Adri, Jaume, Rafel y Víctor, por haber formado parte de este camino.

Al Grupo de Espectroscopía Raman, y en especial a María José, por facilitarme todos los recursos necesarios para la realización de este proyecto.

A Rosanna, por la confianza depositada en mí, por su esfuerzo y por toda la ayuda prestada. Por comprender mi compleja estructura mental y por su ilusión en el trabajo, que se ve reflejada en cada una de estas páginas.

A mis amigos de siempre, Laura, Juan Antonio y Lorenzo. Por todo lo que hemos vivido juntos y por seguir estando a mi lado aun la distancia.

A mis hermanos, Loly y Sebas, y a mis padres, por su apoyo incondicional y por creer siempre en mí, por sus sabios consejos y por hacer todo lo posible para permitirme cumplir mis sueños. Porque si estoy donde estoy es gracias a ellos.

A todos, muchas gracias.

ÍNDICE GENERAL

RESUMEN - RESUM - ABSTRACT	5
0. INTRODUCCIÓN.....	7
0.1. Justificación	7
0.2. Objetivos	12
0.3. Contenido	14
1. ESPECTROSCOPIA RAMAN APLICADA AL ARTE	17
1.1. Espectro Raman	17
1.2. Señal y ruido en Espectroscopía Raman	19
1.3. Aplicación a la identificación de pigmentos artísticos.....	21
2. ANÁLISIS POR COMPONENTES PRINCIPALES. APLICACIÓN AL TRATAMIENTO DE ESPECTROS RAMAN.....	27
2.1. Fundamentos del ACP. ACP normado y no normado	27
2.2. Calidad de la representación: contribución y coseno cuadrado	32
2.3. Aplicación al tratamiento de espectros Raman	34
3. RECONOCIMIENTO/IDENTIFICACIÓN DE ESPECTROS RAMAN	41
3.1. Planteamiento general: Metodología para la identificación.....	41
3.2. Cuantificación del parecido entre espectros: Distancia Euclídea. Selección de candidatos.....	49
3.3. Requerimientos/criterios para la identificación	52
4. ESTUDIO TEÓRICO DE LA METODOLOGÍA	59
4.1. Librería espectral de referencia. Importancia de la elección de la librería de Espectros Raman patrones	59
4.2. Estudio de identificación de Espectros Raman	69
4.2.1. Espectro Raman desconocido igual a un espectro patrón	69
4.2.2. Espectro Raman desconocido distinto a cualquier espectro patrón	70
4.3. Estudio de los efectos sobre la identificación de distintos problemas prácticos. Límite para la identificación	72

4.3.1. Influencia del ruido. SNR mínima para la identificación.....	72
4.3.2. Influencia de la anchura de banda de las Bandas Raman.....	84
4.3.3. Influencia de la variación de las intensidades relativas de las bandas	92
4.3.4. Bandas Raman secundarias enmascaradas: Identificación con solo la Banda Raman fundamental	98
4.3.5. Influencia del desplazamiento de las Bandas Raman.....	103
4.3.6. Presencia de Bandas Raman adicionales: Mezclas de pigmentos.....	109
4.4. Estudio para la identificación de mezclas de pigmentos.....	111
5. APLICACIÓN DE ACP A LA IDENTIFICACIÓN DE ESPECTROS RAMAN DE PIGMENTOS ARTÍSTICOS. CASOS EXPERIMENTALES	117
5.1. Librería espectral de referencia: ¿Cómo escoger los pigmentos?	117
5.2. Identificación de espectros de pigmentos de una paleta básica	117
5.3. Identificación de Espectros Raman de pigmentos con un número elevado de Bandas Raman.....	132
6. RESUMEN Y CONCLUSIONES	147
BIBLIOGRAFÍA Y REFERENCIAS	151
ANEXOS	155
A1. Resultados obtenidos en la reducción de la anchura de las bandas	155
A2. Resultados obtenidos en el desplazamiento de las bandas.....	159
B. Principales funciones implementadas en <i>Matlab</i>	161
C1. Paleta básica: Características de la librería espectral de referencia	169
C2. Paleta de pigmentos amarillos orgánicos: Características.....	179

Resumen

La Espectroscopía Raman es una técnica fotónica que permite caracterizar unívocamente compuestos químicos a partir de la luz dispersada por éstos al incidir sobre ellos un haz de luz monocromática. Debido principalmente a su carácter no-destructivo, esta técnica ha incrementado exponencialmente su aplicación en la identificación de materiales artísticos. Con el objetivo de facilitar y agilizar el proceso de identificación, este Proyecto Final de Carrera presenta un sistema automático de identificación de Espectros Raman de pigmentos mediante la técnica de reducción dimensional del Análisis por Componentes Principales. Para comprobar el comportamiento de la metodología propuesta se simula una librería espectral y se analiza el efecto de distintos problemas prácticos sobre la identificación. Por último, se verifica el correcto funcionamiento del sistema implementado en casos experimentales de Espectros Raman obtenidos sobre obras de arte.

Resum

L'Espectroscòpia Raman és una tècnica fotònica que permet caracteritzar unívocament compostos químics a partir de la llum dispersada per aquests quan incideix sobre ells un feix de llum monocromàtic. Degut principalment al seu caràcter no-destructiu, aquesta tècnica ha incrementat exponencialment la seva aplicació en la identificació de materials artístics. Amb l'objectiu de facilitar i agilitzar el procés d'identificació, aquest Projecte Final de Carrera presenta un sistema automàtic d'identificació d'Espectres Raman de pigments mitjançant la tècnica de reducció dimensional d'Anàlisi per Components Principals. Per tal de comprovar el comportament de la metodologia proposada se simula una llibreria espectral i s'analitza l'efecte sobre la identificació de diferents problemes pràctics. Finalment, es verifica el correcte funcionament del sistema implementat en casos experimentals d'Espectres Raman obtinguts sobre obres d'art.

Abstract

Raman Spectroscopy is a photonic technique that allows to univocally characterize a chemical compound, providing a signal known as Raman Spectrum. This technique has increased its application at identifying constituent materials in artworks thanks to its nondestructive nature mainly. With the aim of speeding up the process of identification, this Master Thesis presents an automatic identification system of Raman Spectra of pigments through the dimensional reduction technique of Principal Components Analysis. In order to validate the behaviour of the proposed methodology, it is simulated a spectral database, studying the effect of different practical problems on spectral identification. Finally, it is verified the correct behaviour of the implemented system at experimental cases of Raman Spectra from pigments obtained on artworks.

0. INTRODUCCIÓN

0.1. Justificación

A lo largo de la historia, el hombre ha utilizado todas las herramientas disponibles para la creación de obras artísticas. En algunas culturas ancestrales, incluso, existía un solo vocablo para expresar los conceptos que hoy en día conocemos como tecnología, ciencia y arte. Así, *arte* proviene del latín “ars” y *técnica* del griego “tekne”, palabras que se refieren a la habilidad para realizar una determinada tarea o alcanzar un cierto objetivo^[1]. De esta forma, aquello que denominamos arte, se ha desarrollado a medida que el hombre ha ido dominando y manejando diversos materiales.

Diferentes civilizaciones han dejado su legado usando técnicas gráficas similares a las que se utilizan hoy en día, desarrollando, desde las eras más primitivas, tintes y pinturas con un alto grado de sofisticación, generando así lo que conocemos como *patrimonio cultural*^[2]. Dicho término describe todo lo referido a la historia de la civilización, incluyendo todo aquello considerado de interés desde puntos de vista tan diversos como el arqueológico, el histórico o el artístico. Estos diferentes puntos de vista confluyen en el interés por el conocimiento, compartido también por la ciencia, fundamentada en el rigor científico con el fin de obtener un determinado resultado que valide, certifique o demuestre cierta hipótesis.

La investigación científica sobre obras de arte se realiza por diferentes razones^[3], entre las que cabe destacar el interés artístico e histórico en conocer los materiales utilizados por un determinado artista, puesto que dichos materiales varían en función de su localización geográfica o cronológica.

Un conocido hallazgo arqueológico en Creta, conocido comúnmente como el Sarcófago Pintado de Haghia Triada (véase *Fig. 0.1*), de inmenso interés arqueológico, es considerado como una de las obras de arte más importantes de la civilización minoica (siglo XXVII A.C.-siglo XV A.C.). Estudios científicos y diferentes investigaciones (como el escáner microscópico o la difracción por rayos X, llevados a cabo por la

Scuola Archeologica Italiana di Atene, en 1998), aportaron información referente a la variedad de materiales pictóricos usados por los antiguos Minoans y qué técnicas fueron usadas en los frescos. Así mismo, se indicó el grado general de tecnología alcanzado por los Minoans, contribuyendo así a un mayor entendimiento de su nivel de civilización.

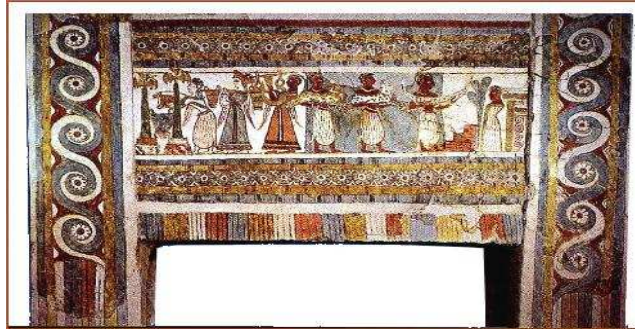


Fig. 0.1: Sarcófago Pintado de Haghia Triada (Creta)

Los materiales pictóricos o *pigmentos* han sido utilizados desde tiempos prehistóricos, y a lo largo de la Historia han sido fundamentales en las artes gráficas. Previamente a la Revolución Industrial, se disponía de una variedad limitada de materiales pictóricos disponibles para el arte y otros usos decorativos, puesto que la mayoría de pigmentos usados eran minerales, siendo de difícil preparación. Tal es el caso del pigmento conocido como Azul Ultramar^[4], asociado a la realeza por su elevado coste, debido a que la única forma de conseguirlo era a partir del *lapislázuli*, una piedra semipreciosa de localización remota. Debido al coste del lapislázuli, se hicieron muchos intentos por encontrar un pigmento azul menos costoso. El Azul de Prusia fue el primer pigmento sintético moderno, descubierto en 1704. A principios del siglo XIX, a las variedades existentes de azules se habían añadido gran cantidad de nuevos pigmentos azules, como, por ejemplo, las diversas formas de Azul de Cobalto.

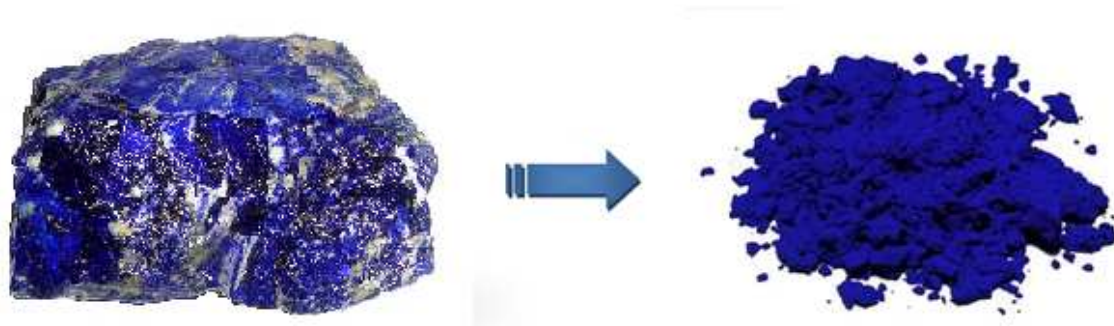


Fig. 0.2: Lapislázuli (a la izquierda) y pigmento Azul Ultramar (a la derecha)

Por todo ello, un estudio detallado de los pigmentos que constituyen una obra de arte es importante para revelar información referente a la caracterización, restauración, conservación, datación y autenticación de dicha obra^[5]. De hecho, este estudio puede ultimar la certificación de autenticidad, verificando la atribución artística e identificando falsificaciones, puesto que la introducción de determinados pigmentos está estrictamente datada. Por ejemplo, el Azul de Prusia no debería aparecer en un cuadro anterior al siglo XVIII, y el Blanco de Titanio, introducido en el siglo XX, sería sospechoso si apareciese en una paleta de colores perteneciente al siglo XIX. Consecuentemente, la identificación de pigmentos utilizados en obras de arte puede dar una indicación de la fecha de realización de la obra e incluso de su autenticidad.

Así, el conocimiento de los materiales pictóricos utilizados en una obra se convierte en primordial en el mundo del arte, donde puede determinar las condiciones óptimas de mantenimiento. Debido al efecto del transcurso del tiempo, la materia está sujeta a procesos de progresiva degradación. Es importante, por tanto, conocer qué procesos causan esta degradación e identificar indicadores que permitan encontrar métodos para detener el curso del daño material, asegurando así que las obras perdurarán el mayor tiempo posible.

La caracterización de los pigmentos es de vital importancia para el diseño de procesos seguros de conservación, así como para determinar condiciones ambientales adecuadas para su exposición, almacenamiento y transporte^[6]. Por ejemplo, una obra de arte que contenga los pigmentos Índigo y/o Carmín, ambos sensibles a la luz, requiere condiciones estrictas en cuanto a exhibición se refiere, como tenues niveles de luz y tiempos controlados de exposición. Cuando sean considerados tratamientos acuosos para estabilizar la celulosa en papel que ha sido degradado por mecanismos de oxidación, es importante identificar pigmentos que pueden ser afectados por estos procesos, como el Azul de Prusia, que puede volverse marrón, o el Amarillo de Cromo, que puede convertirse en naranja.

Durante los últimos años, la *Espectroscopía Raman* ha incrementado exponencialmente su aplicación en el mundo del arte. En este marco, esta técnica analítica, al utilizar un haz de luz como herramienta de estudio, brinda múltiples ventajas, proporcionando información cualitativa y cuantitativa en forma de una señal específica (cuya

representación es llamada *Espectro Raman*) sobre la composición molecular de los pigmentos de una determinada obra, sin necesidad de que ésta sufra ninguna alteración, no produciendo ningún daño sobre el objeto de estudio. El Espectro Raman es único para cada material y se puede obtener de gran parte de compuestos químicos. Además, la utilización de la fibra óptica para el guiado de la señal, permite el estudio de muestras *in situ*. Tanto por su flexibilidad como por la caracterización material unívoca y su carácter no destructivo, la Espectroscopía Raman se convierte en una técnica ideal para el análisis aplicado al mundo del arte.

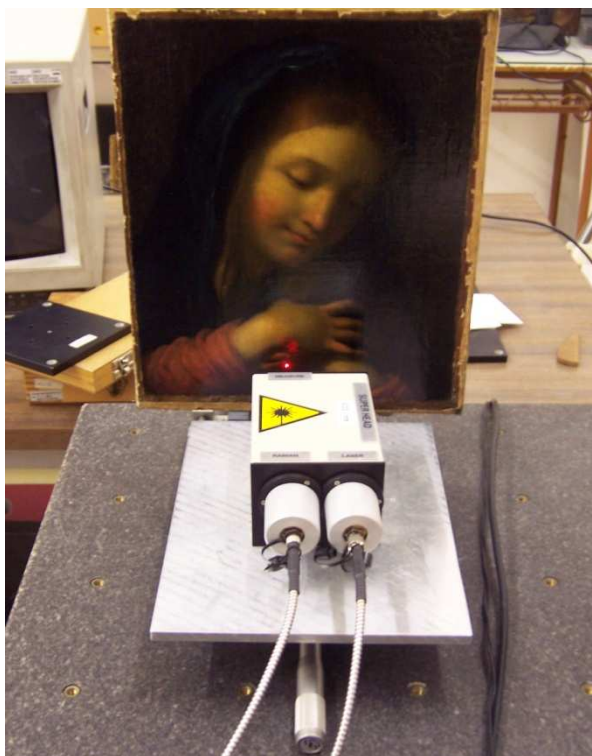


Fig.0.3: Ejemplo de obtención del Espectro Raman de un pigmento de una obra artística

La identificación de materiales a partir de la Espectroscopía Raman está basada en la comparación del Espectro Raman de un pigmento obtenido en una determinada obra con librerías espectrales de los materiales de referencia. Así, surgen diversos problemas relacionados con dicha comparación, puesto que cada laboratorio trabaja con diferentes características en los instrumentos de medida, como las longitudes de onda de los láseres o las rutinas de calibración. Además, la obtención del Espectro Raman puede resultar complicada, puesto que la señal contiene, a parte de la información referente al pigmento analizado, diferentes perturbaciones, como el *ruido shot* o la *fluorescencia*, que incluso pueden enmascarar la señal útil.

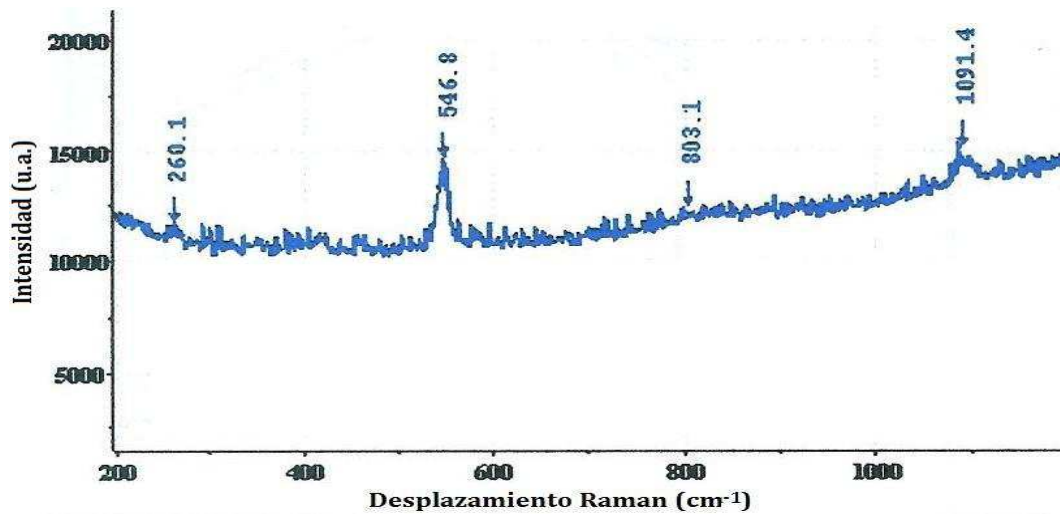


Fig.0.4: Espectro Raman del pigmento Azul Ultramar obtenido de una obra artística

Por otro lado, la utilización de algoritmos de comparación espectral puede resultar de gran complejidad computacional, debido a la alta cantidad de datos presentes en un Espectro Raman, apareciendo la necesidad de poder comprimir los datos, ocupando menos memoria y evitando la redundancia, sin por ello perder información. Por ello, se ha incrementado la utilización de métodos quimiométricos aplicados a la Espectroscopía Raman en el desarrollo de algoritmos de búsqueda espectral, como por ejemplo, el *Análisis por Componentes Principales* (ACP), método basado en la reducción de la dimensionalidad.

Este análisis es un método lineal óptimo para comprimir un conjunto de vectores de alta dimensión a otro de una dimensión muy inferior, sin pérdida de información y permitiendo la reconstrucción de los datos originales^[7]. Además, se trata de una herramienta estadística que permite detectar las similitudes y diferencias entre los datos, obteniendo una nueva expresión reducida respecto a la expresión inicial.

La arqueometría y el análisis de obras de arte es un campo de investigación desafiante, donde el principal objetivo es la extracción de la mayor cantidad de información posible de una obra única. Por ello, la Espectroscopía Raman es una de las técnicas que mejor se adapta a este tipo de estudios, por su examen molecular no-destructivo, que, conjuntamente con el método del *Análisis por Componentes Principales* y el algoritmo de búsqueda aquí propuesto, permiten la identificación de los materiales pictóricos, pudiendo proporcionar información útil para la caracterización, restauración, conservación, datación y autenticación de obras de arte.

Este proyecto se enmarca dentro de la investigación que realiza el Grupo de Comunicaciones Ópticas (GCO) del Departamento de Teoría de la Señal y Comunicaciones (TSC) en Espectroscopía Raman aplicada a la identificación de pigmentos. Dicha investigación es posible gracias a la subvención recibida de la Comisión Interministerial de Ciencia y Tecnología para su desarrollo (TEC 2009-07855), en el laboratorio de Espectroscopía Raman de la Universidad Politécnica de Catalunya (UPC).

0. 2. Objetivos

A la hora de analizar el espectro Raman de un pigmento de una obra de arte medido en un determinado laboratorio, surgen distintos obstáculos relacionados tanto con la medida de dicho espectro como con su comparación con los Espectros Raman de referencia. Así, a continuación se presentan los objetivos a tratar a lo largo de este proyecto:

- Diseñar un método de identificación automático de Espectros Raman utilizando la técnica de reducción dimensional basada en el Análisis por Componentes Principales y la herramienta matemática de la Distancia Euclídea:

Debido a la complicación del proceso de identificación por comparación espectral, puesto que existen multitud de pigmentos de referencia en las diferentes librerías espectrales, se propone diseñar un sistema de reconocimiento de Espectros Raman de materiales pictóricos, automatizando el proceso de identificación, independizándolo, de esta forma, de la subjetividad que pueda introducir el analista en base a su juicio y experiencia. Es decir, se pretende identificar objetivamente los pigmentos de la paleta artística utilizada en una obra de arte determinada, permitiendo así agilizar las tareas de mantenimiento, conservación y estudio de la obra.

- Estudiar la influencia de diversos problemas prácticos en la identificación:
 - **Influencia del ruido *shot***: En Espectroscopía Raman se considera ruido a todo lo que no depende de la estructura molecular del material al

obtenerse el espectro de una muestra determinada. De entre los diferentes tipos de ruido existentes, uno de los más habituales es el ruido *shot*, puesto que aparece ligado a la propia naturaleza aleatoria de la luz. Por ello, se realiza un estudio de la influencia de este ruido en la expresión reducida de los espectros generada por el ACP, estableciéndose una cota máxima de ruido para la identificación.

- **Influencia de la anchura de las Bandas Raman:** Durante el proceso de pretratado espectral, la anchura de las Bandas Raman puede verse reducida, pudiéndose dar el caso en el que un sistema de reconocimiento espectral basado en el ACP no llegue a identificar a un espectro desconocido afectado por dicho efecto. Por ello, se lleva a cabo un análisis de la reducción de la anchura de las Bandas Raman y su influencia en la expresión espectral generada por el ACP para establecer un límite máximo de reducción de la anchura de las bandas para un reconocimiento apropiado.
- **Variación de las intensidades relativas:** Las intensidades relativas entre Bandas Raman pueden variar de un espectro a otro del mismo material pero obtenido en un laboratorio diferente, puesto que la intensidad depende, entre otros factores, de las características y/o los parámetros de los instrumentos de medida utilizados a la hora de llevar a cabo el análisis sobre cierta muestra. A partir del estudio de la influencia de la variación de las intensidades relativas en la expresión de los espectros proporcionada por el ACP, se establece una relación límite entre la intensidad de la Banda Raman fundamental y las Bandas Raman secundarias que asegura la identificación.
- **Influencia del desplazamiento de las Bandas Raman:** Durante la etapa de pretratado espectral, puede ocurrir que las Bandas Raman se vean desplazadas respecto a su posición original. Por ello, se realiza un análisis del desplazamiento de las Bandas Raman y su influencia en la representación espectral del ACP, estableciéndose un límite para la correcta identificación.

- **Presencia de Bandas Raman adicionales: Mezcla de pigmentos:** Al obtenerse un Espectro Raman pueden aparecer Bandas Raman adicionales, que pueden ser generadas debido a que en la realización de obras de arte se utilizan frecuentemente mezclas de pigmentos. Se pretende que el sistema de identificación a implementar sea capaz de reconocer los pigmentos implicados en la mezcla.

El Análisis por Componentes Principales permite una reducción en la dimensionalidad de los Espectros Raman. Al aplicar dicho método sobre una librería espectral de referencia de pigmentos, se obtiene un nuevo espacio matemático de dimensión muy inferior a la del espacio original. No obstante, la proyección del Espectro Raman de un pigmento a analizar sobre este nuevo espacio puede implicar la pérdida de información al reducir éste su dimensión y/o al verse influido por los diferentes problemas prácticos descritos anteriormente. Por ello, se presentan, también, parámetros referentes a la calidad de la representación.

Así mismo, se cuantifica la similitud espectral mediante herramientas matemáticas basadas en la Distancia Euclídea, proporcionando requisitos y criterios para la identificación respecto a una determinada librería espectral. Finalmente se desarrolla un algoritmo de identificación, facilitando, también, un factor de confiabilidad para dicha identificación.

0.3. Contenido

El contenido de este proyecto se estructura en seis capítulos, en los que se desarrollan los objetivos descritos en el anterior apartado.

- **Capítulo 1: Espectroscopía Raman aplicada al arte.** En este primer capítulo se exponen brevemente los fundamentos del efecto Raman (fenómeno en el cual se basa la Espectroscopía Raman) haciendo énfasis en la representación del Espectro Raman y en los ruidos que pueden encontrarse presentes en él. De entre las múltiples aplicaciones de la Espectroscopía Raman, se trata la identificación de materiales artísticos, ya que es una de las aplicaciones más importantes y

adecuadas de esta técnica y es en la que se centra la investigación del grupo en el que se ha realizado este proyecto.

- **Capítulo 2: Análisis por Componentes Principales (ACP): Aplicación al tratamiento de Espectros Raman.** En este capítulo se presentan las bases matemáticas de la técnica de reducción dimensional conocida como Análisis por Componentes Principales (ACP) y se introduce su aplicación al tratamiento de Espectros Raman.
- **Capítulo 3: Reconocimiento de Espectros Raman.** Capítulo en el que se plantea la metodología para la identificación de Espectros Raman y se establecen determinados criterios para su reconocimiento, permitiendo implementar un sistema automático de identificación espectral.
- **Capítulo 4: Estudio teórico de la metodología.** En este capítulo se lleva a cabo un análisis teórico de la metodología expuesta y los efectos sobre la identificación de distintos problemas prácticos, obteniéndose una serie de resultados que permiten acotar las condiciones favorables para el reconocimiento espectral.
- **Capítulo 5: Aplicación del ACP a la identificación de Espectros Raman de pigmentos artísticos. Casos experimentales.** En el capítulo quinto se muestra la aplicación de la metodología desarrollada sobre casos experimentales.
- **Capítulo 6: Conclusiones.** Del trabajo desarrollado se extraen una serie de conclusiones expuestas en este sexto capítulo.

Finalmente, se adjunta un capítulo de bibliografía y referencias que recoge las fuentes consultadas, y un último capítulo de anexos que contiene información adicional, como resultados numéricos o el código *Matlab* implementado durante la realización de este proyecto.

1. ESPECTROSCOPIA RAMAN APLICADA AL ARTE

1. 1. Espectro Raman

Cuando un haz de luz monocromática incide sobre un material, gran parte de la radiación es absorbida, pero también existe una parte que es dispersada por la materia. Si se utiliza radiación monocromática o con una banda estrecha de frecuencia, ν_i , la frecuencia de prácticamente toda la energía dispersada coincide con la de la radiación incidente, dispersión elástica conocida como dispersión Rayleigh, que, aún siendo intensa, no aporta ningún tipo de información sobre la composición de la muestra analizada y por tanto no permite caracterizarla. Sin embargo, existe una pequeña parte de la luz dispersada que sí es característica del material, originando una dispersión no elástica, puesto que existe un intercambio de energía entre el fotón y la molécula durante la colisión. Esta dispersión, conocida como dispersión Raman, se presenta a ciertas frecuencias discretas situadas por encima y por debajo de la frecuencia incidente, $\nu_i \pm \nu_r$, siendo estas frecuencias $\pm \nu_r$ específicas para cada material, puesto que van ligadas a su estructura molecular y a la naturaleza de sus enlaces químicos^[8].

Para este tipo de dispersión no elástica se distinguen dos casos. Si la radiación es dispersada a una frecuencia menor a la de los rayos incidentes se conoce como radiación Raman Stokes. Si, por el contrario, la radiación es dispersada a una frecuencia mayor que la incidente, se llama radiación Raman Anti-Stokes (ver *Fig. 1.1* y *Fig. 1.2*).



Fig. 1.1: Interacción luz-materia

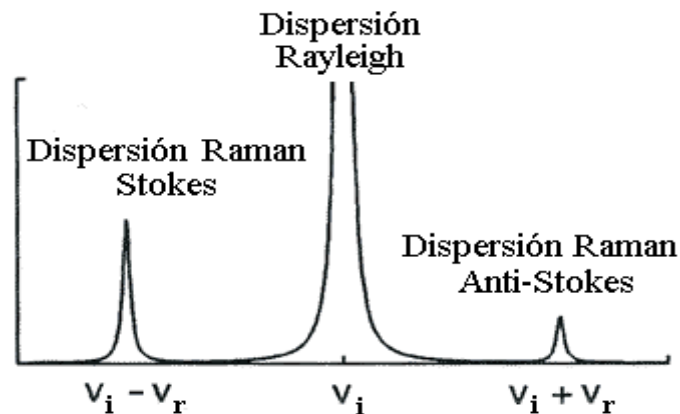


Fig. 1.2: Dispersión Rayleigh, Stokes y Anti-Stokes

La Espectroscopía Raman es la técnica fotónica que proporciona la señal dispersada por la materia al hacerle incidir un haz de luz monocromática (láser), conocida como *Espectro Raman*. De esta forma, un Espectro Raman consiste en la representación gráfica de la radiación Raman (o intensidad, en unidades arbitrarias, [u.a.]) en función del desplazamiento Raman (o número de onda, en $[\text{cm}^{-1}]$) relativo a la frecuencia de la luz incidente (ver Fig. 1.3).

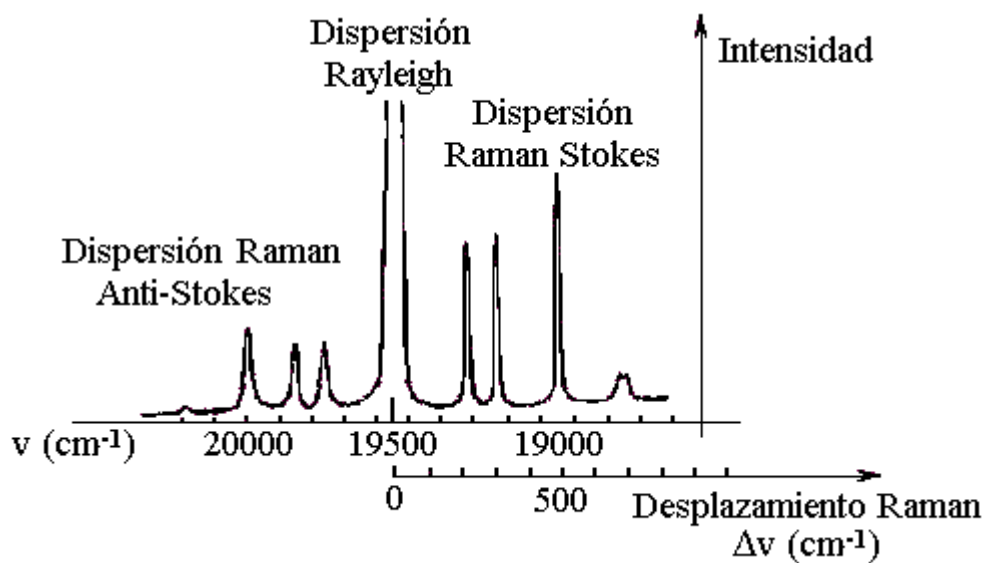


Fig. 1.3: Desplazamiento Raman

Ya que la intensidad de la dispersión Stokes es superior a la de la dispersión Anti-Stokes, habitualmente se representa únicamente el efecto Stokes, situándolo en la parte positiva del eje de ordenadas, y puesto que el desplazamiento de las frecuencias Raman respecto a la frecuencia incidente es independiente de dicha frecuencia, se acostumbra a situar la banda Rayleigh como origen.

El Espectro Raman es único para cada material y se puede obtener de casi cualquier compuesto químico. Esta representación permite visualizar unas bandas espectrales (Bandas Raman) centradas a las frecuencias Raman características de cada material (ver *Fig. 1.4*). En un Espectro Raman la información se encuentra principalmente en la posición de cada una de sus bandas que permiten identificar de manera inequívoca el material analizado.

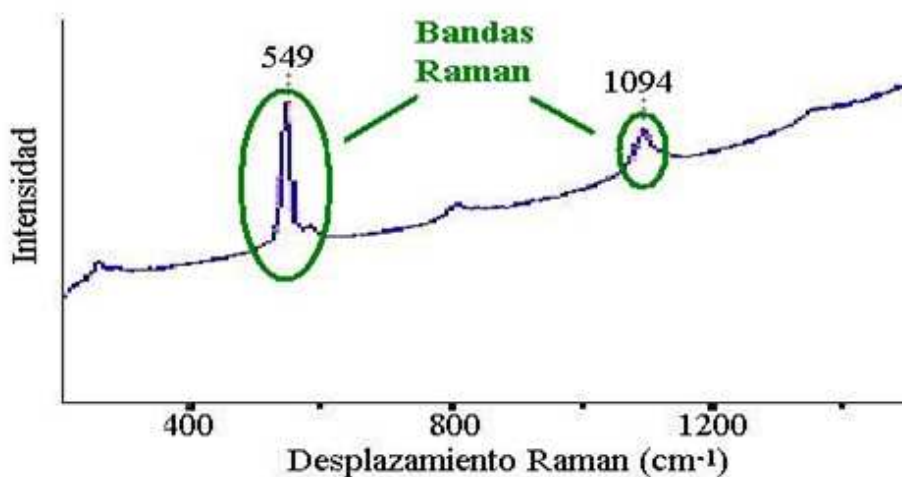


Fig. 1.4: Ejemplo de Espectro Raman obtenido en una obra artística con sus Bandas Raman señalizadas. Espectro correspondiente al pigmento Azul Ultramar

1. 2. Señal y ruido en Espectroscopía Raman

El Espectro Raman puede dividirse en dos partes: la señal útil y el ruido. La señal útil es la parte del Espectro Raman que contiene la información deseada (Bandas Raman), mientras que el ruido se considera como todo aquello que no corresponde a la dispersión Raman, es decir, que no depende de la estructura molecular del material, como por ejemplo las fluctuaciones aleatorias de intensidad, que representan la mayor fuente de incertidumbre en el análisis de una señal Raman.

Para llevar a cabo el análisis de un Espectro Raman es conveniente el aislamiento de la señal respecto al ruido, así como maximizar la señal Raman y minimizar el ruido en el espectro. Una importante métrica para describir la calidad del Espectro Raman es la relación señal a ruido (*SNR*).

El ruido en un Espectro Raman puede clasificarse en cinco grupos diferentes: Ruido *shot*, ruido generado por la muestra, ruido generado por el instrumento, ruido computacional y ruido generado por fuentes externas^[9]:

- **Ruido Shot.** El ruido shot es resultado de la naturaleza aleatoria de la luz. Se produce en el receptor debido a la conversión óptico-eléctrica. Su intensidad es igual a la raíz cuadrada del número de fotones detectados. Este ruido es ineludible en Espectroscopía Raman.
- **Ruido generado por la muestra.** Este tipo de ruido hace referencia a los cambios de intensidad Raman que son provocados por cambios en la muestra a analizar. Características como la forma de una Banda Raman, su posición o su intensidad, pueden verse modificadas por efecto de la temperatura de la muestra. La no homogeneidad de la muestra también puede producirlo, puesto que el análisis que se realiza en un punto concreto de la muestra no tiene por qué ser representativo de la muestra completa. También se consideran en este grupo de ruido las emisiones ópticas no deseadas que se generan por la propia naturaleza de la muestra, siendo la *fluorescencia* la más importante. La fluorescencia es una emisión óptica no deseada en Espectroscopía Raman, independiente de la frecuencia de radiación y que se produce si hay absorción del fotón incidente por parte del material.
- **Ruido generado por el instrumento.** Este tipo de ruido depende de las especificaciones de diseño del instrumento utilizado para analizar la muestra. Incluye diferentes ruidos del detector óptico.
- **Ruido computacional.** Este ruido hace referencia al ruido añadido al Espectro Raman durante el proceso de digitalización de la señal a la salida del detector.
- **Ruido generado por fuentes externas.** Se trata de un ruido no deseado que no está generado ni por la muestra ni por el instrumento de medida. Habitualmente, está causado por una fuente de luz externa que contamina la

señal en algún punto del proceso de medida. Por ejemplo, puede estar causada por la luz solar o la luz de los fluorescentes. En general, si el conjunto de los instrumentos Raman de medición está correctamente diseñado, el equipo suele ser inmune a las radiaciones externas. Otra fuente que genera este tipo de ruido son los rayos cósmicos que llegan a la órbita terrestre, puesto que suelen liberar un número elevado de electrones que son indistinguibles de los fotoelectrones. Estos rayos cósmicos crean unos picos muy estrechos y de gran intensidad en posiciones aleatorias del Espectro Raman, aunque estos picos no suelen confundirse con Bandas Raman, ya que presentan características diferentes.

De entre los ruidos detallados, los que se pueden encontrar más comúnmente en un Espectro Raman son el ruido shot, y la fluorescencia. En la siguiente figura (*Fig. 1.5*) se presenta un espectro en el que se pueden apreciar los ruidos descritos.

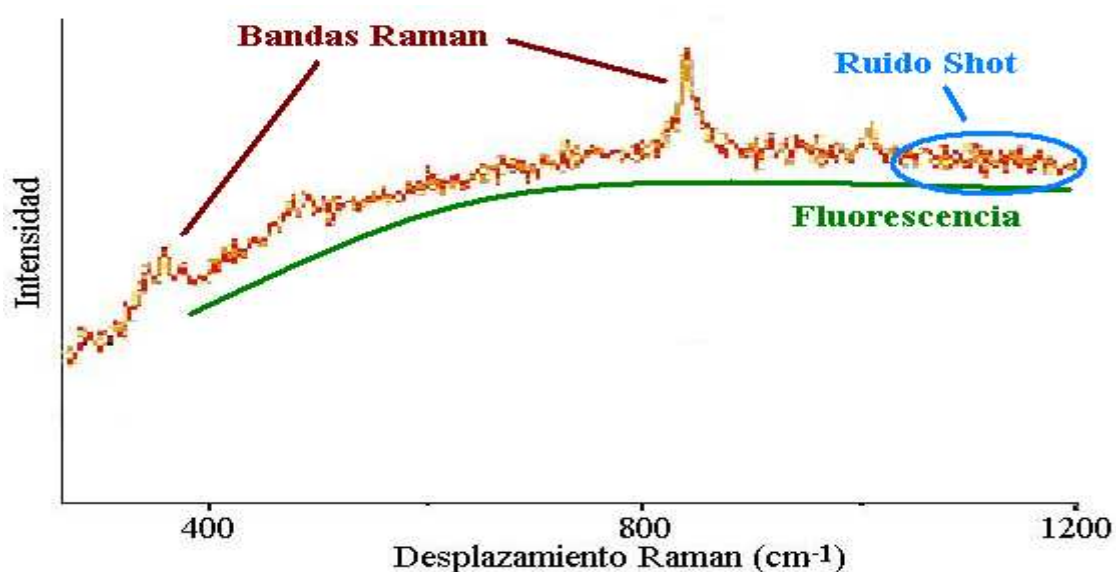


Fig. 1.5: Principales ruidos en Espectroscopía Raman

1. 3. Aplicación de la Espectroscopía Raman a la identificación de pigmentos artísticos

La Espectroscopía Raman es una de las técnicas analíticas que ha generado un mayor interés en los últimos años, siendo objeto de gran cantidad de publicaciones de carácter

científico y convirtiéndose en una técnica cuyo ámbito de aplicación se ha visto incrementado exponencialmente^[10]. Así, la Espectroscopía Raman encuentra su uso en aplicaciones tan diversas como la biomedicina, las ciencias medioambientales, la geología o la exploración espacial, entre muchos otros.

La Espectroscopía Raman se ha convertido en una de las herramientas analíticas de referencia en el campo de la investigación artística. El incremento de la importancia de la Espectroscopía Raman en la investigación de antigüedades y objetos artísticos se hace patente en el número de artículos publicados por año, así como en el incremento de grupos de investigación: En 1995 tan solo cinco grupos publicaron artículos de investigación en este campo, mientras que en 2005 ya eran veinte grupos.

En el análisis de obras de arte, uno de los objetivos principales de la Espectroscopía Raman es la identificación de los materiales pictóricos o *pigmentos*. Un pigmento es un componente químico y sólido en forma de pequeñas partículas separadas que presentan unas ciertas características dependiendo del grupo químico al que pertenecen. Se trata del material más importante en el proceso pictórico. Según su origen químico, se clasifican en dos grandes categorías: Los orgánicos y los inorgánicos.

La identificación de pigmentos mediante la Espectroscopía Raman es una de las investigaciones más activas en el análisis de arte, debido al carácter no-destructivo de la técnica, que permite obtener información molecular sobre una determinada muestra a analizar sin que ésta necesite ningún tipo de preparación, y a la posibilidad de realizar el análisis *in situ*, gracias al uso de un haz de luz como herramienta de trabajo y, por ello, a la tecnología basada en la fibra óptica.

La Espectroscopía Raman permite conocer la composición molecular de los pigmentos que aparecen en una obra, caracterizándolos a partir de la posición frecuencial de las bandas espectrales. El conjunto de estas bandas es único para cada material, y por ello, permite la identificación unívoca del mismo.

Habitualmente, la identificación de un material en Espectroscopía Raman se lleva a cabo por el propio analista a partir de la comparación de su Espectro Raman con una serie de Espectros Raman de referencia, previamente almacenados (ver *Fig. 1.6*). Esta

metodología, aún siendo sencilla e intuitiva, puede convertirse en un proceso largo, y puede llevar implícito un razonamiento subjetivo, según la experiencia del analista. Además, puede resultar un trabajo laborioso, donde se hace necesario un observador experimentado que, según su experiencia y conocimiento, realice las tareas de comparación e identificación. En este proceso se introduce inevitablemente un componente de subjetividad ligado a la intervención del investigador.

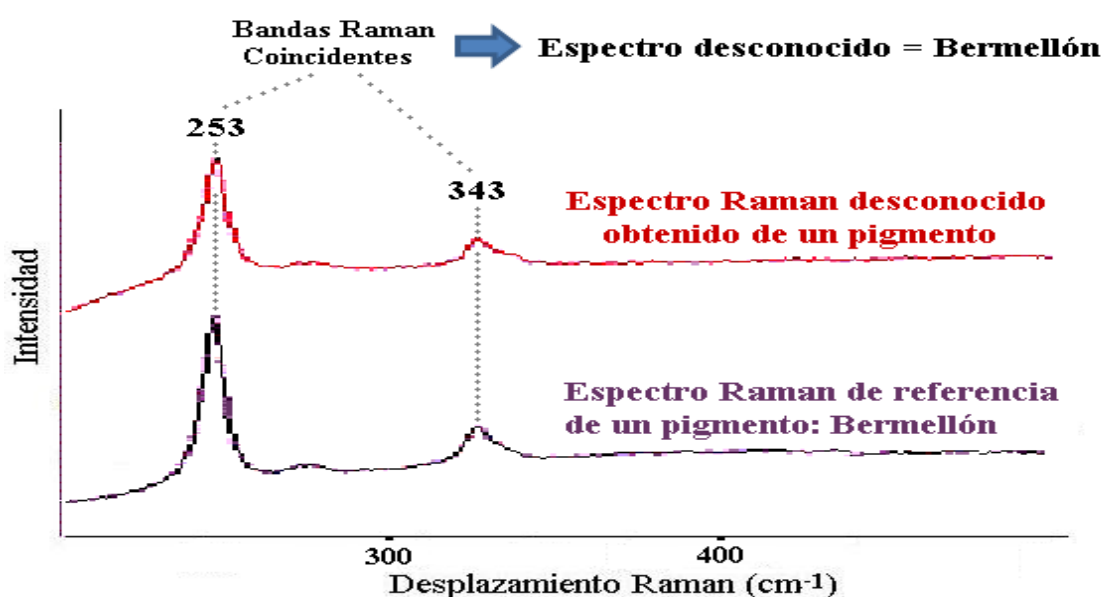


Fig. 1.6: Ejemplo del Espectro Raman de un pigmento tomado en un cuadro y Espectro Raman de referencia (o patrón) del pigmento Bermellón

El analista, cuando compara los espectros Raman, se puede basar o bien en comprobar si las bandas Raman de los dos espectros coinciden en posición, o bien valorar el parecido global de los espectros en cuestión, teniendo en cuenta entonces toda la información contenida en los espectros. La primera opción implica localizar, previamente a la comparación, todas las bandas Raman espectrales. Este requisito puede llegar a ser complicado cuantas más bandas tengan los espectros. En este caso, resulta más apropiada la segunda opción, comparar los espectros de forma global, valorando cada una de sus características para la identificación.

Puesto que el Espectro Raman es único para cada material, el pigmento analizado queda identificado de manera inequívoca, ofreciendo, así, información referente a la caracterización de obras de arte, y, ya que la introducción de determinados pigmentos está estrictamente datada, la identificación de materiales pictóricos puede aportar una indicación de la fecha de la obra e incluso de su autenticidad (ver *Tabla 1.1*^[11]).

	Prehistoria	Antigüedad	V-XV	XVI	XVII	XVIII	XIX	XX
BLANCOS								
Carbonato de calcio	—————	—————	—————	—————	—————	—————	—————	—————
Kaolin	—————	—————	—————	—————	—————	—————	—————	—————
Yeso	—————	—————	—————	—————	—————	—————	—————	—————
Blanco de plomo	—————	—————	—————	—————	—————	—————	—————
Blanco de zinc	—————	—————	—————	—————	—————	—————	—————	—————
Lithopone	—————	—————	—————	—————	—————	—————	—————	—————
Blanco de titanio	—————	—————	—————	—————	—————	—————	—————	—————
Sulfato de bario	—————	—————	—————	—————	—————	—————	—————
NEGROS								
Carbón vegetal	—————	—————	—————	—————	—————	—————	—————	—————
Negro de marfil	—————	—————	—————	—————	—————	—————	—————	—————
Negro de humo	—————	—————	—————	—————	—————	—————	—————	—————
Grafito	—————	—————	—————	—————	—————	—————	—————	—————
AZULES								
Azul egipcio	—————	—————	—————	—————	—————	—————	—————	—————
Azurita	—————	—————	—————	—————	—————	—————	—————
Lapislázuli	—————	—————	—————	—————	—————	—————	—————
Ultramar sintético	—————	—————	—————	—————	—————	—————	—————	—————
Esmalte de cobalto	—————	—————	—————	—————	—————	—————	—————
Azul de Prusia	—————	—————	—————	—————	—————	—————	—————
Azul de cobalto	—————	—————	—————	—————	—————	—————	—————
Cerúleo	—————	—————	—————	—————	—————	—————	—————
Índigo	—————	—————	—————	—————	—————	—————
Azul de ftalocianina	—————	—————	—————	—————	—————	—————	—————	—————
AMARILLOS								
Ocre amarillo	—————	—————	—————	—————	—————	—————	—————	—————
Amarillo Massicot	—————	—————	—————	—————	—————	—————	—————
Oropimente	—————	—————	—————	—————	—————	—————	—————
Amarillo plomo-estaño I	—————	—————	—————	—————	—————	—————	—————
Amarillo de Nápoles	—————	—————	—————	—————	—————	—————
Amarillo indio	—————	—————	—————	—————	—————	—————
Amarillo de cromo	—————	—————	—————	—————	—————	—————	—————
Amarillo de zinc	—————	—————	—————	—————	—————	—————	—————
Amarillo de bario	—————	—————	—————	—————	—————	—————	—————
Amarillo de cadmio	—————	—————	—————	—————	—————	—————	—————
Laca gambogue	—————	—————	—————	—————	—————	—————
Óxidos de hierro sintéticos	—————	—————	—————	—————	—————	—————	—————	—————

	Prehistoria	Antigüedad	V-XV	XVI	XVII	XVIII	XIX	XX
ROJOS								
Ocres rojos		—————	—————	—————	—————	—————	—————	—————
Cinabrio		—————	—————					
Vermellón sintético			—————	—————	—————	—————	—————	—————
Minio		—————	—————	—————	—————	—————	—————	—————
Laca Kermes		—————	—————					
Carmín de cochinilla			—————	—————	—————	—————	—————	—————
Laca de granza			—————	—————	—————	—————	—————	—————
Laca d'alizarine							—————	—————
Goma laca	
Caput mortuum		
Rojo de cadmio								—————
Naranja de cromo							—————	—————
Óxidos de hierro sintéticos							—————	—————
VERDES								
Tierras verdes		—————	—————	—————	—————	—————	—————
Crisocolia		—————						
Malaquita		—————	—————	—————	—————
Verdigris		—————	—————	—————	—————
Resinato de cobre			—————	—————				
Verde de Scheele					
Verde de Schweinfurt					
Verde de cobalto						
Verde de cromo						
Hidróxido de cromo verde						
Óxido de cromo verde						
Verde de ftalocianina								—————
VIOLETAS								
Violeta de cobalto						
Violeta de magnesio						
MARRONES								
Ocres	—————	—————	—————	—————	—————	—————	—————	—————
Marrón van Dyck				

Tabla 1.1: Periodos de utilización de los principales pigmentos utilizados históricamente en el mundo del arte. En línea continua, los períodos de utilización que se conocen con exactitud. En discontinua, los periodos de aparición y desaparición de algunos pigmentos

2. ANÁLISIS POR COMPONENTES PRINCIPALES (ACP). APLICACIÓN AL TRATAMIENTO DE ESPECTROS RAMAN

2. 1. Fundamentos del *Análisis por Componentes Principales* (ACP).

ACP normado y no normado

Paralelamente al aumento en los últimos años del uso de diferentes espectroscopias como herramientas experimentales y analíticas, aparece la necesidad de poder comprimir los datos sin por ello perder información. El objetivo a alcanzar es mantener el contenido informativo ocupando menos memoria y evitando la redundancia^[12].

El *Análisis por Componentes Principales* (o ACP) es un método multivariante de la familia de la quimiometría, siendo ésta una disciplina química que utiliza métodos matemáticos y estadísticos tanto para diseñar o seleccionar procedimientos de medida y experimentos óptimos como para obtener la máxima información química de los análisis^[13].

El ACP está extendido sobretodo como herramienta para la reducción dimensional de datos. Dado un conjunto de datos, esta técnica facilita un nuevo espacio de dimensión inferior al original, conocido como espacio de los *Componentes Principales* o PCs (del inglés *Principal Components*), donde los datos se representan de tal manera que se resalta la información. Se trata, así, de una herramienta estadística que permite detectar las similitudes y las diferencias entre los datos, obteniendo una nueva expresión reducida respecto a la expresión inicial.

Esta característica convierte al ACP en una herramienta muy potente en los procesos de clasificación o reconocimiento de patrones, ya que los agiliza y optimiza el resultado. Así, esta técnica se ajusta y adapta a estudios donde se analizan datos espectrales, ya que cualquier tipo de espectro está generado por un gran número de variables

(frecuencia, periodo, longitud de onda normalizada, etc.), de manera que pueden resultar ser señales difíciles de interpretar. Por ello, aplicar esta técnica de reducción dimensional permite obtener resultados muy satisfactorios en materia de identificación espectral.

La principal motivación a la hora de aplicar una herramienta de reducción dimensional es conseguir expresar un conjunto de datos definido por n variables como un conjunto de menor dimensión, k , con $k \ll n$, pero equivalente en cuanto a contenido informativo. Así, el ACP se encuadra en la categoría de métodos de transformación de variables, puesto que se basa en sustituir las n variables iniciales (U_i) que explican los datos, por otras n variables (u_i), independientes entre sí y ordenadas de tal manera que las primeras tienen más relevancia que las últimas.

El ACP es un proceso que se puede interpretar geoméricamente como un cambio de ejes en la representación de los datos. Su principal objetivo es obtener un espacio k -dimensional, siendo éste un Espacio Vectorial Euclídeo, donde el conjunto inicial de datos se exprese conservando la estructura y la información inicial. Normalmente existe un sistema de coordenadas óptimo para representar cada conjunto de datos, considerando como óptimo aquel donde se puedan diferenciar claramente todas las observaciones. Este sistema de coordenadas es precisamente el que se obtiene mediante la transformación implementada por el ACP. El nuevo sistema de coordenadas es de dimensión igual al inicial, n , pero su configuración permite llevar a cabo *a posteriori* la reducción de n a k dimensiones.

Puesto que la transformación, o cambio de base, asegura obtener un espacio ortogonal, la varianza total de los datos se mantiene en ambos espacios (tanto en el original como en el transformado). Esta transformación permite obtener unas nuevas variables generadas de tal manera que las primeras ($i=1, \dots, k$) tienen más relevancia, es decir, mayor contenido informativo, que las últimas ($i=k+1, \dots, n$). Esta propiedad permite seleccionar un valor, k , de estas nuevas variables, siendo $k \ll n$, en función de la cantidad de información que se quiera ignorar, de forma que finalmente se consigue convertir un conjunto de datos definido por n variables, U_i , a uno descrito por un valor inferior, k , de variables, u_i . El ACP se comporta, pues, como una técnica de selección de variables, puesto que se selecciona un subconjunto de las nuevas variables para la

representación de los datos iniciales. Esta reducción debe asegurar la mínima pérdida de información y aportar la menor distorsión posible.

Los Componentes Principales, u_i , resultan de la combinación lineal de las variables originales, U_j , tal y como se muestra a continuación:

$$u_i = c_{1i}U_1 + \dots + c_{ji}U_j + \dots + c_{ni}U_n$$

donde las coordenadas c_{ji} , conocidas como *loads*, representan la aportación o peso de la variable U_j a la variable u_i . Es importante destacar que las variables u_i son las variables, incorreladas entre sí, que eliminan la redundancia que pueda existir entre las variables iniciales U_j . A partir de los *loads* se obtiene una matriz de transformación:

$$C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix}$$

donde c_{ji} es el peso de la variable U_j respecto a la variable u_i , con $i=1,\dots,n$ y $j=1,\dots,n$. Esta matriz cuadrada explica la relación lineal entre las nuevas variables, PC_i , y las antiguas, los números de onda normalizados. En concreto, cada una de sus columnas define las coordenadas de un Componente Principal en el espacio de las variables u_i .

La ventaja de este nuevo sistema de coordenadas es que se conoce la porción de la varianza total (v_{total}) que aporta cada uno de sus ejes (PC_i), y que estos están ordenados en función de esta aportación, de mayor a menor. Como consecuencia de esta ordenación, se puede saber qué PCs son los más significativos para expresar los datos. También resulta posible valorar qué porcentaje de v_{total} se tiene en cuenta si sólo se consideran k Componentes Principales para definir los datos. De esta forma, se puede calcular la varianza acumulada cada vez que se tiene en cuenta un PC de más.

Es en el momento de escoger el valor k de Componentes Principales a contemplar cuando se produce la reducción dimensional, es decir, cuando se decide cuantos PCs son suficientes para representar los datos de manera fiable. Existen diversos criterios que dictaminan como seleccionar dicho número. Cuando k Componentes Principales concentran gran parte o la totalidad de la varianza original (v_{total}), los datos tratados se

pueden expresar en el espacio definido por estas k nuevas variables sin que la pérdida de información sea crítica. De esta manera se consigue diferenciar los datos y se consigue una reducción dimensional muy considerable. Una vez seleccionado el número de PCs a contemplar se obtiene la matriz reducida de cambio de base:

$$C_{red} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nk} \end{pmatrix}$$

Así, si se considera un conjunto inicial de referencia de P elementos, contemplando, por ejemplo, el criterio de concentrar el 100% de la varianza de las variables originales, se obtienen $k=P-1$ Componentes Principales.

Finalmente, una vez obtenida la matriz reducida de transformación, el conjunto inicial puede ser expresado en el nuevo espacio generado por los PCs aplicando el cambio de base definido por C sobre la matriz que recoge las variables originales, siendo ésta X :

$$S = X \cdot C$$

donde cada fila de S (S_i) corresponde a una de las observaciones iniciales expresada en el espacio de los Componentes Principales. Puesto que el objetivo a alcanzar es el de la reducción de dimensionalidad, la proyección de las variables originales en el nuevo espacio definido por los PCs se realiza mediante:

$$S_{red} = X \cdot C_{red}$$

de donde resulta, suponiendo un conjunto inicial formado por P observaciones:

$$S_{red} = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1k} \\ s_{21} & s_{22} & \cdots & s_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pk} \end{pmatrix}$$

Esta matriz, por filas, permite conocer cuales son las coordenadas (conocidas como *scores*) de cada una de las observaciones en el espacio generado por las nuevas coordenadas. Es decir, la matriz S es equivalente a la matriz X pero definida en el espacio de PCs. Como se puede ver, para convertir un conjunto determinado en otro conjunto mediante el ACP, la notación matricial se hace imprescindible.

Como paso previo a la proyección del conjunto inicial en el espacio definido por los PCs, las variables originales deben ser pretratadas, existiendo dos posibilidades: *centrar las variables iniciales* o bien *estandarizarlas* (o *tipificarlas*). La elección entre realizar un centrado o una estandarización de variables genera dos tipos de ACPs: el *ACP no normado* y el *ACP normado*, respectivamente.

Sea X la matriz inicial de las observaciones, colocadas por filas:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{pmatrix}$$

El *centrado* se basa en la sustracción del valor medio de las variables originales (o columnas de la matriz X):

$$x_{cij} = x_{ij} - \bar{x}_j$$

donde, para un espectro de n variables:

$$\bar{x}_j = \frac{1}{P} \sum_{i=1}^P x_{ij}$$

obteniendo, así, la matriz X_c :

$$X_c = \begin{pmatrix} x_{c11} & x_{c12} & \cdots & x_{c1n} \\ x_{c21} & x_{c22} & \cdots & x_{c2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{cp1} & x_{cp2} & \cdots & x_{cpn} \end{pmatrix}$$

En cambio, la *estandarización*, además de centrar las variables, les impone varianza unitaria:

$$x_{estij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j}$$

siendo:

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^P (x_{ij} - \bar{x}_j)^2}{P}}$$

generando la matriz X_{est} :

$$X_{est} = \begin{pmatrix} x_{est11} & x_{est12} & \cdots & x_{est1n} \\ x_{est21} & x_{est22} & \cdots & x_{est2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{estp1} & x_{estp2} & \cdots & x_{estpn} \end{pmatrix}$$

Así, la principal diferencia entre ambos ACPs se hace patente en el espacio de los PCs generado. Mientras el *ACP no normado* pondera las variables originales a la hora de crear el espacio de PCs, otorgando un mayor peso a ciertas variables, en los rangos donde existe mayor dispersión, el *ACP normado* uniformiza el valor de todas las variables a la hora de generar este nuevo espacio.

2. 2. Calidad de la representación: contribución y coseno cuadrado

Con el fin de poder cuantificar la calidad en la representación una vez realizado el cambio de coordenadas al nuevo espacio de los PCs generado por un conjunto inicial de referencia determinado, existen dos parámetros que permiten medir dicha calidad: la *contribución* y el *coseno cuadrado*^[14].

Se define la *contribución*, $CTR_{\alpha}(i)$, como la influencia del *alfa*-ésimo PC del elemento *i*-ésimo sobre el *alfa*-ésimo PC de todos los elementos:

$$CTR_{\alpha}(i) = \frac{I_{O,U_{\alpha}}(i)}{I_{O,U_{\alpha}}(I)} = \frac{(\langle X_{c/est}^i, C_{red \alpha} \rangle)^2}{\lambda_{\alpha}}$$

donde $I_{O,U_{\alpha}}(i)$ es la coordenada *alfa*-ésima del elemento *i*-ésimo al cuadrado y $I_{O,U_{\alpha}}(I)$ es la suma de las *alfa*-ésima coordenadas al cuadrado de todos los elementos del conjunto.

Como su propio nombre indica, la contribución hace referencia a cómo contribuye una determinada coordenada (la *alfa*-ésima) de un determinado elemento del conjunto (el elemento *i*-ésimo) a la generación del nuevo espacio de PCs.

Contemplando un conjunto inicial de P elementos, atendiendo al criterio de concentración del 100% de la varianza de las variables originales, se obtienen $P-1$ Componentes Principales, con lo cual se pueden calcular $P-1$ contribuciones por

elemento. Es interesante, pues, definir la contribución *total* de un elemento determinado, puesto que implica conocer cómo contribuye dicho elemento a la generación del espacio de PCs. De esta forma, se define la contribución total de un elemento como la suma de todas las contribuciones de dicho elemento, es decir:

$$CTR_{total}(i) = \sum_{\alpha=1}^{P-1} CTR_{\alpha}(i)$$

Así, para un conjunto inicial de referencia de P elementos, la contribución total de cada uno de dichos elementos es constante y de valor $CTR_{total}(i) = (P - 1)/P$, con $i=1, \dots, P$.

La contribución, aun ofreciendo información sobre la influencia de las coordenadas de los elementos iniciales en la generación del espacio de PCs, no se contempla por las siguientes razones:

- No aplica su cálculo a elementos que no pertenezcan al conjunto inicial, puesto que un elemento externo al conjunto inicial no contribuye a la generación del espacio de PCs.
- La información ofrecida por la contribución *alfa-ésima* de un elemento determinado es redundante, ya que dicho valor es proporcional al cuadrado de la *alfa-ésima* coordenada en PCs (o *score*) del elemento contemplado.
- El valor de la contribución total es constante para todo elemento del conjunto inicial generador del espacio de PCs.

Se define el *coseno cuadrado*, $COS2_{\alpha}(i)$, como la influencia del *alfa-ésimo* PC del elemento i -ésimo sobre dicho elemento:

$$COS2_{\alpha}(i) = \frac{I_{O,U_{\alpha}}(i)}{I_O(I)} = \frac{(\langle X_{c/est}^i, C_{red \alpha} \rangle)^2}{I_O(I)} = \frac{[S_{\alpha}(i)]^2}{\|X^i\|^2} = \frac{(\|X^i\| \cos \varphi)^2}{\|X^i\|^2} = \cos^2 \varphi$$

donde φ es el ángulo que forman entre sí los vectores $X_{c/est}^i$ y $C_{red \alpha}$, y $I_O(I)$ es la suma de las coordenadas al cuadrado del elemento i -ésimo. Puesto que C es una matriz unitaria, $\|C_{red \alpha}\| = 1$.

Este parámetro ofrece información referente a la calidad de la representación de un cierto *score* de cualquier elemento respecto al espacio de PCs generado por un

determinado conjunto inicial. En definitiva, se trata de la relación entre las normas de un elemento calculadas en el espacio original y en el espacio transformado, que, de forma general, se toma en tanto por cien. Así, de igual forma que para la contribución, es interesante definir el coseno cuadrado *total*, ya que permite conocer la calidad en la representación de un determinado elemento al ser proyectado en el espacio de PCs:

$$COS2_{total}(i) = \sum_{\alpha=1}^{P-1} COS2_{\alpha}(i)$$

Al tratarse de un coseno, se cumple:

$$0 \leq COS2_{total}(i) \leq 1$$

donde $COS2_{total}(i) = 0$ indica que la calidad de la representación del elemento *i*-ésimo es desfavorable y $COS2_{total}(i) = 1$ indica una representación perfecta o ideal.

Por ello, el coseno cuadrado total es el parámetro que realmente ofrece información interesante. Puesto que la transformación de un elemento perteneciente al conjunto inicial mantiene su norma en el espacio transformado, el valor de su coseno cuadrado total es del 100%.

Por otra parte, el cálculo del coseno cuadrado total sobre elementos que no pertenecen al conjunto inicial generador del espacio de PCs, depende exclusivamente de cómo se ve afectado a la hora de ser proyectado sobre dicho espacio, mostrando, así, la calidad de la representación de dicho elemento en el espacio transformado. Excepto para el caso ideal, en el que un nuevo elemento externo a los elementos generadores del espacio de PCs coincida con alguno de dichos elementos del conjunto inicial, el valor del coseno cuadrado total es inferior al 100%.

2. 3. Aplicación al tratamiento de espectros Raman

En el caso del procesado de Espectros Raman de pigmentos artísticos, el conjunto inicial a considerar es el constituido por los espectros de los pigmentos de referencia, espectros obtenidos directamente de los pigmentos puros, generando las librerías espectrales de referencia, como por ejemplo la formada por los pigmentos inorgánicos del siglo XVII. Tomando una de estas librerías, mediante el ACP, se genera un nuevo

espacio sobre el cual, una vez determinada la reducción dimensional, los Espectros Raman de pigmentos analizados de una obra artística pueden ser proyectados para su posterior comparación e identificación^[15].

Para aplicar el ACP es necesario adaptar los datos que se quieren reducir a la nomenclatura matemática requerida. En concreto, se deben expresar los datos en forma de matriz, donde las filas sean las expresiones de los datos en sí (observaciones, espectros de los pigmentos contemplados) y las columnas las variables (números de onda normalizados). Los Espectros Raman, desde el punto de vista de la formulación, son vectores de n coordenadas definidos en lo que se conoce como “espacio de señal de los números de onda normalizados”:

$$E_i = [e_{i1}, \dots, e_{ij}, \dots, e_{in}]$$

donde e_{ij} son las intensidades Raman del espectro E_i para un número de onda normalizado (desplazamiento Raman) U_j .

La secuencia de pasos a seguir para realizar el ACP sobre una librería espectral de referencia es la siguiente:

1. Escoger los P espectros patrones o de referencia (de n variables) a utilizar como librería espectral, creando una matriz E (de coordenadas e_{ij}):

$$E = \begin{pmatrix} e_{11} & e_{12} & \dots & e_{1n} \\ e_{21} & e_{22} & \dots & e_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ e_{P1} & e_{P2} & \dots & e_{Pn} \end{pmatrix}$$

2. Con el objetivo de minimizar problemas asociados con la comparación de espectros de diferentes intensidades, que pueden sufrir modificaciones por distintos factores, como por ejemplo, diferencias en el tiempo de medida, potencia láser, concentración, etc., los espectros son normalizados mediante la expresión:

$$e_{ij\text{norm}} = \frac{e_{ij} - e_{ij\text{min}}}{e_{ij\text{max}} - e_{ij\text{min}}} \equiv x_{ij}$$

donde $e_{ij_{norm}} \equiv x_{ij}$ es la intensidad normalizada para cada valor, e_{ij} es la intensidad original, y por último, $e_{ij_{min}}$ y $e_{ij_{max}}$ son las intensidades mínimas y máximas de cada espectro, respectivamente.

Así, se hace imprescindible *homogeneizar* dicha librería espectral de referencia, obteniendo la matriz X (de coordenadas x_{ij}):

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{pmatrix}$$

3. En función de qué ACP se desea llevar a cabo (si *ACP no normado* o *ACP normado*), en este punto se procede a pretratar las variables iniciales, existiendo dos posibilidades: *Centrar las variables originales* para el *ACP no normado* o bien *estandarizarlas* (o *tipificarlas*) para el *ACP normado*.

Se obtiene la matriz X_c para el *ACP no normado*:

$$X_c = \begin{pmatrix} x_{c11} & x_{c12} & \cdots & x_{c1n} \\ x_{c21} & x_{c22} & \cdots & x_{c2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{cp1} & x_{cp2} & \cdots & x_{cpn} \end{pmatrix}$$

En cambio, se obtiene la matriz X_{est} para el *ACP normado*:

$$X_{est} = \begin{pmatrix} x_{est11} & x_{est12} & \cdots & x_{est1n} \\ x_{est21} & x_{est22} & \cdots & x_{est2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{estp1} & x_{estp2} & \cdots & x_{estpn} \end{pmatrix}$$

La decisión entre utilizar un ACP u otro es crucial a la hora de tratar determinadas librerías espectrales, puesto que dependiendo de la forma de los espectros patrones, como se verá más adelante, un ACP se adapta mejor o peor.

4. El siguiente paso es el cálculo de la matriz de covarianzas de X_c o de X_{est} , según la elección determinada entre centrado o estandarización de variables originales, obteniendo la matriz X_{cov} :

$$X_{cov} = cov(X_{c/est_i}, X_{c/est_j}) = E[(X_{c/est_i} - E(X_{c/est_i}))(X_{c/est_j} - E(X_{c/est_j}))] =$$

$$= \begin{pmatrix} E[(X_{c/est_1} - E(X_{c/est_1}))(X_{c/est_1} - E(X_{c/est_1}))] & \cdots & E[(X_{c/est_1} - E(X_{c/est_1}))(X_{c/est_n} - E(X_{c/est_n}))] \\ E[(X_{c/est_2} - E(X_{c/est_2}))(X_{c/est_1} - E(X_{c/est_1}))] & \cdots & E[(X_{c/est_2} - E(X_{c/est_2}))(X_{c/est_n} - E(X_{c/est_n}))] \\ \vdots & \ddots & \vdots \\ E[(X_{c/est_n} - E(X_{c/est_n}))(X_{c/est_1} - E(X_{c/est_1}))] & \cdots & E[(X_{c/est_n} - E(X_{c/est_n}))(X_{c/est_n} - E(X_{c/est_n}))] \end{pmatrix}$$

5. Se debe proceder a la diagonalización de dicha matriz de covarianzas, obteniendo sus valores propios (*vap*'s) y sus vectores propios (*vep*'s). Puesto que la matriz X_{cov} tiene dimensión $n \times n$, se obtienen n *vep*'s y n *vap*'s. La información proporcionada por cada *vap* es extremadamente interesante ya que dicho valor coincide con la varianza de las variables originales que contempla su respectivo *vep* (convertido en PC).
6. A continuación, se ordenan de forma descendente dichos *vep*'s, según el valor de sus respectivos *vap*'s, obteniendo los Componentes Principales por orden de importancia, generando así la matriz de cambio de base, C .
7. En este punto, se selecciona el número de PCs ($k < n$) suficiente para caracterizar los datos de manera fiable. Para ello existen diversos criterios basados en la obtención de un determinado valor de varianza mínimo. Puesto que se desea la mínima pérdida de información, se selecciona como criterio que el valor mínimo de varianza a contemplar sea el del 100%, generando así la matriz de cambio de base reducida C_{red} .
8. Una vez obtenida dicha matriz de cambio de base, se expresa la librería espectral en el nuevo espacio obtenido, el espacio de los PCs, de dimensión inferior al original, mediante el cambio expresado por $S_{red} = X_c \cdot C_{red}$, o bien, en su caso, por $S_{red} = X_{est} \cdot C_{red}$:

$$S_{red} = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1k} \\ S_{21} & S_{22} & \cdots & S_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pk} \end{pmatrix} = \begin{pmatrix} x_{c/est_{11}} & x_{c/est_{12}} & \cdots & x_{c/est_{1n}} \\ x_{c/est_{21}} & x_{c/est_{22}} & \cdots & x_{c/est_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{c/est_{p1}} & x_{c/est_{p2}} & \cdots & x_{c/est_{pn}} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nk} \end{pmatrix}$$

Se pueden considerar, pues, tres tipos de coordenadas para los espectros que forman parte de un ACP: Sus coordenadas en las variables originales normalizadas (x_{ij}), sus coordenadas en las variables centradas o estandarizadas (x_{cij} o x_{estij}), y, por último, sus coordenadas en el nuevo sistema de referencia (S(i)). Las matrices que recogen estas coordenadas son, respectivamente X , X_c o X_{est} y S .

La siguiente tabla muestra las características más destacables de estos tres sistemas de coordenadas para el caso espectral:

Matriz	X	X_c	X_{est}	S
Media de las columnas	\bar{x}_j	0	0	0
Varianza de las columnas	σ_j	σ_j	1	λ_j
Correlación entre columnas	$Corr(j, j')$	$Corr(j, j')$	$Corr(j, j')$	0

Así, dado un conjunto de datos, el ACP genera un sistema de coordenadas que depende de este conjunto analizado, donde se define la dirección del primer eje en la dirección de la máxima varianza de los datos, la dirección del segundo eje, perpendicular al primero, como la que maximiza la varianza de los datos, y así sucesivamente hasta obtener n ejes ortogonales. Estos ejes, conocidos como Componentes Principales o PCs, definen un nuevo espacio de dimensión n obtenido mediante una transformación del espacio inicial.

La reducción dimensional que permite el ACP se basa en escoger el valor k ($k \ll n$), es decir, determinar cuantos PCs son suficientes para describir los datos de manera fiable. Como se ha comentado en el punto 7, se escoge el criterio que sugiere contemplar el número de PCs tal que se aporte el 100% de la varianza total. Dado que las librerías espectrales suelen estar formadas por un número determinado de espectros de referencia, P , muy inferior al número de variables originales, n , se obtiene que los $P-1$ primeros PCs aportan el 100% de la varianza total. Puesto que $P-1$ es mucho menor que n , se toma $k=P-1$ como el espacio reducido a considerar.

Ahora, dado un nuevo espectro, w , que no pertenezca a los espectros de la librería de referencia, para proyectar dicho espectro en este nuevo espacio de los PCs generado por la librería de referencia, se deben seguir los siguientes pasos:

1. Normalizar el espectro, w' .
2. Centrar o estandarizar w' en función de los valores medios y las varianzas de las variables normalizadas utilizadas en la obtención del espacio de PCs como $v_{ij} = w'_{ij} - \bar{x}_j$ o bien $v_{ij} = \frac{w'_{ij} - \bar{x}_j}{\sigma_j}$, generando el vector v .
3. Realizar el cambio de base: $v_{PCs}^T = C_{red} \cdot v^T$.

Para realizar el proceso inverso y obtener nuevamente el espectro expresado en las variables originales, se debe multiplicar v_{PCs} por la matriz de cambio de base inversa, C^{-1} . No obstante, al trabajar con la matriz reducida, C_{red} , se obtiene una matriz no cuadrada, con lo cual, no se puede obtener su inversa. Sin embargo, puesto que la matriz cambio de base, C , es unitaria, y lo es también la matriz reducida de cambio de base, C_{red} , su inversa coincide con su traspuesta:

$$\text{Una matriz, } A, \text{ es unitaria } \Leftrightarrow A \cdot A^T = A^T \cdot A = Id$$

Por tanto,

$$v^T = C_{red}^T \cdot v_{PCs}^T$$

de donde, finalmente, se obtiene nuevamente w' :

1. Teniendo en cuenta el centrado de variables:

$$w'^T = C_{red}^T \cdot v_{PCS}^T + x_{med}^T = v^T + x_{med}^T$$

2. Teniendo en cuenta la estandarización de variables:

$$w'^T = \sigma_x \cdot v^T + x_{med}^T$$

donde, de forma esquemática, se ha llamado x_{med} y σ_x a los vectores de medias y desviaciones de las variables originales normalizadas.

Aplicando las siguientes propiedades de matrices transpuestas:

$$(A + B)^T = A^T + B^T$$

$$(A \cdot B)^T = B^T \cdot A^T$$

se obtiene, para el ACP no normado:

$$w' = v_{PCS} \cdot C_{red} + x_{med}$$

o bien, para el ACP normado:

$$w' = \sigma_x \cdot v_{PCS} \cdot C_{red} + x_{med}$$

3. RECONOCIMIENTO/IDENTIFICACIÓN DE ESPECTROS RAMAN

3. 1. Planteamiento general: Metodología para la identificación

Un determinado material es caracterizado por su Espectro Raman de forma única ya que éste depende de la configuración molecular de dicho material. Así, en el momento de identificar un material desconocido, su Espectro Raman permite reconocerlo de manera inequívoca. Esta identificación se lleva a cabo por comparación entre el espectro desconocido y una serie de espectros conocidos, llamados espectros patrones. De esta forma, se hace necesario disponer de una librería espectral en la que previamente se hayan almacenado los espectros de referencia.

Un Espectro Raman puede ser interpretado como un vector de entre 1000 y 2000 componentes, típicamente, de manera que la librería espectral puede contener dimensiones elevadas. Además, puede estar formado por un gran número de bandas, con lo que el reconocimiento del espectro desconocido por comparación visual resulta ser una tarea complicada y costosa para el analista. Asimismo, la técnica de identificación basada en la inspección visual del espectro desconocido y la librería de referencia implica un proceso complejo y sujeto a la interpretación subjetiva de un experto. Por ello, uno de los objetivos de este proyecto se basa en implementar un sistema automático que agilice el procedimiento de identificación, de forma que éste sea más rápido y objetivo, y, puesto que el reconocimiento es realizado de forma automática, minimice la intervención del analista.

Una vez seleccionados los espectros constituyentes de la librería espectral de referencia, se les aplica la técnica de reducción dimensional establecida por el ACP, siguiendo el procedimiento expuesto en el *Capítulo 2*. No obstante, para llevar a cabo dicho análisis, es necesario que la librería inicial cumpla una serie de condiciones de homogeneidad en el formato de los datos. Por ello, los Espectros Raman deben ser tratados previamente con el fin de encontrarse definidos bajo las mismas condiciones, tanto para los número

de onda, definidos dentro de un mismo margen frecuencial, como para sus intensidades arbitrarias, definidas en la misma escala. Además, los espectros deben medirse en las mejores condiciones posibles, siendo filtrados con el fin de eliminar la línea de base que puede, incluso, llegar a enmascarar bandas Raman.

Para ello, se debe asegurar que los espectros compartan un mismo margen de número de onda, realizando una interpolación espectral con el fin de asegurar que las n coordenadas de los espectros originales se correspondan con las intensidades Raman en los mismos número de onda. Asimismo, se debe independizar al máximo la intensidad de las bandas Raman de las condiciones de medida mediante la normalización espectral, de manera que los espectros compartan el mismo rango de intensidades. Una vez realizado el pretratado espectral, los espectros se encuentran en condiciones de ser expresados en el espacio de dimensión reducida que genera el ACP y que, además, facilita el reconocimiento de un espectro desconocido.

En el siguiente diagrama se recoge, de forma esquemática, el procedimiento global formulado para el procesado de Espectros Raman mediante el Análisis por Componentes Principales y su aplicación para el reconocimiento de espectros desconocidos una vez obtenida la matriz de transformación, C_{red} :



Fig. 3.1: Diagrama del procesado de Espectros Raman mediante el ACP

A continuación, a modo de ejemplo, se presenta una librería espectral simulada formada por cuatro espectros, normalizados en intensidad y estandarizados, definidos en el rango de $[0, \dots, 1999] \text{ cm}^{-1}$. A dichos espectros patrones se les aplica la técnica de reducción dimensional propuesta por el ACP normado. En el siguiente diagrama se muestra la librería simulada y su proyección en el dominio transformado generado por el ACP:

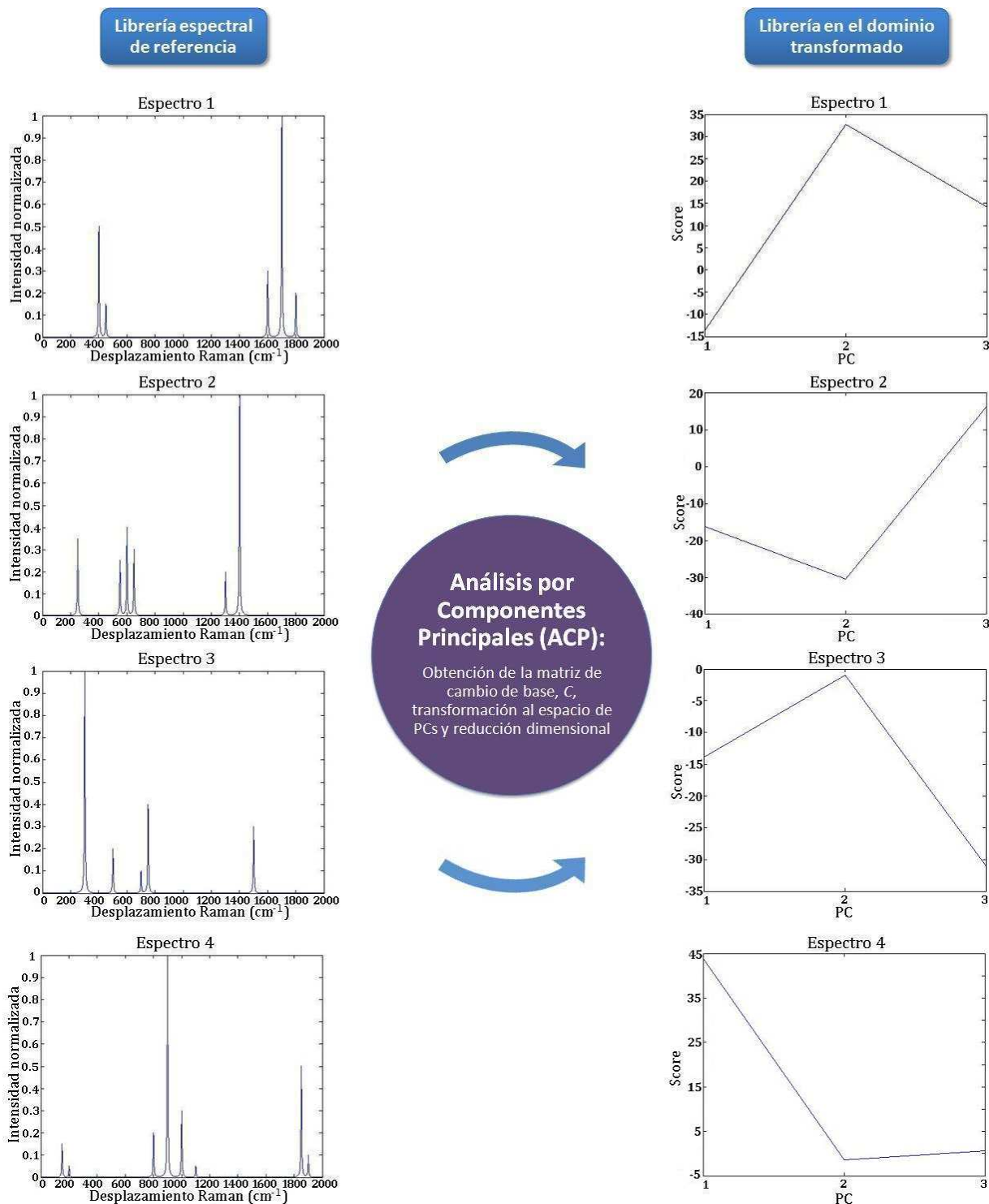


Fig. 3.2: Ejemplo de aplicación sobre una librería compuesta de 4 espectros simulados

La librería espectral simulada consta de cuatro espectros ($P = 4$) y al encontrarse definidos en el rango de $[0, \dots, 1999] \text{ cm}^{-1}$ implica que constan de un total 2000 variables ($n = 2000$). Una vez aplicada la técnica de reducción dimensional propuesta por el ACP normado a dichos espectros, se genera el espacio transformado de PCs y se obtiene la matriz reducida de cambio de base, C_{red} .

Siguiendo el criterio de contemplar el número de PCs que comporte el 100% de la varianza total de las variables originales, para la expresión de los P espectros iniciales simulados de la librería (de n componentes) en el espacio k -dimensional definido por los PCs, se tiene que:

$$k = P - 1 \ll n$$

Concretamente:

$$n = 2000$$

$$P = 4$$

$$k = P - 1 = 4 - 1 = 3 \ll n = 2000$$

Los 3 primeros PCs comportan el 100% de la varianza total acumulada de las variables originales:

PC	Varianza por PC	Varianza acumulada
1	42.77%	42.77%
2	33.38%	76.16%
3	23.84%	100%

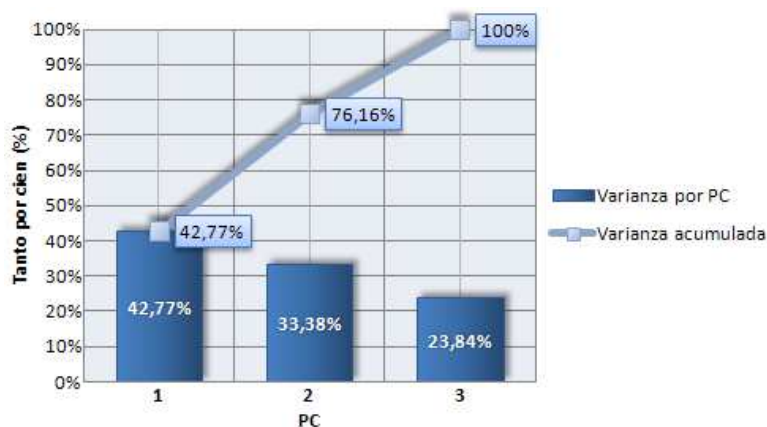


Tabla 3.1: Varianza por PC y varianza acumulada

Así, el ACP genera un espacio tridimensional ($k = 3$), consiguiendo una reducción de 2000 dimensiones del espacio inicial a 3 dimensiones del espacio transformado. Esta reducción permite una expresión simplificada de los espectros patrones en un espacio donde sus variables, incorreladas entre sí, permiten la detección de similitudes y diferencias entre espectros de una forma considerablemente más sencilla, lo cual convierte al ACP en un proceso óptimo para el reconocimiento espectral.

Si se procede a representar la librería de espectros patrones conjuntamente en ambos espacios, se puede comprobar que, efectivamente, los Espectros Raman son más fáciles de diferenciar en su representación en el espacio de PCs que en el espacio original. Se evidencia, así, que la reducción dimensional que permite el Análisis por Componentes Principales facilita de forma considerable la inspección visual de los espectros.

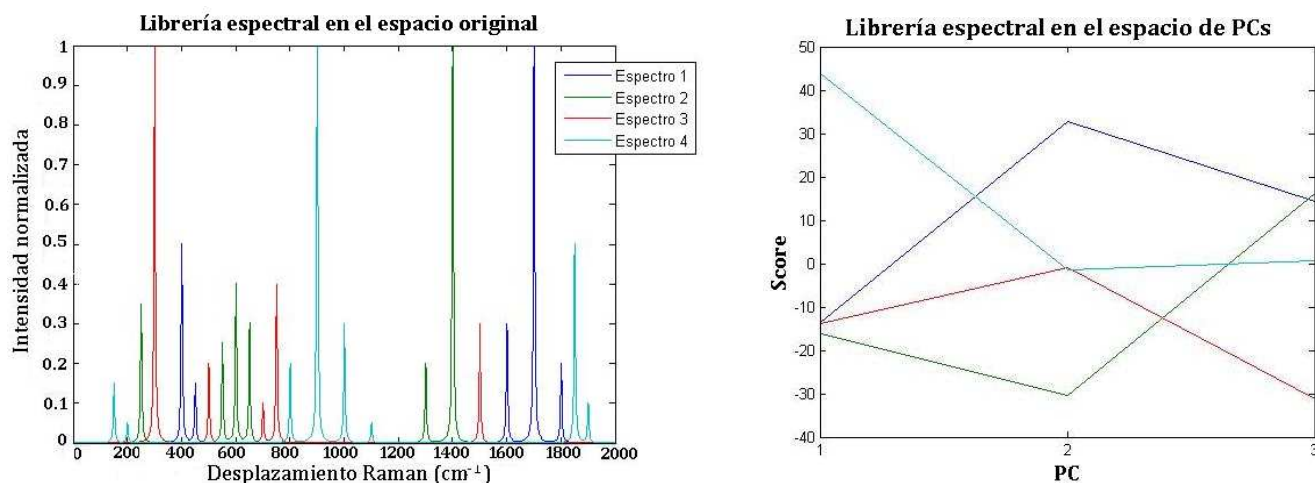


Fig. 3.3: Librería espectral en el espacio original (izquierda) y en el espacio de PCs (derecha)

Puesto que el número escogido de espectros patrones conformadores de la librería para este ejemplo es cuatro ($P = 4$), según el criterio del 100% de varianza, el ACP genera un espacio de $k = P - 1 = 3$ dimensiones, lo cual posibilita la representación de la librería proyectada en este nuevo espacio tridimensional, convirtiendo las 2000 variables de los espectros iniciales en un único punto en el espacio transformado:

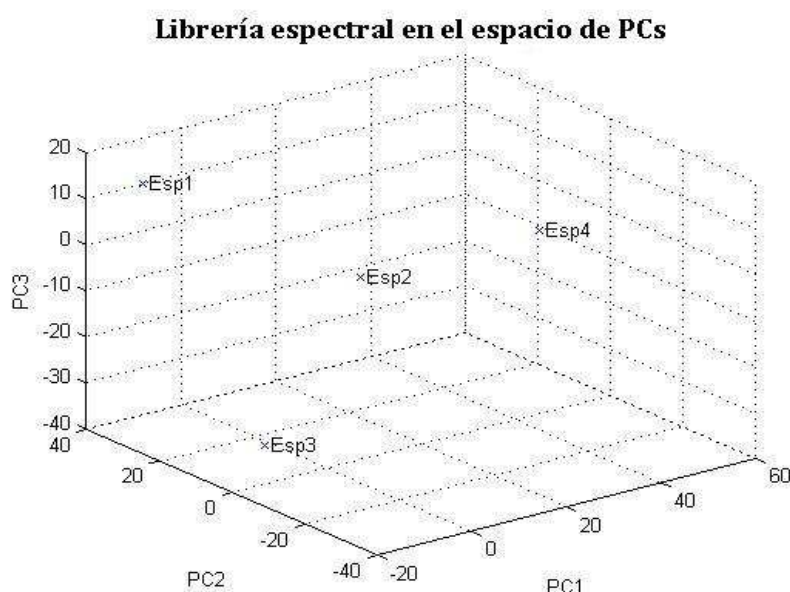


Fig. 3.4: Librería espectral en PC1-PC2-PC3

A continuación se define un nuevo espectro a tratar como espectro desconocido, aunque, para la comprobación del funcionamiento en cuanto a identificación espectral se refiere, se ha creado a partir de uno de los espectros de la librería, en concreto, a partir del Espectro 3. No obstante, para la generación del espectro desconocido, a dicho espectro patrón se le han realizado una serie de modificaciones:

- Desplazamiento de entre -3 y 3 cm^{-1} para todas las bandas
- Eliminación de la menor de las bandas secundarias
- Modificación de la intensidad relativa entre bandas en 2 u. a.
- Variación en los anchos de banda

obteniéndose el siguiente espectro:

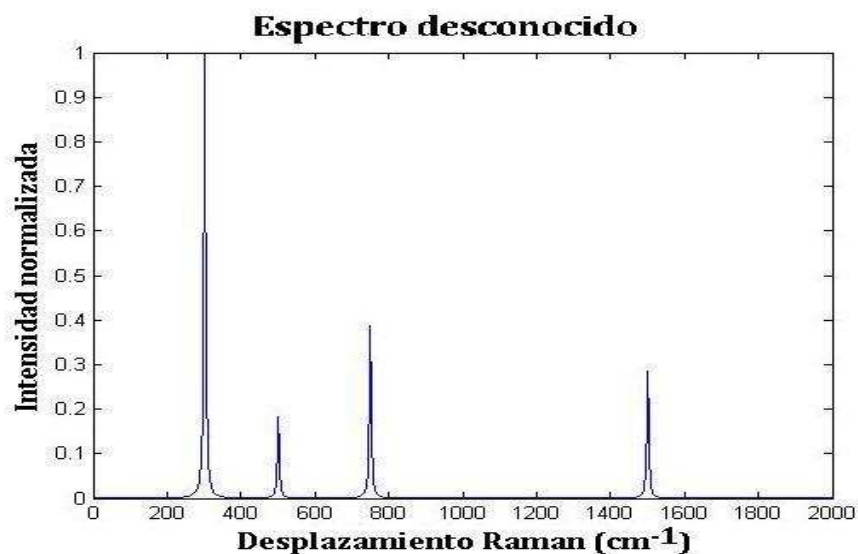


Fig. 3.5: Espectro desconocido simulado

Con la finalidad de comprobar el funcionamiento del ACP aplicado al reconocimiento de espectros, una vez obtenida la matriz de transformación, C_{red} , resultado de aplicar el ACP a los espectros patrones, se procede a proyectar este nuevo espectro en el espacio generado por la librería espectral simulada de ejemplo, siguiendo el esquema expuesto:

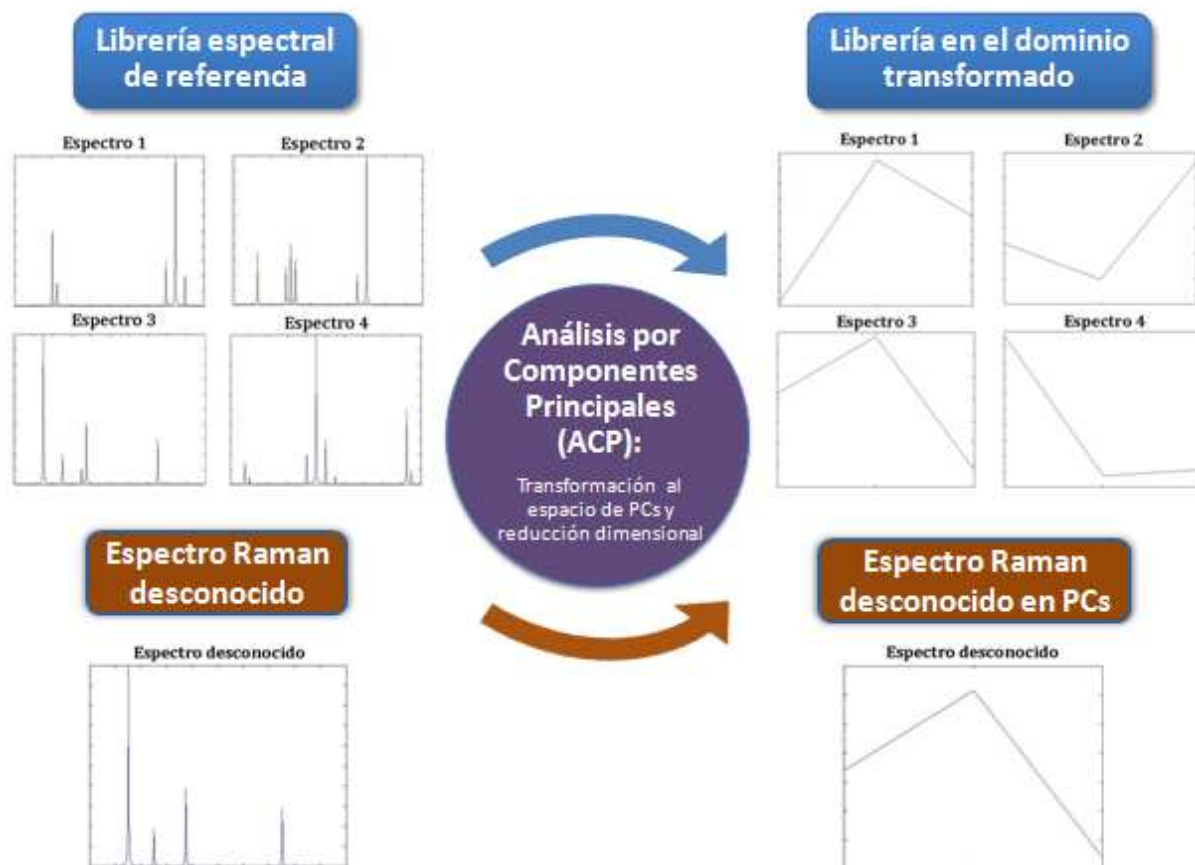


Fig. 3.6: Ejemplo de utilización del ACP aplicado a la identificación de Espectros Raman

Una vez proyectado en el espacio de los PCs generado por la librería espectral de referencia al aplicarle la técnica del ACP, se obtiene la siguiente representación del espectro desconocido:

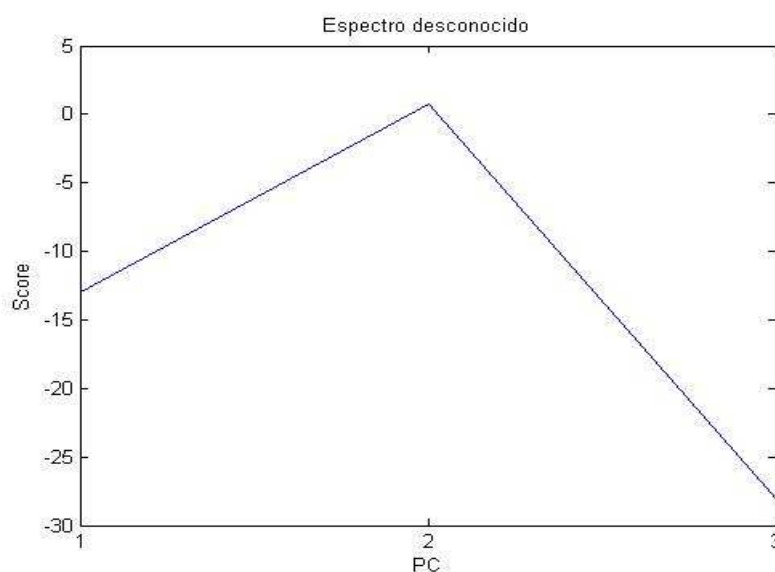


Fig. 3.7: Espectro desconocido simulado en el espacio de PCs

Llegados a este punto es cuando se procede al reconocimiento del espectro desconocido por comparación con los espectros patrones de la librería. Así, por simple inspección visual, el Espectro 3 resulta ser el más parecido al espectro desconocido:

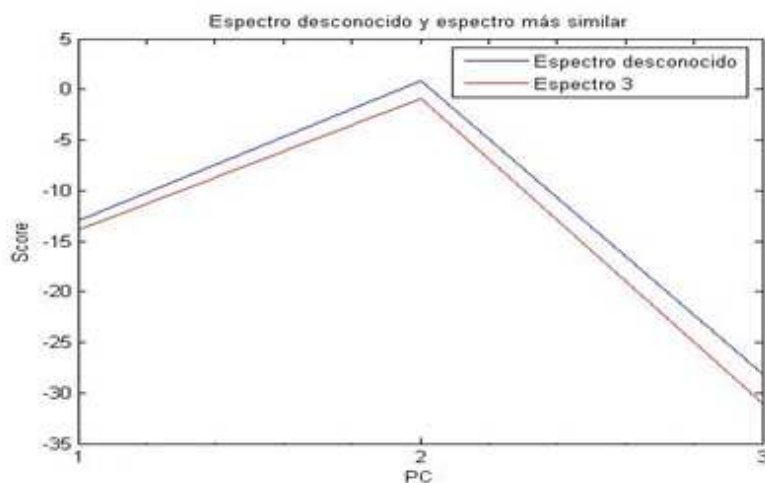


Fig. 3.8: Espectro desconocido y espectro más “parecido”

Si no se hubiese llevado a cabo ninguna modificación sobre el Espectro 3 para la creación del espectro desconocido, la representación en PCs tanto del Espectro 3 como del espectro desconocido sería la misma.

Puesto que el espacio reducido generado por el ACP es un espacio de tres dimensiones, resulta interesante la representación de la librería de referencia y del espectro desconocido como puntos en dicho espacio, revelando, así, que el espectro patrón más cercano al espectro desconocido es el Espectro 3:

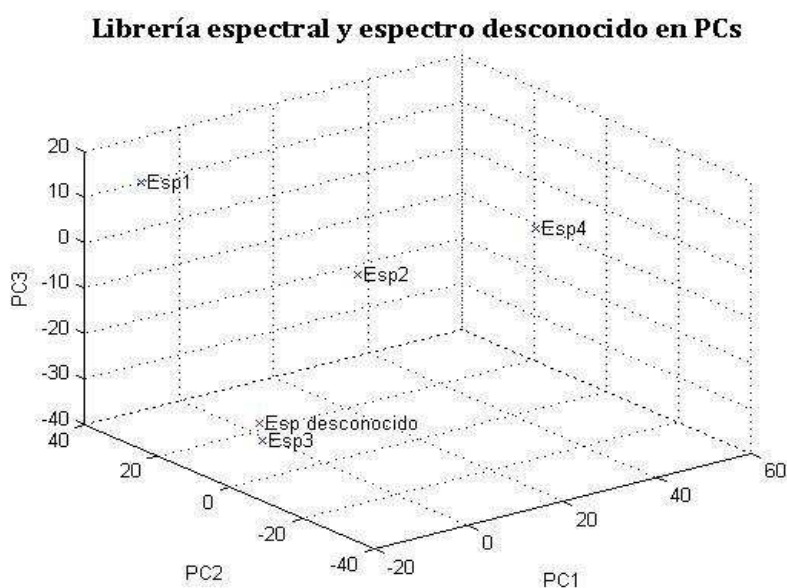


Fig. 3.9: Espectro desconocido y librería espectral en PC1-PC2-PC3

En la siguiente figura se puede observar el espectro desconocido y el Espectro 3 de forma conjunta en su representación en el espacio original:

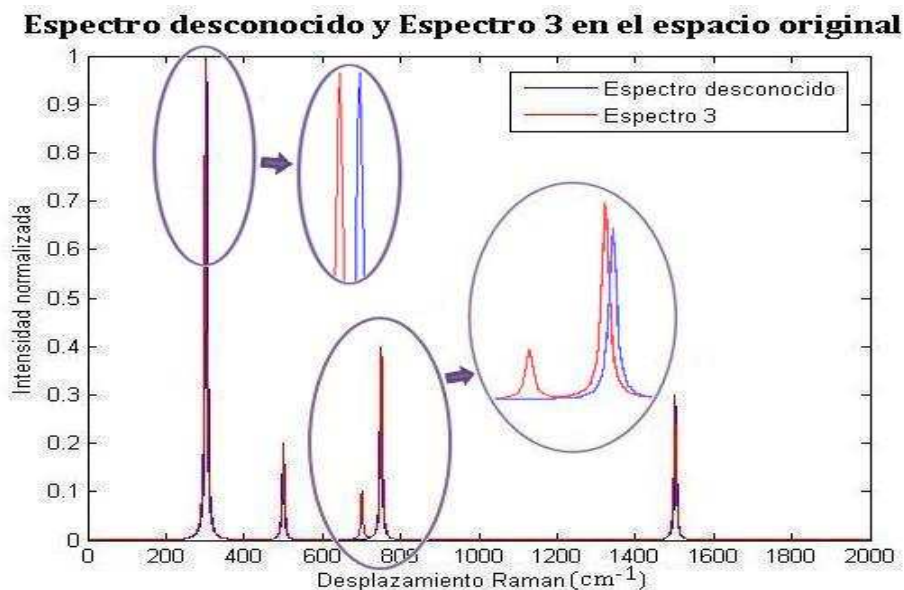


Fig. 3.10: Espectro desconocido y Espectro 3

En este ejemplo es sencillo identificar, por simple inspección visual, a qué espectro patrón se corresponde el espectro desconocido, tanto en el espacio original como en el espacio transformado (siendo más sencilla dicha identificación en el espacio de PCs). Con todo, para evitar cualquier componente subjetiva ligada a la intervención del analista a la hora de realizar el reconocimiento espectral, así como simplificar dicho procedimiento, se propone la utilización de una herramienta matemática que permita cuantificar de forma objetiva la semejanza entre espectros, teniendo en cuenta, también, la calidad de su representación.

3. 2. Cuantificación del parecido entre espectros: Distancia Euclídea.

Selección de candidatos

Dado un Espectro Raman, obtenido de un pigmento analizado sobre una obra artística, el objetivo principal es poder identificarlo de forma fiable. Para ello, se hace necesario contar con herramientas matemáticas que permitan la comparación espectral entre el nuevo espectro y los previamente almacenados en una librería espectral de referencia.

Así, con el fin de poder cuantificar el parecido entre espectros, se opta por la *Distancia Euclídea* (medida en unidades de distancia, [u]) definida formalmente como^[16]:

$$dist(P, Q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

donde $P = (p_1, p_2, \dots, p_n)$ y $Q = (q_1, q_2, \dots, q_n)$ son dos puntos de un Espacio Vectorial Euclídeo n -dimensional. Interesa, pues, que la Distancia Euclídea sea la menor posible, puesto que así se asegura que el punto P es cercano o parecido al punto Q. El caso ideal, es decir, aquel en el que la distancia es nula, indica que el punto P y el punto Q son el mismo.

En el caso particular de trabajar con Espectros Raman de pigmentos artísticos, los puntos P y Q son los espectros expresados en el espacio generado por los PCs:

$$S^{patrón_i} = [s_1^{patrón_i}, s_2^{patrón_i}, \dots, s_n^{patrón_i}]$$

Se comprueba que, para el *ACP no normado*, se cumple que:

$$dist(X_{variables\ originales}) = dist(S_{variables\ en\ PCs})$$

no siendo así para el *ACP normado*, debido al escalado que introduce imponer varianza unitaria a las variables originales normalizadas. No obstante, es siempre interesante trabajar en el espacio de PCs puesto que éste, gracias a su dimensión reducida, ofrece sencillez a la hora de realizar la identificación espectral y una mayor agilidad computacional.

Se puede calcular la Distancia Euclídea de cada espectro patrón respecto a los espectros de la librería:

$$dist(S^{patrón_j}, S^{patrón_k}) = \sqrt{\sum_{i=1}^n (s_i^{patrón_j} - s_i^{patrón_k})^2}$$

Para su posterior comparación, es útil almacenar dichos valores de Distancia Euclídea en matrices o tablas como la mostrada a continuación:

$$D = \begin{pmatrix} \text{dist}(S^{\text{patrón}_1}, S^{\text{patrón}_1}) & \text{dist}(S^{\text{patrón}_1}, S^{\text{patrón}_2}) & \dots & \text{dist}(S^{\text{patrón}_1}, S^{\text{patrón}_P}) \\ \text{dist}(S^{\text{patrón}_2}, S^{\text{patrón}_1}) & \text{dist}(S^{\text{patrón}_2}, S^{\text{patrón}_2}) & \dots & \text{dist}(S^{\text{patrón}_2}, S^{\text{patrón}_P}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{dist}(S^{\text{patrón}_P}, S^{\text{patrón}_1}) & \text{dist}(S^{\text{patrón}_P}, S^{\text{patrón}_2}) & \dots & \text{dist}(S^{\text{patrón}_P}, S^{\text{patrón}_P}) \end{pmatrix}$$

Puesto que la Distancia Euclídea entre un espectro consigo mismo es nula, resulta que:

$$\text{dist}(S^{\text{patrón}_i}, S^{\text{patrón}_i}) = 0$$

con $i=1, \dots, P$, con lo que la diagonal de la matriz de distancias de los espectros patrones, D , es nula:

$$D = \begin{pmatrix} 0 & \text{dist}(S^{\text{patrón}_1}, S^{\text{patrón}_2}) & \dots & \text{dist}(S^{\text{patrón}_1}, S^{\text{patrón}_P}) \\ \text{dist}(S^{\text{patrón}_2}, S^{\text{patrón}_1}) & 0 & \dots & \text{dist}(S^{\text{patrón}_2}, S^{\text{patrón}_P}) \\ \text{dist}(S^{\text{patrón}_3}, S^{\text{patrón}_1}) & \text{dist}(S^{\text{patrón}_3}, S^{\text{patrón}_2}) & \dots & \text{dist}(S^{\text{patrón}_3}, S^{\text{patrón}_P}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{dist}(S^{\text{patrón}_P}, S^{\text{patrón}_1}) & \text{dist}(S^{\text{patrón}_P}, S^{\text{patrón}_2}) & \dots & 0 \end{pmatrix}$$

Además, como

$$\text{dist}(S^{\text{patrón}_i}, S^{\text{patrón}_j}) = \text{dist}(S^{\text{patrón}_j}, S^{\text{patrón}_i})$$

con $i=1, \dots, P$ y $j=1, \dots, P$, la matriz D resulta ser simétrica, es decir:

$$D = D^T$$

Ahora bien, interesa cuantificar el parecido entre un espectro desconocido obtenido de una obra artística, con los espectros de la librería. De esta forma, el punto P es el nuevo espectro a analizar y el punto Q es un espectro patrón de referencia. Así, se pueden obtener tantos valores para la Distancia Euclídea como espectros contenga la librería espectral respecto al nuevo espectro desconocido:

$$\text{dist}(S^{\text{desconocido}}, S^{\text{patrón}_j}) = \sqrt{\sum_{i=1}^n (S_i^{\text{desconocido}} - S_i^{\text{patrón}_j})^2}$$

para $j=1, \dots, P$, si se tiene una librería de P espectros. Análogamente que para el caso de espectros patrones, es útil almacenar dichos valores en un vector para cada espectro desconocido, con el fin de poder ser comparados posteriormente:

$$d = [d_1 \quad d_2 \quad \dots \quad d_i \quad \dots \quad d_p] =$$

$$= [\text{dist}(S^{desc}, S^{patrón_1}) \quad \text{dist}(S^{desc}, S^{patrón_2}) \quad \dots \quad \text{dist}(S^{desc}, S^{patrón_p})]$$

Las relaciones que se establezcan entre la matriz de distancias entre los espectros patrones, D , y el vector de distancias entre un espectro desconocido y los espectros de la librería de referencia, d , darán lugar a los espectros candidatos para la identificación del espectro desconocido.

3. 3. Requisitos y criterios para la identificación

Una vez determinada la herramienta matemática a utilizar para cuantificar el parecido entre espectros, a partir de la Distancia Euclídea, se establecen una serie de criterios para llevar a cabo el reconocimiento espectral.

En primer lugar, se define el parámetro min_lib como la mínima distancia entre patrones (sin contemplar la Distancia Euclídea entre un patrón consigo mismo):

$$min_lib = \min\{D\} = \min\{\text{dist}(S^{patrón_i}, S^{patrón_j})\}$$

con $i = 1, \dots, P$, $j = 1, \dots, P$ y $i \neq j$.

Se define, además, el parámetro min_i como la menor distancia entre el patrón i -ésimo y el resto de patrones de la librería de referencia:

$$min_i = \min\{\text{dist}(S^{patrón_i}, S^{patrón_j} |_{j=1, \dots, P, i \neq j})\}$$

Resulta cómodo almacenar los valores de la menor distancia entre un patrón y el resto de la librería en un vector:

$$v_{min_i} = [min_1 \quad min_2 \quad \dots \quad min_i \quad \dots \quad min_p] =$$


$$= [\min\{\text{dist}(S^{patrón_1}, S^{patrón_j} |_{j=1, \dots, P, i \neq j})\} \quad \dots \quad \min\{\text{dist}(S^{patrón_p}, S^{patrón_j} |_{j=1, \dots, P, i \neq j})\}]$$

de donde, naturalmente:

$$min_lib = \min\{v_{min_i}\}$$

Una vez definidos dichos parámetros, se procede a exponer los criterios de la selección de candidatos para la identificación espectral:


1. Si la distancia entre el espectro desconocido y el espectro patrón k -ésimo es inferior a la mínima distancia de la librería, el espectro k -ésimo es candidato. Es decir:

Si $dist(S^{desconocido}, S^{patrón_k}) < min_lib$  El espectro k -ésimo es candidato

Puesto que este criterio puede resultar insuficiente y/o poco restrictivo para la obtención de espectros candidatos, se define un segundo criterio basado en contemplar la calidad de la representación espectral en el espacio de PCs, introducida por el coseno cuadrado:

2. Si para el espectro desconocido, el valor del coseno cuadrado es superior al de una *cota*, definida como el cociente entre la mínima distancia de la librería y la mínima distancia entre el espectro k -ésimo, se comprueba si la distancia entre el espectro desconocido y el espectro k -ésimo es inferior al valor de la menor distancia entre el espectro k -ésimo y el resto de espectros patrones. Es decir:

Sea $cota_k = min_lib/min_k$,

Si $COS2(S^{desconocido}) > cota_k$ 

Si $dist(S^{desconocido}, S^{patrón_k}) < min_k$  El espectro k -ésimo es candidato

Estos dos criterios, bajo ciertas condiciones de calidad en la representación espectral, permiten la obtención de espectros candidatos para identificar el espectro desconocido. No obstante, se pueden dar casos en que aparezcan varios espectros candidatos para un mismo espectro desconocido, con lo cual, se define un parámetro, llamado *Factor de Confianza*, FC_k , para valorar la calidad de la identificación de un determinado candidato. Para ello, se parte de la siguiente premisa del caso ideal:

$$\text{Si } E_{desc} = E_{patrón_i} \Leftrightarrow \text{dist}(S^{desc}, S^{patrón_i}) = 0$$

Así, se define el Factor de Confianza atendiendo al espectro desconocido y al espectro patrón k -ésimo:

$$FC_k = 1 - \frac{\text{dist}(S^{desc}, S^{patrón_k})}{\min_k}$$

En caso de aparecer varios candidatos para un determinado espectro desconocido y con el fin de homogeneizar el Factor de Confianza, éste se obtiene como:

$$FC_k = 1 - \frac{\text{dist}(S^{desc}, S^{patrón_k})}{\max \{ \min_{cands} \}}$$

donde, $\max \{ \min_{cands} \}$ es el valor máximo de las Distancias Euclídeas entre los espectros patrones candidatos obtenidos y el resto de patrones.

Se genera, por tanto, un Factor de Confianza de la identificación entre el espectro desconocido y el espectro k -ésimo, FC_k , que cumple:

$$0 \leq FC_k \leq 1$$

donde $FC_k = 0$ indica que el espectro desconocido no se corresponde con el espectro patrón k -ésimo y $FC_k = 1$ indica el caso ideal, en el que el espectro desconocido es exactamente el espectro patrón k -ésimo. Sin embargo, para que este parámetro resulte más intuitivo, se ofrece en tanto por cien.

Con el objetivo de clarificar los conceptos expuestos, y atenderlos de forma práctica, se toma el ejemplo anterior propuesto, obteniéndose los siguientes resultados para la matriz de distancias de la librería espectral, D , el vector de mínimos, v_{\min_i} , y la Distancia Euclídea mínima de la librería:

$$D = \begin{pmatrix} 0 & 63.2925 & 56.4403 & 68.2592 \\ 63.2925 & 0 & 55.7893 & 68.5687 \\ 56.4403 & 55.7893 & 0 & 65.8417 \\ 68.2592 & 68.5687 & 65.8417 & 0 \end{pmatrix}$$

$$v_{min_i} = [56.4403 \quad 55.7893 \quad 55.7893 \quad 65.8417]$$

de donde:

$$min_lib = min\{D\} = min\{v_{min_i}\} = 55.7893$$

distancia mínima que se da entre el Espectro 2 y el Espectro 3.

Asimismo, se obtienen los siguientes resultados de Distancias Euclídeas entre el espectro desconocido y la librería de referencia, que se almacenan en el vector d :

$$d = [53.0706 \quad 54.3644 \quad 3.5095 \quad 63.7243]$$

y el siguiente coseno cuadrado, que indica una calidad de la representación en el espacio de PCs relativamente alta:

$$COS2 = 0.8078$$

Puesto que los valores de la distancia entre el espectro desconocido y los espectros patrones son inferiores a la distancia mínima de la librería:

$$\left. \begin{array}{l} d_1 = 53.0706 \\ d_2 = 54.3644 \\ d_3 = 3.5095 \end{array} \right\} < min_lib = 55.7893$$

se aplica el primer criterio, obteniéndose tres de los cuatro espectros de la librería de referencia como espectros candidatos para identificar al espectro desconocido, el Espectro 1, el Espectro 2 y el Espectro 3. No obstante, se puede comprobar que la distancia entre el espectro desconocido y el Espectro 3, $d_3 = 3.5095$, es la menor, así como considerablemente pequeña, indicando que el Espectro 3 es el más parecido al espectro desconocido, como se comprobó de forma gráfica en el primer apartado de este capítulo.

Puesto que se han obtenido tres espectros candidatos, se calculan tres Factores de Confianza, uno para cada candidato, facilitando estos resultados, ya expresados en tanto por cien, a partir de:

$$\begin{aligned} max \{min_{cands}\} &= max\{min_1 \quad min_2 \quad min_3\} = \\ &= max\{56.4403, 55.7893, 55.7893\} = 56.4403 \end{aligned}$$

$$FC_1 = 5.97\%$$

$$FC_2 = 3.43\%$$

$$FC_3 = 93.78\%$$

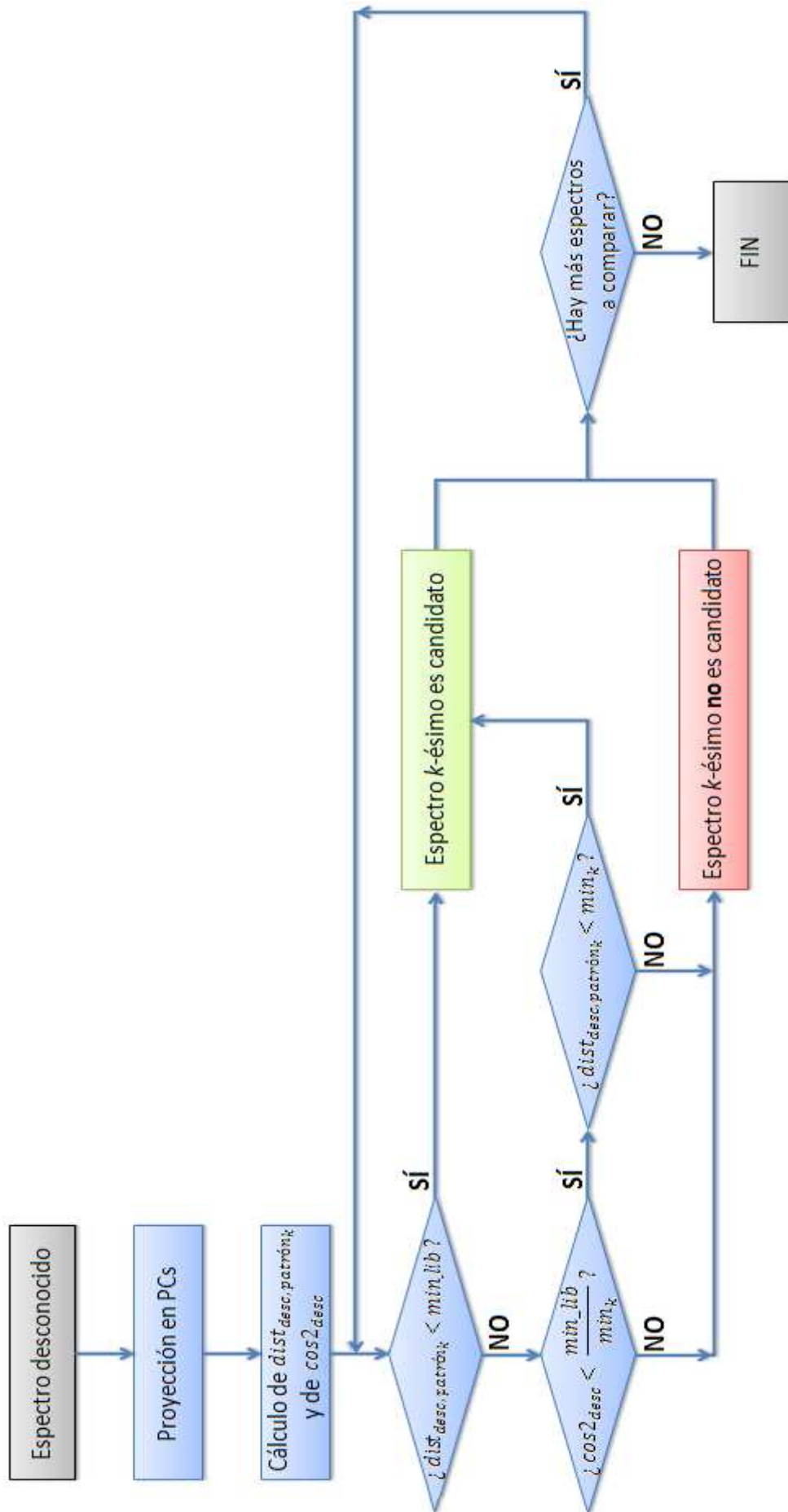
El mayor Factor de Confianza es el correspondiente al Espectro 3, con un valor cercano al ideal, mientras que los valores de Factor de Confianza para el resto de espectros patrones son relativamente cercanos al peor caso, corroborando, a partir de esta valoración, que el espectro desconocido se corresponde con el Espectro 3.

A lo largo de estos primeros capítulos se ha presentado la técnica del Análisis por Componentes Principales y su aplicación al procesado de Espectros Raman. Se ha expuesto, también, la metodología para la identificación espectral a partir de la proyección de un espectro desconocido en el espacio generado por el ACP al ser aplicado a una librería espectral, haciendo uso de la herramienta de cuantificación de parecido entre espectros basada en la Distancia Euclídea.

A continuación, se procede a analizar, de forma teórica, la metodología para la identificación espectral a partir de una librería espectral simulada y la modificación de diferentes factores en dichos espectros (como la adición de ruido, la variación del ancho de banda o de la intensidad relativa, etc.), obteniéndose una serie de resultados que permiten acotar las condiciones favorables para la identificación espectral.

Finalmente, se pretende analizar el comportamiento global de la identificación de Espectros Raman de materiales artísticos a partir del ACP y el sistema de reconocimiento automático implementado en casos experimentales.

El organigrama de la página siguiente muestra la metodología implementada para la identificación de Espectros Raman de pigmentos a aplicar a partir de su expresión en el espacio reducido generado por el ACP.



4. ESTUDIO TEÓRICO DE LA METODOLOGÍA

4. 1. Librería espectral de referencia. Importancia de la elección de la librería de Espectros Raman patrones

En este capítulo se lleva a cabo un estudio teórico del comportamiento del ACP normado aplicado a la identificación espectral. Se escoge la versión normada del Análisis por Componentes Principales puesto que esta versión trata a todas las variables espectrales de forma uniforme, contemplándolas a todas por igual al imponerles varianza unitaria, siendo idónea, por ello, para resaltar similitudes y diferencias entre un espectro desconocido y los espectros patrones.

Con el fin de llevar a cabo dicho estudio, se simula una librería espectral generada por 10 espectros patrones ($P = 10$) definidos en el rango $[0, \dots, 100] \text{ cm}^{-1}$, lo que implica un total de 100 variables ($N = 100$). Se parte de una situación simplificada respecto a una situación real en cuanto a número de espectros y rango espectral, pero, con el fin de extraer conclusiones acerca del reconocimiento espectral a partir del ACP normado, se fuerzan situaciones relacionadas con el parecido entre espectros tanto por la situación de determinadas bandas como por sus anchuras. La librería generada consta de espectros con múltiples bandas (entre 2 y 5) definidas a lo largo de todo el rango espectral contemplado, con diferentes amplitudes y anchos de banda.

Para realizar la simulación de un espectro se propone utilizar la función Lorentziana para implementar las bandas Raman en los espectros simulados, definida como:

$$f(x, x_0, A, B) = \frac{A}{1 + \left(\frac{x - x_0}{B/2}\right)^2}$$

donde x_0 es el valor central de una banda, A la amplitud, y B el ancho cuando la amplitud de la banda ha caído a la mitad.

Los espectros, en general, están formados por una banda de mayor intensidad respecto al resto de bandas, que se conoce como “Banda Fundamental”, y el resto de bandas de menor intensidad, en caso de haberlas, son conocidas como “Bandas Secundarias”. La siguiente tabla recoge la posición de las bandas de los espectros patrones simulados, ordenando las Bandas Secundarias de mayor a menor en intensidad:

Espectro	Banda fundamental	Bandas secundarias			
Espectro 1	23cm ⁻¹	83 cm ⁻¹	64 cm ⁻¹		
Espectro 2	72 cm ⁻¹	34 cm ⁻¹	53 cm ⁻¹	87 cm ⁻¹	
Espectro 3	87 cm ⁻¹	5 cm ⁻¹	30 cm ⁻¹		
Espectro 4	57 cm ⁻¹	38 cm ⁻¹	11 cm ⁻¹	91 cm ⁻¹	
Espectro 5	91 cm ⁻¹	11 cm ⁻¹	38 cm ⁻¹	57 cm ⁻¹	
Espectro 6	46 cm ⁻¹	96 cm ⁻¹	79 cm ⁻¹	61 cm ⁻¹	20 cm ⁻¹
Espectro 7	17 cm ⁻¹	75 cm ⁻¹			
Espectro 8	10 cm ⁻¹	70 cm ⁻¹			
Espectro 9	81 cm ⁻¹	14 cm ⁻¹			
Espectro 10	27 cm ⁻¹	7 cm ⁻¹	84 cm ⁻¹		

Tabla 4.1: Posición de las bandas de los espectros patrones simulados

En la figura siguiente se pueden observar, de forma conjunta, los espectros generadores de la librería espectral de referencia:

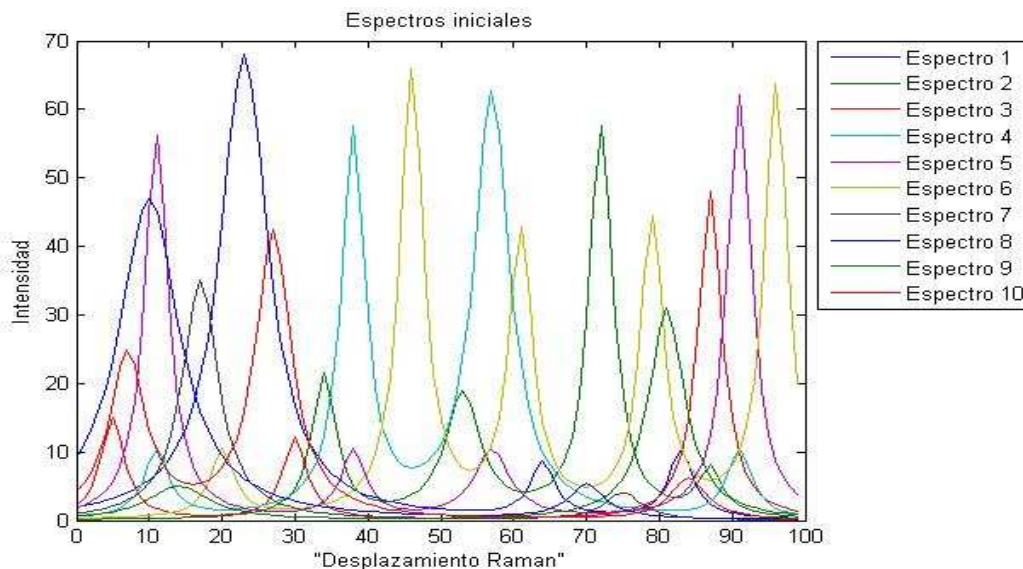


Fig. 4.1: Representación conjunta de la librería generada

La figura Fig. 4.2 muestra gráficamente la librería espectral de referencia con los espectros normalizados:

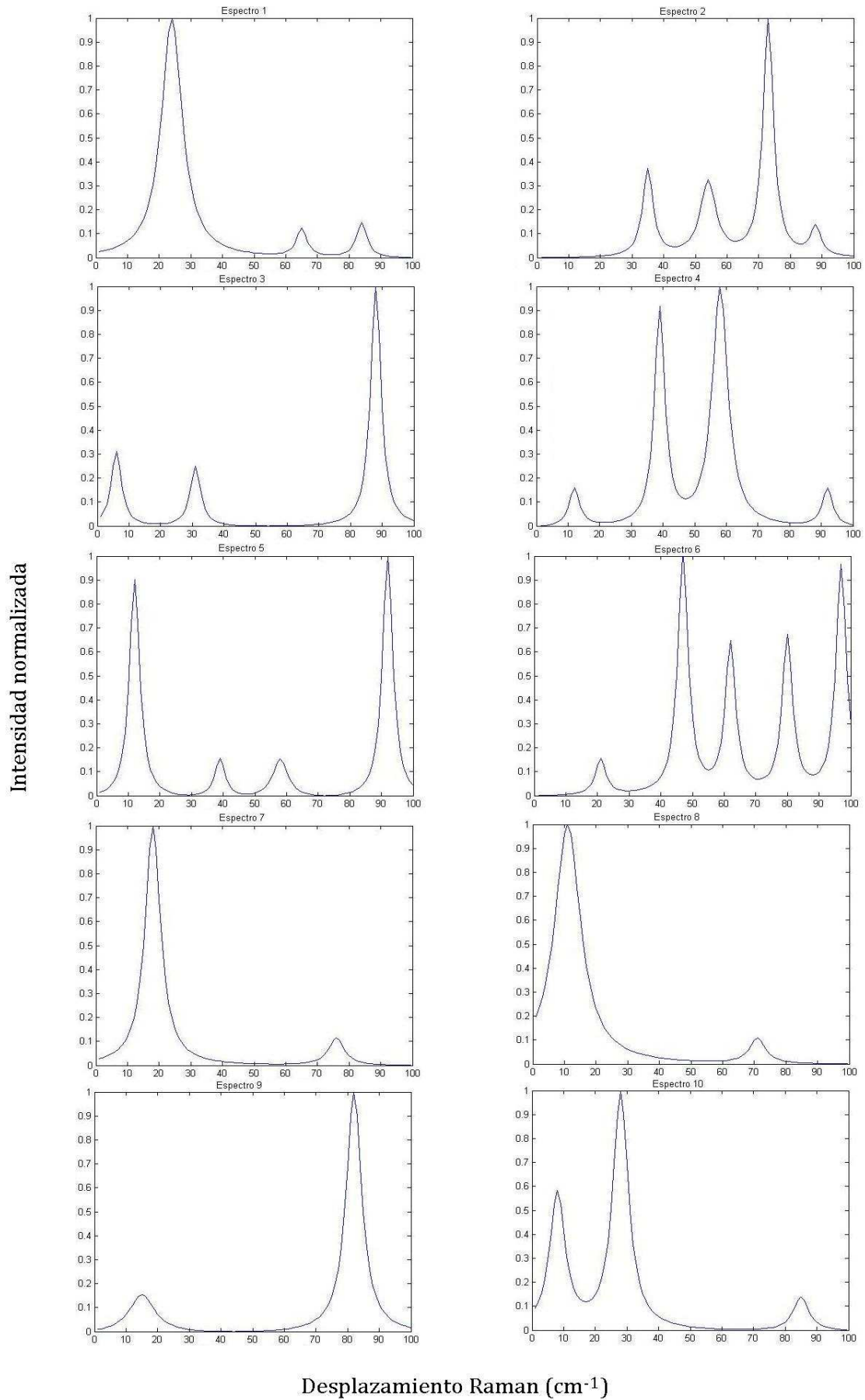


Fig. 4.2: Representación de los espectros patrones en el espacio original

Como ya se ha explicado en el capítulo anterior, los espectros son estandarizados, y se crea la matriz X_{est} . Sobre esta matriz se aplica la técnica del Análisis por Componentes Principales, calculando la matriz de covarianzas, X_{cov} , para diagonalizarla con el objetivo de generar la matriz de cambio de base, C , una vez ordenados los vep 's en función de sus respectivos vap 's. Con el fin de conseguir la reducción de dimensionalidad, se escoge el criterio de trabajar con tantos PCs como se contemple el 100% de la varianza total de las variables originales, dando lugar a $k = P - 1 = 9$ ($k \ll N$, con $P = 10$ patrones y $N = 100$ variables). Una vez calculada la matriz de transformación reducida, C_{red} , se procede a proyectar los espectros normalizados y estandarizados en el nuevo espacio generado por los PCs a partir de:

$$S_{red} = X_{est} \cdot C_{red}$$

La siguiente tabla recoge los valores de los *scores* de cada uno de los espectros patrones una vez realizada la proyección en el espacio de los PCs:

$S_{red} \backslash PC$	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
S_{red_1}	-2,6523	-1,4142	-0,68908	5,4843	2,1201	-2,6979	1,4193	-1,1085	-2,2189
S_{red_2}	5,0534	-4,401	-7,0044	-1,0634	-3,1205	1,7662	1,9114	-0,2956	0,076114
S_{red_3}	-3,1197	2,0962	-1,6282	-3,3203	4,2571	1,9051	0,75769	3,194	-0,89352
S_{red_4}	7,6892	-6,6608	3,6139	-1,9374	2,1335	-1,8987	-2,4466	0,43998	-0,007988
S_{red_5}	-1,5378	1,648	3,5364	-5,2212	0,33462	-1,0174	3,466	-2,374	0,38205
S_{red_6}	10,245	6,2192	2,6412	3,5828	-0,77623	2,1564	0,6661	0,43824	0,15056
S_{red_7}	-3,6403	0,61019	0,054817	1,629	-2,9535	-4,8573	0,69905	2,2688	1,713
S_{red_8}	-6,0783	-1,0484	3,5035	-0,58012	-5,4495	2,8577	-1,8258	0,23667	-1,3377
S_{red_9}	-0,40517	5,3028	-4,2333	-1,9326	0,39727	-1,8624	-3,9963	-1,7202	0,0036574
$S_{red_{10}}$	-5,5538	-2,352	0,20522	3,3589	3,057	3,6483	-0,65101	-1,0795	2,1327

Tabla 4.2: Scores de los espectros patrones

es decir, S_{red_i} corresponde al espectro patrón P_i representado en el espacio de los Componentes Principales generado por la librería espectral.

La figura Fig. 4.3 muestra cada uno de los espectros de la librería de referencia expresados en el espacio transformado de los PCs, representando sus *scores*:

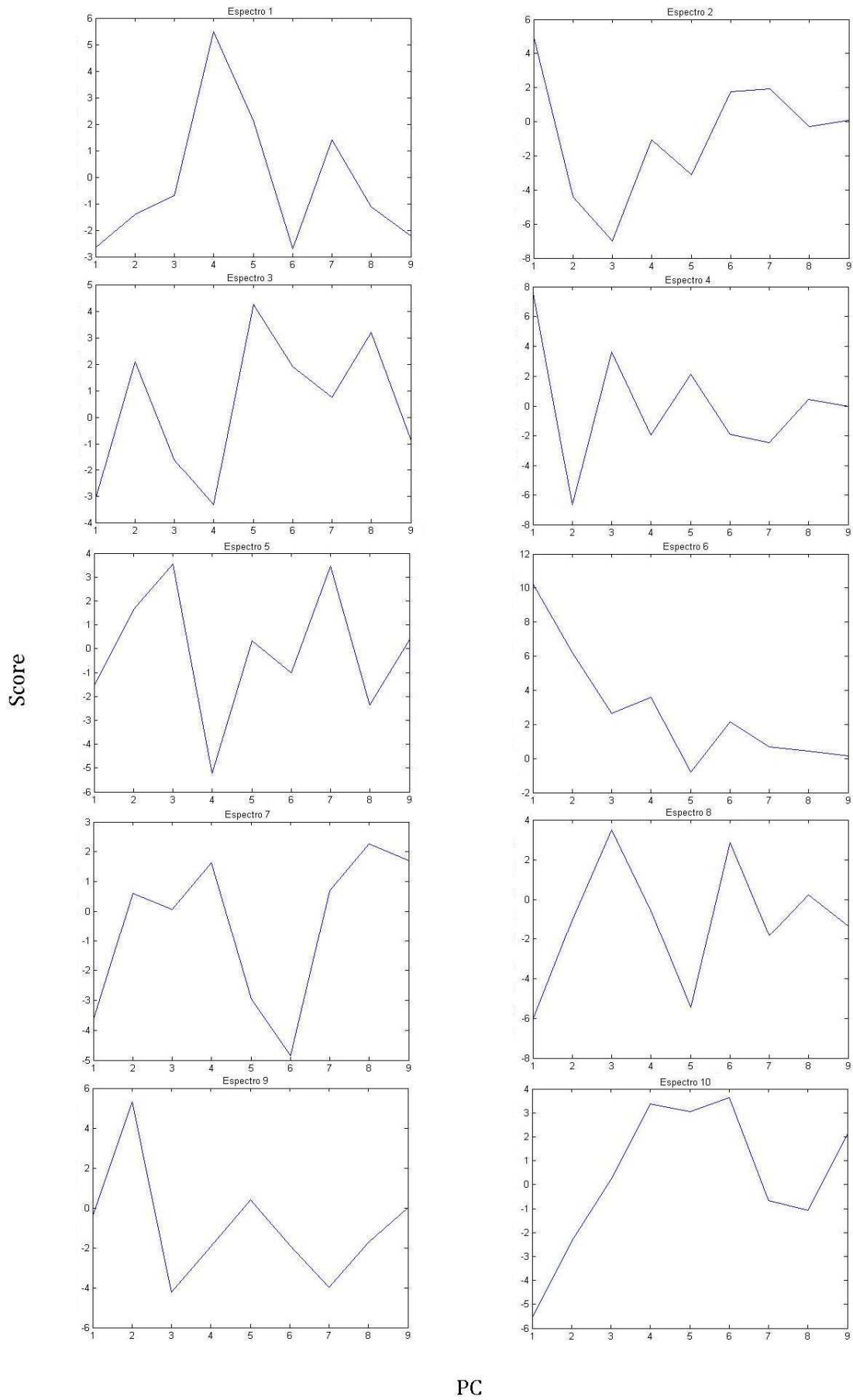


Fig. 4.3: Representación de la librería espectral de referencia en PCs

La representación conjunta de los espectros expresados en PCs se recoge en la figura Fig. 4.4:

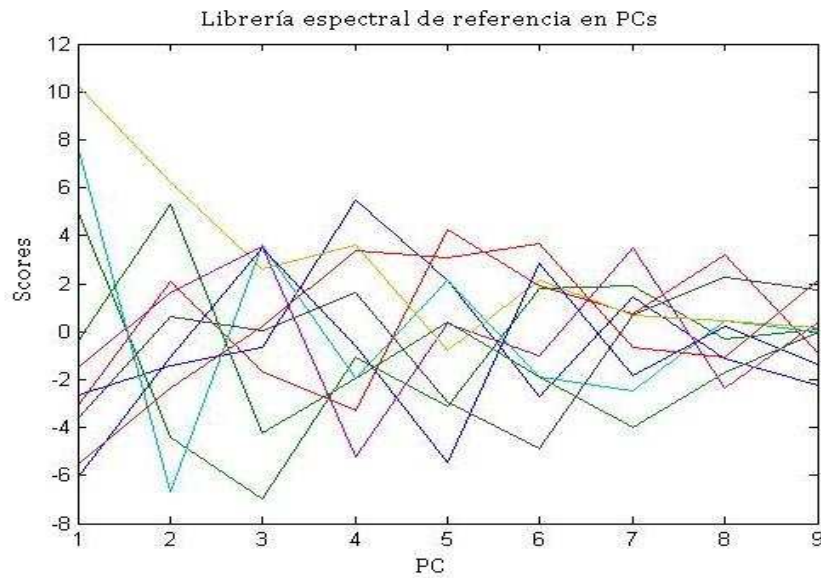


Fig. 4.4: Representación conjunta de los espectros patrones en PCs

El gráfico siguiente muestra los valores de varianza de las variables originales que recogen cada PC y la varianza acumulada cuando se contempla la suma de las varianzas por PC.

PC	Varianza por PC	Varianza acumulada
1	32.22%	32.22%
2	16.29%	48.51%
3	12.77%	61.29%
4	11.57%	72.85%
5	9.50%	82.36%
6	7.98%	90.34%
7	4.99%	95.33%
8	2.99%	98.31%
9	1.69%	100%

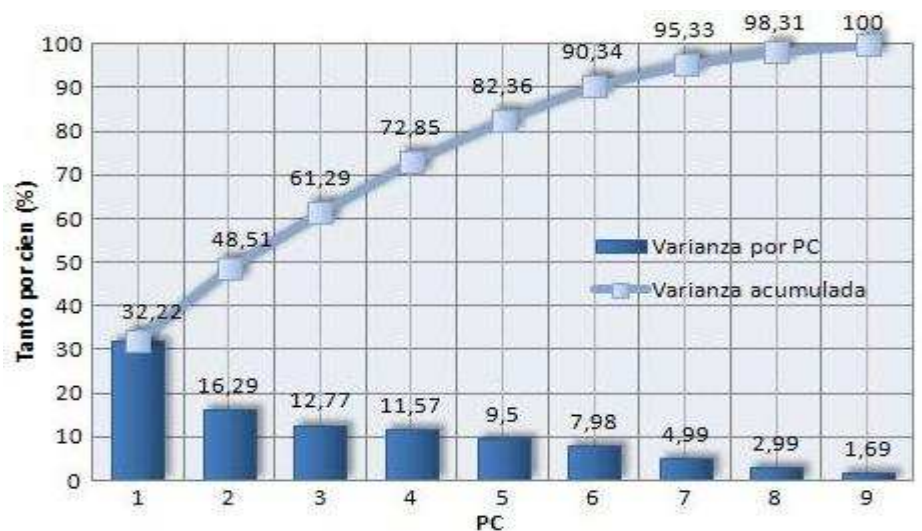


Tabla 4.3: Valores de varianza por PC y varianza acumulada

Puesto que los tres primeros PCs recogen un total de un 61.29% de varianza de las variables originales, la representación tridimensional de los espectros en el espacio formado por PC1, PC2 y PC3 muestra gráficamente relaciones de cercanía y parecido entre espectros:

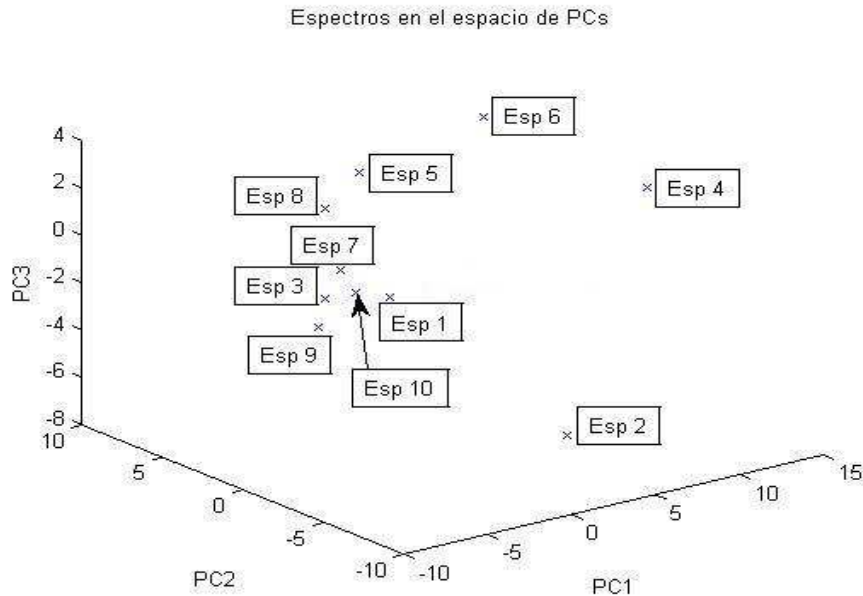


Fig. 4.5: Representación de los espectros patrones en el espacio PC1, PC2 y PC3

Además de representar los espectros en el espacio de los Componentes Principales, con el fin de analizar qué variables originales tienen un mayor peso en la generación de los PCs, resulta interesante representar gráficamente la relación que existe entre los dos conjuntos de variables, las originales y las nuevas (los PCs). Esta representación se consigue mediante los *biplots*, que permiten visualizar los coeficientes de las variables originales que generan los PCs (*loads*) y analizar la magnitud de la influencia de cada variable a los PCs, siendo especialmente representativo para los tres primeros PCs, dado el relativamente alto porcentaje de varianza acumulada recogida por éstos.

Los *biplots* dan información sobre qué variables iniciales aportan más peso en la generación de los PCs, y por tanto, se puede deducir qué zonas de los espectros son las más valoradas en la distribución de los datos en este espacio. En las siguientes figuras (Fig. 4.6a, Fig. 4.6b y Fig. 4.7) se muestran los *biplots* para PC1 y PC2, así como para PC2 y PC3 y un *biplot 3D* (para PC1, PC2 y PC3).

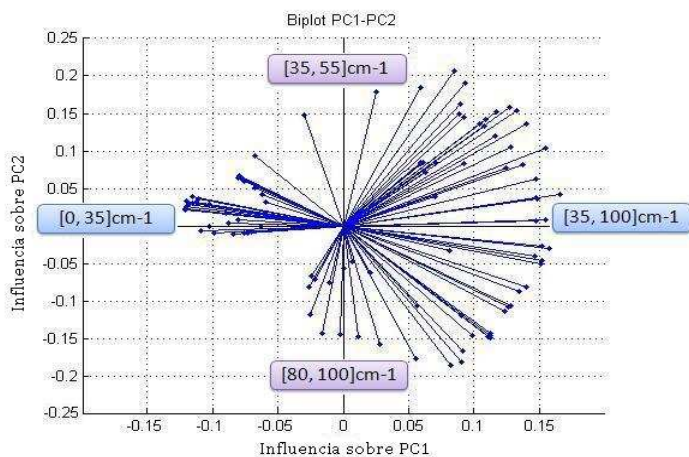


Fig. 4.6a: Biplot de PC1-PC2

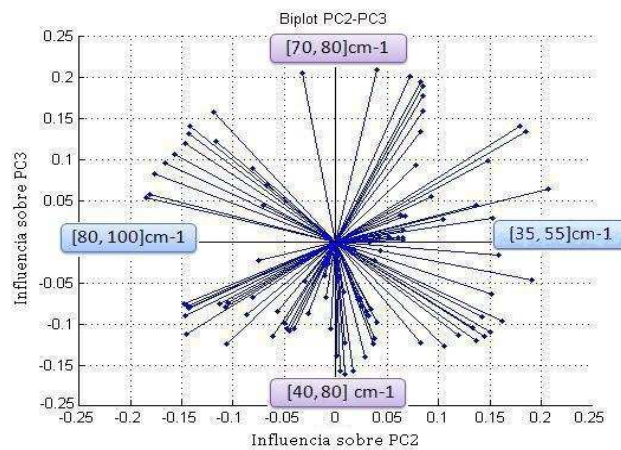


Fig. 4.6b: Biplot de PC2-PC3

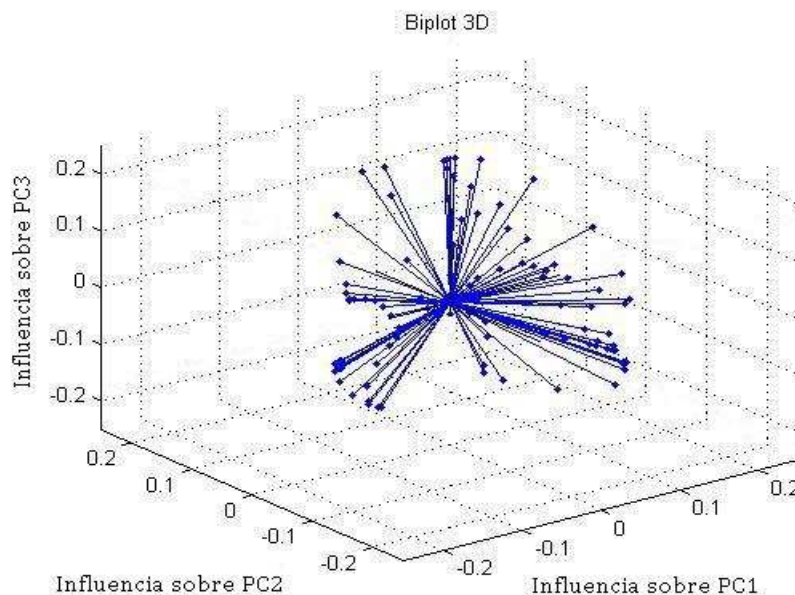


Fig. 4.7: Biplot 3D (PC1, PC2 y PC3)

En la figura Fig. 4.6a, se han marcado los rangos de influencia de las variables originales sobre PC1 (en azul) y PC2 (en morado). Así, el rango de variables originales que influyen en el cálculo de PC1 es, por un lado $[0 - 35] \text{ cm}^{-1}$, y, por otro, $[35 - 100] \text{ cm}^{-1}$, resultando, por tanto, todo el rango espectral, $[0 - 100] \text{ cm}^{-1}$, con lo que, todas las variables originales influyen en PC1. Para el cálculo del resto de PCs la influencia de variables originales va decreciendo; por ejemplo, para el cálculo de PC2, el rango de variables originales que influyen en su cálculo es $[35 - 55, 80 - 100] \text{ cm}^{-1}$, o para el PC3 (Fig. 4.6b, en morado), cuyo rango es $[40 - 80] \text{ cm}^{-1}$, corroborando así que los primeros PCs son los más representativos.

Una vez obtenidas las expresiones de los espectros de la librería de referencia en el espacio transformado de los PCs, se calculan las Distancias Euclídeas entre espectros patrones, D , recogidas en la *Tabla 4.4*:

Espectros	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10
Esp 1	0	14,304	11,722	15,197	12,714	16,728	8,8476	12,942	12,492	8,8877
Esp 2	14,304	0	14,557	13,643	15,328	16,19	14,574	16,391	13,96	15,536
Esp 3	11,722	14,557	0	16,141	9,8429	17,192	11,627	12,776	10,149	10,372
Esp 4	15,197	13,643	16,141	0	14,56	15,442	16,045	17,47	16,742	16,594
Esp 5	12,714	15,328	9,8429	14,56	0	16,283	11,004	11,648	11,95	12,95
Esp 6	16,728	16,19	17,192	15,442	16,283	0	17,161	19,187	15,359	18,813
Esp 7	8,8476	14,574	11,627	16,045	11,004	17,161	0	10,537	11,164	11,71
Esp 8	12,942	16,391	12,776	17,47	11,648	19,187	10,537	0	14,178	10,794
Esp 9	12,492	13,96	10,149	16,742	11,95	15,359	11,164	14,178	0	13,652
Esp 10	8,8877	15,536	10,372	16,594	12,95	18,813	11,71	10,794	13,652	0

Tabla 4.4: Distancia Euclídea entre espectros de la librería de referencia

de donde se pueden extraer las distancias mínimas entre un espectro patrón y el resto de espectros patrones, v_{min_i} , que se recogen en la *Tabla 4.5*:

Espectro patrón	Resto de espectros
Espectro 1	8,8476
Espectro 2	13,6434
Espectro 3	9,8429
Espectro 4	13,6434
Espectro 5	9,8429
Espectro 6	15,3585
Espectro 7	8,8476
Espectro 8	10,5371
Espectro 9	10,1494
Espectro 10	8,8877

Tabla 4.5: Distancia Euclídea mínima entre un espectro patrón y el resto de patrones

La Distancia Euclídea mínima de la librería espectral de referencia, min_{lib} , caracteriza a la librería y permite establecer uno de los límites para el reconocimiento de espectros desconocidos. Como se comentó en el *Capítulo 3*, el valor mínimo de la Distancia Euclídea entre espectros de la librería (*Tabla 4.4*) coincide con el valor mínimo de la Distancia Euclídea mínima entre un espectro patrón y el resto de patrones (*Tabla 4.5*), es decir:

$$min_lib = min\{D\} = min\{v_{min_i}\}$$

En el caso bajo análisis, la Distancia Euclídea mínima de la librería se obtiene entre el Espectro 1 y el Espectro 7, lo cual implica que éstos son los espectros más parecidos dentro de la librería espectral. El valor de dicho parámetro es:

$$min_lib = 8.8476u$$

Las figuras Fig. 4.8 a) y b) recogen la representación conjunta de los espectros más parecidos, tanto en su expresión original como en su expresión en el espacio de los PCs, corroborando, por la posición de sus bandas, que son espectros relativamente parecidos.

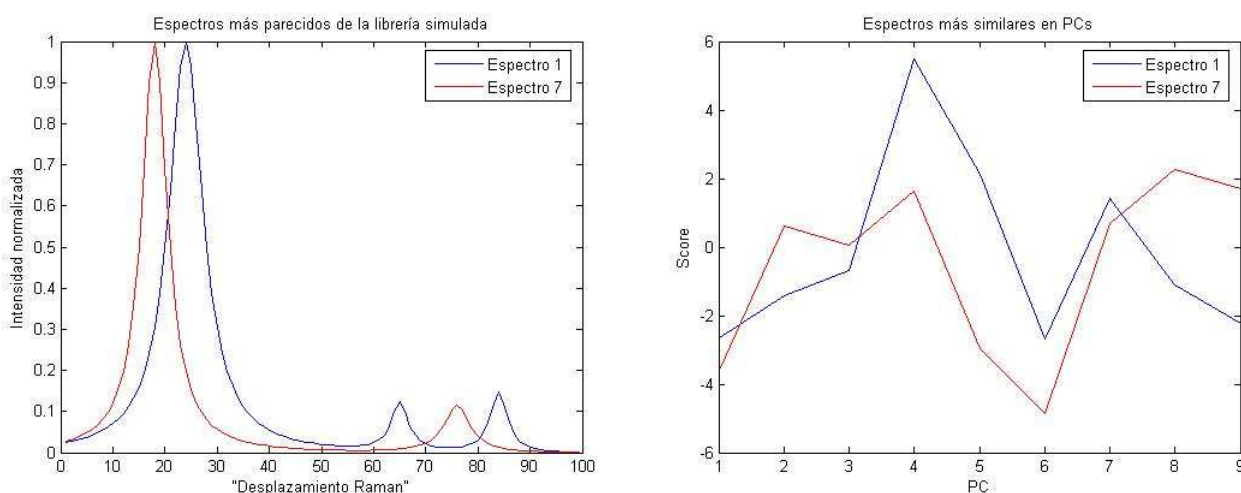


Fig. 4.8 a) y b): Espectros más parecidos de la librería de referencia simulada tanto en el espacio original (izquierda) como en el espacio transformado de los PCs (derecha)

En los siguientes apartados se analiza la identificación de un espectro desconocido respecto a la librería de referencia, cuando este espectro desconocido es perfectamente coincidente con un determinado espectro patrón, o, por el contrario, no coincide total o parcialmente con ningún espectro de la librería. Esta “no coincidencia” puede deberse a que el espectro desconocido no se corresponde con la librería de referencia seleccionada o a diferentes problemas prácticos que dificultan el reconocimiento espectral, como ruido *shot*, estrechamiento de bandas, posibles bandas enmascaradas, variación de intensidades relativas entre bandas o desplazamiento entre bandas, proporcionando un límite para el reconocimiento en cada uno de dichos casos.

4. 2. Estudio de identificación de Espectros Raman

En este apartado se realiza el análisis del reconocimiento espectral en dos situaciones diferentes: Primeramente, cuando un espectro desconocido es igual a un espectro patrón, caso totalmente teórico e ideal, y, finalmente, cuando un espectro desconocido es diferente al resto de espectros de la librería de referencia.

En el primer caso, se espera obtener que el sistema implementado encuentre únicamente un candidato con un factor de confianza máximo (del 100%) y que sea el espectro patrón correspondiente. En cambio, en el segundo caso, se espera no obtener ningún candidato, o bien, en caso de obtenerse, que el factor de confianza sea relativamente pequeño (inferior, por ejemplo, al 10%).

4. 2. 1. Espectro Raman desconocido igual a un Espectro Raman patrón

Se simula un nuevo espectro mediante la función Lorenziana para la creación de sus bandas. Este espectro es totalmente coincidente con un espectro determinado de la librería de referencia, en concreto, con el Espectro 2.

Una vez proyectado el espectro desconocido en el espacio generado al aplicar el ACP a la librería espectral de referencia, se obtiene que la distancia entre el espectro desconocido y el espectro patrón 2 es nula (fila 2 de la *Tabla 4.4*), siendo éste, atendiendo a los criterios de identificación expuestos en el *Capítulo 3*, el único candidato, con un factor de confianza del 100% en la identificación. Este resultado se puede extrapolar a cualquier espectro desconocido que coincida totalmente con un espectro de la librería de referencia.

Cabe resaltar que esta situación es prácticamente imposible en la práctica pero tiene el interés de comprobar la “repetitividad” de la metodología y que el método es consistente.

4. 2. 2. Espectro Raman desconocido distinto a cualquier Espectro Raman patrón

Se simula un nuevo espectro desconocido (*Fig. 4.9 a)*) cuyas bandas, al contrario que en el caso anterior, no coinciden con ningún espectro de la librería de referencia. La *Tabla 4.6* recoge la posición de las bandas de este nuevo espectro y en la *Fig. 4.9 b)* se puede ver la proyección del espectro desconocido en el espacio de los PCs.

Espectro	Banda fundamental	Banda secundaria
Espectro desconocido	32cm ⁻¹	68 cm ⁻¹

Tabla 4.6: Posición de las bandas del espectro desconocido

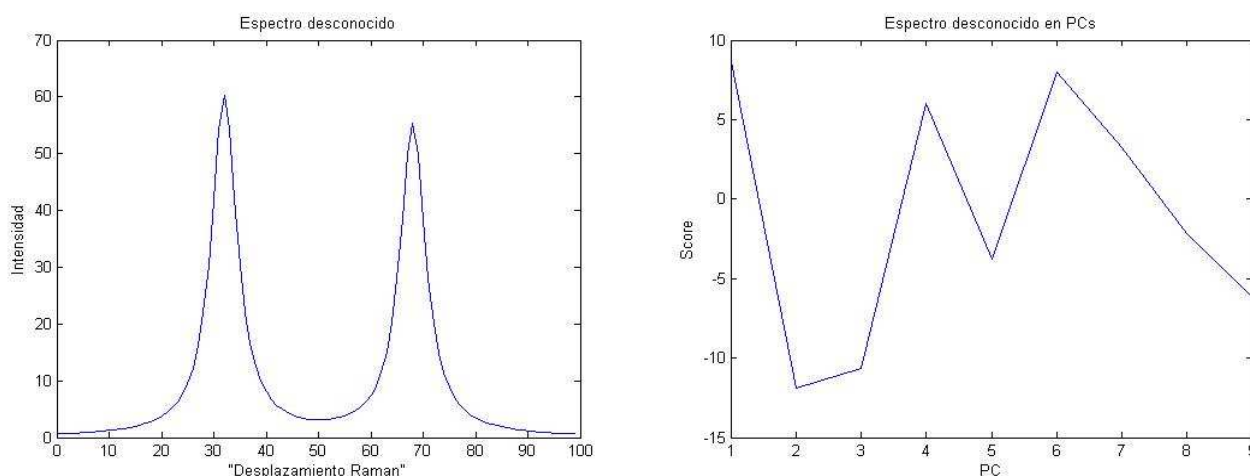


Fig. 4.9 a) y b): Espectro desconocido en el espacio original (a la izquierda) y en el espacio de los PCs (a la derecha)

Con el fin de analizar de forma gráfica la cercanía entre espectros, dado la alta varianza de las variables originales que recogen los tres primeros PCs, se procede a representar el espectro desconocido y los espectros de la librería de referencia en el espacio tridimensional formado por PC1, PC2 y PC3:

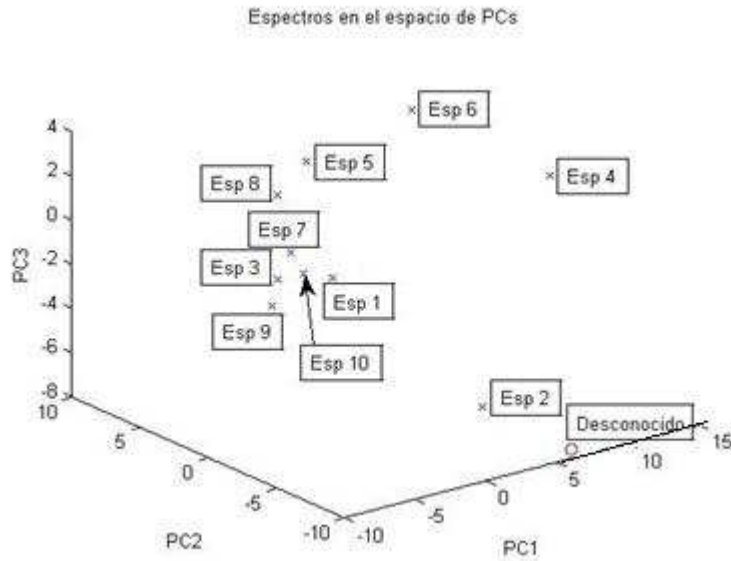


Fig. 4.10: Espectro desconocido y espectros de la librería en PC1-PC2-PC3

Una vez que el espectro desconocido ha sido proyectado en el espacio de los PCs generado por el ACP aplicado sobre la librería de referencia, se calcula la Distancia Euclídea entre dicho espectro desconocido y los espectros patrones, recogida en la *Tabla 4.7*:

Espectro	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10
Espectro desconocido	22,612	14,766	25,943	22,506	27,581	24,719	26,478	25,861	26,303	23,986

Tabla 4.7: Distancia Euclídea entre el espectro desconocido y la librería de referencia

Como puede verse, la distancia entre el espectro desconocido y el espectro patrón i -ésimo, d_i , es mayor que la distancia mínima de la librería:

$$d_i > \min_{lib} = 8.8476u$$

con $i = 1, \dots, 10$.

Además, la distancia entre el espectro desconocido y el espectro i -ésimo es superior al valor de la menor distancia entre el espectro patrón i -ésimo y el resto de espectros patrones:

$$d_i > v_{\min_i}$$

Concretamente, la distancia mínima entre el espectro desconocido y la librería de referencia se da respecto al espectro patrón 2, siendo igual a 14.766u, mientras la

mínima distancia de este patrón con el resto es 13.6434u. No cumpliéndose ninguno de los dos requisitos para la identificación espectral, el sistema no establece reconocimiento ya que no se obtiene ningún espectro candidato.

4. 3. Estudio de los efectos sobre la identificación de distintos problemas prácticos. Límite para la identificación

En este apartado se desarrolla el estudio de la influencia sobre la identificación de un espectro (a partir del ACP y los criterios expuestos basados en la Distancia Euclídea) cuando éste tiene diferentes problemas prácticos, como son:

- Ruido en el espectro
- Estrechamiento de las Bandas Raman
- Bandas Raman secundarias enmascaradas
- Variación de la intensidad relativa de Bandas Raman
- Desplazamiento de las Bandas Raman
- Bandas Raman “adicionales”

De este estudio, se deducen unos límites para que el sistema propuesto reconozca un espectro correctamente en cada caso.

4. 3. 1. Influencia del ruido. SNR mínima para la identificación

Uno de los ruidos más importantes en Espectroscopía Raman es el ruido *shot*, puesto que aparece ligado a la propia naturaleza aleatoria de la luz. La influencia de dicho ruido es perjudicial para la identificación espectral, ya que modifica la representación original del espectro, lo cual también modifica su expresión en el espacio de los PCs.

Para la simulación del ruido, se ha utilizado la función *random* de *Matlab*, de manera que un espectro con ruido se simula añadiendo al espectro “ideal” compuesto por lorentzianas (como se ha descrito en apartados anteriores) el ruido de una amplitud determinada. Cuando a un espectro de la librería se le añade ruido se observa que la

expresión del espectro en el espacio de los PCs es diferente, esto es, el valor de los *scores* varía.

Con el fin de observar esta variación, se simula un espectro tomando el espectro patrón 2 y se le añade ruido de amplitud igual al 5% de la amplitud de la banda fundamental del espectro, como esta es de $57u.a.$, la amplitud del ruido añadido es de $2,9u.a.$ Se llevan a cabo 100 realizaciones de espectros diferentes y para cada uno de ellos se realiza el Análisis de Componentes Principales. En la figura *Fig. 4.11* se muestran los *scores* para cada uno de ellos.

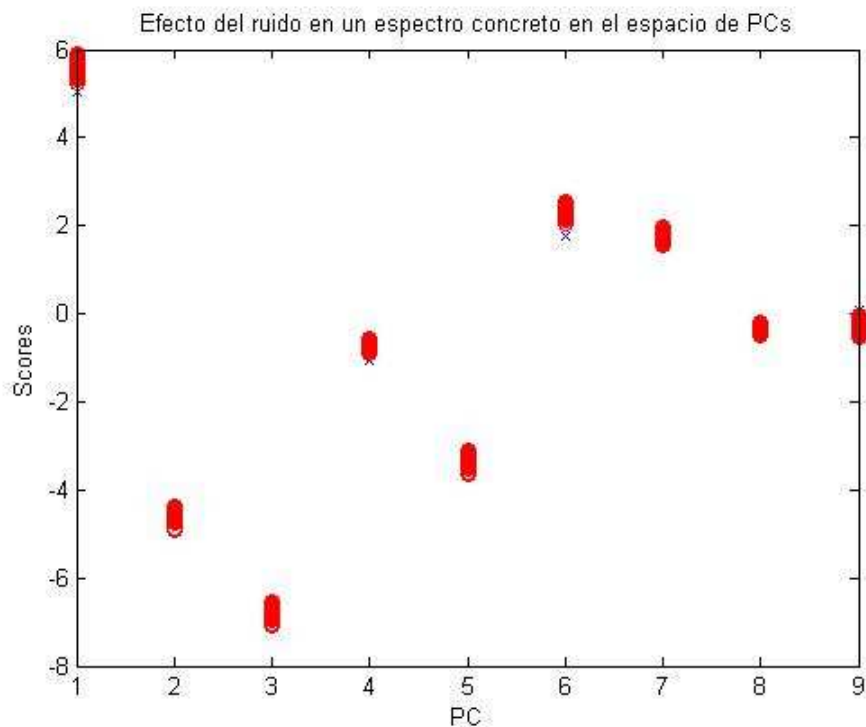


Fig. 4.11: Representación del Espectro 2 expresado en PCs y 100 realizaciones de ruido (amplitud del ruido del 5% respecto a la fundamental)

En la figura *Fig. 4.12* se han representado en PC1-PC2-PC3 los espectros de la librería de referencia y los 100 espectros a los que, a partir del espectro patrón 2, se les ha añadido ruido.

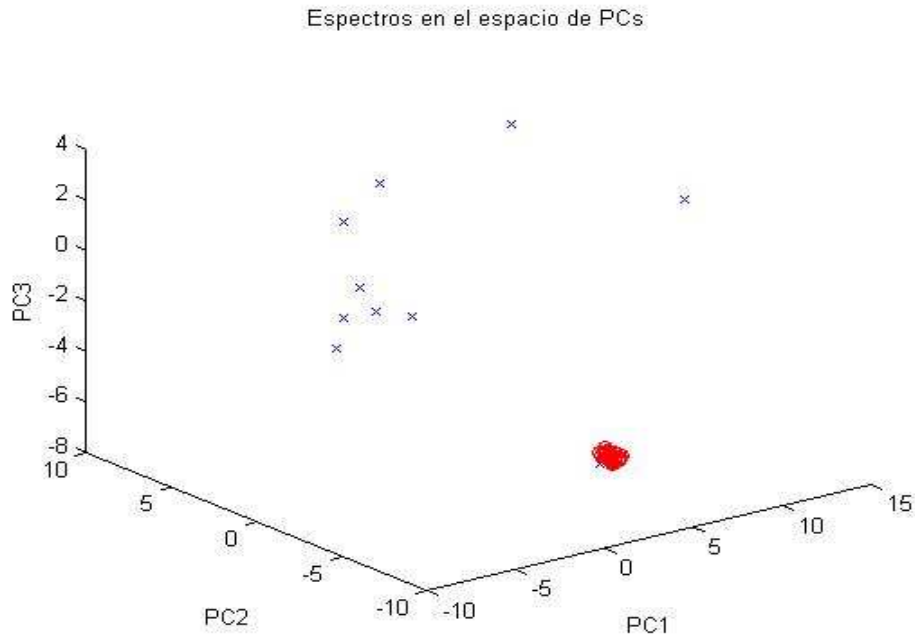


Fig. 4.12: Espectros con ruido a partir del Espectro 2 (amplitud del ruido del 5% respecto a la fundamental) y resto de espectros de la librería en PC1-PC2-PC3

Siguiendo el mismo proceso, en las figuras Fig. 4.13 y Fig. 4.14 se representan el espectro patrón 7 y 100 realizaciones del mismo con ruido. Como la amplitud de la banda fundamental del Espectro 7 es de $35u.a.$, la amplitud del ruido añadido es de $1,75u.a.$

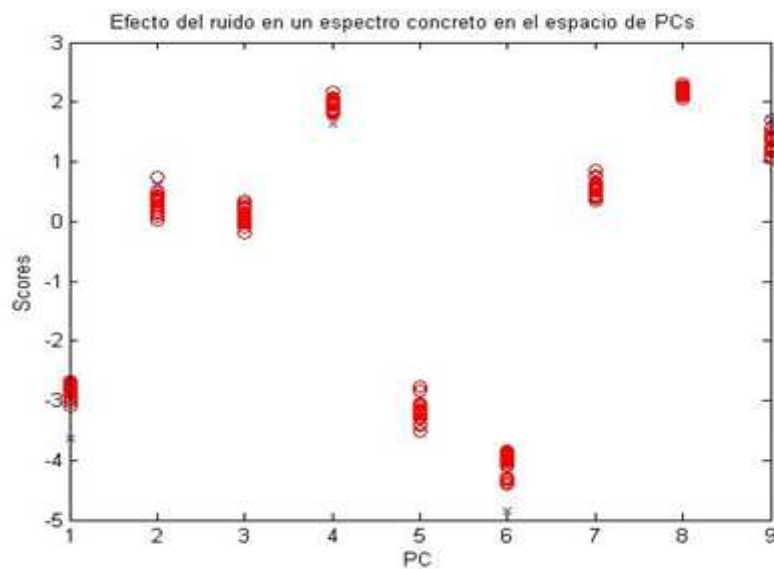


Fig. 4.13: Representación del Espectro 7 expresado en PCs y 100 realizaciones de ruido (amplitud del ruido del 5% respecto a la fundamental)

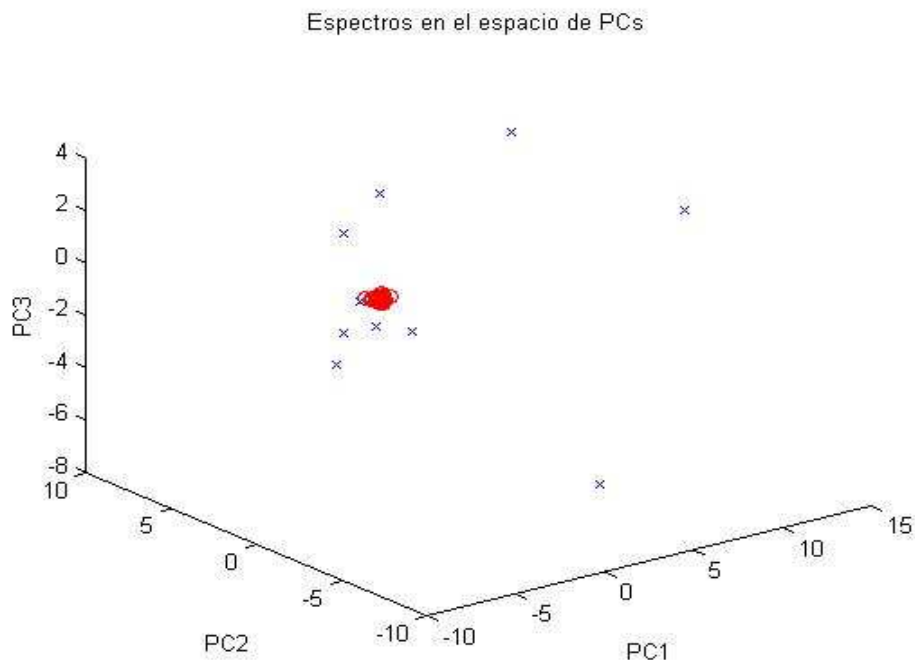


Fig. 4.14: Espectros con ruido a partir del Espectro 7 (amplitud del ruido del 5% respecto a la fundamental) y resto de espectros de la librería en PC1-PC2-PC3

Se puede observar que cuando la amplitud del ruido añadido es de un 5% respecto a la amplitud de la banda fundamental, la variación en PCs se puede acotar en 1 unidad.

Se ha realizado dicho proceso sobre todos los espectros de la librería de referencia, y a partir de los resultados obtenidos, se establece que la variación en los *scores* ocasionado por el ruido está acotada por:

$$cota = M.D. \cdot (\%ruido)$$

donde M. D. es el Margen Dinámico o excursión total de los *scores* de la librería generadora del espacio de los PCs.

Para el caso de la librería simulada, atendiendo a los valores recogidos en la *Tabla 4.2* y la representación de la *Fig. 4.4*, los *scores* toman valores entre -7.0044 y 10.245. Por lo tanto:

$$M.D. = 10.245 - (-7.0044) = 17.25$$

Siendo $\%ruido=5\%$,

$$cota = M.D. \cdot \%ruido = 17.25 \cdot 5\% = 0.86 \approx 1u$$

Para verificar la cota obtenida, se lleva a cabo el análisis inverso, es decir, una vez obtenida la expresión en PCs de los espectros patrones, se les suma un valor, emulando la variación en los *scores* que tendría el efecto del ruido.

Así, atendiendo a la definición de Distancia Euclídea resulta que:

$$dist^2(desconocido, patrón) = \sum_{i=1}^k (score_i^{patrón} - score_i^{desconocido} \pm cota)^2$$

donde k es el número de PCs que contemplan el 100% de la varianza, y **cota** es el valor (máximo) que pueden desplazarse los *scores* por efecto del ruido.

En el caso en que el espectro desconocido coincida con el espectro patrón:

$$dist^2(desconocido, patrón) = \sum_{i=1}^k (\pm cota)^2 = k \cdot cota^2$$

Para considerar que el espectro desconocido *puede ser* candidato, se parte de la siguiente hipótesis:

$$dist(desconocido, patrón) < min_lib$$

donde min_lib es la Distancia Euclídea mínima entre espectros de la librería.

Así, ahora:

$$\begin{aligned} \sqrt{k \cdot cota^2} &= min_lib \\ cota &= \pm \frac{min_lib}{\sqrt{k}} \end{aligned}$$

de donde, considerando la definición de la cota, se obtiene que:

$$(\%ruido) = \pm \frac{min_lib}{\sqrt{k} \cdot M.D.} = 17.0968\% = 17.1\%$$

Para el caso de la librería simulada se obtiene que:

$$cota = \pm \frac{min_lib}{\sqrt{k}} = \pm \frac{8.8476}{\sqrt{9}} = \pm 2,9492$$

lo que implica que para su correcta identificación, un espectro desconocido que se corresponde con un espectro patrón de la librería, puede tener un ruido de amplitud 17,1% respecto a la amplitud de la banda fundamental.

No obstante, no todos los *scores* se desplazan en un mismo sentido respecto al valor del espectro correspondiente original, como puede verse en las figuras *Fig. 4.11* y *Fig. 4.13*. Es decir, el movimiento de los *scores* es aleatorio, igual que el propio ruido. Así, si se analiza el resultado hallado de forma teórica para la cota,

$$cota = \pm 2,9492$$

se observa que el signo \pm indica que el valor numérico considera el desplazamiento de los PCs en un solo sentido. Puesto que el desplazamiento puede darse en los dos sentidos, la cota real es el doble:

$$cota = 2 \cdot 2,9492 = 5,8984$$

imponiendo que el ruido límite que soporta el sistema es del $2 \cdot 17.1\% = 34.2\%$.

A continuación se describen los resultados obtenidos de identificación para el caso de espectros desconocidos correspondientes a un espectro patrón con el máximo ruido (34.2%). Se simulan 100 espectros con ruido para cada patrón y se les aplica la metodología de reconocimiento diseñada. La siguiente tabla recoge los espectros candidatos generados (entre paréntesis se indica el número de veces que aparece el espectro candidato correspondiente) y sus Factores de Confianza:

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 1 con ruido	1 ⁽⁹⁸⁾	[0.29%-48.82%]	10 ⁽⁴⁾	[0.14%-3.57%]		
Esp 2 con ruido	2 ⁽¹⁰⁰⁾	[25.93%-85.14%]	4 ⁽¹⁷⁾	[0.48%-16.81%]	6 ⁽³⁶⁾	[1.57%-10.42%]
Esp 3 con ruido	3 ⁽⁸⁶⁾	[12.44%-49.54%]				
Esp 4 con ruido	4 ⁽¹⁰⁰⁾	[35.21%-74.83%]	2 ⁽³²⁾	[0.27%-20.88%]	6 ⁽²⁷⁾	[0.63%-9.82%]
Esp 5 con ruido	5 ⁽⁸⁸⁾	[14.16%-45.05%]				
Esp 6 con ruido	6 ⁽¹⁰⁰⁾	[41.35%-76.94%]				
Esp 7 con ruido	7 ⁽⁸⁴⁾	[8.24%-38.97%]	8 ⁽²⁸⁾	[0%-20.61%]	1 ⁽¹¹⁾	[0.71%-7.36%]
Esp 8 con ruido	8 ⁽⁹⁹⁾	[21.71%-57.79%]	6 ⁽²⁾	[0.11%-1.48%]		
Esp 9 con ruido	9 ⁽⁹²⁾	[14.58%-48.36%]				
Esp 10 con ruido	10 ⁽⁹⁴⁾	[1.19%-51.37%]				

Tabla 4.8: Espectros candidatos y Factores de Confianza

A la vista de los resultados, se puede concluir que con esta cota máxima de ruido se obtiene siempre como candidato el espectro patrón que corresponde. Aunque pueden ser generados otros candidatos, en estos casos el Factor de Confianza es relativamente mayor en el candidato correcto respecto al resto. Por lo que la cota obtenida no es un límite para cuando el sistema puede equivocarse resolviendo candidatos erróneos con mayor Factor de Confianza, sino que hace referencia a un límite para cuando el espectro desconocido deja de ser identificado. Cabe destacar que si no se obtiene el candidato correcto tampoco se obtienen otros candidatos, lo cual es equivalente a considerar que el sistema no se equivoca.

Para determinadas realizaciones, en el caso del “Espectro 7 con ruido” se obtiene un segundo candidato, el Espectro 8, con un Factor de Confianza comparable al del Espectro 7, aun siendo siempre superior el de este último espectro. Este hecho se explica al observar los espectros en su expresión original, puesto que son relativamente parecidos. Al haberse enmascarado la Banda Secundaria del Espectro 7 por el ruido, es sólo la Banda Fundamental la que permite la identificación, y entre la del Espectro 7 y la del Espectro 8 existe un desplazamiento inferior a 10cm^{-1} (ver *Fig. 4.15*).

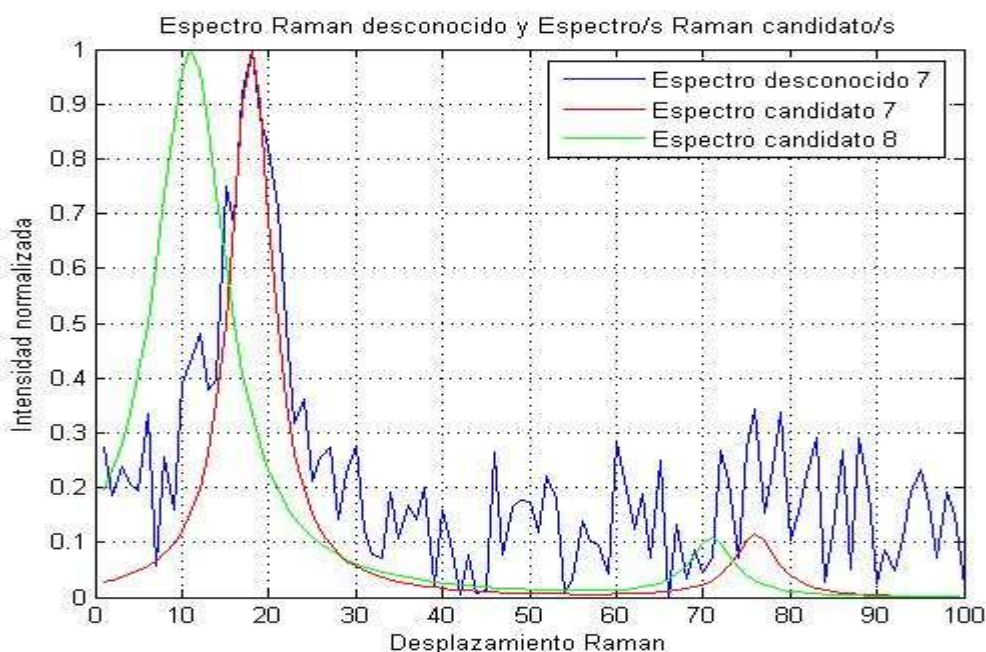


Fig. 4.15: Espectro desconocido (Espectro 7) y espectros candidatos

A continuación, se adjunta una figura donde se han representado los espectros patrones en el espacio formado por $PC1-PC2-PC3$, 100 realizaciones de cada uno de los espectros de la librería al añadirles ruido de amplitud un 34.2% la amplitud de la banda fundamental del espectro, y el límite de movimiento de los *scores* por efecto del ruido asociado a la cota hallada, generando un hexaedro regular de 5,8984u de arista centrado en la posición de los espectros patrones.

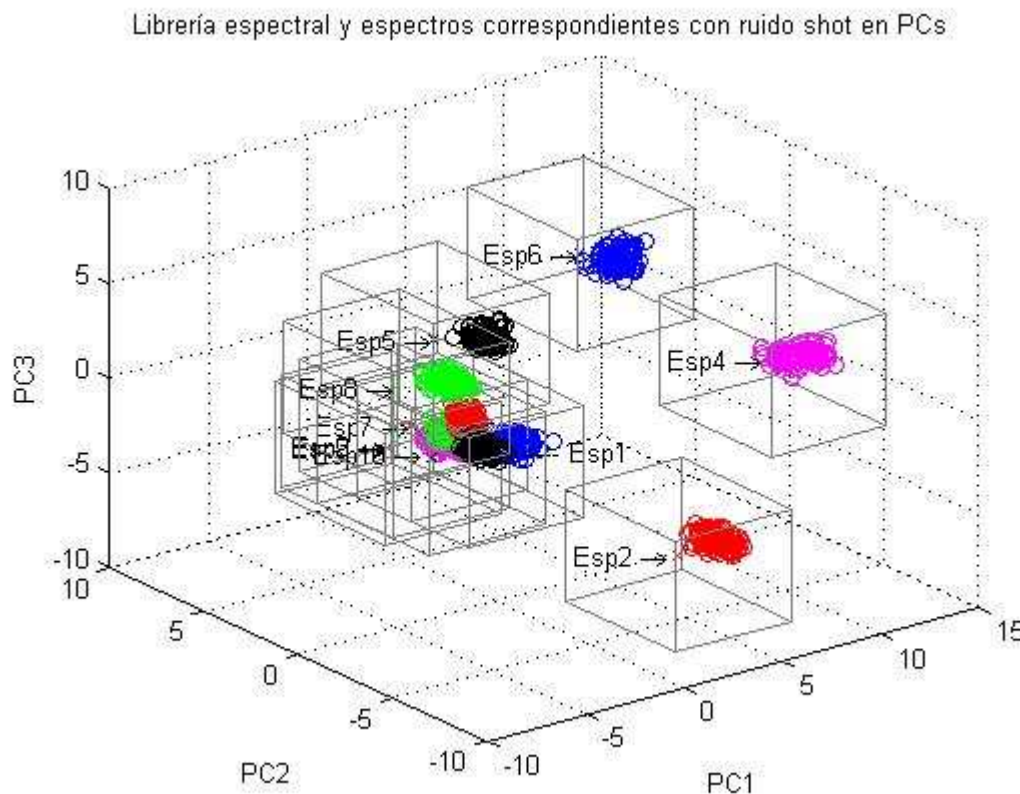
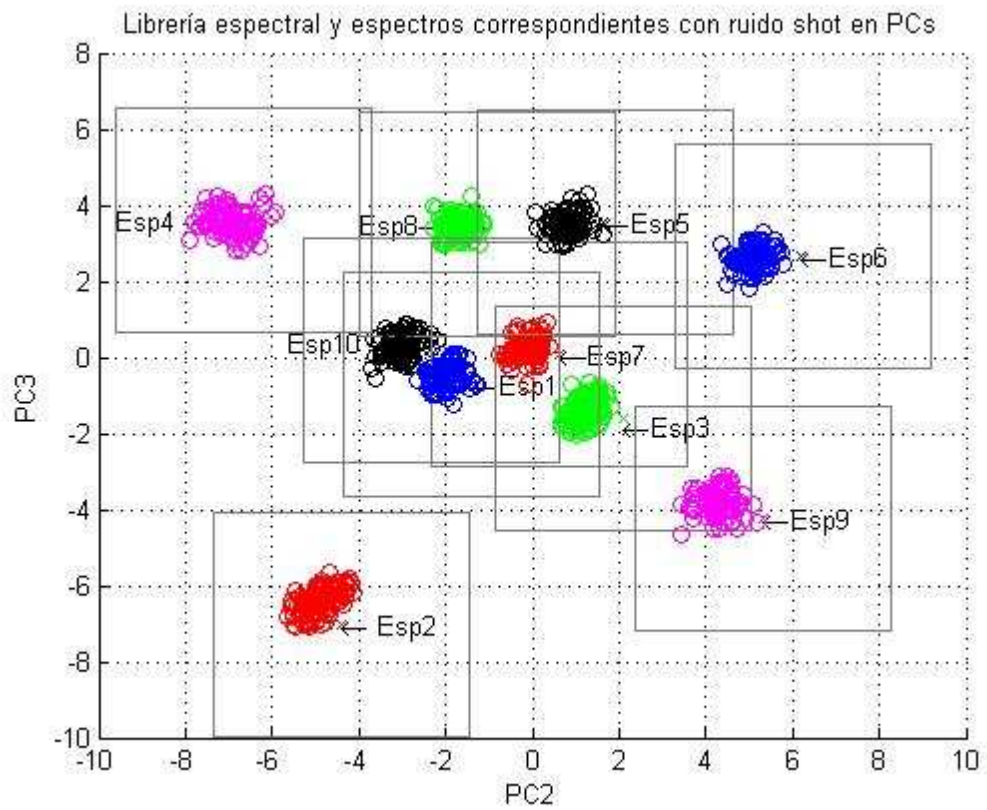
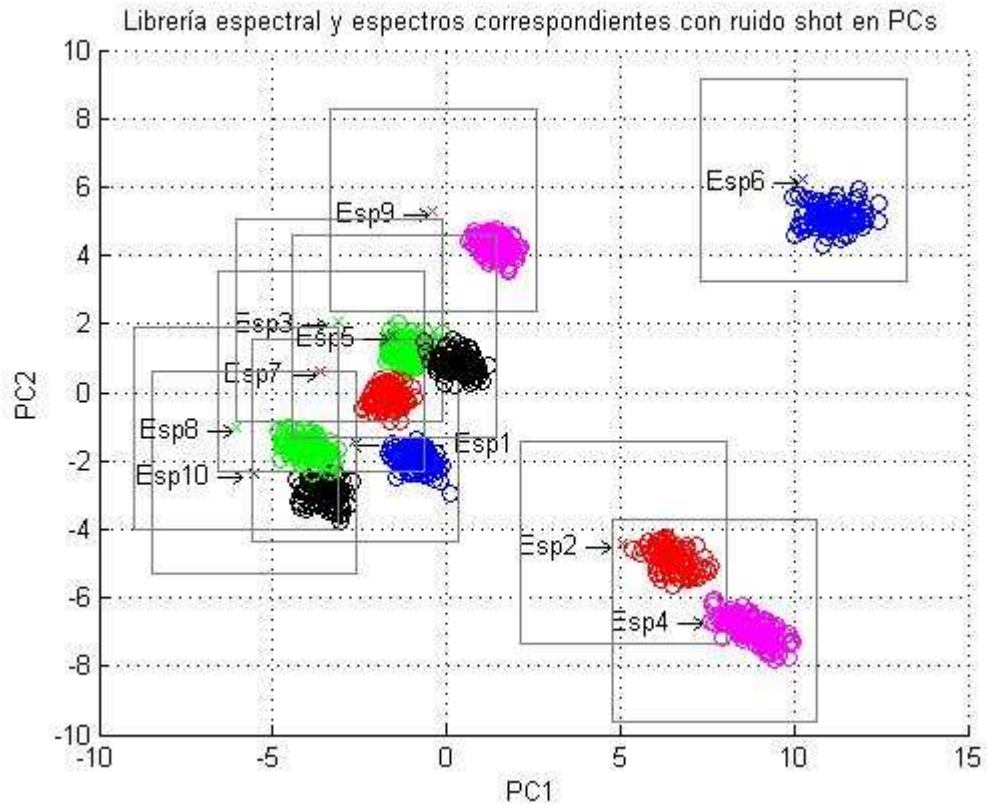


Fig. 4.16: Librería espectral de referencia, 100 realizaciones de los espectros patrones con ruido y límite del movimiento de los scores asociado a la cota hallada debido al ruido. Representación en $PC1-PC2-PC3$

Para una mejor visualización del movimiento de los *scores* y poder así verificar que éstos quedan siempre dentro del límite establecido por la cota hallada, se adjuntan las proyecciones de la anterior figura en $PC1-PC2$, en $PC2-PC3$ y en $PC1-PC3$:



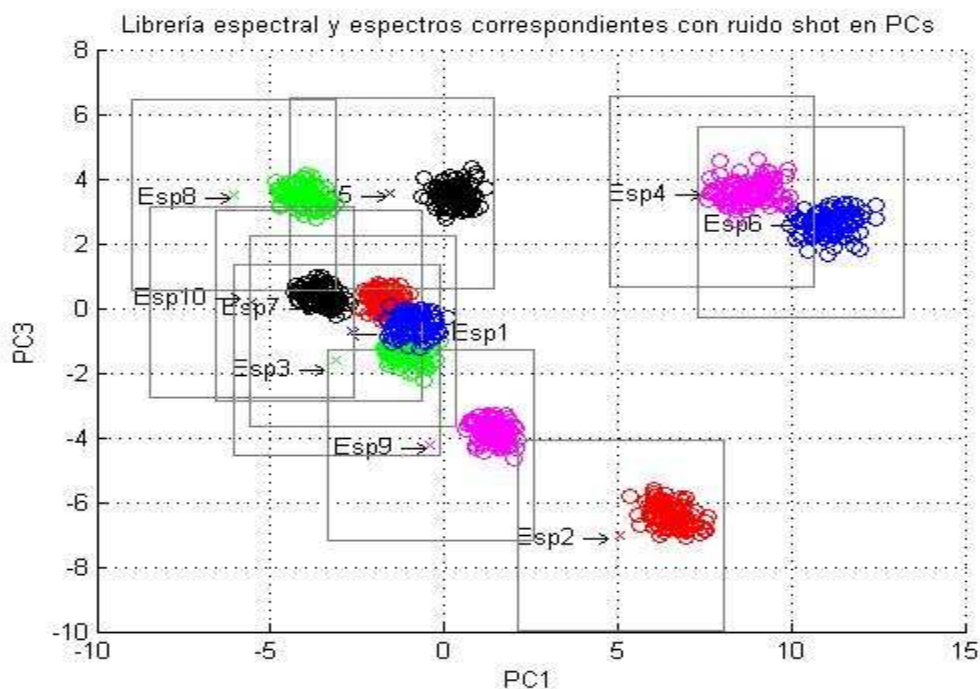


Fig. 4.17: Proyecciones en PC1-PC2, en PC2-PC3 y en PC1-PC3

Una importante métrica para describir la calidad del Espectro Raman es la relación señal a ruido (SNR). Para una determinada medida, la SNR se define formalmente como la relación entre la potencia de la señal (información deseada) y la potencia del ruido (señal no deseada):

$$SNR = \frac{P_{señal}}{P_{ruido}}$$

siendo P la potencia media. Dicha medida se ofrece comúnmente en dBs.

Basado en la SNR , se define un *Parámetro de Calidad*, Q_{dB} , que pondera la calidad del Espectro Raman en función del máximo valor que toma, el de la Banda Fundamental, y la distorsión introducida por el ruido^[17]. Así, para un espectro determinado, E_i , y el ruido asociado a éste, n_{E_i} , se tiene que:

$$Q_{dB}^i = 10 \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right)$$

donde $\max(E_i)$ representa el valor de la Banda Fundamental del espectro i -ésimo y $\sigma_{n_{E_i}}$ la desviación típica del ruido asociado al espectro i -ésimo. En este estudio teórico, para cada espectro E_i , el ruido está definido como:

$$n_{E_i} = A_{n_{E_i}} \cdot \max(E_i) \cdot random$$

donde $A_{n_{E_i}}$ es el factor de escala de la amplitud de la banda fundamental del espectro E_i para generar la amplitud del ruido asociado a dicho espectro, n_{E_i} . Asimismo, y teniendo en cuenta que,

$$\%ruido_{max} = 2 \cdot \frac{min_lib}{\sqrt{k} \cdot M.D.} = \frac{2 \cdot 8.8476}{\sqrt{9} \cdot (10.245 - (-7.0044))} = 34.2\%$$

se establece un umbral para la identificación,

$$Q_{dB}^{umbral} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{A_{n_{E_i}}=34.2\%B.F.}$$

Y, normalizando:

$$Q_{dB}^{umbral} = \frac{1}{\sigma\{0.342 \cdot random\}} = 10.47dB$$

De esta forma,

*Si $Q_{dB}^{desconocido} > Q_{dB}^{umbral} \rightarrow$ El espectro **puede ser identificado***

*Si $Q_{dB}^{desconocido} < Q_{dB}^{umbral} \rightarrow$ El espectro **no es identificado***

Puesto que se ha añadido ruido de amplitud límite para la identificación, pueden darse casos en los que no surjan candidatos. Por ejemplo, en 8 realizaciones del “Espectro 9 con ruido” (según la *Tabla 4.8*), la componente de ruido ha enmascarado las Bandas Secundarias del espectro original, ya que la amplitud del ruido es mayor que la de dichas bandas, como puede observarse en la figura *Fig. 4.18*, y no se han obtenido espectros candidatos:

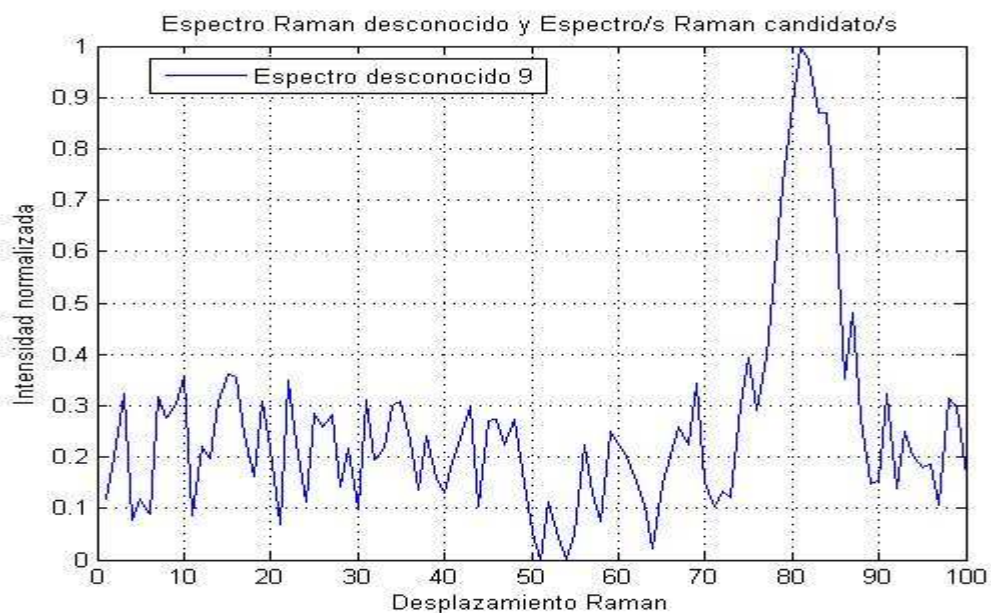


Fig. 4.18: Espectro desconocido (espectro patrón 9) sin candidatos

Si se procede a calcular el Parámetro de Calidad de una de las realizaciones en las que no se han obtenido espectros candidatos se tiene que,

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{nE_i}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.098} \right) = 10.098dB$$

y, como dicho valor es inferior al valor umbral obtenido, $Q_{dB}^{umbral} = 10.47dB$, el sistema no reconoce al espectro, es decir,

$$Q_{dB}^{desconocido} < Q_{dB}^{umbral} \rightarrow \text{El espectro **no** es identificado}$$

En los casos en los que sí se produce el reconocimiento espectral, el Parámetro de Calidad asociado es siempre superior al Q_{dB}^{umbral} .

En un caso experimental, en cambio, no se puede obtener matemáticamente la desviación típica del ruido, puesto que éste es intrínseco al espectro medido. Es por ello que, para obtener el Q_{dB} de un espectro experimental, se propone calcular la desviación típica del ruido sobre una zona del espectro que no contenga información útil, es decir, donde no existan Bandas Raman (ver Fig. 4.19).

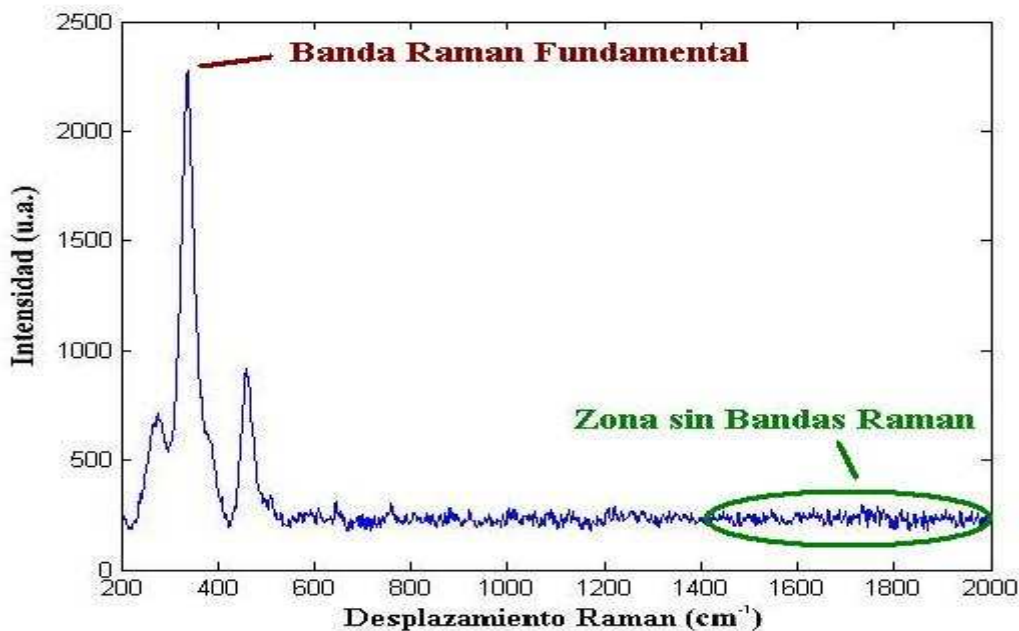


Fig. 4.19: Banda Fundamental y zona sin Bandas Raman

4.3.2. Influencia de la anchura de banda de las Bandas Raman

Durante el proceso de pretratado espectral, la anchura de las Bandas Raman puede verse reducida. La influencia de dicho efecto es perjudicial para la identificación espectral, ya que la expresión del espectro se ve alterada, modificando también su expresión en el espacio de los PCs. Así, puede darse el caso en el que un sistema de reconocimiento espectral basado en el ACP no llegue a identificar un espectro desconocido afectado por el estrechamiento de sus bandas.

Para analizar el efecto de la reducción de la anchura de las bandas en los espectros se ha simulado como espectros desconocidos los espectros patrones de la librería reduciendo la anchura de sus bandas hasta tener sólo rayas espectrales (es decir, $BW = 0$). Como ejemplo, en la figura *Fig. 4.20* se muestra esta reducción para el Espectro 4.

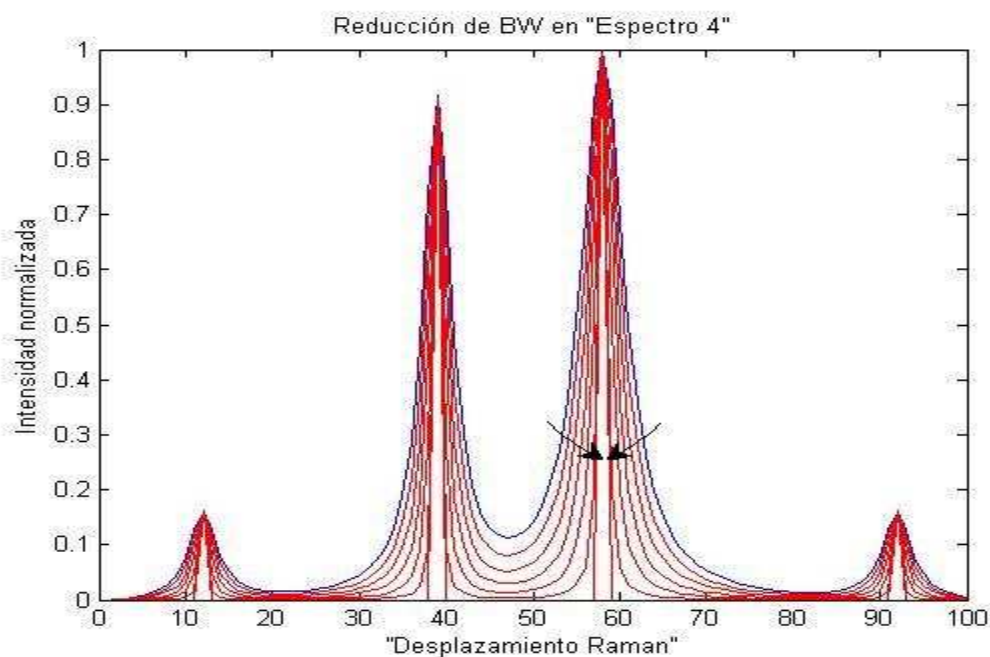


Fig. 4.20: Efecto de la reducción de la anchura de bandas en el espectro patrón 4

Esta reducción de la anchura de las bandas de un espectro modifica también su expresión en el espacio de PCs. En la figura *Fig. 4.21* puede verse, en el espacio tridimensional representado por PC1-PC2-PC3, el efecto que tiene sobre la librería espectral de referencia el reducir la anchura de sus bandas.

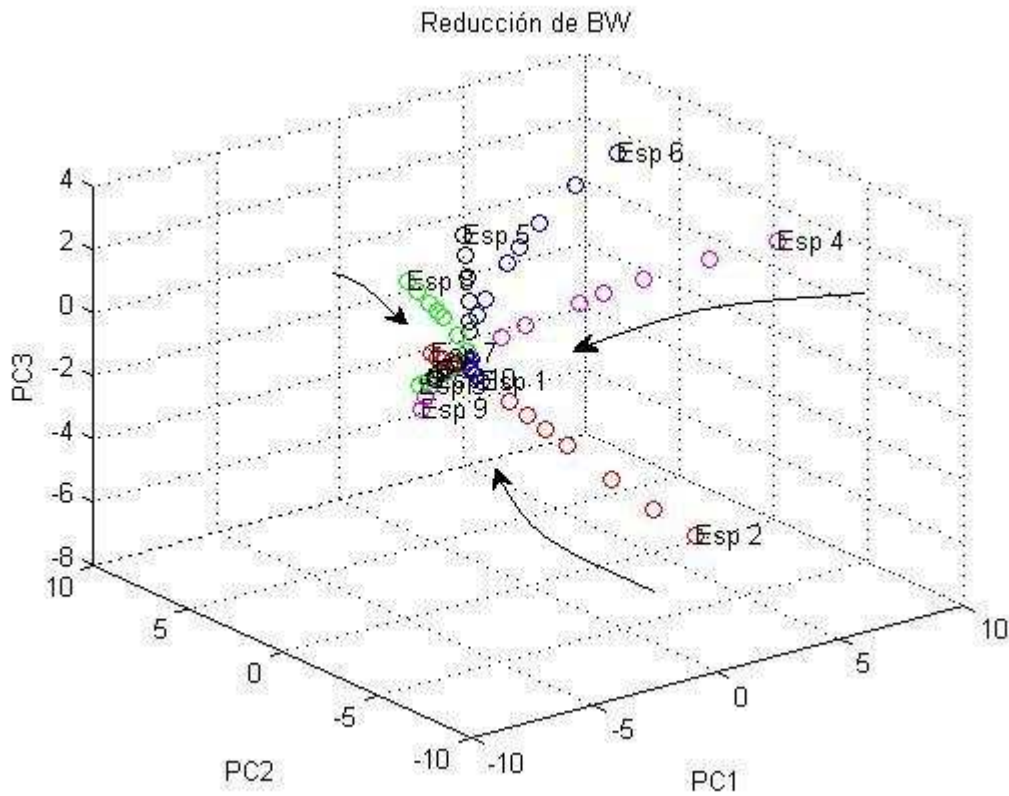
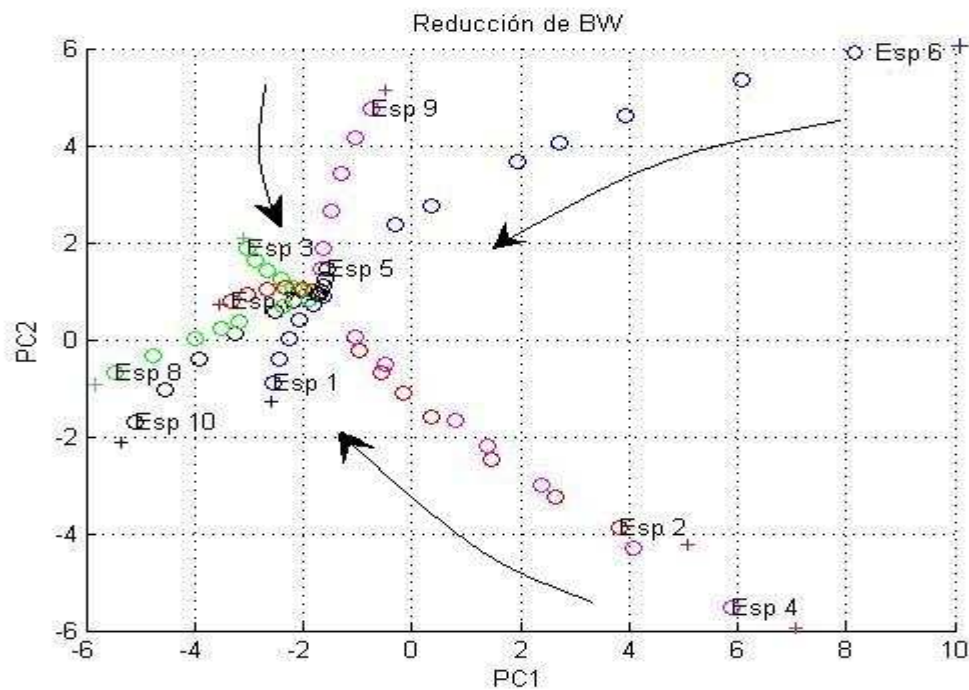


Fig. 4.21: Librería espectral en PC1-PC2-PC3 y estrechamiento de sus bandas

Para una mejor visualización del efecto, las siguientes figuras recogen las proyecciones en PC1-PC2, en PC2-PC3 y en PC1-PC3.



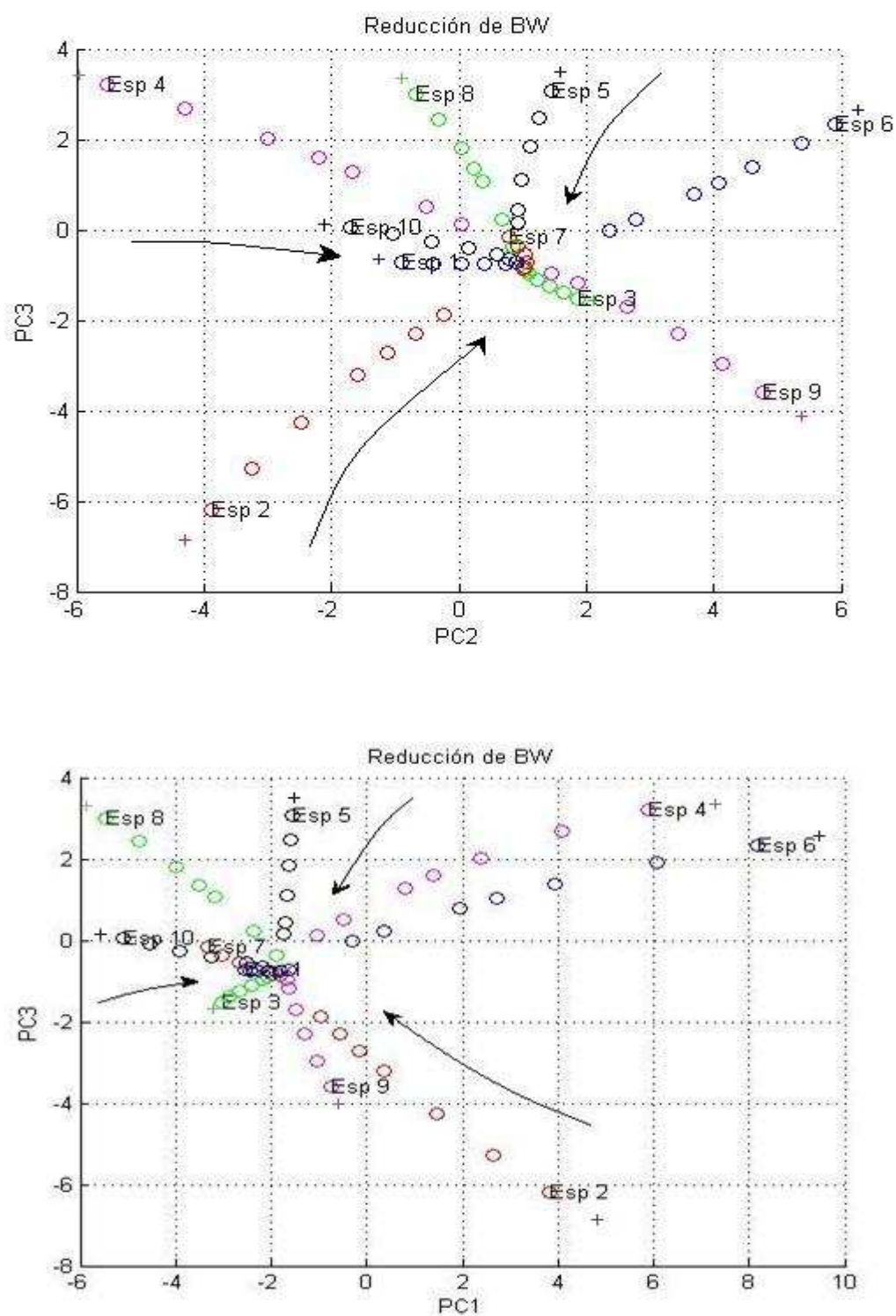


Fig. 4.22: Reducción del ancho de banda de los espectros de la librería de referencia

A la vista de las gráficas anteriores, todos los espectros, al reducir la anchura de sus bandas, tienden a desplazarse a un punto determinado. Concretamente, dicho punto en PCs es el que se corresponde con el espectro nulo, es decir, un espectro cuya expresión en el espacio original es una línea recta que pasa por cero y de pendiente nula. El valor al que tienden a desplazarse los espectros en PCs no es cero (o el origen de coordenadas) puesto que el espectro nulo también es estandarizado.

En la siguiente figura puede verse la representación en PC1-PC2 del espectro nulo (señalado con una X) y los espectros de la librería al estrecharse la anchura de sus bandas:

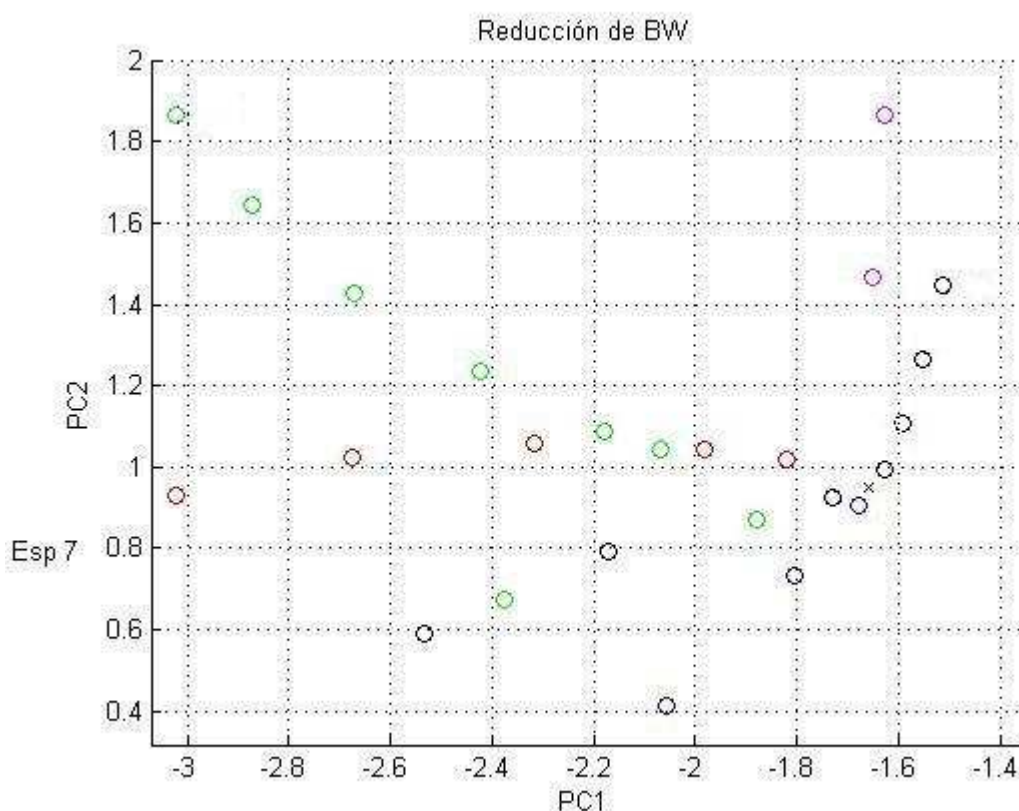


Fig. 4.23: Reducción de la anchura de las bandas de los espectros patrones y espectro nulo (correspondiente a una ampliación de la figura Fig. 4.22 en PC1-PC2)

A continuación se realiza el análisis de la identificación cuando se produce una reducción de la anchura de las bandas en un espectro. Como ejemplo, se muestran los resultados para el Espectro 4 y para el Espectro 6, marcándose en tonalidad verdosa los valores de Distancia Euclídea que cumplen alguno de los criterios para la identificación

en las tablas *Tabla 4.9* y *Tabla 4.10*. Para ver los resultados de los demás espectros de la librería de referencia, consultar el *Anexo A1*.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	COS2
Esp 4 BW'=5/6·BW	13,7	13,04	14,274	2,2796	12,656	15,134	14,153	15,934	14,94	15,069	0,9769
Esp 4 BW'=4/6·BW	12,274	12,616	12,421	4,5732	10,829	14,979	12,268	14,481	13,162	13,623	0,8815
Esp 4 BW'=3/6·BW	10,959	12,358	10,625	6,8543	9,1715	14,949	10,43	13,16	11,451	12,301	0,6905
Esp 4 BW'=2/6·BW	9,83	12,244	8,9818	9,0667	7,8546	15,012	8,7333	12,056	9,9063	11,182	0,4528
Esp 4 BW'=1/6·BW	9,0421	12,257	7,7358	10,996	7,115	15,136	7,4324	11,307	8,7557	10,411	0,2849
Esp 4 BW'=0·BW	8,7715	12,308	7,2667	11,888	6,9513	15,223	6,9383	11,053	8,3292	10,145	0,2367

Tabla 4.9: Distancia Euclídea para el Espectro 4 con estrechamiento de sus bandas

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	COS2
Esp 6 BW'=5/6·BW	14,867	15,14	15,137	14,565	14,327	2,3347	15,017	17,364	13,438	16,901	0,9735
Esp 6 BW'=4/6·BW	12,993	14,148	13,009	13,817	12,346	4,7845	12,796	15,538	11,523	14,969	0,8623
Esp 6 BW'=3/6·BW	11,206	13,285	10,89	13,272	10,44	7,3227	10,579	13,804	9,7447	13,113	0,6314
Esp 6 BW'=2/6·BW	9,7125	12,659	8,9807	13,014	8,8241	9,8392	8,5734	12,343	8,3597	11,524	0,3447
Esp 6 BW'=1/6·BW	8,8049	12,368	7,6639	13,051	7,8272	11,955	7,1869	11,411	7,6691	10,492	0,1762
Esp 6 BW'=0·BW	8,5512	12,323	7,2324	13,143	7,5472	12,847	6,7349	11,122	7,5471	10,167	0,1435

Tabla 4.10: Distancia Euclídea para el Espectro 6 con estrechamiento de sus bandas

Se observa que, a medida que se reduce el ancho de las bandas de los espectros, atendiendo al coseno cuadrado de las tablas anteriores (última columna), la calidad de la representación descende, alcanzando valores inferiores a 0.5 a partir de una reducción del ancho de banda de 2/6.

En las tablas *Tabla 4.11* y *Tabla 4.12* se adjuntan los espectros candidatos generados por el sistema y los respectivos Factores de Confianza asociados:

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 4 BW'=5/6·BW	4	85.16%	2	15.09%	6	1.46%
Esp 4 BW'=4/6·BW	4	70.22%	2	17.86%	6	2.47%
Esp 4 BW'=3/6·BW	4	55.37%	2	19.54%	6	2.67%
Esp 4 BW'=2/6·BW	5	20.2%	7	11.27%		
Esp 4 BW'=1/6·BW	5	29.89%	3	23.78%	7	26.77%
Esp 4 BW'=0·BW	7	31.64%	5	31.51%	3	28.41%

Tabla 4.11: Factores de Confianza para el Espectro 4 con estrechamiento de sus bandas

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 6 BW'=5/6·BW	6	84.8%				
Esp 6 BW'=4/6·BW	6	68.85%				
Esp 6 BW'=3/6·BW	6	52.32%				
Esp 6 BW'=2/6·BW	9	17.63%	5	15.53%	7	13.06%
Esp 6 BW'=1/6·BW	7	29.19%	3	24.49%	9	24.44%
Esp 6 BW'=0·BW	7	33.64%	3	28.74%	9	25.64%

Tabla 4.12: Factores de Confianza para el Espectro 6 con estrechamiento de sus bandas

A raíz de los resultados, se puede ver que, llegados a un determinado estrechamiento, deja de obtenerse el candidato que corresponde. Este hecho marca un límite para el reconocimiento espectral. Por ello, a partir del resultado obtenido, se define un límite para el estrechamiento de Bandas Raman hasta el cual el sistema ofrece el candidato correcto, identificando al espectro con reducción en la anchura de sus bandas con el espectro patrón correspondiente. Dicho límite para la librería simulada corresponde a:

$$BW' = \frac{2}{5}BW = 40\%BW$$

siendo BW' el del espectro patrón con estrechamiento en sus bandas.

A continuación se adjuntan los resultados para el caso límite sobre la librería de referencia:

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	COS2
Esp 1 BW'=2/5·BW	3,9205	13,228	9,0357	14,153	9,694	15,591	7,3307	11,796	9,5911	9,4557	0,6042
Esp 2 BW'=2/5·BW	9,8623	6,9493	8,9677	12,229	9,6826	15,258	8,9785	12,298	9,7103	11,033	0,4929
Esp 3 BW'=2/5·BW	9,7434	13,171	3,074	14,499	8,4695	16,085	8,7651	11,501	9,0205	9,627	0,644
Esp 4 BW'=2/5·BW	10,252	12,273	9,6096	8,197	8,3255	14,978	9,3836	12,464	10,494	11,598	0,5463
Esp 5 BW'=2/5·BW	10,305	13,494	8,2085	13,502	3,3043	15,746	8,5097	10,995	9,6482	11,191	0,638
Esp 6 BW'=2/5·BW	10,256	12,872	9,6985	13,079	9,4156	8.8473	9,3285	12,879	8,8525	12,11	0,455
Esp 7 BW'=2/5·BW	8,969	13,272	9,3585	14,723	9,3128	16,077	3,0383	11,067	9,3076	11,131	0,7336
Esp 8 BW'=2/5·BW	9,7598	13,192	8,6873	14,386	6,6491	16,321	7,1861	6,2428	9,6562	9,3879	0,5288
Esp 9 BW'=2/5·BW	9,7601	13,035	8,0832	14,539	8,8947	15,077	7,8623	11,744	4,148	11,097	0,6105
Esp 10 BW'=2/5·BW	7,6457	13,29	7,2015	14,387	8,9825	16,234	8,1451	10,025	9,3785	5,3112	0,4983

Tabla 4.13: Distancia Euclídea para los espectros patrones con estrechamiento límite

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 1 BW'=2/5·BW	1	55.69%	7	17.14%		
Esp 2 BW'=2/5·BW	2	49.07%				
Esp 3 BW'=2/5·BW	3	68.77%	5	13.95%	7	10.95%
Esp 4 BW'=2/5·BW	4	39.92%	5	38.9%		
Esp 5 BW'=2/5·BW	5	66.43%	3	16.6%	7	13.55%
Esp 6 BW'=2/5·BW	6	~0%				
Esp 7 BW'=2/5·BW	7	77.73%	2	2.72%		
Esp 8 BW'=2/5·BW	8	40.75%	5	36.89%	7	31.81%
Esp 9 BW'=2/5·BW	9	72.99%	7	48.81%	3	47.37%
Esp 10 BW'=2/5·BW	10	46.04%	3	26.84%	1	22.32%

Tabla 4.14: Factores de Confianza para los espectros patrones con estrechamiento límite

En el caso límite, el sistema funciona correctamente, obteniéndose el candidato correspondiente en cada caso con Factores de Confianza de entre el 39.92% y el 77.73%, aunque con una calidad en la representación relativamente baja, dada por el coseno cuadrado, con valores de entre el 45.5% y el 73.36%. Para el “Espectro 6 con $BW' = 2/5 \cdot BW$ ”, en cambio, se obtiene un candidato (el Espectro 6) con un Factor de Confianza muy bajo, puesto que la Distancia Euclídea, de valor 8.8473u, es muy cercana al valor mínimo de la librería, $min_lib = 8.8476u$ (primer criterio de identificación), lo que implica un Factor de Confianza próximo a cero. Además, no se aplica el segundo criterio de identificación, puesto que la calidad de la representación es relativamente baja, inferior a la cota establecida en los criterios de identificación:

$$COS2 = 0.455 < cota = \frac{min_lib}{min_6} = \frac{8.8476}{15.3585} = 0.5761$$

La siguiente figura muestra el Espectro 6 (en azul) y el Espectro 6 con ancho de banda límite (en rojo):

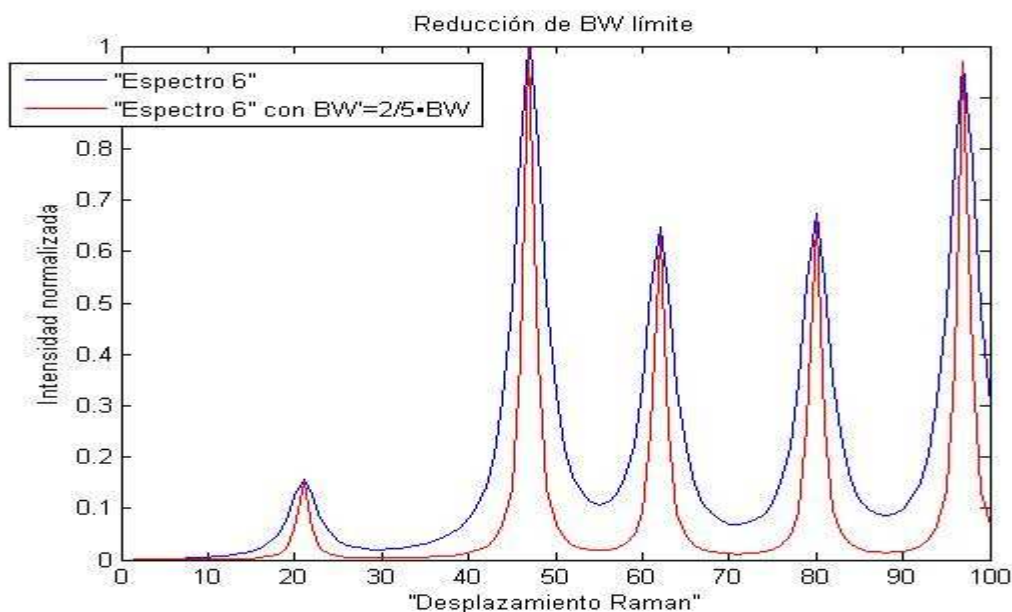


Fig. 4.24: Espectro 6 y Espectro 6 con estrechamiento máximo

Cabe destacar, también, que para el Espectro 8 con estrechamiento de bandas se generan varios candidatos, con unos Factores de Confianza del mismo orden, de igual forma que para el Espectro 4 con estrechamiento de bandas.

A la vista de los resultados, se establece el límite para el reconocimiento espectral que soporta el sistema implementado en un estrechamiento máximo de las bandas del 60%.

4. 3. 3. Influencia de la variación de las intensidades relativas de las Bandas Raman

Las intensidades relativas entre Bandas Raman pueden variar de un espectro a otro obtenido del mismo material, debido a diferentes causas, por ejemplo, cuando son medidos en distintos laboratorios, puesto que la intensidad depende, entre otros factores, de las características de los instrumentos de medida utilizados a la hora de llevar a cabo el análisis sobre cierta muestra. Como consecuencia, los *scores* de los espectros también varían respecto a los del patrón correspondiente, perjudicando la identificación espectral.

Con el fin de analizar este efecto, se estudian los resultados de la identificación del sistema variando la relación entre la amplitud de la Banda Fundamental (I_1) y la amplitud de la Banda Secundaria de mayor intensidad (I_2). Para ello, se define la diferencia entre amplitudes en un espectro patrón, ΔI , como

$$\Delta I \equiv I_1 - I_2$$

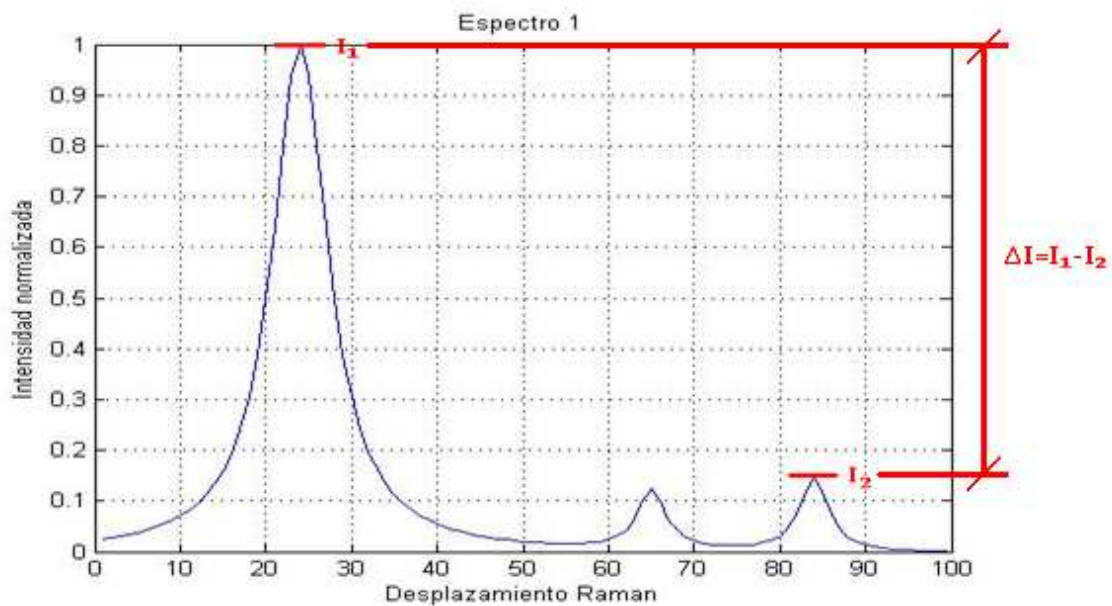


Fig. 4.25: Amplitudes I_1 e I_2 en el espectro patrón 1

Para cada espectro patrón se simula un nuevo espectro modificando la amplitud de la Banda Secundaria, I'_2 , como

$$I'_2 = I_2 + C \cdot \Delta I$$

con $0 \leq C \leq 1$, donde la igualdad $C = 0$ indicaría que la amplitud de la Banda Secundaria no ha sido variada, manteniendo la relación original entre I_1 e I_2 , y, en cambio, $C = 1$ correspondería al caso en que el espectro tuviese dos Bandas Fundamentales (caso muy improbable en la práctica).

Se obtiene una nueva relación entre amplitudes, $\Delta I'$:

$$\Delta I' \equiv I_1 - I'_2 = k \cdot \Delta I$$

donde k es el factor de proporcionalidad entre, ΔI , y, $\Delta I'$. Desarrollando dicha ecuación,

$$\begin{aligned} \Delta I' = I_1 - I'_2 &= I_1 - (I_2 + C \cdot \Delta I) = I_1 - I_2 - C \cdot \Delta I = \\ &= \Delta I - C \Delta I = (1 - C) \cdot \Delta I \end{aligned}$$

con lo que

$$k = (1 - C)$$

En primer lugar, se analiza el efecto sobre la identificación cuando aumenta la amplitud de las Bandas Secundarias (denotadas por B.S.). La *Tabla 4.15* recoge los resultados para el caso en que la amplitud de la Banda Secundaria aumenta generando una reducción de un cuarto en la relación entre I_1 e I_2 , es decir:

$$C = \frac{1}{4}, k = \frac{3}{4}: I'_2 = I_2 + \frac{1}{4} \Delta I \rightarrow \Delta I' = \frac{3}{4} \Delta I$$

Espectro	ΔI	$\Delta I'$	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Coseno cuadrado
Esp 1 B.S.↑	0,8529	0,6397	1	57.38%			0,7971
Esp 2 B.S.↑	0,6315	0,4736	2	80.51%			0,9399
Esp 3 B.S.↑	0,6875	0,5156	3	75.12%			0,9597
Esp 4 B.S.↑	0,0967	0,0725	4	88.21%			0,9918
Esp 5 B.S.↑	0,0967	0,0725	5	79.24%	4	11.15%	0,9864
Esp 6 B.S.↑	0,0307	0,0231	6	91.52%			0,9885
Esp 7 B.S.↑	0,8857	0,6642	7	81.49%	2	2.79%	0,8931
Esp 8 B.S.↑	0,8936	0,6702	8	64.57%			0,9075
Esp 9 B.S.↑	0,8387	0,6291	9	82.22%			0,9691
Esp 10 B.S.↑	0,4285	0,3214	10	79.60%			0,9662

Tabla 4.15: Candidatos y Factores de confianza. $C = \frac{1}{4}$

Puesto que los resultados de identificación son favorables en todos los casos, se sigue aumentando la Banda Secundaria. Los resultados de la *Tabla 4.16* hacen referencia al

caso en que la Banda Secundaria aumenta dando una reducción de un medio en la relación entre I_1 e I_2 :

$$C = k = \frac{1}{2}: I'_2 = I_2 + \frac{1}{2}\Delta I \rightarrow \Delta I' = \frac{1}{2}\Delta I$$

Espectro	ΔI	$\Delta I'$	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Coseno cuadrado
Esp 1 B.S.↑	0,8529	0,4264	1	14.48%			0,6122
Esp 2 B.S.↑	0,6315	0,3157	2	61.31%			0,8444
Esp 3 B.S.↑	0,6875	0,3437	3	50.16%	10	14.31%	0,8882
Esp 4 B.S.↑	0,0967	0,0483	4	76.31%			0,9686
Esp 5 B.S.↑	0,0967	0,0483	5	57.96%	4	28.03%	0,9544
Esp 6 B.S.↑	0,0307	0,0153	6	83.06%			0,9598
Esp 7 B.S.↑	0,8857	0,4428	7	62.88%	2	9.62%	0,7221
Esp 8 B.S.↑	0,8936	0,4468	8	29.01%			0,7714
Esp 9 B.S.↑	0,8387	0,4193	9	64.37%			0,8838
Esp 10 B.S.↑	0,4285	0,2142	10	59.44%			0,8836

Tabla 4.16: Candidatos y Factores de confianza. $C = \frac{1}{2}$

Atendiendo a los resultados, los espectros son identificados de forma correcta en este caso, pero para los espectros 1 y 8, el Factor de Confianza se ha visto sensiblemente reducido (primera y octava fila de la *Tabla 4.16*). Ello se debe a que las amplitudes de las Bandas Secundarias de dichos espectros son las menores de entre las amplitudes de las secundarias de la librería de referencia, y, por tanto, al modificar su amplitud, han aumentado más que en el resto de espectros, pareciéndose menos a sus respectivos patrones, al contrario que los espectros con elevadas amplitudes de sus secundarias, como el Espectro 6 (ver figura *Fig. 4.26*).

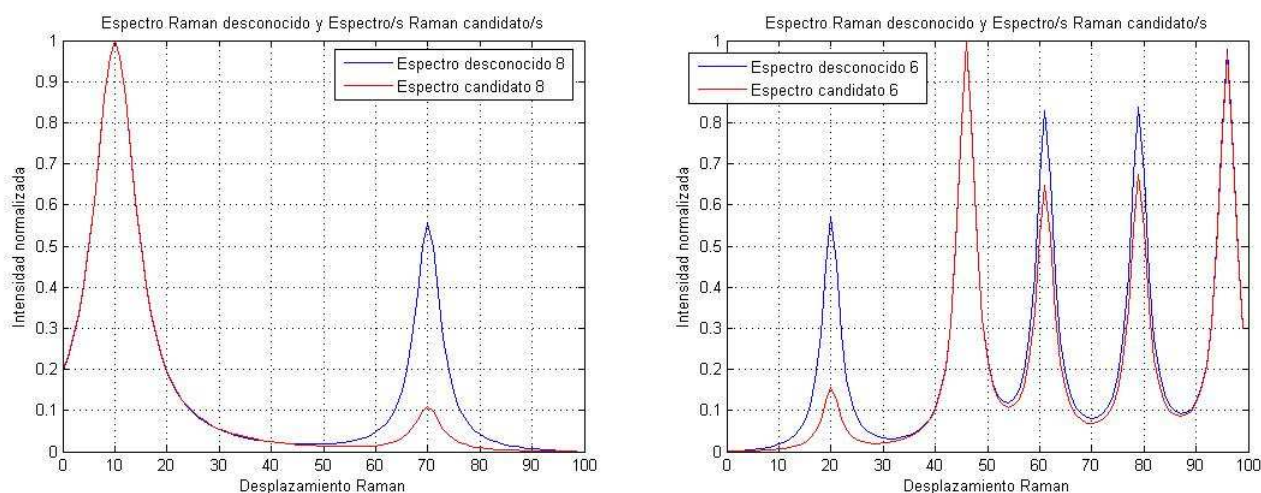


Fig. 4.26: Efecto de variar las intensidades relativas en el Espectro 8 (a la izquierda) y en el Espectro 6 (a la derecha)

La *Tabla 4.17* muestra los resultados cuando aumenta la Banda Secundaria, proporcionando una reducción de tres cuartos en la relación entre I_1 e I_2 , es decir:

$$C = \frac{3}{4}, k = \frac{1}{4} : I'_2 = I_2 + \frac{3}{4}\Delta I \rightarrow \Delta I' = \frac{1}{4}\Delta I$$

Espectro	ΔI	$\Delta I'$	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Coseno cuadrado
Esp 1 B.S.↑	0,8529	0,2132	-	-			0,5285
Esp 2 B.S.↑	0,6315	0,1578	2	42.37%			0,7646
Esp 3 B.S.↑	0,6875	0,1718	3	25.15%	10	14.43%	0,8286
Esp 4 B.S.↑	0,0967	0,0241	4	64.25%			0,9353
Esp 5 B.S.↑	0,0967	0,0241	4	42.91%	5	35.79%	0,9196
Esp 6 B.S.↑	0,0307	0,0076	6	74.61%			0,9221
Esp 7 B.S.↑	0,8857	0,2214	7	13.91%			0,6127
Esp 8 B.S.↑	0,8936	0,2234	-	-			0,6872
Esp 9 B.S.↑	0,8387	0,2096	9	46.43%	7	15.11%	0,7843
Esp 10 B.S.↑	0,4285	0,1071	10	39.24%			0,7963

Tabla 4.17: Candidatos y Factores de confianza. $C = \frac{3}{4}$

En este caso, los espectros 1 y 8 han dejado de ser identificados. Como caso límite, la *Tabla 4.18* recoge los resultados del caso teórico (y muy improbable en la práctica) cuando la amplitud de la Banda Secundaria es la misma que la de la Banda Fundamental, es decir:

$$C = 1, k = 0 : I'_2 = I_2 + 1 \cdot \Delta I = I_2 + (I_1 - I_2) = I_1 \rightarrow \Delta I' = 0$$

Espectro	ΔI	$\Delta I'$	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Coseno cuadrado
Esp 1 B.S.↑	0,8529	0	-	-			0,4881
Esp 2 B.S.↑	0,6315	0	2	23.71%			0,7051
Esp 3 B.S.↑	0,6875	0	-	-			0,7877
Esp 4 B.S.↑	0,0967	0	4	52.54%			0,9019
Esp 5 B.S.↑	0,0967	0	4	52.54%			0,9019
Esp 6 B.S.↑	0,0307	0	6	66.69%			0,8821
Esp 7 B.S.↑	0,8857	0	-	-			0,5541
Esp 8 B.S.↑	0,8936	0	-	-			0,6415
Esp 9 B.S.↑	0,8387	0	9	28.42%	7	17.25%	0,7036
Esp 10 B.S.↑	0,4285	0	-	-			0,4881

Tabla 4.18: Candidatos y Factores de confianza. $C = 1$

A la vista de los resultados obtenidos, se puede concluir que la identificación depende de la diferencia entre amplitudes del espectro patrón, es decir, cuanto mayor sea la diferencia entre I_1 e I_2 , al aumentar la amplitud de las Bandas Secundarias, el espectro es menos parecido al espectro patrón, modificando sus *scores*, pudiéndose dar el caso en

que el espectro no sea reconocido por el sistema implementado. En cambio, casos como el del Espectro 6 son menos sensibles al aumento de la amplitud de las Bandas Secundarias, puesto que se trata de un tipo de espectros con un número elevado de bandas y todas con amplitud cercana a la de la Banda Fundamental.

Una vez estudiado el caso en que aumente la amplitud de las Bandas Secundarias, se analiza el efecto sobre la identificación cuando su amplitud decrece. En este caso se modifica directamente el valor de la amplitud de la Banda Secundaria:

$$I'_2 = C' \cdot I_2$$

con $0 \leq C' \leq 1$, donde $C' = 1$ correspondería al caso en que la amplitud de la Banda Secundaria no ha sido variada, y $C' = 0$ indicaría que la amplitud de la Banda Secundaria es nula, caso límite analizado en el siguiente apartado, *Apartado 4.3.4*.

De esta forma, ahora:

$$\Delta I' = I_1 - I'_2 = I_1 - C' \cdot I_2$$

La *Tabla 4.19* muestra los resultados cuando la Banda Secundaria se ha reducido un factor $\frac{3}{4}$, es decir:

$$C' = \frac{3}{4} : I'_2 = \frac{3}{4} I_2$$

Espectro	ΔI	$\Delta I'$	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Coseno cuadrado
Esp 1 B.S.↓	0,8529	0,8896	1	94.13%	7	3.24%	0,9935
Esp 2 B.S.↓	0,6315	0,7237	2	89.81%			0,9782
Esp 3 B.S.↓	0,6875	0,7656	3	90.61%	5	6.46%	0,9926
Esp 4 B.S.↓	0,0967	0,3226	4	90.71%	2	17.44%	0,9723
Esp 5 B.S.↓	0,0967	0,3226	5	88.73%	3	61.04%	0,9857
Esp 6 B.S.↓	0,0307	0,2731	6	84.65%			0,9638
Esp 7 B.S.↓	0,8857	0,9144	7	96.91%	1	16.49%	0,9979
Esp 8 B.S.↓	0,8936	0,9203	8	95.79%			0,9983
Esp 9 B.S.↓	0,8387	0,8792	9	97.74%	6	0.31%	0,9991
Esp 10 B.S.↓	0,4285	0,5715	10	85.59%	1	8.65%	0,9913

Tabla 4.19: Candidatos y Factores de confianza. $C' = \frac{3}{4}$

La *Tabla 4.20* recoge los resultados cuando la amplitud de la Banda Secundaria se ha reducido a la mitad:

$$C' = \frac{1}{2} : I'_2 = \frac{1}{2} I_2$$

Espectro	ΔI	$\Delta I'$	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Coseno cuadrado
Esp 1 B.S.↓	0,8529	0,9264	1	88.28%			0,9747
Esp 2 B.S.↓	0,6315	0,8158	2	79.55%	3	11.44%	0,9045
Esp 3 B.S.↓	0,6875	0,8438	3	81.22%	5	9.06%	0,9712
Esp 4 B.S.↓	0,0967	0,5484	4	81.41%	2	23.11%	0,8706
Esp 5 B.S.↓	0,0967	0,5484	5	77.89%	7	5.75%	0,9466
Esp 6 B.S.↓	0,0307	0,5154	6	69.21%	2	13.31%	0,8072
Esp 7 B.S.↓	0,8857	0,9429	7	93.83%	1	16.84%	0,9921
Esp 8 B.S.↓	0,8936	0,9468	8	91.58%			0,9936
Esp 9 B.S.↓	0,8387	0,9194	9	95.49%	6	0.63%	0,9962
Esp 10 B.S.↓	0,4285	0,7143	10	71.13%	1	15.73%	0,9634

Tabla 4.20: Candidatos y Factores de confianza. $C' = \frac{1}{2}$

Como último caso, la Tabla 4.21 muestra los resultados cuando la amplitud de la Banda Secundaria se ha reducido a un cuarto su amplitud original:

$$C' = \frac{1}{4} : I'_2 = \frac{1}{4} I_2$$

Espectro	ΔI	$\Delta I'$	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Coseno cuadrado
Esp 1 B.S.↓	0,8529	0,9632	1	82.43%	7	7.96%	0,9455
Esp 2 B.S.↓	0,6315	0,9079	2	72.71%	6	0.83%	0,7839
Esp 3 B.S.↓	0,6875	0,9219	3	71.89%	5	10.76%	0,9392
Esp 4 B.S.↓	0,0967	0,7742	4	72.09%	2	28.03%	0,6902
Esp 5 B.S.↓	0,0967	0,7742	5	67.12%	3	15.67%	0,8899
Esp 6 B.S.↓	0,0307	0,7577	6	53.65%			0,5131
Esp 7 B.S.↓	0,8857	0,9715	7	90.76%	1	17.07%	0,9826
Esp 8 B.S.↓	0,8936	0,9734	8	87.37%			0,9859
Esp 9 B.S.↓	0,8387	0,9597	9	93.21%	6	1.09%	0,9924
Esp 10 B.S.↓	0,4285	0,8572	10	56.62%	1	20.81%	0,9191

Tabla 4.21: Candidatos y Factores de confianza. $C' = \frac{1}{4}$

Atendiendo a los resultados obtenidos, se puede concluir que, cuando las Bandas Secundarias decrecen, el sistema no se equivoca en ningún caso. Si el espectro se compone de un número elevado de bandas, se ve más afectado por el decrecimiento de las secundarias, como en el Espectro 6, cuya calidad de la representación disminuye sensiblemente.

Por ello, el siguiente apartado analiza en profundidad el caso límite de decrecimiento de las Bandas Secundarias ($C' = 0$), referente a la identificación con únicamente la Banda Fundamental. El que no aparezcan en el espectro las bandas secundarias, en caso de tenerlas, sobretodo si la diferencia con la Banda Fundamental es elevada, es menos importante en la identificación que el caso contrario estudiado antes.

4. 3. 4. Bandas Raman secundarias enmascaradas: Identificación con solo la Banda Raman fundamental

Las Bandas Raman secundarias pueden verse enmascaradas, por ejemplo, por efecto del ruido, dejando únicamente visible la Banda Fundamental del espectro a analizar, como se ha visto en el *Apartado 4.3.1*. Por ello, se estudia a continuación, el comportamiento del sistema implementado cuando el espectro a identificar sólo se compone de la Banda Fundamental (denotada por B.F.) del espectro patrón correspondiente, como caso límite del apartado anterior, *Apartado 4.3.3*, cuando la amplitud de las Bandas Secundarias es nula por efecto de la variación de las intensidades relativas. Como ejemplo, la siguiente figura muestra el Espectro 5 de la librería de referencia (en azul) y el Espectro 5 con solo la Banda Fundamental (en rojo):

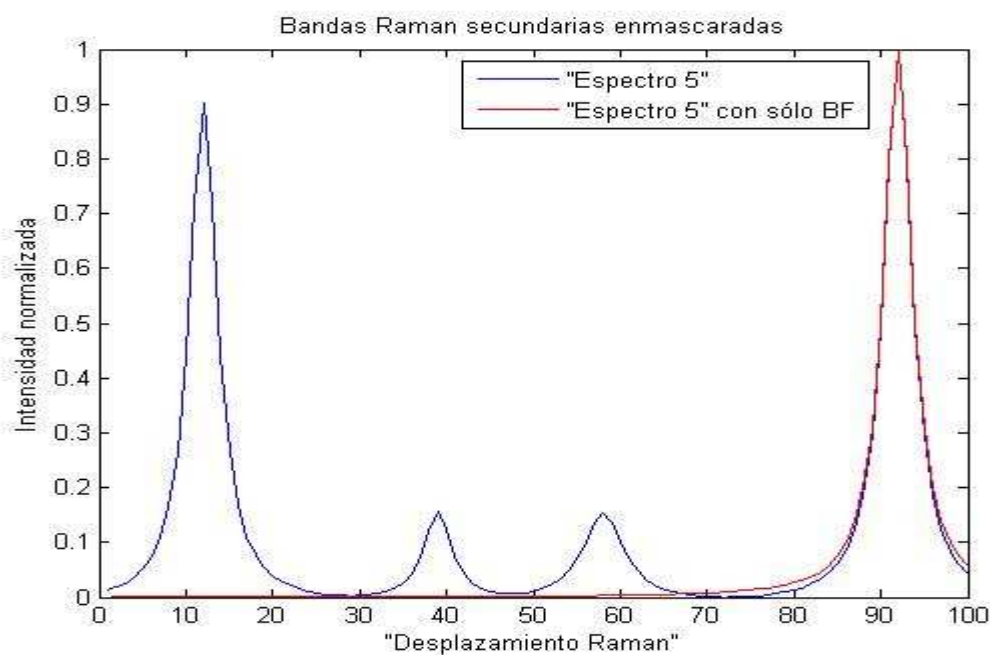


Fig. 4.27: Espectro 5 de la librería de referencia y Espectro 5 con sólo la B.F.

La siguiente tabla recoge las Distancias Euclídeas entre los espectros con sólo la Banda Fundamental y los espectros patrones. Se ha marcado en tonalidad verdosa las distancias que cumplen alguno de los requisitos para el reconocimiento espectral. Además, se adjunta la tabla correspondiente a los Factores de Confianza de la identificación respecto a los espectros candidatos hallados.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	COS2
Esp 1 solo con B.F.	2,0806	14,777	11,457	15,587	12,226	17,551	8,0135	12,58	12,861	8,0532	0,9088
Esp 2 solo con B.F.	12,068	5,602	11,773	15,044	12,312	15,311	10,767	13,352	9,8778	13,987	0,6469
Esp 3 solo con B.F.	11,94	14,233	3,7664	15,715	8,9643	16,364	11,063	14,121	8,868	12,84	0,9043
Esp 4 solo con B.F.	11,32	10,438	11,857	5,7184	10,998	12,542	11,434	13,869	11,938	13,65	0,4764
Esp 5 solo con B.F.	11,972	14,435	7,9994	15,265	4,281	15,479	10,81	13,901	10,463	12,968	0,8275
Esp 6 solo con B.F.	10,035	11,695	9,6638	10,952	9,5883	9,5075	9,2463	12,513	10,532	11,386	0,2341
Esp 7 solo con B.F.	8,7246	15,366	11,52	16,04	10,798	17,579	1,2975	10,462	11,805	11,503	0,9702
Esp 8 solo con B.F.	12,97	17,424	12,508	17,707	11,33	19,576	10,347	1,7729	14,28	10,423	0,9757
Esp 9 solo con B.F.	12,92	13,655	10,237	16,575	12,353	15,041	12,112	15,165	1,4172	14,14	0,9889
Esp 10 solo con B.F.	6,8059	13,95	10,398	15,122	12,34	17,638	10,52	13,683	12,899	5,149	0,8673

Tabla 4.22: Distancia Euclídea entre los espectros con sólo la B.F. y los espectros patrones

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 1 con solo B.F.	1	76.59%	7	9.84%	10	9.39%
Esp 2 con solo B.F.	2	63.53%	6	0.31%		
Esp 3 con solo B.F.	3	62.89%	9	12.63%	5	11.67%
Esp 4 con solo B.F.	4	58.09%				
Esp 5 con solo B.F.	5	56.51%	3	18.73%		
Esp 6 con solo B.F.	-	-				
Esp 7 con solo B.F.	7	87.69%	1	17.2%	8	0.71%
Esp 8 con solo B.F.	8	83.18%				
Esp 9 con solo B.F.	9	90.73%	6	2.07%		
Esp 10 con solo B.F.	10	42.07%	1	23.42%		

Tabla 4.23: Espectros candidatos y Factores de Confianza de la identificación

Como ejemplo, la figura *Fig. 4.28* muestra gráficamente el resultado para el espectro con más bandas, el Espectro 6, para el que, a la vista de la *Tabla 4.23*, no se obtiene ningún espectro candidato a la hora de ser identificado sólo a partir de la Banda Fundamental. Este hecho se explica a partir de su representación en las variables originales, puesto que el Espectro 6 de la librería de referencia consta de cinco bandas, cuatro de las cuales tienen una intensidad relativamente alta, por lo que, al considerar únicamente la Banda Fundamental, se pierde demasiada información, viéndose reflejado en la calidad de la representación ($COS2 = 23.41\%$).

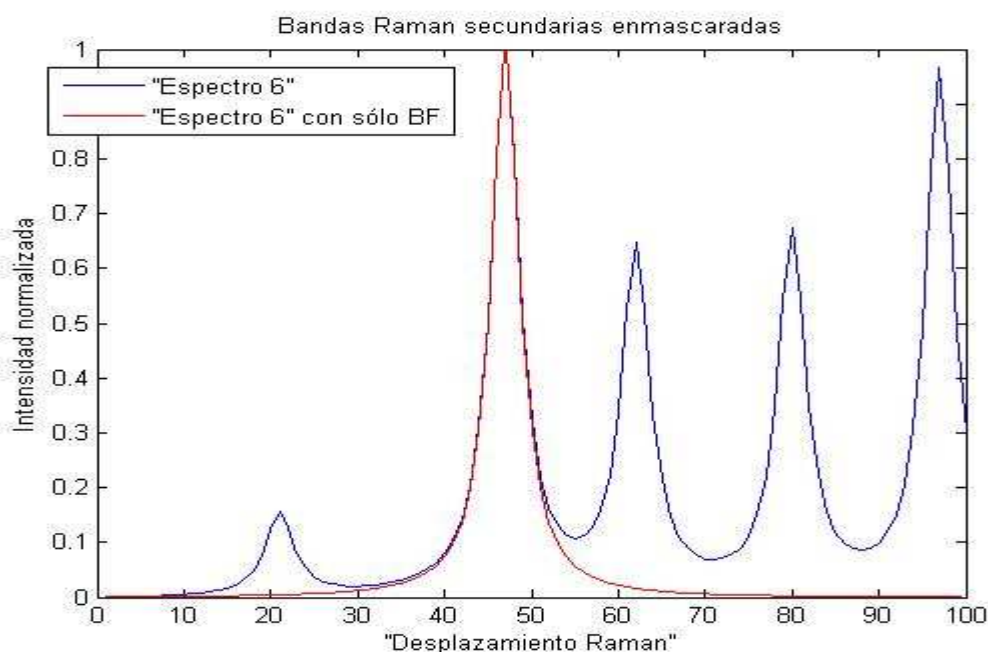


Fig. 4.28: Espectro 6 de la librería de referencia y Espectro 6 con sólo la B.F.

En cambio, si se considera el Espectro 9, con tan sólo dos bandas, al ser analizado a partir de únicamente la Banda Fundamental, sí se obtiene el espectro candidato correspondiente, así como con un Factor de Confianza relativamente alto, del 90.73%. En este caso, la calidad de la representación es considerablemente elevada ($COS2 = 98.89\%$), indicando, para este caso, una mínima pérdida de información al considerar tan sólo la Banda Fundamental. Además, se obtiene otro espectro candidato, el Espectro 6, puesto que tiene una banda coincidente (alrededor de 80cm^{-1}), aunque con un Factor de Confianza muy bajo, del 2.07%.

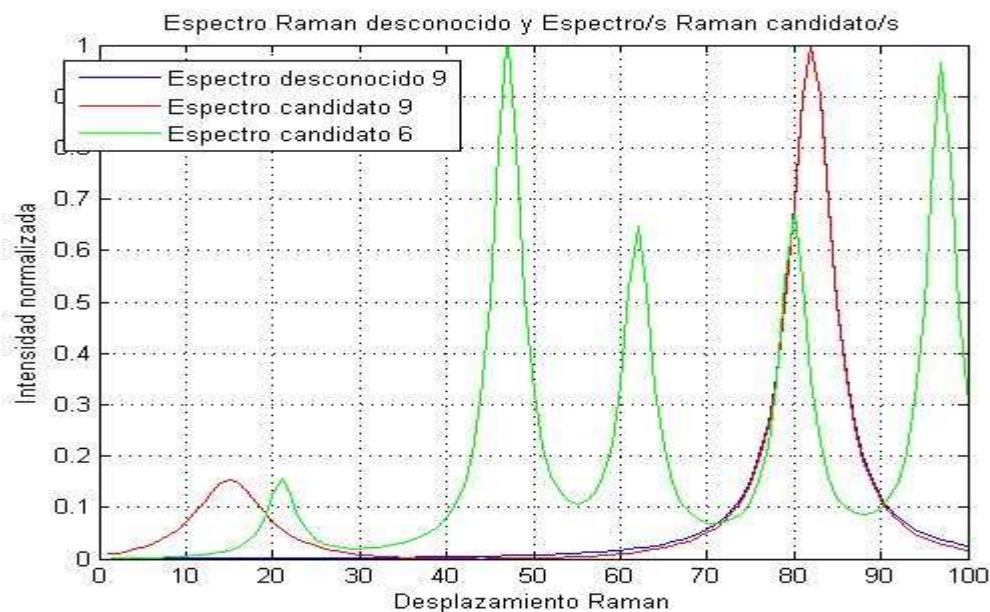


Fig. 4.29: Espectro 9 con sólo la B.F. y espectros candidatos

En el caso en que las bandas sean relativamente anchas, como por ejemplo en el Espectro 8, a la hora de ser identificado a partir de sólo la Banda Fundamental, se obtiene un único candidato, siendo éste el correcto, con un elevado Factor de Confianza, del 83.18%, y una alta calidad en la representación ($COS2 = 97.57\%$).

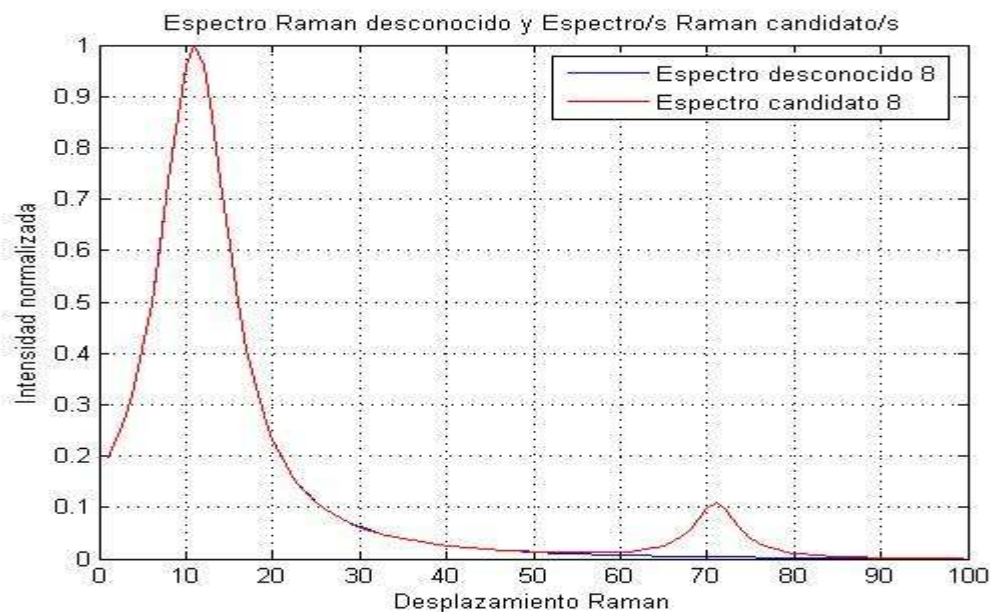


Fig. 4.30: Espectro 8 de la librería de referencia y Espectro 8 con sólo la B.F.

Si en cambio se considera el espectro con las bandas más estrechas, por ejemplo el Espectro 3, a la hora de ser analizado contemplando únicamente la Banda Fundamental, se obtienen tres espectros candidatos, aunque el correcto, se obtiene con un Factor de Confianza del 62.89%, mientras que los otros candidatos son generados con unos Factores de Confianza inferiores (un 12.63% para el Espectro 9 y un 11.67% para el Espectro 5). Para este caso, la calidad de la representación es también elevada ($COS2 = 90.43\%$).

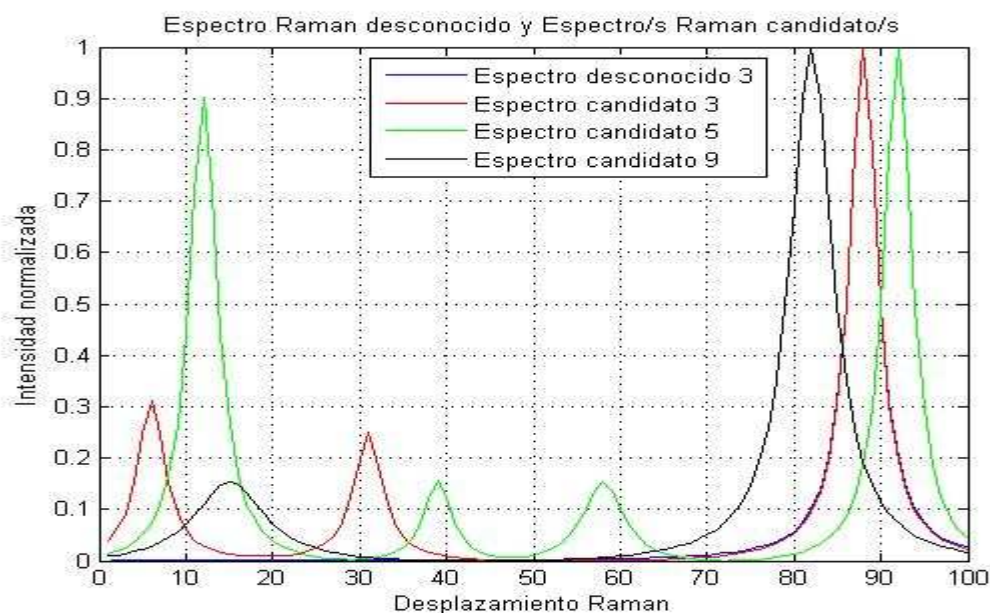


Fig. 4.31: Espectro 3 con sólo la B.F. y espectros candidatos

Con sólo la Banda Fundamental el efecto sobre la identificación depende del número de Bandas Raman del espectro patrón y de la importancia relativa de las Bandas Secundarias respecto a la Banda Fundamental.

El Factor de Confianza es mayor si las Bandas Secundarias tienen menor relevancia relativa respecto a la Banda Fundamental, como por ejemplo en el Espectro 1, el Espectro 8 y el Espectro 9.

4.3.5. Influencia del desplazamiento de las Bandas Raman

Debido a una mala calibración del equipo de medida, puede ocurrir que las Bandas Raman se vean desplazadas respecto a su posición original. Este efecto, provoca que la expresión original de un espectro se vea modificada respecto a la del patrón correspondiente y se modificará también su expresión en el espacio de PCs, pudiendo dificultar el reconocimiento espectral.

Es de esperar que exista un límite de desplazamiento de las bandas hasta el cual el sistema implementado sea capaz de identificar correctamente. Para establecerlo, se considera como espectros desconocidos, todos los espectros patrones con sus bandas desplazadas $\pm 1\text{cm}^{-1}$. A continuación se muestran los resultados del análisis para desplazamientos en un sentido concreto, puesto que los resultados para desplazamientos en el sentido contrario son similares, pudiéndose consultar en el *Anexo A2*.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	COS2
Esp 1 con despl.	1,3616	14,494	11,86	15,256	12,51	16,586	7,8123	12,495	12,342	9,4476	0,9652
Esp 2 con despl.	15,045	2,0527	15,644	14,61	16,643	16,707	15,883	17,029	15,356	16,297	0,8673
Esp 3 con despl.	11,776	14,691	1,809	16,322	10,726	17,177	11,812	12,556	9,2867	10,383	0,8663
Esp 4 con despl.	15,454	12,657	16,313	1,9833	15,087	16,257	16,484	17,719	17,182	16,447	0,8884
Esp 5 con despl.	12,184	14,178	8,2057	13,751	2,2747	16,017	11,001	10,794	11,384	11,923	0,8333
Esp 6 con despl.	16,109	15,145	16,321	13,968	14,931	2,1838	16,111	18,389	14,294	17,935	0,7817
Esp 7 con despl.	9,3832	14,391	11,559	15,879	10,54	17,154	1,1764	9,6218	11,113	11,666	0,9439
Esp 8 con despl.	13,794	16,804	13,227	17,899	12,705	19,535	11,961	1,6873	14,944	11,117	0,9804
Esp 9 con despl.	12,606	13,105	10,845	16,446	12,215	14,585	11,151	13,917	1,5985	13,715	0,9245
Esp 10 con despl.	8,2386	15,471	10,251	16,488	12,828	18,421	11,255	9,6934	13,272	1,7479	0,9283

Tabla 4.24: Desplazamiento de $+1\text{cm}^{-1}$ de las bandas: Distancia Euclídea

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 1 con despl.	1	84.6%	7	11.7%		
Esp 2 con despl.	2	84.93%				
Esp 3 con despl.	3	81.68%				
Esp 4 con despl.	4	85.53%	2	7.24%		
Esp 5 con despl.	5	77.05%	3	16.57%		
Esp 6 con despl.	6	87.71%				
Esp 7 con despl.	7	88.84%	8	8.68%		
Esp 8 con despl.	8	83.98%				
Esp 9 con despl.	9	89.48%	6	5.28%	2	14.65%
Esp 10 con despl.	10	83.41%	1	21.81%	8	8.01%

Tabla 4.25: Desplazamiento de las bandas: Candidatos y Factores de Confianza

Cuando el espectro desconocido es el Espectro 1, por ejemplo, con un desplazamiento de 1cm^{-1} , el sistema genera dos espectros candidatos, el Espectro 1 y el Espectro 7, aunque con un Factor de Confianza considerablemente más elevado para el primero. El Espectro 7 tiene su Banda Fundamental separada menos de 10cm^{-1} de la Banda Fundamental del Espectro 1 (ver Fig. 4.32). De esta forma, a medida que la fundamental del Espectro 1 se desplace hacia la posición de la del Espectro 7, el Factor de Confianza bajará para el candidato Espectro 1 y aumentará para el Espectro 7. Asimismo, la obtención de candidatos diferentes al espectro correspondiente se debe a la cercanía entre Bandas Fundamentales, corroborando así el resultado del estudio del Apartado 4.3.4 sobre la importancia de la Banda Fundamental respecto a la identificación.

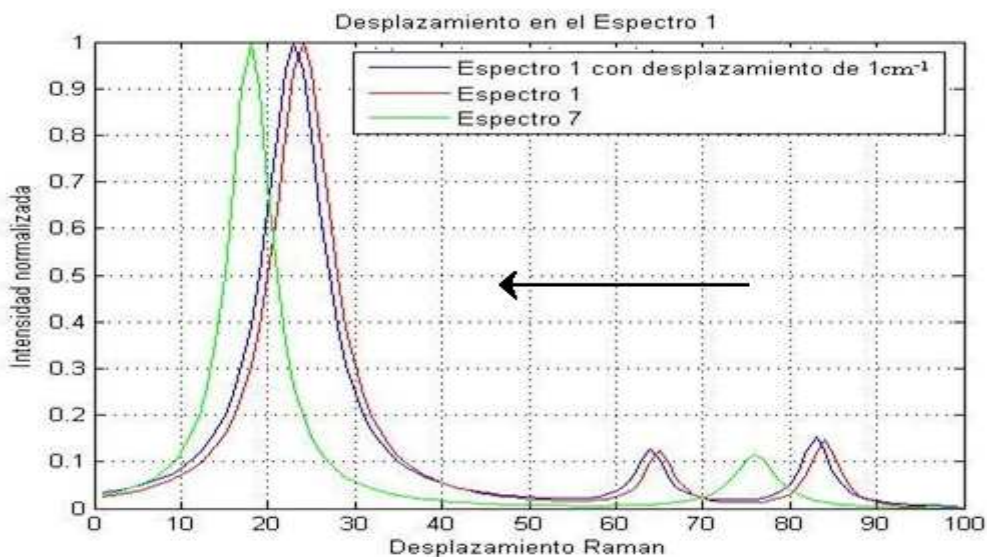


Fig. 4.32: Desplazamiento de 1cm^{-1} en el Espectro 1

A la vista de estos resultados, se puede concluir que un desplazamiento de las bandas de 1cm^{-1} puede hacer incrementar el número de espectros candidatos obtenidos por el sistema, pero siempre aparece el candidato correcto con un Factor de Confianza del orden del 80%. Por ello, se lleva a cabo el estudio cuando el desplazamiento es de 2cm^{-1} :

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	COS2
Esp 1 con despl.	2,8553	14,718	11,925	15,407	12,271	16,617	6,4821	12,036	12,194	9,9182	0,9032
Esp 2 con despl.	16,129	4,3405	17,123	15,806	18,235	17,206	17,413	17,898	16,788	17,57	0,652
Esp 3 con despl.	11,759	14,893	3,718	16,558	11,283	17,171	11,904	12,329	8,0034	10,426	0,6685
Esp 4 con despl.	15,922	11,623	16,737	4,3573	15,979	17,138	17,172	18,312	17,925	16,489	0,7046
Esp 5 con despl.	11,997	13,651	6,26	13,699	4,5748	16,17	11,056	10,323	10,979	11,095	0,644
Esp 6 con despl.	16,437	14,761	16,394	13,039	14,361	4,0739	16,085	18,525	14,325	18,009	0,5126
Esp 7 con despl.	9,7179	14,26	11,401	15,72	9,8935	17,167	2,4394	8,5096	11,015	11,508	0,8268
Esp 8 con despl.	14,762	17,522	13,818	18,562	13,922	20,107	13,344	3,3095	15,855	11,66	0,9319
Esp 9 con despl.	13,07	12,577	11,616	16,63	12,717	14,269	11,344	14,061	2,9021	14,111	0,7605
Esp 10 con despl.	7,8619	15,721	10,313	16,702	12,901	18,396	10,911	8,7418	13,181	3,56	0,7885

Tabla 4.26: Desplazamiento de las bandas: Distancia Euclídea

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 1 con despl.	1	67.72%	7	26.73%		
Esp 2 con despl.	2	68.17%				
Esp 3 con despl.	3	63.41%	9	21.16%		
Esp 4 con despl.	4	68.09%	2	14.82%		
Esp 5 con despl.	5	53.62%	3	36.39%		
Esp 6 con despl.	6	74.54%				
Esp 7 con despl.	7	76.85%	8	19.24%		
Esp 8 con despl.	8	68.59%				
Esp 9 con despl.	9	81.03%	2	18.09%	6	7.25%
Esp 10 con despl.	10	66.21%	1	25.39%	8	17.04%

Tabla 4.27: Desplazamiento de las bandas: Candidatos y Factores de Confianza

Puede verse que para este desplazamiento el sistema sigue funcionando correctamente, identificando al espectro correspondiente para cada espectro con desplazamiento. En relación al caso anterior, los Factores de Confianza han decrecido para los espectros candidatos correspondientes y han aumentado para los espectros candidatos erróneos, aunque el mayor Factor de Confianza sigue siendo el del espectro correspondiente. De esta forma, se procede al análisis cuando el desplazamiento es de 3cm^{-1} .

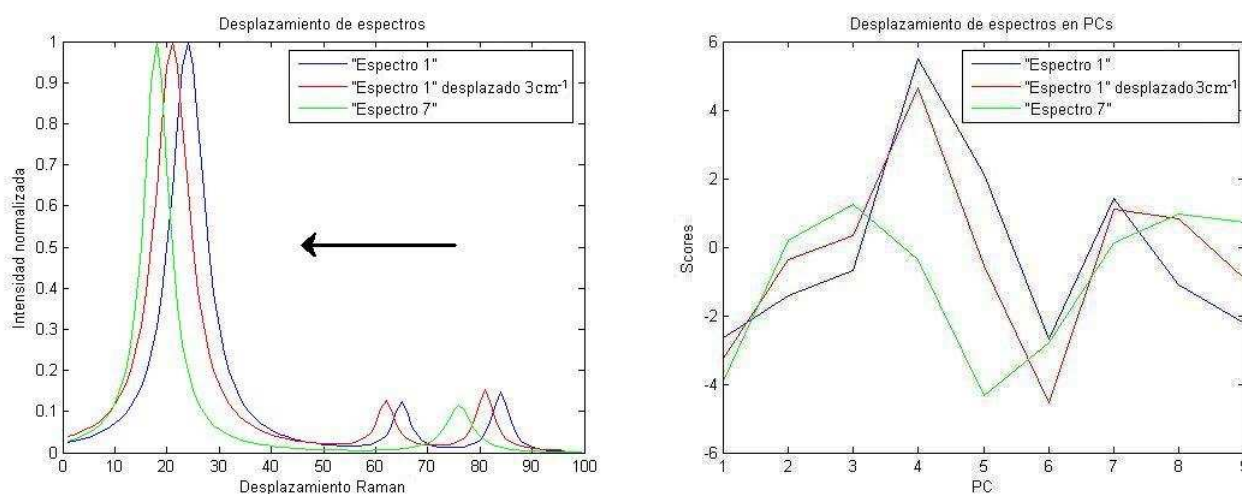


Fig. 4.33: Desplazamiento de 3cm^{-1} en el Espectro 1

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	COS2
Esp 1 con despl.	4,3978	14,883	11,998	15,534	12,06	16,689	5,0336	11,561	12,08	10,348	0,8687
Esp 2 con despl.	17,584	7,1762	19,122	17,155	20,142	17,764	19,267	19,12	18,461	19,469	0,4957
Esp 3 con despl.	11,522	14,896	5,6935	16,591	11,542	16,99	11,746	11,929	6,5418	10,418	0,5716
Esp 4 con despl.	16,478	10,56	17,312	7,1124	17,151	17,853	18,022	19,051	18,831	16,579	0,5843
Esp 5 con despl.	12,002	13,339	4,5714	13,984	7,1	16,381	11,317	10,072	10,769	10,357	0,6285
Esp 6 con despl.	16,457	13,828	16,276	11,275	13,669	6,4411	15,919	18,445	14,385	17,877	0,353
Esp 7 con despl.	9,9111	14,06	11,166	15,5	9,1482	17,099	3,853	7,2777	10,892	11,246	0,7364
Esp 8 con despl.	15,82	18,463	14,557	19,37	15,253	20,811	14,737	4,9468	16,901	12,408	0,8695
Esp 9 con despl.	13,724	11,976	12,496	16,988	13,452	14,219	11,746	14,419	4,3676	14,695	0,6073
Esp 10 con despl.	7,9232	16,12	10,718	17,088	13,228	18,545	10,772	7,9733	13,388	5,4366	0,6681

Tabla 4.28: Desplazamiento de las bandas: Distancia Euclídea

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 1 con despl.	1	50.29%	7	43.11%		
Esp 2 con despl.	2	47.38%				
Esp 3 con despl.	3	43.94%	9	35.57%		
Esp 4 con despl.	4	47.88%				
Esp 5 con despl.	3	53.55%	5	27.94%		
Esp 6 con despl.	6	59.01%				
Esp 7 con despl.	7	56.45%	8	30.93%		
Esp 8 con despl.	8	53.05%				
Esp 9 con despl.	9	71.51%	6	7.53%		
Esp 10 con despl.	10	48.41%	1	24.81%	8	24.33%

Tabla 4.29: Desplazamiento de las bandas: Candidatos y Factores de Confianza

En la figura *Fig. 4.33* se representa el espectro desconocido (Espectro 1 con desplazamiento de 3cm^{-1} de sus Bandas Raman) y los espectros candidatos detectados por el sistema, en donde se aprecia que la Banda Fundamental del desconocido está entre las Bandas Fundamentales de los candidatos. Los Factores de Confianza proporcionados son parecidos aunque ligeramente superior para el Espectro 1 debido al resto de Bandas Secundarias. Para el caso del Espectro 5 con desplazamiento, se produce una inversión en los valores de los Factores de Confianza proporcionado por el sistema para los candidatos (Espectro 5 y Espectro 3). Esto es, el Factor de Confianza es menor para el patrón candidato correspondiente.

No obstante, no puede decirse que el sistema esté actuando de manera errónea observando la representación de los espectros implicados (*Fig. 4.34*), puesto que la Banda Fundamental del Espectro 5 desplazado es aproximadamente coincidente con la B. F. del Espectro 3:

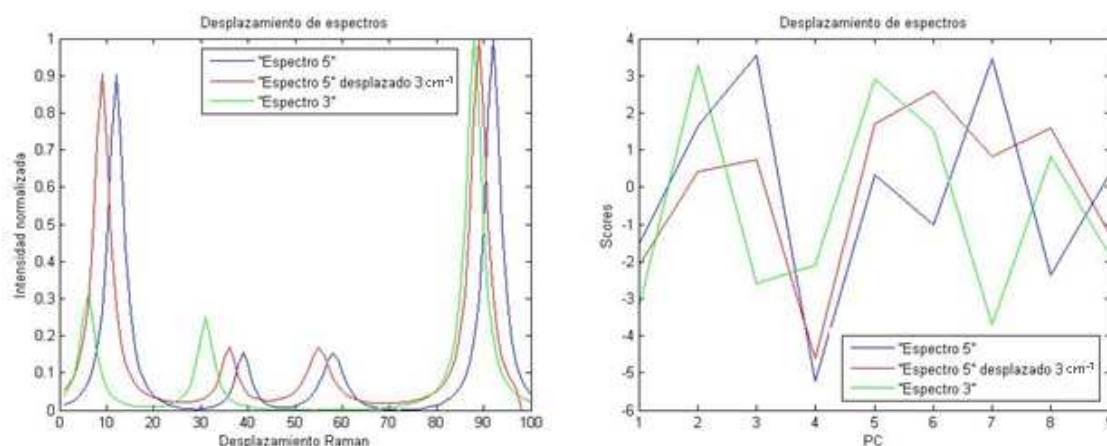


Fig. 4.34: Espectro 5 desplazado y espectros candidatos

Este resultado confirma la importancia de la identificación a partir de la Banda Fundamental analizada en el *Apartado 4.3.4*. El espectro patrón con la fundamental más cercana a la del Espectro 5 (situada en 91cm^{-1}) es el Espectro 3 (cuya fundamental está en 87cm^{-1}). Es por ello que, se establece que el límite máximo de desplazamiento de Bandas Raman hasta el cual el espectro desplazado es reconocido por el sistema implementado como el espectro original con el máximo Factor de Confianza, es función de la separación entre la Banda Fundamental del espectro patrón correspondiente y la del patrón más cercano a éste.

4. 3. 6. Presencia de Bandas Raman adicionales: Mezclas de pigmentos

Durante el proceso pictórico pueden generarse mezclas de pigmentos, lo cual puede implicar la aparición de Bandas Raman adicionales en el espectro medido. Para el análisis de dicha situación, se genera un nuevo espectro a partir del espectro patrón 7 con dos bandas adicionales cuyas posiciones no coinciden con las de ningún espectro de la librería de referencia simulada (ver *Tabla 4.30* y *Fig. 4.35*).

Espectro	Bandas del Espectro 7		Bandas adicionales	
Esp. Bandas adicionales	17 cm ⁻¹	75 cm ⁻¹	37 cm ⁻¹	58 cm ⁻¹

Tabla 4.30: Posición de las Bandas Raman del espectro desconocido

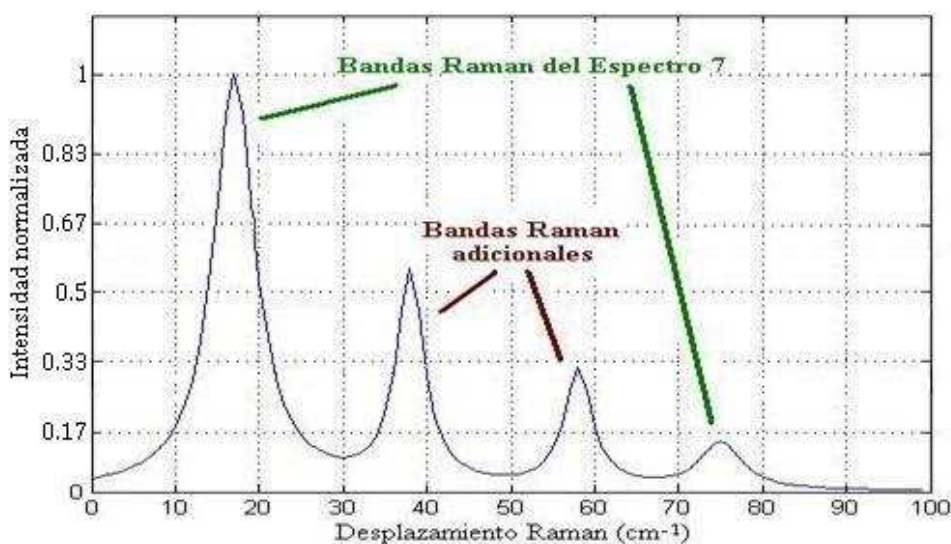


Fig. 4.35: Espectro patrón 7 con bandas adicionales

Como se puede ver en la *Tabla 4.31*, se obtienen tres espectros candidatos, el Espectro 7 con el mayor Factor de Confianza y otros dos espectros (el Espectro 4 y el Espectro 2) con unos Factores de Confianza inferiores al 10%, con lo que se decidiría que el espectro desconocido se corresponde con el Espectro 7. Una posible razón por la cual aparece el Espectro 4 y el Espectro 2 es que poseen una banda secundaria cercana a una de las bandas adicionales del espectro desconocido, el Espectro 4 en 38 cm⁻¹ y el Espectro 2 en 36 cm⁻¹.

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp. bandas adicionales	7	74.58%	4	9.41%	2	1.06%

Tabla. 4.31: Candidatos y Factores de Confianza para el espectro con bandas adicionales

Se concluye, así, que la aparición de bandas adicionales, que no pertenecen a ningún espectro de la librería de referencia, provoca que el sistema pueda generar varios candidatos, y que el candidato correcto obtenga el mayor Factor de Confianza, aún cuando éste se vea disminuido respecto al caso ideal.

La generación de otros espectros candidatos además del correcto, con bandas en posiciones cercanas a las bandas adicionales, motiva el análisis sobre qué sucede cuando dichas bandas adicionales sí se corresponden con algún espectro de la librería de referencia. Como ejemplo, se escogen los espectros 1 y 5 para generar un espectro mezcla (ver Fig. 4.36).

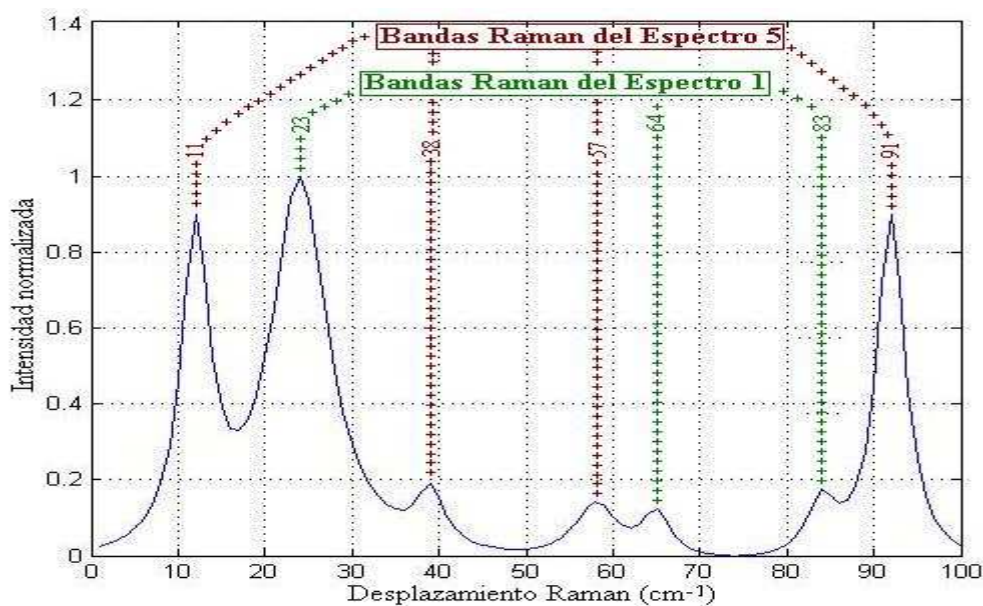


Fig. 4.36: Espectro mezcla a partir del Espectro 1 y el Espectro 5

La Tabla 4.32 muestra el resultado obtenido por el sistema. Se puede ver que se obtienen como espectros candidatos los espectros implicados en el espectro mezcla.

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp. bandas adicionales	1	33.38%	5	28.24%

Tabla. 4.32: Candidatos y Factores de Confianza para el espectro mezcla

Los Factores de Confianza obtenidos para los espectros candidatos, aunque relativamente bajos, son de un orden semejante, lo que puede llevar a pensar que el espectro desconocido sea en realidad un Espectro Raman de una mezcla de pigmentos, cosa que, *a priori*, no se conoce. Por ello, el siguiente apartado presenta un procedimiento para la identificación de un espectro mezcla.

4. 4. Estudio para la identificación de mezclas de pigmentos

En el proceso de creación de una obra artística es común la utilización de mezclas de pigmentos. Por ejemplo, las tonalidades verdosas son obtenidas a partir de la mezcla de un pigmento amarillo y otro azul. Los diferentes pigmentos que componen una mezcla hacen que el Espectro Raman de dicha mezcla pueda contener, en general, las Bandas Raman de los espectros de los pigmentos implicados.

Como se ha visto hasta este punto, el sistema implementado no es capaz de determinar si el espectro desconocido se corresponde con un espectro de una mezcla de pigmentos, aun pudiendo generar los espectros implicados como candidatos. Por ello, se presenta el siguiente procedimiento para la identificación automática de mezclas.

Los espectros desconocidos son proyectados sobre el espacio reducido generado por el Análisis por Componentes Principales aplicado sobre la librería de referencia. Puesto que los espectros mezcla no son identificados como tal por el sistema, una estrategia para su reconocimiento podría ser *crear* (en el espacio de las variables originales) el espectro mezcla a partir de los espectros candidatos proporcionados para el espectro desconocido (cuando éstos obtienen unos Factores de Confianza de un orden parecido), y aplicar nuevamente el ACP sobre la misma librería de referencia pero añadiéndole ahora el nuevo espectro mezcla creado. Esta estrategia, aunque viable, se ha descartado por el consumo de tiempo y recursos que implicaría volver a aplicar el ACP, lo cual, además, modificaría todos los parámetros involucrados en la metodología de identificación (C_{red} , min_{lib} , $COS2$, ...), es decir, significaría volver a comenzar todo el proceso desde el principio.

La estrategia que se propone se basa en trabajar directamente sobre el espacio reducido obtenido por el ACP, *creando* el espectro mezcla a partir de la expresión en PCs de los espectros candidatos (si tienen Factores de Confianza de orden similar). Esta estrategia es adecuada ya que no requiere volver a aplicar el ACP, aprovechando los parámetros ya obtenidos implicados en la metodología de identificación y ahorrando tiempo y recursos.

Sea E_M el espectro mezcla que se crearía en las variables originales generado a partir del espectro patrón i y el j normalizados, es decir,

$$E_M^{norm} = E_i^{norm} + E_j^{norm}$$

Dicho espectro sería normalizado, según el apartado 2.3 (Capítulo 2), de forma que,

$$E_M^{norm} = \frac{E_M - \min(E_M)}{\max(E_M) - \min(E_M)} \Big|_{\min(E_M)=0} = \frac{E_M}{\max(E_M)} = k \cdot E_M \rightarrow k = \frac{1}{\max(E_M)}$$

Según la estrategia propuesta, el espectro de la mezcla en el espacio reducido de PCs se puede expresar como:

$$\begin{aligned} S_{M_{red}} &= \left(\frac{kE_M^{norm} - m_{lib}}{\sigma_{lib}} \right) \cdot C_{red} = \left(\frac{k(E_i^{norm} + E_j^{norm}) - m_{lib} + k \cdot m_{lib} - k \cdot m_{lib}}{\sigma_{lib}} \right) \cdot C_{red} = \\ &= k \cdot \left(\frac{(E_i^{norm} + E_j^{norm}) - m_{lib} + \left(m_{lib} - \frac{m_{lib}}{k}\right)}{\sigma_{lib}} \right) \cdot C_{red} \end{aligned}$$

Siguiendo la misma idea, sumando y restando m_{lib} en el numerador,

$$S_{M_{red}} = k \cdot \left(\frac{(E_i^{norm} + E_j^{norm}) - m_{lib} - m_{lib} + m_{lib} + \left(m_{lib} - \frac{m_{lib}}{k}\right)}{\sigma_{lib}} \right) \cdot C_{red}$$

Puesto que,

$$m_{lib} + \left(m_{lib} - \frac{m_{lib}}{k}\right) = 2m_{lib} - \frac{m_{lib}}{k} = m_{lib} \left(2 - \frac{1}{k}\right) = m_{lib} \left(\frac{2k - 1}{k}\right)$$

Resulta que,

$$S_{M_{red}} = k \cdot \left(\frac{(E_i^{norm} + E_j^{norm}) - m_{lib} - m_{lib} + m_{lib} \left(\frac{2k - 1}{k}\right)}{\sigma_{lib}} \right) \cdot C_{red} =$$

$$= k \cdot \left(\frac{E_i^{norm} - m_{lib}}{\sigma_{lib}} + \frac{E_j^{norm} - m_{lib}}{\sigma_{lib}} + \frac{m_{lib} \left(\frac{2k-1}{k} \right)}{\sigma_{lib}} \right) \cdot C_{red} =$$

$$= k \cdot S_{i_{red}} + k \cdot S_{j_{red}} + (2k-1) \frac{m_{lib}}{\sigma_{lib}} C_{red}$$

El espectro de la mezcla de dos pigmentos se puede expresar en el espacio reducido de PCs en función de las expresiones de los pigmentos componentes de la mezcla en dicho espacio reducido y otras constantes almacenadas durante la aplicación del ACP a la librería de referencia.

Con el resultado matemático obtenido, se analiza nuevamente el ejemplo del apartado anterior, construyendo el espectro mezcla con el Espectro 1 y el Espectro 5 de la librería. Con este espectro mezcla construido se calcula la Distancia Euclídea respecto a los espectros patrones y el coseno cuadrado, aplicando los criterios descritos en el *Capítulo 3*. El sistema ahora, sí es capaz de determinar que se trata de una mezcla, obteniendo un Factor de Confianza muy superior al de los espectros por separado (ver *Tabla 4.33*).

Espectro	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp. mezcla	1	33.38%	5	28.24%	Mezcla 1 y 5	96.12%

Tabla 4.33: Espectros candidatos para ser mezcla y Factor de Confianza

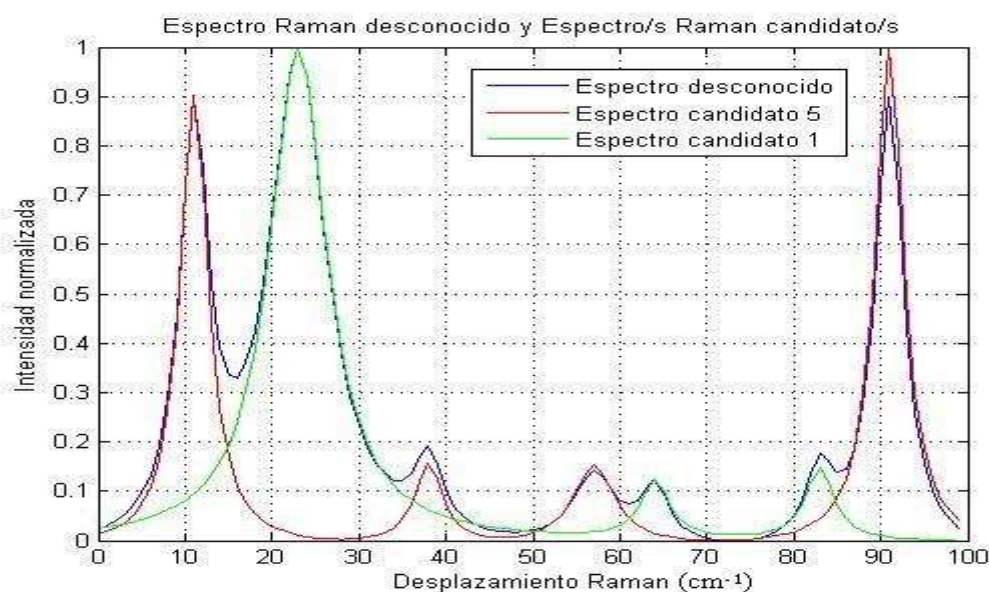
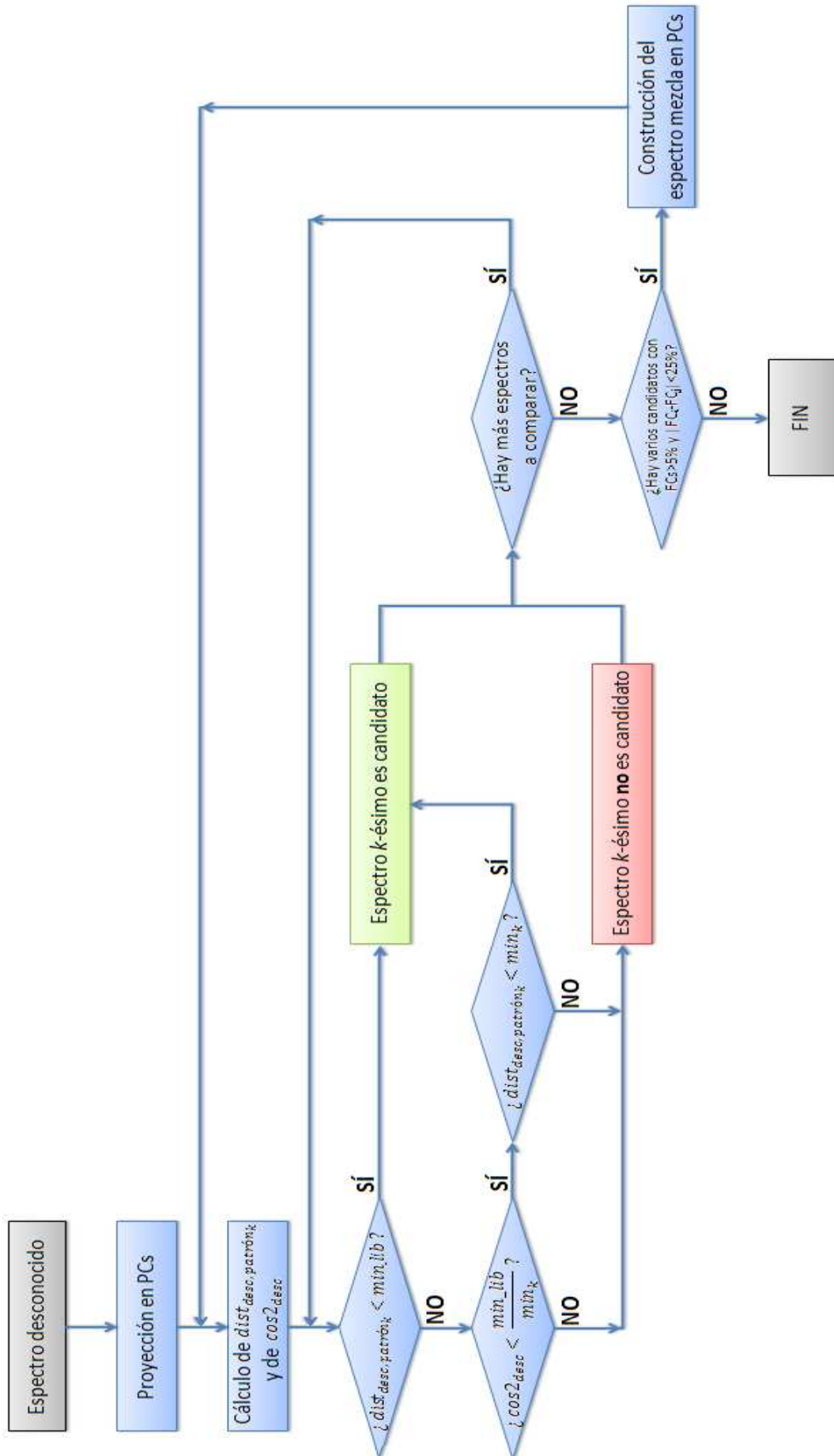


Fig. 4.37: Espectros candidatos generados por el sistema implementado para el espectro mezcla simulado a partir del Espectro 1 y el Espectro 5

Si para un determinado espectro desconocido el sistema genera varios espectros candidatos con unos Factores de Confianza de orden similar, pudiera resultar que dicho espectro desconocido perteneciera a una mezcla de pigmentos. En este caso, el sistema construye la mezcla binaria de los candidatos y se le aplica la metodología de identificación. Es por ello que se plantea que para considerar que el espectro desconocido puede ser una mezcla, los Factores de Confianza deben ser superiores a un determinado umbral (Δ_1), y, además, deben tener una diferencia menor que una cierta cantidad ($|FC_i - FC_j| < \Delta_2$). Se propone que el umbral sea de un $\Delta_1 = 5\%$ y la diferencia de un $\Delta_2 = 25\%$.

El diagrama de bloques de la página siguiente muestra, de forma esquemática, la metodología final implementada para la identificación de Espectros Raman de materiales pictóricos a partir del Análisis por Componentes Principales, contemplando la posibilidad de que el espectro desconocido pueda ser una mezcla binaria de pigmentos.



5. APLICACIÓN DEL ACP A LA IDENTIFICACIÓN DE ESPECTROS RAMAN DE PIGMENTOS ARTÍSTICOS. CASOS EXPERIMENTALES

5. 1. Librería espectral de referencia. ¿Cómo escoger los pigmentos?

A la hora de llevar a cabo un análisis espectroscópico sobre obras de arte, se deben tener en cuenta diferentes criterios para generar la librería espectral de referencia, seleccionando los pigmentos de forma estratégica: si se sospecha, por ejemplo, que la obra pudo ser realizada por un autor determinado se deberían incluir espectros de pigmentos que utilizase dicho autor, o si se cree que la obra se enmarca en un determinado movimiento artístico, incluir espectros de pigmentos utilizados en dicha época, siempre según documentación histórico-artística, como, por ejemplo, las tablas *Tabla 1.1* presentadas en el *Capítulo 1*.

Para el análisis del comportamiento de la metodología implementada en casos experimentales a partir de Espectros Raman de pigmentos medidos sobre obras artísticas, en este proyecto se proponen dos paletas. En primer lugar, se simula una situación hipotética en la que se querrían efectuar tareas de catalogación y/o conservación de una obra de arte supuestamente, por ejemplo, anterior al S.XX, caso en el que el estudio espectroscópico se realizaría sobre una paleta básica de pigmentos inorgánicos, elegidos estratégicamente; y, en segundo lugar, se parte del supuesto de conocer, *a priori*, que se trata de un pigmento amarillo moderno, un compuesto orgánico, ya que se dispone de una base de datos amplia de pigmentos amarillos orgánicos que generan espectros con un número elevado de Bandas Raman en posiciones próximas y muy semejantes.

5. 2. Identificación de espectros de pigmentos de una paleta básica

En el *Anexo C1* se pueden consultar los 21 Espectros Raman seleccionados en esta paleta básica, generadores de la librería espectral de referencia, todos ellos medidos en el laboratorio de Espectroscopía Raman de la UPC, así como los valores característicos de dicha librería, como son la Distancia Euclídea entre patrones, la mínima distancia y el Parámetro de Calidad referente al ruido umbral.

Ejemplo 1:

Este primer ejemplo recoge el resultado de aplicar la metodología implementada a un Espectro Raman desconocido de un pigmento obtenido al medir sobre una obra de arte. El sistema genera un candidato con un elevado Factor de Confianza, identificando, así, al espectro desconocido como el Espectro Raman del pigmento **Bermellón**.

El Factor de Confianza, de un 86.42%, no es ideal puesto que el espectro medido sobre la obra con el patrón no mantiene una correspondencia perfecta en la posición de las Bandas Raman, como se puede observar en la figura *Fig. 5.1*. Esto es debido, probablemente, a una mala calibración en el equipo de medida, lo cual modifica la posición de los *scores* al proyectar dicho espectro en el espacio de PCs (ver figura *Fig. 5.2*).

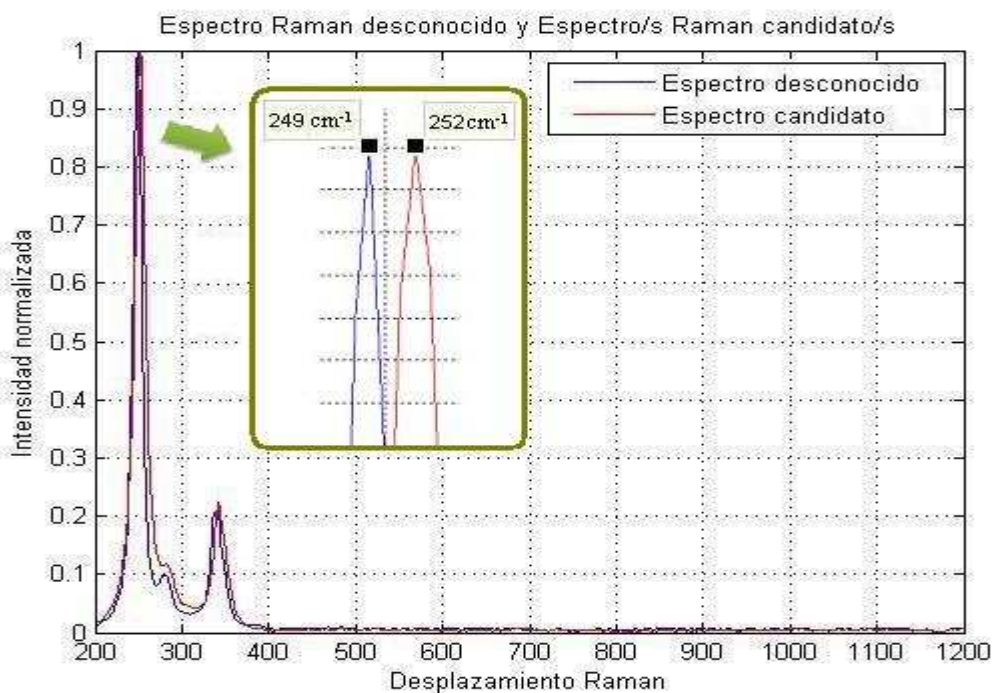


Fig. 5.1: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón Bermellón

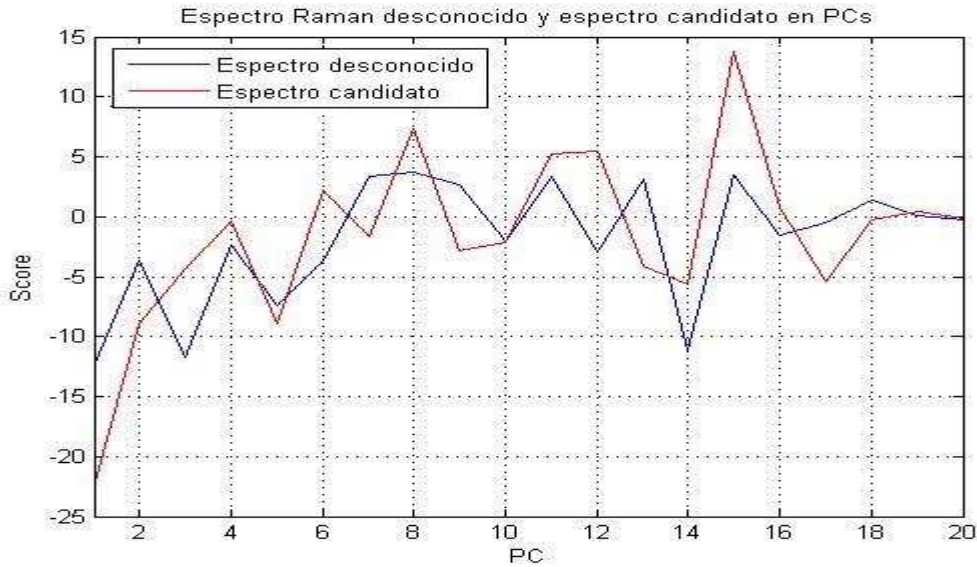


Fig. 5.2: Espectro desconocido y espectro candidato en PCs: Bermellón

En la *Tabla 5.1* pueden consultarse los valores obtenidos, tanto para la Distancia Euclídea como para el coseno cuadrado, al aplicar la metodología de identificación implementada a este primer espectro.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	40,847	58,383	45,919	25,291	27,507	23,216	31,884	45,358	48,894	42,661	40,349
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	37,663	45,89	23,297	59,469	33,055	24,997	24,145	52,443	39,662	3,487	0,9141

Tabla 5.1: Distancia Euclídea y coseno cuadrado

Siguiendo la definición del cálculo del Parámetro de Calidad referente al ruido, Q , se obtiene una relación relativamente elevada respecto al umbral de la librería espectral de referencia, lo que permite que el espectro desconocido pueda ser identificado:

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.0013} \right) = 28.86dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 14.35dB$$

Aún así, como se ha podido observar en la figura *Fig. 5.1*, se ha producido un desplazamiento de las Bandas Raman, efecto principal, en este caso, que provoca la disminución del Factor de Calidad:

$$\Delta = 252cm^{-1} - 249cm^{-1} = 3cm^{-1}$$

Ejemplo 2:

En este segundo ejemplo se presenta el resultado de la identificación de un nuevo Espectro Raman desconocido obtenido al realizar el análisis de una obra de arte. El sistema genera un espectro candidato con un elevado Factor de Confianza, del 79.73%, identificando el espectro desconocido como el Espectro Raman del pigmento **Amarillo de Cromo**. El Factor de Confianza no es ideal puesto que, como se puede observar en la figura Fig. 5.3, se ha producido un desplazamiento en sus Bandas Raman, lo cual modifica la posición de los *scores* (ver Fig. 5.4), dificultando la identificación espectral.

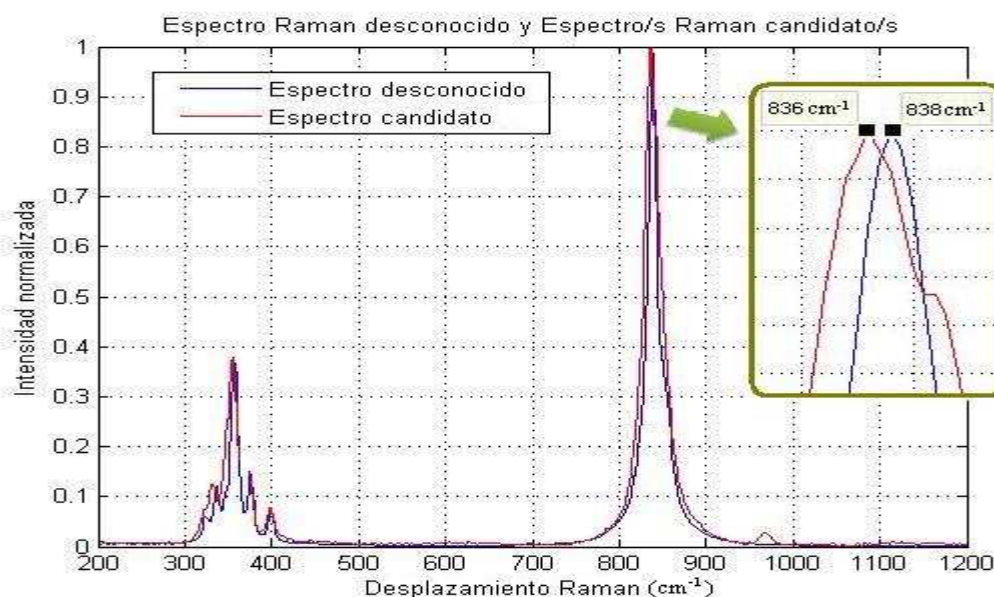


Fig. 5.3: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón Amarillo de Cromo

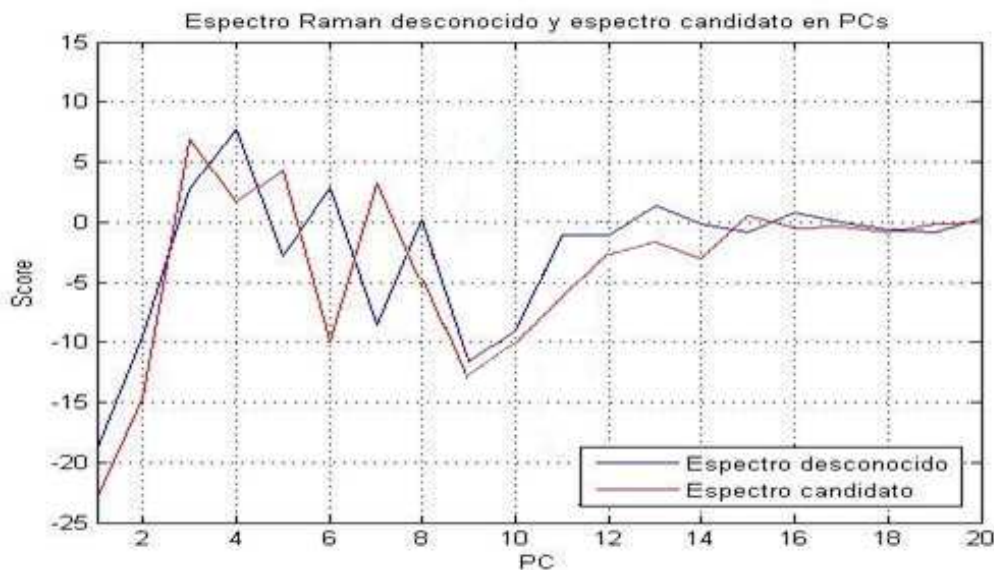


Fig. 5.4: Espectro desconocido y espectro candidato en PCs: Amarillo de Cromo

La *Tabla 5.2* muestra los valores de Distancia Euclídea y coseno cuadrado obtenidos al aplicar la metodología de identificación implementada sobre este segundo ejemplo.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	42,263	60,904	44,299	31,56	33,083	29,832	35,469	41,135	53,896	47,288	11,095
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	42,658	39,962	30,16	64,911	40,928	31,651	29,17	58,429	46,234	34,404	0,9493

Tabla 5.2: Distancia Euclídea y coseno cuadrado

A partir del cálculo del Parámetro de Calidad referente al ruido se concluye que el espectro desconocido puede ser identificado, ya que mantiene una relación elevada respecto al umbral de la librería de referencia:

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.0052} \right) = 22.84dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 14.35dB$$

Como se ha podido observar en la figura Fig. 5.3, se ha producido un desplazamiento de las Bandas Raman:

$$\Delta = 838cm^{-1} - 836cm^{-1} = 2cm^{-1}$$

Este desplazamiento modifica el valor de los *scores* al proyectarse en el espacio de PCs, obstaculizando el reconocimiento espectral.

Aunque el desplazamiento de las Bandas Raman en este ejemplo es inferior al desplazamiento visto en el primer ejemplo, el Factor de Confianza en la identificación es menor. Esto es debido a que la componente de ruido es mayor en este segundo ejemplo que en el primero, como se extrae del cálculo de los Parámetros de Calidad referentes al ruido:

$$Q_{dB}^{Ej.1} = 28.86dB > Q_{dB}^{Ej.2} = 22.84dB \rightarrow Ruido_{Ej.1} < Ruido_{Ej.2}$$

lo cual implica que el ruido dificulta el reconocimiento espectral en mayor medida que el desplazamiento de Bandas Raman.

Ejemplo 3:

Este tercer ejemplo recoge el resultado del reconocimiento de un Espectro Raman obtenido sobre una obra artística. El sistema genera un espectro candidato, identificando al espectro desconocido nuevamente como el espectro del pigmento **Amarillo de Cromo**. No obstante, se obtiene un bajo Factor de Confianza, del 8.96%, puesto que se dan diferentes problemas prácticos (desplazamiento de las Bandas Raman y diferencia en las intensidades relativas), como puede observarse en la figura Fig. 5.5, que modifican la posición de los *scores* (ver Fig. 5.6), dificultándose, así, la identificación espectral.

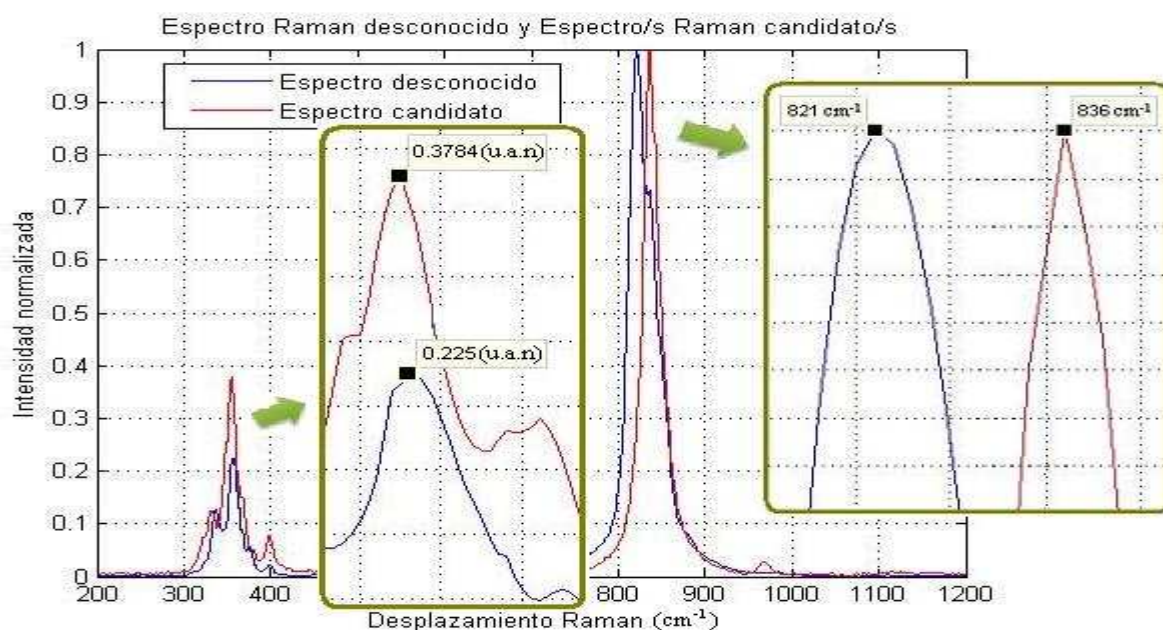


Fig. 5.5: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón Amarillo de Cromo

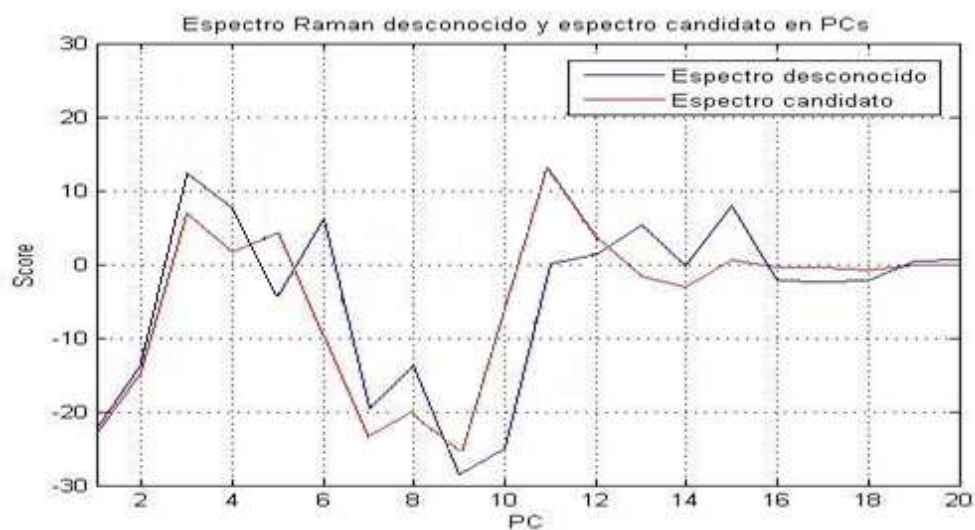


Fig. 5.6: Espectro desconocido y espectro candidato en PCs: Amarillo de Cromo

La tabla *Tabla 5.3* muestra los valores de Distancia Euclídea y coseno cuadrado obtenidos por el sistema implementado.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	71,066	83,246	64,131	68,797	66,925	67,134	68,179	59,928	77,935	74,529	33,376
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	71,591	65,942	66,867	86,973	72,377	68,985	69,029	83,79	74,852	71,318	0,6109

Tabla 5.3: Distancia Euclídea y coseno cuadrado

El Parámetro de Calidad referente al ruido se muestra elevado respecto al umbral, permitiendo la identificación del espectro:

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.0006} \right) = 32.22dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 14.35dB$$

Como se ha podido observar en la *Fig. 5.5*, se ha producido una diferencia en las intensidades relativas, concretamente la amplitud de la Banda Secundaria del espectro desconocido ha disminuido:

$$I_2' \downarrow \rightarrow I_2' = C \cdot I_2 \rightarrow C = \frac{I_2'}{I_2} = \frac{0.225}{0.3784} = 0.595 \approx 0.6 \rightarrow I_2' = 0.6 \cdot I_2$$

Según los resultados obtenidos en la librería simulada, el sistema identifica correctamente cuando se da una disminución de la amplitud de las Bandas Secundarias (excepto en el caso límite de detección con sólo la Banda Fundamental si el espectro está formado por un número elevado de bandas), con lo que se puede concluir que esta disminución de amplitud no afecta excesivamente en el reconocimiento espectral.

El problema práctico que más afecta a la identificación de este espectro y que, por ello, hace que el Factor de Confianza se vea reducido a un valor tan bajo, es el elevado desplazamiento de las Bandas Raman respecto al espectro patrón:

$$\Delta = 836cm^{-1} - 821cm^{-1} = 15cm^{-1}$$

Ejemplo 4:

A continuación se recoge un ejemplo de no identificación. El sistema no es capaz de reconocer el Espectro Raman desconocido obtenido a partir de un pigmento de una obra artística. El principal problema práctico que afecta a este espectro es el ruido, de una intensidad considerable, como puede verse en la figura Fig. 5.7, lo cual modifica los scores (ver Fig. 5.8), haciendo imposible la identificación.

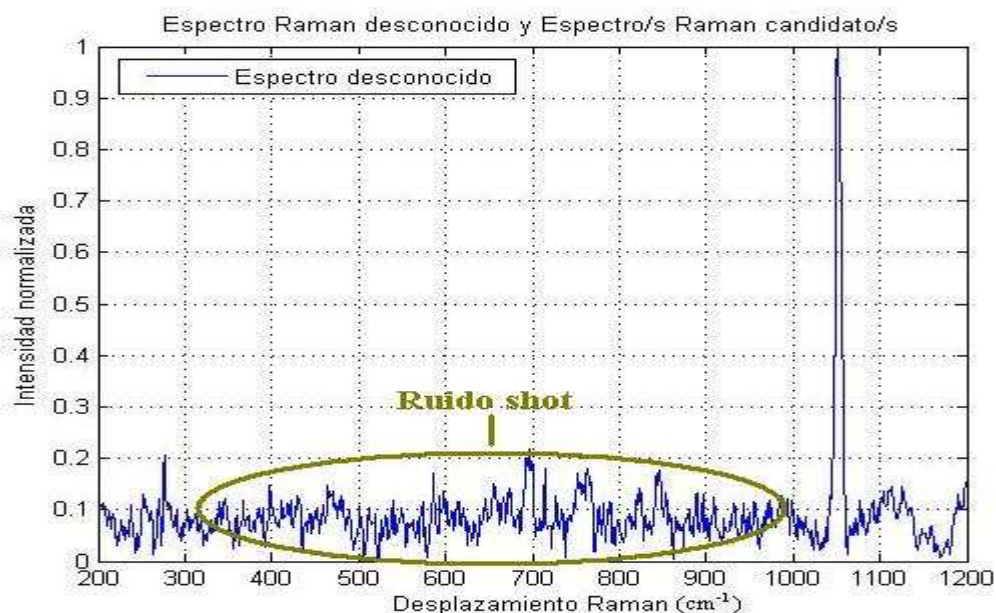


Fig. 5.7: Espectro desconocido no identificado (Blanco de Plomo)

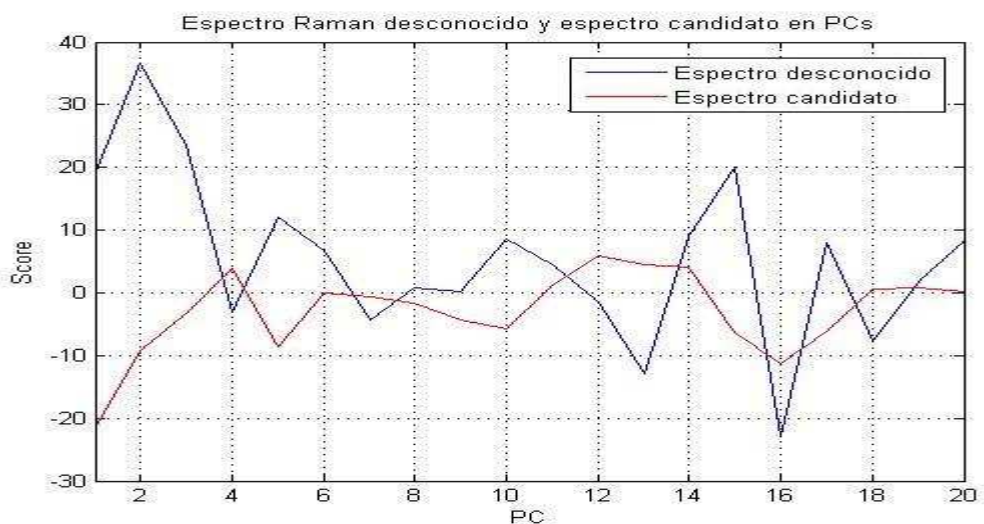


Fig. 5.8: Espectro desconocido no identificado en PCs

La siguiente tabla muestra los valores de Distancia Euclídea y coseno cuadrado obtenidos al aplicar la metodología implementada.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	80,88	69,994	73,154	67,613	62,926	70,919	67,237	67,103	64,245	58,8	77,866
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	59,255	67,208	67,45	52,806	70,13	74,053	80,125	81,966	68,022	79,859	0,7296

Tabla 5.4: Distancia Euclídea y coseno cuadrado

Si se procede a calcular el Parámetro de Calidad referente al ruido, se observa que éste es inferior al umbral, implicando que la componente de ruido en este espectro desconocido es superior al límite establecido y que, por tanto, el espectro no puede ser identificado.

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.04} \right) = 13.97dB$$

$$Q_{dB}^{desconocido} < Q_{dB}^{umbral} = 14.35dB \rightarrow \text{El espectro no es identificado}$$

Ejemplo 5:

En este quinto ejemplo, al igual que en el anterior, no se produce reconocimiento espectral. El sistema no es capaz de identificar al espectro desconocido puesto que se ha producido una reducción en la anchura de sus bandas, como puede verse en la figura Fig. 5.9, modificando los scores (ver Fig. 5.10), haciéndose imposible la identificación.

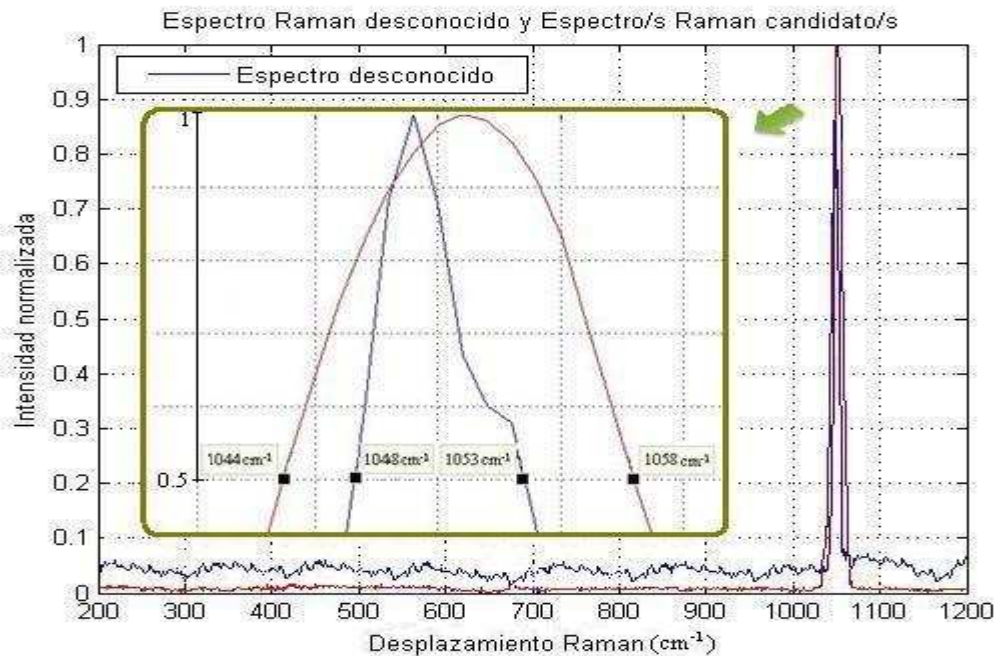


Fig. 5.9: Espectro desconocido no identificado (en azul) y Espectro Raman del pigmento Blanco de Plomo (en rojo)

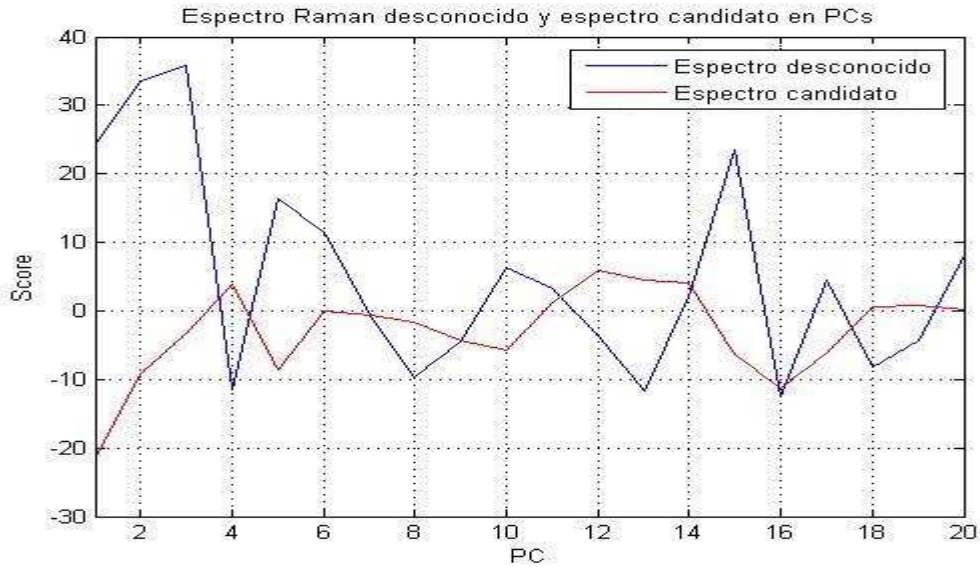


Fig. 5.10: Espectro desconocido no identificado en PCs

La siguiente tabla muestra los valores de Distancia Euclídea y coseno cuadrado obtenidos por el sistema de identificación espectral.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	86,075	71,478	75,18	78,963	68,954	77,749	72,9	70,626	73,366	58,797	81,581
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	64,026	71,209	75,343	61,316	74,744	81,474	87,632	85,495	69,788	86,83	0,6251

Tabla 5.5: Distancia Euclídea y coseno cuadrado

El cálculo del Parámetro de Calidad referente al ruido concluye que el espectro puede ser identificado:

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.0121} \right) = 19.17dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 14.35dB$$

No obstante, como se ha podido observar en la Fig. 5.9, se ha producido una reducción de la anchura de las Bandas Raman:

$$BW_{desconocido} = 1053cm^{-1} - 1048cm^{-1} = 5cm^{-1}$$

$$BW_{patrón} = 1058cm^{-1} - 1044cm^{-1} = 14cm^{-1}$$

$$BW_{desconocido} = K \cdot BW_{patrón} \rightarrow K = \frac{BW_{desconocido}}{BW_{patrón}} = \frac{5cm^{-1}}{14cm^{-1}} = 0.357 = 35.7\% \rightarrow$$

$$Reducción\ del\ 64.3\% > Reducción\ límite\ (60\%)$$

Esta reducción del 64.3% es superior a la reducción límite del 60%, hallada sobre la librería simulada, con lo cual, esta es la razón por la que el espectro desconocido no es identificado.

Ejemplo 6:

En este sexto ejemplo se presenta el resultado de la identificación de un Espectro Raman desconocido obtenido al analizar una obra de arte. El sistema genera un espectro candidato con un Factor de Confianza, del 38.79%, identificando el espectro desconocido como el del pigmento **Azul Ultramar**. El Factor de Confianza es bajo ya que, como se puede observar en la figura *Fig. 5.11*, se han producido nuevamente problemas prácticos en la medida del espectro, lo cual modifica la posición de los *scores* (ver *Fig. 5.12*).

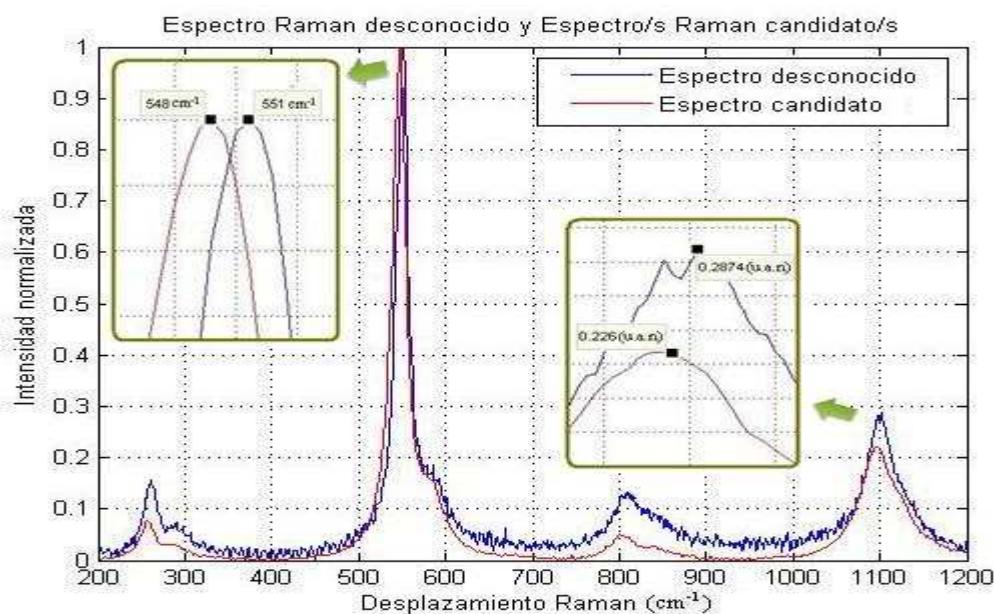


Fig. 5.11: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón Azul Ultramar

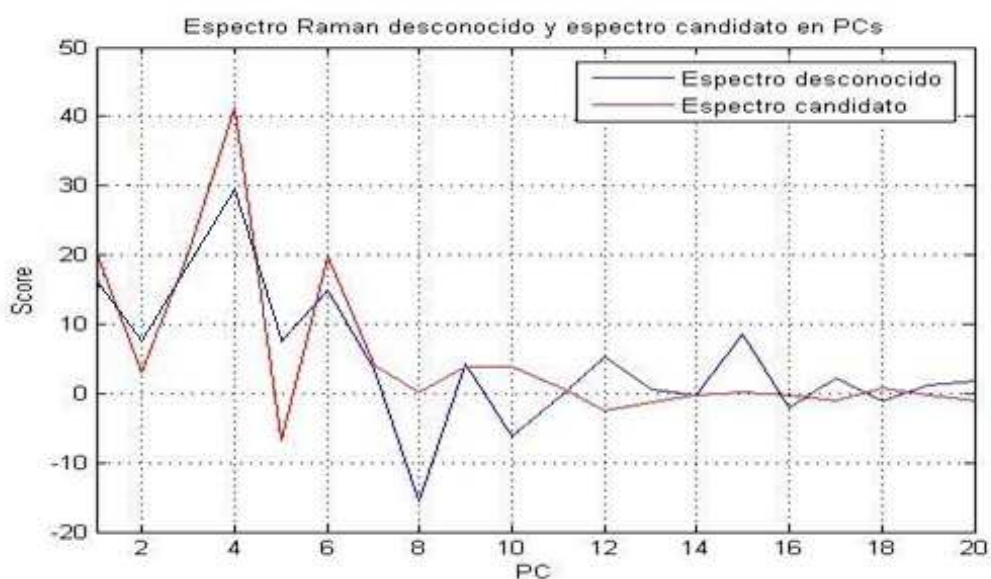


Fig. 5.12: Espectro desconocido y espectro candidato en PCs: Azul Ultramar

La tabla *Tabla 5.6* muestra los valores de Distancia Euclídea y coseno cuadrado obtenidos por el sistema implementado.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	69,57	64,619	62,609	63,368	52,065	59,877	59,014	59,633	30,153	52,223	63,86
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	54,433	60,456	56,226	48,281	60,424	57,61	68,158	72,597	62,317	67,902	0,8748

Tabla 5.6: Distancia Euclídea y coseno cuadrado

El Parámetro de Calidad referente al ruido indica que el espectro puede ser identificado, ya que su valor es un poco superior al umbral.

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{nE_i}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.016} \right) = 17.96dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 14.35dB$$

Asimismo, se ha producido un desplazamiento de las Bandas Raman, como ha podido observarse en la *Fig. 5.11*,

$$\Delta = 551cm^{-1} - 548cm^{-1} = 3cm^{-1}$$

Además, se ha producido una diferencia de intensidades relativas, concretamente, la amplitud de la Banda Secundaria ha aumentado:

$$I_2' \uparrow \rightarrow I_2' = I_2 + C \cdot \Delta I \rightarrow C = \frac{I_2' - I_2}{\Delta I} = \frac{0.2874 - 0.226}{1 - 0.226} = 0.079 \approx 0.08 = 8\% \rightarrow$$

$$\rightarrow k = (1 - C) = 1 - 0.08 = 0.92 \rightarrow \Delta I' = k \cdot \Delta I = 0.92 \cdot \Delta I$$

lo cual quiere decir que la relación de intensidades entre la Banda Fundamental y la Banda Secundaria de mayor amplitud, ΔI , se ha visto disminuida un 8% respecto al espectro patrón.

El conjunto de todos estos problemas prácticos hace que el Factor de Confianza en la identificación se vea afectado y se aleje del valor ideal.

Ejemplo 7:

Este ejemplo se corresponde al estudio del comportamiento del sistema al aplicarse sobre un Espectro Raman **mezcla** de pigmentos obtenido al analizar una obra de arte.

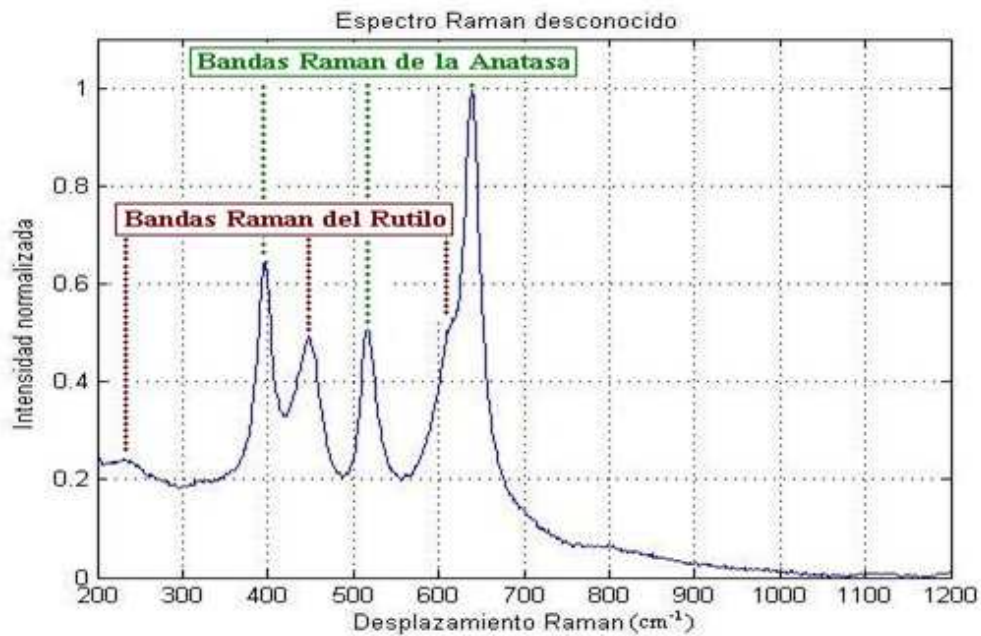


Fig. 5.13: Espectro mezcla desconocido

El sistema genera tres espectros candidatos, concretamente el blanco Anatasa, con un Factor de Confianza del 26.43%, el Rutilo, con un Factor de Confianza del 21.31%, y el Azul de Cobalto, con un Factor de Confianza del 14.19%.

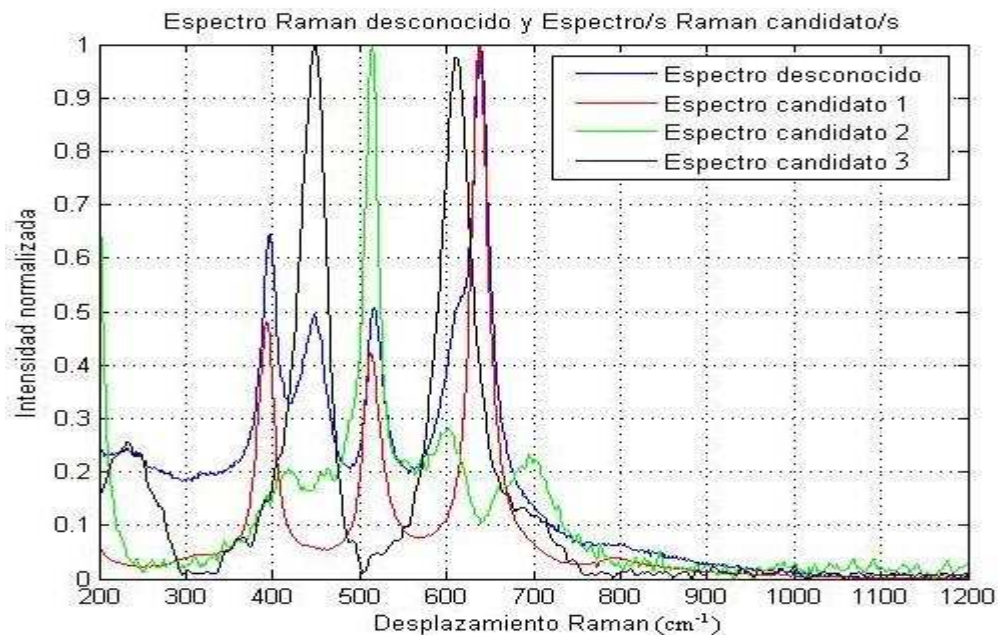


Fig. 5.14: Espectro desconocido y espectros candidatos generados por el sistema implementado (1: Anatasa, 2: Azul de Cobalto y 3: Rutilo)

Como puede verse, dichos Factores de Confianza son no muy altos y de un orden similar, con lo que puede llevar a pensar que el espectro desconocido se trate en realidad de una mezcla de los Espectros Raman candidatos obtenidos. Por ello, el sistema, automáticamente, crea los espectros mezcla binarios: Anatasa-Rutilo, Anatasa-Azul de Cobalto y Rutilo-Azul de Cobalto y vuelve a aplicar la metodología implementada. Se obtiene ahora un único espectro mezcla candidato, el Anatasa-Rutilo, con un Factor de Confianza del 59.38%, identificando, así, al espectro desconocido como el Espectro Raman correspondiente a la mezcla de los pigmentos Anatasa y Rutilo (ver Fig. 5.15 y Fig. 5.16).

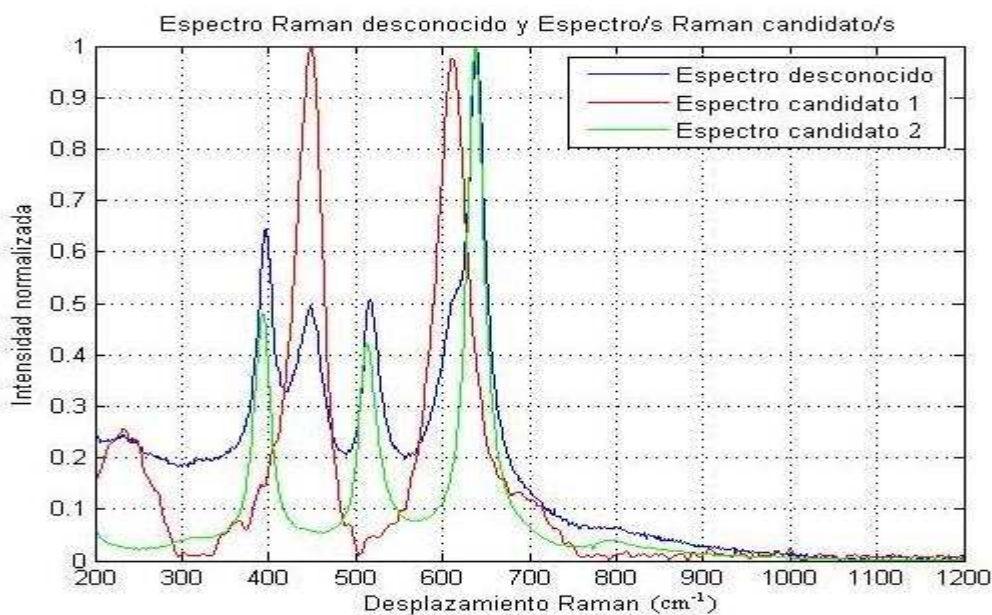


Fig. 5.15: Espectro desconocido y espectros candidatos generados por el sistema implementado contemplando el espectro mezcla (1: Rutilo y 2: Anatasa)

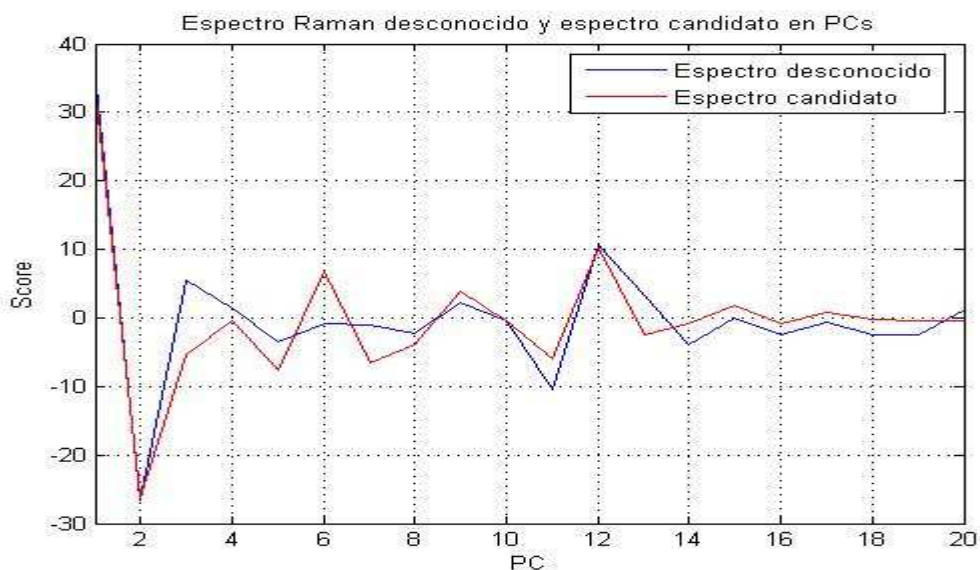


Fig. 5.16: Espectro desconocido y espectro candidato en PCs: Mezcla de Rutilo y Anatasa

La *Tabla 5.7* recoge los valores de Distancia Euclídea entre los espectros patrones y el espectro desconocido, así como el coseno cuadrado de éste último. La *Tabla 5.8* muestra la Distancia Euclídea entre los Espectros Raman mezcla de patrones y el espectro desconocido.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	32,827	38,289	52,534	62,738	59,705	56,757	53,71	53,505	68,906	59,803	64,901
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	55,444	65,851	56,314	66,623	51,471	57,458	61,795	35,112	60,416	61,437	0,9501

Tabla 5.7: Distancia Euclídea y coseno cuadrado

DISTANCIA	Anatasa-Rutilo	Anatasa-Azul de Cobalto	Rutilo-Azul de Cobalto
Espectro desconocido	17,871	24,596	39,702

Tabla 5.8: Distancia Euclídea respecto a los espectros patrones mezcla

Como puede observarse, el espectro desconocido posee un coseno cuadrado relativamente elevado, indicando, así, una representación cercana a la ideal. Asimismo, se puede observar que el valor menor de Distancia Euclídea es el que se corresponde con la mezcla Anatasa-Rutilo.

El Factor de Confianza se ha visto afectado puesto que existe una cierta componente de ruido, que se hace patente si se calcula el Parámetro de Calidad referente al ruido:

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.0073} \right) = 21.37dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 14.35dB$$

Dicho efecto juntamente con el hecho de no corresponderse perfectamente el espectro desconocido con la mezcla creada hacen que el Factor de Confianza en la identificación disminuya respecto al caso ideal. No obstante, con este ejemplo, se comprueba que el sistema es capaz de reconocer Espectros Raman correspondientes a mezclas binarias de pigmentos.

5. 3. Identificación de Espectros Raman de pigmentos con un número elevado de bandas

En el *Anexo C2* se pueden consultar los 21 Espectros Raman seleccionados en esta paleta de pigmentos amarillos orgánicos (*Pigment Yellow, PY*), generadores de la librería espectral de referencia, así como los valores característicos de dicha librería, como son la Distancia Euclídea entre espectros patrones, distancia mínima entre un espectro patrón y el resto de espectros patrones, la mínima distancia y el Parámetro de Calidad referente al ruido umbral.

Esta paleta genera un interés especial debido a que los Espectros Raman procedentes de pigmentos orgánicos (en este caso amarillos), poseen un número elevado de Bandas Raman y en posiciones muy cercanas y similares. Se pretende, pues, analizar el comportamiento de la metodología implementada en el caso de espectros muy parecidos y con muchas bandas.

Los Espectros Raman utilizados como espectros de la librería de referencia fueron medidos sobre obras de arte del Museo de Arte Moderno de Ljubljana por Polonca Ropret, investigadora en el Instituto para la Protección del Patrimonio Cultural de Ljubljana, Eslovenia, mientras que el conjunto de espectros desconocidos fueron medidos en otro laboratorio por Nadim C. Scherrer, de la Universidad de Arte, Conservación y Restauración de Berna, Suiza^[6,25]. Espectros de un mismo material, al haber sido medidos en laboratorios distintos, pueden presentar características espectrales diferentes.

Es por ello que en este apartado se pretende verificar la metodología implementada comprobando el correcto funcionamiento y reconocimiento espectral al analizar una serie de Espectros Raman de los que se conoce, *a priori*, a qué pigmentos se corresponden. Como se puede observar a lo largo de cada uno de los ejemplos incluidos en este apartado, se puede concluir que el sistema de identificación automático de Espectros Raman funciona correctamente.

Ejemplo 1:

En este primer ejemplo se analiza el Espectro Raman del pigmento amarillo correspondiente al **PY111** según el laboratorio de origen. El sistema reconoce el pigmento de forma correcta generando un candidato (el **SPY111**) con un Factor de Confianza elevado, del 76.42%. Este Factor de Confianza no es el ideal puesto que, como puede verse en la figura Fig. 5.17, se ha producido la pérdida de algunas Bandas Raman o débiles diferencias de intensidades relativas, lo que repercute en el valor de los *scores* (ver Fig. 5.18).

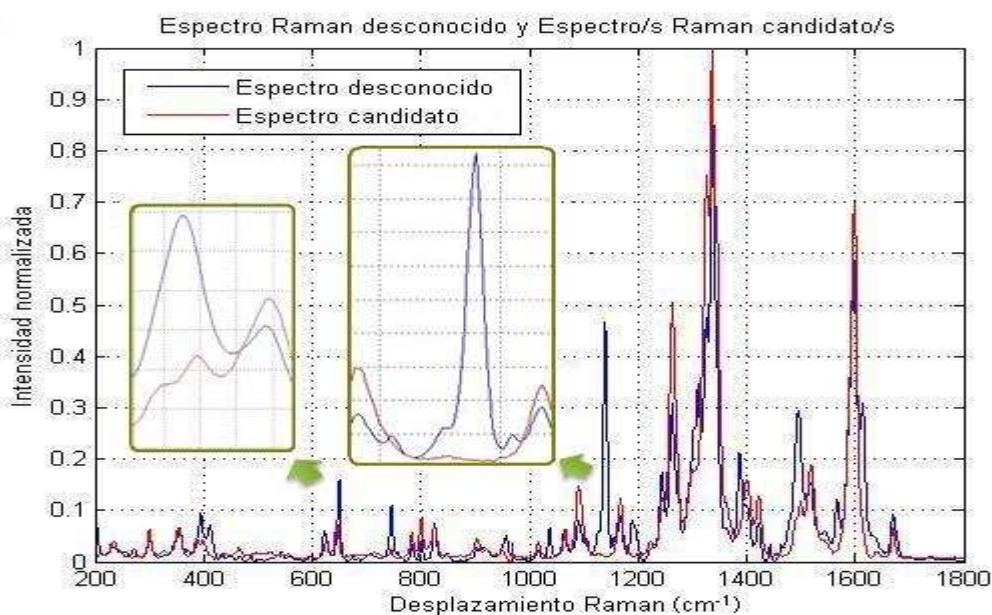


Fig. 5.17: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón SPY111

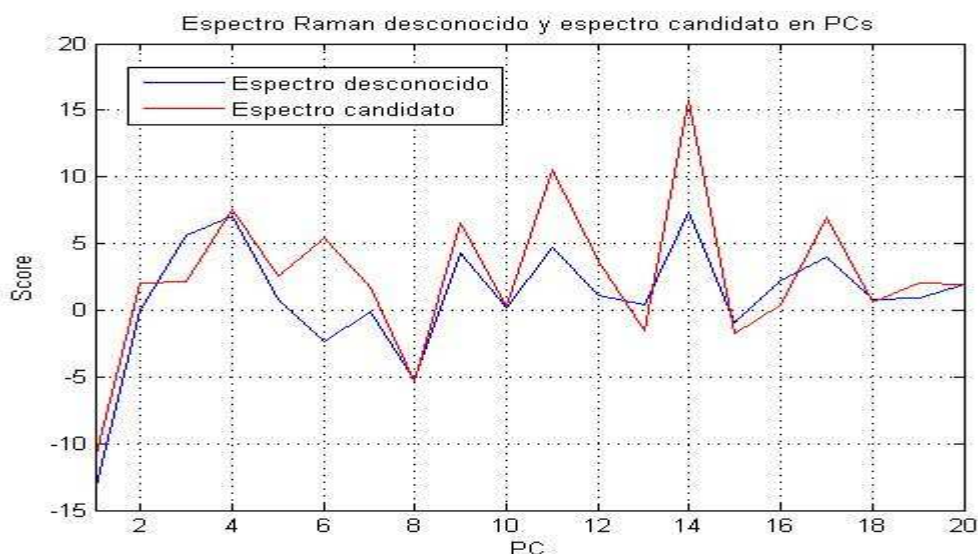


Fig. 5.18: Espectro desconocido y espectro candidato en PCs

En la *Tabla 5.9* pueden verse los valores obtenidos al aplicar la metodología implementada para la Distancia Euclídea y el coseno cuadrado.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	29,298	25,81	24,378	29,559	31,924	26,562	24,098	29,507	43,634	45,389	14,984
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	34,156	37,324	49,035	63,331	52,222	37,362	29,599	63,125	98,098	53,237	0,6907

Tabla 5.9: Distancia Euclídea y coseno cuadrado

El valor del Parámetro de Calidad referente al ruido es elevado en comparación con el umbral, lo que indica una baja componente de ruido.

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.004} \right) = 23.98dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 16.28dB$$

Como se puede ver en la siguiente figura, se ha producido una diferencia en las intensidades relativas, concretamente, las Bandas Secundarias han aumentado:

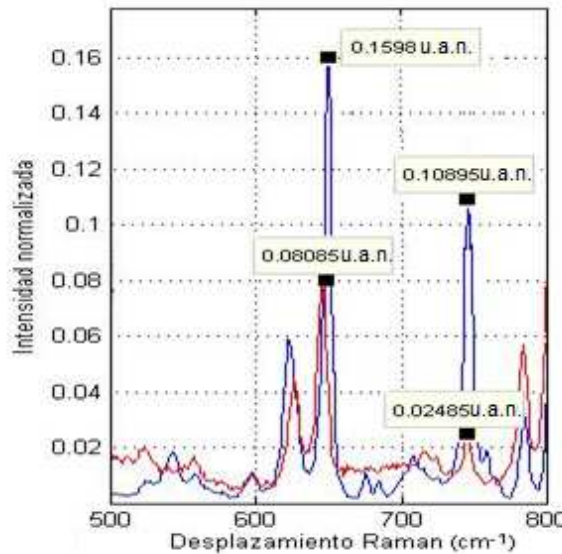


Fig. 5.19: Ampliación de la Fig. 5.17 en el rango [500, 800]cm⁻¹

$$I_2' \uparrow \rightarrow I_2' = I_2 + C \cdot \Delta I \rightarrow C = \frac{I_2' - I_2}{\Delta I} = \frac{0.1598 - 0.0809}{1 - 0.0809} = 0.086 \approx 0.09 = 9\% \rightarrow$$

$$\rightarrow k = (1 - C) = 1 - 0.09 = 0.91 \rightarrow \Delta I' = k \cdot \Delta I = 0.91 \cdot \Delta I$$

lo cual quiere decir que la relación de intensidades entre la Banda Fundamental y la Banda Secundaria de mayor amplitud, ΔI , se ha visto disminuida un 9% respecto al espectro patrón.

El conjunto de estos problemas prácticos provoca que el Factor de Confianza en la identificación se vea afectado.

Ejemplo 2:

El siguiente ejemplo se corresponde al análisis del pigmento amarillo **PY129** según el laboratorio de origen. El sistema lo identifica de forma correcta con un Factor de Confianza del 84.1%. En este caso, el valor del Factor de Confianza en la identificación viene condicionado por la diferencia en las intensidades relativas, como puede verse en la figura Fig. 5.19, modificando ligeramente el valor de los *scores* (ver Fig. 5.20).

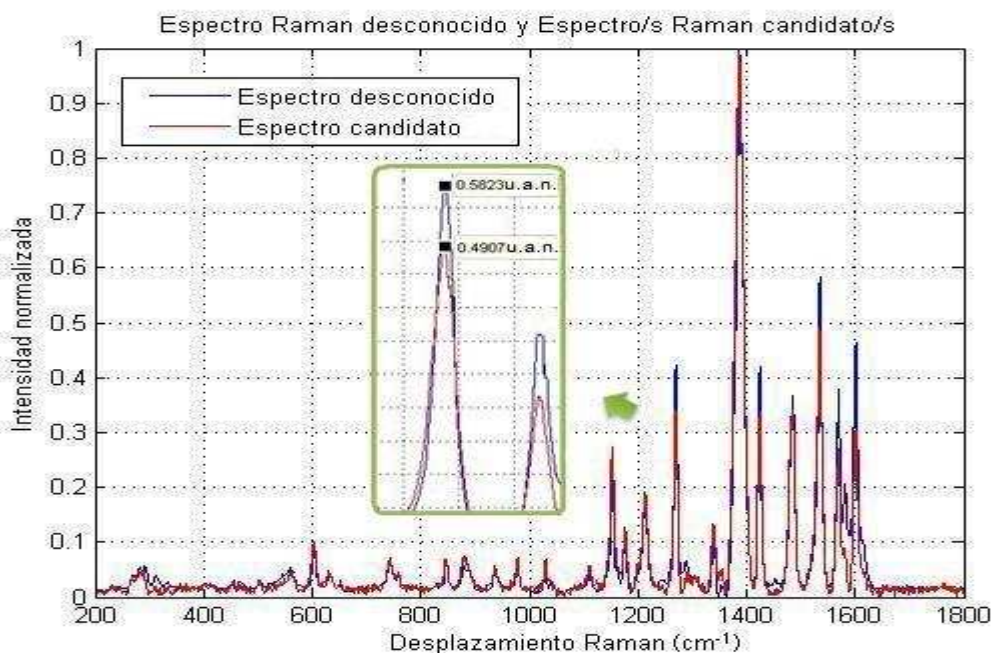


Fig. 5.19: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón SPY129

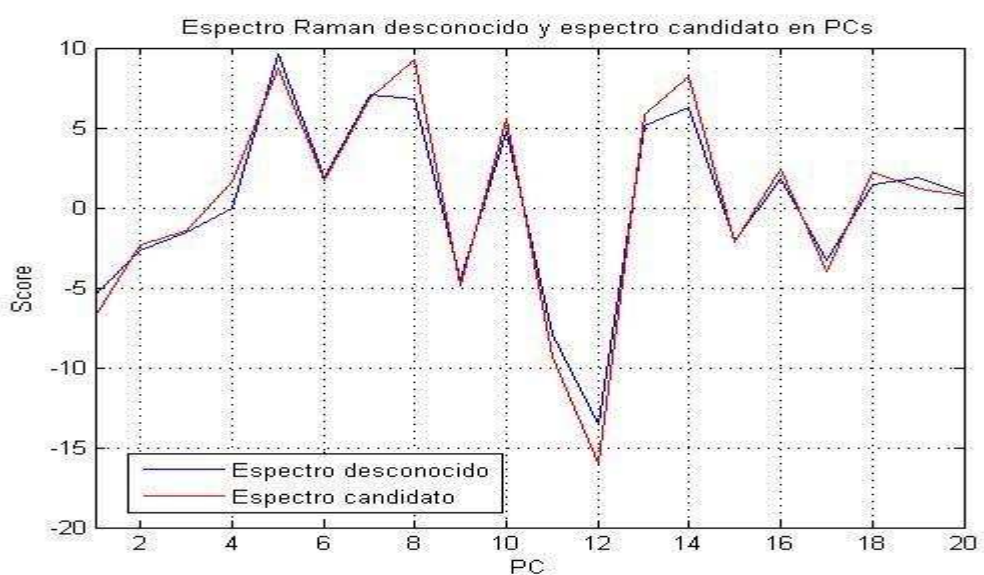


Fig. 5.20: Espectro desconocido y espectro candidato en PCs

La siguiente tabla muestra los valores de Distancia Euclídea y coseno cuadrado obtenidos para este espectro al aplicar la metodología implementada.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	38,999	36,003	34,182	41,733	41,167	33,538	29,486	38,164	46,139	47,349	37,259
	5,1651	39,332	46,512	64,224	55,961	39,489	29,92	61,657	90,454	52,813	0.8885

Tabla 5.10: Distancia Euclídea y coseno cuadrado

Si se procede al cálculo del Parámetro de Calidad, se observa que éste es elevado, indicando una baja componente de ruido, lo que permite el reconocimiento espectral.

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.0035} \right) = 24.56dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 16.28dB$$

Como ya se ha comentado, se ha producido una diferencia de intensidades relativas, concretamente, la amplitud de la Banda Secundaria ha aumentado:

$$I_2' \uparrow \rightarrow I_2' = I_2 + C \cdot \Delta I \rightarrow C = \frac{I_2' - I_2}{\Delta I} = \frac{0.5823 - 0.4907}{1 - 0.4907} = 0.179 \approx 0.18 = 18\% \rightarrow$$

$$\rightarrow k = (1 - C) = 1 - 0.18 = 0.82 \rightarrow \Delta I' = k \cdot \Delta I = 0.82 \cdot \Delta I$$

lo cual quiere decir que la relación de intensidades entre la Banda Fundamental y la Banda Secundaria de mayor amplitud, ΔI , se ha visto disminuida un 9% respecto al espectro patrón, haciendo que el Factor de Confianza se aleje del caso óptimo.

Ejemplo 3:

Este tercer ejemplo se corresponde al análisis del pigmento amarillo **PY151** según el laboratorio de origen. El sistema lo identifica correctamente con un Factor de Confianza del 80.41%. En este caso, el Factor de Confianza se ve afectado nuevamente por la diferencia en las intensidades relativas, como puede verse en la figura Fig. 5.21, modificando los *scores* (ver Fig. 5.22).

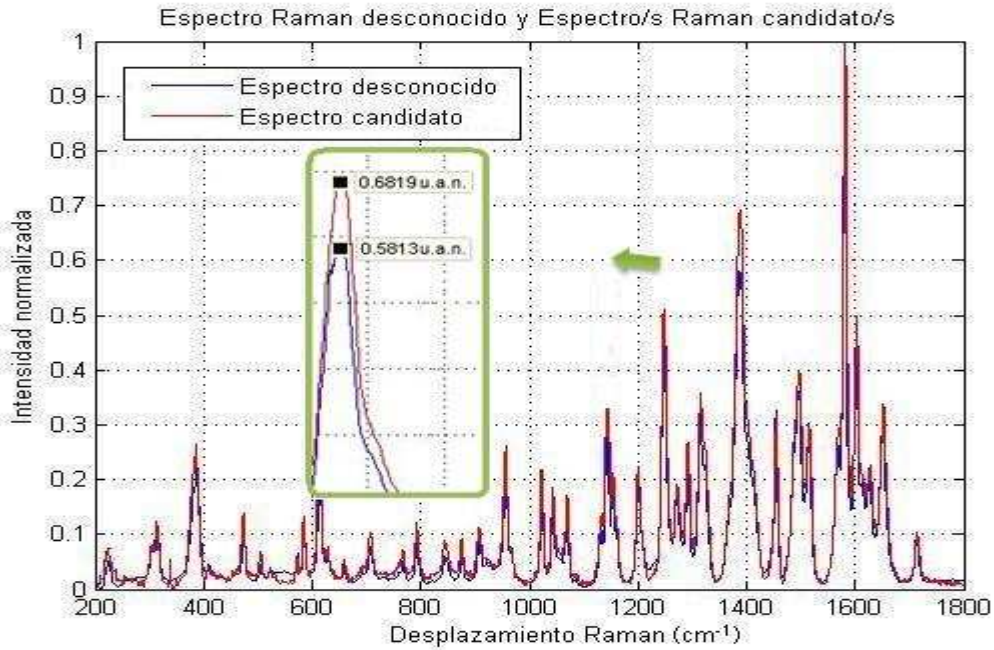


Fig. 5.21: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón SPY151

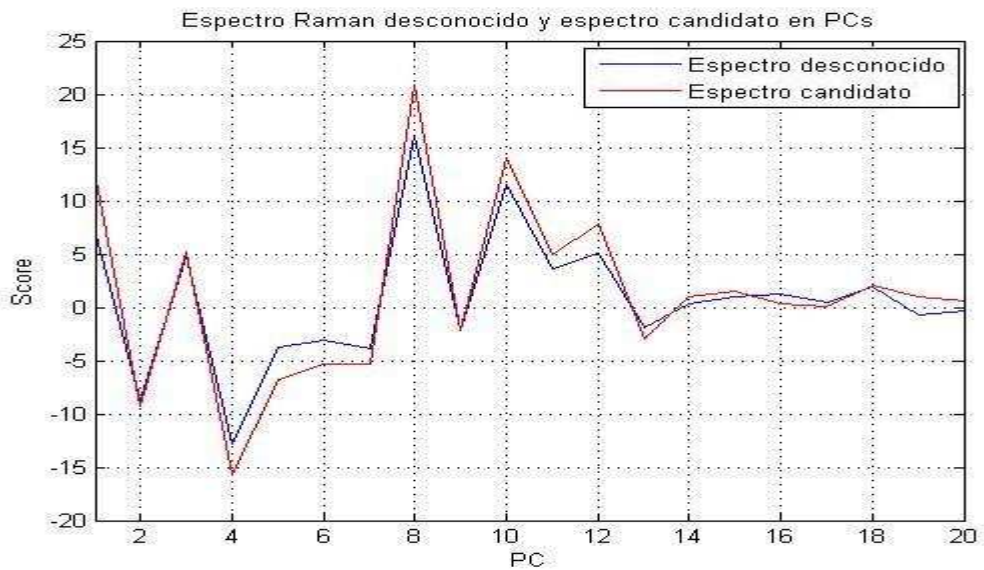


Fig. 5.22: Espectro desconocido y espectro candidato en PCs

En la *Tabla 5.11* puede verse los valores obtenidos al aplicar la metodología implementada para la Distancia Euclídea y el coseno cuadrado.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	39,619	43,193	45,102	44,122	41,929	45,507	36,265	41,764	46,675	53,576	45,087
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	40,479	51,337	9,8604	65,6	46,996	41,958	40,316	60,402	82,424	52,385	0,8702

Tabla 5.11: Distancia Euclídea y coseno cuadrado

El cálculo del Parámetro de Calidad indica que la componente de ruido es pequeña respecto al umbral establecido.

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.0016} \right) = 27.96dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 16.28dB$$

Respecto a la diferencia entre las intensidades relativas, la amplitud de la Banda Secundaria más intensa ha aumentado:

$$I_2' \uparrow \rightarrow I_2' = I_2 + C \cdot \Delta I \rightarrow C = \frac{I_2' - I_2}{\Delta I} = \frac{0.6819 - 0.5813}{1 - 0.5813} = 0.2402 \approx 24\% \rightarrow$$

$$\rightarrow k = (1 - C) = 1 - 0.24 = 0.76 \rightarrow \Delta I' = k \cdot \Delta I = 0.76 \cdot \Delta I$$

lo cual quiere decir que la relación de intensidades entre la Banda Fundamental y la Banda Secundaria de mayor amplitud, ΔI , se ha visto disminuida un 24% respecto al espectro patrón, haciendo que el Factor de Confianza en la identificación disminuya.

Ejemplo 4:

El siguiente ejemplo se corresponde al análisis del pigmento amarillo **PY154** según el laboratorio de origen. El sistema lo identifica de forma correcta con un Factor de Confianza del 75.38%. En este caso, el Factor de Confianza en la identificación no es ideal puesto que, como puede verse en la figura Fig. 5.23, se han producido ciertos problemas prácticos, como la pérdida de Bandas Secundarias o diferencia en las intensidades relativas, modificando el valor de los *scores* (ver Fig. 5.24).

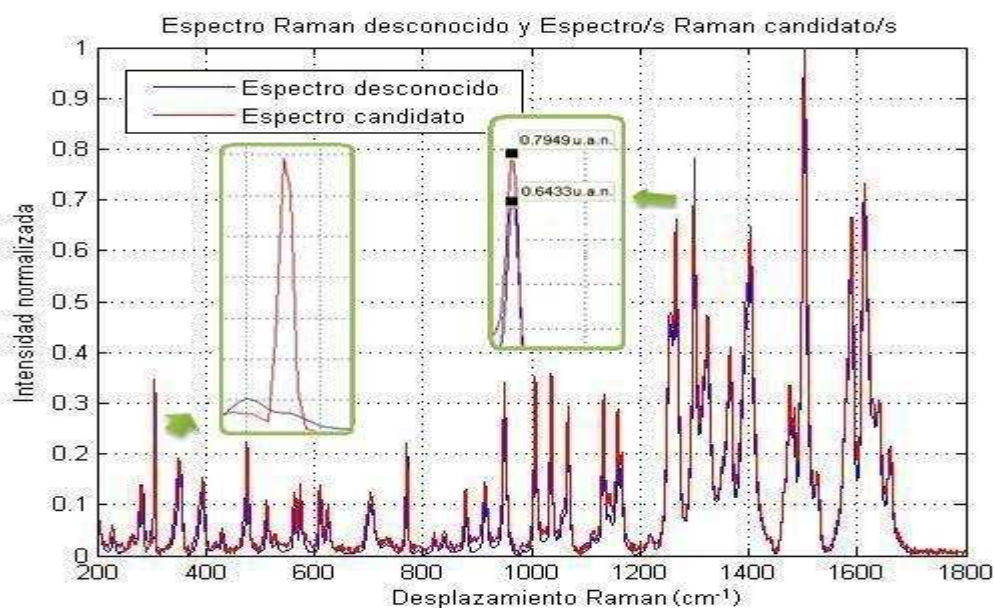


Fig. 5.23: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón SPY154

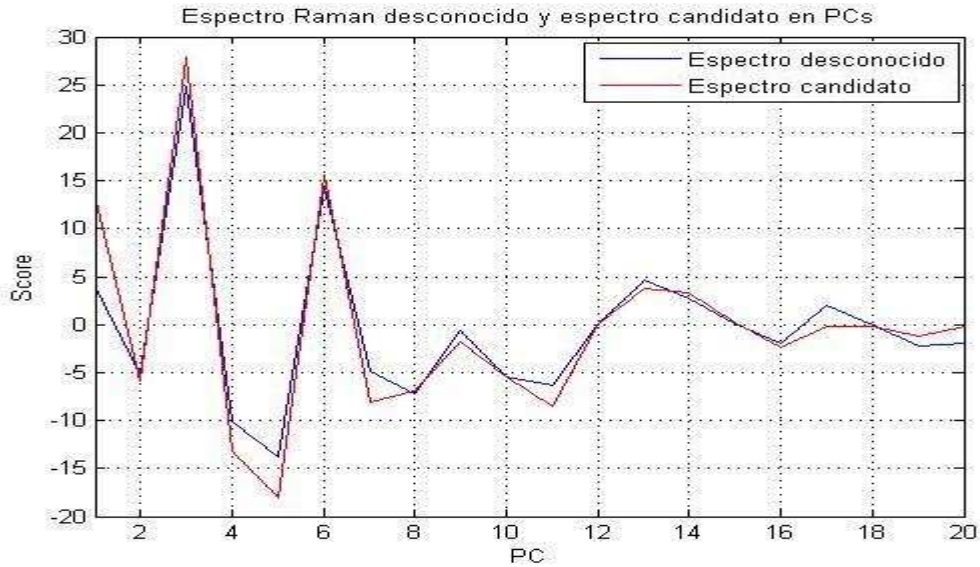


Fig. 5.24: Espectro desconocido y espectro candidato en PCs

En la *Tabla 5.12* pueden verse los valores obtenidos al aplicar la metodología implementada para la Distancia Euclídea y el coseno cuadrado.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	43,995	45,323	48,725	46,221	49,434	47,374	44,433	42,192	58,305	62,773	46,417
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	50,5	56,849	48,755	68,803	12,394	47,275	47,54	64,74	90,294	62,335	0,8561

Tabla 5.12: Distancia Euclídea y coseno cuadrado

El valor del Parámetro de Calidad referente al ruido es elevado, permitiendo que el espectro pueda ser identificado.

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.00097} \right) = 30.13dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 16.28dB$$

Pero se ha producido una diferencia de intensidades relativas, concretamente, la amplitud de la Banda Secundaria ha aumentado:

$$I_2' \uparrow \rightarrow I_2' = I_2 + C \cdot \Delta I \rightarrow C = \frac{I_2' - I_2}{\Delta I} = \frac{0.7949 - 0.6433}{1 - 0.6433} = 0.425 \approx 0.43 = 43\% \rightarrow$$

$$\rightarrow k = (1 - C) = 1 - 0.43 = 0.57 \rightarrow \Delta I' = k \cdot \Delta I = 0.57 \cdot \Delta I$$

lo cual quiere decir que la relación de intensidades entre la Banda Fundamental y la Banda Secundaria de mayor amplitud, ΔI , se ha visto disminuida un 43% respecto al espectro patrón, siendo éste el principal motivo por el que el Factor de Confianza se aleja del caso ideal.

Ejemplo 5:

Este quinto ejemplo se corresponde al análisis del pigmento amarillo *PY16* según el laboratorio de origen. El sistema lo identifica correctamente con un Factor de Confianza del 83.12%. En este caso, el Factor de Confianza se ve afectado por un desplazamiento de las Bandas Raman, como puede verse en la figura *Fig. 5.25*, que modifica la posición de los scores (ver *Fig. 5.26*).

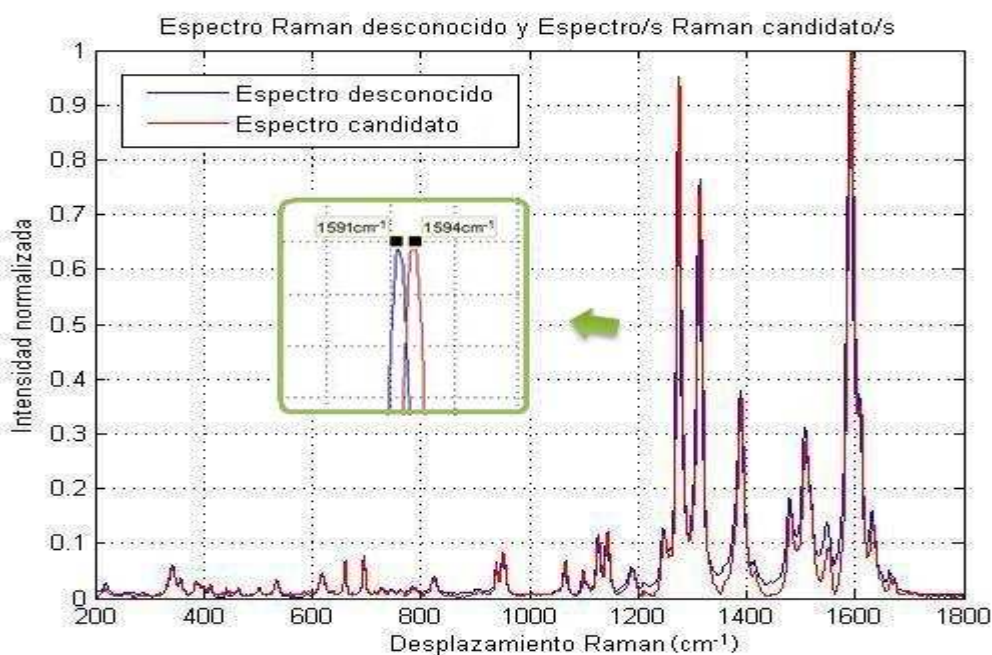


Fig. 5.25: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón SPY16

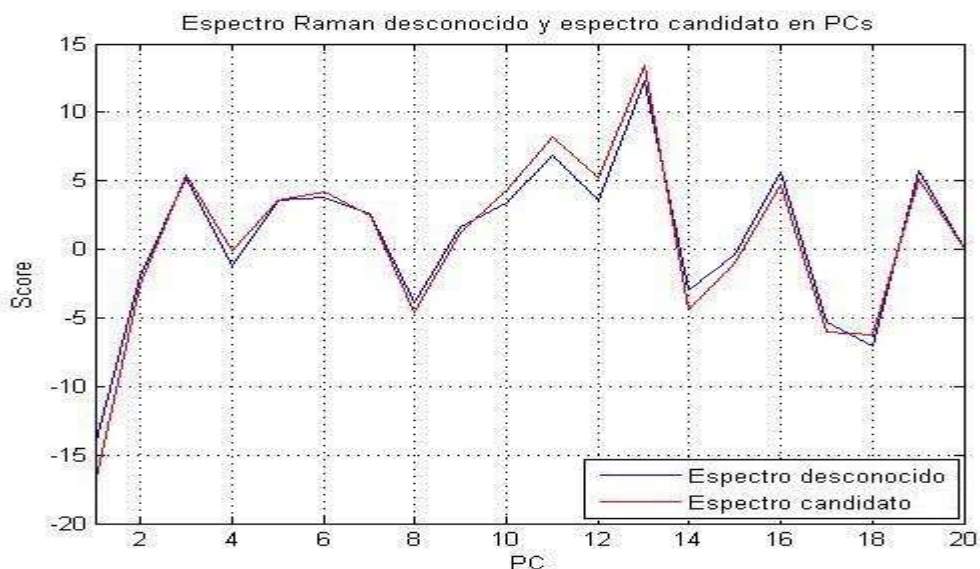


Fig. 5.26: Espectro desconocido y espectro candidato en PCs

La siguiente tabla muestra los valores de Distancia Euclídea y coseno cuadrado obtenidos para este espectro al aplicar la metodología implementada.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	37,864	4,6628	27,225	39,622	42,174	28,551	26,25	34,451	46,762	48,299	31,025
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	36,239	39,601	49,03	65,67	52,905	40,03	30,562	67,231	98,786	54,509	0,9137

Tabla 5.13: Distancia Euclídea y coseno cuadrado

El Parámetro de Calidad, al ser superior al umbral, indica que el espectro puede ser identificado.

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.0026} \right) = 25.85dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 16.28dB$$

El desplazamiento de las Bandas Raman, que ha podido observarse en la Fig. 5.25, resulta ser de $3cm^{-1}$,

$$\Delta = 1594cm^{-1} - 1591cm^{-1} = 3cm^{-1}$$

lo cual provoca que el Factor de Confianza en la identificación disminuya.

Ejemplo 6:

En este ejemplo se analiza el Espectro Raman del pigmento amarillo correspondiente al **PY73** según el laboratorio de origen. El sistema reconoce el pigmento de forma correcta generando un candidato (el **SPY73**) con un Factor de Confianza elevado, del 72.97%. En este caso, el valor del Factor de Confianza en la identificación, se ve afectado por una diferencia en las intensidades relativas (ver Fig. 5.27), lo que repercute en el valor de los scores (ver Fig. 5.28).

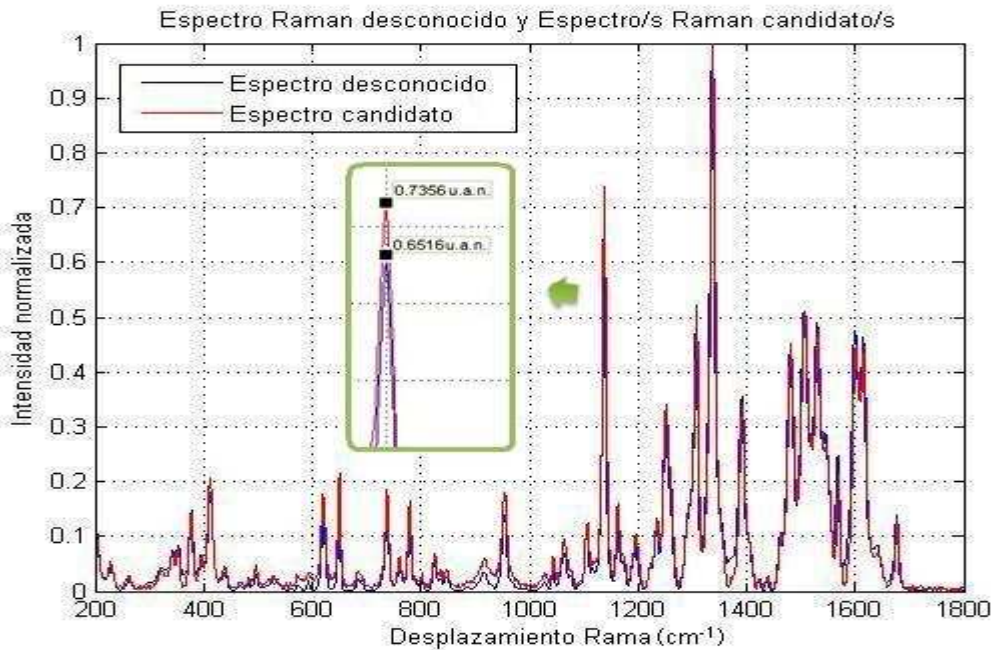


Fig. 5.27: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón SPY73

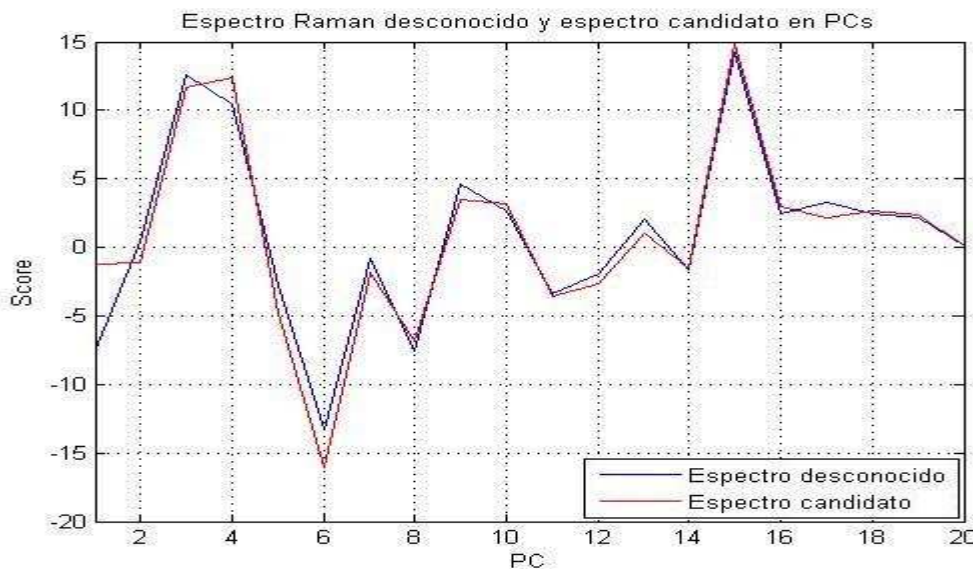


Fig. 5.28: Espectro desconocido y espectro candidato en PCs

En la *Tabla 5.12* pueden verse los valores obtenidos al aplicar la metodología implementada para la Distancia Euclídea y el coseno cuadrado.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	31,711	37,007	37,149	8,1268	32,382	37,645	33,921	38,967	51,315	54,341	36,747
	Esp 12	Esp 13	Esp 14	Esp 15	Esp 16	Esp 17	Esp 18	Esp 19	Esp 20	Esp 21	COS2
	41,84	46,71	51,834	65,044	52,913	43,944	38,281	64,328	96,366	58,673	0,8867

Tabla 5.14: Distancia Euclídea y coseno cuadrado

El Parámetro de Calidad presenta un valor elevado, lo cual indica una baja componente de ruido, permitiendo que el espectro pueda ser identificado.

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{nE_i}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.00095} \right) = 30.22dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 16.28dB$$

Además, la amplitud de la Banda Secundaria ha aumentado:

$$I_2' \uparrow \rightarrow I_2' = I_2 + C \cdot \Delta I \rightarrow C = \frac{I_2' - I_2}{\Delta I} = \frac{0.7356 - 0.6516}{1 - 0.6516} = 0.241 \approx 24\% \rightarrow$$

$$\rightarrow k = (1 - C) = 1 - 0.24 = 0.76 \rightarrow \Delta I' = k \cdot \Delta I = 0.76 \cdot \Delta I$$

lo cual quiere decir que la relación de intensidades entre la Banda Fundamental y la Banda Secundaria de mayor amplitud, ΔI , se ha visto disminuida un 24% respecto al espectro patrón, siendo este factor el responsable principal de que disminuya el Factor de Confianza.

Ejemplo 7:

El presente ejemplo recoge el análisis sobre el espectro del pigmento amarillo **SPY129B**, medido por Polonca Ropret sobre una obra artística^[6], en el cual, por efecto de un mal filtrado a la hora de eliminar la fluorescencia, se produjo una anomalía en el rango $[1280, 1370]cm^{-1}$, como puede observarse en las figuras Fig. 5.35 y Fig. 5.36. Esta diferencia provoca un movimiento en los *scores* que dificulta el reconocimiento espectral (ver Fig. 5.37). No obstante, el sistema proporciona el espectro candidato correcto con un Factor de Confianza en la identificación del 89.87%.

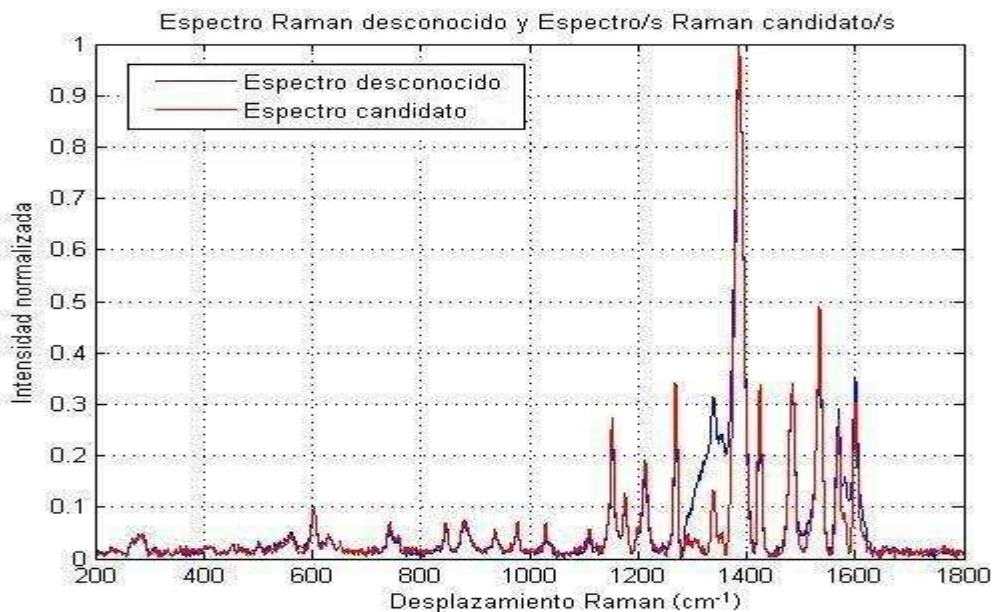


Fig. 5.35: Espectro desconocido y espectro candidato generado por el sistema implementado correspondiente al espectro patrón SPY129

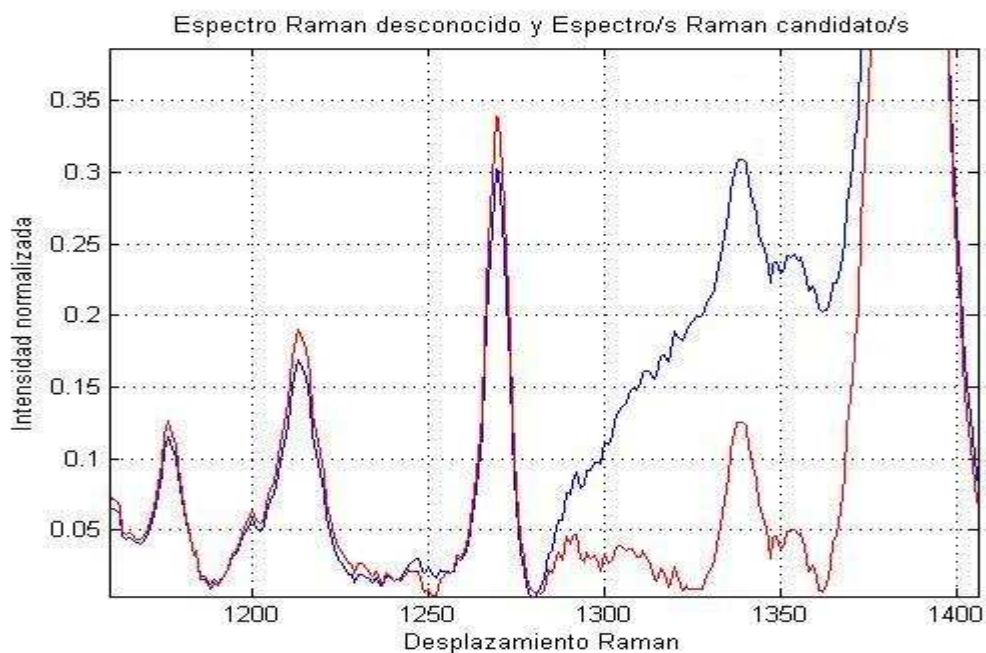


Fig. 5.36: Ampliación del rango que presenta diferencias

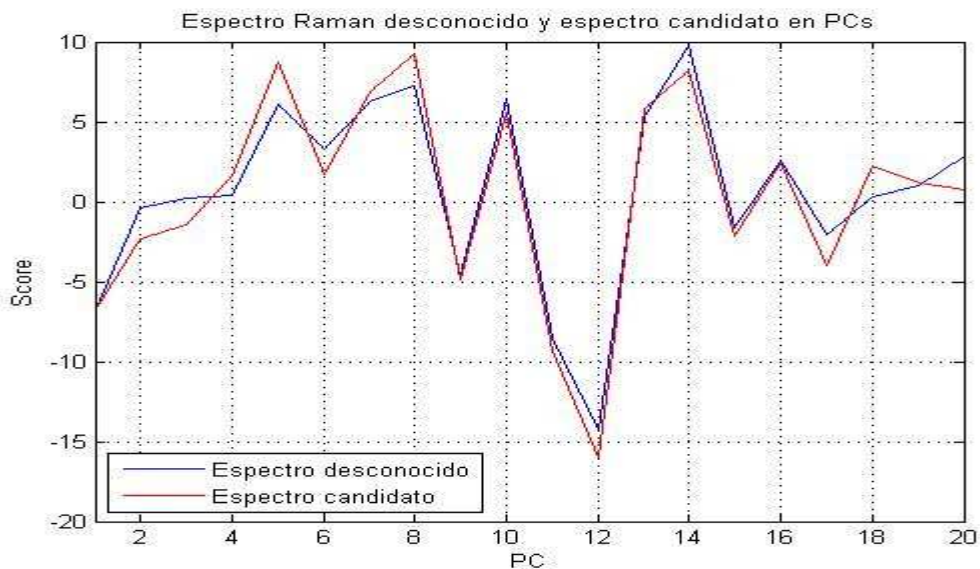


Fig. 5.37: Espectro desconocido y espectro candidato en PCs

En la *Tabla 5.15* puede verse los valores obtenidos al aplicar la metodología implementada para la Distancia Euclídea y el coseno cuadrado.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	Esp 11
Espectro desconocido	39,5	36,831	35,403	41,771	42,015	35,206	30,951	38,248	47,864	48,086	35,968
	6,438	40,158	46,908	63,243	54,535	41,595	30,726	61,784	93,221	53,522	0,8799

Tabla 5.15: Distancia Euclídea y coseno cuadrado

El cálculo del Parámetro de Calidad indica que el espectro puede ser identificado, puesto que es superior al valor umbral, implicando que la componente de ruido es baja:

$$Q_{dB}^{desconocido} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{normalización} = 10 \cdot \log \left(\frac{1}{0.0048} \right) = 23.19dB$$

$$Q_{dB}^{desconocido} > Q_{dB}^{umbral} = 16.28dB$$

Así, pues, la diferencia que presenta el Espectro Raman *SPY129B* respecto al pigmento patrón *SPY129* en el rango $[1280, 1370]cm^{-1}$ es, principalmente, la responsable de la disminución en el valor del Factor de Confianza en la identificación.

A lo largo de este capítulo, se ha llevado a cabo un análisis de la metodología desarrollada en diferentes casos experimentales de identificación de Espectros Raman de pigmentos. Dicho análisis se ha efectuado en dos etapas: En primer lugar, el estudio se ha realizado sobre una paleta básica de Espectros Raman de pigmentos inorgánicos, con pocas bandas, obtenidos sobre obras de arte, todos ellos medidos en el Laboratorio de Espectroscopía Raman de la UPC. En segundo lugar, se ha estudiado el comportamiento del sistema implementado en Espectros Raman de pigmentos orgánicos de una determinada tonalidad, muy parecidos y con un elevado número de bandas, medidos en diferentes laboratorios.

En base a los resultados, se comprueba que la metodología de identificación se comporta correctamente, tanto en espectros de pigmentos orgánicos como en espectros de pigmentos inorgánicos, siendo capaz de realizar el reconocimiento espectral según los límites establecidos en el *Capítulo 4* sobre la librería de referencia simulada. De esta forma se muestra la extensibilidad del sistema implementado a la identificación de cualquier tipo de material.

Finalmente, del trabajo desarrollado durante la realización del presente Proyecto Final de Carrera se han extraído una serie de conclusiones que se recogen en forma de resumen en el siguiente capítulo.

6. RESUMEN Y CONCLUSIONES

El objetivo principal del presente Proyecto Final de Carrera ha sido el desarrollo de un sistema automático para la identificación de Espectros Raman de pigmentos, con el fin de agilizar dicho proceso, y evitar la componente subjetiva que pudiera comportar la tarea llevada a cabo por un analista.

El sistema implementado permite la identificación de espectros sin necesidad de localizar la posición de sus bandas Raman, valiéndose de la expresión reducida del espectro completo proporcionada por el Análisis por Componentes Principales, la cual es una representación simplificada que evita la redundancia sin que se produzca pérdida de información. De esta forma se consigue un ahorro significativo de tiempo y recursos.

El *ACP normado* ha proporcionado mejores resultados que el *ACP no normado*, ya que trata a todas las variables de forma uniforme, lo cual lo convierte en una práctica óptima para resaltar similitudes y diferencias entre un espectro desconocido y los espectros de referencia.

El coseno cuadrado ha demostrado ser una herramienta válida para evaluar la calidad de la representación espectral en el nuevo espacio transformado generado por el ACP, ya que la representación de un espectro desconocido puede verse afectada al proyectarse sobre el espacio reducido generado por la librería de referencia.

La Distancia Euclídea ha resultado una herramienta óptima para cuantificar el parecido entre espectros. Partiendo únicamente de esta herramienta, aplicada sobre la expresión de los Espectros Raman en el espacio de PCs, se han definido unos criterios que permiten el reconocimiento espectral. Cuando un espectro desconocido cumple alguno de dichos criterios, el sistema lo genera como candidato, asociándole un parámetro relacionado con el grado de confiabilidad en la identificación, en función del parecido espectral.

En definitiva, la implementación del sistema se ha llevado a cabo mediante la técnica de reducción dimensional del ACP y la Distancia Euclídea, resultando una metodología computacionalmente simple, objetiva, rápida y eficaz.

Para llevar a cabo el estudio del comportamiento de la identificación en el nuevo espacio de PCs se ha simulado una librería espectral y se han hecho una serie de modificaciones sobre sus características espectrales relacionadas con problemas prácticos que se dan en espectros reales: ruido *shot*, reducción de la anchura de sus bandas, diferencia en las intensidades relativas, pérdida de bandas y desplazamiento de sus bandas. Este análisis ha permitido establecer los límites que soporta el sistema diseñado cuando se producen estos efectos en un caso real:

- Se ha comprobado que existe una relación proporcional entre la amplitud del ruido y la variación que sufre la representación del espectro en el espacio transformado, lo cual permite acotarla en función de la expresión de la librería en el espacio de PCs. El límite se establece para cuando el sistema deja de generar espectros candidatos, no generando candidatos incorrectos.
- Por efecto del filtrado en la etapa de pretatado espectral, puede darse una reducción de la anchura de las bandas. Se ha comprobado que el sistema implementado permite identificar espectros con una reducción máxima del 60% respecto al espectro patrón correspondiente.
- Si los espectros son medidos en condiciones diferentes se puede dar una variación en las intensidades relativas de las bandas Raman, provocando que las bandas secundarias o bien disminuyan, efecto que tiene poca repercusión en la identificación, o bien que aumenten, hecho que provoca una disminución en la confiabilidad de la identificación y puede generar otros candidatos incorrectos.
- Por la misma causa descrita en el punto anterior, se puede producir una pérdida de bandas secundarias, lo cual, en el caso límite, podría implicar la identificación con sólo la Banda Fundamental. Se ha comprobado que el resultado de la identificación en dicho caso es siempre favorable, convirtiendo la Banda Fundamental en la banda más importante para el reconocimiento espectral llevado a cabo por el sistema implementado.
- Debido a una calibración errónea de los equipos de medida del laboratorio puede producirse un desplazamiento de las Bandas Raman. Este efecto incrementa el número de candidatos generados por el sistema. Se ha comprobado que el desplazamiento hasta el cual el sistema genera el candidato correcto es función de la separación de Bandas Fundamentales de la librería de referencia.

- La aparición de bandas adicionales en espectros desconocidos ha motivado el estudio para la identificación de mezclas de pigmentos a partir de su creación en el espacio de PCs, estrategia que ha resultado apropiada puesto que aprovecha los parámetros obtenidos sin necesidad de repetir el ACP y se obtiene una identificación correcta.

Por último, se ha comprobado el correcto funcionamiento de la metodología para la identificación de Espectros Raman de pigmentos en casos experimentales, es decir, en espectros medidos sobre obras de arte, presentando una serie de ejemplos y justificando los resultados a partir de las conclusiones extraídas sobre la librería simulada.

Por tanto, la metodología implementada resulta ser útil tanto para espectros de pigmentos inorgánicos, con pocas bandas y en posiciones distintas, como para espectros de pigmentos orgánicos, muy similares entre sí y con un elevado número de bandas situadas en posiciones muy cercanas unas a otras. Este hecho muestra la versatilidad del sistema implementado, pudiendo ser extrapolado a la identificación de cualquier tipo de material.

BIBLIOGRAFÍA

Richard L. McCreery, “*Raman Spectroscopy for Chemical Analysis*”, Wiley-Interscience, John Wiley & Sons, Inc, 2000, ISBN 0471252875

Ewen Smith, Geoffrey Dent, “*Modern Raman Spectroscopy – A Practical Approach*”, John Wiley & Sons, Ltd, 2005, ISBN 0471496685

Daniel Peña, “*Análisis de datos multivariantes*”, McGraw Hill, 2002, ISBN: 8448136101

Mireia Castanys Tutzó, “*Estudi, anàlisi, disseny i implementació de tècniques de reconeixement d’Espectres Raman. Aplicació a la identificació de materials pictòrics*”, 2009, Tesis doctoral. *Departament de Teoria del Senyal i Comunicacions (UPC)*

Seymour Lipschutz, “*Linear Algebra*”, McGraw Hill, 1992, ISBN 8476157584

REFERENCIAS

- [1] Joaquín Fargas, “*El encuentro del arte, la ciencia y la tecnología*” Razón y palabra, 2008, ISSN 1605-4806, N° 65
- [2] Enrico Ciliberto, “*Modern Analytical Methods in Art and Archaeology*”, Chemical Analysis Series, Volume 155, John Wiley & Sons, Inc.
- [3] Howell G. M. Edwards, “*Probing history with Raman spectroscopy*”, i-Section Tutorial Review, The analyst
- [4] Michel Pastoureau, “*Blue: The History of a Color*”, Princenton University Press
- [5] Franco Cariati, “*Raman Spectroscopy*”, Modern Analytical Methods in Art and Archaeology, Chemical Analysis Series, Volume 155, John Wiley & Sons, Inc.
- [6] Polanca Ropret, S.A. Centeno, P. Bukovec, “*Raman identification of yellow synthetic organic pigments in modern and contemporary paintings: Reference spectra and case studies*”, Spectrochimica Acta Part A

- [7] M. C. M. Grimbergen, “*Signal-to-Noise Contribution of Principal Component Loads in Reconstructed Near-Infrared Raman Tissue Spectra*” Volume 64, Number 1, Applied Spectroscopy, American Chemical Society
- [8] M. J. Pelletier, “*Introduction to Applied Raman Spectroscopy*”, Analytical Applications of Raman Spectroscopy, Ed. M. J. Pelletier, 1999
- [9] M. J. Pelletier, “*Quantitative Analysis Using Raman Spectrometry*”, Volume 57, Number 1, Applied Spectroscopy
- [10] Peter Vandenabeele, “*A decade of Raman Spectroscopy in Art and Archaeology*” Volume 107, Number 3, Chemical Reviews, American Chemical Society
- [11] Gonzalo Barja Becker, “*Diseño e implementación de una librería documentada de espectros de pigmentos con espectroscopía Raman. Consideraciones y problemas experimentales*”, Proyecto Final de Carrera, UPC
- [12] A.M.C. Davies, “*Back to basics: the principles of principal component analysis*”, Spectroscopy Europe
- [13] Lindsay I Smith, “*A tutorial on Principal Components Analysis*”, February 26, 2002
- [14] Peter Vandenabeele, “*Evaluation of a Principal Components-Based Searching Algorithm for Raman Spectroscopic Identification of Organic Pigments in 20th Century Artwork*” Volume 55, Number 5, Applied Spectroscopy, American Chemical Society
- [15] Salvador Carrasco Arroyo, “*Estadística Multivariada*”, Capítulo 2: Análisis de Componentes Principales, Universitat Politècnica de València
- [16] Mike J. Baxter, “*Modern Analytical Methods in Art and Archeology*”, Chapter 20: Data handling and statistical analysis, Chemical Analysis Series, Volume 155, John Wiley & Sons, Inc.
- [17] Richard L. McCreery, “*Raman Spectroscopy for Chemical Analysis*” Chapter 4: “*Signal-to-noise in Raman Spectroscopy*”, Wiley-Interscience
- [18] Max Doerner, “*Los materiales de pintura y su empleo en el arte*”. Editorial Reverté, 1998
- [19] “*Artists’ Pigments: A handbook of their history and characteristics*”, Vol. 3. Elisabeth West Fitzhugh, Editor, 1997

- [20] W. Stanley, J.W. Mayer, “*The science of paintings*”. Springer-Verlag New York, Inc, 2000
- [21] Ralph Mayer, “*Materiales y técnicas del arte*”. Tursen Hermann Blume Ediciones, 1993
- [22] “*Artists’ Pigments: A handbook of their history and characteristics*”, Vol. 2. Ashok Roy, Editor, 1997
- [23] Nicholas Eastaugh et al, “*The Pigment Compendium*”, Elsevier Butterworth-Heinemann, UK, 2004
- [24] R.J.H. Clark, “*Synthesis, Structural Characterisation and Raman Spectroscopy of the Inorganic Pigments Lead Tin Yellow: their Identification on Medieval Painting and Manuscripts*”. J. Chem. Soc. Dalton Trans, 1995, pp. 2577-2583
- [25] N. C. Scherrer, S. Zumbuehl, F. Delavy, A. Fritsch, R. Kuehnen, “*Synthetic organic pigments of the 20th and 21st century relevant to artist’s paints: Raman spectra reference collection*”, Spectrochimica Acta Part A-Molecular and Biomolecular Spectroscopy, 73 (3), 505-524, 2009

ANEXO A1: Resultados obtenidos en la reducción de la anchura de las bandas: Distancia Euclídea, Candidatos y Factores de Confianza

Se presentan, a continuación, los resultados obtenidos en el apartado de la influencia de la anchura de las Bandas Raman sobre la librería espectral de referencia simulada. Dichos resultados incluyen los valores para la Distancia Euclídea cuando un espectro de la librería ha visto reducido la anchura de sus bandas (a 5/6, a 4/6, a 3/6, a 2/6, a 1/6 y en el caso límite, a 0, del anchura original de sus bandas), así como los espectros candidatos generados por el sistema y sus respectivos Factores de Confianza.

DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10
Esp 1 BW'=5/6-BW	0,9875	14,117	11,118	15,01	12,022	16,447	8,4546	12,719	11,778	9,0922
Esp 1 BW'=4/6-BW	2,0296	13,857	10,426	14,752	11,235	16,141	8,0573	12,434	11,004	9,2745
Esp 1 BW'=3/6-BW	3,165	13,497	9,6051	14,406	10,317	15,805	7,6221	12,064	10,148	9,406
Esp 1 BW'=2/6-BW	4,4687	13,027	8,6253	13,968	9,251	15,445	7,1258	11,603	9,206	9,4809
Esp 1 BW'=1/6-BW	6,0074	12,511	7,5813	13,502	8,1458	15,106	6,6564	11,133	8,2787	9,5711
Esp 1 BW'=0-BW	7,0394	12,252	7,0478	13,273	7,5893	14,955	6,5025	10,93	7,8324	9,7058
Esp 2 BW'=5/6-BW	12,997	1,8577	12,974	13,003	13,747	15,843	13,01	15,165	12,721	14,164
Esp 2 BW'=4/6-BW	11,714	3,7783	11,383	12,52	12,147	15,56	11,429	13,981	11,507	12,847
Esp 2 BW'=3/6-BW	10,508	5,754	9,8335	12,255	10,572	15,346	9,868	12,884	10,35	11,65
Esp 2 BW'=2/6-BW	9,4845	7,7331	8,4424	12,274	9,1384	15,218	8,4288	11,953	9,3243	10,682
Esp 2 BW'=1/4-BW	9,0933	8,6626	7,8756	12,398	8,5471	15,196	7,8232	11,59	8,9087	10,329
Esp 2 BW'=1/6-BW	8,8121	9,4678	7,4446	12,573	8,0955	15,201	7,3503	11,323	8,5931	10,084
Esp 2 BW'=0-BW	8,5974	10,246	7,0908	12,808	7,7256	15,232	6,9502	11,11	8,3341	9,9031
Esp 3 BW'=5/6-BW	11,272	14,251	0,6855	15,766	9,5661	16,977	10,995	12,486	9,9514	10,168
Esp 3 BW'=4/6-BW	10,739	13,878	1,4877	15,323	9,2065	16,686	10,237	12,14	9,6637	9,9488
Esp 3 BW'=3/6-BW	10,129	13,445	2,4353	14,817	8,7626	16,324	9,3473	11,745	9,2823	9,7346
Esp 3 BW'=2/6-BW	9,4926	12,993	3,5177	14,291	8,2755	15,927	8,3738	11,347	8,8438	9,5729
Esp 3 BW'=1/6-BW	8,9768	12,628	4,5469	13,86	7,8746	15,593	7,5202	11,046	8,4707	9,5199
Esp 3 BW'=0-BW	8,7779	12,488	5,0131	13,689	7,7226	15,461	7,1635	10,939	8,3259	9,5311
Esp 4 BW'=4/6-BW	13,7	13,04	14,274	2,2796	12,656	15,134	14,153	15,934	14,94	15,069
Esp 4 BW'=4/6-BW	12,274	12,616	12,421	4,5732	10,829	14,979	12,268	14,481	13,162	13,623
Esp 4 BW'=3/6-BW	10,959	12,358	10,625	6,8543	9,1715	14,949	10,43	13,16	11,451	12,301
Esp 4 BW'=2/6-BW	9,83	12,244	8,9818	9,0667	7,8546	15,012	8,7333	12,056	9,9063	11,182
Esp 4 BW'=1/6-BW	9,0421	12,257	7,7358	10,996	7,115	15,136	7,4324	11,307	8,7557	10,411
Esp 4 BW'=0-BW	8,7715	12,308	7,2667	11,888	6,9513	15,223	6,9383	11,053	8,3292	10,145
Esp 5 BW'=5/6-BW	12,185	14,904	9,5206	14,271	0,7326	16,207	10,475	11,523	11,451	12,577
Esp 5 BW'=4/6-BW	11,545	14,41	9,0947	13,967	1,5925	16,074	9,8194	11,344	10,839	12,11

Identificación automática de Espectros Raman de pigmentos mediante ACP

Esp 5 BW'=3/6-BW	10,793	13,848	8,5627	13,664	2,6135	15,882	9,031	11,127	10,116	11,552
Esp 5 BW'=2/6-BW	9,9819	13,264	7,9727	13,411	3,785	15,654	8,1592	10,916	9,3382	10,954
Esp 5 BW'=1/6-BW	9,2926	12,788	7,4795	13,286	4,9025	15,465	7,4005	10,787	8,6824	10,458
Esp 5 BW'=0-BW	9,013	12,603	7,2886	13,275	5,41	15,394	7,0887	10,758	8,4197	10,263
Esp 6 BW'=5/6-BW	14,867	15,14	15,137	14,565	14,327	2,3347	15,017	17,364	13,438	16,901
Esp 6 BW'=4/6-BW	12,993	14,148	13,009	13,817	12,346	4,7845	12,796	15,538	11,523	14,969
Esp 6 BW'=3/6-BW	11,206	13,285	10,89	13,272	10,44	7,3227	10,579	13,804	9,7447	13,113
Esp 6 BW'=2/6-BW	9,7125	12,659	8,9807	13,014	8,8241	9,8392	8,5734	12,343	8,3597	11,524
Esp 6 BW'=1/6-BW	8,8049	12,368	7,6639	13,051	7,8272	11,955	7,1869	11,411	7,6691	10,492
Esp 6 BW'=0-BW	8,5512	12,323	7,2324	13,143	7,5472	12,847	6,7349	11,122	7,5471	10,167
Esp 7 BW'=5/6-BW	8,9762	14,319	11,178	15,793	10,67	16,932	0,7518	10,782	10,745	11,709
Esp 7 BW'=4/6-BW	9,0466	13,989	10,609	15,462	10,251	16,653	1,5553	10,963	10,26	11,598
Esp 7 BW'=3/6-BW	9,0302	13,568	9,883	15,032	9,7102	16,312	2,4442	11,058	9,6914	11,352
Esp 7 BW'=2/6-BW	8,9071	13,063	8,9727	14,5	9,0177	15,913	3,4654	11,054	9,0392	10,955
Esp 7 BW'=1/6-BW	8,7203	12,561	7,9738	13,942	8,2498	15,519	4,5924	10,984	8,4056	10,479
Esp 7 BW'=0-BW	8,6314	12,329	7,4506	13,664	7,8522	15,336	5,2295	10,95	8,118	10,233
Esp 8 BW'=5/6-BW	12,05	15,466	11,64	16,593	10,198	18,371	9,4877	1,5928	12,924	10,227
Esp 8 BW'=4/6-BW	11,144	14,542	10,484	15,707	8,7186	17,552	8,4923	3,3058	11,638	9,7655
Esp 8 BW'=3/6-BW	10,261	13,667	9,3442	14,856	7,3331	16,763	7,6207	5,1195	10,373	9,4681
Esp 8 BW'=2/6-BW	9,4455	12,906	8,2681	14,1	6,3088	16,046	6,9404	7,0005	9,2099	9,3786
Esp 8 BW'=1/6-BW	8,7752	12,335	7,3387	13,529	6,1035	15,464	6,5007	8,8821	8,2848	9,5081
Esp 8 BW'=0-BW	8,4832	12,092	6,8959	13,309	6,554	15,197	6,3813	10,073	7,9012	9,7133
Esp 9 BW'=5/6-BW	11,778	13,793	9,6172	16,184	11,186	15,277	10,322	13,533	1,0221	12,969
Esp 9 BW'=4/6-BW	11,016	13,556	9,0508	15,574	10,346	15,191	9,4094	12,85	2,136	12,254
Esp 9 BW'=3/6-BW	10,228	13,248	8,452	14,928	9,4449	15,112	8,4439	12,15	3,3586	11,525
Esp 9 BW'=2/6-BW	9,4622	12,887	7,8386	14,286	8,5374	15,061	7,4901	11,491	4,6945	10,825
Esp 9 BW'=1/6-BW	8,8396	12,531	7,2825	13,738	7,78	15,057	6,7442	11,008	6,0346	10,254
Esp 9 BW'=0-BW	8,5835	12,349	7,0081	13,491	7,4805	15,069	6,515	10,875	6,7228	10,013
Esp 10 BW'=5/6-BW	8,3766	14,93	9,403	16,005	11,887	18,101	10,78	10,455	12,5	1,3226
Esp 10 BW'=4/6-BW	7,962	14,3	8,4645	15,384	10,771	17,368	9,7873	10,189	11,294	2,7619
Esp 10 BW'=3/6-BW	7,7021	13,663	7,622	14,755	9,64	16,645	8,7579	10,04	10,078	4,3206
Esp 10 BW'=2/6-BW	7,6569	13,053	6,9731	14,154	8,5694	15,979	7,7528	10,054	8,9451	5,9898
Esp 10 BW'=1/6-BW	7,8559	12,559	6,6504	13,673	7,7366	15,465	6,9392	10,261	8,0994	7,6397
Esp 10 BW'=0-BW	8,0576	12,363	6,6284	13,483	7,4287	15,267	6,6285	10,449	7,8056	8,4939

Espectros candidatos y Factores de Confianza:

Espectro desconocido	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de confianza
Esp 1 BW'=5/6-BW	1	0,88839	7	0,044419	-	-
Esp 1 BW'=4/6-BW	1	0,7706	7	0,089317	-	-
Esp 1 BW'=3/6-BW	1	0,76802	7	0,44133	2	0,010702

Anexo A1: Resultados en la reducción de la anchura de las bandas

Esp 1 BW'=2/6-BW	1	0,546	3	0,1237	7	0,27604
Esp 1 BW'=1/6-BW	1	0,4081	3	0,25303	5	0,19741
Esp 1 BW'=0-BW	1	0,30642	3	0,3056	5	0,25224
Esp 2 BW'=5/6-BW	2	0,86384	4	0,046965	-	-
Esp 2 BW'=4/6-BW	2	0,72307	4	0,082349	-	-
Esp 2 BW'=3/6-BW	2	0,62535	6	0,00078586	-	-
Esp 2 BW'=2/6-BW	2	0,4332	7	0,38221	3	0,38121
Esp 2 BW'=1/6-BW	3	0,2665	5	0,20237	1	0,13176
Esp 2 BW'=0-BW	3	0,30135	5	0,23881	1	0,15292
Esp 3 BW'=5/6-BW	3	0,93246	5	0,057472	9	0,019507
Esp 3 BW'=4/6-BW	3	0,85342	5	0,092899	9	0,047857
Esp 3 BW'=3/6-BW	3	0,8215	5	0,35774	2	0,014562
Esp 3 BW'=2/6-BW	3	0,65341	5	0,18463	7	0,17494
Esp 3 BW'=1/6-BW	3	0,552	7	0,25905	5	0,22413
Esp 3 BW'=0-BW	3	0,50607	5	0,2391	1	0,13512
Esp 4 BW'=5/6-BW	4	0,85158	2	0,15094	6	0,014611
Esp 4 BW'=4/6-BW	4	0,70223	2	0,17857	6	0,024703
Esp 4 BW'=3/6-BW	4	0,55371	2	0,1954	6	0,026637
Esp 4 BW'=2/6-BW	5	0,202	7	0,11273	-	-
Esp 4 BW'=1/6-BW	3	0,23781	5	0,29897	7	0,2677
Esp 4 BW'=0-BW	1	0,13576	3	0,28402	5	0,3151
Esp 5 BW'=5/6-BW	5	0,92557	3	0,032742	-	-
Esp 5 BW'=4/6-BW	5	0,8382	3	0,076011	-	-
Esp 5 BW'=3/6-BW	3	0,13006	5	0,73448	-	-
Esp 5 BW'=2/6-BW	5	0,61546	3	0,19001	7	0,17106
Esp 5 BW'=1/6-BW	5	0,51697	7	0,27084	3	0,26306
Esp 5 BW'=0-BW	5	0,46696	7	0,30156	3	0,28186
Esp 6 BW'=5/6-BW	6	0,84799	-	-	-	-
Esp 6 BW'=4/6-BW	6	0,68848	-	-	-	-
Esp 6 BW'=3/6-BW	6	0,52322	-	-	-	-
Esp 6 BW'=2/6-BW	5	0,13058	7	0,15528	9	0,17633
Esp 6 BW'=1/6-BW	1	0,13247	3	0,24489	5	0,2288
Esp 6 BW'=0-BW	1	0,15747	3	0,28741	5	0,25639
Esp 7 BW'=5/6-BW	7	0,91503	-	-	-	-
Esp 7 BW'=4/6-BW	7	0,82421	-	-	-	-
Esp 7 BW'=3/6-BW	7	0,82085	2	0,0055458	-	-

Identificación automática de Espectros Raman de pigmentos mediante ACP

Esp 7 BW'=2/6-BW	7	0,746	2	0,042559	-	-
Esp 7 BW'=1/6-BW	3	0,21435	5	0,18716	-	-
Esp 7 BW'=0-BW	3	0,2659	5	0,22634	1	0,14956
Esp 8 BW'=5/6-BW	8	0,84884	-	-	-	-
Esp 8 BW'=4/6-BW	8	0,68628	7	0,19406	5	0,17258
Esp 8 BW'=3/6-BW	8	0,51415	5	0,30407	7	0,27677
Esp 8 BW'=2/6-BW	5	0,40128	7	0,34134	3	0,21534
Esp 8 BW'=1/6-BW	5	0,39864	3	0,27693	1	0,13539
Esp 8 BW'=0-BW	5	0,35424	3	0,32056	1	0,16417
Esp 9 BW'=5/6-BW	9	0,93345	3	0,37382	6	0,0053194
Esp 9 BW'=4/6-BW	9	0,86092	2	0,11739	6	0,010934
Esp 9 BW'=3/6-BW	9	0,78132	7	0,45021	3	0,44969
Esp 9 BW'=2/6-BW	9	0,53751	3	0,22768	7	0,26201
Esp 9 BW'=1/6-BW	9	0,40542	3	0,28247	5	0,23345
Esp 9 BW'=0-BW	9	0,33761	3	0,30951	5	0,26296
Esp 10 BW'=5/6-BW	10	0,87448	1	0,20504	3	0,10763
Esp 10 BW'=4/6-BW	10	0,73789	1	0,24439	3	0,1967
Esp 10 BW'=3/6-BW	10	0,51391	3	0,22564	1	0,21749
Esp 10 BW'=2/6-BW	10	0,32606	3	0,29156	1	0,22209
Esp 10 BW'=1/6-BW	3	0,34475	5	0,23772	10	0,14041
Esp 10 BW'=0-BW	3	0,34692	5	0,26806	10	0,04431

ANEXO A2: Resultados obtenidos en el desplazamiento de las Bandas Raman: Distancia Euclídea, Candidatos y Factores de Confianza

Se presentan, a continuación, los resultados obtenidos en el apartado de la influencia del desplazamiento de las Bandas Raman sobre la librería espectral de referencia simulada. Dichos resultados incluyen los valores para la Distancia Euclídea cuando un espectro se ha desplazado -1cm^{-1} , -2cm^{-1} , y -3cm^{-1} , así como los espectros candidatos generados por el sistema y sus respectivos Factores de Confianza.

DESP=-1											
DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	COS2
Esp 1	1,6052	13,669	11,554	14,846	12,924	16,591	9,7777	13,264	12,675	8,1441	0,9527
Esp 2	14,01	1,8149	14,011	13,003	14,537	15,559	13,69	16,054	12,696	15,267	0,8462
Esp 3	11,415	13,906	1,9495	15,513	8,5123	16,751	11,406	12,813	10,543	10,244	0,8727
Esp 4	15,572	14,478	16,855	2,0819	15,181	14,359	16,569	17,817	17,178	17,308	0,8864
Esp 5	12,515	15,041	10,782	14,095	1,8804	14,954	10,745	11,635	12,11	13,193	0,8258
Esp 6	17,311	16,592	18,231	16,131	17,66	2,2433	18,366	20,004	16,498	19,758	0,8469
Esp 7	7,82	14,47	11,534	15,902	11,215	16,865	1,2799	11,127	11,077	11,487	0,9473
Esp 8	11,642	15,489	12,099	16,579	10,333	18,342	8,561	2,1666	13,088	10,345	0,8852
Esp 9	12,217	14,016	9,3448	16,52	11,789	15,418	11,337	14,262	1,1946	13,431	0,927
Esp 10	9,4073	15,146	10,672	16,339	13,242	18,843	12,328	11,981	14,202	1,6991	0,9009
DESP=-2											
DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	COS2
Esp 1	3,2476	13,12	11,337	14,677	13,096	16,719	10,454	13,564	12,934	7,2517	0,8701
Esp 2	13,983	3,9281	13,754	12,843	14,081	14,826	12,996	15,942	11,254	15,279	0,5856
Esp 3	11,25	13,448	4,1548	15,028	6,8263	16,448	11,197	12,869	10,767	10,386	0,7253
Esp 4	16,256	15,49	17,917	4,4115	16,279	13,305	17,438	18,595	18,016	18,345	0,6984
Esp 5	12,281	15,089	11,265	14,047	3,5148	13,8	10,056	11,699	12,112	13,378	0,5896
Esp 6	18,304	17,025	19,694	16,845	19,34	4,921	19,991	21,258	18,016	21,177	0,6227
Esp 7	6,5415	14,452	11,39	15,799	11,309	16,691	2,7061	11,57	11,016	11,163	0,8546
Esp 8	10,732	15,069	11,876	16,158	9,6811	17,925	6,7608	4,1216	12,429	10,365	0,7886
Esp 9	12,167	14,345	8,34	16,628	11,648	15,803	11,542	14,553	2,6505	13,4	0,8042
Esp 10	10,009	14,965	11,108	16,338	13,709	19,236	13,002	13,157	14,969	3,2474	0,7317
DESP=-3											
DISTANCIA	Esp 1	Esp 2	Esp 3	Esp 4	Esp 5	Esp 6	Esp 7	Esp 8	Esp 9	Esp 10	COS2
Esp 1	5,0962	12,794	11,311	14,726	13,414	17,085	11,197	14,003	13,428	6,4818	0,8263
Esp 2	14,05	6,2696	13,668	12,969	13,862	13,888	12,562	15,951	9,6891	15,387	0,4342
Esp 3	11,239	13,173	6,3086	14,6	5,2067	16,098	11,028	12,906	10,94	10,81	0,7328
Esp 4	17,154	16,49	19,262	7,3097	17,828	12,097	18,622	19,655	19,151	19,639	0,5787
Esp 5	11,721	14,857	11,402	13,897	5,6	12,344	8,9655	11,673	11,906	13,287	0,4768
Esp 6	19,127	16,731	20,936	17,078	20,796	8,0227	21,395	22,29	19,263	22,426	0,4877
Esp 7	5,1099	14,403	11,203	15,662	11,328	16,578	4,2202	11,903	10,981	10,734	0,8001

Identificación automática de Espectros Raman de pigmentos mediante ACP

Esp 8	10,102	14,917	11,991	16,019	9,6568	17,723	5,1448	5,8706	12,113	10,655	0,7404
Esp 9	12,137	14,542	7,073	16,671	11,379	16,064	11,68	14,762	4,3278	13,367	0,7256
Esp 10	10,591	14,656	11,715	16,297	14,272	19,681	13,679	14,207	15,806	4,7412	0,5749

Espectros candidatos y Factores de Confianza:

DESP=-1						
Espectro desconocido	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 1	1	0,81939	10	0,083662	-	-
Esp 2	2	0,86697	4	0,046936	-	-
Esp 3	3	0,80194	5	0,13519	-	-
Esp 4	4	0,86444	6	0,065062	-	-
Esp 5	5	0,87756	6	0,026336	-	-
Esp 6	6	0,85394	-	-	-	-
Esp 7	7	0,85533	1	0,11614	-	-
Esp 8	8	0,79438	7	0,18754	-	-
Esp 9	9	0,88229	3	0,079273	-	-
Esp 10	10	0,80882	-	-	-	-
DESP=-2						
Espectro desconocido	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 1	1	0,76197	10	0,46849	2	0,038387
Esp 2	2	0,74424	6	0,034685	-	-
Esp 3	3	0,69547	5	0,49966	2	0,014356
Esp 4	4	0,71276	6	0,13371	-	-
Esp 5	5	0,77115	6	0,10147	-	-
Esp 6	6	0,67959	-	-	-	-
Esp 7	7	0,69414	1	0,26065	-	-
Esp 8	8	0,60885	7	0,35838	-	-
Esp 9	9	0,73885	3	0,17828	-	-
Esp 10	10	0,63461	-	-	-	-
DESP=-3						
Espectro desconocido	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza	Espectro candidato	Factor de Confianza
Esp 1	1	0,62647	10	0,52491	2	0,062262
Esp 2	2	0,54047	-	-	-	-
Esp 3	5	0,61837	3	0,53761	2	0,034477
Esp 4	4	0,52406	6	0,21234	-	-
Esp 5	5	0,43106	-	-	-	-
Esp 6	6	0,47764	-	-	-	-
Esp 7	7	0,52301	1	0,42245	-	-
Esp 8	7	0,51175	8	0,44286	-	-
Esp 9	9	0,57359	3	0,30311	-	-
Esp 10	10	0,46654	-	-	-	-

ANEXO B: Funciones implementadas en *Matlab*

Se adjunta, a continuación, el código correspondiente a las principales funciones desarrolladas durante la realización de este proyecto, implementadas en *Matlab*.

Función *Lorentziana*:

```
function y=lorentz(x, x0, A, B)
    y=A./(1+((x-x0).^2)/B^2);
end
```

PCA: Generación del espacio de PCs:

```
function [Sred,Cred,k]=PCAnorm(t,x)
    [m,n]=size(x);
    figure(1);
    plot(t,x);
    title('Librería espectral de referencia');
    xlabel('Desplazamiento Raman');
    ylabel('Intensidad');

    %Normalización de las intensidades:
    for v=1:m
        Imin=min(x(v,:));
        Imax=max(x(v,:));
        dif=Imax-Imin;
        x(v,:)=(x(v,)-Imin)/dif;
    end

    %Estandarización de las variables:
    sum=0;
    for k=0:n-1
        for i=1+(k*m):m+(k*m)
            sum=sum+x(i);
        end
        sum=sum/m;
        vAr=sqrt(var(x(:,k+1)));
        for i=1+(k*m):m+(k*m)
            x(i)=(x(i)-sum)/vAr;
        end
        sum=0;
    end

    %Cálculo de la matriz de covarianzas:
    c=cov(x);
```

```

%Diagonalización:
[V,D]=eig(c);

%Ordenación de los vep's:
[a,b]=size(D);
vect(1:n)=0;
for vari=1:n
    c=0;
    pos=1;
    gran=0;
    for i=1:a+1:(a*b)
        if abs(D(i))>=gran
            for var2=1:n
                if i ~= vect(var2)
                    c=c+1;
                end
                if c==n
                    pos=i;
                end
            end
            gran=D(i);
            c=0;
        end
    end
    vect(vari)=pos;
end

%Cálculo de los PC's:
C(:,:)=0;
for j=1:n
    c=ceil(vect(j)/n);
    for i=1:n
        C(i,j)=V(i,c);
    end
end
C=C';
S=x*C';

%Selección del número de PCs deseado (k<<n)
%mediante el criterio del 100% de varianza:
vtot=0;
for i=1:n*n
    vtot=vtot+D(i);
end
v=0;
k=1;
vacum(:,1)=0;
for index=1:100
    vacum(index,1)=index;

```

```

end
%Varianza acumulada < 100% varianza total:
a=roundn(v,-4);
while a<1.0000
    v=v+D(vect(k))/vtot;
    vacum(k,3)=v;
    vacum(1,2)=vacum(1,3);
    if k>1
        vant=vacum(k-1,3);
        vacum(k,2)=v-vant;
    end
    k=k+1;
    a=roundn(v,-4);
end
k=k-1;
disp('  Varianzas por PC y varianza acumulada:');
disp(' PC      var_PC  var_acum')
disp(vacum(1:k,:));
Sred=S(:,1:k);
Cred=C(1:k,:);

%Plot del resultado:
figure(2);
plot(Sred');
title('PCs');
xlabel('eje x');
ylabel('eje y');

end

```

Cálculo de la Distancia Euclídea:

```

function y=distancia(x)
    y=squareform(pdist(x));
end

```

Cálculo del Coseno Cuadrado:

```

function co2=CO2(S,a,K,C)
    %como contribuye la alfa-ésima PC del
    %espectro i-ésimo sobre dicho espectro // filas
    [m1,n1]=size(a);
    x=tipifica(a(1:K+1,:));
    a=[x;a(K+2:m1,:)];
    [m,n]=size(S(:,,:));
    co2(:,:)=0;
    for j=1:m
        sum=norm(a(j,:))^2;
    end

```

```

    for i=1:K
        co2(j,i)=(S(j,i)^2)/sum;
    end
end
for i=1:m
    co2(i,K+1)=suma(co2(i,1:K));
end
end

```

Parámetro *Factor de Confianza*:

```

function fc=FC(dist,mins)
    fc=(1-(dist/mins));
end

```

Cálculo de las mezclas binarias:

```

function [SM,am]=mezclas(a,S,Cred,K,meds,vars)
[m,n]=size(a);
e=1;
for i=1:1:m
    for j=1:1:m
        for k=1:1:n
            am(e,k)=a(i,k)+a(j,k);
        end
        e=e+1;
    end
end
[m2,n2]=size(am);
for i=1:1:m2
    Const(i)=1/max(am(i,:));
end
e=1;
aux=((meds./vars)*Cred');
for i=1:1:m
    for j=1:1:m
        for k=1:1:K
            aux1=((2*Const(e))-1)*aux;
            SM(e,k)=Const(e)*(S(i,k)+S(j,k))+aux1(k);
        end
        e=e+1;
    end
end
end

```

Cálculo de la SNR:

```

function y = snr(esc,a);

```

```
[m,n]=size(a);
per=esc*17.0968;
y=per*a;
for i=1:1:m
    snrmean(i)=min(y(i,:));
end
y=[y,snrmean'];
end
```

Algoritmo de Identificación Espectral:

```
function [res,res1,res2,dist,cos]=decide(adesc,a,S,C,t)

[m,meds,vars]=proyecta(adesc,a,C);
[K1,K]=size(S);
dist=distancia([S;m]);
cos=CO2([S;m],[a;estandariza(adesc,a,C)],K);
mins=buscaminbase(dist(1:K+1,1:K+1));
minbase=min(mins);
maxbase=max(mins);
[m1,n1]=size(dist);
fc=0;
num=0;
guarda=0;

[SM,am]=mezclas(normaliza(a),S,C,K,meds,vars);
dist1=distancia([SM;m]);

for i=K+2:1:m1
    k=1;
    for j=1:1:K+1
        if dist(i,j)<minbase
            candidatos(i-(K+1),k)=j;
            candist(k)=dist(i,j);
            guarda(i-(K+1),k)=mins(j);
            %fc(i-(K+1),k)=FC(candist(k),mins(j),cos(i,K+1));
            k=k+1;
        else
            candidatos(i-(K+1),k)=0;
        end
    end
end
for q=1:1:K+1
    cota=minbase/mins(q);
    if cos(i,K+1)>cota
        if dist(i,q)<mins(q)
            if dist(i,q)>min(mins)
                candidatos(i-(K+1),k)=q;
                guarda(i-(K+1),k)=mins(q);
                %fc(i-(K+1),k)=FC(dist(i,q),mins(q),cos(i,K+1));
```

```

        k=k+1;
    end
end
end
end

k=1;
for j=1:1:K+1
    if dist(i,j)<minbase
        fc(i-(K+1),k)=FC(candist(k),max(guarda(i-(K+1),:)),cos(i,K+1));
        k=k+1;
    end
end
for q=1:1:K+1
    cota=minbase/mins(q);
    if cos(i,K+1)>cota
        if dist(i,q)<mins(q)
            if dist(i,q)>min(mins)
                fc(i-(K+1),k)=FC(dist(i,q),max(guarda(i-(K+1),:)),cos(i,K+1));
                k=k+1;
            end
        end
    end
end
end

val=i-(K+1);
plotea(val, adesc, k, candidatos);

[m2,n2]=size(fc);
res(:,:)=0;
k=1;
for s=1:1:n2
    for t1=1:1:m2
        res(t1,k)=candidatos(t1,s);
        res(t1,k+1)=fc(t1,s);
    end
    k=k+2;
end
end

[m3,n3]=size(res);
res1=0;
%dist1=0;
UMBRAL1=0.05;
UMBRAL2=0.5;
h=1;
if n2>1
    for i=1:1:m2
        for j=1:1:n2-1
            for k=j:1:n2

```


ANEXO C1: Paleta básica: Características de la librería espectral de referencia

A continuación se adjunta una descripción de los pigmentos utilizados en la librería espectral de referencia aquí llamada “paleta básica”:

Tonalidad Blanca:

- Anatasa [TiO_2]: El anatasa es un pigmento que proviene de una cristalización especial del blanco de titanio. Apareció en pintura después de la primera Guerra Mundial y se dejó de utilizar por ser tóxico y por la aparición del Rutilo en el año 1938^[18,19,20].
- Rutilo [TiO_2]: Los blancos de titanio aparecieron hacia 1920 en su versión anatasa. A partir de 1940 se comercializó en Estados Unidos la versión rutilo (de igual composición química pero distinta estructura cristalina) que substituyó a la anatasa debido a sus mejores propiedades en pintura. Su introducción en Europa como pigmento no se produjo hasta 1945^[18,21].
- Blanco de Plomo [$2\text{PbCO}_3(\text{OH})_2$]: Este pigmento es considerado el pigmento blanco más importante y se trata de uno de los pigmentos artificiales más antiguo. Se tiene noticia de este pigmento desde el año 300 a.C., aproximadamente, y desde entonces es mencionado en todos los catálogos de pigmentos. Hasta mitad del siglo XIX fue el único pigmento utilizado en pintura al óleo, y se siguió utilizando, aunque en menor medida, tras la aparición de nuevos pigmentos blancos^[18,22].
- Blanco de Creta [CaCO_3]: La Creta (Carbonato Cálcico) se empleaba ya de muy antiguo en la técnica pictórica. Se utilizaba como pigmento en las imprimaciones de cuadros, en las pinturas de cola al temple y de caseína como pigmento de mezcla, así como en otras pinturas de emulsión diluibles en agua, como imprimación en dorados, en masillas para restauraciones y en la fabricación de barras para pastel^[18].
- Blanco de Zinc [ZnO]: Fue introducido como color de acuarela en el 1834, pero, debido a ciertas dificultades, no se empezó a utilizar en pintura al óleo hasta el 1844 por LeClaire (París). En 1845 se comenzó a producir a gran escala y cinco años después se empezó a manufacturar por toda Europa. Actualmente aún se utiliza, aunque no en su forma pura, y se ha ido substituyendo por la gama de

Blancos de Titanio. Tiene ventajas respecto al Blanco de Plomo ya que no es tóxico, no se oscurece con el sulfuro del aire y es más económico^[18,22].

Tonalidad Amarilla

- Amarillo de Nápoles [Pb₃(SbO₄)]: La procedencia y antigüedad de esta combinación de Plomo y Antimonio llamada amarillo de Nápoles, o también amarillo de Antimonio, es dudosa, aunque se puede afirmar casi con seguridad que hasta el siglo XVII no fue utilizada como pigmento. En las pinturas del siglo XVII su presencia es prácticamente nula, mientras que en las del siglo XVIII aparece frecuentemente. Hoy en día, no se utiliza en la técnica de pintura artística por el peligro que comporta para la salud^[18].
- Amarillo de Cromo [PbCrO₄]: Fue presentado por primera vez en el año 1805, aproximadamente, pero no fue hasta 10 años más tarde que se comenzó a fabricar en grandes cantidades como pigmento. En el siglo XIX ya estaba ampliamente extendido como color. Hoy en día la industria de pigmentos suministra amarillos de Cromo muy estables a la luz, hecho que lo convierte en miembro de la clase de pigmentos amarillos más sólidos y estables, y es muy utilizado en pinturas al óleo, acrílicas o acuarelas^[18].
- Amarillo de Cadmio [CdS]: El metal de cadmio fue descubierto en el año 1817 por Stromeyer y propuesto como pigmento 12 años más tarde por Melandri. En 1830 se pintaba ya con este pigmento, pero no fue hasta el 1849 cuando se fabricó en grandes cantidades^[18,23].
- Amarillo de Zinc [ZnCrO₄]: Henri Vauquelin, descubridor del cromo, descubrió en 1809 por primera vez el cromato de cinc, y desde 1850 se ha fabricado como pigmento, aunque actualmente ha dejado de fabricarse puesto que se ha constatado que provoca cáncer^[18].
- Oropimente [As₂S₃]: El oropimente se ha empleado como pigmento en pintura desde la antigüedad hasta el Renacimiento Plinio y Vitrubio lo describen con el nombre de *auripigmentum*. En el siglo XIX se elaboró artificialmente y se encontraba en los comercios con los nombres de *operment*, *amarillo embriagador* y *amarillo real*, y se empleaba principalmente en pintura al temple. El oropimente es un producto natural (muy venenoso, debido a la presencia de arsénico) cuya coloración varía desde el amarillo limón hasta el amarillo anaranjado^[18].

- Amarillo de Plomo-Estaño Tipo I [Pb_2SnO_4]: Este pigmento fue utilizado por los pintores europeos entre 1450 y 1750^[24], aunque Ashok Roy^[22] sitúa casi el 100% de su uso hasta finales del siglo XVII.
- Amarillo de Plomo, Estaño y Antimonio [triple óxido de plomo, estaño y antimonio con posible adicción de sílice]: Este tipo de amarillo apareció en la industria cerámica italiana a lo largo del siglo XVI, y fue introducido desde Venecia en la pintura barroca italiana durante el siglo XVII^[18], siendo desplazado a lo largo del siglo XVIII por el amarillo de Nápoles^[24].

Tonalidad Roja

- Bermellón [HgS]: Este compuesto existe como mineral (cinabrio) desde la antigüedad. Era muy popular en la antigua China y Egipto, y se cree que hasta el siglo XVI fue utilizado en la pintura artística. Ahora bien, su versión sintética (que data del siglo VIII en las referencias europeas) es la que típicamente se utiliza desde hace muchos siglos. A principios del siglo XX este pigmento fue desplazado por el rojo de Marte por su alta toxicidad^[18,22].
- Minio de Hierro [Fe_2O_3]: Los óxidos de hierro naturales son utilizados en pintura desde la más remota antigüedad y se siguen utilizando hoy en día. Se fabrican de forma artificial desde el siglo XIX^[18].
- Minio de Plomo [$\text{Pb}_3\text{O}_4 \cdot 2\text{PbO} \cdot \text{PbO}_2$]: Se utiliza desde la Antigüedad, aunque actualmente no es aconsejable su uso para técnicas artísticas, puesto que, como derivado del plomo, es tóxico. No obstante, es cubriente, aunque no es resistente al aire ni a la luz, ya que se oscurece; mezclado con el aceite es más estable, por lo que sólo es aconsejable en la técnica del óleo^[18].

Tonalidad Azul

- Azurita [$2\text{CuCO}_3\text{Cu}(\text{OH})_2$]: Se empezó a utilizar como material pictórico en el antiguo Egipto, y fue el pigmento azul más importante en la pintura europea durante la Edad Media y el Renacimiento. En obras posteriores también se puede encontrar, aunque con el tiempo, el azul de ultramar toma mayor protagonismo. A partir del siglo XIX cae en desuso y es sustituido de la paleta europea por el Azul de Prusia, inventado en el siglo XVIII^[18].

- Azul Ultramar [$\text{Na}_{6-10}\text{Al}_6\text{Si}_6\text{O}_4\text{S}_{2-4}$]: Guimet fue el primero en sintetizarlo en Francia en el año 1828, y a partir de 1830 estuvo disponible comercialmente en su versión sintética, mucho más barata que su versión natural, obtenida a partir del lapislázuli, de localización remota y poco abundante. Este pigmento moderno forma parte de uno de los grupos de pigmentos más importantes e indispensables tanto en la pintura al óleo, como a su vez en la acuarela^[18,21].
- Azul de Cobalto [$\text{CoO-Al}_2\text{O}_3$]: Descubierta en el año 1777 en ensayos químicos con soplete, fue fabricado en 1802 por el químico francés Thénard, calentando hidrato de arcilla con fosfato de cobalto al rojo. Desde mediados del siglo XIX se emplea en pintura. La solidez frente a la luz y la estabilidad respecto a los álcalis y a los ácidos son perfectos. Estas magníficas propiedades técnico-pictóricas y su tonalidad azul clara y fresca han abierto a este pigmento las puertas para su aplicación en las técnicas pictóricas pese a su elevado precio^[18,23].

Tonalidad Verde

- Verde de Scheele [CuHAsO_3]: De origen tanto mineral como sintético, creado en 1775 por el químico sueco Carl Wilhelm Scheele. Se ha empleado intensamente durante el siglo pasado, sobre todo en pintura al óleo. Se altera con el contacto con plomo y se descompone con el ácido^[18].
- Verdigrís [$n\text{Cu}(\text{OH})_2-(\text{CH}_3\text{COO})_2-n\text{H}_2\text{O}$]: Se empezó a utilizar como pigmento en el siglo XIV y se extendió su uso a lo largo de los siglos XVI y XVII, mezclado habitualmente con amarillo de Plomo-Estaño tipo I. Cayó en desuso a comienzos del siglo XVIII^[24].
- Malaquita [$2\text{CuCO}_3\text{Cu}(\text{OH})_2$]: Se encuentra comúnmente como mineral junto a la azurita. Se diferencia únicamente de la azurita por el contenido de componentes básicos. Se empleaba principalmente con aglutinantes del temple o en el fresco. A finales del siglo XVIII desapareció de la pintura europea y no se conoce su fabricación sintética, seguramente por la aparición de otros pigmentos verdes^[18].

La notación utilizada tanto en este anexo como en el *Apartado 5.2* se corresponde con:

Esp 1:	Anatasa	Esp 6:	Amarillo de Plomo-Estaño Tipo I	Esp 11:	Amarillo de Cromo	Esp 16:	Minio de Hierro	Esp 21:	Bermellón
Esp 2:	Azul de Cobalto	Esp 7:	Amarillo de Plomo, Estaño y Antimonio	Esp 12:	Amarillo de Nápoles	Esp 17:	Minio de Plomo		
Esp 3:	Azurita	Esp 8:	Verde de Scheele	Esp 13:	Amarillo de Cinc	Esp 18:	Oropimente		
Esp 4:	Blanco de Plomo	Esp 9:	Azul Ultramar	Esp 14:	Blanco de Cinc	Esp 19:	Rutilo		
Esp 5:	Blanco de Creta	Esp 10:	Amarillo de Cadmio	Esp 15:	Malaquita	Esp 20:	Verdigrís		

Tabla C1.1: Notación de los espectros de la librería y correspondencia con los pigmentos

A partir de los espectros de dichos pigmentos, se crea la librería espectral siguiente (ver *Fig. C1.1* y *C1.2* y su representación en el espacio de PCs en *Fig. C1.3*):

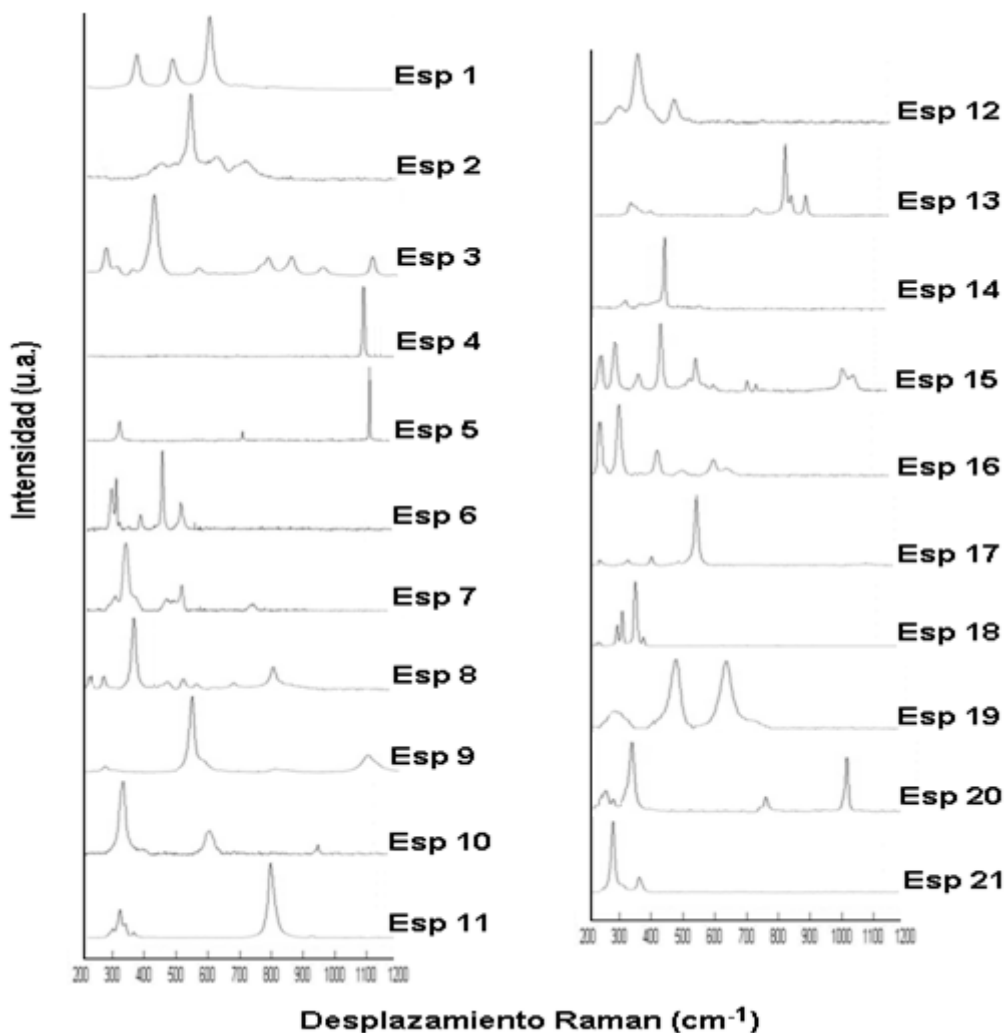


Fig. C1.1: Representación de los Espectros Raman de la paleta básica

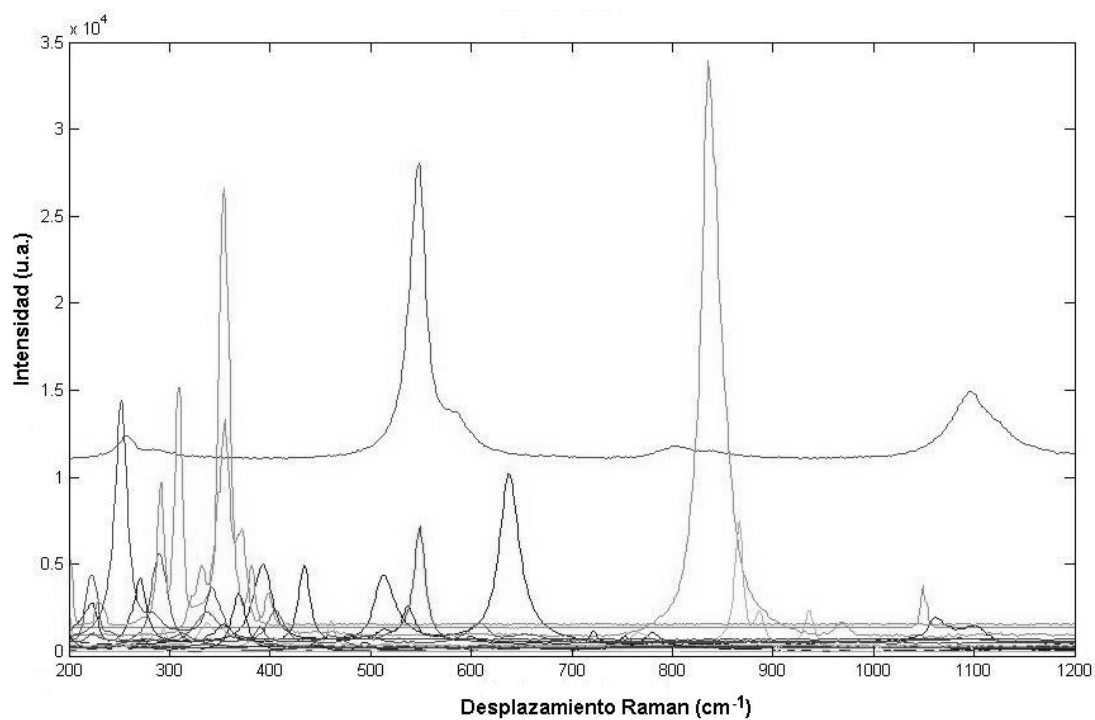


Fig. C1.2: Representación conjunta de los Espectros Raman de la paleta básica

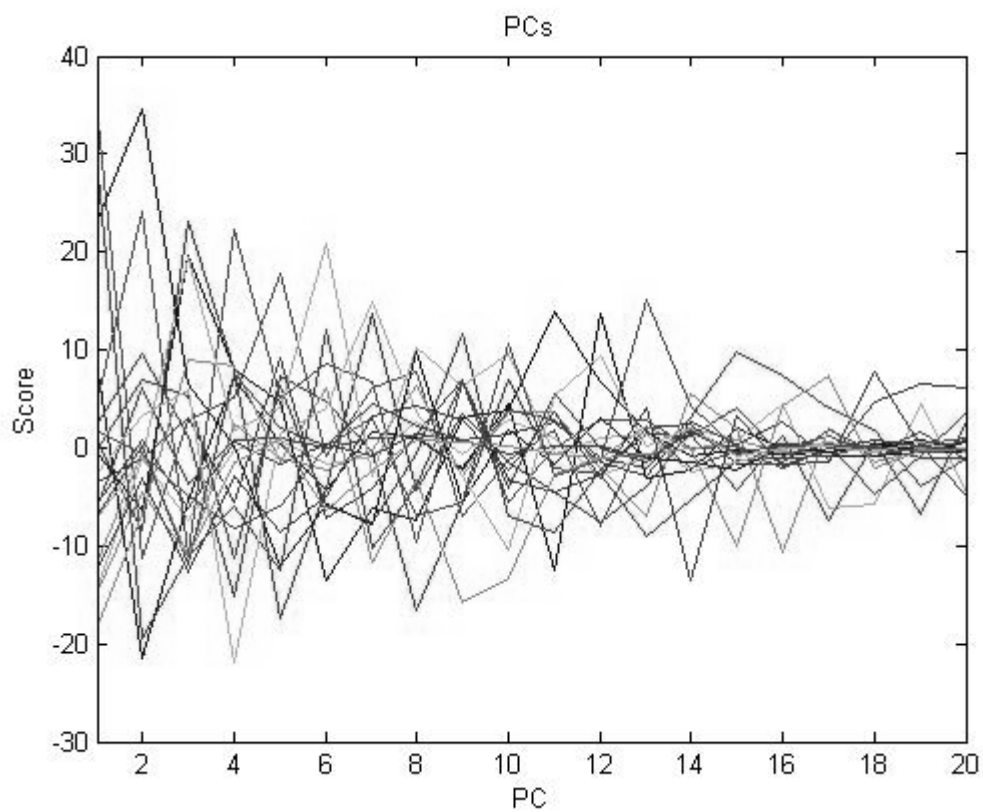


Fig. C1.3: Representación conjunta de los espectros patrones en PCs

La *Tabla C1.2* muestra los valores de varianza por PC y la varianza acumulada, recogidos en forma de diagrama en la figura *Fig. C1.4*:

PC	Varianza por PC	Varianza acumulada
PC1	0.1956	0.1956
PC2	0.1564	0.3520
PC3	0.1226	0.4747
PC4	0.0911	0.5658
PC5	0.0645	0.6303
PC6	0.0571	0.6874
PC7	0.0483	0.7357
PC8	0.0432	0.7789
PC9	0.0382	0.8171
PC10	0.0366	0.8538
PC11	0.0295	0.8832
PC12	0.0253	0.9085
PC13	0.0223	0.9308
PC14	0.0161	0.9469
PC15	0.0132	0.9602
PC16	0.0114	0.9716
PC17	0.0090	0.9805
PC18	0.0075	0.9880
PC19	0.0064	0.9945
PC20	0.0055	1.0000

Tabla C1.2: Varianza por PC y varianza acumulada (en tanto por uno)

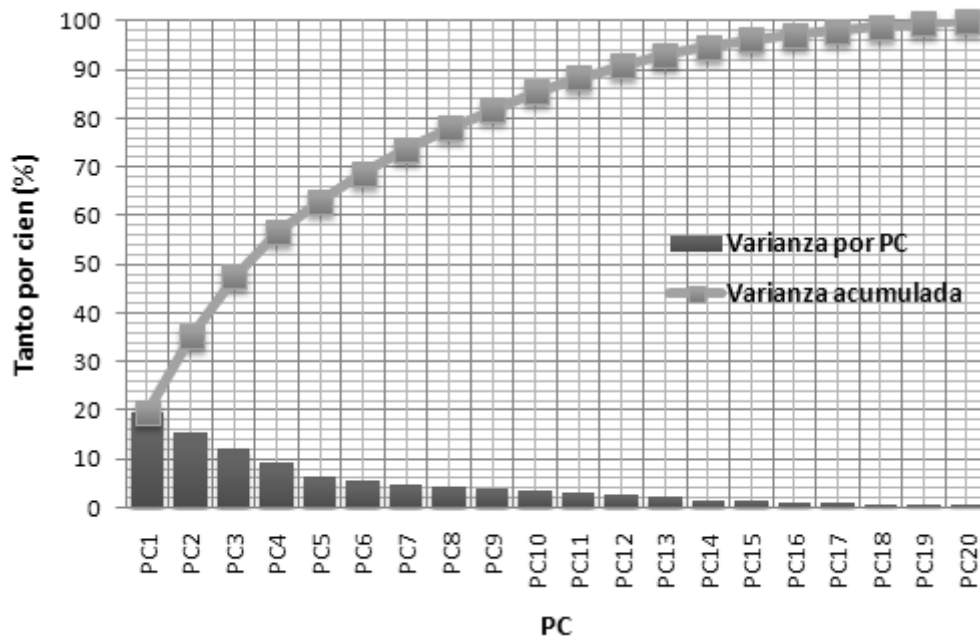


Fig. C1.4: Diagrama de varianzas

La *Tabla C1.3* recoge los valores de las mínimas Distancias Euclídeas para cada espectro patrón y el resto de espectros de la librería.

Espectro	Menor distancia al resto de patrones
Esp 1	36.6179
Esp 2	44.6218
Esp 3	40.0916
Esp 4	21.4311
Esp 5	19.2559
Esp 6	17.0100
Esp 7	21.5343
Esp 8	35.3017
Esp 9	37.9033
Esp 10	33.7501
Esp 11	36.6586
Esp 12	21.5343
Esp 13	39.5932
Esp 14	17.0100
Esp 15	49.2658
Esp 16	28.8965
Esp 17	19.9617
Esp 18	22.1205
Esp 19	44.0004
Esp 20	36.5615
Esp 21	25.6785

Tabla C1.3: Mínima Distancia Euclídea entre los espectros patrones y el resto de espectros de la librería de referencia

De la anterior tabla se obtiene que la mínima distancia de la librería es **17.01u**.

Las figuras siguientes muestran los *biplots* bidimensionales de la librería de referencia:

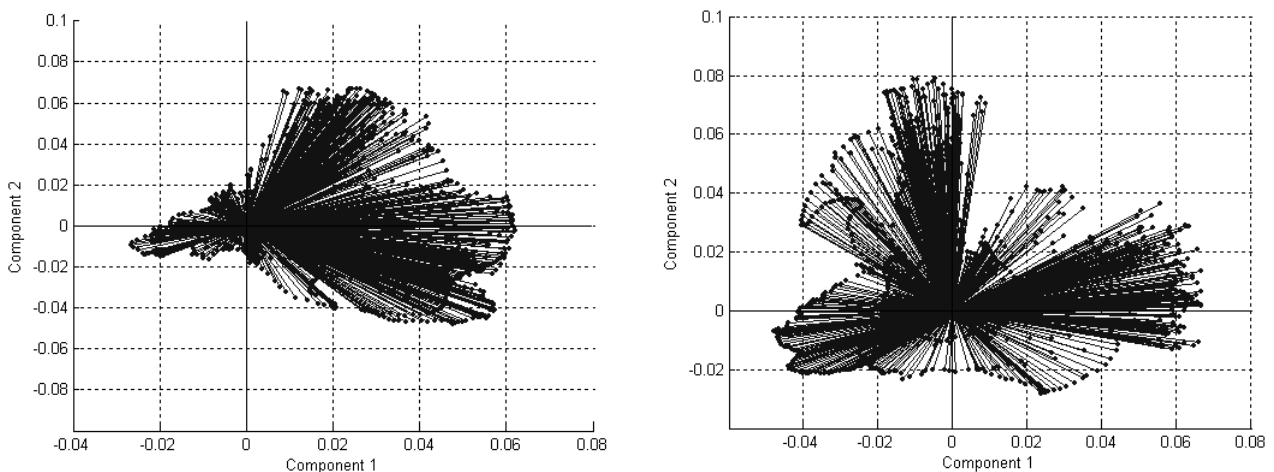


Fig. C1.5: Biplot PC1-PC2 a la izquierda y biplot PC2-PC3 a la derecha

Se adjunta, también, el cálculo del Parámetro de Calidad referente al ruido umbral para la paleta básica:

$$\%ruido_{max} = 2 \cdot \frac{min_lib}{\sqrt{k} \cdot M.D.} = \frac{2 \cdot 17.01}{\sqrt{20} \cdot (34.55 - (-21.86))} = 13.49\%$$

$$Q_{dB}^{umbral} = 10 \cdot \log \left(\frac{max(E_i)}{\sigma_{nE_i}} \right) \Bigg|_{A_{nE_i}=13.49\%}; \text{ Normalización } \rightarrow$$

$$Q_{dB}^{umbral} = \frac{1}{\sigma\{0.1349 \cdot random\}} = 14.35dB$$

Si $Q_{dB}^{desconocido} > Q_{dB}^{umbral} \rightarrow$ El espectro **puede ser identificado**

Si $Q_{dB}^{desconocido} < Q_{dB}^{umbral} \rightarrow$ El espectro **no es identificado**

Por último, la *Tabla C1.4* recoge las posiciones de las Bandas Fundamentales de cada uno de los espectros de la librería de referencia.

Posición de la Banda Fundamental (cm ⁻¹)									
Esp 1:	638	Esp 6:	462	Esp 11:	837	Esp 16:	291	Esp 21:	253
Esp 2:	515	Esp 7:	338	Esp 12:	338	Esp 17:	550		
Esp 3:	404	Esp 8:	370	Esp 13:	868	Esp 18:	355		
Esp 4:	1050	Esp 9:	549	Esp 14:	440	Esp 19:	450		
Esp 5:	1092	Esp 10:	306	Esp 15:	435	Esp 20:	327		

Tabla C1.4: Posiciones de las Bandas Fundamentales de los espectros patrones

ANEXO C2: Paleta de pigmentos amarillos orgánicos: Características de la librería espectral de referencia

La paleta aquí analizada ha sido generada a partir de pigmentos amarillos. La codificación utilizada en dichos pigmentos se compone de dos letras (*PY – Pigment Yellow*) y un número que refieren su color, composición, su origen y su aplicación industrial, según la sociedad internacional *The Society of Dyers and Colourists* (SDC). A continuación se desglosa la librería de referencia por sus características químicas^[6]:

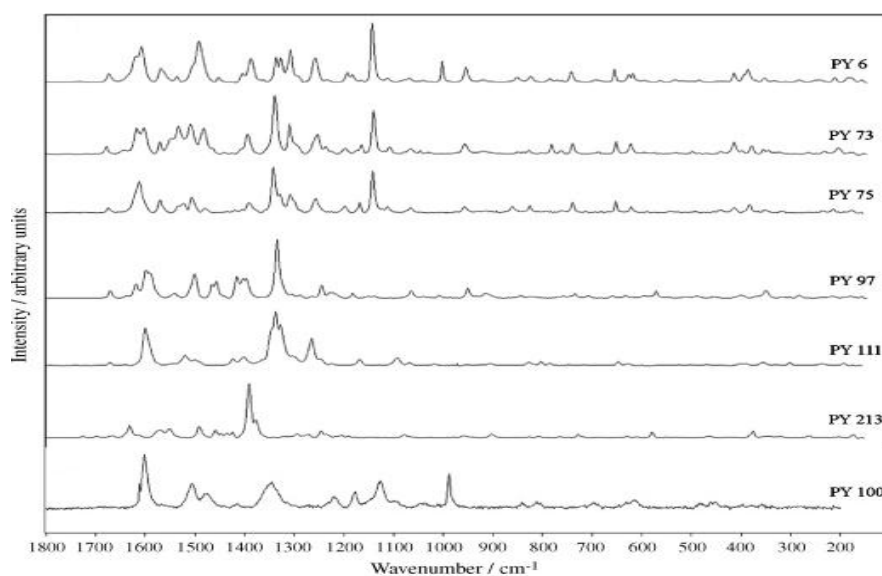


Fig. C2.1.a: Pigments Yellows tipo Monoazo

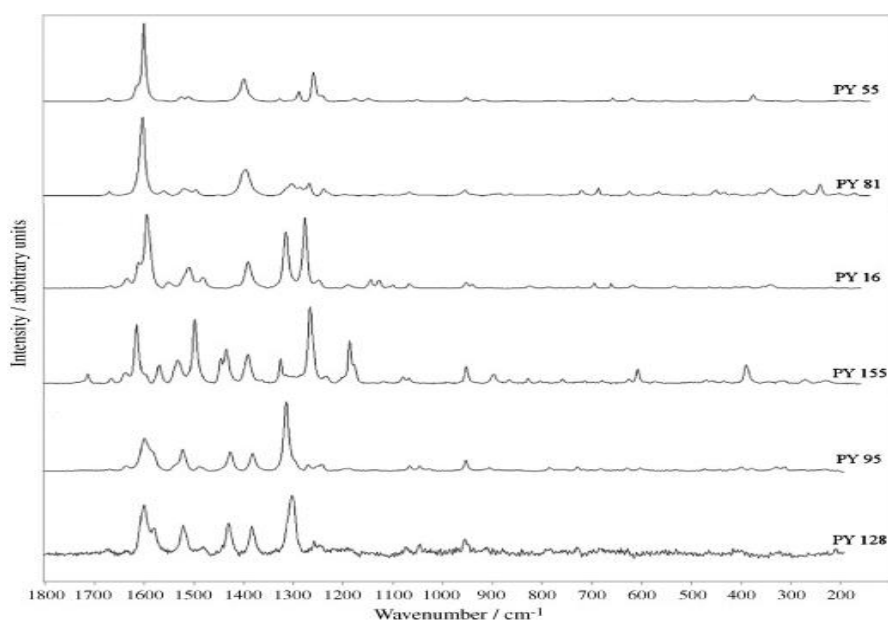


Fig. C2.1.b: Pigments Yellows tipo Diarilida

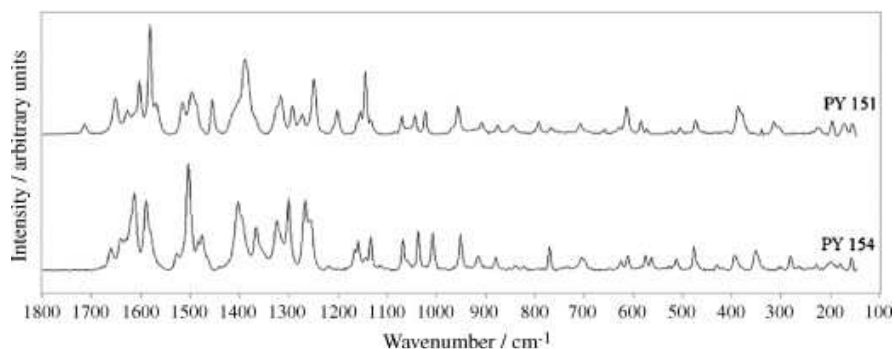


Fig. C2.1.c: Pigments Yellows tipo Benzimidazolone

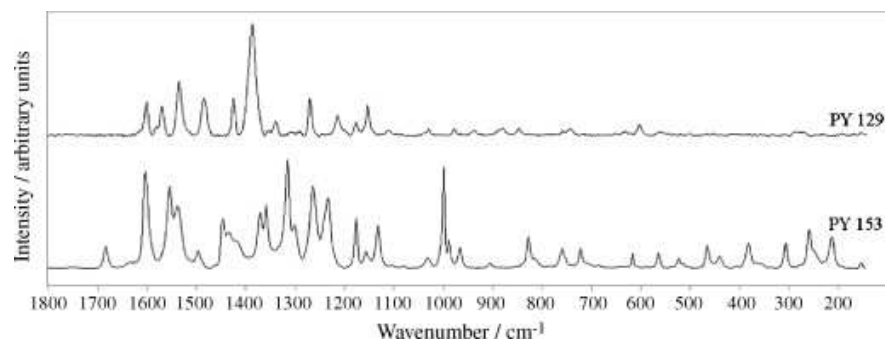


Fig. C2.1.d: Pigments Yellows tipo Azomethine

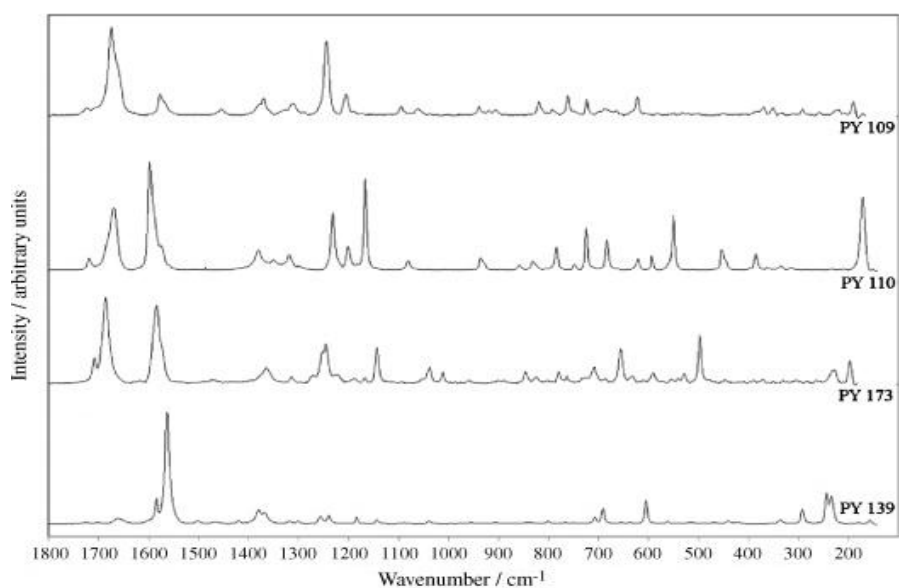


Fig. C2.1.e: Pigments Yellows tipo Isoindoline e isoindoline

La notación utilizada en el *Apartado 5.3* se corresponde con:

Esp 1:	PY 6	Esp 6:	PY 81	Esp 11:	PY 111	Esp 16:	PY 154	Esp 21:	PY 173
Esp 2:	PY 16	Esp 7:	PY 95	Esp 12:	PY 129	Esp 17:	PY 155		
Esp 3:	PY 55	Esp 8:	PY 97	Esp 13:	PY 139	Esp 18:	PY 213		
Esp 4:	PY 73	Esp 9:	PY 109	Esp 14:	PY 151	Esp 19:	PY 100		
Esp 5:	PY 75	Esp 10:	PY 110	Esp 15:	PY 153	Esp 20:	PY 128		

Tabla C2.1: Notación de los Espectros Raman PYs

En la figura Fig. C2.2, se adjuntan las características químicas de los pigmentos usados:

C.I. name and group	Generic group formula / Pigment formula	Substituents
AZO TYPE		
Monoazo		Y_D^2 Y_D^4 Y_D^5 Y_K^2 Y_K^4 Y_K^5
PY 6		NO ₂ Cl H H H H
PY 73		NO ₂ Cl H OCH ₃ H H
PY 75		NO ₂ Cl H H OC ₂ H ₅ H
PY 97		OCH ₃ SO ₂ NH-C ₆ H ₅ OCH ₃ OCH ₃ Cl OCH ₃
PY 111	OCH ₃ NO ₂ H OCH ₃ H Cl	
PY 213		
Monoazo pigment lake: PY 100		
Diarylide		X Y Z^2 Z^4 Z^5
PY 55		Cl H H CH ₃ H
PY 81	Cl Cl CH ₃ CH ₃ H	
Bisacetoacetarylide		
PY 16		
PY 155		

C.I. name and group	Generic group formula / Pigment formula	Substituents / Pigment formula												
Disazo condensation		<table border="1"> <thead> <tr> <th>X</th> <th>Y</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> </tr> <tr> <td></td> <td></td> </tr> </tbody> </table>	X	Y										
X		Y												
PY 95														
PY 128		<table border="1"> <tbody> <tr> <td></td> <td></td> </tr> </tbody> </table>												
Benzimidazolone		<table border="1"> <thead> <tr> <th>X₂</th> <th>X₃</th> <th>X₄</th> <th>X₅</th> </tr> </thead> <tbody> <tr> <td>COOH</td> <td>H</td> <td>H</td> <td>H</td> </tr> <tr> <td>CF₃</td> <td>H</td> <td>H</td> <td>H</td> </tr> </tbody> </table>	X ₂	X ₃	X ₄	X ₅	COOH	H	H	H	CF ₃	H	H	H
X ₂	X ₃	X ₄	X ₅											
COOH	H	H	H											
CF ₃	H	H	H											
PY 151														
PY 154														
AZOMETHINE and METHINE TYPES														
Azomethine metal complex														
PY 129 PY 153	<p>PY 129</p>	<p>PY 153</p>												
Isoindolinone (azomethine type) and isoindoline (methine type)														
PY 109 PY 110 PY 173 PY 139	<p>PY 109</p> <p>Azomethine type</p>	<p>PY 110</p> <p>Azomethine type</p>												
	<p>PY 173</p> <p>Azomethine type</p>	<p>PY 139</p> <p>Methine type</p>												

Fig. C2.2: Características químicas de los espectros PY

Se crea, así, una librería con un elevado número de Bandas Raman (ver figura *Fig. C2.3*).

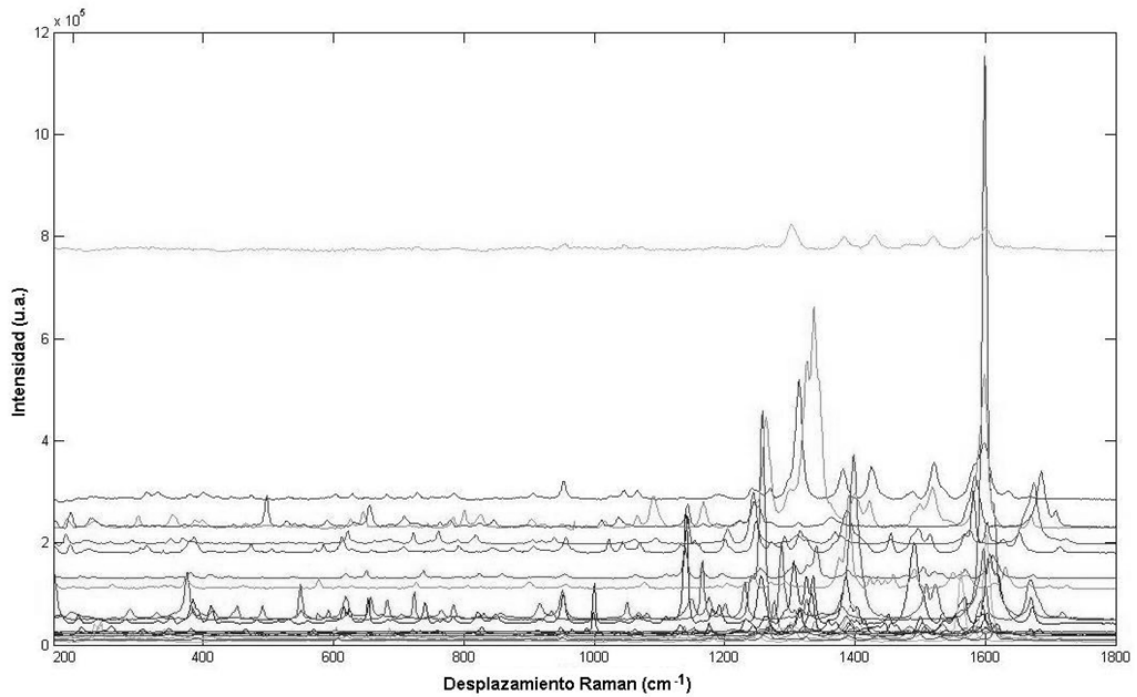


Fig. C2.3: Representación conjunta de los Espectros Raman Pigment Yellow (PYs)

Al aplicar la técnica de reducción dimensional basada en el Análisis por Componentes Principales se obtiene la representación espectral mostrada en la figura *Fig. C2.4*.

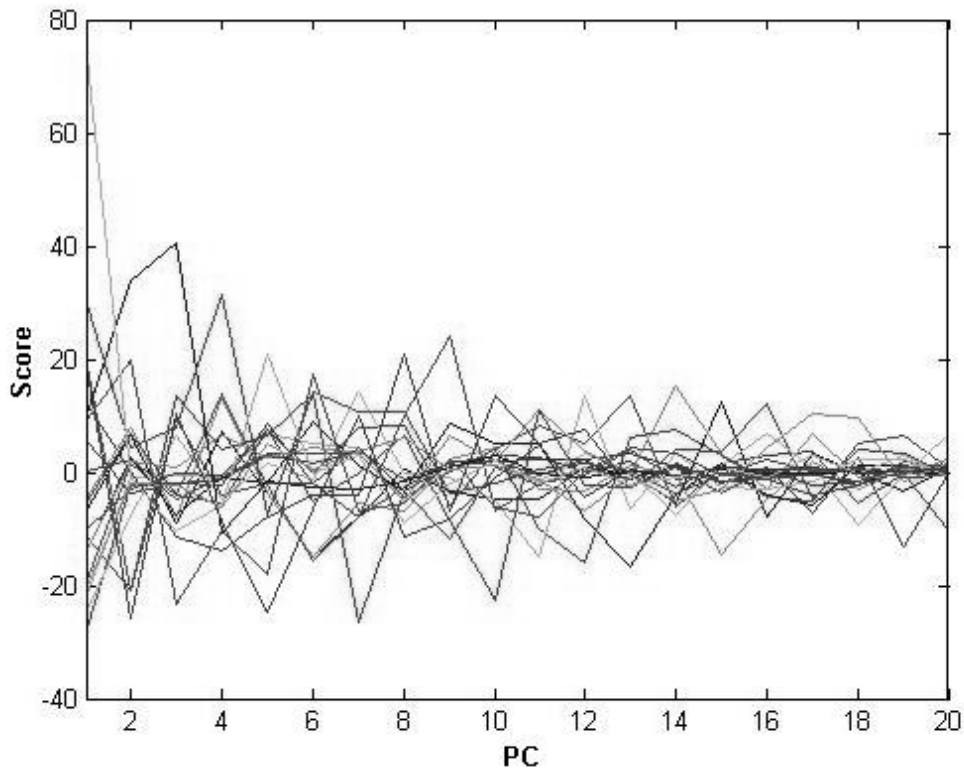


Fig. C2.4: Representación en PCs paralelos de la librería de referencia

La *Tabla C2.2* muestra los valores de varianza por PC y la varianza acumulada, recogidos en forma de diagrama en la figura *Fig. C2.5*:

PC	Varianza por PC	Varianza acumulada
PC1	0.3416	0.3416
PC2	0.1020	0.4436
PC3	0.0971	0.5407
PC4	0.0640	0.6047
PC5	0.0571	0.6619
PC6	0.0497	0.7116
PC7	0.0481	0.7597
PC8	0.0370	0.7966
PC9	0.0335	0.8302
PC10	0.0294	0.8595
PC11	0.0256	0.8851
PC12	0.0217	0.9068
PC13	0.0191	0.9259
PC14	0.0160	0.9419
PC15	0.0139	0.9557
PC16	0.0114	0.9671
PC17	0.0110	0.9781
PC18	0.0088	0.9869
PC19	0.0082	0.9952
PC20	0.0048	1.0000

Tabla C2.2: Varianza por PC y varianza acumulada (en tanto por uno)

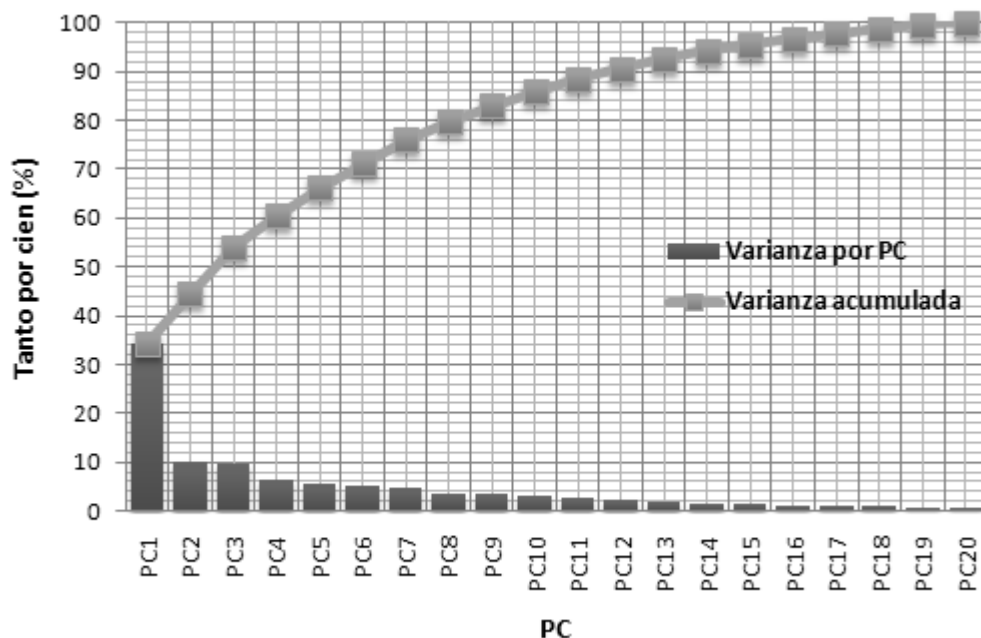


Fig. C2.5: Diagrama de varianzas

La *Tabla C2.3* recoge los valores de las mínimas Distancias Euclídeas para cada espectro patrón y el resto de espectros de la librería.

Espectro	Menor distancia al resto de patrones
Esp 1	30.7170
Esp 2	26.8206
Esp 3	21.5067
Esp 4	30.0692
Esp 5	30.0692
Esp 6	21.5067
Esp 7	27.6298
Esp 8	33.1897
Esp 9	43.7238
Esp 10	45.1922
Esp 11	30.3757
Esp 12	32.4861
Esp 13	33.4061
Esp 14	45.2573
Esp 15	63.5475
Esp 16	50.3388
Esp 17	36.9832
Esp 18	25.6046
Esp 19	56.3382
Esp 20	73.3557
Esp 21	51.2820

Tabla C2.3: Mínima Distancia Euclídea entre los espectros patrones y el resto de espectros de la librería de referencia

De la anterior tabla se obtiene que la mínima distancia de la librería es **21.5067u**.

Las figuras siguientes muestran los *biplots* bidimensionales de la librería de referencia:

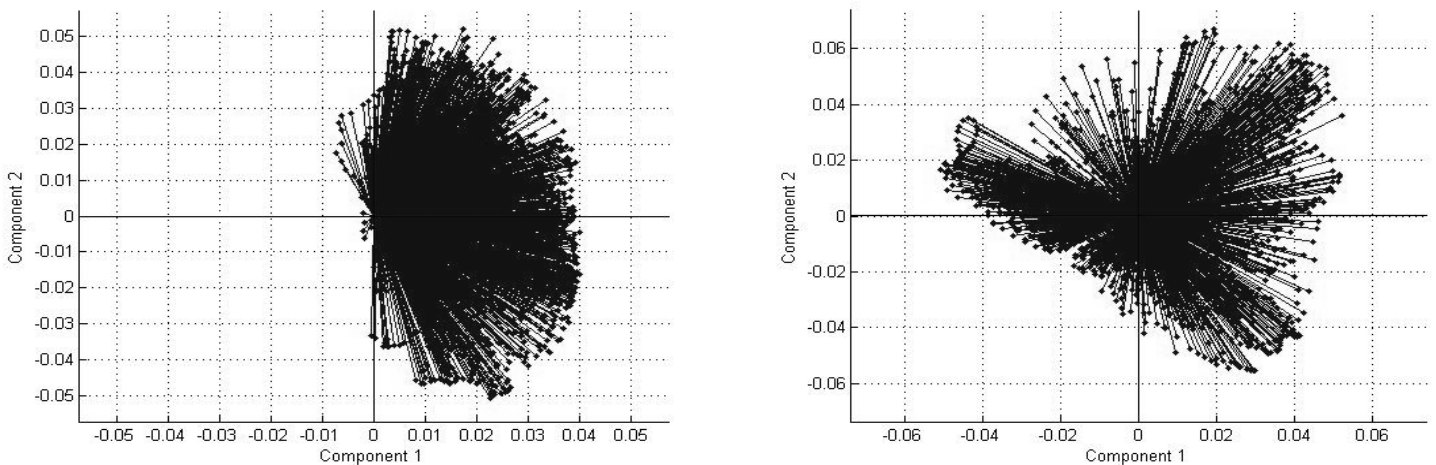


Fig. C2.5: Biplot PC1-PC2 a la izquierda y biplot PC2-PC3 a la derecha

Se adjunta, a continuación, el cálculo del Parámetro de Calidad referente al ruido umbral para la paleta de elevado número de Bandas Raman:

$$\%ruido_{max} = 2 \cdot \frac{min_lib}{\sqrt{k} \cdot M.D.} = \frac{2 \cdot 21.5067}{\sqrt{20} \cdot (79.8944 - (-27.7015))} = 8.94\%$$

$$Q_{dB}^{umbral} = 10 \cdot \log \left(\frac{\max(E_i)}{\sigma_{n_{E_i}}} \right) \Bigg|_{A_{n_{E_i}}=8.94\%} ; \text{ Normalización } \rightarrow$$

$$Q_{dB}^{umbral} = \frac{1}{\sigma\{8.94 \cdot random\}} = 16.28dB$$

Si $Q_{dB}^{desconocido} > Q_{dB}^{umbral} \rightarrow$ El espectro **puede ser identificado**

Si $Q_{dB}^{desconocido} < Q_{dB}^{umbral} \rightarrow$ El espectro **no es identificado**

Por último, la *Tabla C2.4* recoge las posiciones de las Bandas Fundamentales de cada uno de los espectros de la librería de referencia.

Posición de la Banda Fundamental (cm ⁻¹)									
Esp 1:	942	Esp 6:	1403	Esp 11:	1138	Esp 16:	1304	Esp 21:	1487
Esp 2:	1392	Esp 7:	1115	Esp 12:	1187	Esp 17:	1067		
Esp 3:	1401	Esp 8:	1134	Esp 13:	1364	Esp 18:	1192		
Esp 4:	1139	Esp 9:	1475	Esp 14:	1383	Esp 19:	1402		
Esp 5:	1142	Esp 10:	1399	Esp 15:	1116	Esp 20:	1104		

Tabla C2.4: Posiciones de las Bandas Fundamentales de los espectros patrones