

Instituto Universitario para el Desarrollo Tecnológico y la Innovación en Comunicaciones



Trabajo Fin de Máster

Prototipo de sistema para la determinación de contaminantes en el agua mediante métodos espectroscópicos

Autor: Dña. María Noemí Armas Déniz

Tutores: Dr. D. Juan Luis Navarro Mesa
Dr. D. Eduardo Hernández Pérez

Fecha: Enero de 2017

Instituto Universitario para el Desarrollo Tecnológico y la Innovación en Comunicaciones

Hoja de Firmas del TFM



Prototipo de sistema para la determinación de contaminantes en el agua mediante métodos espectroscópicos

Autor:

Fdo.: Dña. María Noemí Armas Déniz

Tutores:

Fdo.: Dr. D. Juan Luis Navarro Mesa

Fdo.: Dr. D. Eduardo Hernández Pérez

Fecha: Enero de 2017

Instituto Universitario para el Desarrollo Tecnológico y la Innovación en Comunicaciones

Hoja de Evaluación del TFM



Prototipo de sistema para la determinación de contaminantes en el agua mediante métodos espectroscópicos

Calificación:

Presidente

Secretario

Vocal

Fdo.:

Fdo.:

Fdo.:

Fecha: Enero de 2017

Índice de Contenidos

Capítulo 1: Introducción.....	15
1.1 Antecedentes.	15
1.1.1 Tipología de métodos analíticos.....	16
1.1.2 Parámetros de calidad.....	20
1.1.3 Estudios de métodos espectroscópicos en determinación de contaminantes.	22
1.2 Objetivos	24
1.3 Estructura de la memoria.....	24
Capítulo 2: Instrumentación.	27
2.1 Fuente de radiación	28
2.2 Cubetas	29
2.3 Selector de longitud de onda	31
2.4 Detector	34
Capítulo 3: Estudio y fabricación de espectrómetro.....	37
3.1 Descripción del láser.	38
3.2 Diseño y estudio del espectrofotómetro.	39
3.3 Características de la cámara.	46
Capítulo 4: Recogida de muestras y generación de la base de datos.....	47
4.1 Recogida de las muestras en escenarios reales.....	47
4.2 Generación de la Base de Datos.	48
Capítulo 5: Extracción de características y clasificador	51
5.1 Extracción de características	51
5.1.1 Características estadísticas y de textura	51
5.1.2 Entropía de Permutación	54
5.2 El Clasificador Utilizado	57
5.2.1 Clasificación binaria mediante SVM	57
5.2.2 Clasificación multiclase	61
Capítulo 6: Resultados y conclusiones.	63
6.1 Resultados	63
6.1.1 Cuantificación de la concentración de contaminante en función del aceite.....	64
6.1.1.1 Iluminación con 1 láser.	64

6.1.1.2 Iluminación con 2 láseres.	73
6.1.2 Cuantificación “ciega” de la concentración de contaminante.	81
6.1.2.1 Iluminación con 1 láser.	81
6.1.2.2 Iluminación con 2 láseres.	85
6.2 Conclusiones.....	89
Capítulo 7: Conclusiones y líneas futuras.....	91
7.1 Conclusiones.....	91
7.2 Líneas futuras	92
Bibliografía.	93
Anexo 1: Lista de acrónimos	97

Índice de Ilustraciones

Ilustración 1: Instrumentación métodos fluorescentes.....	19
Ilustración 2: Esquema montaje detección clorofila.....	22
Ilustración 3: Descubrimiento exoplanetas.....	23
Ilustración 4: Instrumentación medición absorción.....	28
Ilustración 5: Lámpara halógena de tungsteno.....	29
Ilustración 6: Lámpara de deuterio.....	29
Ilustración 7: Lámpara de arco de xenón.....	29
Ilustración 8: Cubetas muestras líquidas.....	30
Ilustración 9: Filtro de vidrio.....	31
Ilustración 10: Dispersión de luz por un prisma.....	33
Ilustración 11: Esquema prisma monocromador.....	34
Ilustración 12: Diagrama fototubo.....	35
Ilustración 13: Fototubo.....	36
Ilustración 14: Diagrama tubo fotomultiplicador.....	36
Ilustración 15: Esquema espectrofotómetro.....	37
Ilustración 16: Estructura puntero láser.....	38
Ilustración 17: Proceso de diseño.....	40
Ilustración 18: Proceso de impresión(1).....	41
Ilustración 19: Proceso de impresión (2).....	41
Ilustración 20: Diagrama 3d espectrofotómetro versión 1 y 2.....	41
Ilustración 21: Alzado 3d espectrofotómetro versión 1 y 2.....	42
Ilustración 22: Perfil 3d espectrofotómetro versión 1 y 2.....	42
Ilustración 23: Planta 3d espectrofotómetro versión 1 y 2.....	42
Ilustración 24: Espectrómetro versión 1(1)	43
Ilustración 25: Espectrómetro versión 1(2)	43
Ilustración 26: Espectrómetro versión 2(1)	43
Ilustración 27: Espectrómetro versión 2(2)	43
Ilustración 28: Diagrama 3d espectrofotómetro versión 3.....	44
Ilustración 29: Corrección pieza espectrofotómetro versión 3.....	45
Ilustración 30: Espectrómetro versión 3(1)	45
Ilustración 31: Espectrómetro versión3(2) 1láser	45
Ilustración 32: Espectrómetro versión3(3) 2láseres	45
Ilustración 33: Cámara web(1)	46
Ilustración 34: Cámara web(2)	46
Ilustración 35: Identificación muestra.....	48

Ilustración 36: Imagen2mLH2O+4mLAc.20.50(1láser).....	50
Ilustración 37: Imagen2mLH2O+4mLAc.20.50(2láseres).....	50
Ilustración 38: Imagen4mLH2O+2mLAc.20.50(1láser).....	50
Ilustración 39: Imagen4mLH2O+2mLAc.20.50(2láseres).....	50
Ilustración 40: Imagen5mLH2O+1mLAc.20.50(1láser).....	50
Ilustración 41: Imagen5mLH2O+1mLAc.20.50(2láseres).....	50
Ilustración 42: SVM linealmente separable.....	58
Ilustración43: SVM no linealmente separable inducida por una función kernell.....	59
Ilustración 44: SVM con margen blando.....	59

Índice de Tablas

Tabla 1: Color absorbido y color complementario.....	32
Tabla 2: Especificaciones técnicas puntero láser.....	38
Tabla 3: Especificaciones técnicas cámara	46
Tabla 4: Resultados cuantificación en función de concentración del contaminante usando solo Características de Textura: a) con fondo, b) sin fondo (1 láser)	64
Tabla 5: Resultados cuantificación en función de concentración del contaminante usando solo parámetros estadísticos: a) con fondo, b) sin fondo (1 láser)	65
Tabla 6: Resultados cuantificación en función de concentración del contaminante usando solo Coseno: a) con fondo, b) sin fondo (1 láser)	66
Tabla 7: Resultados cuantificación en función de concentración del contaminante usando solo Permution Entropy: a) con fondo, b) sin fondo (1láser).....	66
Tabla 8: Resultados cuantificación en función de concentración del contaminante usando solo Ángulo: a) con fondo, b) sin fondo (1 láser).....	67
Tabla 9: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura y parám. estadísticos: a) con fondo, b) sin fondo (1 láser)	67
Tabla 10: Resultados cuantificación en función de concentración del contaminante combinando Coseno y Permution Entropy: a) con fondo, b) sin fondo (1 láser)	68
Tabla 11: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura, parám. estadísticos y Coseno: a) con fondo, b) sin fondo (1 láser)	69
Tabla 12: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura, parám. estadísticos y Permution Entropy: a) con fondo, b) sin fondo (1 láser).....	69
Tabla 13: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura, parám. estadísticos y Ángulo: a) con fondo, b) sin fondo (1 láser)	70
Tabla 14: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura, parám. estadísticos, Permution Entropy y Ángulo: a) con fondo, b) sin fondo (1 láser)	70
Tabla 15: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura, parám. estadísticos, Permution Entropy, Ángulo y Coseno: a) con fondo, b) sin fondo (1 láser)	71
Tabla 16: Resumen Cuantificación en función de concentración del contaminante (1 láser)	71
Tabla 17: Resultados cuantificación en función de concentración del contaminante usando solo Característ. de textura: a) con fondo, b) sin fondo (2 láseres)	73
Tabla 18: Resultados cuantificación en función de concentración del contaminante usando solo Parámetros estadísticos: a) con fondo, b) sin fondo (2 láseres).....	74

Tabla 19: Resultados cuantificación en función de concentración del contaminante usando solo Coseno: a) con fondo, b) sin fondo (2 láseres)	74
Tabla 20: Resultados cuantificación en función de concentración del contaminante usando solo Permution Entropy: a) con fondo, b) sin fondo (2 láseres)	75
Tabla 21: Resultados cuantificación en función de concentración del contaminante usando solo Ángulo: a) con fondo, b) sin fondo (2 láseres)	75
Tabla 22: Resultados cuantificación en función de concentración del contaminante usando con característ. de textura y parám. estadísticos: a) con fondo, b) sin fondo (2 láseres).....	76
Tabla 23: Resultados cuantificación en función de concentración del contaminante combinando Coseno y Pemutiton Entropy: a) con fondo, b) sin fondo (2 láseres)	77
Tabla 24: Resultados cuantificación en función de concentración del contaminante combinando características de textura, parámetros estadísticos y Coseno: a) con fondo, b) sin fondo (2 láseres)	77
Tabla 25: Resultados cuantificación en función de concentración del contaminante combinando características de textura, parámetros estadísticos y Permution Entropy: a) con fondo, b) sin fondo (2 láseres).....	78
Tabla 26: Resultados cuantificación en función de concentración del contaminante combinando características de textura, parámetros estadísticos y Ángulo: a) con fondo, b) sin fondo (2 láseres)	78
Tabla 27: Resultados cuantificación en función de concentración del contaminante combinando características de textura, parámetros estadísticos Permution Entropy y Ángulo: a) con fondo, b) sin fondo (2 láseres).....	79
Tabla 28: Resultados cuantificación en función de concentración del contaminante combinando características de textura, parámetros estadísticos Permution Entropy, Ángulo y Coseno: a) con fondo, b) sin fondo (2 láseres)	79
Tabla 29: Resumen Cuantificación en función de concentración del contaminante (2 láseres)	80
Tabla 30: Resultados cuantificación “ciega” usando solo Características de textura: a) con fondo, b) sin fondo (1láser).....	82
Tabla 31: Resultados cuantificación “ciega” usando solo Parámetros estadísticos: a) con fondo, b) sin fondo (1láser)	82
Tabla 32: Resultados cuantificación “ciega” usando solo Coseno: a) con fondo, b) sin fondo (1láser)	82
Tabla 33: Resultados cuantificación “ciega” usando solo Permution Entropy: a) con fondo, b) sin fondo (1láser)	83
Tabla 34: Resultados cuantificación “ciega” usando solo Ángulo: a) con fondo, b) sin fondo (1láser)	83

Tabla 35: Resultados cuantificación “ciega” usando con característ. de textura y parám. estadísticos: a) con fondo, b) sin fondo (1láser).....	83
Tabla 36: Resultados cuantificación “ciega” combinando Coseno y Permution Entropy: a) con fondo, b) sin fondo (1 láser)	83
Tabla 37: Resultados cuantificación “ciega” usando característ. de textura, parám. estadísticos y Coseno: a) con fondo, b) sin fondo (1láser)	83
Tabla 38: Resultados cuantificación “ciega” usando característ. de textura, parám. estadísticos y Permution Entropy: a) con fondo, b) sin fondo (1láser).....	83
Tabla 39: Resultados cuantificación “ciega” combinando característ. de textura, parám. estadísticos y Ángulo: a) con fondo, b) sin fondo (1láser).....	83
Tabla 40: Resultados cuantificación “ciega” combinando característ. de textura, parám. estadísticos, Permution Entropy y Ángulo: a) con fondo, b) sin fondo (1 láser).....	83
Tabla 41: Resultados cuantificación “ciega” combinando característ. de textura, parám. estadísticos, Permution Entropy, Ángulo y Coseno: a) con fondo, b) sin fondo (1 láser).....	84
Tabla 42: Resumen Cuantificación “ciega” (1 láser).....	84
Tabla 43: Resultados cuantificación “ciega” usando solo Características de textura: a) con fondo, b) sin fondo (2láseres).....	86
Tabla 44: Resultados cuantificación “ciega” usando solo Parámetros estadísticos: a) con fondo, b) sin fondo (2láseres).....	86
Tabla 45: Resultados cuantificación “ciega” usando solo Coseno: a) con fondo, b) sin fondo (2 láseres)	86
Tabla 46: Resultados cuantificación “ciega” usando solo Permution Entropy: a) con fondo, b) sin fondo (2 láseres).....	87
Tabla 47: Resultados cuantificación “ciega” usando solo Ángulo: a) con fondo, b) sin fondo (2 láseres)	87
Tabla 48: Resultados cuantificación “ciega” usando característ. de textura y parám. estadísticos: a) con fondo, b) sin fondo (2 láseres)	87
Tabla 49: Resultados cuantificación “ciega” combinando Coseno y Pemutiton Entropy: a) con fondo, b) sin fondo (2 láseres).....	87
Tabla 50: Resultados cuantificación “ciega” usando característ. de textura, parám. estadísticos y Coseno: a) con fondo, b) sin fondo (2 láseres).....	87
Tabla 51: Resultados cuantificación “ciega” usando característ. de textura, parám. estadísticos y Permution Entropy: a) con fondo, b) sin fondo (2 láseres).....	87
Tabla 52: Resultados cuantificación “ciega” combinando característ. de textura, parám. estadísticos y Ángulo: a) con fondo, b) sin fondo (2 láseres)	87

Tabla 53: Resultados cuantificación “ciega” combinando característ. de textura, parám. estadísticos, Permutation Entropy y Ángulo: a) con fondo, b) sin fondo (2 láseres)	87
Tabla 54: Resultados cuantificación “ciega” combinando característ. de textura, parám. estadísticos, Permutation Entropy, Ángulo y Coseno: a) con fondo, b) sin fondo (2 láseres)	88
Tabla 55: Resumen Cuantificación “ciega” (2 láseres)	88
Tabla 56: Conclusiones Cuantificación en función de la concentración del contaminante	89
Tabla 57: Conclusiones Cuantificación "ciega"	90

Capítulo 1: Introducción.

En este apartado se realiza una breve presentación acerca de los métodos analíticos en los que se basa el presente documento. Se describen precedentes, clasificación de los métodos de análisis y parámetros de calidad; nombrando además varios estudios en los que se han utilizado dichos análisis. Asimismo, se indican cuáles son los objetivos a alcanzar en este proyecto a la vez que se presenta cuál es la estructura que conforma del presente documento.

1.1 Antecedentes

La absorción y emisión de energía radiante que realizan las moléculas y los átomos constituye el fundamento de muchos procedimientos en Química Analítica, obteniendo tanto resultados cuantitativos como cualitativos. El Trabajo Fin de Máster que presentamos viene de una confluencia entra la química analítica que la estudiante trae como formación de grado y la formación adquirida en este Máster dentro en el campo de las Tecnologías de la Información y las Comunicaciones.

Los métodos ópticos son métodos instrumentales basados en la interacción de la radiación electromagnética con un analito para identificarlo o determinar su concentración. A su vez, estos métodos se pueden separar en dos subgrupos: espectroscópicos y no espectroscópicos, siendo el primer grupo el que se va a estudiar e implementar en este trabajo. La principal característica de éste procedimiento es que el analito sufre procesos de absorción, emisión o luminiscencia [1].

El funcionamiento de este aparato analítico basado en esta técnica es el de hacer pasar un haz de luz a través de la muestra problema y cuya composición queremos determinar. Cuando esto ocurra las moléculas de dicha sustancia se excitarán al absorber parte de esta radiación, por lo que transmite un haz de menor energía radiante. Dicha luminiscencia se descompondrá en diferentes colores al llegar al difractor, siendo ésta la respuesta que se recoge al finalizar todo el proceso [2].

En función de los elementos químicos que esté compuesto el analito a estudiar se obtendrán diferentes espectros de absorción; ya que cada componente absorbe a una determinada longitud de onda y no depende del compuesto del que forma parte.

La espectrofotometría se usa para diversas aplicaciones tales como análisis cuantitativo y cualitativo de soluciones desconocidas en un laboratorio de investigación, estandarización de colores de diversos materiales (plásticos y pinturas), detección de niveles de contaminación en aire y agua y determinación de trazas de impurezas en alimentos y en reactivos.

Finalmente, se puede decir que las principales ventajas de este procedimiento son su precisión, versatilidad, rapidez, facilidad de uso y bajo costo. Dichos aparatos han mejorado considerablemente en las dos primeras características en los últimos años con los avances de la tecnología, considerándose indispensables en un laboratorio de química analítica [3].

1.1.1 Tipología de métodos analíticos

Los análisis químicos proporcionan información acerca de la composición de una muestra, pudiendo clasificarse en cualitativos y cuantitativos. En ambos procesos los resultados se obtienen por la medida de una propiedad física que está relacionada con el componente de interés.

Los análisis cualitativos se caracterizan por aportar información útil sobre especies moleculares o atómicas o reconocer en ella determinados grupos funcionales; mientras que los cuantitativos presentan los resultados con datos numéricos, expresándolos mediante porcentajes.

A su vez los métodos cuantitativos se pueden dividir en dos grandes grupos, clásicos e instrumentales. Los primeros se fundamentan en las propiedades químicas de la especie a analizar, mientras que los segundos se fundamentan en propiedades físico-químicas.

Los **métodos clásicos** son aquellos que se fundamentan en las propiedades químicas de la especie química que se quiere analizar, basándose en las reacciones cuantitativas en

las que se alcanza el equilibrio químico con la intervención del analito. Se pueden clasificar en dos grupos:

- Análisis gravimétricos. Consiste en separar la especie química que se quiere determinar del resto de los componentes de la muestra y proceder a su pesada. La instrumentación básica está compuesta por balanzas analíticas, estufas de secado regulables a las temperaturas que indiquen los métodos de aplicación.
- Análisis volumétricos. Se determina cuánto volumen de una disolución se necesita para reaccionar exactamente con otra de concentración y volumen conocidos. En este tipo de determinaciones es necesario añadir un indicador que nos indique el instante en el que la reacción llegue al equilibrio.

Las ventajas de utilizar este método analítico es que no se hace necesario usar un equipo o una cualificación técnica sofisticada, pero suelen suministrar límites de cuantificación bajos y las incertidumbres son por el contrario elevadas [4].

Los **métodos instrumentales** están basados en propiedades químico-físicas, pueden dividirse en tres subgrupos:

- Métodos ópticos. Se basan en la interacción de la radiación electromagnética con la materia. A su vez, se pueden clasificar en métodos espectroscópicos y no espectroscópicos.
- Métodos eléctricos. Se caracterizan por la determinación de cualquier variación en las propiedades físicas como consecuencia del paso de una corriente eléctrica.
- Métodos cromatográficos. Se aplica este procedimiento en aquellas muestras en las que junto al analito a analizar existen especies semejantes. Los componentes de la muestra se distribuyen entre una fase móvil, arrastra los componentes de la muestra y puede ser líquido o gas, y fase estacionaria, retienen los componentes de la muestra y por la que fluye la fase móvil, pudiendo ser líquido o sólido.

Los métodos espectroscópicos se dividen a su vez en métodos de absorción en los que se produce una absorción de energía por parte de la materia, y métodos de emisión en los que se produce una emisión de energía o radiación por parte de una materia que ha absorbido energía previamente.

En los métodos de emisión atómica, las moléculas que son excitadas a altos niveles de energía que pueden caer a niveles inferiores emitiendo radiación. Si estas partículas son excitadas por una fuente de energía a alta temperatura se denomina espectroscopía de emisión atómica o fotometría de llama si la forma de excitación es a través de una llama; mientras que si se utiliza una luz para excitar a dichas partículas se le denomina espectroscopía de fluorescencia [5].

La Espectroscopía de Fluorescencia Atómica/*Atomic Fluorescence Spectroscopy* (AFS) es una técnica de elementos múltiples utilizada para el análisis de trazas de metales en agua de mar, sustancias biológicas y muestras agrícolas. De estos espectros atómicos se puede obtener información cualitativa y cuantitativa de los elementos de la muestra [6].

En lo concerniente al instrumental necesario para realizar este tipo de análisis (ver **Ilustración 1**), encontramos dos tipos: métodos dispersivos y no dispersivos. Los primeros están compuestos por una fuente de luz, un atomizador, un analizador, un detector, un procesador de señal y un dispositivo de lectura. La fuente ideal para la fluorescencia atómica es el láser, pero la fuente más común es la lámpara de descarga sin electrodo.

Los no dispersivos se componen por una fuente de luz, un atomizador y un detector. No se necesita un analizador. Cuando una lámpara de descarga sin electrodo sirve como fuente de excitación la radiación emitida es la de un elemento simple [7].

Las regiones que proporcionan datos atómicos espectrales son la región del ultravioleta visible (UV-V), ya que tiene energía suficiente para promover los electrones de valencia a estados mayores de energía, y región de rayos X, siendo éste el más energético al promover los electrones internos del átomo.

Las regiones que proporcionan datos atómicos espectrales son la región del ultravioleta visible (UV-V), ya que tiene energía suficiente para promover los electrones de valencia a estados mayores de energía, y región de rayos X, siendo éste el más energético al promover los electrones internos del átomo.

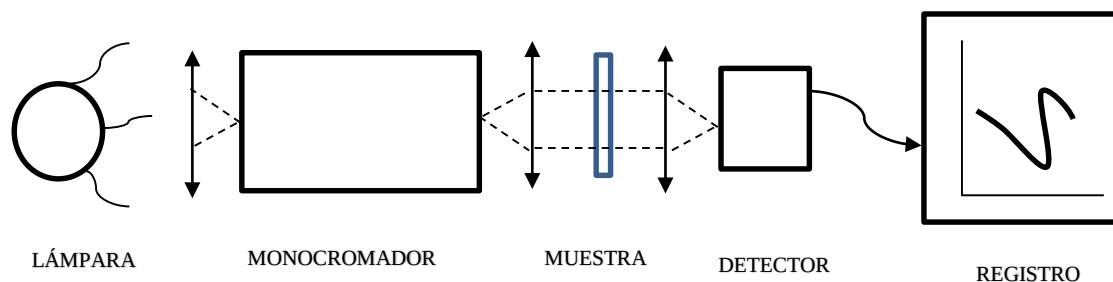
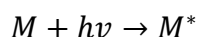


Ilustración 1: Instrumental métodos fluorescentes

Veamos algunos conceptos de interés en el análisis espectroscópico. Éste se apoya en el cambio en los estados energéticos característicos de la materia como consecuencia de la radiación electromagnética interactuando con la materia. En el momento en el que esta radiación interacciona con la materia se produce absorción de energía, pasando de un estado de energía fundamental a otro de energía superior denominado excitado. Veamos una expresión que representa este fenómeno [8].



donde M representa el nivel de energía fundamental, $h\nu$ simboliza la radiación electromagnética y M^* es el nivel de energía superior.

En dicho proceso de absorción, la materia tiende a perder la energía en exceso, generalmente a través de la colisión de unas partículas con otras produciendo calor. También se puede eliminar el exceso emitiendo energía, originándose en ese caso métodos fluorescentes y fluminoscentes.

En los métodos de absorción un factor importante que debe de cumplirse es que la energía de la radiación incidente debe ser igual a la diferencia de energía entre dos estados permitidos de la sustancia.

Cualquier sistema elemental posee unos estados de energía cuantizados.

$$h_{\mu} = E_f - E_i$$

donde h_{μ} es la energía de radiación incidente, E_f representa el valor de energía final y E_i indica el valor de energía inicial o fundamental.

Una vez expuesta la tecnología de referencia, veamos qué opción hemos elegido para este Trabajo de Fin de Máster (TFM) para la determinación de contaminantes en agua.

Hemos optado por una técnica analítica basada en el método óptico de tipo espectroscópico. Esta decisión se ha tomado por los siguientes motivos:

- Se pueden aplicar en números campos. Por ejemplo, hay numerosos compuestos orgánicos e inorgánicos que absorben en UV-V.
- Ofrecen alta sensibilidad, pudiendo determinar concentraciones molares (M) mínimas en amplios rangos.
- Ofrecen alta selectividad, se puede encontrar longitudes de onda a la que sólo absorbe la especie química que nos interesa, eliminándose de esta forma interferencias.
- Poseen buena precisión y exactitud.
- Son fáciles de manejar, cómodos y rápidos.
- Hay fuentes ópticas comerciales en el mercado que permiten desarrollar prototipos de bajo coste.

1.1.2 Parámetros de calidad

En primer lugar, se presenta la **exactitud**, siendo el grado de concordancia entre un resultado obtenido experimentalmente y el valor real. Se puede cuantificar mediante error absoluto (E_a) o error relativo (E_r), de tal manera que:

$$E_a = \mu - x \quad E_r = E_a / x$$

donde μ es la media de las medidas y x valor verdadero. En función del valor de E_r se puede considerar el método analítico como veraz (si E_r es $<1\%$), moderadamente veraz si se cumple $1\% < E_r < 5\%$, y poco veraz si E_r es $>5\%$.

El siguiente parámetro es la **precisión**, que representa el grado de dispersión de una serie de resultados obtenidos con replicados de la misma muestra; por lo que cuanto más grande sea este valor, el valor de su reproducibilidad aumentará. De forma matemática se puede expresar mediante:

$$CV(\%) = (s/\mu) * 100$$

donde CV es el coeficiente de variación, s representa la desviación normal de una serie de medidas y μ es la media.

La tercera característica es la **sensibilidad**, que relaciona la variación de la señal analítica medida con la variación de la concentración del analito que la provoca:

$$S = \Delta x / \Delta C$$

donde Δx es el incremento de la señal medida frente a un incremento, ΔC , de la concentración del analito.

Se define al **límite de detección (LD)** como la cantidad mínima de analito que origina una señal analítica (X_{LD}) estadísticamente distinta a la generada por el blanco (μ_B), de tal manera que

$$X_{LD} = \mu_B + 3s_B$$

donde s_B es el valor de la desviación estándar de la señal del blanco.

El **límite de cuantificación (LC)** es la concentración mínima de analito que origina una señal analítica (X_{LD}) diez veces superior a la generada por el blanco (μ_B)

$$X_{LD} = \mu_B + 10s_B$$

El **intervalo de linealidad** es el rango de concentraciones del analito en el que el método es aplicable, lo que significa que la relación concentración/señal es ajustable a la ecuación de una recta. El límite inferior de este rango viene dado por el límite de cuantificación.

Otra característica es la **incertidumbre**, que marca el intervalo de valores dentro del que se espera encontrar el valor verdadero de la magnitud medida. Se trataría del rango en el que, con una probabilidad determinada, se encuentra el valor real del analito.

El último parámetro es la **selectividad o especificidad**, es la capacidad de un método para originar resultados que sólo dependan de la cantidad de analito presente en la muestra. Se encuentra relacionada con las interferencias que puedan producir los diferentes componentes de la matriz de la muestra: cuanto menos influyan en la medida más selectivo será el método [9].

1.1.3 Estudios de métodos espectroscópicos en determinación de contaminantes.

Como se ha mencionado con anterioridad, los métodos espectroscópicos son utilizados para determinar diversas sustancias debido a la facilidad de uso y baja complejidad de su instrumentación; tal y como se demuestra en los siguientes estudios en los que se ha utilizado este análisis.

-Detección de nitrógeno en un humedal artificial.

Su objetivo fue determinar la absorción de este elemento utilizando técnicas espectroscópicas de fluorescencia y reflectancia. Las plantas utilizan el nitrógeno para sintetizar proteínas, enzimas y clorofila; por lo que es posible estimar la concentración de este elemento en la hoja al establecer una relación de ésta con la fluorescencia de la molécula de clorofila.

El montaje seguido para determinar la presencia y concentración de estas sustancias fue la que aparece en la **Ilustración 2**. Se tomaron 6 espectros en diferentes zonas la hoja, irradiándola con un diodo emisor de luz (LED) de alta luminiscencia para obtener la excitación molecular. El diodo y su voltaje de operación, permanecieron sin cambio a lo largo de todo el periodo de medición.

Se obtuvieron valores del contenido porcentual de nitrógeno, que son compatibles con los encontrados con la utilización de una técnica química estándar. Estos resultados alientan el diseño de un método óptico, aplicable in situ, y en tiempo real para la determinación del nitrógeno acumulado en las plantas de un humedal artificial [10].



Ilustración 2: Esquema montaje detección clorofila

- Determinación de materia orgánica disuelta en suelos sometidos a secado.

También se ha utilizado este método de análisis en la determinación de materia orgánica disuelta en suelos sometidos a secado; ya que este tipo de material extracelular soluble en agua emite fluorescencia por unidad de carbono orgánico cuantitativamente a un menor grado que el material humidificado soluble en agua en el suelo [11].

- Descubrimiento de exoplanetas.

Por definición un exoplaneta se encuentra fuera del sistema solar, por lo que detectarlos no es algo sencillo. Una de las técnicas utilizadas para detectarlos es discernir las variaciones en el espectro de luz emitidas por la estrella.

A través de estas medidas podemos determinar la presencia y el movimiento descrito por la estrella, así como las características del posible planeta que la acompaña. Para ello se observa las variaciones cíclicas de brillo de ciertas estrellas; puesto que el paso de un planeta entre la Tierra y la estrella lleva una modificación en la luz que nos llega de esta última.

Se explica este concepto en la **Ilustración 3**, donde se explica que el movimiento de las líneas oscuras en el espectro de la estrella es la firma del movimiento del planeta y la estrella en torno a su centro de gravedad [12].

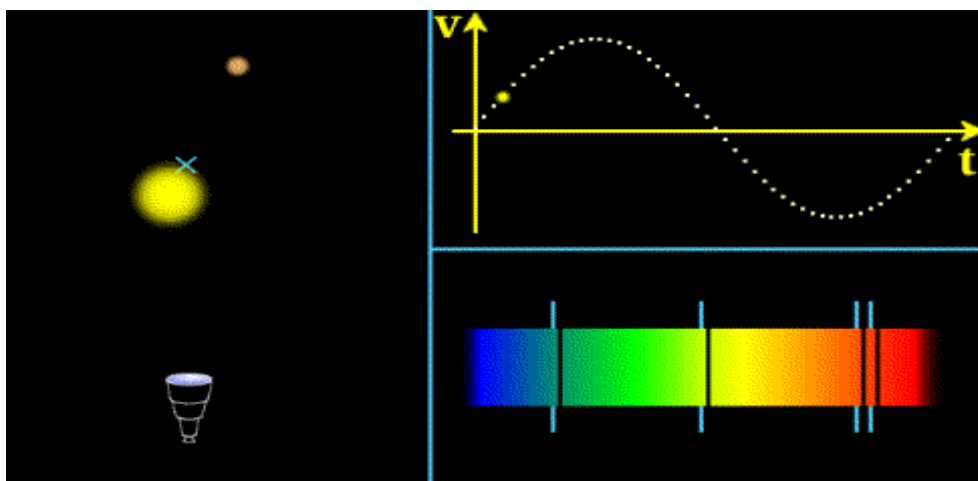


Ilustración 3: Descubrimiento de exoplanetas

1.2 Objetivos

Los objetivos principales que se quieren conseguir con la ejecución de este proyecto son la fabricación del fotómetro con el que se va a obtener los espectros de las sustancias problemáticas, estudiar la presencia de aceites como agentes contaminantes en el agua e identificar los diferentes tipos mediante análisis de imágenes espectrográficas.

Del mismo, se procurará estudiar las respuestas de estos elementos con distintas fuentes de iluminación, así como analizar su comportamiento al variar la frecuencia de estas fuentes.

1.3 Estructura de la memoria

En lo referente al esqueleto de esta memoria, se estructura de la siguiente forma:

- **Capítulo 1: Antecedentes.** Se describen las principales características de las técnicas analíticas, así como los parámetros de calidad que nos proporcionan los resultados obtenidos de dichos procedimientos. Asimismo, se indican varios estudios en los que se emplean dichas técnicas.
- **Capítulo 2: Instrumentación.** Se explica cada uno de los componentes que conforman la maquinaria necesaria para realizar los análisis espectroscópicos.
- **Capítulo 3: Estudio y fabricación de espectrómetro.** Se indica el procedimiento con el que se ha construido el dispositivo con el que se va a analizar.
- **Capítulo 4: Recogida de muestras y generación de la Base de Datos.** Se explican los pasos a seguir para recopilar las muestras de aguas y el modo en que se recopiló las imágenes para crear la Base de Datos (BBDD).
- **Capítulo 5: Extracción de características y clasificador.** En este capítulo se indica los parámetros utilizados para comparar las imágenes, así como las características del clasificador utilizado.
- **Capítulo 6: Resultados y conclusiones.** Se reflejan los resultados tras seguir los pasos descritos en los capítulos anteriores y se indica cuál es la mejor característica para obtener el mejor resultado posible.

- **Capítulo 7: Conclusiones y líneas futuras.** Se detallan las conclusiones a las que se ha llegado con el desarrollo del trabajo, así como se plantean líneas futuras de investigación.
- **Bibliografía.**

Capítulo 2: Instrumentación.

Los espectrómetros son aparatos que pueden realizar mediciones de absorción molecular en el intervalo UV-visible. Dichos dispositivos pueden ser muy simples a la hora de indicar la presencia de compuestos en trabajos rutinarios. También pueden ser sofisticados para utilizarlos en trabajo muy especializado.

Antes de proceder al estudio de cómo fabricar nuestro prototipo, veamos la instrumentación necesaria para medir la absorción de radiación. Para este propósito es necesario contar con los siguientes elementos (ver **Ilustración 4**):

- Fuente de radiación. Tiene la misión de proporcionar una radiación suficiente y constante de energía. En función de la zona del espectro en la que nos encontremos, se usarán distintos tipos de lámparas.
- Dispositivo para seleccionar la longitud de onda. Es el encargado de seleccionar la longitud de onda nos va a permitir conseguir un mejor análisis. Este proceso consiste en dispersar la radiación policromática en sus longitudes de onda constituyentes, logrando así que no se produzcan interferencias. Los instrumentos más usados son el sistema de filtros y el monocromador.
- Dispositivo para la muestra. Está compuesto por celdas que en función de la región en la que se vaya a realizar el análisis estará formado por un material diferente.
- Detección y lectura. Estos elementos transforman la energía radiante en energía eléctrica; debiendo responder a la energía radiante en un amplio rango de longitudes de onda, medir la intensidad de radiaciones bajas y producir una señal eléctrica que se pueda amplificar.

En cuanto al rango de luz que usaremos para la fuente de radiación, en este TFM nos hemos decantado por ultravioleta visible. El motivo es que la luz a esta frecuencia es capaz de dar lugar a un espectro rico al interactuar con la materia de interés, en particular, al producirse absorción. Además, existe en el mercado una serie de dispositivos comerciales que nos permite tener ultravioleta a potencias útiles para nuestros experimentos [13].

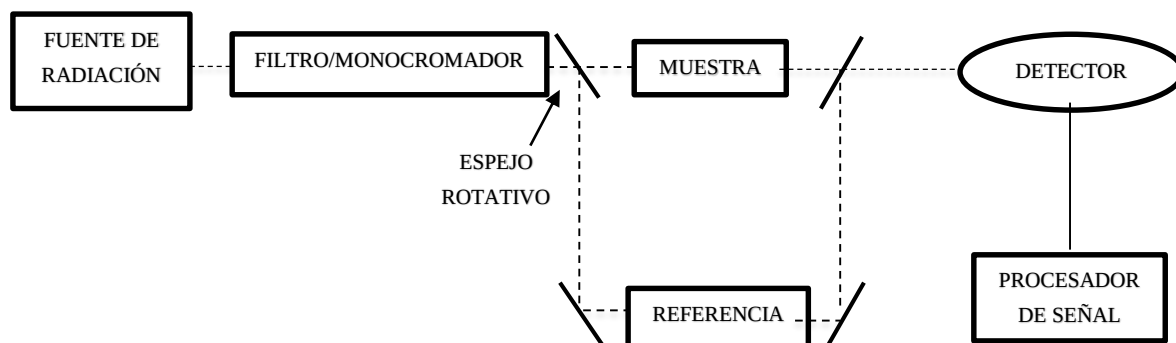


Ilustración 4: Instrumental medición de absorción

La base conceptual de la espectroscopía que utiliza este tipo de luz consiste en poder comparar la intensidad de la radiación absorbida a una longitud de onda específica, o un margen determinado, comparándola con otras disoluciones de concentración conocida que contengan la misma especie absorbente.

Para establecer esta relación se emplea la Ley de Lambert-Beer [8,14], por la que se establece que la absorbancia, magnitud que representa la fracción de luz absorbida, es directamente proporcional a la concentración; de tal manera que:

$$\log \frac{I_0}{I} = A = a * b * c$$

donde a es la cte. de proporcionalidad dada en $l/(cm \cdot g)$ ó $l/(cm \cdot mol)$, b es el espesor del material dado en cm , y c es la concentración de la muestra dada en g/l .

2.1 Fuente de radiación

La fuente de radiación debe proporcionar una salida estable de alta energía en un intervalo suficientemente amplio de longitudes de onda. Se utilizan normalmente fuentes continuas para las mediciones de absorción molecular, mientras que las fuentes *spectral-line* se emplean para la calibración de longitud de onda de un espectrómetro. Las que más se suelen utilizar son:

- Lámparas halógenas de tungsteno. Se caracterizan principalmente por tener una temperatura radiante de 3000K, por lo que tienen un espectro radiante máximo a una longitud de onda de aproximadamente $1,2 \mu m$ (ver **Ilustración 5**).
- Lámparas de deuterio. Emiten fuertemente en la región UV por debajo de los 330 nm (ver **Ilustración 6**).

- Lámparas de arco de xenón. Ofrecen un continuo que va desde menos de 190 nm y por encima de 1000 nm (ver **Ilustración 7**).

2.2 Cubetas

Los compartimientos de muestras varían su composición según el estado físico de la muestra a analizar. Para muestras líquidas se utilizan celdas fotométricas, estando ubicados en contenedores que se pueden calentar o enfriar con el fin de controlar la temperatura del líquido en la celda de muestra. La forma de la cubeta puede ser rectangular o cilíndrica, siendo una característica importante el volumen y la sección transversal a la vez que el material de la ventana (ver **Ilustración 8**) [5,6,7,8].

Un par de celdas con propiedades ópticas muy similares se conocen como celdas emparejadas. Una celda es la celda de la muestra, mientras que la otra, es la de referencia (o blanco), que contiene el disolvente o una solución de referencia. En espectrómetros de doble haz, la radiación atraviesa de forma simultánea o alternativamente a través de las dos celdas. En los instrumentos de un solo haz las celdas se mueven secuencialmente en el haz de radiación.



Ilustración 5: Lámpara halógena de tungsteno



Ilustración 6: Lámpara de deuterio



Ilustración 7: Lámpara de arco de xenón



Ilustración 8: Cubeta muestras líquidas

Para muestras gaseosas se utilizan celdas similares a las anteriores, teniendo en cuenta que a cualquier presión los gases se encuentran en celdas cerradas para su medición; mientras que las muestras sólidas se mantienen en compartimientos especiales. Para obtener unos adecuados resultados se debe realizar un correcto uso y mantenimiento de estos elementos, y para conseguirlo se seguirán las indicaciones que se a continuación:

- Inspeccionar si hay imperfecciones en las ventanas, sin tocar las superficies ópticas de las celdas con los dedos.
- Utilizar una única celda para toda una serie de mediciones, o asegurar que todas las celdas que se utilizan tienen la misma longitud de camino, b , para transmitancias similares.
- Al llenar las celdas, enjuáguelos a fondo con la solución a medir, y luego rellenar, asegurando que no haya burbujas de aire que se adhieran a las ventanas.
- Colocar las celdas dentro del instrumento con cuidado, asegurando que estén bien sentados en el compartimiento, utilizando la misma orientación de la celda cada vez que realice la medida.
- Si se utiliza un disolvente volátil, colocar una tapa sobre la celda para reducir la evaporación.
- Nunca almacenar soluciones en las celdas. Enjuagar a fondo cuando haya terminado y dejar secar en un lugar libre de polvo. Las soluciones básicas corroen y dejan grabados en las celdas si no se limpian correctamente.
- Mantener siempre las celdas en soportes o cajas para cubetas cuando no están dentro del instrumento.

2.3 Selector de longitud de onda

Los selectores de longitud de onda son necesarios en las mediciones espectrofotométricas para mejorar la selectividad y la sensibilidad de los instrumentos. Las máquinas más económicas usan un filtro para aislar la energía radiante y proporcionar una amplia banda de longitudes de onda, mientras que en aquellas donde es necesario variarla continuamente se usan monocromadores [5,6,7,8].

El método más básico en la utilización de un filtro que separe las longitudes de onda consiste en colocar una piedra de vidrio de color (ver **Ilustración 9**), absorbiendo la luz a cierta longitud de onda y permitiendo a las otras pasar.

Un dato importante que no hay que olvidar es la composición de la luz blanca, formada por siete colores diferentes (violeta, índigo, azul, verde, amarillo, naranja y rojo). En el momento en el que ésta cae sobre un objeto una parte se absorbe, mientras que otra es absorbida.

El componente de absorción y el color que se observa se suman a su vez para obtener la luz blanca; siendo dichos colores complementarios. En otras palabras, un objeto de un color particular muestra su color porque este color se transmite y su color complementario es absorbido.



Ilustración 9: Filtro de vidrio

Para seleccionar el filtro adecuado para medir la absorbancia de una sustancia es necesario entender los colores complementarios, de tal manera que el color del filtro debe ser complementario al color de la solución a medir, **Tabla 1**. Para comprender mejor esta idea supongamos que disponemos de una solución naranja, lo que implica que la luz naranja no está siendo absorbida pero sí lo hace su complementario (azul); por lo que el filtro que se va a utilizar en una solución naranja es azul.

Los filtros de absorción son simples y totalmente adecuados para muchas aplicaciones en el rango visible, aunque para rangos amplios se necesitan filtros de interferencia. Están formados por dos placas de vidrio (simples o en combinación), que se encuentran recubiertas con sustancias colorantes o pigmentos que absorben radiación en determinados rangos de longitudes de onda.

Estos filtros de interferencia están disponibles para la región del ultravioleta, visible e infrarrojo cercano. Las características de rendimiento de estos filtros son significativamente superiores a los filtros de absorción (coloreados).

Otro sistema que nos puede proporcionar selectivamente la radiación de una longitud de onda deseada es a través del monocromador, que a diferencia del filtro dispersa la radiación policromática para seleccionar un solo tipo de radiación.

Región de λ (nm)	Color absorbido	Color complementario de la luz residual
400-450	Violeta	Amarillo-verdoso
450-480	Azul	Amarillo
480-490	Verde-azul	Naranja
490-500	Azul-verde	Rojo
500-560	Verde	Púrpura
560-580	Amarillo-verde	Violeta
580-600	Amarillo	Azul
600-650	Naranja	Verde-azul
650-750	Rojo	Azul-verde

Tabla 1: Color absorbido y color complementario

Dentro de esta técnica de separación se encuentran el prisma y la rejilla. En el primer sistema se dispersa la luz en los colores que la conforman debido a la refracción de ésta al atravesar dicho prisma.

Las radiaciones de diferentes colores tienen distintas longitudes de onda que se refractan a distintas medidas, debido a la diferencia en el índice de refracción del vidrio para diferentes longitudes de onda.

En un monocromador de prisma, como se muestra en la **Ilustración 10**, se obtiene un fino haz de luz de la fuente cuando se hace pasar a través de una rendija de entrada. Este es entonces colimado en el prisma con la ayuda de una lente. Los rayos refractados se enfocan en una rendija de salida. El prisma se gira entonces en una forma predeterminada para proporcionar la longitud de onda deseada de la rendija de salida.

El otro sistema monocromador anteriormente mencionado es la rejilla, que se construye cortando o grabando una serie de ranuras paralelas espaciadas estrechamente de tal manera que forman una superficie densamente poblada (cerca de 500 a 2000 dientes por milímetro cuadrado) sobre la superficie lisa reflectante de un material sólido. La superficie se hace reflectante mediante una capa delgada de aluminio en él y el ataque químico se realiza con la ayuda de una herramienta de diamante de forma adecuada.

Su funcionamiento comienza en la rendija de entrada, donde un fino haz de la luz de la fuente cae sobre un espejo cóncavo, reflejándose en la rejilla en donde se dispersa para después dirigirse hacia la rendija de salida (ver **Ilustración 11**).

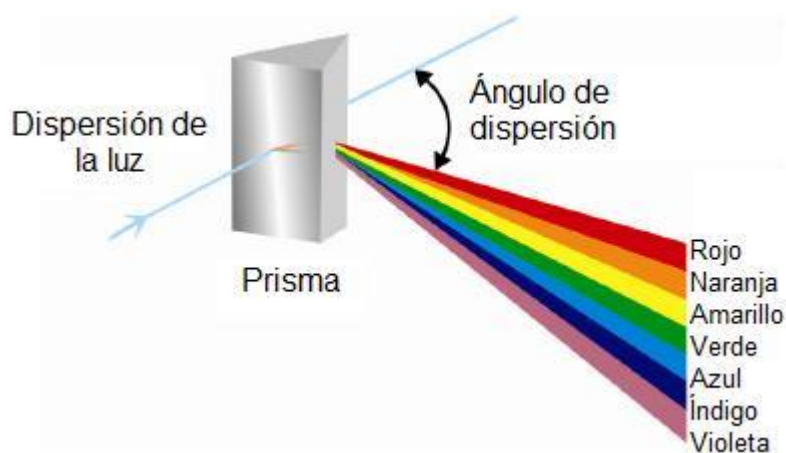


Ilustración 10: Dispersión de luz por un prisma

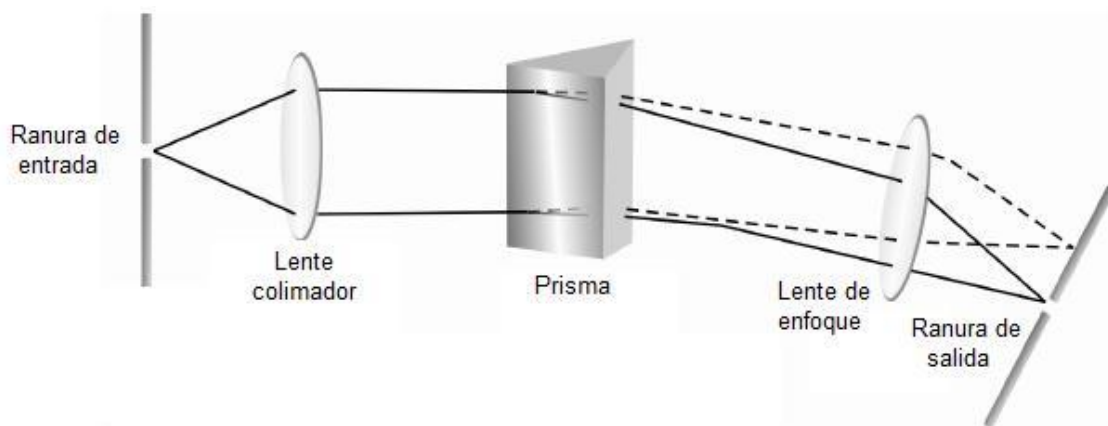


Ilustración 11: Esquema prisma monocromador

La gama de longitudes de onda aisladas por el monocromador se determina por el grado de dispersión de la rejilla y la anchura de la rendija de salida. La rotación de la rejilla de una manera predeterminada puede ser utilizada para obtener la longitud de onda deseada de la rendija de salida.

2.4 Detector

El siguiente elemento a describir es el detector. Dicho dispositivo se utiliza para convertir una señal luminosa a una señal eléctrica que puede ser adecuadamente medida y transformada en una salida [5,6,7,8].

Los detectores usados en la mayoría de los instrumentos generan una señal, que es lineal en la transmitancia, es decir, que responden linealmente a la potencia radiante que incide sobre ellos. Los valores de transmitancia se pueden cambiar logarítmicamente en unidades de absorbancia por una disposición eléctrica o mecánica en la señal de lectura del dispositivo. Hay tres tipos de detectores que se utilizan en espectrofotómetros modernos: fototubo, tubo fotomultiplicador y detector arreglo de diodos.

El fototubo (ver **Ilustración 12** y **13**) está compuesto por un cátodo fotoemisor y un ánodo en un tubo de vacío con una ventana de cuarzo, estando ambos sometidos a una alta diferencia de voltaje (100 V). Cuando un fotón entra en el tubo y golpea el cátodo, un electrón es expulsado y es atraído hacia el ánodo, lo que resulta en un flujo de corriente

que puede ser amplificado y medido. La respuesta del material fotoemisor es dependiente de la longitud de onda y fototubos diferentes están disponibles para las diferentes regiones del espectro.

Un tubo fotomultiplicador (ver **Ilustración 14**) se compone de una serie de electrodos llamados dínodos. El voltaje de los electrodos sucesivos se mantiene 50 a 90 voltios más positivo que el anterior. Cuando una radiación cae sobre el cátodo un electrón se emite a partir de ella.

Esto se acelera hacia el siguiente electrodo fotoemisor por la diferencia de potencial entre los dos. Aquí, libera muchos más electrones secundarios. Estas a su vez son acelerados hacia el electrodo de al lado donde cada electrón secundario libera más electrones. El proceso continúa hasta aproximadamente 10 etapas de amplificación. La salida final del tubo fotomultiplicador da una señal mucho mayor que la fotocelda.

El último sistema de detección es el arreglo de diodos. Cuando una radiación UV-VIS cae en el diodo, su conductividad aumenta significativamente y sabiendo que este aumento es proporcional a la intensidad de la radiación, se puede medir fácilmente. Dado que un gran número de diodos pueden ser organizados juntos, la intensidad en un número de longitudes de onda se puede medir de forma simultánea.

Aunque el arreglo de fotodiodos no es tan sensible como el tubo fotomultiplicador, la posibilidad de ser capaz de medir un gran número de longitudes de onda hace que sea un detector de elección en los instrumentos rápidos modernos [15,16].



Ilustración 12: Diaarama fototubo



Ilustración 13: Fototubo

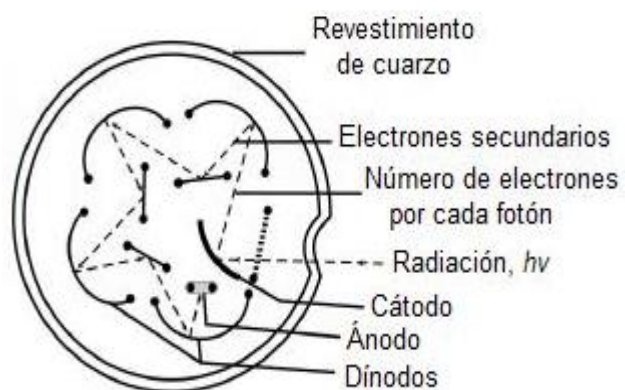


Ilustración 14: Diagrama tubo fotomultiplicador

Capítulo 3: Estudio y fabricación de espectrómetro.

El esquema genérico que hemos usado para construir nuestro espectrómetro es el que se muestra en la **Ilustración 15**, donde **A** representa el láser que excita el resplandor fluorescente en la muestra, **B** es el haz de luz que entra por la rendija del espectrómetro, **C** comprende la rejilla de difracción que separa la luz por el color y **D** muestra un ejemplo de espectro de intensidades obtenido a partir de una imagen tomada con cámara y software de gráficos de colores [2].

Para explicar de una forma más concisa cuál ha sido el proceso de construcción del dispositivo, se va a dividir este capítulo en los siguientes apartados:

- Descripción del láser utilizado.
- Diseño y estudio del espectrofotómetro creado.
- Características de la cámara utilizada durante los experimentos.

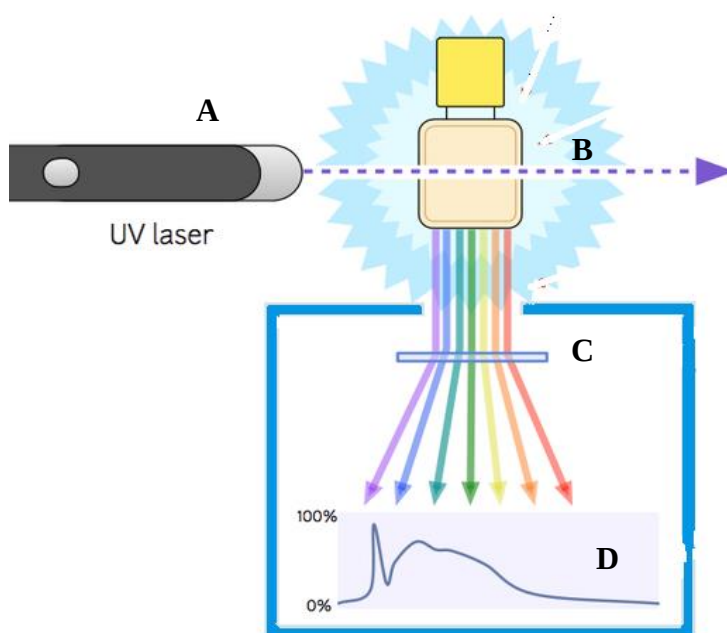


Ilustración 15: Esquema espectrofotómetro

3.1 Descripción del láser.

Como se ha explicado en capítulos anteriores, en la determinación de analitos mediante los métodos espectroscópicos es fundamental el paso de un haz de luz por la muestra a analizar. Con este proceso se consigue excitar a las moléculas y así obtener la descomposición en colores de la luminiscencia al llegar al difractor, siendo esta la respuesta que se recoge al finalizar con el experimento.

En este TFM el haz de luz procede de un puntero láser de ultravioleta-visible (**Ilustración 16**) y cuyas características técnicas se presentan en la **Tabla 2 [13]**. Durante el desarrollo del ensayo se ha estudiado además el comportamiento de las sustancias al ser excitadas por dos láseres, estando los resultados obtenidos en los siguientes capítulos.

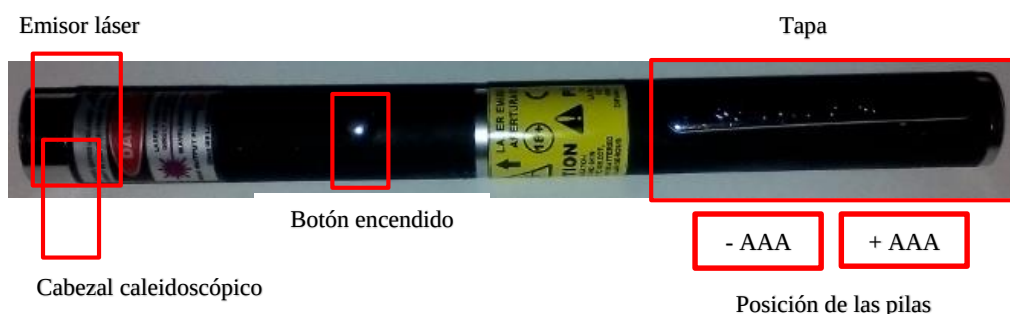


Ilustración 16: Estructura puntero láser

Color	Violeta
Potencia	100mW
Longitud de onda	405nm
Clase Producto Láser	3B
Divergencia del haz	2mRad
Diámetro del haz	6.0mm
O.D. Gafas de seguridad	3.0+
Alimentación	2xAAA 1.5v
Ciclo de trabajo	Continuo: 30 segundos encendido / 20 segundos apagado
Medidas	162x10mm
Peso	40g
Material del cuerpo	Aluminio y goma

Tabla 2: Especificaciones técnicas puntero láser

3.2 Diseño y estudio del espectrofotómetro.

El proceso de diseño del espectrómetro (**Ilustración 17**) comienza buscando información y consejos de las personas que hayan fabricado estos dispositivos con los que poder identificar una sustancia por un espectrograma.

Después de explorar en diferentes medios hemos seleccionado el portal Web *Spectral Workbench* [2] como guía, puesto que existe una sección en la que se aportan ideas y recomendaciones para que cualquier persona pueda crear uno de estos dispositivos. En este TFM hemos realizado un espectrómetro con una impresora 3D.

Indagando en este portal encontramos que uno de los usuarios compartía un modelo 3D del espectrómetro que había construido, por lo que cogimos dicho fichero y realizado las modificaciones que considerásemos necesarias. Dichos cambios los hemos efectuado utilizando el programa **3ds Max** [17], ya que el formato del archivo utilizado era 3ds.

Una vez se hayan efectuado los cambios en la pieza se debe convertir al formato STL para poder así utilizar el programa **Netfabb** [18], que optimiza y repara las capas con dicho formato. El siguiente paso consiste en convertir el archivo reparado a GCODE mediante los programas **Slicer** [19] o **Cura** [20]. Se logra así laminar la pieza y poder imprimirla. También en este momento se introducen las características de la impresora 3D que se va a utilizar, así como del material con el que se va a elaborar el dispositivo.

Cuando se haya introducido esta información ya se puede imprimir el espectrómetro. Después de que se finalice el proceso de impresión, es necesario hacer varias pruebas y comprobar si se debe hacer alguna corrección. Siguiendo este procedimiento se han realizado tres iteraciones, cuyas características se mostrarán a continuación.

Tras el proceso de impresión comenzamos a realizar las primeras pruebas con disoluciones de diferentes concentraciones en aceite y agua contenidas en pequeñas probetas de cristal.

La respuesta recogida durante estos experimentos no ha sido la esperada, puesto que las imágenes son demasiado tenues y no es posible sacar información útil de ellas. Por este motivo se va a realizar una segunda pieza con diferente material y de mayor tamaño, cuyas características se presentan en el siguiente apartado.

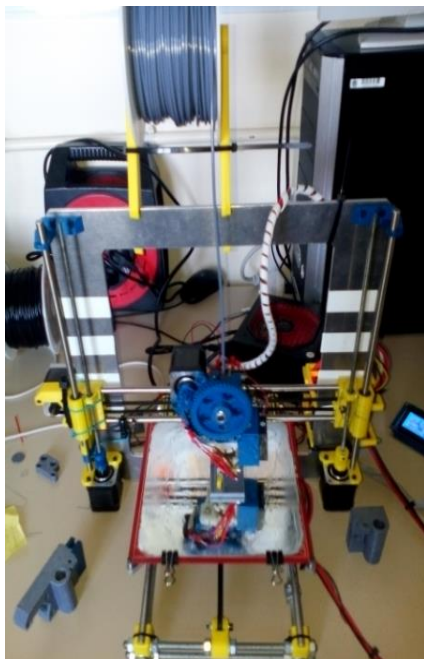


Ilustración 18: Proceso de impresión (1)

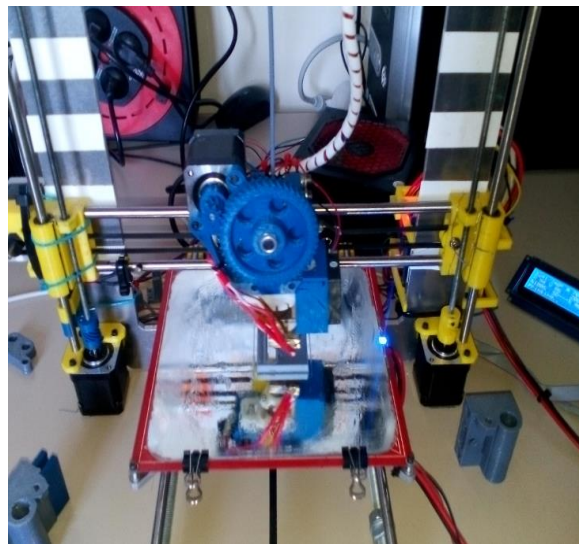


Ilustración 19: Proceso de impresión (2)

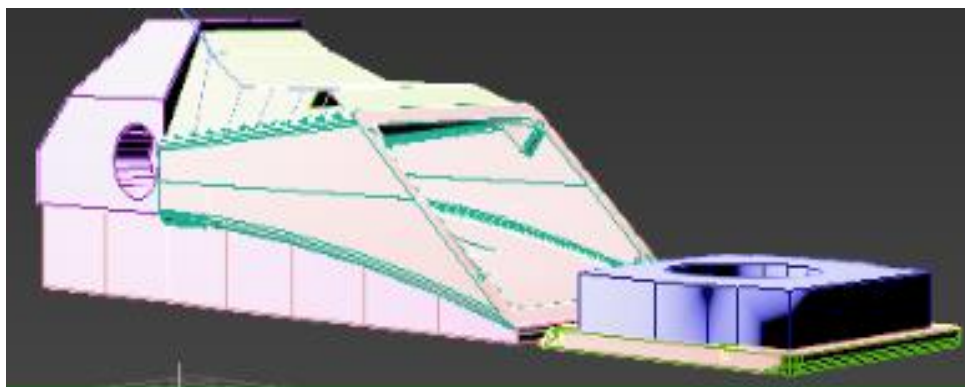


Ilustración 20: Diagrama 3d espectrofotómetro versión 1 y 2

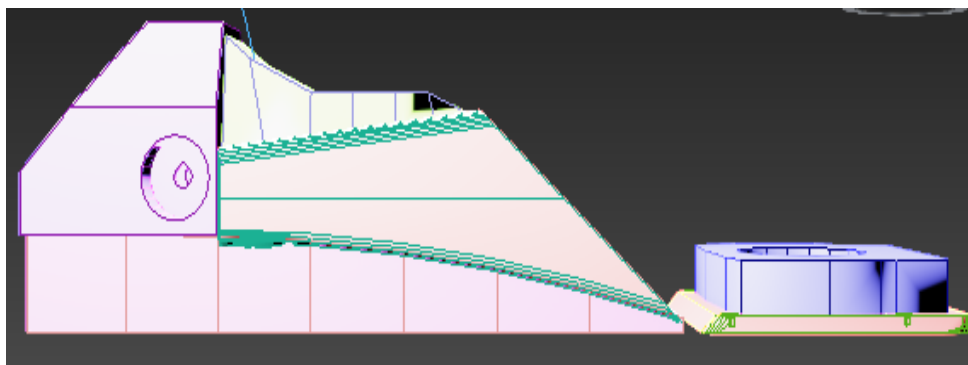


Ilustración 21: Alzado 3d espectrofotómetro versión 1 y 2

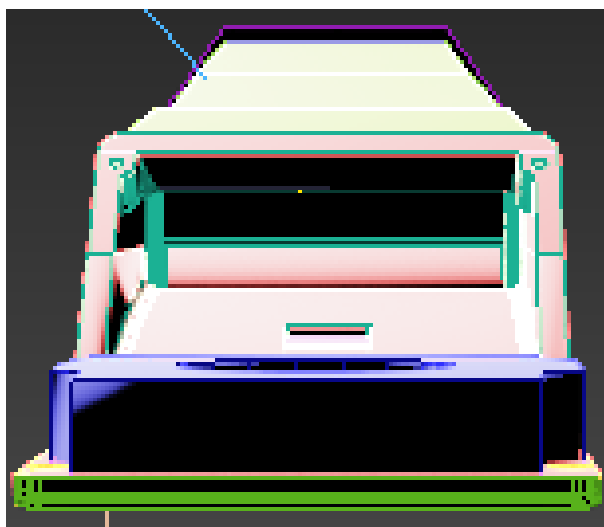


Ilustración 22: Perfil 3d espectrofotómetro versión 1 y 2

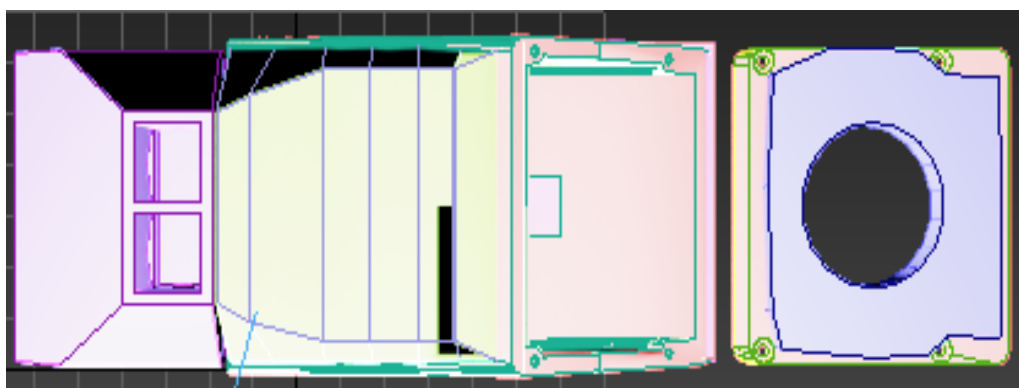


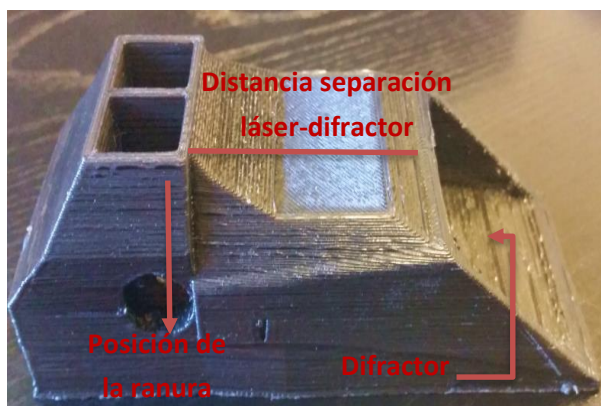
Ilustración 23: Planta 3d espectrofotómetro versión 1 y 2

*Ilustración 24: Espectrómetro versión 1(1)**Ilustración 25: Espectrómetro versión 1(2)*

2ª Iteración.

Ante los problemas detectados con el primer tipo de plástico, en la segunda iteración (ver **Ilustración 26** y **27**) se trabajó con plástico de tipo PLA (color negro). Aprovechamos este nuevo diseño para aumentar el tamaño un 12% para dar mejor cabida a los recipientes donde se pondrán las sustancias a analizar.

Una vez implementada esta nueva versión hemos comprobado que la calidad de las imágenes obtenidas era mejorable pues seguían resultando muy tenues e indefinidas. El problema estaba en que es necesario un ajuste muy fino de la distancia de la cámara al difractor y del ángulo de inclinación. Se impuso, pues, un nuevo diseño que tuviera en cuenta estos aspectos del diseño.

*Ilustración 26: Espectrómetro versión 2(1)**Ilustración 27: Espectrómetro versión 2(2)*

3ª Iteración.

La tercera iteración creada se divide en varias piezas, logrando así aumentar o disminuir la distancia entre el láser y la cámara; además de modificar el ángulo en el que se inclina la misma.

Estas cuatro partes se ven de forma clara en el diagrama 3-D de la **Ilustración 28**, correspondiéndose la **Ilustración 29** a la corrección de los errores en las capas de la primera pieza mediante el programa **Netfabb**. Dicho proceso se ha llevado a cabo en todos los componentes del espectrómetro para evitar posibles fallos que no se han podido apreciar al utilizar el programa de diseño.

Después de efectuar las mismas pruebas a las que se ha sometido a los dispositivos anteriores, se comprueba que los resultados obtenidos son mejores. Por ello se ha seleccionado esta tercera pieza (ver **Ilustración 30 a 32**) para efectuar los análisis de contaminantes en el agua.

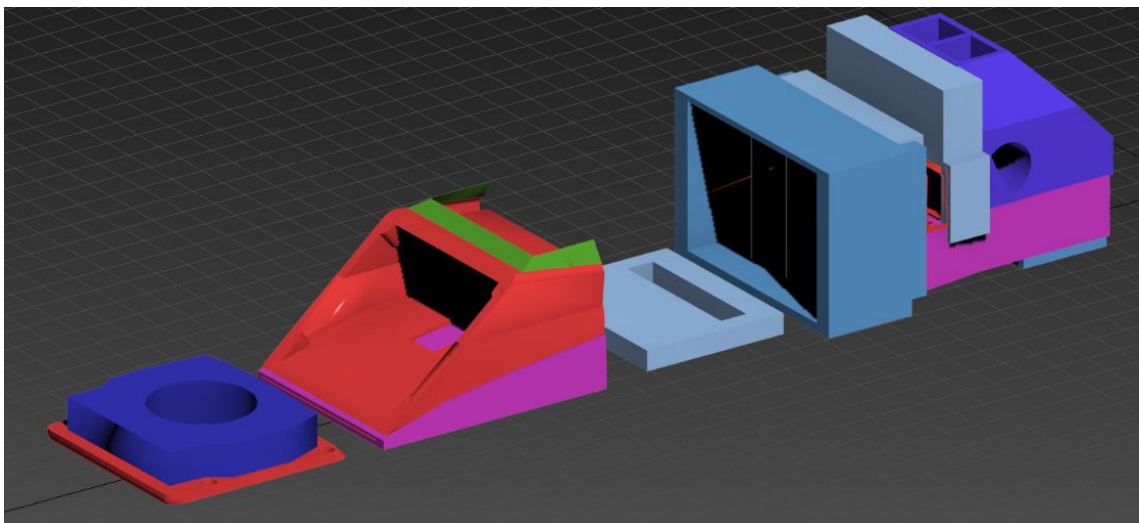


Ilustración 28: Diagrama 3d espectrofotómetro versión 3

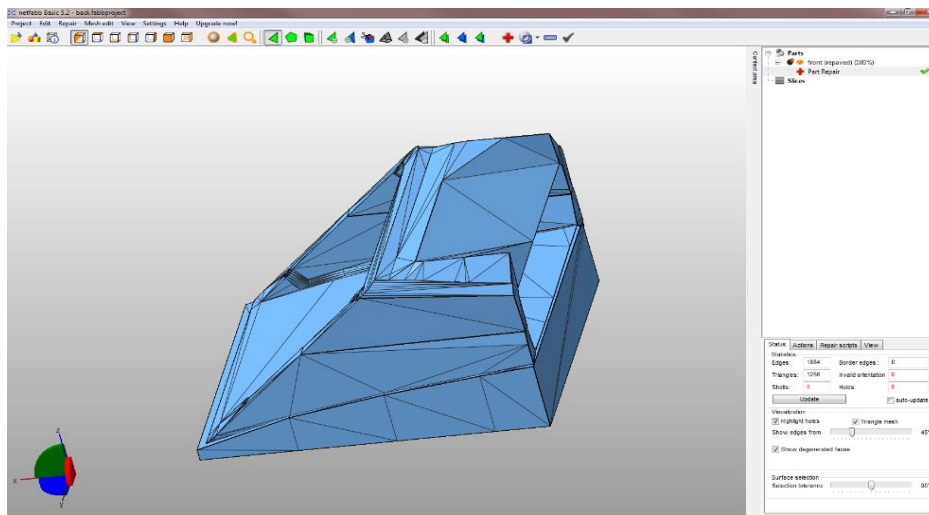


Ilustración 29: Corrección pieza espectrómetro versión 3

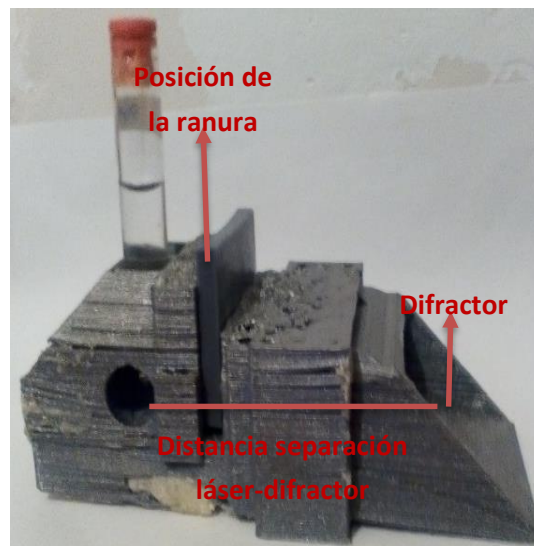


Ilustración 30: Espectrómetro versión 3(1)



Ilustración 31: Espectrómetro versión 3(2)
1 láser



Ilustración 32: Espectrómetro versión 3(3)
2 láseres

3.3 Características de la cámara.

Como se ha indicado en los capítulos anteriores, después de hacer pasar el haz de luz por la disolución a analizar es necesario capturar la respuesta de la sustancia fluorescente a dicho paso.

En esta ocasión se ha estudiado el comportamiento de estos elementos usando la cámara del teléfono móvil o de una web-cam llegando a la conclusión de que es mejor utilizar la segunda opción. Esta decisión se debe a que se han obtenido imágenes de una mayor calidad y de las que podemos obtener información útil en el momento de procesarlas. Las características de la cámara que hemos elegido para nuestro prototipo (ver **Ilustraciones 33 y 34**) se presentan en la **Tabla 3 [21]**.



Ilustración 33: Cámara web (1)



Ilustración 34: Cámara web (2)

Información de la versión	
Nombre producto	Microsoft LifeCam Studio™
Versión producto	Microsoft LifeCam Studio
Versión Webcam	Microsoft LifeCam Studio
Dimensiones del producto	
Longitud	114mm
Ancho	60.0mm
Altura	45.0mm
Peso	128g
Longitud cable	1829mm+152/-0mm
Características de la captura	
Sensor	Tecnología de sensor CMOS
Resolución	Resolución del sensor: 1920x1080 Still Image: 5 megapixel (2560x2048 pixel, interpolado)
Campo de visión	75 ° Campo de visión diagonal

Tabla 3: Especificaciones técnicas cámara

Capítulo 4: Recogida de muestras y generación de la base de datos.

Una vez fabricado el espectrómetro ya estamos en disposición de comenzar con los análisis de las diferentes disoluciones; siendo importante destacar que todos los experimentos deben efectuarse en las mismas condiciones, y que las muestras se recogen en cámara cerrada de tal forma que la única luz presente es la generada por el láser.

De esta manera, nos aseguraremos de que los resultados obtenidos son correctos al realizar los mismos pasos en todos los experimentos. Por ende, éstos pueden ser reproducidos de forma fidedigna. Por ello, es primordial definir los pasos que se van a seguir para recoger las muestras, así como para elaborar la base de datos.

4.1 Recogida de las muestras en escenarios reales.

En este apartado describimos cómo se recogen muestras en situaciones reales, de tal forma que sabemos cómo recoger las muestras en la fase de diseño del prototipo y que este sea realista. Esto es, generar una base de datos bien ajustada a la realidad.

El primer paso consiste en identificar las botellitas que contienen la muestra elaborando un pequeño registro con la fecha, hora y lugar de recogida de éstas; siendo aconsejable tomar una foto tal y como se indica en la **Ilustración 35**.

En el momento de preparar la muestra usaremos un frotis de algodón o un pequeño pincel, restregándolo en la zona contaminada para después disolverla en los recipientes que contendrán las muestras. En todo este proceso la persona que esté en contacto estas soluciones deberá utilizar guantes para evitar que la muestra se contamine y que el analista se infecte de algún posible contaminante.

De igual modo, es recomendable disponer de muestras adicionales en frascos de vidrio y almacenarlos en una zona fresca y oscura para poder realizar las pruebas en otros medios y así poder comparar los resultados si se cree conveniente [22]



Ilustración 35: Identificación muestra

Cuando tengamos preparada la muestra podremos medir la fluorescencia de ella por el paso del haz de luz que la traspasa, según se describe en el apartado 2.1. Para poder obtener dichas medidas es necesario realizar el montaje que se muestra en la **Ilustración 31 y 32**, con uno o dos láseres.

4.2 Generación de la Base de Datos.

Una vez que hemos descrito cómo es la recogida de muestras en escenarios reales, procedemos a explicar cómo se ha generado la base de datos usada para este Trabajo.

Los lubricantes que se han estudiado en este TFM presentan la siguiente denominación, siendo establecida por la SAE (*Society of Automotive Engineers/Sociedad de Ingenieros del Automóvil*). Dicho Organismo ha implantado un sistema de códigos numéricos para categorizar los aceites de motor en función de la viscosidad cinemática, midiéndose esta según el tiempo que tarda una cantidad determinada de aceite en fluir a través de un orificio a una temperatura estándar. Esta afirmación implica que a mayor tiempo aumenta la viscosidad, y por consiguiente mayor es el código SAE [23,24].

Para este TFM usaremos los siguientes:

- 10.40.
- 15.40.
- 20.50.
- 75.80.
- ATF (*Automatic Transmission Fluids* o Fluidos de Transmisión Automática).

Con el objetivo de conocer la capacidad de cuantificación de nuestro prototipo, se generan disoluciones con concentraciones de aceite variables; es necesario realizar varias mezclas que contengan grasa y agua en diferentes proporciones. De esta forma se estudia el comportamiento del haz de luz al atravesar las disoluciones de diferente concentración de aceite; siendo éstas las siguientes:

- 5ml agua + 1ml aceite.
- 4ml agua + 2ml aceite.
- 3ml agua + 3ml aceite.
- 2ml agua + 4ml aceite.

Una vez hayamos dispuesto las concentraciones de las disoluciones, montamos la estructura de las **Ilustraciones 31 y 32**; sacando varias imágenes de cada solución. Para representar la posible variabilidad en cuanto al efecto de la sesión de toma de datos, hemos realizado dos sesiones y en cada una se han sacado tres imágenes. De esta forma hay seis imágenes por concentración.

Del mismo modo, con objeto de determinar la ausencia de aceites (u otro tipo de sustancia química) es necesario tomar imágenes con la respuesta del agua cuando pasa por ella un haz de luz. En este caso la cantidad de agua era de 6ml.

Asimismo, se deben tomar varias fotos con la cámara sin encender el láser, para así conocer el ruido de fondo en las imágenes, por ejemplo, el procedente de este dispositivo. También hay que destacar que se realizó este proceso utilizando tanto uno como dos láseres; para así estudiar si se encuentran diferencias entre las respuestas por utilizar el doble de potencia.

Una vez hayamos desarrollado los pasos indicados anteriormente obtenemos imágenes como la que se presentan en las **Ilustraciones 36 hasta 41**, realizando a cada una de ellas un procesado en el que se seleccionan las zonas de interés para evitar que haya imagen inútil para su posterior análisis mediante las características y el clasificador descritos en el capítulo 5.

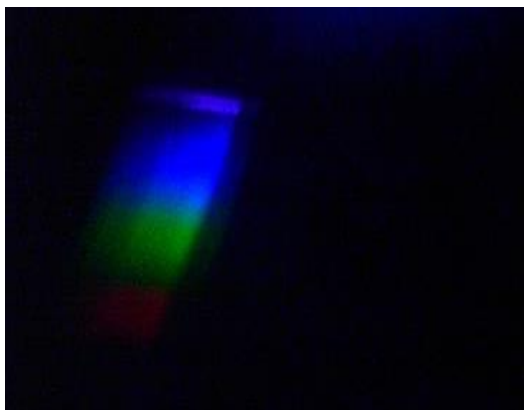


Ilustración 36: Imagen
2mlH2O+4mlAc.20,50 (1 láser)

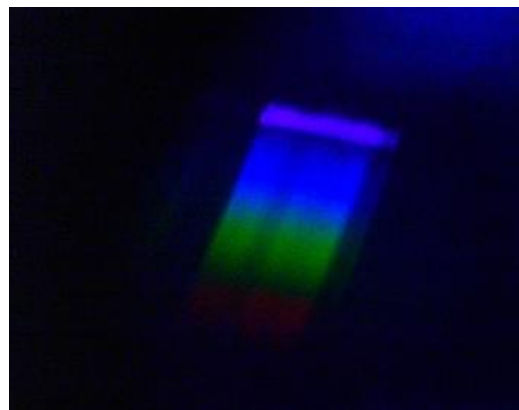


Ilustración 37: Imagen
2mlH2O+4mlAc.20,50(2 láseres)

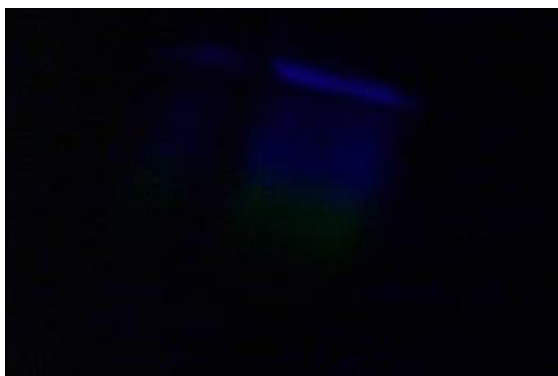


Ilustración 38: Imagen 4mlH2O+2mlAc.20,50
(1 láser)



Ilustración 39: Imagen 4mlH2O+2mlAc.20,50
(2 láseres)

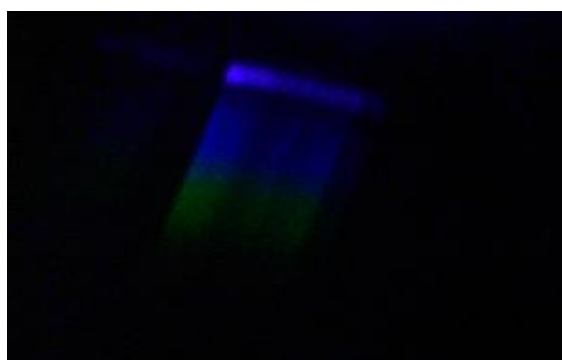


Ilustración 40: Imagen 5mlH2O+1mlAc.20,50
(1 láser)

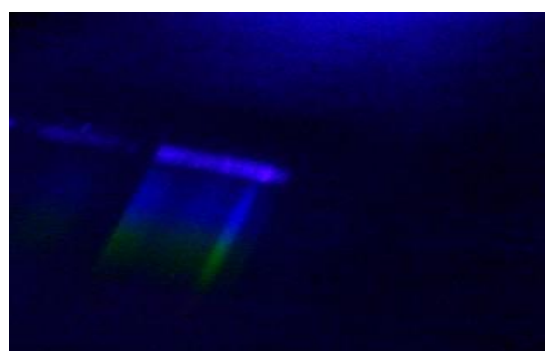


Ilustración 41: Imagen 5mlH2O+1mlAc.20,50
(2 láseres)

Capítulo 5: Extracción de características y clasificador

En este capítulo expondremos los aspectos más relevantes del prototipo que presentamos en términos de lo que es la extracción de características relevantes y el clasificador que nos permitirán determinar los contaminantes y su concentración. Una vez explicados estos elementos pasaremos en el siguiente capítulo a verificar el funcionamiento del prototipo desarrollado mediante una serie de experimentos.

5.1 Extracción de características

En lo referente a la extracción de características estudiamos cuáles son más relevantes para facilitar una correcta estimación del tipo de contaminante (por ejemplo, aceite en nuestros experimentos) o de su cuantificación en el agua. Entre las múltiples opciones de que disponemos en la literatura, hemos elegido las siguientes por ser de interés y contar con trabajos de referencia [25,26,27,28,29].

Así, las características que aplicamos a nuestro prototipo son de tres tipos diferentes que usaremos de forma individual o combinadas, y que se describen a continuación son:

1. **Características estadísticas:** mean, desviación estándar, varianza, *smoothness*, *skewness* y *kurtosis*.
2. **Características de textura:** contraste, correlación, energía, homogeneidad.
3. **Ángulo y coseno:** Medidos estos entre la imagen en escala de gris (Y) y las de colores (R, G, B), vistas las imágenes como una secuencia de valores ordenadas en vectores.
4. **Entropía de Permutación:** Medidas en las imágenes, vistas estas como una secuencia de valores ordenadas en vectores.
5. **Combinaciones de características.**

5.1.1 Características estadísticas

Los **parámetros estadísticos** utilizados en este TFM han sido los siguientes:

- **Media.** Es el parámetro estadístico más básico, utilizándose sobre todo en geometría y en análisis de imágenes, ya que crea cierto grado de borrosidad que logra reducir el ruido de ésta. Existen diferentes tipos, pero en este trabajo hemos

utilizado el filtro de media aritmética o promedio, trabajando en una ventana 'm × n' en la que calcula el promedio de todos los valores de píxeles dentro de ésta y reemplaza el valor del píxel central en la imagen de destino por el resultado. Se calcula mediante la siguiente fórmula:

$$f(x, y) = \frac{1}{mn} \sum_{(r,c) \in W} g(r, c)$$

donde g es la imagen con ruido, $f(x, y)$ es la imagen restaurada, r y c son las filas y columnas respectivamente, dentro de una ventana 'W' de tamaño 'm × n'.

- **Desviación estándar.** Este segundo criterio indica cuánta variación o dispersión existe del promedio. Si obtenemos una desviación estándar baja indica que los puntos de datos tienden a estar muy cerca de la media, mientras que si es alta indica que los puntos de datos se extienden a lo largo de una amplia gama de valores. De forma matemática la calculamos aplicando la siguiente ecuación:

$$f(x, y) = \sqrt{\frac{1}{mn - 1} \sum_{(r,c) \in W} (g(r, c) - \frac{1}{mn - 1} \sum_{(r,c) \in W} g(r, c))^2}$$

- **Varianza.** Este parámetro indica la dispersión de los valores alrededor de la media y se puede utilizar para determinar la posición del borde en el procesamiento de imágenes. Matemáticamente se calcula si aplicamos la fórmula siguiente:

$$f(x, y) = \frac{1}{mn - 1} \sum_{(r,c) \in W} (g(r, c) - \frac{1}{mn - 1} \sum_{(r,c) \in W} g(r, c))^2$$

- **Skewness.** Dicha característica es una medida de asimetría de la distribución de probabilidad de una variable aleatoria de valor real, pudiendo ser negativo, positivo o no definido. Si resultado es negativo indica que la mayor parte de los valores se encuentra a la derecha de la media, mientras que si es positivo se hayan éstos a la izquierda. Cuando el valor es nulo, los valores se distribuyen uniformemente a ambos lados de la media. Matemáticamente se calcula aplicando la siguiente fórmula,

$$f(x, y) = \frac{\frac{1}{mn-1} \sum_{(r,c) \in W} \left(\frac{1}{mn-1} \sum_{(r,c) \in W} (g(r, c) - \frac{1}{mn-1} \sum_{(r,c) \in W} (g(r, c))) \right)^3}{\left(\frac{1}{mn-1} \sum_{(r,c) \in W} \left(\frac{1}{mn-1} \sum_{(r,c) \in W} (g(r, c) - \frac{1}{mn-1} \sum_{(r,c) \in W} (g(r, c)))^2 \right) \right)^{3/2}}$$

- **Kurtosis.** Los valores de procesamiento digital de esta característica se interpretan en combinación con el ruido y la medición de resolución. Valores altos de *Kurtosis* deben ir de la mano de bajo ruido y baja resolución. Se calcula aplicando la siguiente fórmula

$$f(x, y) = \frac{\frac{1}{mn-1} \sum_{(r,c) \in W} \left(\frac{1}{mn-1} \sum_{(r,c) \in W} (g(r, c) - \frac{1}{mn-1} \sum_{(r,c) \in W} (g(r, c))) \right)^4}{\left(\frac{1}{mn-1} \sum_{(r,c) \in W} \left(\frac{1}{mn-1} \sum_{(r,c) \in W} (g(r, c) - \frac{1}{mn-1} \sum_{(r,c) \in W} (g(r, c)))^2 \right) \right)^2}$$

- **Smoothness.** El *smoothness* de la imagen se explica por característica *Inverse Difference Moment/ Momento de Diferencia Inversa (IDM)*, esperando que el IDM sea alto si los niveles de gris del píxel son similares. Esta medida se relaciona inversamente con la medida de contraste. Así, una expresión matemática de cálculo es: $1 - (1/(1+a))$, donde 'a' es la suma de todos los valores de la imagen bajo análisis.

Los parámetros estadísticos se calculan mediante las funciones correspondientes de Matlab: mean2, std2, var, skewness y kurtosis.

5.1.2 Características de textura

Los **parámetros de textura** son los que se describen a continuación:

- **Contraste.** Otro criterio utilizado es el contraste, que diferencia en las propiedades visuales que hace que un objeto se distinga de otros objetos y el fondo, estando determinado por la diferencia en el color y el brillo del objeto y otros objetos dentro del mismo campo de visión. El contraste es 0 para una imagen constante.
- **Correlación.** Este parámetro establece la dependencia lineal del nivel de gris entre los píxeles en las posiciones especificadas entre sí; así como la energía que mide la uniformidad de la distribución del nivel de intensidad. La correlación tiene un valor de 1 o -1 para una imagen perfectamente positiva o negativamente correlacionada.
- **Energía.** La característica energética mide la uniformidad de la distribución del nivel de intensidad. Si el valor es alto, entonces la distribución corresponde a pequeño número de niveles de intensidad; por lo que la energía de una imagen es alta cuando

la imagen es homogénea. El rango es [0 1] correspondiendo este último valor para una imagen constante.

- **Homogeneidad.** La proximidad de la distribución de elementos en el *GLCM* a la *GLCM* diagonal y rango = [0 1]. La homogeneidad es 1 para un *GLCM* diagonal.

Hemos de precisar que para obtener para extraer las características de textura usamos la función *Graycoprops (GLCM, PROPERTIES)* de Matlab, que normaliza la matriz de co-ocurrencia de nivel de gris (*GLCM*) de manera que la suma de sus elementos sea una. Cada elemento en el *GLCM* normalizado (r, c), es la probabilidad conjunta de ocurrencia de pares de píxeles con una relación espacial definida que tiene valores de nivel de gris r y c en la imagen. *Graycoprops* utiliza el *GLCM* normalizado para calcular PROPIEDADES. *GLCM* puede ser una matriz $m \times n \times p$ de matrices de co-ocurrencia de nivel de gris válidas. Cada matriz de co-ocurrencia de nivel de gris se normaliza de modo que su suma sea una.

5.1.3 Entropía de Permutación

La detección de los cambios dinámicos cualitativos y cuantitativos en los sistemas complejos, tanto naturales como artificiales, es uno de los problemas más importantes de la física, ingeniería, ciencias económicas, y medicina, donde la detección precisa de las transiciones de un estado normal a un estado anormal puede mejorar el diagnóstico y el tratamiento de un problema dado [31]. Previamente ha sido probado con buenos resultados en un TFG [30] aplicado a la detección de apnea del sueño. A continuación, exponemos un extracto de los fundamentos de este tipo de características.

Bandt y Pompe en el 2002 presentaron el concepto de entropía de permutación (PE), como un método análisis que mide la complejidad de series de tiempo. La PE es conceptualmente simple y computacionalmente muy rápido. Estas dos características motivan a explorar si este concepto puede ser utilizado eficazmente para detectar cambios dinámicos en series temporales complejas.

En el contexto de la teoría de la probabilidad, la entropía cuantifica la complejidad de una serie temporal dada, tal que $X = \{x_i\}_{i=1 \dots N}$, donde N es la longitud de la serie temporal. La PE se define a partir de la entropía de Shannon tal como se expresa en la siguiente fórmula

$$H_{Shannon} = - \sum_{i=1}^n p_i * \log_2 p_i$$

donde n es el orden de permutación y p_i representa la probabilidad de ocurrencia de los posibles patrones de la secuencia en la serie temporal. La entropía de permutación por cada uno de los patrones viene definida por la expresión siguiente. Usando el término $(n-1)$ se normaliza el valor con el fin de comparar varias definiciones de PE, de tal forma que $0 \leq h_{Shannon} \leq 1$.

$$h_{Shannon} = -\frac{1}{(n-1)} * \sum_{i=1}^{n!} p_i * \log_2 p_i$$

Alrededor de esta definición, basada en la entropía de Shannon, han surgido otras que se pueden considerar generalizaciones de ésta. Rényi presenta una familia de entropías que convergen en una entropía diferencial cuando $\alpha \rightarrow 1$, definida en

$$H_{Rényi} = \frac{1}{1-\alpha} * \log(\sum_{i=1}^N p_i^\alpha) \quad \text{Si } \alpha \neq 1$$

$$H_{Shannon} \quad \text{Si } \alpha = 1$$

A su vez, Tsallis propuso otra familia de entropías que nos puede ser útil y se presenta en

$$H_{Tsallis} = \frac{1}{1-\alpha} * (\sum_{i=1}^N p_i^\alpha - 1) \quad \text{Si } \alpha \neq 1$$

$$H_{Shannon} \quad \text{Si } \alpha = 1$$

Cabe mencionar que la entropía exponencial y la exponencial generalizada fueron presentadas por Campbell, Koski y Persson, para ser aplicada en compresión de datos, Tsallis y la entropía exponencial convergen en una entropía diferencial cuando $\alpha \rightarrow 1$.

Comparativa cualitativa de las definiciones

La entropía de Shannon se puede usar para evaluar la entropía de las distribuciones de densidad de probabilidad alrededor de algunos puntos. Esta noción de entropía se puede generalizar para proporcionar información adicional acerca de la importancia de eventos específicos, como por ejemplo los valores atípicos o eventos raros.

Medidas de entropía definidas con este factor $\sum_{i=1}^N p_i^\alpha$, dependen del valor de la probabilidad y el valor de α , proporciona control sobre dicho valor de la probabilidad:

- Si $\alpha \gg 0$ la medida es más sensible a los acontecimientos que se producen con frecuencia.
- Si $\alpha \ll 0$, la medida es más sensible a los acontecimientos que ocurren rara vez.

Partiendo de la definición de Shannon, Constantino Tsallis y Alfred Rényi proponen nuevas definiciones de entropía. Estas propuestas tienen propiedades similares a la entropía de Shannon, cabe destacar:

- Son aditivas.
- Tienen un máximo = $\ln(n)$ para probabilidades $p_i = 1/n$.

Pero contienen α que representa el grado de sensibilidad a la forma de las distribuciones de probabilidad.

Cálculo numérico detallado

Como se ha venido indicando, la PE mide la información basada en que aparezcan o no ciertos patrones de permutación, [33] en el rango de valores de la una serie temporal. Para computar o calcular la PE se debe tener una serie temporal $\{x_i\}$ de longitud N y aplicarle los siguientes pasos:

1. Definir el orden de permutación n , esto nos lleva a $\pi_j = 1 \dots n!$. patrones de permutación, n queda definido como $n=1\dots n$.
2. Se inicializa $i=1$, siendo este el índice de la serie temporal que se considera $\{x_i\}$ $i=1\dots N$ y el contador $z_j = 0$, para cada patrón de permutación π_j .
3. Se calcula los valores en cada posición, de la secuencia $x_i, \dots, x_{i+n-1}$, los cuales se van almacenando en $r_{i,\dots,r_{i+n-1}}$, esta nueva secuencia es el valor de cada índice en orden ascendente.

4. Se compara lo que se ha obtenido en el paso 3 con todos los patrones de permutación, y se va aumentando el contador ($z_k = z_k + 1$) si hay patrones iguales $\pi_k = r_{i,...,i+n-1}$.
5. Si $i \leq N-n$, entonces $i=i+1$ y se comienza en el paso 3 de nuevo, si $i \geq N-n$ se va al paso 6.
6. Se calcula la frecuencia relativa de todas las permutaciones π_j , por medio de $P'_j = \frac{z_j}{\sum z_k}$, de tal manera que se estima su probabilidad P_j .
7. Se selecciona todos los valores $P'_j > 0$, y se calcula su PE, sin contar símbolos con valor "0".(non zero).

5.2 El clasificador utilizado

El clasificador utilizado en este TFM es un SVM con kernel polinomial de orden 3. Dado que los SVM son por naturaleza clasificadores biclase, hemos aplicado una estrategia de clasificación multiclase a partir de los biclase. A continuación, expondremos los fundamentos del SVM y, posteriormente, presentaremos varias opciones para el clasificador multiclase indicando cuál hemos tomado para nuestro prototipo.

5.2.1 Clasificación binaria mediante SVM

La teoría de Máquinas de Vectores de Soporte (*Support Vector Machine*, SVM, en su acepción inglesa) fue desarrollada por Vapnik apoyándose en el principio de Minimización del Riesgo Estructural (SRM); ya que en muestra un mejor desempeño que al utilizar el principio de Minimización del Riesgo Empírico (ERM). Esto se debe a que la SVM minimiza un límite superior al riesgo esperado, mientras que el ERM minimiza el error sobre los datos de entrenamiento [32].

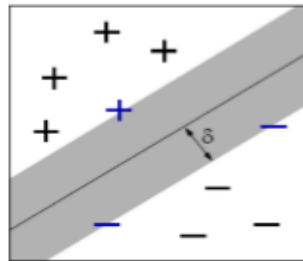
Se trata de un clasificador lineal que mapea los puntos de entrada a un espacio de características de una dimensión mayor (introducidos por funciones núcleo o kernel), para luego encontrar el hiperplano que los separe y así maximice el margen entre las clases. Se pueden utilizar para resolver tanto problemas lineales como no lineales.

Los SVM han sido desarrollados como una técnica robusta para clasificación y regresión aplicado a grandes conjuntos de datos complejos con ruido; es decir, con

variables inherentes al modelo que para otras técnicas aumentan la posibilidad de error en los resultados pues resultan difíciles de cuantificar y observar.

Separador lineal

Para este caso el SVM crea hiperplanos que separan los datos de entrada en dos subgrupos que poseen etiqueta propia (ver **Ilustración 42**). En medio de todos los posibles planos de separación de las dos clases etiquetadas como $\{-1, +1\}$, existe sólo un hiperplano de separación óptimo, por lo que la distancia entre el hiperplano óptimo y el valor de entrada más cercano sea máxima (maximización del margen) para que se generalice la máquina que se está creando los puntos en los que se apoya el margen máximo se denominan vectores soporte.



***Ilustración 42:** SVM Linealmente separable*

Separador no lineal

Existen dos posibles soluciones para resolver este problema:

- El primero es usar una función kernel cuando los datos se pueden separar con margen máximo en un espacio de características, siendo este de mayor dimensión y que se obtiene mediante una transformación a las variables de entrada.
- El segundo caso es utilizar el *Soft Margin* o Margen Blando, que se emplea cuando no se puede transformar los datos para conseguir una separación lineal, ya sea en el espacio de entrada o de características.

SVM con margen máximo en el espacio de características.

Como se ha mencionado, la transformación de los datos de un espacio inicial a otro de mayor dimensión se logra mediante el uso de la función kernel (por ejemplo, ver la **Ilustración 43**). La función núcleo o kernel es un producto interno en el espacio de características que tiene su equivalente en el espacio de entrada (Gunn, 1997)

$$k(x, x') = \langle \Phi(x), \Phi(x') \rangle$$

donde k , es una función simétrica positiva definida que cumple las condiciones de Mercer.

De manera gráfica se puede observar en la figura como la función kernel permite realizar la separación y el traslado de los datos al espacio de características.

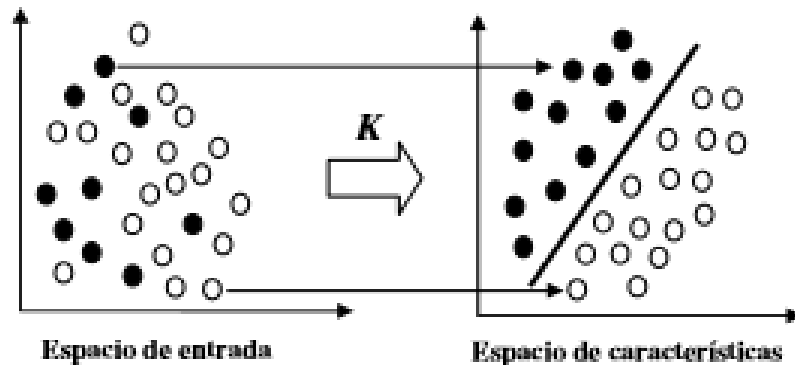


Ilustración 43: SVM no linealmente separable inducida por una función kernel

Entre los kernels más comunes, se encuentran: la función lineal, polinomial, RBF (Radial Basis Function), ERBF (Exponential Radial Basis Function), entre otros.

SVM con Margen Blando

Este tipo particular de los SVM trata aquellos casos donde existe datos de entrada erróneos, ruido o alto solapamiento de las clases en los datos de entrenamiento, donde se puede ver afectado el hiperplano clasificador, por esta razón se cambia un poco la perspectiva y se busca el mejor hiperplano clasificador que pueda tolerar el ruido en los datos de entrenamiento. En la **Ilustración 44** podemos ver un ejemplo de caso de este tipo.

Según Benavidez, et al., (2006) esto se logra relajando las restricciones presentadas en el caso lineal, introduciendo variables de holgura no-negativas.

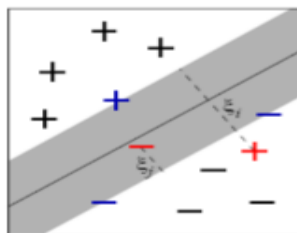


Ilustración 44: SVM con margen blando

De manera que, matemáticamente el problema queda definido como:

$$\text{Minimizar} \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i$$

$$\text{sujeto a} \quad y_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i \quad \xi_i \geq 0, \quad i = 1, \dots, l$$

Se incluye un término de regularización que depende de las variables de holgura, que a su vez depende de la magnitud de las mismas y del margen. Este término incluye también una constante C , que determina la holgura del margen blando.

El valor de este parámetro C , que debe ser estimado a priori, depende del evaluador. La elección de un valor para este parámetro y el tipo de función kernel influyen en el desempeño de las MVS (Benavidez, et al ,2006).

Siguiendo el mismo procedimiento utilizado en el caso linealmente separable, la resolución de este problema viene dada por la búsqueda de los multiplicadores de Lagrange, para esto se construye un Lagrangiano y se resuelve en el problema dual:

$$\text{Maximizar} \quad \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N y_i y_j \alpha_i \alpha_j K(x_i, x_j)$$

$$\text{sujeto a} \quad \sum_{i=1}^N y_i \alpha_i = 0$$

$$0 \leq \alpha_i \leq C \quad 1 \leq i \leq N$$

La función a maximizar es la misma que para el caso de margen máximo, a diferencia de la restricción $0 \leq \alpha_i \leq C$. Esto implica que los datos o patrones que cumplen la condición de tener valores $\alpha_i \geq C$ tienen el mismo comportamiento en la SVM con margen máximo. Es decir, que las SVM con margen máximo se pueden obtener con $C = \infty$. Esto significa que cuanto más grande es el valor de C , más estricta es la SVM al momento de permitir errores, penalizándolos con mayor rigurosidad (Betancourt, 2005).

Las ventajas de SVM son que pertenece a disciplinas de aprendizaje automático o estadístico; consiguiendo así que las máquinas vayan aprendiendo las salidas correctas para ciertas entradas mediante ejemplos.

Asimismo, este clasificador aplica un nuevo principio inductivo, que busca minimizar el riesgo estructural, además del uso de una función núcleo o kernel, atribuyendo una gran capacidad de generalización.

Por otro lado, uno de los problemas que presenta este clasificador es el *Overtraining* o sobre entrenamiento, que da a una mala generalización del modelo. Este inconveniente consiste en aprender los datos de entrenamiento, pero es capaz de clasificar bien ejemplos nuevos (datos de verificación). Otro problema es realizar una mala clasificación al no aprender de forma correcta los datos de entrenamiento.

5.2.2 Clasificación multiclase

Tal como hemos avanzado, el SVM es un clasificador binario en el que las etiquetas de clase únicamente pueden tomar dos valores: ± 1 . Si bien hay muchos ejemplos de la necesidad de contar con clasificadores biclase, lo cierto es que no son pocas las aplicaciones prácticas reales donde se trabaja con más de dos clases. Un ejemplo claro de esta situación es en reconocimiento óptico de características [33] o el prototipo que estamos presentando en este TFM. En estos casos estamos ante la necesidad de una clasificación multiclase, y a continuación veremos las principales estrategias para llegar a un clasificador multiclase a partir de uno biclase. Particularmente, hemos estudiado las estrategias de “uno contra todos” y “por pares” siendo esta última la elegida para nuestro prototipo.

- **Uno contra todos:** Para obtener clasificadores de M -clases, se construyen conjuntos de clasificadores binarios $f^1, f^2, \dots, f^M$, cada uno entrenado para separar una **clase del resto**.

Una vez contruidos estos clasificadores ‘uno contra todos’ se combinan para obtener una **clasificación multiclase** acorde a la máxima salida después de aplicar la función siguiente:

$$\operatorname{argmax} g^j(x), \text{ donde } g^j(x) = \sum_{i=1}^m y_i \alpha_i^j k(x, x_i) + b^j$$

$g^j(x)$ devuelve una señal de valor real que se interpreta como la distancia desde hiperplano de separación hasta el punto x .

El valor también puede interpretarse como un **valor de confianza**. Cuanto mayor sea el valor, más fiable es que el punto x pertenezca a la clase positiva. Por tanto, hay que asignar al punto x la clase cuya fiabilidad sea mayor para dicho punto.

- **Clasificación SVM por pares:** Se entrena a un clasificador para cada posible par de clases. Así, para M clases, da lugar a $\binom{M}{2} = \frac{(M-1)M}{2}$ clasificadores binarios.

Clasifica un punto desconocido X aplicando cada uno de los clasificadores binarios $\binom{M}{2}$ y cuenta cuántas veces al punto X se le asigna la etiqueta de clase. De esta forma, la clase con el valor más alto es considerada la etiqueta para el punto desconocido X .

Capítulo 6: Resultados y conclusiones.

En este capítulo se indican los resultados obtenidos tras desarrollar los montajes previamente descritos en el capítulo 3 y 4, y hacer los programas que permiten determinar los contaminantes a partir del clasificador alimentado con las características extraídas (capítulo 5). Además, exponemos las conclusiones a las que llegamos una vez analizamos los resultados obtenidos de nuestros experimentos.

6.1 Resultados

Para verificar el funcionamiento del prototipo desarrollado en este TFM hemos llevado a cabo una serie de experimentos en los que hemos probado múltiples opciones con la finalidad de conocer cuáles son las mejores condiciones de trabajo de cara a un producto final.

El contaminante usado es aceite de varios tipos. Dado que el número de ejemplos que tenemos de cada aceite y, particularmente, de cada cuantificación de la concentración es limitado, cada experimento de clasificación se repite cincuenta (50) veces, en cada una se toma el 75% de los ejemplos para entrenar y el resto para clasificar.

En estos experimentos nos hemos preguntado sobre los siguientes aspectos relativos a la información básica o primaria que tomamos:

- Qué efecto tiene utilizar uno o dos láseres como fuente iluminación.
- Qué colores son más relevantes como fuente primaria de información y, si procede utilizar una combinación de ellos. Así, distinguimos Red (R), Green (G) u Blue (B) o escala de grises (Y) mediante una combinación de los colores básicos con igual coeficiente de participación, esto es, 1/3.
- Qué efecto tiene el ruido de imagen de fondo, esta es, la imagen que se obtiene sin iluminar con láser.

Vinculados al objetivo final del prototipo, nos preguntamos sobre la capacidad de este de proporcionar buenos resultados en estos contextos:

- Desconocido el aceite, y dentro de la gama que tenemos caracterizada, estimar cuál es y su concentración. Este experimento se realiza en una competición de 'todos contra todos'.
- Conocido el aceite, estimar la concentración del mismo en la disolución.

Digamos, para facilitar una buena comprensión de los datos, que en las tablas que presentamos en los siguientes apartados, se muestran para tipo de aceite (Ver capítulo 4, apartado 2), los resultados medios (η) de acierto y su desviación típica (σ) para cada experimento de clasificación. Distinguimos si se utiliza una extracción del ruido de fondo de imagen ($Si_fondo=1$) o no ($Si_fondo=0$), y en función del color elegido (R, G, B, Y) para extraer características. Destacamos mejor resultado en función del color.

6.1.1 Cuantificación de la concentración de contaminante en función del aceite.

6.1.1.1 Iluminación con 1 láser.

En la Tabla 4 se presentan los datos conseguidos al usar únicamente las Características de Textura entre las distintas imágenes, deduciendo así que:

- La peor tasa de acierto se presenta cuando no se elimina el fondo en el Ac.15.40 (71.500%); mientras que el mejor resultado es de 90.875% correspondiente al Ac. 75.80 sin quitar tampoco el ruido de fondo.
- El color con el que se consiguen los resultados anteriores es el verde, al igual que vemos que para cada aceite se utilizan los mismos colores, ya sea si restamos o no el fondo.
- Obtenemos mejores resultados al retirar el fondo en las imágenes de todas las disoluciones excepto en el Ac. 75.80.

En términos generales podemos considerar que las tasas de acierto altas; dando a entender que en las condiciones de estos experimentos tenemos buena capacidad de cuantificación de la disolución.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	85.375% $\pm$ 7.32%	2		86.500%$\pm$6.36%	2
15.40	71.500% $\pm$ 10.80%	2		72.125%$\pm$9.46%	2
20.50	72.625% $\pm$ 11.28%	3		72.625%$\pm$10.55%	3
75.80	90.875%$\pm$6.46%	2		87.750% $\pm$ 7.78%	2
ATF	81.625% $\pm$ 7.72%	3		83.750%$\pm$8.56%	3

Tabla 4: Resultados cuantificación en función de concentración del contaminante usando solo Características de Textura: a) con fondo, b) sin fondo (1 láser)

En la Tabla 5 se disponen los resultados tras utilizar los parámetros estadísticos explicados anteriormente. De dicho cuadro concluimos lo siguiente:

- Sin eliminar el fondo el mejor resultado que se consigue es de 99.125% mientras que el peor es de 89.750%. Del mismo modo cuando sustraemos dicho ruido, el mejor porcentaje de acierto es de 95.250% y el peor de 91.250%.
- El aceite que mejor respuesta presenta al utilizar estas características es el Ac. 75.80 usando el color verde tanto si quitamos como si no el fondo.
- Los resultados de aciertos obtenidos (η) tanto si retiramos como si no el fondo son muy similares.

Se aprecia así que las tasas de acierto son bastante altas al compararlas con las de la Tabla 4; por lo que las condiciones de estos experimentos son mejores que al emplear únicamente las características de textura.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	94.875% $\pm$ 5.16%	3		95.250%$\pm$4.47%	3
15.40	89.750% $\pm$ 7.55%	3		91.250%$\pm$8.28%	2
20.50	95.125%$\pm$5.96%	3		93.250% $\pm$ 6.29%	3
75.80	99.125%$\pm$2.53%	2		97.750% $\pm$ 5.89%	2
ATF	92.500% $\pm$ 6.56%	2		92.750%$\pm$7.07%	2

Tabla 5: Resultados cuantificación en función de concentración del contaminante usando solo parámetros estadísticos: a) con fondo, b) sin fondo (1 láser)

En la Tabla 6 se indican los datos al usar el Coseno entre las imágenes en escala de grises y de colores, sacando de aquí las siguientes conclusiones:

- El aceite que da una mejor respuesta es el Ac. 75.80 (95.750%) al eliminar el ruido de fondo; mientras que el peor lo obtenemos en el Ac. 10.40 (90.375%) sin quitar el fondo.
- Los resultados más significativos se obtienen sustrayendo el fondo de las imágenes tomadas.
- El color predominante tanto si eliminamos como si no el ruido es el verde (2), mientras que la escala de grises es el que menos información da.

En términos generales podemos considerar que las tasas de acierto son altas, dando a entender que en las condiciones de estos experimentos tenemos buena capacidad de cuantificación de la disolución.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	90.375% $\pm$ 6.21%	2		91.625%$\pm$5.73%	Y
15.40	92.625% $\pm$ 7.33%	2		93.875%$\pm$5.12%	2
20.50	91.875%$\pm$6.60%	1		91.625% $\pm$ 9.07%	1
75.80	92.875% $\pm$ 5.78%	2		95.750%$\pm$3.44%	3
ATF	95.500%$\pm$6.19%	2		94.125% $\pm$ 6.24%	2

Tabla 6: Resultados cuantificación en función de concentración del contaminante usando solo Coseno: a) con fondo, b) sin fondo (1 láser)

En la Tabla 7 se disponen los resultados tras utilizar únicamente la característica *Permutation Entropy* de imágenes explicada anteriormente. De dicho cuadro concluimos lo siguiente:

- Sin eliminar el fondo el mejor resultado que se consigue es de 87.125% mientras que el peor es de 60.500%. Del mismo modo cuando sustraemos dicho ruido, el mejor porcentaje de acierto es de 88.000% y el peor de 62.750%.
- El aceite que mejor respuesta presenta al utilizar solo esta variable es el Ac. 75.80 usando el color azul cuando no se quita el fondo y el color verde al eliminarlo.
- Los resultados de aciertos obtenidos (η) tanto si retiramos como si no el fondo son muy similares.

Se aprecia así que las tasas de acierto son bastante bajas al compararlas con las conseguidas usando el Coseno; por lo que las condiciones de estos experimentos son malas para obtener buena capacidad de cuantificación.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	73.625%$\pm$11.17%	3		73.000% $\pm$ 10.59%	Y
15.40	60.500%$\pm$9.65%	3		59.000% $\pm$ 10.94%	2
20.50	66.250%$\pm$10.86%	1		62.750% $\pm$ 10.78%	2
75.80	87.125% $\pm$ 8.87%	3		88.000%$\pm$6.66%	2
ATF	68.000% $\pm$ 10.61%	1		68.500%$\pm$10.48%	2

Tabla 7: Resultados cuantificación en función de concentración del contaminante usando solo *Permutation Entropy*: a) con fondo, b) sin fondo (1 láser)

En la Tabla 8 se presentan los datos conseguidos al usar la variable Ángulo entre las distintas imágenes, deduciendo así que:

- La peor tasa de acierto se presenta cuando no se elimina el fondo en el Ac.10.40 (91.625%); mientras que el mejor resultado es de 95.750% correspondiente al Ac. ATF al quitar dicho fondo.

- La escala de grises y el verde son los colores con los que se consiguen los resultados anteriores respectivamente.
- Obtenemos mejores resultados al retirar el fondo en las imágenes de todas las disoluciones.

Vemos que los resultados presentan un alto nivel de acierto y bastante similares cuando se usa solo el Coseno (Tabla 6); por lo que también en este caso conseguimos una alta capacidad de cuantificación para estos ensayos.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	91.625% $\pm$ 5.30%	Y		92.125%$\pm$5.18%	Y
15.40	93.250% $\pm$ 5.18%	2		94.750%$\pm$5.12%	2
20.50	93.500%$\pm$6.43%	1		93.375% $\pm$ 8.02%	1
75.80	92.000% $\pm$ 5.79%	2		95.500%$\pm$3.58%	3
ATF	94.625% $\pm$ 6.92%	Y		95.750%$\pm$5.27%	2

Tabla 8: Resultados cuantificación en función de concentración del contaminante usando solo Ángulo: a) con fondo, b) sin fondo (1 láser)

En la Tabla 9 se muestran los datos obtenidos tras combinar los parámetros estadísticos con las características de textura, concluyendo así que:

- En todos los aceites el color que mejor resultados da es el verde al usar dicha combinación de variables.
- El mejor y el peor resultado son 97.625% y 93.250% respectivamente. En el primer caso no se ha eliminado el ruido, mientras que en el otro si se suprime.
- El aceite que mejor responde es el Ac. 10.40 mientras que el peor es el Ac.15.40.

Estudiando estas cifras vemos que las tasas de acierto son bastante altas, por lo que al usar esta combinación disponemos de buena capacidad de cuantificación de la disolución.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	97.625%$\pm$3.31%	2		96.875% $\pm$ 3.84%	2
15.40	93.625%$\pm$6.62%	2		93.250% $\pm$ 6.29%	2
20.50	94.125%$\pm$6.49%	2		94.000% $\pm$ 7.36%	2
75.80	97.250%$\pm$5.07%	2		96.750% $\pm$ 5.24%	2
ATF	94.125% $\pm$ 6.24%	2		95.000%$\pm$6.68%	2

Tabla 9: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura y parám. estadísticos: a) con fondo, b) sin fondo (1 láser)

En el siguiente cuadro (Tabla 10) están los resultados obtenidos al combinar el Coseno y la *Permutation Entropy*, deduciendo así lo siguiente:

- El aceite Ac.ATF es el que presenta una mayor tasa de aciertos cuando no se suprime el ruido de fondo (93.125%), obteniendo dicho valor cuando usamos el color azul.
- Por el contrario, el aceite Ac.20.50 es el que posee el más bajo número de aciertos (78.625%) cuando se retira el fondo; logrando esta cifra al utilizar el color verde.
- Asimismo, vemos que los mejores resultados se obtienen sin suprimir el ruido del fondo.

Comparando estos valores con los representados en las tablas anteriores vemos que, a pesar de ser las tasas de acierto altas, éstas solo mejoran cuando se usa exclusivamente la variable *Permutation Entropy* mientras que para los otros casos son peores estos resultados.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	86.500%±8.41%	Y		86.250%±8.56%	3
15.40	83.625%±10.17%	2		83.000%±7.68%	Y
20.50	80.750%±10.77%	2		78.625%±9.28%	2
75.80	90.500%±7.05%	3		90.250%±5.81%	Y
ATF	93.125%±5.82%	3		92.375%±6.22%	3

Tabla 10: Resultados cuantificación en función de concentración del contaminante combinando Coseno y *Permutation Entropy*: a) con fondo, b) sin fondo (1 láser)

En la Tabla 11 están los resultados obtenidos al combinar el característ. de textura, parám. estadísticos y Coseno, deduciendo así lo siguiente:

- El aceite 75.80 es el que presenta una mayor tasa de aciertos al suprimir el ruido de fondo (98.375%), obteniendo dicho valor cuando usamos el color verde.
- Por el contrario, el aceite Ac.15.40 es el que posee el más bajo número de aciertos (93.625%) cuando se retira el fondo; logrando esta cifra al utilizar también el color verde.
- También vemos que todos los aceites tienen el mejor resultado al emplear el verde excepto en el 10.40, que se obtiene con el azul.

Comparando estos valores con los representados en las tablas anteriores vemos que esta combinación es una de las que mejores resultados nos da, por lo que buena capacidad de cuantificación de la disolución.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	97.250%±3.82%	3		96.375%±4.90%	3
15.40	90.750%±8.20%	2		93.625%±6.62%	2
20.50	94.875%±6.64%	2		94.500%±5.74%	2
75.80	98.125%±3.62%	2		98.375%±4.52%	2
ATF	94.875%±5.88%	2		94.250%±5.48%	2

Tabla 11: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura, parám. estadísticos y Coseno: a) con fondo, b) sin fondo (1 láser)

En la Tabla 12 se presentan los datos tras utilizar una combinación de las variables característ. de textura, parám. estadísticos y *Permutation Entropy* descritas con anterioridad. De aquí sacamos que:

- El valor medio de acierto más alto lo presenta el Ac.75.80 (99.000%) sin eliminar el fondo. Asimismo, el Ac. 15.40 es el que tiene el peor resultado (89.875%) cuando no se retira tampoco el ruido de fondo.
- El color que predomina y con el que se obtiene los mejores resultados es el verde.

Comparando estas cifras con los resultados utilizando las otras características apreciamos que hay un ligero incremento en la tasa de aciertos (η) usando una combinación de éstas que al emplear cada uno de forma individual.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	96.750%±4.04%	3		97.000%±4.04%	2
15.40	89.875%±7.87%	2		92.750%±5.41%	2
20.50	91.250%±7.78%	2		93.000%±7.21%	2
75.80	99.000%±2.31%	2		98.750%±3.34%	2
ATF	93.125%±7.05%	3		96.000%±4.51%	2

Tabla 12: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura, parám. estadísticos y *Permutation Entropy*: a) con fondo, b) sin fondo (1 láser)

En la Tabla 13 se muestran los datos obtenidos tras combinar el característ. de textura, parám. estadísticos y Ángulo, concluyendo así que:

- El mejor y el peor resultado son 98.625% y 91.500% respectivamente. En el primer caso no se ha eliminado el ruido, mientras que en el otro si se suprime.
- El aceite que mejor responde es el Ac. 75.80 mientras que el peor es el Ac.15.40.
- Las peores tasas de acierto se presentan cuando eliminamos el ruido de fondo.

Estudiando estas cifras vemos que las tasas de acierto son bastante altas, por lo que al usar esta combinación disponemos de buena capacidad de cuantificación de la disolución.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	96.875%±5.54%	2		96.375%±4.74%	3
15.40	91.250%±6.80%	2		91.500%±7.00%	2
20.50	94.625%±5.65%	3		93.875%±6.85%	2
75.80	98.625%±3.63%	2		98.000%±4.96%	2
ATF	94.250%±6.29%	2		94.125%±5.56%	2

Tabla 13: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura, parám. estadísticos y Ángulo: a) con fondo, b) sin fondo (1 láser)

En el siguiente cuadro (Tabla 14) están los resultados obtenidos al combinar característ. de textura, parám. estadísticos, *Permutation Entropy* y Ángulo, deduciendo así lo siguiente:

- El aceite Ac.75.80 es el que presenta una mayor tasa de aciertos cuando no se suprime el ruido de fondo (99.000%), obteniendo dicho valor cuando usamos el color verde.
- Por el contrario, el aceite Ac.15.40 es el que posee el más bajo número de aciertos (91.375%) cuando no quitamos el fondo al emplear el color verde.
- Asimismo, vemos que los mejores resultados se obtienen sin suprimir el ruido del fondo.

Comparando estos valores con los indicados en las tablas anteriores vemos que presentan valores bastante altos, incluidos los resultados más bajos, por lo que al emplear dicha combinación dispondremos de buena capacidad de cuantificación.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	97.750%±3.74%	3		97.750%±3.52%	3
15.40	91.375%±6.42%	2		91.000%±6.20%	2
20.50	91.750%±8.79%	2		92.500%±7.03%	3
75.80	99.000%±2.63%	2		96.875%±5.68%	2
ATF	95.875%±4.30%	2		94.875%±5.00%	1

Tabla 14: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura, parám. estadísticos, *Permutation Entropy* y Ángulo: a) con fondo, b) sin fondo (1 láser)

En la Tabla 15 están los resultados obtenidos al combinar el característ. de textura, parám. estadísticos, *Permutation Entropy*, Ángulo y Coseno, deduciendo así lo siguiente:

- El aceite 10.40 es el que presenta una mayor tasa de aciertos sin suprimir el ruido de fondo (97.750%), obteniendo dicho valor cuando usamos el color azul.
- Por el contrario, el aceite Ac.15.40 es el que posee el más bajo número de aciertos (90.750%) cuando se retira el fondo; logrando esta cifra al utilizar el color verde.
- También vemos que casi todos los aceites obtienen mejores resultados al usar el color azul cuando no se suprime el fondo, excepto en el Ac.15.40 que utiliza el verde.

Comparando estos valores con los representados en las tablas anteriores vemos que esta combinación es una de las que mejores resultados nos da, por lo que buena capacidad de cuantificación de la disolución.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	97.750%±3.73%	3		96.500%±4.22%	0
15.40	89.125%±6.66%	2		90.750%±6.71%	2
20.50	93.375%±5.42%	3		91.750%±7.83%	3
75.80	97.125%±3.62%	3		97.500%±4.37%	3
ATF	95.500%±6.19%	3		95.875%±4.48%	0

Tabla 15: Resultados cuantificación en función de concentración del contaminante combinando característ. de textura, parám. estadísticos, *Permutation Entropy*, Ángulo y Coseno: a) con fondo, b) sin fondo (1 láser)

Finalmente, en la Tabla 16 se indica para cada aceite cuál es la característica o combinación de éstas en la que se obtiene un mejor resultado, así como si la sustracción o no del fondo y el color con el que se consiguen dichos datos al utilizar solo un láser.

Aceite	Característica	Fondo	RGB	$\eta \pm \sigma$
10.40	Textura, estadísticos, <i>PE</i> y Ángulo	Sin fondo	B	97.750%±3.52%
15.40	Solo Ángulo	Sin fondo	G	94.750%±5.12%
20.50	Parámetros estadísticos	Con fondo	B	95.125%±5.96%
75.80	Parámetros estadísticos	Con fondo	G	99.125%±2.53%
ATF	Textura, estadísticos y <i>Permutation Entropy</i>	Sin fondo	G	96.000%±4.51%

Tabla 16: Resumen Cuantificación en función de concentración del contaminante (1 láser)

6.1.1.2 Iluminación con 2 láseres.

En la Tabla 17 se presentan los datos conseguidos al usar únicamente las Características de Textura entre las distintas imágenes, deduciendo así que:

- La peor tasa de acierto se presenta cuando no se elimina el fondo en el Ac.20.50 (72.750%); mientras que el mejor resultado es de 91.000% correspondiente al Ac. 75.80 quitando en este caso el ruido de fondo.
- En los resultados de los aceites 10.40 y 15.40 encontramos una mayor similitud, ya sea eliminando o no el fondo, usando en ambos casos la combinación de grises.
- En las otras disoluciones se obtienen mejores tasas de acierto al emplear el color azul.

En términos generales podemos considerar que las tasas de acierto altas; dando a entender que en las condiciones de estos experimentos tenemos buena capacidad de cuantificación de la disolución.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	89.000%±6.86%	Y		89.625%±7.21%	Y
15.40	83.500%±8.72%	Y		83.000%±8.38%	Y
20.50	72.750%±12.16%	3		73.750%±9.70%	3
75.80	87.000%±9.43%	3		91.000%±8.29%	3
ATF	86.625%±7.98%	3		86.000%±9.07%	3

Tabla 17: Resultados cuantificación en función de concentración del contaminante usando solo Característ. de textura: a) con fondo, b) sin fondo (2 láseres)

En la Tabla 18 se disponen los resultados tras utilizar los parámetros estadísticos explicados anteriormente. De dicho cuadro concluimos lo siguiente:

- La mejor respuesta en todos los aceites se encuentra al no eliminar el ruido de fondo; coincidiendo todos en utilizar el azul para hallar esos resultados excepto del Ac. 20.50, que emplea el verde.
- El aceite que mejor respuesta presenta al utilizar estas características es el Ac. ATF usando el color azul sin quitar el fondo.
- Los resultados de aciertos obtenidos (η) tanto si retiramos como si no el fondo, son muy similares.

Se aprecia así que las tasas de acierto son bastante altas al compararlas con las de la Tabla 17; por lo que las condiciones de estos experimentos son mejores que al emplear únicamente las características de textura.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	96.000%±4.85%	3		95.625%±5.54%	3
15.40	94.250%±5.48%	3		93.250%±7.23%	3
20.50	94.375%±5.24%	2		92.875%±7.14%	2
75.80	96.500%±4.02%	3		96.125%±3.97%	3
ATF	96.875%±4.04%	3		96.750%±4.93%	3

Tabla 18: Resultados cuantificación en función de concentración del contaminante usando solo Parámetros estadísticos: a) con fondo, b) sin fondo (2 láseres)

En la Tabla 19 se establecen los datos correspondientes al utilizar únicamente el Coseno, concluyendo así que:

- De forma genérica, los mejores datos se logran cuando se suprime el fondo en todas las imágenes.
- El aceite Ac. 10.40 presenta la mayor tasa de acierto de todas las combinaciones realizadas (98.000%), mientras que el peor resultado lo posee el aceite Ac.15.40 (92.375%).
- Para cada aceite se utiliza el mismo color tanto si se sustrae o no el fondo.

Se aprecia que las tasas de acierto son bastantes altas, dando a entender que las condiciones de estos experimentos generan una buena capacidad de cuantificación de la disolución. Asimismo, si comparamos estos datos con los indicados en la Tabla 6, vemos que se logran mejores resultados usando dos láseres que solo uno.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	97.500%±3.79%	2		98.000%±3.44%	2
15.40	92.375%±7.71%	Y		94.125%±6.96%	Y
20.50	93.875%±4.80%	2		94.375%±5.24%	2
75.80	95.125%±4.94%	3		96.125%±3.54%	3
ATF	96.750%±5.09%	1		95.500%±7.03%	1

Tabla 19: Resultados cuantificación en función de concentración del contaminante usando solo Coseno: a) con fondo, b) sin fondo (2 láseres)

En el siguiente cuadro (Tabla 20) están los resultados obtenidos al usar solo *Permutation Entropy*, deduciendo así lo siguiente:

- El aceite Ac. ATF es el que presenta una mayor tasa de aciertos cuando se suprime el ruido de fondo (88.625%), obteniendo dicho valor cuando usamos la escala de grises.

- Por el contrario, el aceite Ac.20.50 es el que posee el más bajo número de aciertos (75.875%) retirando también el fondo; logrando esta cifra al utilizar el color azul.
- Asimismo, vemos que los peores resultados se obtienen sin suprimir el ruido del fondo.

Al igual que ocurre cuando usamos un solo láser, los resultados obtenidos son peores al trabajar solo con la misma variable. Aun así, estos valores mejoran al usar dos láseres que al hacerlo con solo uno.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	70.125% $\pm$ 8.95%	1		82.500%$\pm$8.37%	2
15.40	78.000% $\pm$ 9.80%	Y		78.125%$\pm$9.74%	1
20.50	76.500%$\pm$12.01%	1		75.875% $\pm$ 9.45%	3
75.80	84.375% $\pm$ 7.91%	1		97.500%$\pm$4.18%	Y
ATF	82.500% $\pm$ 10.18%	3		88.625%$\pm$8.53%	Y

Tabla 20: Resultados cuantificación en función de concentración del contaminante usando solo Permutation Entropy: a) con fondo, b) sin fondo (2 láseres)

En la Tabla 21 se muestran los datos obtenidos usar únicamente el Ángulo, deduciendo así que:

- El mejor y el peor resultado son 98.125% y 91.500% respectivamente; sin suprimir en ambos casos el ruido de fondo.
- El aceite que mejor responde es el Ac. 10.40 mientras que el peor es el Ac.15.40.
- Los únicos aceites que utilizan el mismo color tanto si se elimina como si no el ruido de fondo son el Ac. 15.40 (escala de grises) y Ac. 75.80 (azul).

Al igual que en los dos casos anteriores, se han mejorado los resultados al usar dos láseres en vez de un láser al trabajar con la misma variable. También vemos que presenta una alta tasa de aciertos, por lo que genera una buena capacidad de cuantificación.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	98.125%$\pm$4.04%	1		97.875% $\pm$ 4.11%	2
15.40	91.500% $\pm$ 7.33%	Y		94.000%$\pm$5.64%	Y
20.50	95.125%$\pm$5.25%	3		94.875% $\pm$ 4.84%	2
75.80	95.250% $\pm$ 4.47%	3		95.750%$\pm$5.12%	3
ATF	96.000%$\pm$6.02%	1		93.875% $\pm$ 5.85%	Y

Tabla 21: Resultados cuantificación en función de concentración del contaminante usando solo Ángulo: a) con fondo, b) sin fondo (2 láseres)

En la Tabla 22 se indican los datos al usar una combinación de Características de Textura y Parámetros estadísticos, sacando de aquí las siguientes conclusiones:

- El aceite que da una mejor respuesta es el Ac. 10.40 (97.000%) sin eliminar el ruido de fondo; mientras que el peor lo obtenemos en el Ac. 15.40 (91.875%) sin quitar el fondo.
- Se obtienen mejores resultados al no eliminar el fondo, a excepción del Ac.20.50.

Se han mejorado los resultados obtenidos al pasar de uno a dos láseres trabajando con las mismas características en ambos casos. Asimismo, las tasas de acierto son altas, dando a entender que en las condiciones de estos experimentos tenemos buena capacidad de cuantificación de la disolución.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	97.000%±4.42%	2		96.125%±5.03%	2
15.40	92.000%±7.47%	Y		91.875%±8.40%	Y
20.50	93.750%±5.92%	2		94.375%±6.09%	3
75.80	97.500%±3.57%	3		96.750%±4.04%	3
ATF	94.375%±7.06%	1		94.250%±5.48%	3

Tabla 22: Resultados cuantificación en función de concentración del contaminante usando con característ. de textura y parám. estadísticos: a) con fondo, b) sin fondo (2 láseres)

En la Tabla 23 se presentan los datos conseguidos al combinar Coseno y *Pemutiton Entropy* entre las distintas imágenes; deduciendo así que:

- La peor tasa de acierto se presenta en el Ac. 20.50 (84.500%); mientras que el mejor resultado es de 94.125% correspondiente al Ac. 75.80 al quitar en ambos casos el ruido del fondo.
- El rojo y el verde son los colores con los que se consiguen los resultados anteriores respectivamente.

Al igual que en los casos anteriormente descritos, se han mejorado los resultados al usar dos láseres en vez de un láser al trabajar con la misma combinación. Además, vemos que presenta una alta tasa de aciertos, por lo que genera una buena capacidad de cuantificación.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	90.375% $\pm$ 7.70%	3		90.875%$\pm$7.16%	Y
15.40	90.625%$\pm$6.47%	Y		87.625% $\pm$ 7.73%	Y
20.50	84.750%$\pm$7.27%	3		84.500% $\pm$ 7.06%	1
75.80	93.625% $\pm$ 5.71%	1		94.125%$\pm$3.87%	2
ATF	90.375%$\pm$9.30%	1		89.625% $\pm$ 8.04%	3

Tabla 23: Resultados cuantificación en función de concentración del contaminante combinando Coseno y Pemititon Entropy: a) con fondo, b) sin fondo (2 láseres)

En la Tabla 24 están los resultados obtenidos al combinar características de textura, parámetros estadísticos y Coseno, deduciendo así lo siguiente:

- El aceite Ac. 10.40 es el que presenta una mayor tasa de aciertos cuando no se suprime el ruido de fondo (98.375%), obteniendo dicho valor cuando usamos el color verde.
- Por el contrario, el aceite Ac.15.40 es el que posee el más bajo número de aciertos (91.750%) sin retirar el fondo; logrando esta cifra al utilizar el color azul.
- Asimismo, vemos que todos los aceites presentan los mejores resultados usando el azul menos en el Ac.10.40 que se obtiene al usar el verde.

Siguiendo con el patrón de los casos anteriores, se ha mejorado la tasa de aciertos usando dos láseres que uno solo (Tabla 11) al trabajar con la misma combinación de características.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	98.375%$\pm$3.04%	2		97.500% $\pm$ 3.34%	2
15.40	92.875%$\pm$6.44%	3		91.750% $\pm$ 5.71%	3
20.50	94.625% $\pm$ 6.80%	3		94.750%$\pm$5.97%	3
75.80	97.125% $\pm$ 4.03%	3		97.875%$\pm$3.48%	3
ATF	96.000%$\pm$4.51%	3		95.750% $\pm$ 5.12%	3

Tabla 24: Resultados cuantificación en función de concentración del contaminante combinando características de textura, parámetros estadísticos y Coseno: a) con fondo, b) sin fondo (2 láseres)

En la Tabla 25 se muestran los datos obtenidos tras usar características de textura, parámetros estadísticos y *Permutation Entropy*, deduciendo así que:

- El mejor y el peor resultado son 96.250% y 92.625% respectivamente; suprimiendo únicamente en el primer caso el ruido de fondo.
- El aceite que mejor responde es el Ac.ATF mientras que el peor es el Ac.15.40.
- Todos los aceites utilizan el mismo color tanto si se elimina como si no el ruido de fondo menos el Ac.10.40.

Al igual que en los experimentos anteriores, se obtienen resultados bastante buenos, por lo que genera una buena capacidad de cuantificación, siendo ligeramente mejores éstos que cuando trabajamos con un solo láser.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	93.375% $\pm$ 6.36%	3		95.750%$\pm$5.27%	2
15.40	92.625%$\pm$5.46%	3		91.875% $\pm$ 7.49%	3
20.50	94.000% $\pm$ 5.35%	2		95.250%$\pm$5.44%	2
75.80	96.500%$\pm$3.61%	3		94.375% $\pm$ 5.09%	3
ATF	94.500% $\pm$ 5.87%	3		96.250%$\pm$4.89%	3

Tabla 25: Resultados cuantificación en función de concentración del contaminante combinando características de textura, parámetros estadísticos y *Permutation Entropy*: a) con fondo, b) sin fondo (2 láseres)

En la Tabla 26 se muestran los datos obtenidos al usar características de textura, parámetros estadísticos y Ángulo, deduciendo así que:

- El mejor y el peor resultado son 97.500% y 92.000% respectivamente; sin suprimir en ambos casos el ruido de fondo.
- El aceite que mejor responde es el Ac. 75.80 mientras que el peor es el Ac.15.40.
- Los únicos aceites que utilizan el mismo color tanto si se elimina como si no el ruido de fondo son el Ac. 20.50 y 75.80, usando en ambos casos el azul.

Al igual que en los dos casos anteriores, se han mejorado los resultados al usar dos láseres en vez de un láser al trabajar con la misma variable. También vemos que presenta una alta tasa de aciertos, por lo que genera una buena capacidad de cuantificación.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	96.750% $\pm$ 5.68%	2		96.875%$\pm$5.10%	Y
15.40	92.000% $\pm$ 7.47%	3		94.375%$\pm$9.38%	1
20.50	94.125% $\pm$ 5.42%	3		95.125%$\pm$5.40%	3
75.80	97.500%$\pm$3.57%	3		97.375% $\pm$ 3.59%	3
ATF	96.375%$\pm$4.39%	3		95.875% $\pm$ 4.82%	1

Tabla 26: Resultados cuantificación en función de concentración del contaminante combinando características de textura, parámetros estadísticos y Ángulo: a) con fondo, b) sin fondo (2 láseres)

En el siguiente cuadro (Tabla 27) están los resultados obtenidos al combinar características de textura, parámetros estadísticos *Permutation Entropy* y Ángulo, de tal manera que:

- El aceite Ac.10.40 es el que presenta una mayor tasa de aciertos tanto si se suprime como si no el fondo. En el primer caso se usa el color azul mientras que para el segundo es verde.

- Por el contrario, el aceite Ac.15.40 es el que posee el más bajo número de aciertos (93.125%) cuando se retira el fondo; logrando esta cifra al utilizar el color azul.

Al igual que en los casos anteriormente descritos, se han mejorado los resultados al usar dos láseres en vez de un láser al trabajar con la misma combinación. Además, vemos que presenta una alta tasa de aciertos, por lo que genera una buena capacidad de cuantificación.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	96.000%±5.47%	2		96.000%±4.85%	3
15.40	93.375%±6.49%	3		93.125%±6.22%	3
20.50	93.625%±5.43%	2		94.750%±5.41%	2
75.80	95.250%±4.65%	3		94.375%±6.08%	2
ATF	95.125%±5.25%	3		94.250%±6.90%	1

Tabla 27: Resultados cuantificación en función de concentración del contaminante combinando características de textura, parámetros estadísticos *Permutation Entropy* y Ángulo: a) con fondo, b) sin fondo (2 láseres)

En la Tabla 28 están los resultados obtenidos al combinar el característ. de textura, parám. estadísticos, *Permutation Entropy*, Ángulo y Coseno, deduciendo así lo siguiente:

- El aceite 75.80 es el que presenta una mayor tasa de aciertos sin suprimir el ruido de fondo (97.875%), obteniendo dicho valor cuando usamos el color azul.
- Por el contrario, el aceite Ac.15.40 es el que posee el más bajo número de aciertos (93.000%) cuando se retira el fondo; logrando esta cifra al utilizar el color rojo.

Comparando estos valores con los presentados anteriormente vemos que esta combinación es una de las que mejores resultados nos da, por lo que buena capacidad de cuantificación de la disolución.

Aceite	$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
10.40	96.500%±5.94%	Y		95.750%±4.80%	Y
15.40	93.125%±6.47%	3		93.000%±8.53%	1
20.50	94.000%±6.05%	2		94.750%±6.48%	2
75.80	97.875%±3.00%	3		95.750%±4.63%	3
ATF	95.500%±5.51%	1		95.000%±6.56%	3

Tabla 28: Resultados cuantificación en función de concentración del contaminante combinando características de textura, parámetros estadísticos *Permutation Entropy*, Ángulo y Coseno: a) con fondo, b) sin fondo (2 láseres)

Tal y como hicimos en el apartado anterior en la Tabla 16 se indican con qué característica se obtienen los mejores resultados para cada aceite, además de especificar si se quita o no el ruido de fondo y el color que se ha utilizado.

Aceite	Característica	Fondo	RGB	$\eta \pm \sigma$
10.40	Textura, Estadística y Coseno	Con fondo	G	98.375%±3.04%
15.40	Textura, Estadísticos y Ángulo	Sin fondo	R	94.375%±9.38%
20.50	Textura, Estadísticos y PE	Sin fondo	G	95.250%±5.44%
75.80	Textura, Estadística, PE y Coseno	Sin fondo	B	97.875%±3.48%
ATF	Solo Estadísticos	Con fondo	B	96.875%±4.04%

Tabla 29: Resumen Cuantificación en función de concentración del contaminante (2 láseres)

6.1.2 Cuantificación “ciega” de la concentración de contaminante.

6.1.2.1 Iluminación con 1 láser.

Los resultados de estos experimentos se presentan en las siguientes tablas dadas en función de las características utilizadas:

- Tabla 30: Características de textura.
- Tabla 31: Parámetros estadísticos.
- Tabla 32: Coseno.
- Tabla 33: PE.
- Tabla 34: Ángulo.
- Tabla 35: Característ. de textura y parám. estadísticos.
- Tabla 36: Coseno y PE.
- Tabla 37: Característ. de textura, parám. estadísticos y Coseno.
- Tabla 38: Característ. de textura, parám. estadísticos y PE.
- Tabla 39: Característ. de textura, parám. estadísticos y Ángulo.
- Tabla 40: Característ. de textura, parám. estadísticos, PE y Ángulo.
- Tabla 41: Característ. de textura, parám. estadísticos, PE, Ángulo y Coseno.

Se aprecia en dichos cuadros que al usar únicamente las características de textura (Tabla 30); se obtienen una baja tasa de acierto (48%) al usar el color azul quitando el ruido de fondo. Asimismo, si estudiamos únicamente los parámetros estadísticos (Tabla 31), vemos que nos da un alto porcentaje de aciertos, así como una baja desviación típica ($83.354\% \pm 3.77\%$), utilizando también el color azul y sin quitar el fondo.

En la Tabla 32 vemos que el porcentaje de acierto también es elevado si usamos exclusivamente el Coseno entre las imágenes, siendo estos valores similares a utilizar también solo el Ángulo entre ellas (Tabla 34). En ambos se utiliza el azul para obtener dichos resultados.

El peor resultado se encuentra al utilizar exclusivamente el *Permutation Entropy* al obtener entorno al 35% de aciertos tanto si se suprime como si no el fondo (Tabla 33), usando en ambos casos también el azul (B).

En la Tabla 35 se indican los resultados tras utilizar características de textura y parámetros estadísticos. En este caso vemos que obtenemos una alta tasa de acierto (83.229%) al eliminar el fondo y usando el color azul.

Al combinar los parámetros Coseno y PE (Tabla 36) se aprecia una disminución del nivel de aciertos, así como un aumento en la desviación al utilizar el mismo color y al suprimir el fondo ($71.396\% \pm 5.09\%$), por lo que no sería tan beneficioso emplear esta combinación comparada con la anterior.

Del mismo modo estudiamos cuál es el mejor resultado al combinar las características de textura y parámetros estadísticos con Coseno, PE y Ángulo de forma independiente. Los resultados obtenidos indican que la mejor combinación es la que incluye el Coseno con un 83.730% de acierto (Tabla 37), seguida por el Ángulo con una tasa del 83.646% (Tabla 39) y finalmente por el PE con un número de aciertos de 82.708% (Tabla 38).

Otras combinaciones que se han estudiado han sido las de estudiar características de textura, parámetros estadísticos, PE y Ángulo en (Tabla 40) y otra en la que se une todas éstas con Coseno (Tabla 41). Los resultados establecen que al incluir esta última variable mejora ligeramente el porcentaje de aciertos ($85.125\% \pm 3.37\%$) aunque el valor de la desviación típica es casi igual; mientras que si no lo incluimos obtenemos una tasa de aciertos del $84.854\% \pm 3.35\%$.

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
$47.041\% \pm 4.21\%$	3		$48.125\% \pm 5.66\%$	3

Tabla 30: Resultados cuantificación "ciega" usando solo Características de textura: a) con fondo, b) sin fondo (1láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
$83.354\% \pm 3.77\%$	3		$81.687\% \pm 3.97\%$	3

Tabla 31: Resultados cuantificación "ciega" usando solo Parámetros estadísticos: a) con fondo, b) sin fondo (1láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
$81.104\% \pm 4.17\%$	3		$80.479\% \pm 4.23\%$	3

Tabla 32: Resultados cuantificación "ciega" usando solo Coseno: a) con fondo, b) sin fondo (1láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
35.208%±3.84%	3		35.479%±4.04%	3

Tabla 33: Resultados cuantificación "ciega" usando solo Permutation Entropy: a) con fondo, b) sin fondo (1láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
81.292%±3.51%	3		82.167%±4.22%	2 3

Tabla 34: Resultados cuantificación "ciega" usando solo Ángulo: a) con fondo, b) sin fondo (1láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
82.708%±3.91%	3		83.229%±3.88%	3

Tabla 35: Resultados cuantificación "ciega" usando con característ. de textura y parám. estadísticos: a) con fondo, b) sin fondo (1láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
70.667%±3.63%	3		71.396%±5.09%	3

Tabla 36: Resultados cuantificación "ciega" combinando Coseno y Permutation Entropy: a) con fondo, b) sin fondo (1 láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
82.312%±4.03%	3		83.730%±3.86%	3

Tabla 37: Resultados cuantificación "ciega" usando característ. de textura, parám. estadísticos y Coseno: a) con fondo, b) sin fondo (1láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
82.708%±4.19%	3		82.208%±3.84%	3

Tabla 38: Resultados cuantificación "ciega" usando característ. de textura, parám. estadísticos y Permutation Entropy: a) con fondo, b) sin fondo (1láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
83.312%±3.83%	3		83.646%±3.48%	3

Tabla 39: Resultados cuantificación "ciega" combinando característ. de textura, parám. estadísticos y Ángulo: a) con fondo, b) sin fondo (1láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
84.854%±3.35%	3		83.583%±3.32%	3

Tabla 40: Resultados cuantificación "ciega" combinando característ. de textura, parám. estadísticos, Permutation Entropy y Ángulo: a) con fondo, b) sin fondo (1 láser)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
84.604% $\pm$ 3.11%	3		85.125%$\pm$3.37%	3

Tabla 41: Resultados cuantificación "ciega" combinando característ. de textura, parám. estadísticos, Permutation Entropy, Ángulo y Coseno: a) con fondo, b) sin fondo (1 láser)

Finalmente podemos concluir que, al desconocer el tipo de aceite, así como su concentración, en la Tabla 42 vemos que al hacer competir 'todos contra todos' la mejor tasa de acierto (85.125%) se obtiene al usar las características de textura, sin eliminar el fondo y usando el color azul (B) al usar un solo láser.

Característica	Fondo	RGB	$\eta \pm \sigma$
Textura, Estadísticos, PE, Ángulo y Coseno	Sin fondo	B	85.125%$\pm$3.37%

Tabla 42: Resumen Cuantificación "ciega" (1 láser)

6.1.2.2 Iluminación con 2 láseres.

Al igual que en el apartado anterior, los resultados de estos experimentos se presentan en las siguientes tablas dadas en función de las características utilizadas:

- Tabla 43: Características de textura.
- Tabla 44: Parámetros estadísticos.
- Tabla 45: Coseno.
- Tabla 46: PE.
- Tabla 47: Ángulo.
- Tabla 48: Característ. de textura y parám. estadísticos.
- Tabla 49: Coseno y PE.
- Tabla 50: Característ. de textura, parám. estadísticos y Coseno.
- Tabla 51: Característ. de textura, parám. estadísticos y PE.
- Tabla 52: Característ. de textura, parám. estadísticos y Ángulo.
- Tabla 53: Característ. de textura, parám. estadísticos, PE y Ángulo.
- Tabla 54: Característ. de textura, parám. estadísticos, PE, Ángulo y Coseno.

En la Tabla 43 se refleja que al tener en cuenta únicamente las características de textura, se obtienen mejores tasas de acierto al utilizar el color azul y eliminando el fondo, consiguiendo así una media de aciertos del 50.000%.

Si estudiamos los parámetros estadísticos vemos que dicha media aumenta hasta alcanzar un nivel de aciertos del 87.625% (Tabla 44) al utilizar el mismo color y sin eliminar el ruido de fondo.

Si comparamos estos resultados con los indicados en la Tabla 45 vemos que el porcentaje de acierto disminuye ligeramente al usar como única variable el Coseno entre las imágenes (85.854%) sin quitar el ruido de fondo y usando el color verde.

En la Tabla 46 están reflejados los resultados tras usar únicamente el parámetro *Permutation Entropy* y, vemos que se obtiene la peor tasa de aciertos de todas las combinaciones estudiadas (57.896%) cuando retiramos el ruido de fondo y con el color verde. Sin quitar el ruido de fondo y usando el color verde conseguimos una de aciertos del 86.083% al emplear únicamente el parámetro Ángulo (Tabla 47).

Asimismo, en la Tabla 48 están reflejados los resultados al combinar características de textura y parámetros estadísticos, obteniendo así una media de acierto alta y con una baja desviación típica ($85.583\% \pm 3.07\%$).

Cuando combinamos los parámetros Coseno y PE (Tabla 49) se aprecia una disminución del nivel de aciertos al utilizar el mismo color y al suprimir el fondo (76.687%), por lo que no es recomendable usar esta combinación si la comparamos con las anteriores.

Del mismo modo estudiamos cuál es el mejor resultado al combinar estas mismas variables con Coseno, PE y Ángulo de forma independiente. Los resultados obtenidos indican que la mejor combinación es la que incluye el Coseno con un 88.208% de acierto (Tabla 50), seguida por el Ángulo con una tasa del 87.980% (Tabla 52) y finalmente por el PE con un número de aciertos de 86.480% (Tabla 51).

Otras combinaciones que se han estudiado han sido las de estudiar características de textura, parámetros estadísticos, PE y Ángulo en (Tabla 53) y otra en la que se une todas éstas con Coseno (Tabla 54). Los resultados establecen que al incluir esta última variable mejora ligeramente el porcentaje de aciertos ($88.666\% \pm 3.48\%$) aunque el valor de la desviación típica es casi igual; mientras que si no lo incluimos obtenemos una tasa de aciertos del $87.980\% \pm 2.62\%$.

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
$49.917\% \pm 4.47\%$	3		$50.000\% \pm 4.26\%$	3

Tabla 43: Resultados cuantificación "ciega" usando solo Características de textura: a) con fondo, b) sin fondo (2láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
$87.625\% \pm 3.27\%$	3		$87.333\% \pm 4.39\%$	3

Tabla 44: Resultados cuantificación "ciega" usando solo Parámetros estadísticos: a) con fondo, b) sin fondo (2láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
$85.854\% \pm 3.68\%$	2		$85.500\% \pm 3.67\%$	2

Tabla 45: Resultados cuantificación "ciega" usando solo Coseno: a) con fondo, b) sin fondo (2 láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
57.521%±5.48%	1		57.896%±4.56%	2

Tabla 46: Resultados cuantificación "ciega" usando solo Permution Entropy: a) con fondo, b) sin fondo (2 láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
86.083%±2.92%	2		85.583%±4.27%	2

Tabla 47: Resultados cuantificación "ciega" usando solo Ángulo: a) con fondo, b) sin fondo (2 láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
85.583%±3.07%	3		85.542%±3.44%	3

Tabla 48: Resultados cuantificación "ciega" usando característ. de textura y parám. estadísticos: a) con fondo, b) sin fondo (2 láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
76.354%±4.70%	3		76.687%±4.22%	3

Tabla 49: Resultados cuantificación "ciega" combinando Coseno y Pemutiton Entropy: a) con fondo, b) sin fondo (2 láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
87.562%±3.18%	3		88.208%±3.94%	3

Tabla 50: Resultados cuantificación "ciega" usando característ. de textura, parám. estadísticos y Coseno: a) con fondo, b) sin fondo (2 láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
86.480%±3.28%	3		85.780%±3.39%	3

Tabla 51: Resultados cuantificación "ciega" usando característ. de textura, parám. estadísticos y Permution Entropy: a) con fondo, b) sin fondo (2 láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
87.792%±2.81%	3		87.980%±2.62%	3

Tabla 52: Resultados cuantificación "ciega" combinando característ. de textura, parám. estadísticos y Ángulo: a) con fondo, b) sin fondo (2 láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
88.666%±3.48%	3		87.648%±3.37%	3

Tabla 53: Resultados cuantificación "ciega" combinando característ. de textura, parám. estadísticos, Permution Entropy y Ángulo: a) con fondo, b) sin fondo (2 láseres)

$\eta \pm \sigma$	RGB		$\eta \pm \sigma$	RGB
87.625%±3.30%	3		87.562%±3.44%	3

Tabla 54: Resultados cuantificación “ciega” combinando característ. de textura, parám. estadísticos, *Permutation Entropy*, Ángulo y Coseno: a) con fondo, b) sin fondo (2 láseres)

Finalmente, en la Tabla 55 vemos que al hacer competir ‘todos contra todos’ la mejor tasa de acierto (88.666%) se obtiene al usar las característ. de textura, parám. estadísticos, *Permutation Entropy* y Ángulo, sin suprimir el fondo y usando el color azul (B) al usar dos láseres.

Característica	Fondo	RGB	$\eta \pm \sigma$
Textura, estadísticos, PE y Ángulo	Con fondo	B	88.666%±3.48%

Tabla 55: Resumen Cuantificación “ciega” (2 láseres)

6.2 Conclusiones

Tras estudiar los resultados indicados en las Tablas 16 y 29 correspondiente al resumen de cuantificación según la concentración del contaminante usando 1 láser y 2 láseres respectivamente, llegamos a las conclusiones siguientes:

- En el aceite 10.40 la mejor combinación es la de características de textura, parámetros estadísticos y Coseno, sin quitar el ruido de fondo, usando el color verde y empleando dos láseres. Así se alcanza un 98.375% de aciertos.
- Para el aceite 15.40 la característica con la que obtenemos la mejor tasa de aciertos es el utilizar únicamente Ángulo, suprimiendo el fondo y utilizando el color verde con 1 láser.
- En el aceite 20.50 usamos características de textura, parámetros estadísticos y PE, eliminando el ruido de fondo al utilizar el verde y empleando dos láseres para conseguir el mejor resultado.
- Conseguimos la mayor tasa de aciertos y una baja desviación típica de todas las combinaciones en el aceite 75.80 (99.125%±2.53%) al usar exclusivamente los parámetros estadísticos, con el ruido de fondo, empleando el color verde y utilizando un láser.
- Con este mismo parámetro hemos obtenido el mejor número de aciertos para el aceite ATF (96.875%) al usar el color azul sin suprimir el fondo y empleando dos láseres.

Todo esto se refleja en la siguiente tabla, de tal manera que se indica para cada tipo de aceite cuál es la mejor característica/combinación, así como si se elimina o no el fondo, el color con el que se obtienen y si usamos uno o dos láseres.

Aceite	Característica	Fondo	RGB	Láser
10.40	Textura, Estadística y Coseno	Con fondo	G	2 láseres
15.40	Solo Ángulo	Sin fondo	G	1 láser
20.50	Textura, Estadísticos y PE	Sin fondo	G	2 láseres
75.80	Solo estadísticos	Con fondo	G	1 láser
ATF	Solo estadísticos	Con fondo	B	2 láseres

Tabla 56: Conclusiones Cuantificación en función de la concentración del contaminante

Del mismo modo estudiamos cuál es la mejor combinación para la cuantificación “ciega” en la que compiten todos contra todos. Comparando los resultados reflejados en las Tablas 42 y 55 en la que se indican las tasas de acierto tanto para un láser como para dos respectivamente, vemos que la mejor combinación es utilizar características de textura, parámetros estadísticos, PE y Ángulo, sin quitar el fondo, usando el azul y empleando dos láseres.

Característica	Fondo	RGB	Láser
Textura, estadísticos, PE y Ángulo	Con fondo	B	2 láseres

Tabla 57: Conclusiones Cuantificación “ciega”

Capítulo 7: Conclusiones y líneas futuras.

En este capítulo se muestran las conclusiones a las que se ha llegado con el desarrollo del trabajo presentado en este documento, así como con la comparativa de resultados obtenidos en los apartados anteriores. Además, presentamos una serie de líneas futuras de trabajo que surgen tras realizar este Trabajo Fin de Máster.

7.1 Conclusiones

Como primera conclusión se ha de destacar que hemos fabricado un prototipo de espectrofotómetro por el que hemos obtenido buenas imágenes para posteriormente extraer información útil que nos sirva para poder identificar diferentes contaminantes.

La segunda conclusión está en que hemos podido diferenciar entre varias clases aceites, pudiendo discriminar a su vez en diferentes concentraciones de éstos. En el capítulo 6 se muestra que obtenemos una alta tasa de acierto (llegando en algunos casos al 99%) con una baja desviación típica; por lo que obtenemos una buena capacidad de cuantificación de las distintas disoluciones.

Como se ha expuesto en este TFM hemos utilizado tanto un láser como dos láseres para estudiar con cuál de los dos métodos obtenemos mejores tasas de acierto. Después de comparar los resultados de ambas combinaciones, se observa que al utilizar dos láseres los resultados son mejores que al emplear solo un láser. Esto es debido a que conseguimos mejor iluminación y, en consecuencia, imágenes donde se aprecia muy bien la riqueza espectral de las sustancias bajo estudio.

Del mismo modo se aprecia que en la cuantificación en función de la concentración del contaminante, se obtienen mayores tasas de acierto al emplear el color verde y sin suprimir el ruido de fondo de la cámara. Asimismo, vemos que en la cuantificación “ciega” al usar el color azul y sin eliminar el fondo obtenemos el mejor resultado.

Como conclusión final se ha de indicar la amplia línea de trabajo que se abre con este proyecto de cara a trabajos más especializados en los diferentes puntos tratados, como se puede observar en el siguiente apartado.

7.2 Líneas futuras

Dentro de las líneas futuras que se plantean con la realización de este trabajo, cabe destacar las siguientes:

- Utilizar otros clasificadores para saber con cuál de ellos obtenemos mayores tasas de acierto.
- Realizar nuevas combinaciones de colores para determinar cuál de ellas es con la que se obtienen mejores resultados.
- Estudiar y fabricar una nueva pieza que permita utilizar al teléfono móvil, de tal manera que obtengamos imágenes de la misma calidad que al usar la Web-cam y así convertir el prototipo en dispositivo portátil.
- Realizar una base de datos que contemple la presencia de un espectro más amplio de contaminantes y una gama más amplia de concentraciones.

Bibliografía.

[1]"Introducción a los métodos instrumentales de análisis". Web. Universidad de Cádiz. <<http://ocw.uca.es/mod/resource/view.php?id=271>>. Fecha de consulta: 27 de julio de 2015.

[2]"Public Lab". Web. <<http://publiclab.org/wiki/spectrometer>>. Fecha de consulta: 27 de julio de 2015.

[3]"Definición de espectrometría". Experiencias de laboratorio. 2010. Web. <<http://www.frlp.utn.edu.ar/materias/qcasis/mostracion2.html>>. Fecha de consulta: 27 de julio de 2015.

[4]"Métodos ópticos analíticos"Olsen, Eugene D. Métodos Ópticos De Análisis. [Ed. en español]. ed. Barcelona: (es): Reverté, 1990. Web. Fecha de consulta: 27 de julio de 2015.

[5]"Métodos instrumentales"H. Willard, Hobart. Métodos Instrumentales De Análisis. Ed. Compañía Editorial Continental. 7ª ed. México: Print. Fecha de consulta: 27 de julio de 2015.

[6]"Espectroscopía de fluorescencia Atómica."Web. <https://es.wikipedia.org/wiki/Espectroscopia_de_fluorescencia>. Fecha de consulta: 27 de julio de 2015.

[7]"Espectrometría de fluorescencia."Web. <http://hiq.aga.cl/International/Web/LG/CL/likelgspgcl.nsf/docbyalias/anal_fluor_more>. Fecha de consulta: 27 de julio de 2015.

[8]"Espectrometría de fluorescencia."Web. <http://hiq.aga.cl/International/Web/LG/CL/likelgspgcl.nsf/docbyalias/anal_fluor_more>. Fecha de consulta: 27 de julio de 2015.

[9]"Parámetros de calidad" Moros García, Pablo José. "Parámetros De Calidad De Un Método Analítico." *Técnicas Analíticas*. Ed. Escuela de Negocio EOI.7. Print. Fecha de consulta: 27 de julio de 2015.

[10] Rosero, Edison, et al. "Detección De Nitrógeno En Un Humedal Artificial De Flujo Subsuperficial Utilizando Métodos Espectroscópicos De Fluorescencia Y Reflectancia." : 4. Print. Fecha de consulta: 27 de julio de 2015.

[11] Zsolnay, Adam, et al. "Determinación de materia orgánica disuelta." *ScienceDirect.Web*.
<<http://www.sciencedirect.com/science/article/pii/S0045653598001660#>>. Fecha de consulta: 27 de julio de 2015.

[12] "Descubrimiento de exoplanetas." *Web*.
<<http://www.astronoo.com/en/exoplanets.html>>. Fecha de consulta: 8 de agosto de 2015.

[13] "Puntero láser." *Web*. <http://www.lasers-pointers.com/es/punteros-laser-violeta/7-puntero-laser-violeta-100mw-405nm.html> Fecha de consulta: 27 de julio de 2015.

[14] "Ley de Lambert-Beer." *Web*. http://www.optek.com/es/Lambert_Beer_Law.asp Fecha de consulta: 27 de julio de 2015.

[15] "Instrumentación." *Web* <http://datateca.unad.edu.co/contenidos/401539/exe-2%20de%20agosto/leccin_5_instrumentacin_de_la_espectrofotometra_uvvis.html>. Fecha de consulta: 27 de julio de 2015.

[16] "Instrumentación." *Web* <<http://bvs.per.paho.org/texcom/cd048437/electroe.pdf>>. Fecha de consulta: 27 de julio de 2015.

[17] "Programa 3ds Max." *Web*. < <http://www.autodesk.es/products/3ds-max/overview>>. Fecha de consulta: 14 de noviembre de 2015.

[18] "Programa Netfabb." *Web*. <<http://www.netfabb.com/>>. Fecha de consulta: 14 de noviembre de 2015.

[19] "Programa Slicer." *Web*. < <http://www.slicer.org/>>. Fecha de consulta: 14 de noviembre de 2015.

[20]"Programa Cura".Web.<<https://ultimaker.com/en/products/cura-software>>. Fecha de consulta: 14 de noviembre de 2015.

[21]"Data sheet cámara".< <http://www.microsoft.com/hardware/es-es/d/lifecam-studio>> .Fecha de consulta: 20 de noviembre de 2015.

[22]"Recogida de muestras".Web<<http://publiclab.org/wiki/oil-testing-kit>>. Fecha de consulta: 27 de julio de 2015.

[23]"Nomenclatura de aceite de automóvil". Web<https://es.wikipedia.org/wiki/Aceite_de_motor> Fecha de consulta: 01 de noviembre de 2015.

[24]"Nomenclatura de aceite de automóvil". Web<http://www.liquimoly.com/liquimoly/web.nsf/id/pa_es_domb8bakrw.html> Fecha de consulta: 01 de noviembre de 2015.

[25] N.Selvarasu, Alamelu Nachaippan, N.M.Nandhitha, *"Image Processing Techniques and Neural Networks for Automated Cancer Analysis from Breast Thermographs-A Review"*, pp 134-135, 2012.

[26] Vijay Kumar, Priyanka Gupta, *"Importance of Statistical Measures in Digital Image Processing"*, *International Journal of Emerging Technology and Advanced Engineering*, Volume 2, Issue 8, 2012.

[27] Fazal Malik, Baharum Baharudin, *"The Statistical Quantized Histogram Texture Features Analysis for Image Retrieval Based on Median and Laplacian Filters in the DCT Domain"*, *The International Arab Journal of Information Technology*, Vol. 10, No. 6, 2013

[28] E. D. Selepci, O. G. Dulu "Image processing and data analysis in computed tomography", 2006.

[29] Robert A. Schowengerdt, *"Remote Sensing, Models, and Methods for Image Processing"*, pp 118, 1997.

[30] María Belén López Puetate, “Caracterización y reconocimiento de pausas respiratorias durante el sueño en tratamiento con CPAP”, Cap.4 Extracción de Características y Clasificación. Escuela de Ingeniería de Telecomunicación y Electrónica de la Universidad de Las Palmas de Gran Canaria, 2015.

[31] M Riedl, A Müller, and N Wessel. *Practical considerations of permutation entropy*. *The European Physical Journal Special Topics*, 222(2):249_262, 2013. Fecha de consulta: 24 de noviembre de 2016.

[32] Gerardo Colmenares, Capítulo IV, Inteligencia Artificial Máquinas de Vectores de Soporte. Web<www.webdelprofesor.ula.ve/economia/gcolmen/.../maquinas_vectores_soporte.pdf>. Fecha de consulta: 24 de noviembre de 2016.

[33] Thomas Stibor. *Malware Detection with Machine Learning Methods*. Fakultät für Informatik. Technische Universität München. Web <<https://www.sec.in.tum.de/assets/lehre/ws0910/ml/slideslecture9.pdf>>. Fecha de consulta: 24 de noviembre de 2016.

Anexo 1: Lista de acrónimos

- TFM. Trabajo Fin de Master
- AFS. *Atomic Fluorescence Spectroscopy*/ Espectroscopía de Fluorescencia Atómica
- UV-V. Ultravioleta Visible
- M. Molar
- E_a . Error absoluto
- E_r . Error relativo
- μ . Media de las medidas
- x . Valor verdadero
- CV. Coeficiente de variación
- s . Desviación normal
- S . Sensibilidad
- Δx . Incremento de la señal medida
- ΔC . Incremento de la concentración del analito
- LD. Límite de detección
- μ_B . Señal analítica generada por el blanco
- s_B . Desviación estándar de la señal del blanco
- LC. Límite de cuantificación
- a . Cte de proporcionalidad en $l/(cm \cdot g)$ ó $l/(cm \cdot mol)$
- b . Espesor del material en cm.
- c . Concentración de la muestra en en g/l.
- SAE. Society of Automotive *Engineers*/ Sociedad de Ingenieros del Automóvil
- ATF. *Automatic Transmission Fluids*/ Fluidos de Transmisión Automática
- R. *Red*/Rojo
- G. *Green*/Verde
- B. *Blue*/Azul
- Y. Escala de grises
- η . Resultados medios
- σ . Desviación típica