



Universidad Nacional de Educación a Distancia
y
Universidad Complutense de Madrid
Master en Ingeniería de Sistemas y Control

**Detección de moho en queso mediante
procesamiento de imágenes**

Alumno: David Molinero González

Director: Gonzalo Pajares Martinsanz

Curso 2013-2014
Convocatoria Junio 2014

Máster en Ingeniería de Sistemas y Control

Detección de moho en queso mediante procesamiento de imágenes

*Proyecto fin de máster de tipo B
proyecto específico propuesto por el alumno.*

Alumno: David Molinero González

Director: Gonzalo Pajares Martinsanz



Autorización

Autorizamos a la Universidad Complutense y a la UNED a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a sus autores, tanto la memoria de este Trabajo Fin de Máster, como el código, la documentación y/o el prototipo desarrollado.

Firmado:

A handwritten signature in black ink, which appears to read 'D. Molinero', is enclosed within a large, hand-drawn oval shape.

David Molinero González

Firma del alumno

RESUMEN DEL PROYECTO

El presente proyecto que en esta memoria se expone, muestra el cómo mediante la aplicación de técnicas de segmentación de imágenes y técnicas de aprendizaje supervisado, es posible detectar la presencia de moho en los palés de queso de pasta prensada de forma desasistida. La solución, nace con el objeto de dar remedio al problema de presencia de grandes niveles de moho en los quesos almacenados en las instalaciones de uno de los fabricantes de queso de pasta prensada de España. El contexto donde se desarrolla el proyecto es una instalación totalmente automatizada, donde se aplican las tecnologías industriales logísticas más punteras, Automatic Guided Vehicles (AGVs), volteadores automáticos, robots y almacenes inteligentes para la última fase de producción de queso.

Es en una de estas estaciones de volteo automática, donde mediante la instalación de cuatro cámaras, ocho puntos de iluminación externa y un sistema de fotocélulas, se le saca una foto a cada una de las caras de palé de quesos, éstas son recogidas por la aplicación servidor, que se encargará de segmentar la imagen, clasificando cada uno de los píxeles de las imágenes recibidas según su color, aplicando técnicas de Visión por Computador, obteniendo así un porcentaje de píxeles de cada imagen valorados como moho o no según proceda, calculando finalmente un indicador de moho del palé de queso en el momento de la operación. Es en este punto donde se investigará sobre las opciones de segmentación de imágenes más apropiadas para resolver el problema propuesto, buscando en la literatura las distintas técnicas existentes y profundizando en dos de ellas, umbralización mediante método de Otsu y agrupamiento borroso mediante Fuzzy K-means, con el objetivo de implantarlas en la problemática que se nos presenta.

Tras el desarrollo y puesta en productivo, el siguiente paso fue analizar los resultados.

Los resultados obtenidos tras la aplicación de los métodos que se proponen, son positivos, siendo posible detectar el nivel de moho de un palé de forma desasistida (automática) mediante técnicas de segmentación de imágenes y Visión por Computador. En cuanto a futuras líneas de trabajo destacan dos principalmente, el desarrollo de técnicas de segmentación más complejas como redes neuronales, que requieren una mayor potencia de computación, y técnicas de segmentación para la mejora de la imagen a procesar eliminando las cajas donde se ubican los quesos.

PALABRAS CLAVE

Instalación automatizada, detección automática, moho, palés de queso, procesamiento de imágenes, Visión por Computador, aprendizaje supervisado, iluminación.

ABSTRACT

This project shows how by applying several techniques based on image segmentation and supervised learning; it is possible to detect the presence of mould on pallets with compressed paste of cheese in an unassisted situation. The aim of this work consists of reducing the presence of high levels of mould in the industrial plant of one of manufacturers of Spain making compressed paste of cheese. The context where the project is developed is an installation that is fully automated, where the most advanced logistic industrial approaches are used to the last stage of cheese production, Automatic Guided Vehicles (AGVs), automatic turners, robots and several smart storage.

In one of these automatic dump stations, where by installing four cameras, eight points of external lighting and a system of photocells, a picture is captured for each of the faces of the pallet of cheese, these are collected by server application, which is responsible for segmenting the image by classifying each of the pixels of the images received by color, in different classes detected during the unsupervised learning cameras applied to the images coming from the camera. Each class has a monitored indication which shows if it is mould or not, therefore a percentage of pixels per image identify the quantity of mould on the cheese of the pallet during of the operation. It is point of process where we will investigate all options for image processing and segmentation and also the applying image computer vision- based techniques.

After installing and starting the project, the next stage was to analyze the results.

The results presented here are positive, it is possible to detect the level of mould of a pallet with cheese in an unattended situation (automatic) using image segmentation approaches and supervised learning. In the future two main lines are to be developed, the first is to apply more complex segmentation approaches, such as neural networks, which require more computing power, and the second the application of approaches for interference suppression in images, improving their quality.

KEYWORDS

Automated installation, automatic detection, mould, pallets of cheese, image processing, computer vision, supervised learning, lighting.

ÍNDICE

1	INTRODUCCIÓN	8
1.1	<i>IDENTIFICACION DEL PROBLEMA</i>	8
1.2	<i>CONTENIDO DE LA MEMORIA</i>	9
1.3	<i>OBJETIVOS</i>	9
1.4	<i>METODOLOGÍA</i>	10
2	ESTADO DEL ARTE	11
2.1	<i>ANTECEDENTES</i>	11
2.2	<i>PÍXEL Y EL COLOR RGB</i>	13
2.3	<i>SEGMENTACION</i>	14
2.4	<i>UMBRALIZACION O THRESHOLDING</i>	15
2.5	<i>AGRUPAMIENTO O CLUSTERING</i>	16
3	REVISION DE METODOS UTILIZADOS	20
3.1	<i>INTRODUCCIÓN</i>	20
3.2	<i>MÉTODO OTSU</i>	20
3.3	<i>FUZZY K-MEANS</i>	22
4	PROBLEMÁTICA Y DISEÑO	25
4.1	<i>INTRODUCCIÓN</i>	25
4.2	<i>ANTECEDENTES</i>	26
4.3	<i>DISEÑO DEL SISTEMA</i>	27
5	ANALISIS DE RESULTADOS	30
5.1	<i>INTRODUCCIÓN</i>	30
5.2	<i>DESCRIPCION DE LA BASE DE CONOCIMIENTO UTILIZADA</i>	30
5.3	<i>METODO UTILIZADO PARA LA APLICACIÓN</i>	31
5.4	<i>FASE DE APRENDIZAJE</i>	31
5.5	<i>CLASIFICACIÓN</i>	41
6	CONCLUSIONES Y TRABAJO FUTURO	53
6.1	<i>INTRODUCCIÓN</i>	53
6.2	<i>CONCLUSIONES</i>	53
6.3	<i>TRABAJO FUTURO</i>	54
7	BIBLIOGRAFÍA	55
8	GLOSARIO	56
8.1	<i>INTRODUCCIÓN</i>	56
8.2	<i>DEFINICIONES</i>	56
	ANEXO I	59
	ANEXO II	61

LISTA DE FIGURAS Y TABLAS

1-1 Situación de los palés en cámaras.....	8
2-1 Proceso de análisis de imágenes.....	11
2-2 Esquema máquina de aprendizaje.....	12
2-3 Imagen descompuesta en sus tres componentes RGB.....	13
2-4 Histograma con un umbral o con dos umbrales	15
2-5 Agrupamiento de nube de puntos en tres clases	17
2-6 Agrupamiento de nube de puntos en dos clusters	18
3-1 Histograma imagen polar.....	20
3-2 Ejemplo histograma imagen palé con los umbrales resultado de aplicar dos veces el método Otsu	22
3-3 Agrupamiento de nube de puntos en 4 clusters	23
4-1 Esquema del proceso de aprendizaje-clasificación.....	25
4-2 Diagrama de despliegue de los distintos subsistemas que intervienen	28
4-3 Esquema de instalación cámaras, pantallas y fotocélulas en volteo simple	28
4-4 Ejemplos de fotografía de un palé obtenidos por el sistema	29
5-1 Ejemplo cabecera y tupla base de datos de conocimiento.....	31
5-2 Clasificación de imagen mediante metodo Fuzzy K-means, con 10 clusters en 10 iteraciones.	33
5-3 Resultado de imagen sin moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 2 clusters.....	33
5-4 Resultado de imagen con nivel medio de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 2 clusters	34
5-5 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 2 clusters	34
5-6 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 4 clusters	35
5-7 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 4 clusters	36
5-8 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 4 clusters	36
5-9 Resultado de imagen sin moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 8 clusters.....	37
5-10 Resultado de imagen con nivel medio de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 4 clusters	38
5-11 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 10 clusters	38
5-12 Resultado de imagen sin moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 8 clusters	39
5-13 Resultado de imagen con nivel medio de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 4 clusters	40
5-14 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 10 clusters	40

5-15	Tabla de resultados de la aplicación de los dos métodos Fuzzy K-means y umbralización Otsu a las 100 imágenes de los 25 palés de estudio, recogidos en la BC	41
5-16	Resultados del método Fuzzy k-means y método Otsu cámara 2, palé 16	42
5-17	Resultados del método Fuzzy k-means y método Otsu cámara 3, palé 3	42
5-18	Valores del indicador de moho para cada una de las 25 imágenes de la Cámara 1.	43
5-19	Resultados del método Fuzzy k-means y método Otsu con un problema de iluminación.....	44
5-20	Resultados del método Fuzzy k-means y método Otsu	44
5-21	Resultados del método Fuzzy k-means y método Otsu con un palé con cajas de color negro.....	45
5-22	Valores de indicador de moho para cada una de las 25 imágenes de la Cámara 2.	46
5-23	Resultados del método Fuzzy k-means y método Otsu	46
5-24	Resultados del método Fuzzy k-means y método Otsu	47
5-25	Valores de indicador de moho para cada una de las 25 imágenes de la Cámara 3.	48
5-26	Resultados del método Fuzzy k-means y método Otsu en el palé número 2 con problema en la fase de aprendizaje	49
5-27	Resultados del método Fuzzy k-means y método Otsu en el palé número 6, con problema en fase de aprendizaje.	49
5-28	Resultados del método Fuzzy k-means y método Otsu en el palé número 13, con alto valor de moho.	50
5-29	Valores de indicador de moho para cada una de las 25 imágenes de la Cámara 4.	50
5-30	Resultados del método Fuzzy k-means y método Otsu en el palé número 4.	51
5-31	Resultados del método Fuzzy k-means y método Otsu en el palé número 20. ..	51
5-32	Valores de indicador de moho para cada una de las 25 imágenes, promediado de las imágenes procedentes de las 4 cámaras.	52
AI-1	Base de conocimiento o“ground truth”	60
All-1	Instalación con un punto de luz y brazo extensible”	61
All-2	Instalación con dos puntos de luz sobre el vallado exterior	61
All-3	Instalación con un punto de luz	62
All-4	Instalación con un punto de luz fluorescente sobre el vallado exterior.....	62
All-5	Instalación final con dos puntos de luz fluorescente sobre el vallado exterior.....	63
All-6	Instalación fotocélulas soporte metálico	63

1 INTRODUCCIÓN

1.1 IDENTIFICACION DEL PROBLEMA

El proyecto “Detección de moho en queso mediante procesamiento de imágenes” que aquí se expone, muestra cómo mediante la aplicación de técnicas de procesamiento de imágenes es posible detectar la presencia de moho en los palés de queso de pasta prensada de forma desasistida, esto es automática.

La solución, nace con el objeto de dar remedio al problema real de presencia de grandes niveles de moho en las instalaciones de una de los principales fabricantes de queso de pasta prensada de España.

El moho es un hongo que se encuentra tanto al aire libre como en lugares húmedos y con baja luminosidad, condiciones en las que se almacenan los quesos para realizar su proceso de curación. La aparición de moho es característico de ciertas variedades de queso muy conocidas, como pueden ser el queso azul, el brie, roquefort, etc., si bien en la modalidad de pasta prensada o queso manchego, la abundancia de moho y sobre todo la prolongación de éste en la corteza del producto, puede originar que se introduzca en su interior, variando sus propiedades organolépticas.

El contexto donde se desarrolla el proyecto es una instalación totalmente automatizada, donde se aplican las tecnologías industriales más punteras como, Automatic Guided Vehicles (AGVs), volteadores automáticos, robots y almacenes inteligentes para la última fase de producción de queso. Los quesos, durante su proceso de maduración se almacenan en las distintas cámaras que componen la instalación, a una temperatura y humedad óptimas para que el proceso de maduración se desarrolle correctamente. Del almacenamiento de los palés de queso en dichas cámaras, se encargan una serie de AGVs que agrupan el producto por tipo y fecha de maduración en filas, compuestas por ocho palés de fondo y tres palés de altura. La separación entre filas es de 30 cm, lo que impide la inspección por parte de personal del estado de los palés y con ello revisar los niveles de moho que éstos tienen. La figura 1.1 muestra la disposición y distribución de los palés dentro de la cámara, así como los quesos en el interior de los mismos.



1-1 Situación de los palés en cámaras

Durante el periodo en que los palés de queso se encuentran almacenados en dichas cámaras, es necesario aplicar, dependiendo del tipo de producto y maduración, volteos periódicamente, con el objeto de que los quesos cambien de posición,

favoreciendo el correcto secado e impidiendo que éstos se deformen. El volteo es una operación que consiste en rotar 180° sobre la vertical el palé de quesos.

Es en una de estas estaciones de volteo automática, ya que son las propias AGVs las que trasladan el queso, donde mediante la instalación de cuatro cámaras, se capturará una imagen de cada una de las caras del palé de quesos. Mediante la aplicación de segmentación de estas imágenes, junto con técnicas de clasificación se determina si el palé tiene moho o no y en qué proporción.

Por lo tanto, el objetivo principal del proyecto que aquí se presenta es detectar la presencia de moho en una instalación real, para ello, se estudiarán las distintas técnicas de segmentación y clasificación de imágenes con el fin de seleccionar de entre todas ellas dos, con el objetivo de implantarlas en la problemática que se nos presenta.

1.2 CONTENIDO DE LA MEMORIA

La memoria se organiza en los siguientes capítulos como sigue:

Capítulo uno: en lo que resta del presente capítulo se exponen los objetivos que se plantearon en la investigación aquí recogida, en base a los planteamientos previos y que tratan de dar solución a la problemática existente. También se realiza una revisión de la metodología empleada durante el desarrollo de la investigación.

Capítulo dos: se realiza la síntesis relativa al estudio bibliográfico llevado a cabo sobre revisión de metodologías relacionadas con la segmentación de imágenes, realizándose una explicación de los métodos y técnicas elegidas.

Capítulo tres: se describe con detalle los dos métodos implantados y la solución que se ha adoptado.

Capítulo cuatro: se expone el diseño del sistema desarrollado en relación al problema planteado.

Capítulo cinco: se diseña una estrategia de pruebas y se analizan los resultados obtenidos con los mismos sobre el conjunto de imágenes disponibles y los métodos aplicados.

Capítulo seis: se exponen las conclusiones y líneas de investigación y trabajo futuras.

1.3 OBJETIVOS

El proyecto persigue varios objetivos, que se sintetizan a continuación:

1. Aprender a manejar referencias bibliográficas, así como la forma de abordar las investigaciones.
2. Identificar los métodos de segmentación de imágenes existentes en la literatura.
3. Determinar los métodos que nos puedan ser más útiles en la resolución del problema planteado.
4. Aplicación de los métodos al problema real presentado. Será el objetivo principal, desarrollar un sistema que mediante el procesamiento de las imágenes de los palés consiga determinar el nivel de moho existente en los quesos contenidos en el palé.
5. Analizar las ventajas e inconvenientes de cada uno de los métodos y técnicas aplicadas y determinar su validez. Se estudiará la validez del estudio, mediante el análisis de los resultados obtenidos.

6. Realizar un aporte de carácter investigador con los resultados obtenidos y las conclusiones finales.

1.4 METODOLOGÍA

A continuación se exponen de forma cronológica los pasos seguidos para llevar a cabo la investigación y el desarrollo de la propuesta, junto con la actividad desarrollada en cada uno de ellos.

En primer lugar se realiza una revisión bibliográfica sobre los métodos de segmentación de imágenes y clasificación existentes en la literatura.

En base a lo anterior, el siguiente paso consiste en la identificación de aquellos métodos simples existentes en la literatura con resultados satisfactorios, para su análisis y el diseño/implementación de las estrategias encaminadas hacia la mejora de resultados.

Selección de los métodos más prometedores que puedan dar solución a nuestra problemática, en este caso utilizaremos uno basado en color y otro en escala de grises. El objetivo consiste en conseguir una automatización máxima de forma que una vez realizada la calibración del algoritmo en la fase de aprendizaje, no haya intervención por parte del usuario.

Se desarrolla el sistema y se implantará en el entorno real cada uno los métodos, con el objeto de obtener resultado que confirmen su validez.

Por último, se estudian los resultados obtenidos, se extraen conclusiones y se exponen las líneas de trabajo futuras.

2 ESTADO DEL ARTE

2.1 ANTECEDENTES.

Una de las áreas de la ciencia de la computación que en las últimas dos décadas ha ido ganando terreno es, sin duda, la Visión por Computador, más en concreto el procesamiento de imágenes digitales y las distintas aplicaciones derivadas, que terminan convergiendo hacia técnicas más complejas de Visión por Computador.

Obviamente esto sucede debido a que cada día la información visual es más importante y abundante debido a la existencia de sensores más potentes y de mayores prestaciones, como ocurre con las cámaras digitales, y es por eso que se necesitan técnicas que manipulen adecuadamente esa información. Es más, gran parte de la información que procesamos los humanos está en forma de imágenes y de alguna manera, nosotros realizamos procesamiento de imágenes sin darnos cuenta, por ejemplo, el simple hecho de convertir dos imágenes registradas por nuestros ojos al cerebro en una representación tridimensional nos da la posibilidad de comprender la profundidad de los objetos.

Dentro del campo del procesamiento de imágenes digitales una de las etapas del proceso es el análisis. Dicho análisis de imágenes, emplea técnicas que extraen información de las imágenes permitiendo identificar y clasificar objetos. La figura 2-1 muestra un esquema simple del proceso de análisis de imágenes, desde su adquisición por medio de las cámaras disponibles, hasta la extracción de las características, que identifican los objetos, previa clasificación de su contenido



2-1 Proceso de análisis de imágenes

En el análisis de las imágenes, dentro del contexto que nos ocupa, es donde surge la necesidad del reconocimiento de patrones, que es la disciplina científica cuyo objetivo es la clasificación de objetos en un cierto número de categorías o clases. La aplicación de esta disciplina no se centra solo en imágenes, como es en el caso de estudio, sino que pueden ser de cualquier otro tipo, tales como formas de ondas de señales o medidas de variables y en general cualquier conjunto de datos que necesitan ser clasificadas. Nos referiremos a esos objetos de forma genérica utilizando el término de patrones (Pajares y Cruz, 2008).

Históricamente los dos enfoques en el reconocimiento de patrones han sido el estadístico (o teoría de la decisión) y el sintáctico (o estructural). Recientemente, el desarrollo de las redes neuronales ha proporcionado un nuevo enfoque. El aumento de la capacidad de almacenamiento ofrecida por los más modernos sistemas ha propiciado el reconocimiento basado en la apariencia. Tanto el enfoque estadístico como el basado en redes neuronales, utilizan patrones de los que se extraen de ellos propiedades de naturaleza cuantitativa, mientras que el enfoque sintáctico se fundamenta en las relaciones geométricas asociadas a la forma de los objetos y el enfoque basado en la apariencia considera distintas formas de vista de los mismos.

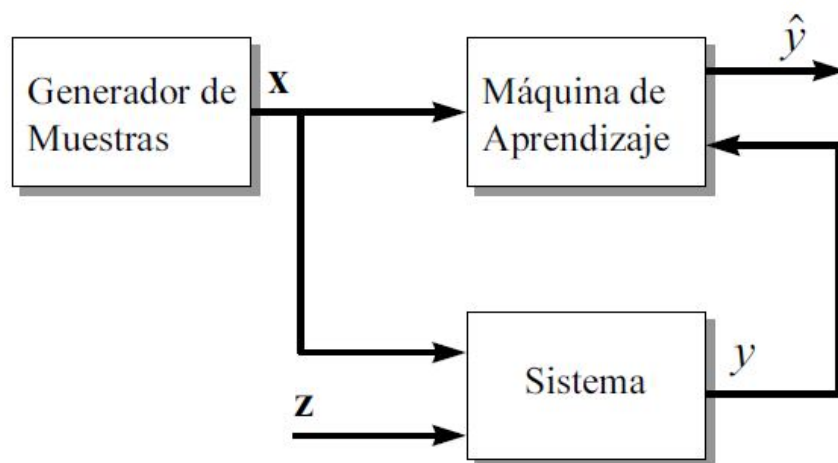
En nuestro caso, el método es el de reconocimiento de patrones, mediante la estimación agrupación y clasificación, de forma que a partir de unas muestras de entrenamiento, previamente seleccionadas se pueden obtener los parámetros de aprendizaje, que luego se utilizan para clasificación, denominado máquina de aprendizaje. Veamos con más detalle los papeles de cada componente de la máquina de aprendizaje en términos matemáticos:

Generador, el generador o distribución de muestras produce vectores aleatorios $x \in \mathfrak{R}^d$, estos vectores determinan una función de densidad de probabilidad $p(x)$, que es desconocida. En terminología estadística esta situación se llama observacional.

Sistema, el objetivo del aprendizaje consiste en estimar las dependencias desconocidas entre las variables de entrada x y de salida y a partir de un conjunto de observaciones pasadas de valores (x,y) . En la figura 2.2 el conjunto de variables etiquetadas como z indican el resto de factores que afectan a las salidas pero que no son observados o controlados. Por ejemplo, en el proceso de captura de imágenes, la presencia de suciedad o imperfecciones en el sistema óptico pueden ser cantidades de este tipo. Por tanto, el conocimiento de los valores observados (x) no especifica de forma unívoca las salidas (y). Esta incertidumbre en las salidas refleja la laguna de conocimiento de los factores no observados (z), de aquí se deriva una dependencia estadística entre las entradas observadas y la salida o salidas. El efecto de las entradas no observadas puede caracterizarse por una distribución de probabilidad condicional $p(y/x)$, que expresa la probabilidad de que ocurra y dada la entrada x . En definitiva, el sistema produce un valor de salida y para cada vector de entrada x de acuerdo a la densidad condicional $p(y/x)$, que es también desconocida. Obsérvese que esta descripción incluye el caso específico de sistemas deterministas donde $y = f(x)$.

Máquina de aprendizaje, en el caso más general, es capaz de implementar un conjunto de funciones $f(x, w)$, $w \in \Omega$ donde Ω es un conjunto de parámetros abstractos utilizados solamente para indexar el conjunto de funciones. En esta formulación el conjunto de funciones implementadas por la máquina de aprendizaje puede ser cualquier conjunto de funciones, elegidos previamente, antes de que el proceso de aprendizaje haya comenzado.

La figura 2-2 muestra un esquema simple de máquina de aprendizaje, donde aparecen las muestras que se suministran al sistema para la obtención de los parámetros y su posterior clasificación.



2-2 Esquema máquina de aprendizaje

Podemos distinguir dos tipos de aprendizaje, conocidos como aprendizaje supervisado y no supervisado.

Aprendizaje supervisado, se utiliza para estimar una relación desconocida (entrada, salida) a partir de muestras conocidas (entrada, salida). La clasificación y la regresión caen en este grupo. El término supervisado se corresponde con el hecho de que los valores de salida para las muestras de entrenamiento son conocidos y por tanto, proporcionados por un supervisor o por un sistema que está siendo modelado.

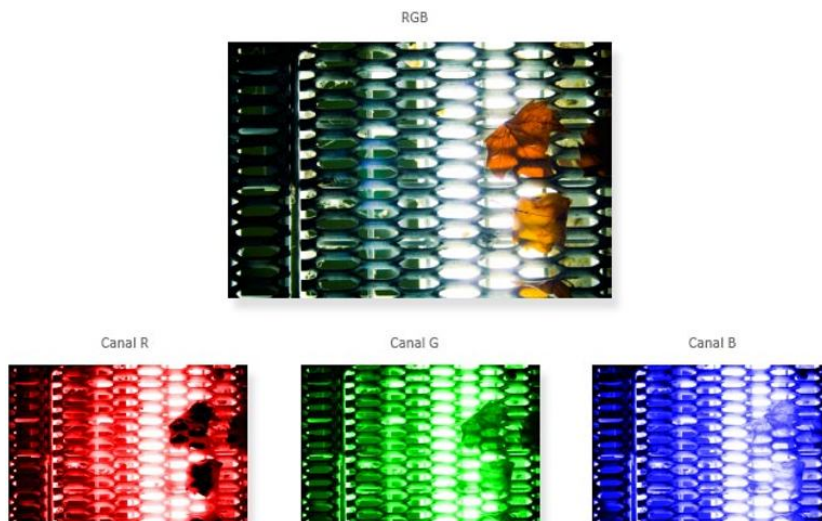
Aprendizaje no supervisado, solamente se proporcionan al sistema de aprendizaje las muestras de entrada y no existe noción alguna de la salida durante el aprendizaje. El objetivo del aprendizaje no supervisado puede ser estimar la distribución de probabilidad de las entradas o descubrir la estructura natural (agrupamientos o *clusters* en inglés) en los datos de entrada.

Finalmente, conviene advertir que la distinción entre aprendizaje supervisado y no supervisado se sitúa sólo a nivel de la declaración del problema. Esto no implica que los métodos desarrollados originalmente para aprendizaje supervisado no puedan usarse con algunas modificaciones para tareas de aprendizaje no supervisado y viceversa. (Pajares y Cruz 2008)

2.2 PÍXEL Y EL COLOR RGB

Entendemos por píxel la superficie homogénea más pequeña de las que componen una imagen digitalizada, que se define por su brillo y color (DRAE, 2001).

En una imagen en una pantalla, ya sea un monitor de ordenador, una televisión o una imagen proyectada, cada píxel de color posee tres componentes: rojo, verde y azul. Estos son los colores primarios de la luz y cualquier color que podamos percibir se puede formar a partir de ellos. Dado que el modo más general que tiene un ordenador de mostrar imágenes es a través de su pantalla, éste es el método más habitual que se suele emplear para almacenar las imágenes, esto es: sus componentes rojo, verde y azul, RGB por sus iniciales en inglés, *red-green-blue*. La Figura 2-3 muestra las respectivas componentes roja, verde y azul de una imagen obtenidas a partir de la imagen de original, si bien las componentes por separado son en gris, aquí se muestran en color por ser más ilustrativa la representación en relación al concepto que se quiere expresar.



2-3 Imagen descompuesta en sus tres componentes RGB

Bien es cierto que este método de representar el color no es el único. Sin ir más lejos, a la hora de imprimir una imagen los colores primarios son otros: cian, amarillo y magenta (CYM en inglés). El ordenador debe realizar una conversión de uno a otro formato para que la impresora pueda representar fielmente la imagen. Sin embargo, así como la pantalla podía mostrar cualquier color “por el mismo precio” si queremos imprimir algo de color negro usando este método la impresora consumirá tinta de tres colores. Como la tinta negra es habitualmente mucho más barata que las de color, las impresoras normalmente tienen cuatro cartuchos: cian, amarillo, magenta y negro (cuatricromía conocida habitualmente como CYMK).

Además de los modelos de color mencionados existen otros varios. Sin necesidad de entrar en detalles, en el caso de imágenes en vídeo se utilizan espacios de color como YUV (en el sistema de televisión PAL), YIQ (en NTSC) y YCbCr (similar a YUV). Todos ellos comparten una división en tres canales, uno de luminancia y dos de información de color. En el caso de YUV y YIQ el canal de luminancia representa la imagen original en escala de grises conservando los detalles de la misma; en el caso de YCbCr, más orientado al vídeo digital, los canales se encuentran “comprimidos” para poder añadir datos adicionales en el flujo de vídeo. Otros espacios de color ampliamente utilizados son HSV (tono, saturación y valor) y HSI (tono, saturación e intensidad), especialmente presentes en las interfaces de usuario de programas de edición de imagen, por la facilidad que presentan para escoger un color en concreto cambiando su tono o su saturación. En el caso de HSI el canal I correspondiente a la intensidad de una copia en escala de grises de la imagen, similar al canal Y en YUV y YIQ, donde la información de color se contiene en las componentes H y S.

2.3 SEGMENTACION

Duda y col. (2001) definen la segmentación como “la extracción de un objeto de interés de una imagen arbitraria”. González y Woods (2002) la definen como “la subdivisión de una imagen en sus partes constituyentes”.

Nosotros lo definiremos en el contexto del campo de la visión artificial, como “el proceso de dividir una imagen digital en varias partes (grupos de píxeles) u objetos”, o de forma más precisa como “el proceso de asignación de una etiqueta a cada píxel de la imagen de forma que los píxeles que compartan la misma etiqueta tengan ciertas características visuales similares”, basándonos para ello en los siguientes principios:

Similitud: cada uno de los píxeles de un elemento tiene valores parecidos con respecto a alguna propiedad determinada.

Discontinuidad: los objetos y más concretamente las regiones que los diferencian destacan del entorno y tienen por tanto unos bordes definidos.

Conectividad: los píxeles pertenecientes al mismo objeto o región tienen que ser contiguos, es decir, tienden a agruparse constituyendo regiones homogéneas.

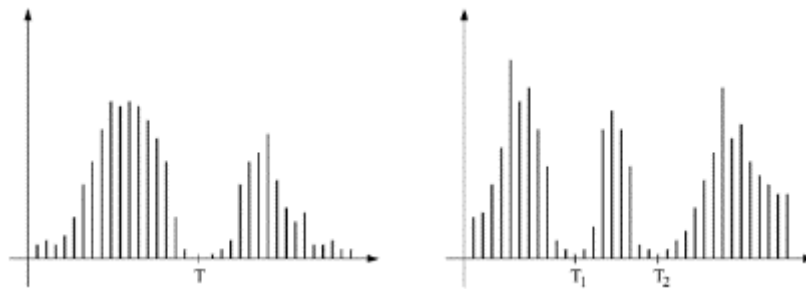
La segmentación se usa tanto para localizar objetos como para encontrar los límites de éstos dentro de una imagen y son varios los métodos que podemos encontrar a lo largo de la literatura, agrupándolos en: umbralización, crecimiento de regiones, división y unión de regiones (*split & merge*), minimización de una función de energía, agrupamiento o *clustering*, transformada Watershed, etiquetado de componentes conexas, estructuras de datos para la segmentación y caracterización de regiones.

Ante la gran cantidad de métodos que encontramos en la bibliografía, que dan solución a la segmentación, nos centraremos en dos familias, métodos de segmentación por umbralización y métodos de agrupamiento o *clustering*, debido a su sencillez computacional para realizar el estudio a fondo dentro del contexto de nuestro problema.

2.4 UMBRALIZACION O THRESHOLDING

La umbralización es una técnica de segmentación ampliamente utilizada en las aplicaciones industriales. Se emplea cuando hay una clara diferencia entre los objetos a extraer respecto del fondo de la escena. Los principios por los que se rige son la similitud entre los píxeles pertenecientes a un objeto y sus diferencias respecto al resto. Por tanto, la escena debe caracterizarse por un fondo uniforme y por objetos parecidos.

Se basa en el histograma de una imagen, permitiéndonos conocer la frecuencia relativa de aparición de cada uno de los posibles niveles de intensidad dentro de la imagen en cuestión. El histograma nos servirá para agrupar los píxeles en función de su nivel de intensidad luminosa. Debido a la propiedad en que se basa, destaca el hecho adverso relativo a los posibles efectos derivados de la iluminación sobre la imagen, que hacen que la segmentación de objetos, tales como de texturas no sea tan sencilla. En la umbralización por texturas, debemos tener en cuenta que dependiendo de la textura que nos encontremos, deberemos realizar un filtrado de media o de varianza previo a la umbralización. Si la textura posee un patrón de repetición, la mejor opción será recurrir al dominio de la frecuencia. La figura 2-4 muestra sendos histogramas, representando la distribución de niveles de intensidad en dos hipotéticas imágenes. En el primer caso se trata de un histograma bimodal, mientras que en el segundo es trimodal. El objetivo estriba en encontrar los umbrales T , en el primer caso y T_1 y T_2 en el segundo, de forma que se separen convenientemente los objetos representados por los diferentes modos.



2-4 Histograma con un umbral o con dos umbrales

Al aplicar un umbral T la imagen en escala de grises, $f(x,y)$, quedará binarizada; etiquetando con '1' los píxeles correspondientes al objeto y con '0' aquellos que corresponden al fondo. El umbral puede depender de $f(x,y)$, de alguna propiedad local del píxel $p(x,y)$, y hasta de su propia posición. Si el umbral sólo depende de $f(x,y)$ se dice que es un umbral global; en el caso de que además dependa de $p(x,y)$, por ejemplo, el valor medio de los píxeles vecinos, el umbral es denominado local; y si depende también de la posición (x,y) del píxel, se denominará dinámico.

Existen dos métodos básicos para abordar el problema de la umbralización global (Maravall, 1993):

- Basada en la búsqueda de mínimos.
- Basada en las técnicas de reconocimiento de patrones.

La umbralización basada en la búsqueda de mínimos o procedimientos no paramétricos, trata de encontrar niveles de intensidad que dividan el histograma de la imagen con el fin de encontrar el mejor valor para la separación de los objetos que se pretende. Por otro lado, la umbralización basada en las técnicas de reconocimiento de patrones o procedimientos paramétricos consiste, precisamente, en buscar patrones o formas en el histograma de la imagen.

La mayoría de las técnicas de umbralización se basan en estadísticas sobre el histograma unidimensional de una imagen.

En particular, el método de Otsu (1979), que será objeto de estudio más adelante, elige el umbral óptimo maximizando la varianza entre clases mediante una búsqueda exhaustiva.

Si bien existen diferentes métodos para hallar un umbral, la mayoría de ellos no dan buenos resultados cuando se trabaja con imágenes del mundo real debido a la presencia de ruido, histogramas planos o una iluminación inadecuada. Por el contrario, el método de Otsu, ampliamente utilizado en la literatura, ha sido uno de los mejores métodos de selección de umbral para imágenes del mundo real, como en el caso que nos ocupa, las propias relativas a la detección del moho. Sin embargo, como ya anticipamos, este método usa una búsqueda exhaustiva para evaluar el criterio de maximización de la varianza entre clases. A medida que el número de clases de una imagen aumenta, el método de Otsu necesita mucho más tiempo para seleccionar un umbral multinivel adecuado. La importancia del método de Otsu radica en que es automático, es decir, no necesita supervisión humana ni información previa de la imagen antes de su procesamiento. Por otra parte, se trata de un método que proporciona un único umbral, de forma que en el supuesto de querer obtener más de un valor umbral, es necesario aplicarlo sucesivamente sobre cada una de las clases que identifican los diferentes modos del histograma.

2.5 AGRUPAMIENTO O CLUSTERING

El problema del agrupamiento es el de separar un conjunto de datos en un cierto número de grupos o *clusters* basándose en alguna medida de similitud. El objetivo es *encontrar* un conjunto de grupos para los cuales, las muestras dentro de un grupo son más similares entre sí que con las muestras de diferentes grupos. Obsérvese que la tarea del agrupamiento puede estar fuera del aprendizaje predictivo, puesto que el objetivo es agrupar los datos principalmente y no caracterizar los datos futuros. En el análisis de grupos la medida de similitud para el agrupamiento se elige subjetivamente basada en su capacidad para crear agrupaciones de interés. Los algoritmos de agrupamiento se pueden dividir en las siguientes categorías principales:

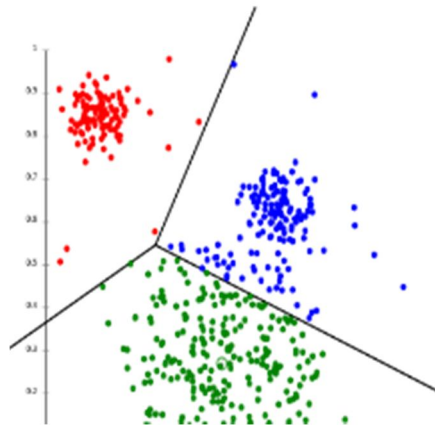
Secuenciales. Estos algoritmos producen un agrupamiento simple. En muchos de ellos los vectores de características se presentan al algoritmo una vez o un número limitado de veces. El resultado final depende generalmente del orden en el que se presentan dichos vectores al algoritmo.

Jerárquicos. Que a su vez pueden ser de aglomeración o de división. Los de aglomeración producen una secuencia de agrupamientos en un orden decreciente de grupos m en cada paso.

El agrupamiento producido en cada paso es consecuencia de la fusión de dos grupos del paso previo. Por el contrario, los de división producen un número de grupos creciente m en cada paso como consecuencia de la división de un simple grupo en dos.

Basados en la optimización de una función de coste. En estos algoritmos se evalúa el agrupamiento basándose en una función de coste J que es optimizada. Generalmente el número de grupos se mantiene constante. En esta categoría se incluyen lo que nosotros denominamos algoritmos puros, donde un vector dado pertenece exclusivamente a un determinado grupo. El algoritmo más famoso de esta categoría es el Isodata. En contraposición a los anteriores están los algoritmos basados en la lógica difusa (*fuzzy*) donde un vector pertenece a más de un grupo, con un determinado grado de pertenencia. También podemos destacar los algoritmos probabilistas, que siguen el criterio de clasificación Bayesiano, de forma que cada vector x se asigna al grupo cuya probabilidad a posteriori sea máxima. La figura 2-5

muestra un ejemplo ilustrativo de agrupamiento de clases, identificadas por colores con sus correspondientes líneas de separación entre clases o dicho de otro modo, funciones de decisión.



2-5 Agrupamiento de nube de puntos en tres clases

Existen otras categorías de algoritmos de agrupamiento tales como algoritmos genéticos, métodos de relajación estocástica, algoritmos basados en transformaciones morfológicas, basadas en grafos, etc. así como abundantes variantes de los mencionados anteriormente (Pajares y Cruz. 2008). La mayoría de ellos se pueden encontrar en Theodoridis y Koutroumbas (1999).

Al agrupamiento o *clustering* puede ser considerado el más importante problema de aprendizaje no supervisado. Estos métodos descomponen el conjunto de objetos en un conjunto de clases o *clusters* disjuntos, minimizando una función criterio que enfatiza la estructura local de los objetos, asignando *clusters* a máximos locales en la estructura global (Jiménez, 2008)

A continuación se describe brevemente uno de los métodos más utilizados en diferentes aplicaciones en imágenes reales para posteriormente centrarnos en una variante de éste, basada en lógica difusa.

K-MEDIAS O K-MEANS

El término "K-means" fue utilizado por primera vez por James MacQueen en 1967, (MacQueen, 1967) aunque la idea se remonta a Hugo Steinhaus en 1957 (Steinhaus, 1957). El algoritmo estándar fue propuesto por primera vez por Stuart Lloyd en 1957 como una técnica para modulación por impulsos codificados, aunque no se publicó fuera de los laboratorios Bell hasta 1982. (Lloyd, 1957). En 1965, Forgy publicó esencialmente el mismo método, por lo que a veces también se le conoce como Lloyd-Forgy (Forgy, 1965). Una versión más eficiente fue propuesta y publicada en Fortran por Hartigan y Wong en 1975/1979 (Hartigan, 1975).

Se trata de un método particional que intenta encontrar un número específico de grupos, los cuales están representados por centroides, aplicable a un grupo de objetos en un espacio n-dimensional.

Se denomina centroide representativo de un agrupamiento, al vector formado por las medias de cada una de las componentes de los elementos pertenecientes a dicho agrupamiento. La técnica general de k-medias es muy simple. A continuación se presenta la descripción del algoritmo básico (Tan y col, 2006)

1º Seleccionar K centroides, donde K es el número de *cluster* deseado.

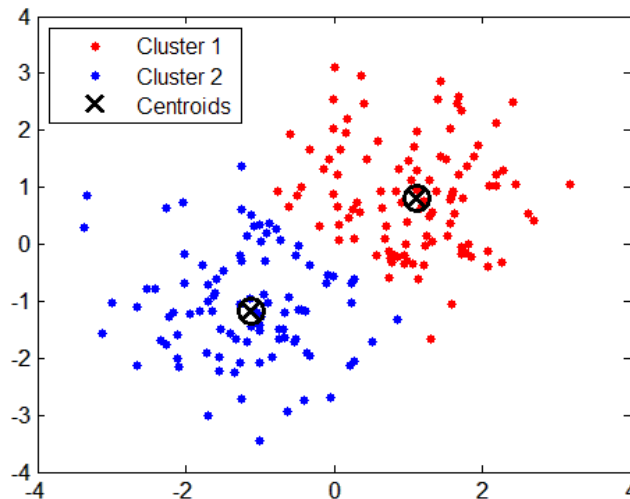
2º Asignar cada punto al centroide más cercano, donde cada colección de puntos asignados a un centroide constituye un *cluster* (región de Voronoi)

3º Actualizar los centroides de cada *cluster*, basados en los puntos asignados al *cluster*.

4º Repetir el proceso de asignación y actualización hasta que ningún punto cambie de *cluster*, o lo que es lo mismo, hasta que los centroides permanezcan iguales.

5º Fin

En la figura 2-6 se muestran dos clústeres formados por los diferentes puntos coloreados, agrupados alrededor de los correspondientes centroides que les representan, resultado de la ejecución del algoritmo descrito.



2-6 Agrupamiento de nube de puntos en dos clusters mediante algoritmo K-means

En general, la determinación de la inicialización de los centroides juega un papel crucial debido a que cuanto mejor sea la partición inicial más rápido convergerá el algoritmo. Aunque se asegura la convergencia, ésta no tiene por qué ser un mínimo global. Desde el punto de vista de la detección del mofo, al tratarse de imágenes una forma de realizar esta elección es calcular el histograma n-dimensional de la imagen y determinar los picos dominantes del mismo. Los picos dominantes se hacen corresponder con los valores de los centroides iniciales. El concepto de histograma n-dimensional hace referencia al hecho de que las imágenes en color poseen generalmente tres canales, bien R, G y B en el modelo de color RGB u otro conjunto en otro modelo. Para cada canal se puede obtener un histograma, que en su conjunto formarían el histograma tri-dimensional en este caso. No obstante, dicha asignación se puede realizar aleatoriamente al principio del proceso, con lo que disminuye la complejidad pero aumenta el tiempo de convergencia del algoritmo.

El problema es computacionalmente complejo y varía según la heurística empleada, siendo de orden polinomial en la mayoría de los casos (Arthur y col., 2009). Sin embargo, hay eficientes heurísticas que se emplean comúnmente y convergen rápidamente a un óptimo local. Es fácil de implementar incluso para largos conjuntos de datos. Por lo que ha sido ampliamente usado en muchas áreas tales como segmentación de mercados, geoestadística y astronomía, y por supuesto en Visión por Computador. También se usa como preprocesamiento para otros algoritmos, por ejemplo para buscar una configuración inicial.

AGRUPAMIENTO BORROSO FUZZY O CLUSTERING

El método de agrupamiento borroso es también conocido como *Fuzzy Clustering* en terminología inglesa. En este trabajo nos referiremos indistintamente a los términos borroso y fuzzy sin ninguna distinción. Su funcionamiento básico consiste en dividir primero el conjunto de muestras en clases, luego aparece una fase de optimización

para reducir el número de parámetros utilizados en cada clase, y finalmente combina las clases locales obteniendo una solución final que encuentra una mínima distancia. Esta parte se puede realizar de varias formas. Algunos métodos intentan utilizar una heurística que genera el mínimo conjunto de reglas de decisión para cada clase (Grzymala-Busse, 2004), si bien, cuando existen atributos que no son significativos, aumenta sustancialmente la complejidad del problema. Otros buscan la mínima distancia mediante un algoritmo teórico-gráfico, obteniendo el camino más corto mediante el método de Dijkstra o alguna de sus variantes (Takagi y col., 2004).

Otro planteamiento mediante la estrategia fuzzy consiste en cubrir con cada regla una región concreta, que constituye el antecedente de un atributo, mientras que en el consecuente se maximiza la calidad de esa regla. Esto hace que la asignación en el consecuente determine para cada regla la clase mayoritaria correspondiente (Alatas y Akin, 2005), sin embargo este método ofrece varios problemas, uno de ellos es que requiere que los datos estén adaptados convenientemente para que la función heurística pueda ser utilizada. Otro problema es que necesita utilizar una regla que haga de transición entre otras reglas, pues la probabilidad de asignar una condición a una regla equivocada es elevada, esto y las actualizaciones constantes de las condiciones ralentizan y dificultan considerablemente el proceso, llegando a producir pérdidas de información. En contraposición al planteamiento anterior, en Fakhrahmad y col. (2007) sugieren impedir la generación de todas las reglas posibles con respecto a todas las combinaciones de antecedentes para mejorar la interoperabilidad, de forma que elaboran su clasificador mediante la generación de reglas con distinto peso específico. Es posible que los resultados no sean mucho mejores porque fácilmente producirá la pérdida de datos correspondientes a una clase minoritaria.

En próximo capítulo nos centraremos en la variación difusa del algoritmo K-Means, denominado Fuzzy K-means también llamado Fuzzy C-means en el que a diferencia del determinismo de K-Means en que cada punto pertenece a una única clase, surgirá la visión probabilística en que un punto puede pertenecer a más de un cluster con cierto valor de pertenencia, denominado grado (Jain y col, 1999).

3 REVISION DE METODOS UTILIZADOS

3.1 INTRODUCCIÓN

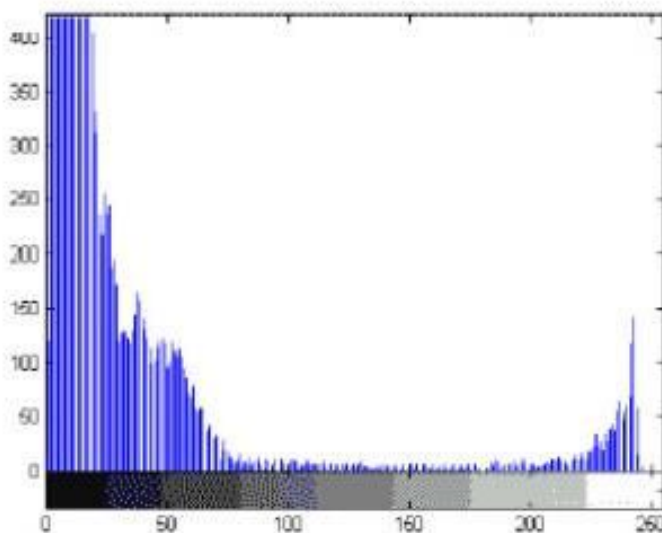
En el capítulo previo se analizaron distintas posibilidades de segmentación y clasificación de imágenes con fines de detección del moho en quesos. En este capítulo ahondaremos en los dos métodos de segmentación que posteriormente se aplican sobre el problema de detección de moho al que nos enfrentamos. De los múltiples métodos de segmentación estudiados en la literatura, que ofrecen resultados favorables en su combinación entre tiempo de ejecución y porcentaje de acierto, se ha seleccionado los siguientes, por ser eficientes para la solución del problema en cuestión y por las garantías que ofrece el hecho de haberse aplicado ampliamente en la literatura con resultados eficientes, en concreto se trata de los siguientes:

- Umbralización mediante método de Otsu
- Agrupamiento borroso mediante *Fuzzy K-Means*

3.2 MÉTODO OTSU

El método de Otsu (1979), utiliza técnicas estadísticas para resolver el problema de detección del valor umbral óptimo en una dispersión. En concreto, se utiliza la variancia, que es una medida de la dispersión de valores, en este caso aplicado a los niveles de gris de la imagen.

El método de Otsu calcula el valor umbral de forma que la dispersión dentro de cada segmento sea lo más pequeña posible, pero al mismo tiempo, la dispersión sea lo más alta posible entre segmentos diferentes. Para ello se calcula el cociente entre ambas variancias y se busca un valor umbral para el que este cociente sea máximo. En la figura 3-1 se muestra el histograma que representa la distribución de niveles de intensidad de grises en una hipotética imagen. En el eje de abscisas se muestra la intensidad de gris que representa cada grupo de apariciones.



3-1 Histograma imagen polar

La importancia del método de Otsu radica en que es automático, es decir, no necesita supervisión humana ni información previa de la imagen antes de su procesamiento.

Descripción del Método de Otsu para un umbral óptimo

Una imagen es una función bidimensional de la intensidad del nivel de gris, y contiene N píxeles cuyos niveles de gris se encuentran entre 1 y L . El número de píxeles con nivel de gris i se denota como f_i , y la probabilidad de ocurrencia del nivel de gris i en la imagen está dada por:

$$p_i = \frac{f_i}{N}$$

En el caso de la umbralización en dos niveles de una imagen conocida también como binarización, los píxeles son divididos en dos clases: C_1 con niveles de gris $[1, \dots, t]$; y C_2 , con niveles de gris $[t + 1, \dots, L]$. Entonces, la distribución de probabilidad de los niveles de gris para las dos clases es:

$$C_1: \frac{p_1}{\omega_1(t)}, \dots, \frac{p_t}{\omega_1(t)}$$

$$C_2: \frac{p_{t+1}}{\omega_2(t)}, \frac{p_{t+2}}{\omega_2(t)}, \dots, \frac{p_L}{\omega_2(t)}$$

donde

$$\omega_1(t) = \sum_{i=1}^t p_i \quad \omega_2(t) = \sum_{i=t+1}^L p_i$$

También, la media para la clase C_1 y la clase C_2 es

$$\mu_1(t) = \sum_{i=1}^t \frac{i \cdot p_i}{\omega_1(t)} \quad \mu_2(t) = \sum_{i=t+1}^L \frac{i \cdot p_i}{\omega_2(t)}$$

Sea μ_T la intensidad media de toda la imagen. Es fácil demostrar que

$$\omega_1 \cdot \mu_1 + \omega_2 \cdot \mu_2 = \mu_T \quad \omega_1 + \omega_2 = 1$$

Usando análisis discriminante, Otsu definió la variancia entre clases de una imagen umbralizada como:

$$\sigma_B^2 = \omega_1 \cdot (\mu_1 - \mu_T)^2 + \omega_2 \cdot (\mu_2 - \mu_T)^2$$

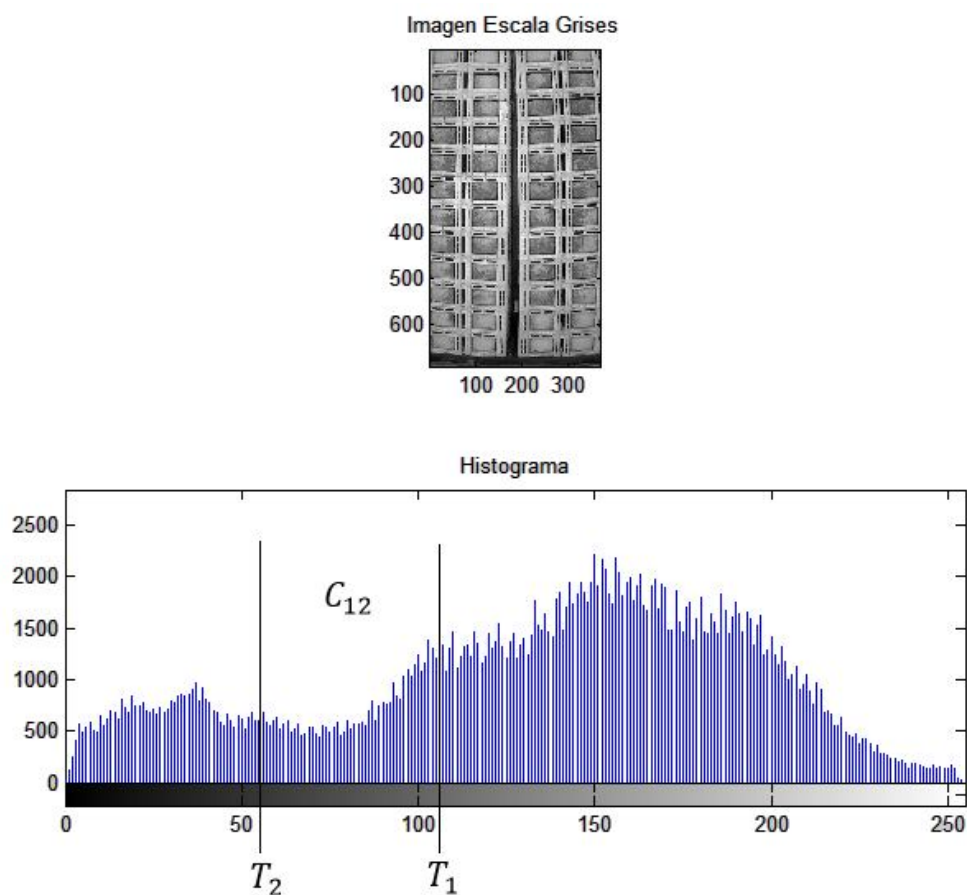
Para una umbralización de dos niveles, Otsu verificó que el umbral óptimo t^* se elige de manera que σ_B^2 sea máxima; esto es

$$t^* = \max_t \{\sigma_B^2(t)\} \quad 1 \leq t \leq L$$

Tiene el inconveniente de que es muy sensible a las variaciones en la luminosidad. Por ello, a menudo se aplican técnicas de reducción de la borrosidad sobre los bordes o lo que es lo mismo del incremento en la nitidez de los bordes. En el anexo II se detallan los distintos sistemas de iluminación que se han instalado y probado para la

obtención de imágenes, y con los que se ha conseguido minimizar el impacto negativo de las variaciones de iluminación que este algoritmo posee.

La aplicación de este método al problema de detección de moho que aquí se estudia, se plantea aplicando dos veces de forma consecutiva el método. La primera vez aplicamos el método Otsu sobre la imagen original. Tal y como se ha descrito, al tratarse de un método de segmentación en dos clases, obtendremos un umbral T_1 que dividirá la imagen en una clase C_1 con niveles de intensidad de gris más oscuros y otra clase C_2 con niveles de gris más claros. Seguidamente sobre la primera de las clases C_1 , se volverá a aplicar el método consiguiendo un nuevo umbral T_2 que dará lugar a dos nuevas clases que llamaremos C_{11} , elementos de C_1 , con niveles de gris más oscuros y C_{12} que contendrá los niveles de gris más claros de C_1 . Se considerará moho los píxeles pertenecientes a esta segunda clase C_{12} , y que por lo tanto contendrá aquellos píxeles con niveles de gris comprendidos entre T_1 y T_2 . En la figura 3-2 se muestra la imagen de un palé en escala de grises, el histograma correspondiente a dicha imagen y los umbrales obtenidos como resultado de aplicar dos veces de forma consecutiva el método de umbralización de Otsu tal y como se acaba de explicar sobre una imagen real.



3-2 Ejemplo histograma imagen palé con los umbrales resultado de aplicar dos veces el método Otsu

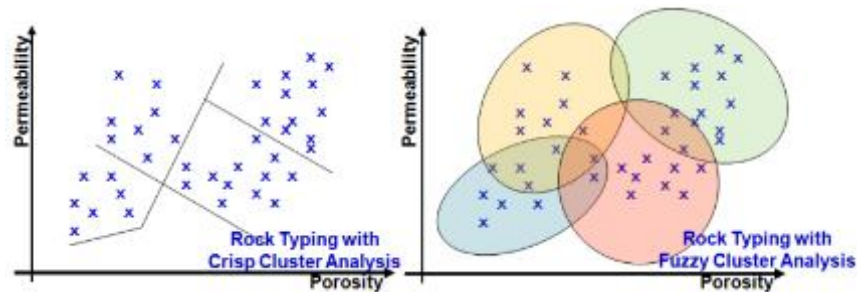
3.3 FUZZY K-MEANS

Fuzzy K-means también llamado Fuzzy C-means, es una extensión del K-mean introducido en el capítulo previo. Este tipo de algoritmos surge de la necesidad de resolver una deficiencia del agrupamiento exclusivo que hace K-means, el cual considera que cada elemento se puede agrupar inequívocamente con los elementos de su *cluster* y que, por lo tanto, no se asemeja al resto de los elementos. Tras la introducción de la lógica difusa por Zadeh en 1965 surgió una solución para este problema, caracterizando la similitud de cada elemento a cada uno de los grupos

(Bezdek y col, 1984). Esto se logra representando la similitud entre un elemento y un grupo por una función, llamada función de pertenencia, que toma valores entre cero y uno. Los valores cercanos a uno indican una mayor similitud, mientras que los cercanos a cero indican una menor similitud. Por lo tanto, el problema del agrupamiento difuso se reduce a encontrar una caracterización de este tipo que sea óptima. Es decir, la teoría *Fuzzy set* es una generalización del álgebra de booleana, por la que la función de pertenencia de un elemento a los grupos se encuentra en el intervalo [0,1], es decir, un elemento puede pertenecer totalmente a una clase, a todas o a ninguna (Ortega y col, 2002)

Como el K-means, la versión Fuzzy k-means trabaja con aquellos objetos que pueden ser representados en un espacio n-dimensional con una medida de distancia definida (Jain y col, 1999)

Los algoritmos Fuzzy K-means son algunos de los principales algoritmos utilizados en el agrupamiento difuso y pertenecen a una clase de algoritmos basados en funciones objetivo. En la figura 3-3 se pueden ver dos gráficos con sendas nubes de puntos, la primera de ellas con agrupamiento determinista y la segunda con agrupamientos difusos.



3-3 Agrupamiento de nube de puntos en 4 clusters con clustering determinista y clustering borroso

Estos algoritmos definen un criterio de agrupamiento en la forma de una función objetivo que depende de la partición difusa. El procedimiento, en sentido general, consiste en minimizar iterativamente esta función hasta obtener una partición difusa óptima. En este caso es la función de error mínimo cuadrático, la que se pretende minimizar:

$$J_m(U, v) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m d_{ik}^2$$

El valor d_{ik}^2 indica la distancia cuadrada entre los elementos de x y los centros de los grupos y puede calcularse utilizando la siguiente expresión:

$$d_{ik}^2 = \|x_k - v_i\|_A^2 = (x_k - v_i)^T A (x_k - v_i)$$

Donde:

$x = (x_1, x_2, \dots, x_n) \in R^s$ son los datos,

$v_i = (v_{i1}, v_{i2}, \dots, v_{ic})$ es el vector centro del grupo i ,

$\|A\|$ es la norma inducida por A

A es una matriz definida positiva de pesos de dimensiones $(n \times n)$.

En particular, si A es la matriz identidad, d_{ik}^2 es el cuadrado de la distancia euclidiana.

El peso asociado a cada distancia cuadrada, $(u_{ik})^m$ es la m -ésima potencia del grado de pertenencia del k -ésimo dato al grupo i . Cuando $m \rightarrow 1$ la partición óptima es cada vez más cercana a una partición exclusiva, mientras que cuando $m \rightarrow \infty$ la partición óptima se aproxima a la matriz con todos sus valores iguales a $(1/c)$. Los valores de m

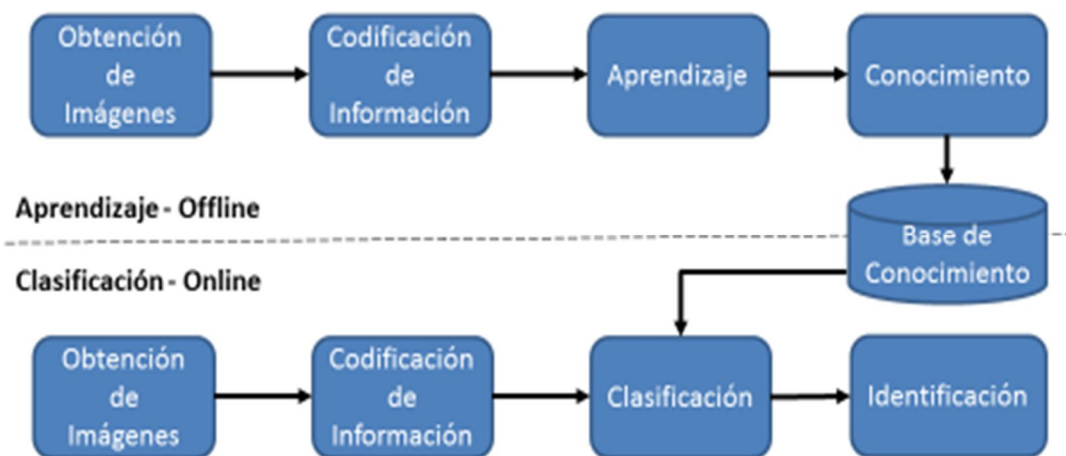
que normalmente se utilizan son valores en el intervalo $[1,30]$. Cada selección de un valor particular de m marca un algoritmo *Fuzzy K-means* específico.

Al tratarse de un problema de optimización iterativo, según se defina el criterio de detención, el algoritmo convergerá con mayor o menor número de iteraciones. Sin embargo, la solución no tiene por qué ser siempre óptima pues puede converger hacia mínimos locales en función de la estimación inicial (elección de centroides en el inicio) (Olivai y col, 2001), la misma desventaja que aparecía en el algoritmo *K-means*, aunque este método mejora el problema de convergencia del *K-means* (Jain y col, 1999).

4 PROBLEMÁTICA Y DISEÑO

4.1 INTRODUCCIÓN

Antes de comenzar el proceso de diseño del sistema que se describe más adelante, conviene destacar las fases o mecanismos involucrados en todo proceso de aprendizaje-clasificación. Destacar, como ya hemos comentado en puntos anteriores, la diferente naturaleza de los dos algoritmos que se aplican para la resolución del problema, Fuzzy K-Means y umbralización mediante el método de Otsu, difieren en que el primero requiere una fase de aprendizaje supervisado, mientras que el segundo de ellos, no requiere intervención humana ni información previa. La figura 4-1 muestra un esquema general del proceso, que se describe seguidamente



4-1 Esquema del proceso de aprendizaje-clasificación

Existen dos procesos perfectamente identificados desde el punto de vista de integración en el proceso de fabricación del queso, el primero de ellos off-line, sólo lo lleva a cabo el algoritmo Fuzzy K-means, y en él se realiza la fase de aprendizaje tal y como hemos comentado en el punto tres. El segundo proceso on-line, que ambos algoritmos poseen, Fuzzy K-means y umbralización mediante el método Otsu, y en el que se realiza la clasificación de las muestras.

A continuación se describen con más detalle cada una de las fases mostradas en la figura 4.1:

Obtención de imágenes: fase que tiene lugar tanto en el proceso de aprendizaje como en el de clasificación, y que en nuestro caso no tienen distinción alguna. Procedentes de la instalación de una serie de cámaras digitales que veremos en el siguiente punto con más detalle, se extraen una serie de imágenes en formato JPEG de resolución de 720x 1280 píxeles. Sobre cada una de estas imágenes se realiza un pre procesamiento en el que se elimina un marco blanco o máscara, que rodea las partes de interés de la escena, y que permite eliminar a su vez las partes de la imagen carentes de información. Estas máscaras son configuradas cuando se realiza el montaje e instalación de cada una de las cámaras. En el Anexo II se muestra exactamente la distribución y montaje de dichas cámaras.

Codificación de la información: una vez que la información se dispone de forma que pueda ser procesada adecuadamente, el color de cada píxel se codifica en el espacio RGB en forma de vector o matriz tridimensional.

Aprendizaje: en el caso del método Fuzzy K-means, se aplica el algoritmo sobre la información obtenida en las dos fases anteriores, proporcionándole para ello los parámetros necesarios para su ejecución, número de clústeres y número máximo de iteraciones para conseguir una solución. Debido a la complejidad y variabilidad del problema, los resultados obtenidos en esta fase (centros de los clústeres), son clasificados por un experto humano externo, en este caso el maestro en el proceso quesero, que clasifica cada uno de los *clústeres* identificándolo como “moho” y “no moho” respectivamente.

Estos parámetros se almacenan en la Base de Conocimiento (BC) con el fin de poderlos recuperar durante la fase de clasificación que como hemos comentado, se ejecuta de forma on-line.

Conocimiento: está compuesto por los parámetros resultantes de la fase de aprendizaje y son almacenados en la BC. En concreto se trata de los centroides de cada cluster y de la consideración como moho para cada uno de ellos. Esta información será utilizada para clasificar las nuevas muestras y cada pixel que las componen.

Clasificación: proceso que se encuentra integrado en el proceso de fabricación del queso. Se obtienen las imágenes procedentes de las cámaras instaladas tal y como se ha descrito en la fase de aprendizaje y sobre las que se aplican de forma totalmente desasistida, cada uno de los clasificadores de estudio. En el caso de Fuzzy K-means, se asignará cada píxel de las imágenes, a cada uno de los *clústeres* cuyos centroides están almacenados en la BC, de tal forma que se minimice la distancia euclidiana entre dichos centroides. Para el caso de la umbralización mediante el método Otsu, se ejecuta el algoritmo dos veces, obteniendo dos umbrales de Otsu. Clasificaremos entonces cada píxel de la imagen según su nivel de gris, considerando como moho aquellos píxeles que se encuentren dentro del intervalo delimitado por ambos umbrales, tal y como se ha explicado en el punto 3.2.

Identificación: a partir de las clasificaciones anteriores, se obtiene para cada imagen un porcentaje de píxeles perteneciente a cada clase o cluster, y por extensión que se consideran moho o no moho. Promediando el valor obtenido de cada una de las imágenes, se obtiene el valor que identifica el nivel de moho del palé.

4.2 ANTECEDENTES

Tal y como se ha explicado en la introducción, la motivación del diseño de un sistema de detección de niveles de moho viene dada por la necesidad de dar solución al problema real de presencia de grandes niveles de moho en las instalaciones de una de los principales fabricantes de queso de pasta prensada de España.

Con anterioridad a la apertura del nuevo centro logístico totalmente automatizado, el proceso de elaboración del queso se llevaba a cabo de forma manual, permitiendo la inspección y detección de forma continua por parte de los usuarios expertos de los niveles de moho. Con el ambicioso proyecto tecnológico de modernización del proceso de producción, en el que se apuesta por una instalación totalmente automatizada, donde se aplican las tecnologías industriales más punteras como, vehículos guiados de forma desasistida, volteadores automáticos, robots y almacenes inteligentes, se hace prácticamente imposible la supervisión humana, del mismo modo que se venía haciendo en el proceso manual.

Hasta donde se ha podido indagar, no se tiene constancia de la aplicación de las técnicas de visión por computador propuestas en este trabajo en un proceso de este tipo o similar, al igual que se desconoce la existencia de sistemas desarrollados implantados que apliquen dichas técnicas o similares a entornos de producción queseros, con fines de detección de niveles de moho.

4.3 DISEÑO DEL SISTEMA

Como ya se ha explicado en los capítulos dos y tres, las técnicas aquí expuestas han sido desarrolladas dando forma a un sistema para su posterior implantación en un entorno real. En el siguiente apartado describiremos brevemente cómo se ha llevado a cabo el diseño y despliegue del sistema, cuyos detalles se incluyen en el Anexo II, así como los distintos subsistemas que intervienen en el proceso de captura de información.

El proceso se inicia cuando se reciben en el centro logístico los distintos palés de producto procedentes de cada una de las fábricas de la quesería. En dicho centro logístico se encuentra instalado el sistema de detección de moho en queso mediante procesamiento de imágenes que implementa los métodos de clasificación descritos en los apartados dos y tres. Dicho sistema consta de dos partes totalmente diferenciadas:

Servidor o sistema de detección de moho: Se trata del subsistema encargado de cubrir las fases de aprendizaje, conocimiento, clasificación e identificación descritas en el punto 4.1. Además, se encargará de realizar el intercambio de información y coordinación con los distintos sistemas que intervienen en el proceso.

Cliente detección de moho: Se trata de una aplicación que permitirá la consulta de la información generada por el servidor para los distintos palés que pasen por el proceso

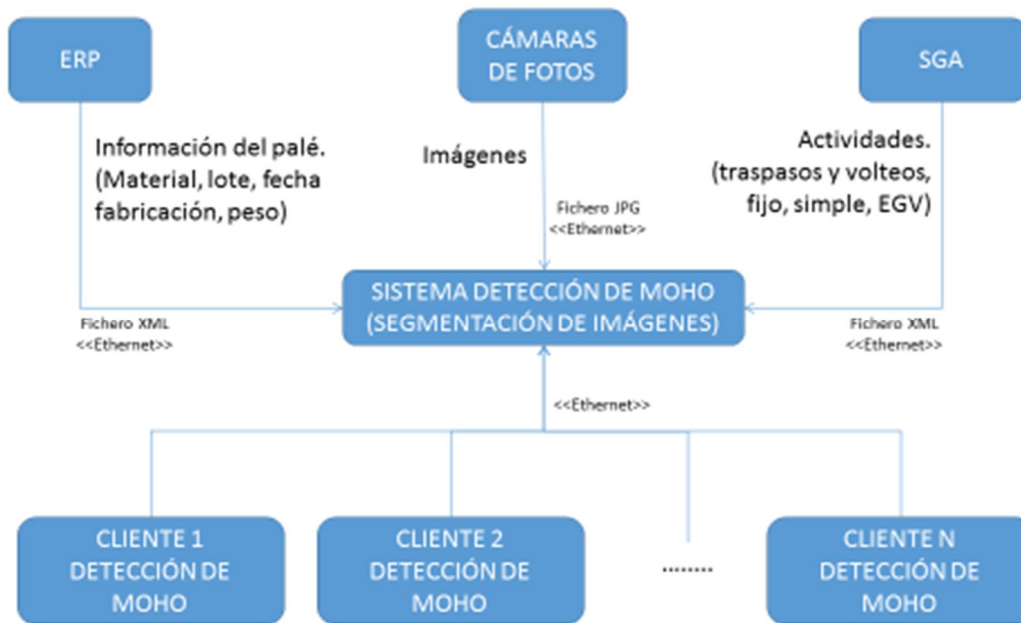
Además de estos sistemas, en el proceso intervienen otros sistemas que se encuentran presentes en la instalación, como es el sistema de planificación de recursos empresariales (con sus iniciales en inglés ERP), el sistema gestor del almacén (con sus iniciales en inglés SGA), el sistema de AGVs, o el sistema de captura de imágenes compuesto por cuatro cámaras.

El flujo de la información se inicia por parte del ERP, enviando un fichero XML con la información de los palés que entran indicando el producto, lote, fecha de fabricación, peso. Esta información la recibe la aplicación servidor y la almacena en la base de datos correspondiente.

Tras esto, los palés son ubicados físicamente en cámaras frigoríficas mediante el sistema de AGVs, reportando dicha información, mediante un fichero XML, a la aplicación servidor que lo almacena.

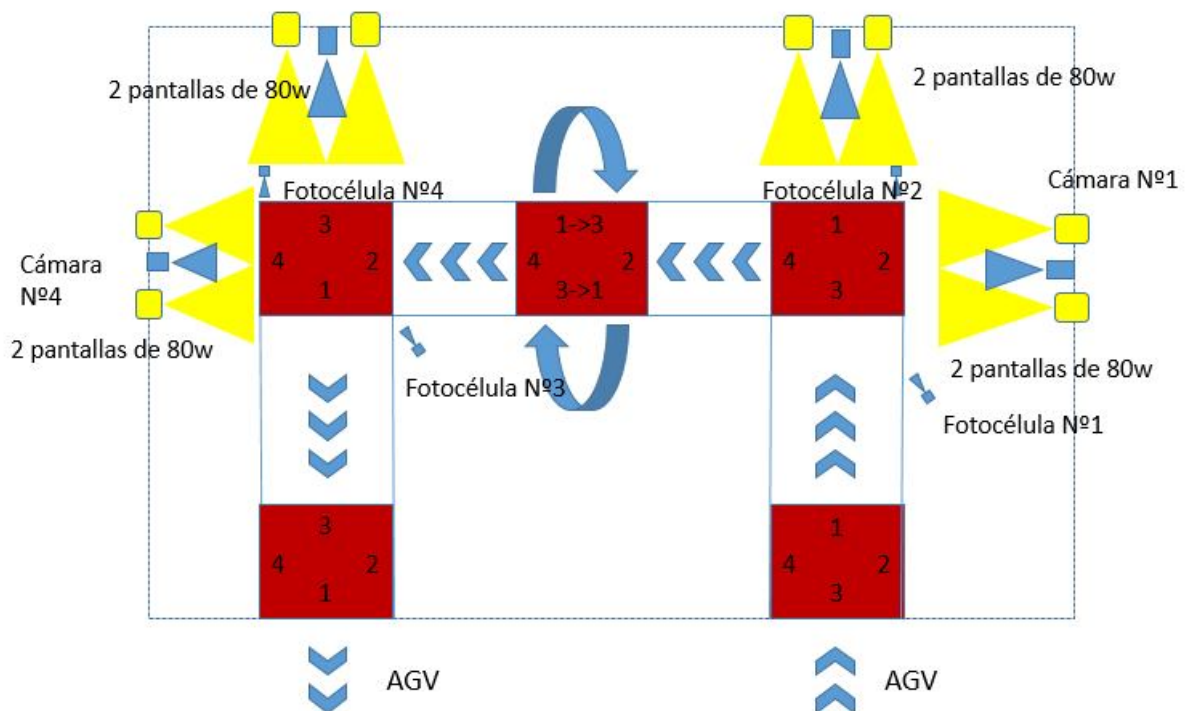
A partir de este momento comienza la fase de maduración del producto, cuya duración estriba entre 1 a 12 meses. Durante este tiempo, el sistema SGA realiza distintas actividades sobre el palé de quesos, distintos tipos de volteo y traslados entre cámaras frigoríficas bajo distintas condiciones. Todos estos procesos, al igual que en el caso anterior, generan una comunicación entre el SGA y la aplicación servidor en forma de fichero XML, manteniendo así la trazabilidad de cada palé durante todo el proceso de curación y en su tránsito por toda la instalación.

Es en una de estas actividades, el volteo simple, que consiste en girar el palé 180° sobre la vertical con los quesos, donde se obtiene las imágenes del palé para la aplicación de las técnicas de segmentación. En la figura 4-2 podemos ver un esquema de los distintos sistemas que tomar parte en el proceso que acabamos de explicar y de la información y formato que aporta cada uno al proceso.



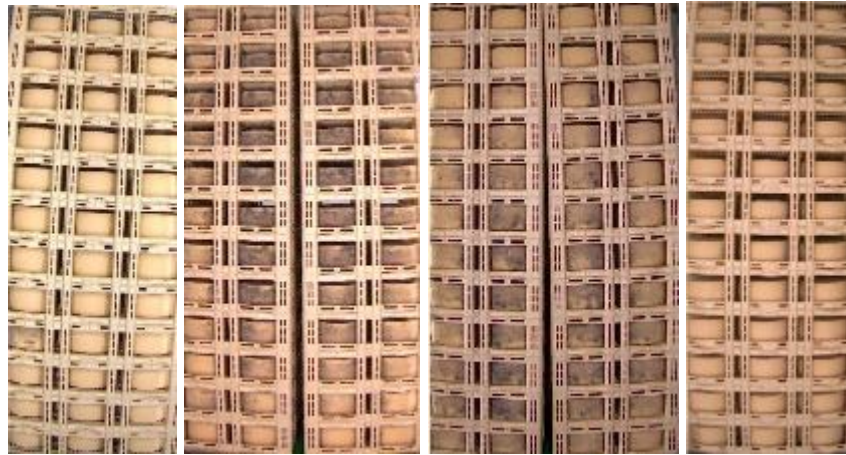
4-2 Diagrama de despliegue de los distintos subsistemas que intervienen

En cuanto al sistema de captura de imágenes, se ha llevado a cabo mediante la instalación de 4 cámaras, 8 puntos de iluminación y 4 fotocélulas para la detección de presencia del palé delante de las cámaras, obteniendo cuatro imágenes, una por cada cara del palé. En la figura 4-3 podemos ver el esquema de la instalación de las cámaras, pantallas de iluminación y fotocélulas. Además se representa el recorrido que el palé lleva por la instalación y como son capturadas cada una de las cuatro caras del palé representadas por los número del 1 al 4. Comentar que el recorrido completo del palé se lleva a cabo en 1 minutos 30 segundos y que la cadencia mínima por cada uno de los puestos de captura de imágenes ronda los 15 segundos.



4-3 Esquema de instalación cámaras, pantallas y fotocélulas en volteo simple

De las cuatro imágenes obtenidas, se extrae información de los 144 quesos que componen el palé y que se encuentran situados dentro de las cajas de plástico de distintos colores, con aperturas laterales en forma de cuadrante. En cada uno de estos cuadrantes podemos ver una cuarta parte del área lateral de un queso de forma cilíndrica, por lo que tendremos capturada un área total de $144 \times 2\pi r h / 4$, frente a una superficie total de 144 quesos $144 \times 2\pi r (h + r)$. De esta forma, con las cuatro imágenes laterales, una por cada lado del palé tendremos una información de algo menos del 15% la superficie total del queso para analizar mediante la aplicación de los métodos de segmentación explicados en el capítulo dos y tres. En la figura 4-4 podemos ver el ejemplo de las cuatro imágenes capturadas para un palé, una por cada cara de mismo.



4-4 Ejemplos de fotografía de un palé obtenidos por el sistema

5 ANALISIS DE RESULTADOS

5.1 INTRODUCCIÓN

Una vez establecidas las distintas estrategias que hemos seguido para la resolución del problema planteado en nuestra investigación, en este capítulo se hace el correspondiente análisis y valoración de las mismas. El objetivo principal de nuestro trabajo es la identificación automática de moho en los palés de queso, mediante la generación de un indicador que mida los niveles de aparición del mismo.

Como se ha comentado reiteradamente el algoritmo Fuzzy K-means consta de dos fases claramente diferenciadas, aprendizaje y clasificación. En una primera etapa evaluaremos la fase de aprendizaje de dicho algoritmo, fijando el número de clases a emplear, estudiando tanto la convergencia del algoritmo, número de iteraciones necesarias y la eficiencia a la hora de identificar el moho, para así poder evaluar la validez de la fase de clasificación.

Por el contrario, el algoritmo de Otsu al carecer de la fase de aprendizaje por tratarse de un método automático, no realizaremos estudio alguno en lo que respecta a la fase de aprendizaje, como por otra parte resulta obvio.





En la segunda etapa, evaluaremos la fase de clasificación, para ello aplicaremos los dos algoritmos sobre una serie de imágenes y comprobaremos los resultados obtenidos con el objetivo de poder determinar cuál es la mejor de las dos técnicas aquí estudiadas.

Para ambas fases utilizaremos la base de datos de imágenes reales que se describe en el siguiente apartado.

5.2 DESCRIPCION DE LA BASE DE CONOCIMIENTO UTILIZADA

Debido a la inexistencia tanto de estudios teóricos como de aplicaciones prácticas de técnicas de Visión por Computador a un proceso similar de detección de moho sobre la superficie del queso, se ha procedido a generar el correspondiente “*ground truth*” para el estudio de los resultados de aplicar las distintas técnicas de clasificación propuestas. Para ello se ha generado una base de datos con 100 imágenes reales procedentes del proceso de volteo simple de 25 palés, elegidas de forma aleatoria como representación de las 39000 imágenes recogidas durante más de 3 meses de forma automática por el sistema. Un experto del proceso de maduración del queso ha evaluado cada una de las imágenes de la muestra según dos criterios, cantidad de superficie del palé afectada por moho y nivel de concentración de moho en la superficie afectada. Estas características se evaluarán de 0 a 3 según su intensidad creciente (0=nada, 1=algo, 2=bastante y 3=mucho). Cada uno de estos niveles se traducirá en un porcentaje que refleje la magnitud con el objeto de aplicar dichos coeficientes como factores de multiplicación y obtener así un indicador de presencia de moho por imagen. Tras obtener el indicador de cada una de las cuatro imágenes que conforman el estudio del paso por volteo simple de un palé, se realizará el promedio, obteniendo así el indicador de moho de dicho palé en ese momento. Este valor obtenido, se enfrentará al obtenido por cada uno de los algoritmos estudiados que al igual que el anterior indicador, se obtendrá a partir del valor promedio obtenido de las cuatro imágenes que conforman el estudio. Dicha base de datos se sintetiza en el Anexo I. La figura 5-1 muestra la cabecera y una de las filas de la tabla almacenada en la base de conocimiento. Como se ha comentado, se puede ver cómo para cada palé estudiado, se recoge la imagen procedente de cada una de las cuatro cámaras, ésta es evaluada sobre dos características, superficie afectada por moho y nivel de concentración del moho. Cada característica se evalúa en cuatro niveles. De la

traducción en porcentaje de estos niveles y el producto de las dos características, se obtiene el nivel de indicador de moho de cada imagen, promediando estos cuatro valores se obtiene el indicador de moho del palé.

Base de Conocimiento de niveles de moho en palés																	
Nº de Palé	Imagen Camara 1	Superficie Afectada	Nivel Concentracion	Indicador Moho Imagen	Imagen Camara2	Superficie Afectada	Nivel Concentracion	Indicador Moho Imagen	Imagen Camara3	Superficie Afectada	Nivel Concentracion	Indicador Moho Imagen	Imagen Camara 4	Superficie Afectada	Nivel Concentracion	Indicador Moho Imagen	Indicador Moho Palé
1		2	3	66%		2	3	66%		2	2	44%		1	1	10%	47%

5-1 Ejemplo cabecera y tupla base de datos de conocimiento

Como se puede observar a partir de las imágenes del ejemplo, éstas varían tanto en posición como en iluminación, debido a la naturaleza tanto del proceso, como de la instalación del sistema de visión en relación a su ubicación con respecto al palé, como se ha explicado en el punto 4.3.

5.3 METODO UTILIZADO PARA LA APLICACIÓN

La implementación se ha realizado en la plataforma .NET, que es un producto de la compañía Microsoft, más en concreto se ha utilizado C# que se trata de un lenguaje orientado a objetos compilado. Además de por el conocimiento y experiencia en la plataforma y el lenguaje, se ha decidido usar un lenguaje compilado con el objeto de optimizar y requerir menor potencia de cálculo que un lenguaje interpretado, como el que utiliza Matlab. No obstante, los tiempos de cómputo no constituyen el objeto clave del presente estudio, siempre y cuando no superen el tiempo del propio proceso de captura de imágenes explicado en el punto 4.3 (15 segundos por imagen, un minuto y medio por palé). Si veremos que en la fase de aprendizaje, la naturaleza iterativa del método *Fuzzy K-means* origina la necesidad de un cálculo intensivo, pudiendo llegar a ser un problema los tiempos de respuesta del sistema.

5.4 FASE DE APRENDIZAJE

Como se ha descrito en el punto 4.3, *Fuzzy K-means* se trata de un sistema de aprendizaje supervisado, por lo que el objetivo de este punto será fijar los parámetros de dicho algoritmo, para su posterior utilización en la fase de clasificación. Para esta fase de aprendizaje, se han utilizado los datos recogidos en la BC, realizaremos distintas aproximaciones del algoritmo de *Fuzzy K-means* con el objeto de obtener el mejor de los ajustes, es decir, vamos a determinar tanto el número de *clusters* como el de las iteraciones máximas y la precisión del algoritmo que consigan una solución válida, entendiendo por solución válida, aquella que permita la identificación de los *clusters* que representan el moho en las imágenes de los palés.

Como se ha explicado en el punto 4-3, el sistema está compuesto por cuatro cámaras independientes con sus respectivos sistemas de iluminación, es por ello por lo que los parámetros utilizados para la ejecución del algoritmo *Fuzzy K-means* son independientes para cada una de las cámaras. Tendremos pues que realizar cuatro fases de aprendizaje, para llevar a cabo la calibración de cada una de las cuatro cámaras.

En esta fase el tiempo no es una limitación, siempre y cuando este se encuentre dentro de unos límites entendibles para el funcionamiento de una aplicación que interactúa con el de usuario. Al tratarse de la fase de aprendizaje, se entiende que usaremos tiempos ociosos de la máquina no dedicada únicamente para este proyecto. El servidor utilizado tiene una potencia de cálculo de 10 GFLOPS y ninguna de las ejecuciones del algoritmo, con las muestras que se describen a continuación, ha superado las 12 horas, tiempo asumible para la aplicación propuesta

Para el aprendizaje, se usará una muestra representativa de la BC de conocimiento compuesta de 10 palés elegidos de tal forma que recojan todo el espectro de variaciones de nivel de moho y casuísticas como pueden ser la presencia de cajas dentro de la estructura del palé de color negro y la presencia de cajas sin queso dentro de los palés. Se usa un total de 40 imágenes, es decir, 10 imágenes para el aprendizaje de cada una de las cuatro cámaras que conforman el sistema de detección de moho, estudiaremos pues un total de 36.864.000 píxeles para la calibración del sistema.

Para la correcta evaluación de los parámetros, además de la base de conocimiento descrita en el apartado 5.2, utilizaremos una funcionalidad que permita al usuario identificar visualmente los agrupamientos sobre la imagen, permitiendo cambiar el color representativo de las clases para facilitar la identificación de las que representan moho. Es importante resaltar que la consideración de que un *cluster* sea moho o no, es realizada por parte de un experto en el proceso, y de la elección de los mismos determinará el índice del valor de moho de la imagen.

Como ejemplo ilustrativo que sirve igualmente para las capturas de imágenes procedentes de todas las cámaras, en el siguiente apartado se analiza la variación de los parámetros de ejecución del algoritmo *Fuzzy K-means* en la fase de aprendizaje, para la cámara 2.

VARIACION DE PARÁMETROS

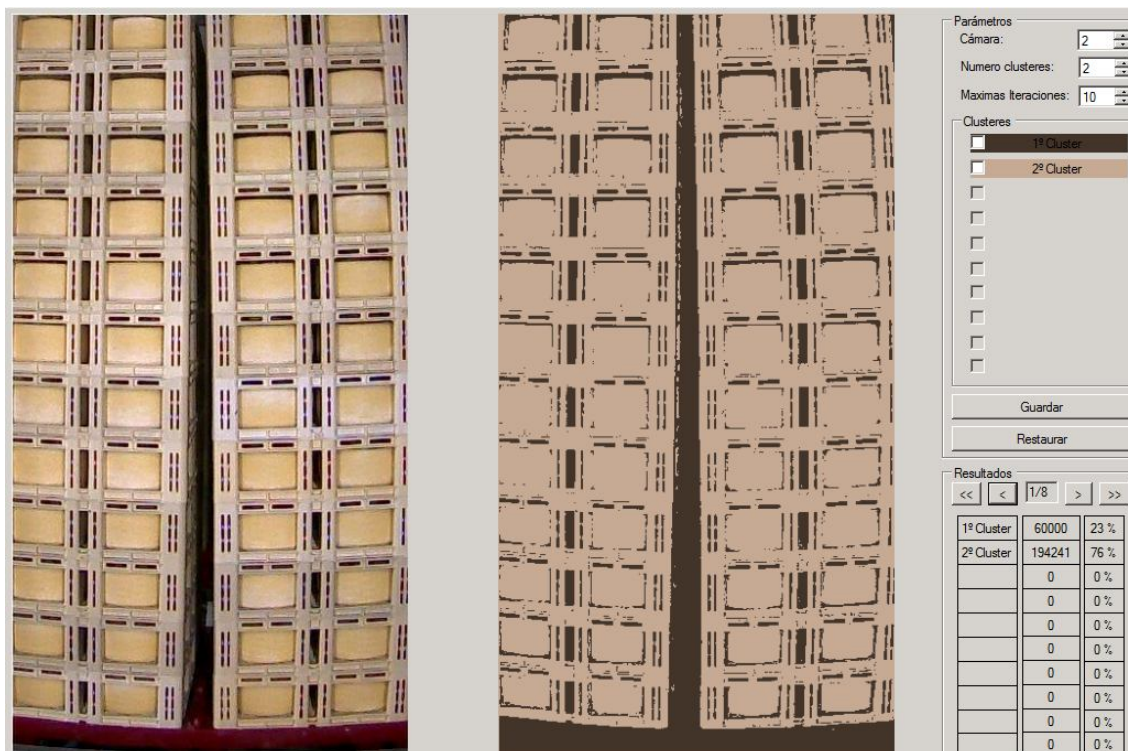
En este apartado se estudia la variación de los parámetros de entrada del algoritmo implementado Fuzzy K-means para ver cómo afectan dichas variaciones a los resultados del mismo. Para ello, de las 10 imágenes que se utilizan por cámara, se muestran los resultados sobre 3 de las mismas con la finalidad de validar el resultado, una con mucha presencia de moho, otra con poca presencia de moho y otra sin presencia de moho. Además, para determinar que la solución converge, idealmente se deberá alcanzar una precisión igual a 0 antes de alcanzar el máximo de iteraciones, es decir, que la diferencia entre los valores de la función objetivo calculada como se describe en el punto 3.3 de dos iteraciones consecutivas sea inferior a 10^{-8} . De no cumplirse esta condición entenderemos que es necesario aumentar el número de iteraciones para alcanzar una solución que converja. Para cada una de las imágenes estudiadas se mostrará una figura compuesta por la imagen original, el resultado de aplicar el aprendizaje realizado, los parámetros del algoritmo y el recuento del número de píxeles pertenecientes a cada uno de los *clusters*. El valor final del moho vendrá dado por el porcentaje de píxeles que sumen las clases denominadas como moho (marcadas con el *check*).



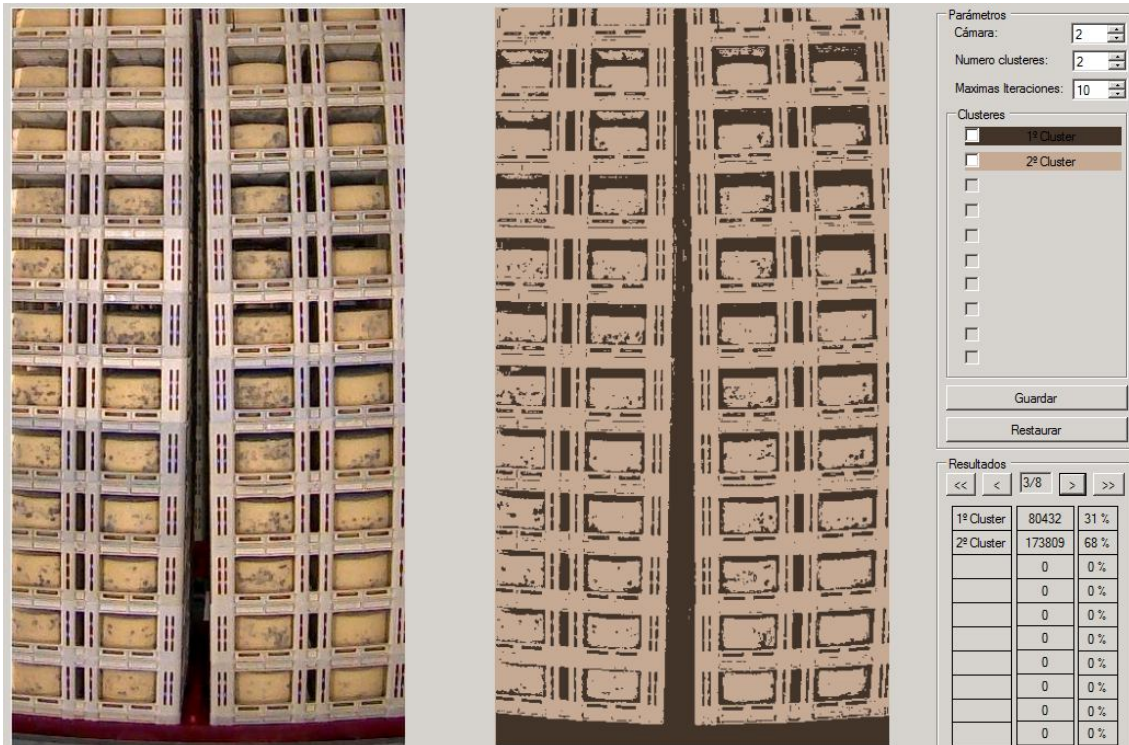
5-2 Clasificación de imagen mediante metodo Fuzzy K-means, con 10 clusters en 10 iteraciones.

1º) Ejecución con parámetros: N° clases =2, Máximas iteraciones = 10.

Tras 5 iteraciones el algoritmo converge, ya que la diferencia en la función objetivo de la 4ª iteración y la 5ª resulta ser prácticamente cero con una tolerancia de 10^{-8} . Tanto en la figura 5-4 como en la 5-5 se puede ver cómo el moho está representado por el 1º cluster, el problema es que tanto en estas figuras como en la 5-3, en dicho cluster se consideran los huecos entre las cajas y la base del palé como pixeles pertenecientes al 1º cluster. Consideraremos por lo tanto la solución como no válida y aumentaremos el número de clusters.



5-3 Resultado de imagen sin moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 2 clusters



5-4 Resultado de imagen con nivel medio de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 2 clusters



5-5 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 2 clusters

2º) Ejecución con parámetros: Nº clases =4, Máximas iteraciones = 10.

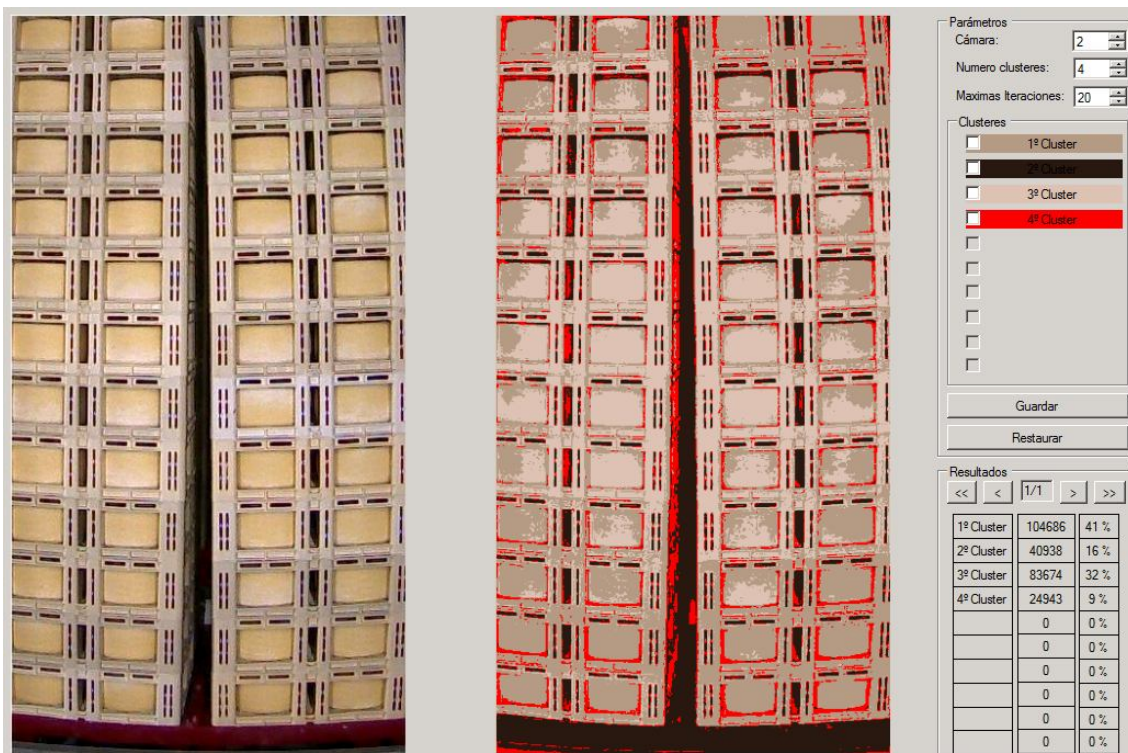
Tras ejecutar las 10 iteraciones el algoritmo concluye, devolviendo una diferencia entre la 9ª iteración y la 10ª de 2454128, por lo que claramente no converge. Recordar que dependiendo de los centroides que se usan en la inicialización para la primera iteración del algoritmo, se consigue una aproximación más rápida del algoritmo hacia

la solución. En nuestro caso como comentábamos en el punto 3.3, que la elección de estos es aleatoria, por lo que debemos ejecutar varias veces el algoritmo de aprendizaje con los mismos parámetros para determinar su validez o no.

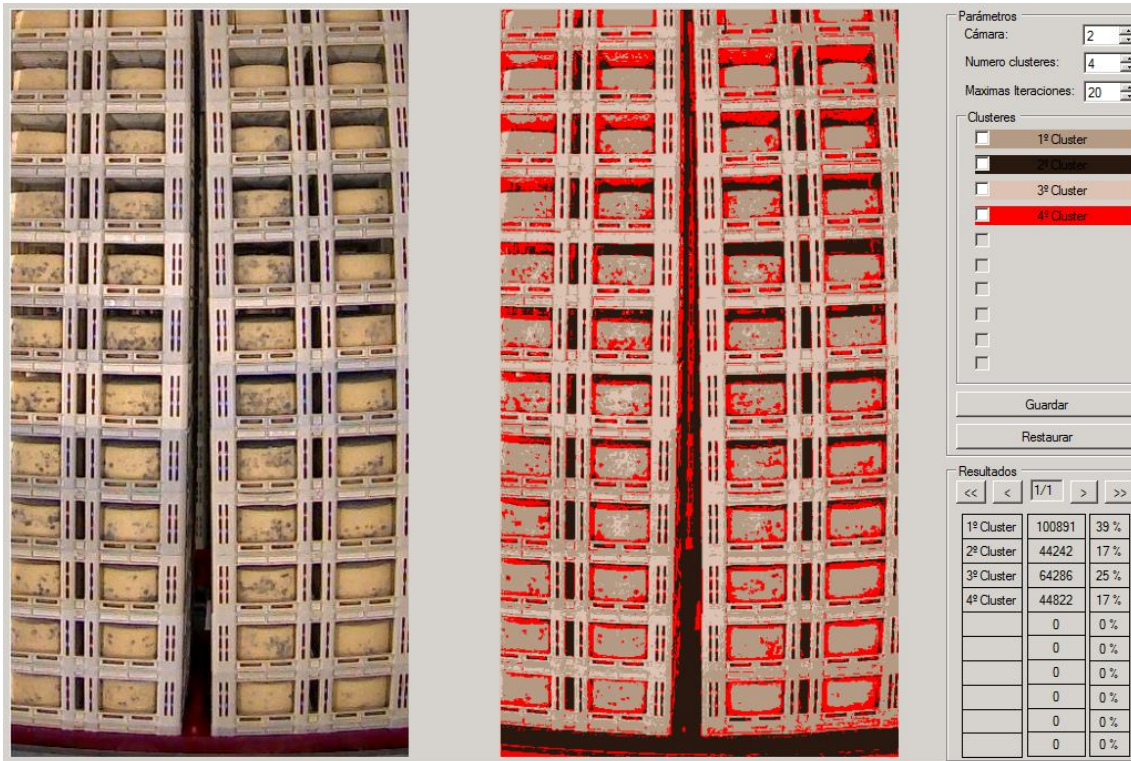
En la siguiente ejecución aumentaremos el número de iteraciones máximas para conseguir la convergencia de la solución.

3º) Ejecución con parámetros: Nº clases =4, Máximas iteraciones = 20

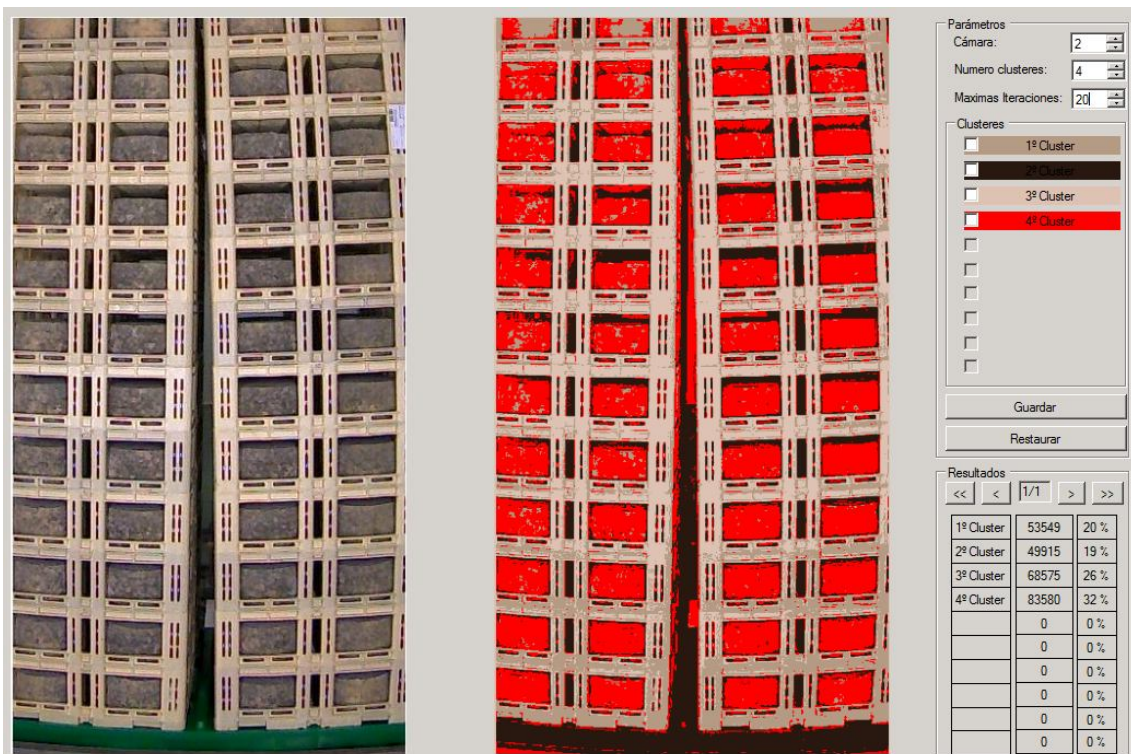
Tras 12 iteraciones el algoritmo converge y finaliza, ya que la diferencia entre la función objetivo de la 11ª y 12ª iteración se sitúa por debajo de la tolerancia fijada. Ayudándonos del experto quesero y el resultado de las 10 imágenes, le solicitamos que indique los *clusters* considerados representativos del moho, resultando ser el 4º cluster. La valoración al respecto aconseja la necesidad de una mayor división para poder aislar mejor los píxeles de moho de los huecos de las cajas. Si revisamos las figuras 5-6, 5-7 y 5-8, podemos visualizar en rojo el *cluster* número 4, considerado moho por el experto quesero. En la figura 5-6 que muestra un palé sin moho, dicho *cluster* agrupa el 9% de los píxeles de la imagen.



5-6 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 4 clusters



5-7 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 4 clusters

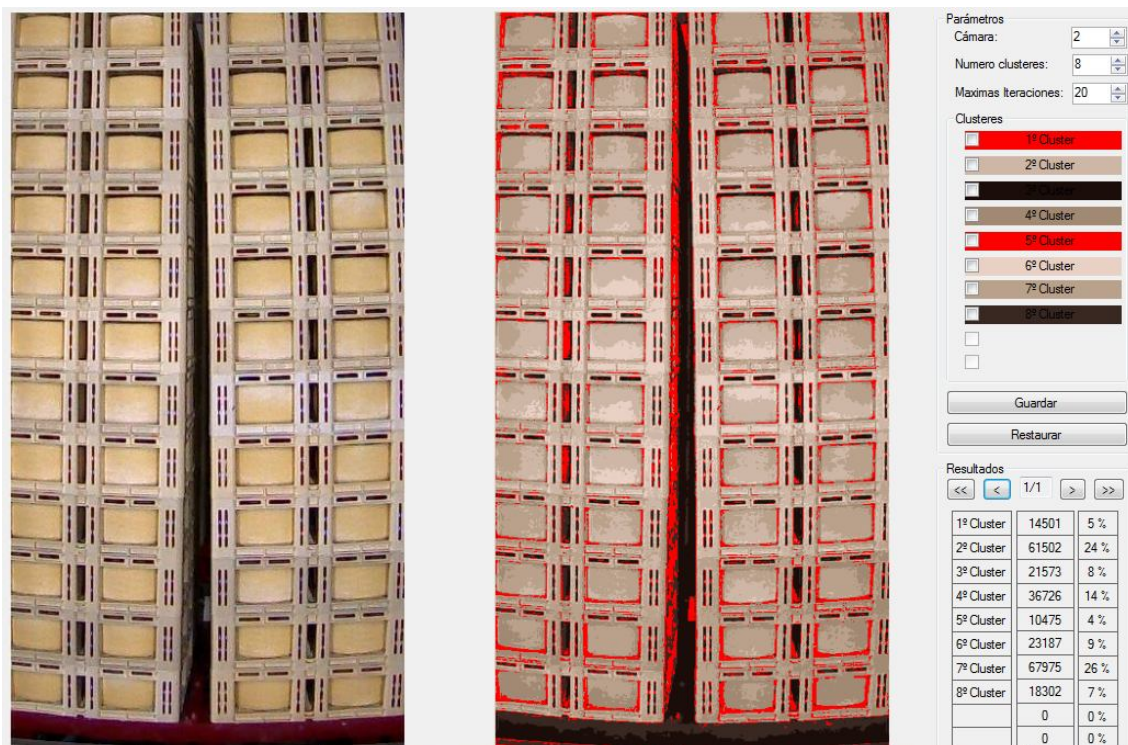


5-8 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 4 clusters

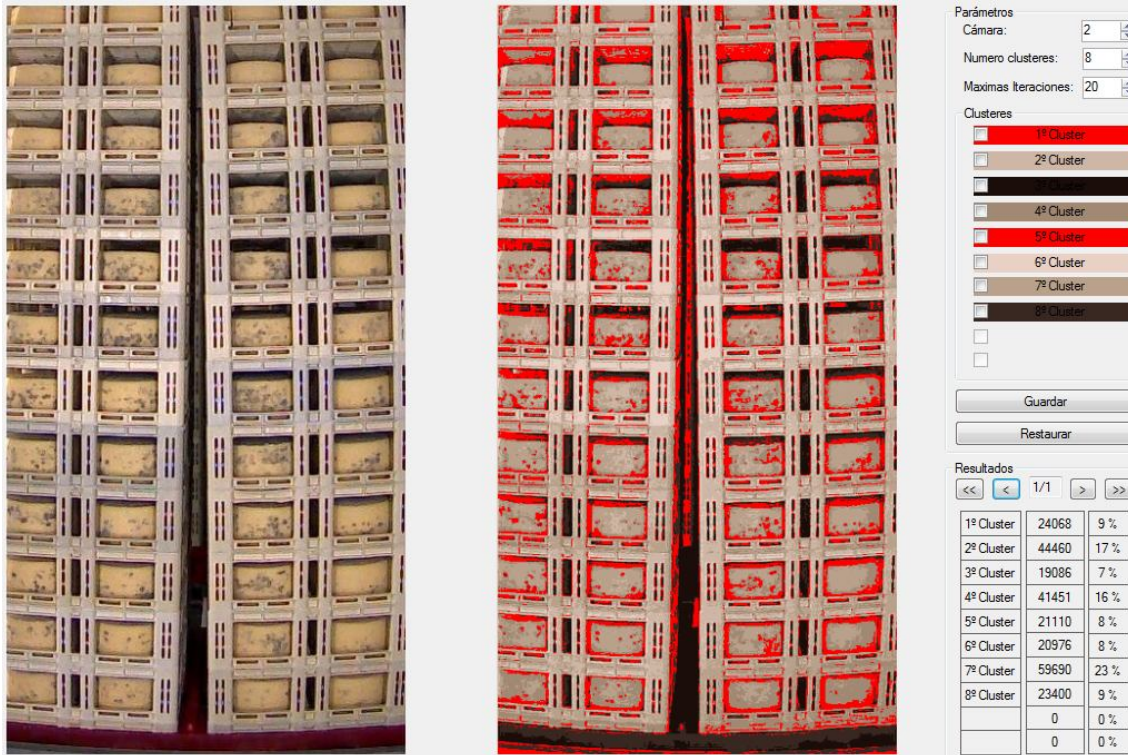
4º) Ejecución con parámetros: Nº clases =8, Máximas iteraciones = 20

Tras ejecutar la primera vez 20 iteraciones el algoritmo concluye, devolviendo una diferencia entre las iteraciones 19 y 20 de 173699, por lo que no converge. Como se ha indicado reiteradamente, dependiendo de los centroides que se usan para la

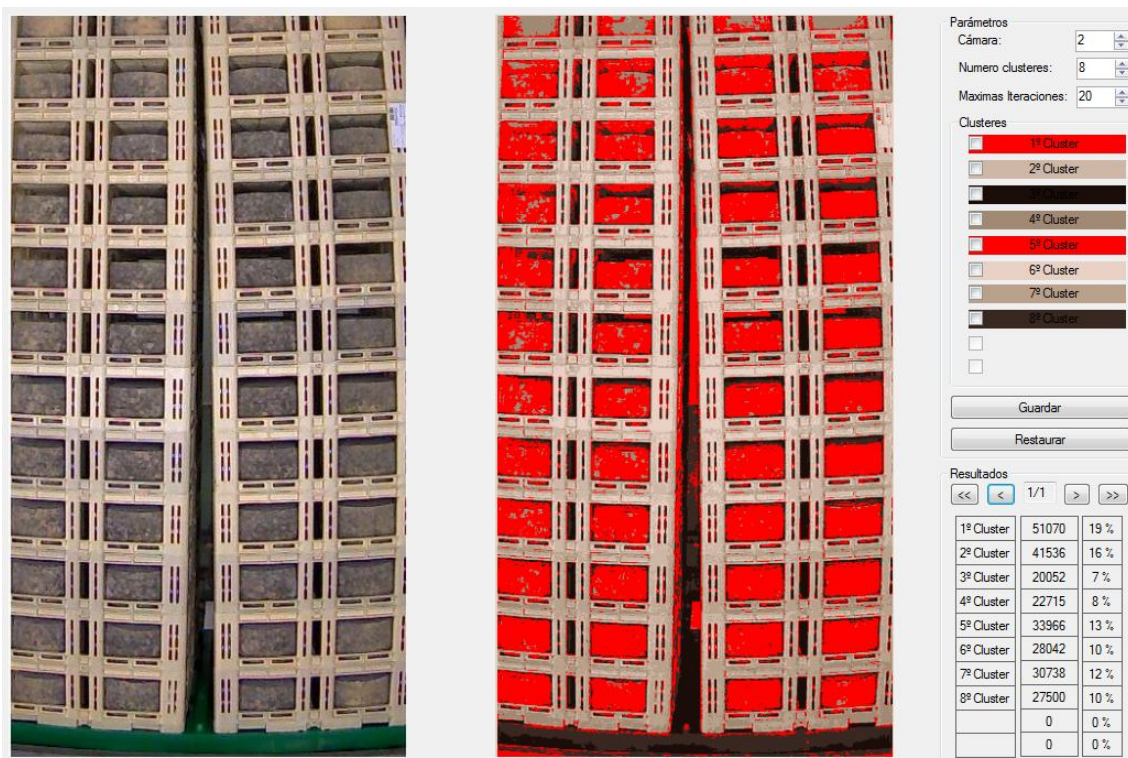
primera iteración del algoritmo determinará una aproximación más rápida del algoritmo hacia la solución. En nuestro caso como comentábamos en el punto 3.3, la elección de estos puntos es aleatoria, por lo que ejecutamos de nuevo el algoritmo con los mismos parámetros para comprobar el comportamiento. En esta ocasión el algoritmo converge en la iteración 18. Ayudándonos del experto quesero y el resultado de las 10 imágenes, le solicitamos que nos indique que *clusters* considera que representan el moho, indicando el 1º y el 5º como los más representativos. Observando las imágenes con poca o ninguna concentración de moho, como son la recogidas en la figura 5-9 y 5-10, comenta el experto que si sería posible obtener una mayor división para distinguir más los huecos de las cajas de los píxeles que realmente representan moho. Ante las 2 horas de ejecución del aprendizaje con un número de clases 8 y 20 iteraciones como máximo, se incrementa el número de clases a 10. Si revisamos las figuras 5-9, 5-10 y 5-11, podemos ver en rojo la representación del *cluster* número 1 y el número 5, considerados moho por el experto quesero. En la figura 5-9 que representa un palé sin moho, la suma de dichos *clústeres* agrupa el 9% de los píxeles de la imagen.



5-9 Resultado de imagen sin moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 8 clusters



5-10 Resultado de imagen con nivel medio de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 4 clusters



5-11 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 10 clusters

5º) Ejecución con parámetros: Nº clases =10, Máximas iteraciones = 20

Tras 19 iteraciones el algoritmo converge y termina, ya que la diferencia en la función objetivo de la 18ª iteración y la 19ª se sitúa por debajo de los límites de tolerancia. Ayudándonos de nuevo del experto quesero y con las 10 imágenes disponibles, le

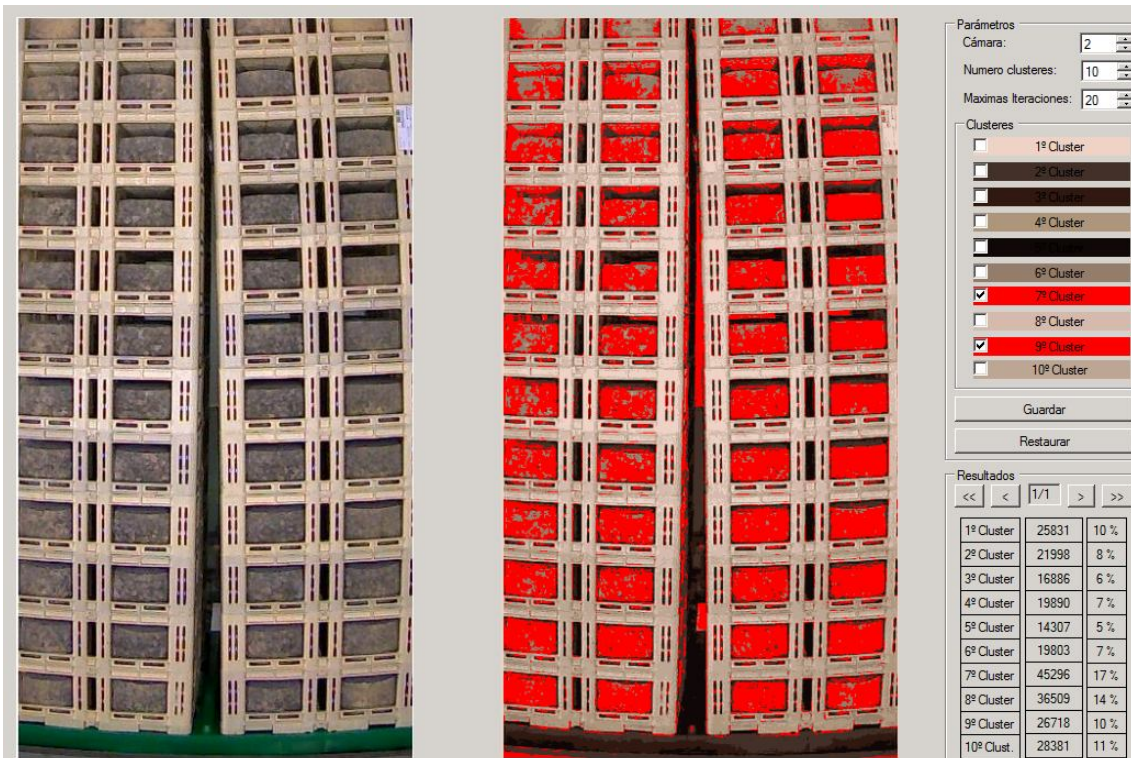
solicitamos que nos indique los *clusters* considerados representativos del moho, indicando los *clusters* 5 y 7. En la figura 5-13 y 5-14 se pueden ver cómo existen clases que representan las porciones de moho del producto en las cuatro imágenes sin incluir grandes porciones de sombras. Aunque el experto quesoero piensa en aumentar el número de divisiones, decidimos considerar una solución válida definitiva. Tomamos esta determinación al observar que el tiempo de ejecución del algoritmo con estos parámetros es superior a 3 horas, haciendo insostenible seguir aumentando los parámetros de prueba.



5-12 Resultado de imagen sin moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 8 clusters



5-13 Resultado de imagen con nivel medio de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 4 clusters



5-14 Resultado de imagen con nivel alto de moho de la fase aprendizaje mediante algoritmo Fuzzy K-means en 10 clusters

5.5 CLASIFICACIÓN

Tras llevar a cabo la fase de aprendizaje para cada una de las cuatro cámaras y conseguir ajustar los parámetros del algoritmo *Fuzzy K-means*, comprobaremos los resultados de la fase de clasificación de los dos algoritmos aplicando las muestras de la BC, tal y como se explica en el punto 5.2. En la figura 5-15 se muestra la tabla de resultados obtenidos de la ejecución de los dos métodos descritos en el apartado 3, junto a la información de los 25 palés, 100 imágenes, que contiene la BC descrita en el ANEXO I. Cada fila representa el análisis de un palé, mostrando los valores de moho obtenidos por cada método para cada una de las cuatro imágenes. En las columnas azules, aparece el valor promedio de las cuatro imágenes, expresado como el indicador de moho del palé.

Indicador nivel de moho															
Nº de Palé	Base de Conocimiento					Metodo Otsu					Fuzzy K-means(K=10)				
	Imagen 1	Imagen 2	Imagen 3	Imagen 4	Indicador Moho Palé	Imagen 1	Imagen 2	Imagen 3	Imagen 4	Indicador Moho Palé	Imagen 1	Imagen 2	Imagen 3	Imagen 4	Indicador Moho Palé
1	66%	66%	44%	10%	47%	27%	29%	20%	22%	25%	36%	26%	30%	16%	27%
2	0%	0%	0%	0%	0%	7%	0%	17%	16%	10%	16%	10%	24%	21%	18%
3	33%	33%	66%	22%	39%	19%	19%	19%	14%	18%	32%	18%	33%	25%	27%
4	10%	10%	0%	0%	5%	19%	13%	12%	15%	15%	24%	18%	32%	28%	26%
5	11%	11%	11%	10%	11%	18%	14%	14%	15%	15%	27%	18%	47%	32%	31%
6	0%	0%	0%	0%	0%	17%	16%	11%	14%	15%	21%	15%	41%	34%	28%
7	11%	11%	22%	11%	14%	20%	18%	18%	0%	14%	21%	16%	29%	31%	24%
8	11%	11%	22%	11%	14%	20%	15%	15%	13%	16%	30%	18%	41%	31%	30%
9	0%	11%	11%	0%	6%	7%	17%	16%	15%	14%	20%	14%	32%	26%	23%
10	0%	22%	22%	0%	11%	11%	19%	16%	15%	15%	23%	16%	32%	24%	24%
11	0%	22%	44%	11%	19%	7%	11%	17%	16%	13%	20%	16%	25%	20%	20%
12	66%	66%	0%	0%	33%	16%	21%	14%	17%	17%	35%	24%	19%	15%	23%
13	100%	100%	100%	100%	100%	37%	32%	42%	47%	40%	40%	39%	42%	45%	42%
14	44%	44%	66%	44%	50%	14%	25%	31%	31%	25%	32%	24%	40%	40%	34%
15	66%	44%	66%	66%	61%	24%	12%	36%	30%	26%	40%	19%	39%	37%	34%
16	22%	66%	44%	22%	39%	17%	18%	12%	18%	16%	23%	23%	37%	35%	30%
17	66%	22%	66%	11%	41%	18%	16%	22%	21%	19%	27%	15%	18%	29%	22%
18	0%	0%	0%	0%	0%	7%	0%	13%	19%	10%	38%	16%	18%	13%	21%
19	22%	66%	66%	22%	44%	15%	27%	20%	19%	20%	29%	35%	35%	12%	28%
20	0%	0%	44%	0%	11%	21%	20%	27%	22%	23%	32%	23%	34%	16%	26%
21	66%	100%	66%	66%	75%	20%	34%	30%	31%	29%	35%	34%	36%	14%	30%
22	66%	100%	100%	22%	72%	27%	39%	36%	21%	31%	37%	34%	38%	12%	30%
23	66%	100%	100%	22%	72%	22%	39%	38%	27%	32%	35%	34%	40%	15%	31%
24	44%	100%	66%	0%	53%	20%	37%	26%	16%	25%	29%	34%	37%	10%	28%
25	66%	100%	66%	66%	75%	26%	35%	30%	30%	30%	42%	34%	35%	17%	32%

5-15 Tabla de resultados de la aplicación de los dos métodos Fuzzy K-means y umbralización Otsu a las 100 imágenes de los 25 palés de estudio, recogidos en la BC

Para una mayor clarificación y comprensión de los resultados mostrados en la tabla 5-15, a continuación realizaremos una representación gráfica de los mismos, agrupados por la cámara de procedencia de las imágenes. En el eje de abscisas se representan los palés analizados del 1 al 25. En el eje de ordenadas aparece el porcentaje de moho detectado para cada uno de los dos métodos aplicados (Otsu y K-means) frente

al *ground-truth*, según el experto quesero. Para cada uno de estos grupos comentaremos los puntos críticos y de discrepancia entre los resultados obtenidos por los métodos y la BC, mostrando para ello tres imágenes por figura: original, resultado de aplicar el método Fuzzy K-mean, donde se resaltará en tonos rojizos lo que según el sistema considera moho y resultado de aplicar el método de Otsu, donde aparece en negro lo que se considera moho. A continuación mostramos en las figuras 5-16 y 5-17, dos ejemplos de imágenes con los resultados de moho detectados por los consabidos métodos aplicados K-means y Otsu.



5-16 Resultados del método Fuzzy k-means y método Otsu cámara 2, palé 16

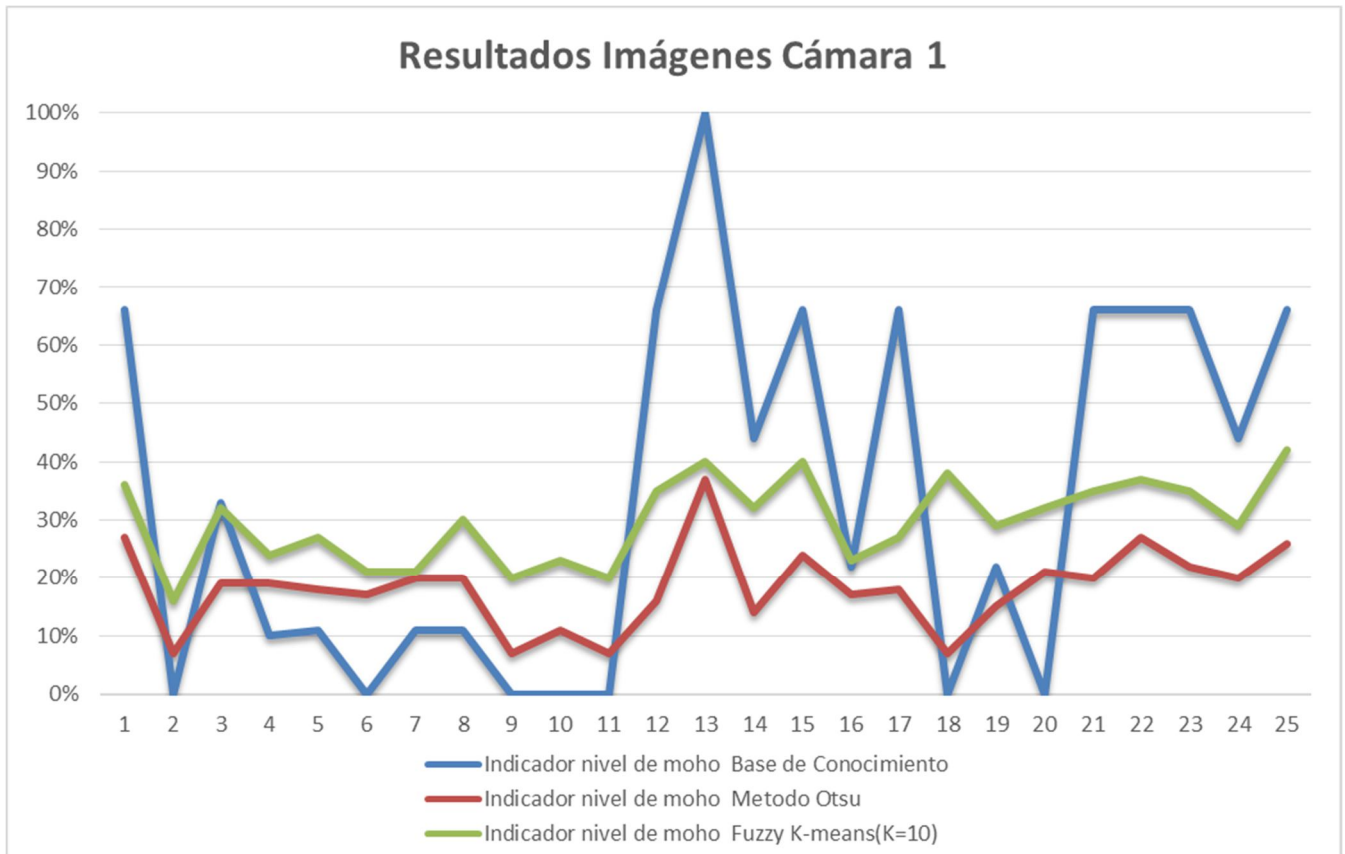


5-17 Resultados del método Fuzzy k-means y método Otsu cámara 3, palé 3

Imágenes procedentes de la cámara 1

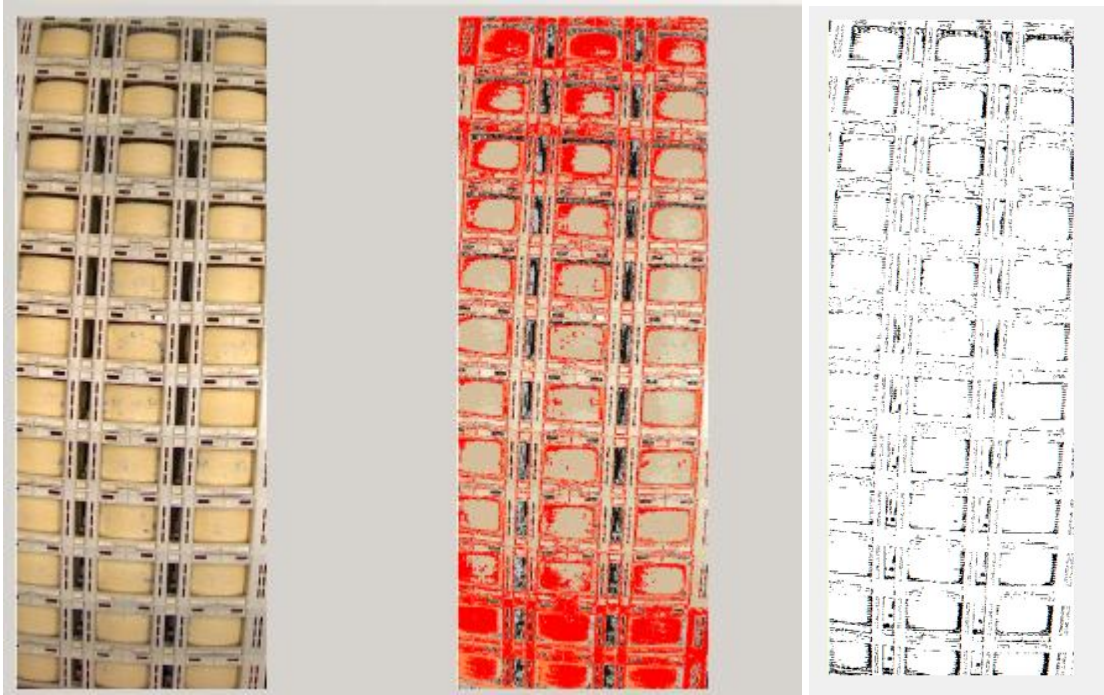
En la figura 5-18 se muestra la representación de los valores de moho obtenidos para las imágenes procedentes de la cámara 1, enfrentando los datos del *ground truth*

basados en el criterio del maestro quesero, el método de umbralización mediante Otsu y el clasificador basado en K-means.



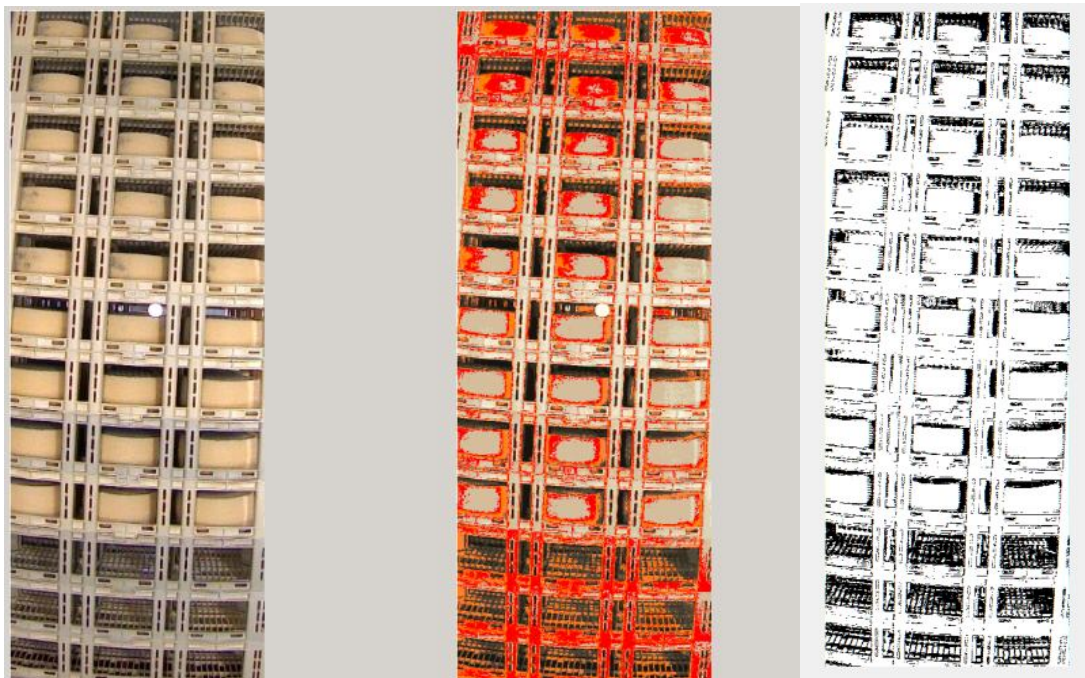
5-18 Valores del indicador de moho para cada una de las 25 imágenes de la Cámara 1.

Observando los perfiles de las tres series, con especial énfasis en las crestas y valles, se puede ver cómo ambos métodos se comportan de forma satisfactoria, ya que aunque se obtienen valores distintos a los existentes en la BC, recogen de forma similar las variaciones de moho, al tener las gráficas el mismo comportamiento. Podemos observar la existencia de una diferencia entre el método Fuzzy K-means ($k=10$) y el método de Otsu en el palé 18. El método Fuzzy K-means ($k=10$) se comporta de forma errónea debido a un problema en la iluminación, proveniente de falta de luz en la parte superior e inferior del palé debido a la temperatura de las lámparas, factor que le afecta en menor medida al método implementado mediante umbralización Otsu. En la figura 5-19 se muestra en la parte izquierda el resultado del algoritmo Fuzzy K-means, resaltado en colores rojizos sobre la imagen los considerados como moho. En contra punto, podemos ver en la parte de la derecha de la figura el resultado de aplicar el método Otsu (imagen blanco y negro), en negro se resaltan los pixeles considerados como moho.



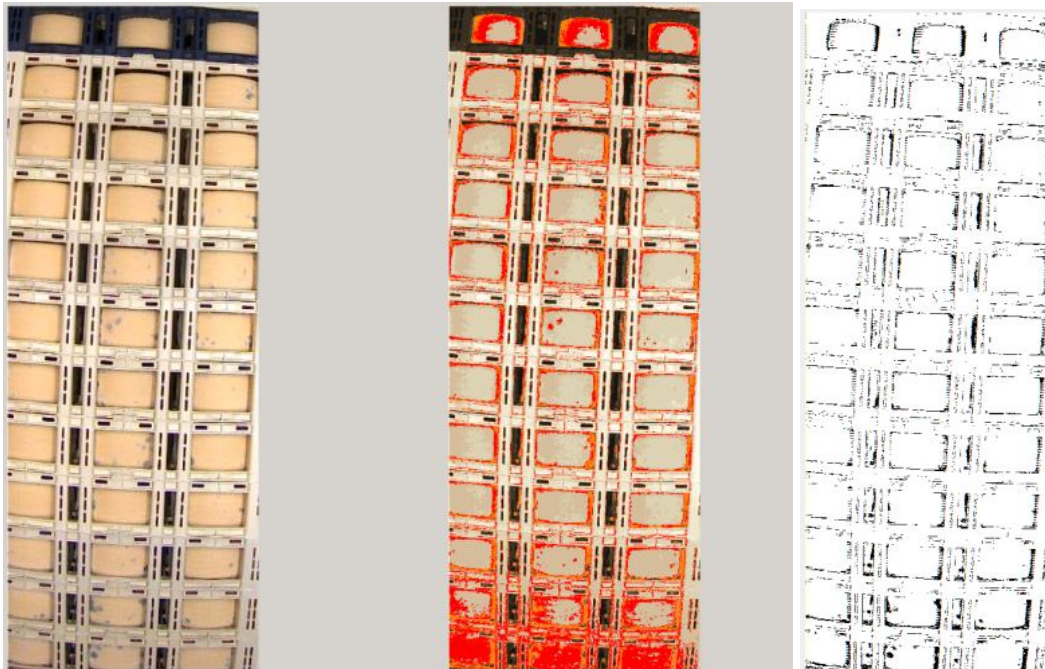
5-19 Resultados del método Fuzzy k-means y método Otsu con un problema de iluminación.

Por otro lado, según la figura 5-18, observamos una gran diferencia en el palé 20, donde ambos métodos fallan, al tratarse este un palé con varias cajas vacías de quesos en su interior. Figura 5-20 se puede ver en colores rojizos el resultado obtenido por el algoritmo Fuzzy K-means y en blanco y negro el resultado del método Otsu. Aunque en la zona de ausencia de quesos ambos fallan al considerar mohos lo que no es, se puede observar cómo en el caso de Fuzzy K-means, se tiene el mismo problema que en los modelos anteriores, la falta de la iluminación en la parte superior e inferior del palé origina que el método falle, considerando mohos píxeles que no lo son.



5-20 Resultados del método Fuzzy k-means y método Otsu con un palé con cajas vacías

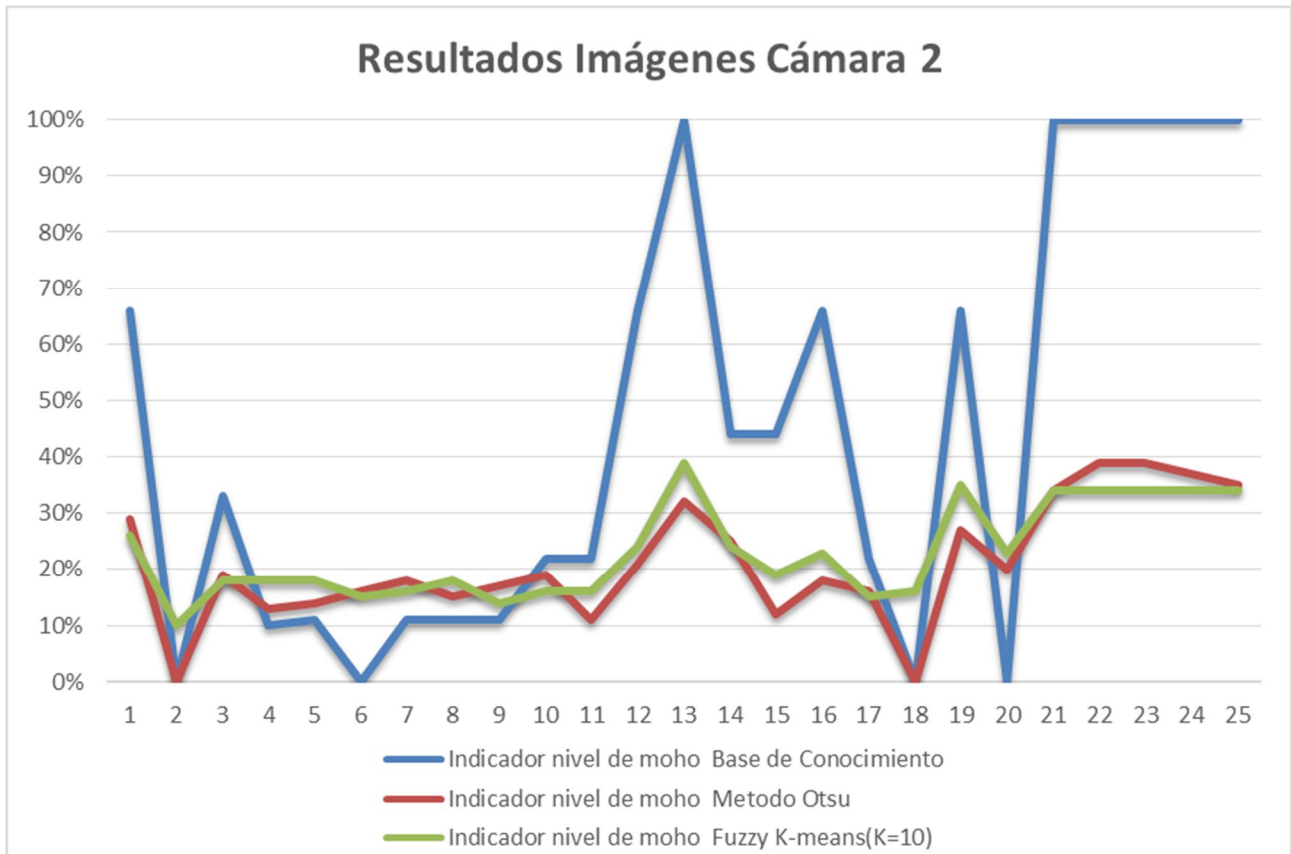
En el caso de los palés 9,10 y 11, que vienen conformados con una de las cajas en color negro, ambos algoritmos se comportan de forma satisfactoria al no considerar moho los píxeles que las conforman. En la figura 5-21 podemos ver en la imagen con colores rojizos cómo la primera caja que es de color negro, nos es considerada moho, de mismo modo que no lo es por el método Otsu en la imagen en blanco y negro.



5-21 Resultados del método Fuzzy k-means y método Otsu con un palé con cajas de color negro.

Imágenes procedentes de la cámara 2

En la figura 5-22 mostramos los valores de moho obtenidos para las imágenes procedentes de la cámara 2, enfrentando los datos del ground truth, umbralización Otsu y K-means.



5-22 Valores de indicador de moho para cada una de las 25 imágenes de la Cámara 2.

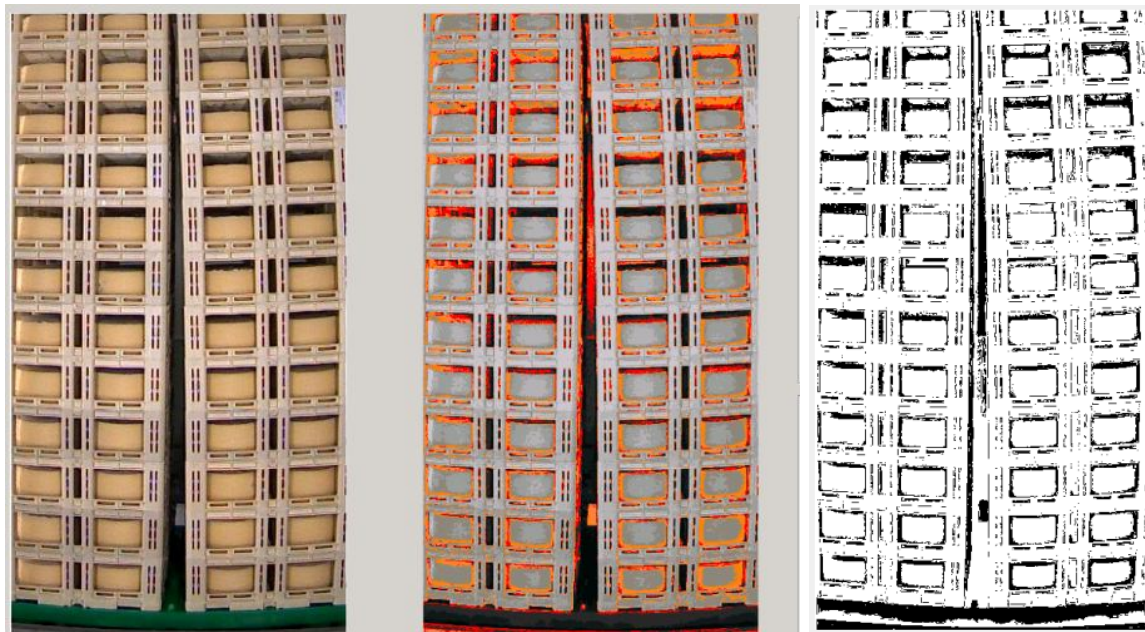
Visualizando los perfiles de las gráficas observamos que las tres series muestran comportamientos similares (crestas y valles). De igual modo que en el caso anterior, podemos observar discrepancias en el comportamiento del palé número 20, compuesto con cajas sin queso, al indicarse en ambos métodos un valor de moho elevado cuando en la BC se refleja como nulo. En la figura 5-23 se puede ver con detalle, la imagen con colores rojizos, clases consideradas moho, tras aplicar el método Fuzzy K-means. En la imagen en blanco y negro, resultado del método Otsu, se pueden ver los píxeles negros considerados como moho.



5-23 Resultados del método Fuzzy k-means y método Otsu

con un palé con cajas vacías

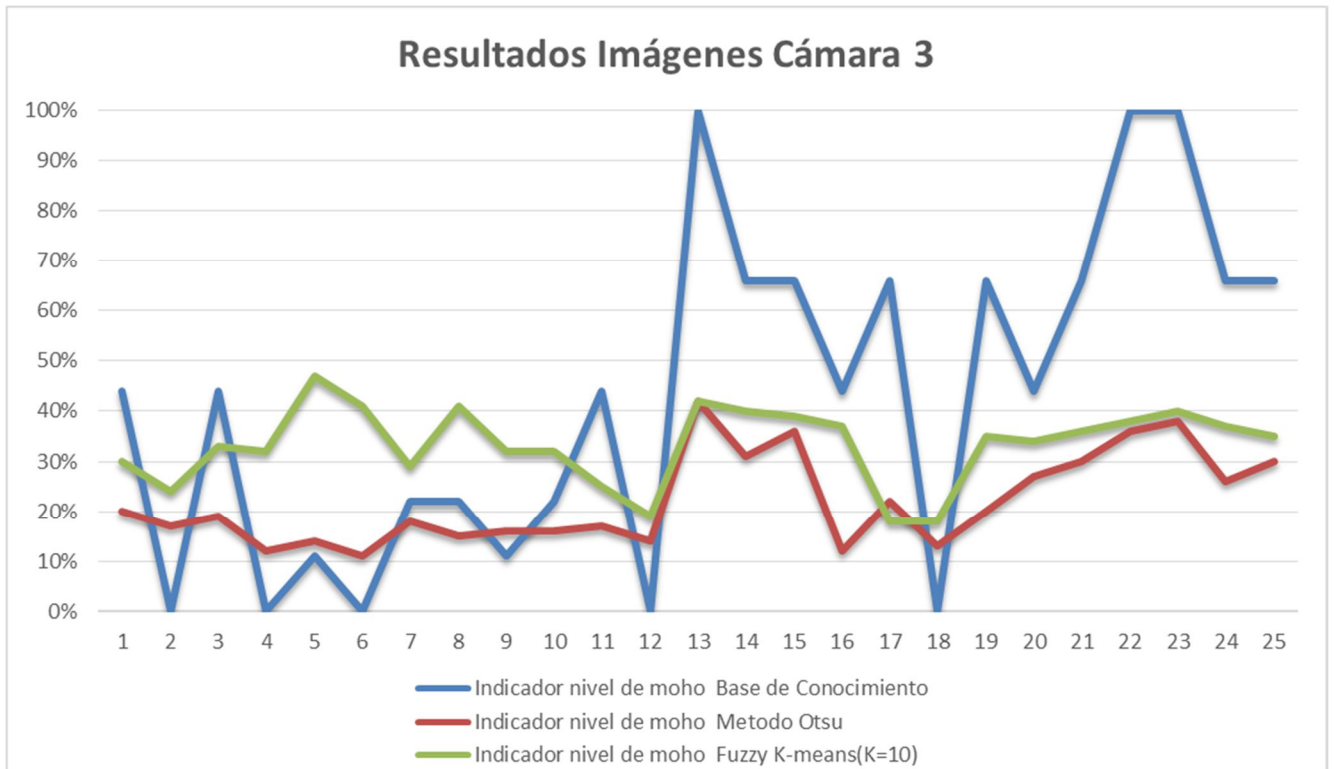
Por otro lado, llama la atención el palé número 6, en el que siendo nulo el valor de moho recogido en la BC, ambos métodos dan como resultado un valor elevado de moho. Observando la imagen con detalle, se determina que dicho fenómeno se debe a un problema en el cuadro del palé dentro de la imagen, debido a la sincronización entre la detección de presencia por parte de la fotocélula y el disparo de la cámara al obtener la foto. En la imagen 5-24 se puede ver en la parte central con colores rojizos, el resultado del método Fuzzy K-means, considerando píxeles de la base del palé como moho. En la imagen en blanco y negro, resultado del método Otsu, se observa el mismo efecto donde los píxeles negros son considerados como moho.



***5-24 Resultados del método Fuzzy k-means y método Otsu
con un palé descuadrado***

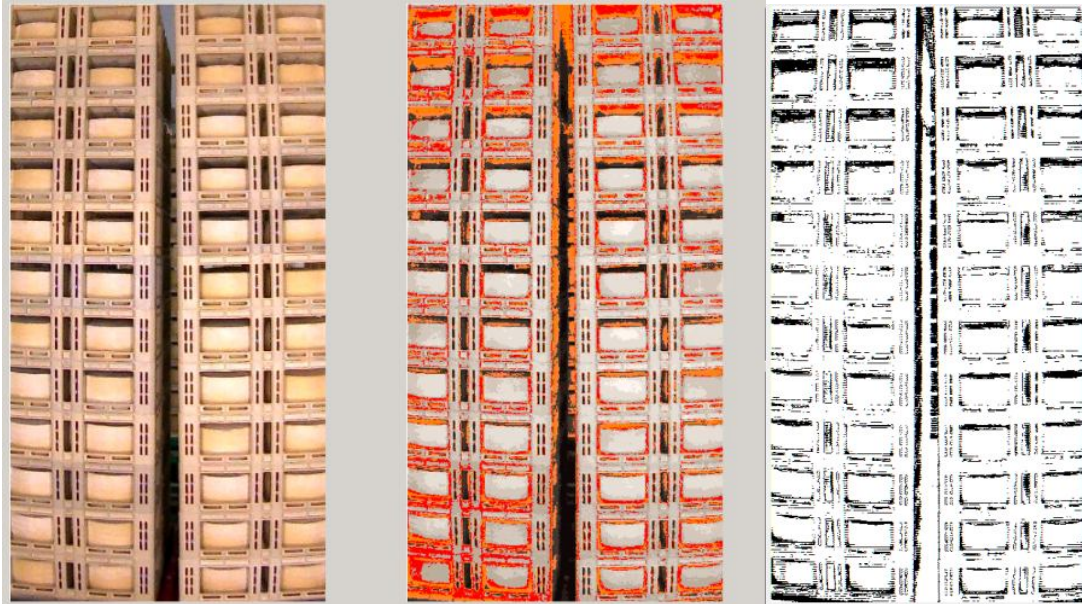
Imágenes procedentes de la cámara 3

En la figura 5-25 mostramos los valores de moho obtenidos para las imágenes procedentes de la cámara 3, enfrentando los datos del ground truth según el criterio del maestro quesero, y los métodos de umbralización mediante el método Otsu y de clasificación K-means.

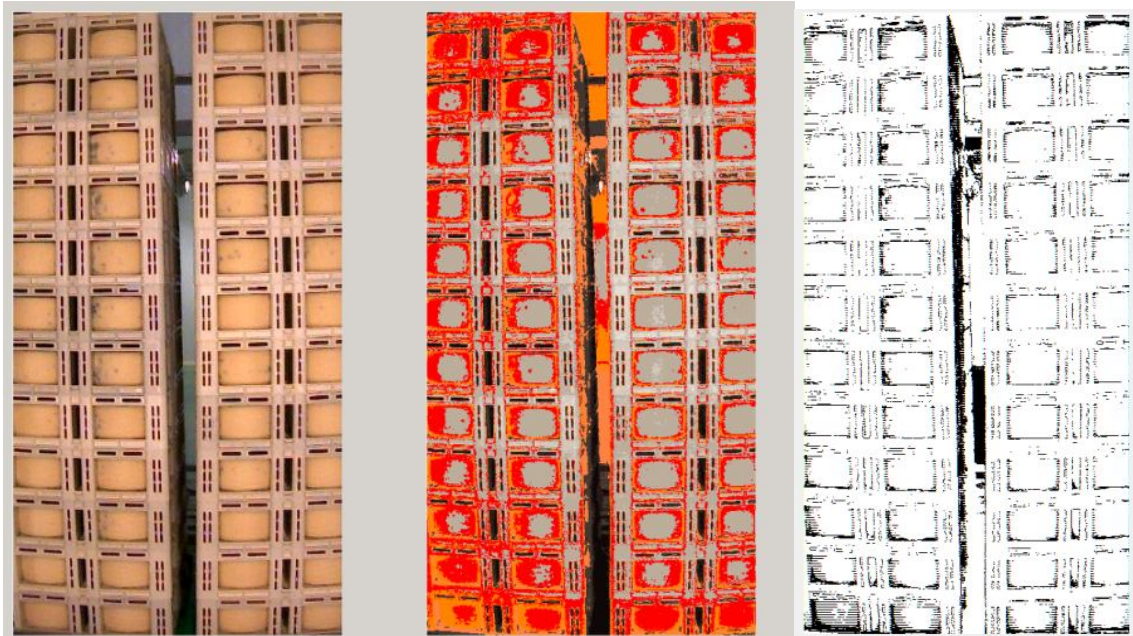


5-25 Valores de indicador de moho para cada una de las 25 imágenes de la Cámara 3.

En la serie verde, perfil del método Fuzzy K-Means, en concreto en las muestras obtenidas de los primeros palés número 2, 4 y 6, la diferencia con el valor de BC, hace pensar en un defecto en la fase de aprendizaje para esta cámara en concreto. Se puede observar cómo el método de Otsu, se comporta mejor aunque en los palés con poco índice de moho, como es el caso del número 2, se consideran como moho múltiples píxeles que no los son. Esto es consecuencia de los problemas derivados de la iluminación y la existencia de sombras en estos palés. En las figuras 5-26 y 5-27 se pueden ver en color rojizo las imágenes resultantes de la aplicación del método Fuzzy K-means. Al no haber presencia de moho en la imagen original, son muchos los píxeles con tonos rojizos considerados moho. En las imágenes en blanco y negro, resultado del método Otsu, se puede observar cómo aun imputando píxeles como moho, en negro, el resultado es mejor.



5-26 Resultados del método Fuzzy k-means y método Otsu en el palé número 2 con problema en la fase de aprendizaje



5-27 Resultados del método Fuzzy k-means y método Otsu en el palé número 6, con problema en fase de aprendizaje.

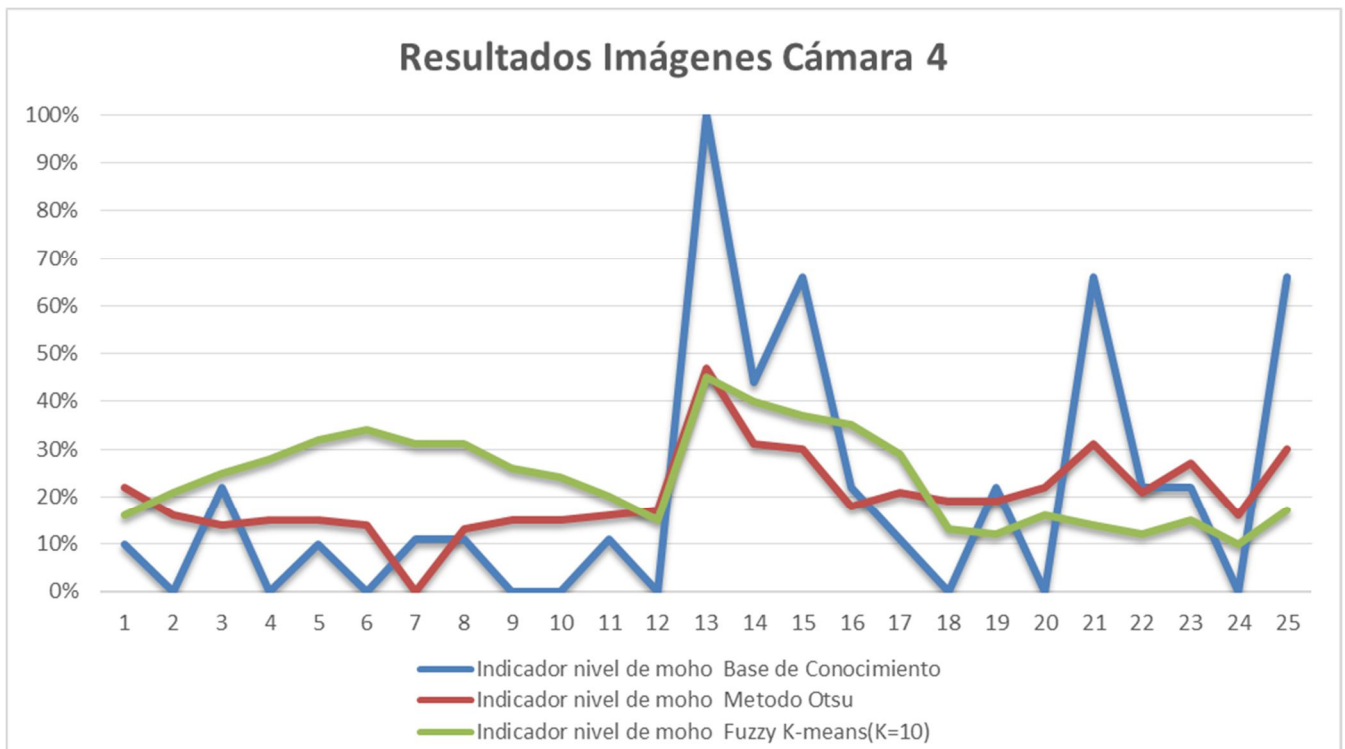
Comentar el resultado obtenido en el palé número 13 que ambos métodos han obtenido el mismo valor. En la imagen de la figura 5-28 podemos observar en tonos rojizos el resultado de aplicar el método Fuzzy K-means y en blanco y negro el resultado obtenido mediante el método de Otsu.



5-28 Resultados del método Fuzzy k-means y método Otsu en el palé número 13, con alto valor de moho.

Imágenes procedentes de la cámara 4

En la figura 5-29 mostramos los valores de moho obtenidos para las imágenes procedentes de la cámara 4, enfrentando los datos del ground truth, según el maestro quesero, y los métodos de umbralización mediante el método Otsu y K-means.



5-29 Valores de indicador de moho para cada una de las 25 imágenes de la Cámara 4.

Observando la gráfica, aparece un comportamiento similar al caso anterior obtenido con la cámara 3. El perfil errático del método Fuzzy K-means en la mayoría de las imágenes de los primeros palés. 4, 5, 6, 9 y 10, donde el valor de moho es bajo según

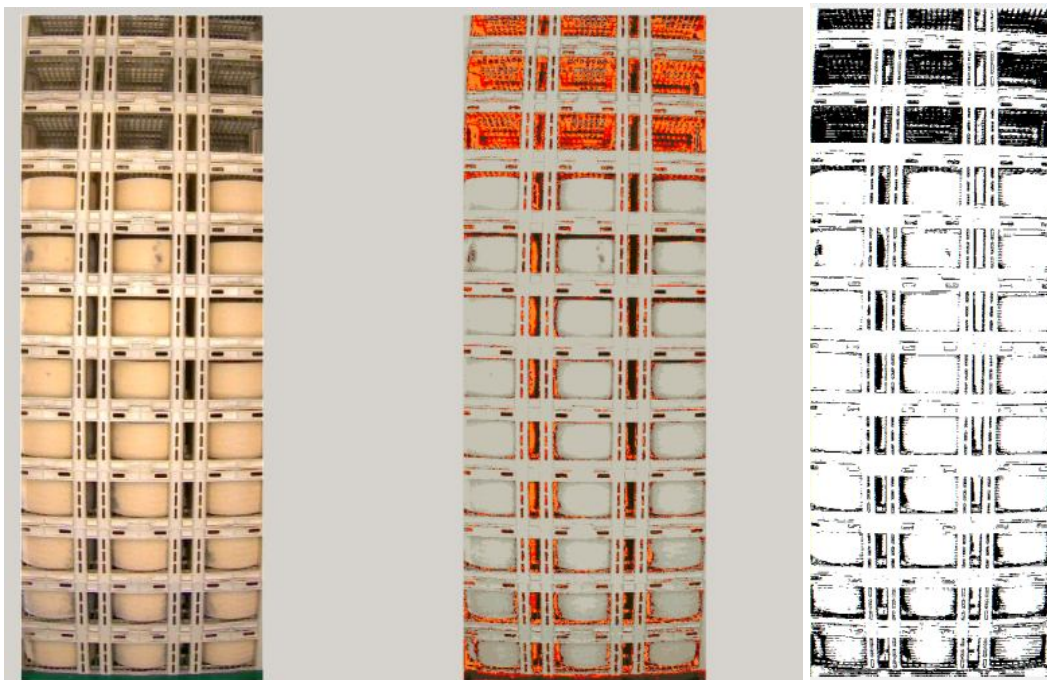
la BC, vuelven a poner de manifiesto una mala fase de aprendizaje para esta cámara. En la figura 5-30 se puede ver en la imagen con colores rojizos el resultado obtenido con el método Fuzzy K-Means, en blanco y negro el resultado del método Otsu.



5-30 Resultados del método Fuzzy k-means y método Otsu en el palé número 4.

En cuanto al método Otsu vemos que no sigue el perfil de forma tan precisa como en las cámaras 1 y 2.

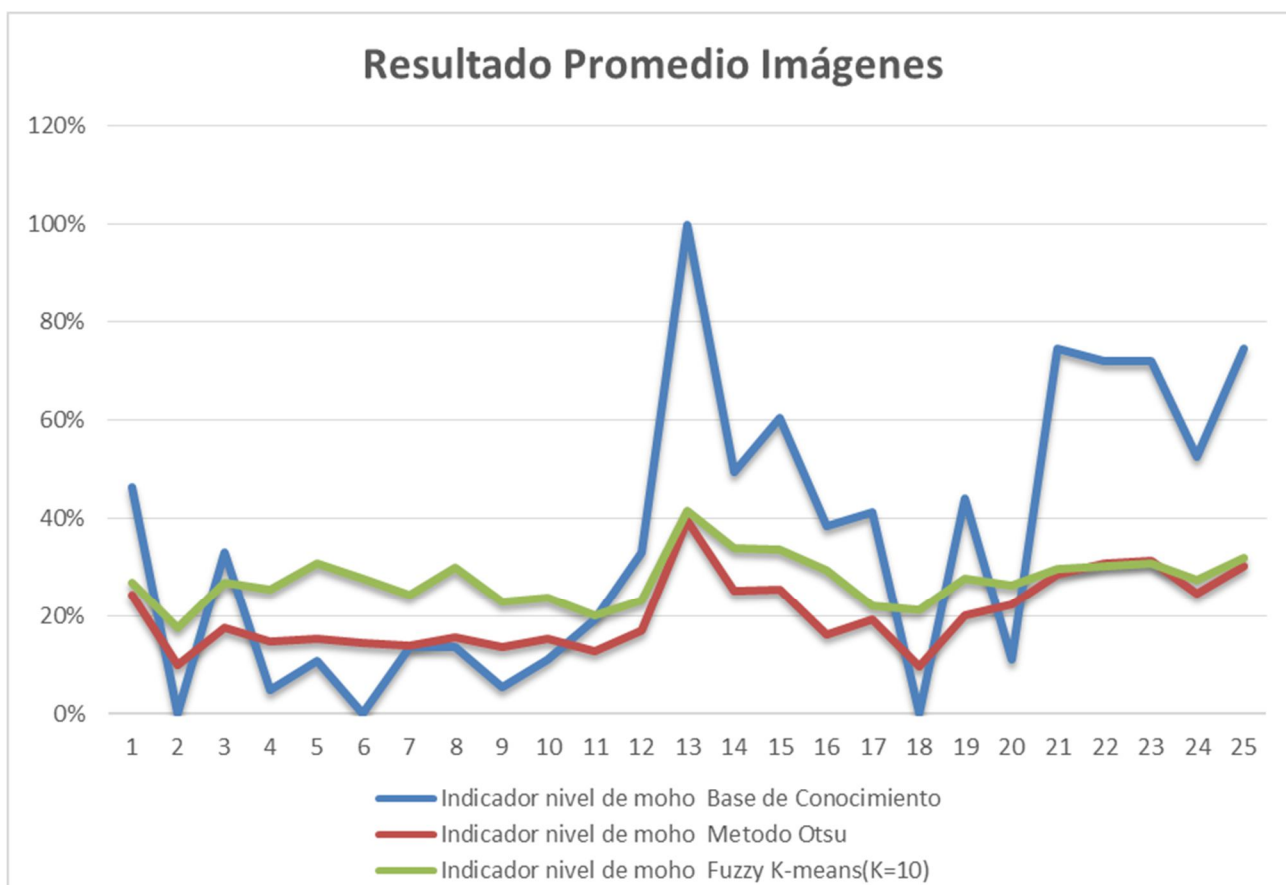
De nuevo falla con el palé número 20 por falta de quesos en las cajas. En la figura 5-31 se observa en colores rojizos el resultado del método Fuzzy K-means y en la imagen en blanco y negro las imágenes obtenidas como resultado de aplicar el método Otsu, resaltando en negro los píxeles considerados como moho, cuando realmente hay ausencia de queso.



5-31 Resultados del método Fuzzy k-means y método Otsu en el palé número 20.

Resultado promedio de las imágenes

En la figura 5-32 mostramos los valores de moho obtenidos del promedio de los resultados de indicador de moho de las 4 imágenes que componen el análisis de cada palé, mostrado en las gráficas anteriores, enfrentando de nuevo los datos al promedio de los datos recogidos por la BC.



5-32 Valores de indicador de moho para cada una de las 25 imágenes, promediado de las imágenes procedentes de las 4 cámaras.

Tras estudiar por separado el comportamiento de cada una de las cámaras, representamos el valor promedio de las cuatro cámaras. Los perfiles ofrecidos por ambos métodos son muy similares al que se extrae de la Base de Conocimiento, podríamos decir que se detecta las variaciones de moho siempre y cuando los valores de moho sean altos, mostrando cadencias en las imágenes con valores de moho bajos, achacables a falta de iluminación y sombras generadas. Además, hemos detectado problemas a la hora de analizar palés con cajas con ausencia de queso, como es el caso del palé número 20. Por otro lado, el funcionamiento de ambos algoritmos en el caso de existencia de cajas de color negro dentro del palé ha sido satisfactorio, al no considerar éstas como moho, palés números 9,10 y 11 En el siguiente punto recogeremos las conclusiones y trabajos futuros que extraemos de los resultados presentados en este apartado.

6 CONCLUSIONES Y TRABAJO FUTURO

6.1 INTRODUCCIÓN

En este capítulo se hace necesario plantear un balance global sobre las líneas de investigación seguidas, que desembocan en las conclusiones expuestas. Por otra parte, se abren una serie de perspectivas de futuro en relación con los temas de investigación abordados y sus posibilidades de ampliación futuras.

6.2 CONCLUSIONES

En este trabajo se ha realizado una revisión de dos métodos de segmentación de imágenes con el fin de analizar sus posibilidades en la aplicación de detección de moho en el queso de forma automática. El interés se ha centrado en la aplicación sobre un sistema real, hasta el momento no aplicado en este tipo de situaciones, lo que constituye un reto importante en este sentido.

Se ha utilizado un conjunto de imágenes capturadas por el sistema dando lugar a una base de datos o *ground-truth*, específica del problema planteado, por no existir información anterior al respecto, con el fin de verificar su eficacia, al disponer de imágenes de prueba que permiten cuantificar la calidad de los métodos de segmentación aplicados.

Tras el estudio realizado en el capítulo cinco, se exponen los resultados obtenidos por los métodos descritos en las secciones 3.2 y 3.3. A continuación se exponen las conclusiones más relevantes derivadas de los mismos.

Los dos métodos muestran resultados satisfactorios, ya que permiten detectar la presencia de niveles de moho en los palés de forma desasistida. De los dos métodos aplicados el de umbralización basado en Otsu, ha mostrado su mayor eficacia por los mejores resultados obtenidos, siendo más robusto en lo que a condiciones lumínicas se refiere. Destacar el mejor comportamiento del método Fuzzy K-means, en situaciones de baja iluminación y de palés completados con cajas con ausencia de queso. Resaltar como características positivas que el método basado en umbralización por el método Otsu, además de ser totalmente automático, es decir que no requiere aprendizaje supervisado como el Fuzzy K-means, que es el método más sencillo de aplicar y con el que se obtienen mejores tiempos computacionales. El algoritmo Fuzzy K-means posee la fase de aprendizaje donde requiere grandes periodos de tiempo, en función de los parámetros configurados y el número de imágenes que se utilice. Además, la existencia de esta fase y la validación de la misma por parte de un experto quesero, origina que un posible punto de fallo sea la calibración, siendo arrastrado a la fase de clasificación como hemos podido ver.

En la fase de clasificación o decisión ambos métodos muestran tiempo de respuesta válidos por ser inferior el tiempo de clasificación a los quince segundos marcados por el proceso de paso del palé por la instalación. Por otro lado, la necesidad de supervisión, provoca una dependencia del experto que dirige la aplicación, resultando un método con un alto grado de dependencia. El principal problema que plantean ambos métodos al basarse en segmentación por color, es la necesidad de aplicar una iluminación uniforme sobre todos los quesos que conforman el palé.

Podría pensarse en eliminar parte de la variación de iluminación aplicando algún tipo de filtro. Si bien en este caso el problema estriba en la naturaleza redonda del queso y la estructura de las cajas de palé, originando siempre ruido en forma de sombras.

6.3 TRABAJO FUTURO

La primera línea de investigación abierta, es la derivada de los resultados obtenidos, sintetizados en la sección previa de conclusiones. En efecto, la investigación en los entornos reales debe continuar por la mejora de la calidad de la imagen, tanto en lo que respecta al sistema de iluminación de los palés, como a la alineación del sistema de las cámaras para obtención de imágenes similares, así como en relación a la eliminación de todo tipo de interferencias que existen dentro de la imagen con el fin de aislar la superficie de interés, los quesos. En esta línea se plantean las siguientes propuestas de investigación futuras:

- 1) Eliminación de las estructuras que conforman las cajas.
- 2) Eliminar la base del palé o al menos los colores de estas.
- 3) Eliminación del hueco entre las dos columnas de cajas que conforman el palé.

Se planteó como objetivo el estudio de la validación de los métodos, para lo cual se construyó una base de conocimiento compuesta por 100 imágenes reales, 25 por cada una de las cámaras, evaluadas por un experto quesero. Dichas imágenes, aunque obtenidas a partir de un entorno real, no plantean la totalidad de la casuística que puede existir realmente. Por lo que una línea de trabajo futuro, pasaría por la generación de una más completa base de datos utilizable tanto en la fase de aprendizaje de Fuzzy K-means como para el análisis de los resultados obtenidos por los dos métodos implementados.

Desarrollo y planteamiento de nuevos métodos de segmentación para resolver el problema que aquí se plantea. Existen una amplia bibliografía con múltiples métodos de segmentación probados, que se pueden aplicar a la detección de moho en los palés de queso y analizar sus resultados frente a los que aquí se han planteado.

7 BIBLIOGRAFÍA

1. Alatas, B. & Akin, E. (2005). FCACO: Fuzzy Classification Rules Mining Algorithm with Ant Colony Optimization.. In L. Wang, K. C. 0001 & Y.-S. Ong (eds.), ICNC (3) (p./pp. 787-797), : Springer. ISBN: 3-540-28320-X .
2. Arthur, D., Manthey, B. & Roeglin, H. (2009). K-means has Polynomial Smoothed Complexity. Arxiv preprint arXiv:0904.1113, .
3. Bezdek, James C.; Ehrlich, Robert; Full, William (1984). FCM: The Fuzzy c-Means Clustering Algorithm. Computers & Geosciences 10 (2-3): pp. 191-203.
4. Duda, R. O., P. E. Hart and D. G. Stork (2001). Pattern Classification and Scene Analysis. John Wiley & Sons. New York.
5. Fakhrahmad S. M., Zare A. and Jahromi M. Z. (2007). Constructing Accurate Fuzzy Rule-Based Classification Systems Using Apriori Principles and Rule-Weighting". LNCS 4881, pp. 547–556.
6. Forgy, E. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics, 21, 768-780.
7. Gonzalez, R. C., Woods, R. E. (2008). Digital image processing. Upper Saddle River, NJ: Prentice-Hall. ISBN: 0-201-18075-8
8. Grzymala-Busse, J. W., Stefanowski, J. & Wilk, S. (2004). A Comparison of Two Approaches to Data Mining from Imbalanced Data.. In M. G. Negoita, R. J. Howlett & L. C. Jain (eds.), KES (p./pp. 757-763), : Springer. ISBN: 3-540-23318-0
9. Hartigan, J. (1975). Clustering Algorithms. New York, NY: John Wiley & Sons.
10. Jain, A. K., Murty, M. N. & Flynn, P. J. (1999). Data clustering: a review. ACM Computing Surveys, 31, 264-323.
11. Jiménez, M. (2008). Introducción al tratamiento digital y clustering de imágenes. Universidad de Alcalá. Madrid.
12. Lloyd, S. P. (1957). Least square quantization in PCM. Bell Telephone Laboratories Paper. Publicado mucho más tarde en la revista: Lloyd., S. P. (1982). Least squares quantization in PCM. IEEE Transactions on Information Theory 28 (2): pp. 129–137. doi:10.1109/TIT.1982.1056489.
13. MacQueen, J. B. (1967). Some Methods for Classification and Analysis of MultiVariate Observations. In L. M. L. Cam & J. Neyman (eds.), Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability (p./pp. 281-297), : University of California Press.
14. Maravall, D. (1993). Algoritmos de agrupación de clases: clustering. RA-MA. Barcelona.
15. Otsu, N. (1979). A Thresholding Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, New York 9.1979, S.62–66. ISSN 1083-4419
16. Ortega, J. A. Foster, W., & Ortega, R. (2002). Definición de Sub-Roles para la Silvicultura de Precisión: Una aplicación del método Fuzzy Kmeans. Santiago, Chile: Facultad de Agronomía e Ingeniería Forestal, Pontificia Universidad Católica de Chile, Casilla 306-22.
17. Oliva i Cuyás, F., de Cáceres Ainsa, M., Font Castell, X., & Cuadras Avellana, C. M. (Noviembre de 2001). Contribuciones desde una perspectiva basada en proximidades a al Fuzzy K-means Clustering. Jaen, Andalucía España: XXVI Congreso Nacional de Estadística e Investigación Operativa: Úbedabeda, Universidad de las Riosa.
18. Pajares, G. and Cruz, J.M (2008). Visión por Computador: Imágenes Digitales y Aplicaciones 2ª ed, Alfaomega , RA-MA, Madrid .
19. Pajares, G. and Cruz, J.M (2008). Ejercicios resueltos de Visión por Computador 2ª ed , Alfaomega , RA-MA, Madrid ..
20. Real Academia Española. (2001). *Diccionario de la lengua española* (22.aed.). Consultado en <http://www.rae.es/rae.html>
21. Steinhaus, H. (1957). «Sur la division des corps matériels en parties» (en francés). Bull. Acad. Polon. Sci. 4 (12): pp. 801–804.
22. Takagi N., Kikuchi H., and Mukaidono M. (2004). “Applications of Fuzzy Logic Functions to Knowledge Discovery in Databases”. Transactions on Rough Sets II, LNCS 3135, pp. 107–128.
23. Tan, P.-N., Steinbach, M., Kumar, V. (2006). Introduction to Data Mining. Addison Wesley.
24. Theodoridis, S., Koutroumbas, K. (2009). *Pattern Recognition, Fourth Edition*. Academic Press. ISBN: 9781597492720

8 GLOSARIO

8.1 INTRODUCCIÓN

El Glosario contiene las definiciones de los términos utilizados en el proyecto “Detección de moho en queso mediante procesamiento de imágenes”. Este capítulo se define y/o aclara aquellos términos que tienen especial significado para este proyecto.

8.2 DEFINICIONES

.NET

.NET es un proyecto de Microsoft para crear una nueva plataforma de desarrollo de software con énfasis en transparencia de redes, con independencia de la plataforma y que permita un rápido desarrollo de aplicaciones. Basado en esta plataforma, Microsoft intenta desarrollar una estrategia horizontal que integre todos sus productos.

AGVs

Automatic Guided Vehicle, o lo que es lo mismo, vehículos de guiado automático. Los sistemas AGVs tienen sus comienzos en 1953, cuando se pensó en hacer realidad un “sueño” con un camión remolque sin conductor. Este primer vehículo precisaba de un cable que enterrado en el suelo de la fábrica creaba un campo magnético que servía de guía al vehículo. Poco a poco, se han ido incorporando al mundo industrial a la vez que crecían en aplicaciones y sofisticación. En el caso descrito en el presente proyecto se trata de un vehículo magneto guiado por pista generadas mediante pequeños imágenes en el suelo, que tiene una capacidad de carga de dos palés por movimiento, 1600 Kg.

C#

C# es un lenguaje de programación orientado a objetos desarrollado y estandarizado por Microsoft como parte de su plataforma .NET, que después fue aprobado como un estándar por la ECMA e ISO.

Curación

También denominado maduración o añejamiento, es un proceso en el que se secan y adicionalmente se aplican técnicas de conservación, como el salado o el ahumado. El proceso de curado hace que obtenga una textura bastante más dura y seca, así como que se incremente la intensidad de su sabor, propiedad muy deseada entre los amantes del queso. Según el grado del proceso pueden distinguirse varias clases: tiernos, semicurados, curados, viejos y añejos.

ERP

Sistemas de planificación de recursos empresariales, o ERP (por sus siglas en inglés, Enterprise Resource Planning) son sistemas de información gerenciales que integran y manejan muchos de los negocios asociados con las operaciones de producción y de los aspectos de distribución de una compañía en la producción de bienes o servicios. Los sistemas ERP típicamente manejan la producción, logística, distribución, inventario, envíos, facturas y contabilidad de la compañía de forma modular. Sin embargo, la Planificación de Recursos Empresariales o el software ERP pueden intervenir en el control de muchas actividades de negocios como ventas, entregas, pagos, producción, administración de inventarios, calidad de administración y la administración de recursos humanos. Entre sus módulos más comunes se encuentran el de manufactura o producción, almacenamiento, logística e información tecnológica, incluyen además la contabilidad, y suelen incluir un sistema de administración de recursos humanos y herramientas de mercadotecnia y administración estratégica.

Fotocélulas

Sensor fotoeléctrico consistente en un dispositivo electrónico que responde al cambio en la intensidad de la luz. Estos sensores requieren de un componente emisor que genera la luz, y un componente receptor que percibe la luz generada por el emisor. Los diferentes modos de sensado se basan en este principio de funcionamiento. Están diseñados especialmente para la detección, clasificación y posicionado de objetos; la detección de formas, colores y diferencias de superficie, incluso bajo condiciones ambientales extremas.

Moho

El moho es un hongo que se encuentra tanto al aire libre como en lugares húmedos y con baja luminosidad, condiciones en las que se almacenan los quesos para realizar su proceso de curación. Son característicos de ciertas variedades de queso muy conocidas como pueden ser el queso azul, el brie, roquefort, etc. Si bien en la modalidad de pasta prensada o queso manchego, la abundancia de moho y sobre todo la prolongación de éste en la corteza del producto, puede originar que invada el interior del queso, variando sus propiedades organolépticas.

Organolépticas

Propiedades organolépticas son todas aquellas descripciones de las características físicas que posee la materia en general, gracias a las cuales pueden percibir los sentidos. En concreto para un alimento, destacan su aspecto (forma), color, olor, sabor y textura. Su estudio es importante en las ramas de la ciencia en las que es habitual evaluar inicialmente las características de la materia sin la ayuda de instrumentos científicos.

Palé

Palés de Quesos, agrupación de quesos formada por 144 unidades, divididas en 24 cajas, con formas de dos torres de 12 cajas cada una, dos tapas en la parte superior de cada torre y un palé base que agrupa todo. Se utiliza como unidad indivisible durante el proceso de producción.

Paletizado/despaletizado

Proceso de creación o digresión del palé. Ciertas acciones del proceso de fabricación de queso se realizan sobre la unidad de agrupamiento del palé, pero otras requieren un proceso individualizado, es en aquí donde la agrupación se deshace, rehaciéndola posteriormente.

Pasta prensada

Modalidad de queso denominada así porque tras el proceso de coagulación de la cuajada, desuerado e introducción de ésta en los moldes, se somete a presión el mismo para facilitar la salida del suero de la masa.

Servicios Web XML

Un servicio Web (en inglés Web service) es una colección de protocolos y estándares que sirven para intercambiar datos entre aplicaciones. Distintas aplicaciones de software desarrolladas en lenguajes de programación diferentes, y ejecutadas sobre cualquier plataforma, pueden utilizar los servicios Web para intercambiar datos en redes de ordenadores como Internet. La interoperabilidad se consigue mediante la adopción de estándares abiertos. Las organizaciones OASIS y W3C son los comités responsables de la arquitectura y reglamentación de los servicios Web.

SGBD

Los Sistemas de gestión de base de datos son un tipo de software muy específico, dedicado a servir de interfaz entre la base de datos, el usuario y las aplicaciones que la utilizan. Se compone de un lenguaje de definición de datos, de un lenguaje de manipulación de datos y de un lenguaje de consulta. En los textos que tratan este tema, o temas relacionados, se mencionan los términos SGBD y DBMS, siendo ambos equivalentes, y acrónimos, respectivamente, de Sistema Gestor de Bases de Datos y DataBase Management System, su expresión inglesa.

SGA

Programa informático destinado a gestionar la operativa de un almacén. Proviene de la traducción del término inglés WMS (Warehouse Management System). Para ser considerado como tal, y no una simple gestión de stocks, el programa no sólo ha de gestionar las ubicaciones de los productos, sino también los movimientos de los operarios y de las máquinas encargadas de la manutención de los artículos.

Volteo

Proceso por el cual trascurrido un periodo de tiempo, al queso se le cambia de posición invirtiendo la posición de las caras planas del cilindro con el objeto de evitar deformaciones provenientes de la acción de la gravedad y conseguir así una curación homogénea en todo el volumen del queso.

ANEXO I

Base de conocimiento

En este anexo se recoge la base de conocimiento o “ground truth” que se ha hecho para el estudio de detección de moho en queso mediante procesamiento de imágenes. Está compuesta por un total de 25 palés, un total de 100 imágenes, agrupadas según la cámara de procedencia de las mismas. Para cada una de las imágenes un experto en el proceso de maduración del queso ha evaluado cada una de las imágenes de la muestra según dos criterios, cantidad de superficie del palé afectada por moho y nivel de concentración de moho en la superficie afectada. Estas características se evaluarán de 0 a 3 según su intensidad creciente (0=nada, 1=algo, 2=bastante y 3=mucho). Cada uno de estos niveles se traduce en un porcentaje que refleja la magnitud, estos factores son multiplicados consiguiendo así un indicador de presencia de moho por imagen. Tras obtener el indicador de cada una de las cuatro imágenes que conforman el estudio del palé, se realizará el promedio, obteniendo así el indicador de moho de dicho palé en ese momento del proceso de maduración.

Base de Conocimiento de niveles de moho en palés																	
Nº de Palé	Imagen Camara 1	Superficie Afectada	Nivel Concentracion	Indicador Moho Imagen	Imagen Camara2	Superficie Afectada	Nivel Concentracion	Indicador Moho Imagen	Imagen Camara3	Superficie Afectada	Nivel Concentracion	Indicador Moho Imagen	Imagen Camara 4	Superficie Afectada	Nivel Concentracion	Indicador Moho Imagen	Indicador Moho Palé
1		2	3	66%		2	3	66%		2	2	44%		1	1	10%	47%
2		0	0	0%		0	0	0%		0	0	0%		0	0	0%	0%
3		1	3	33%		1	3	33%		2	2	44%		1	2	22%	33%
4		1	1	10%		1	1	10%		0	0	0%		0	0	0%	5%
5		1	1	11%		1	1	11%		1	1	11%		1	1	10%	11%
6		0	0	0%		0	0	0%		0	0	0%		0	1	0%	0%
7		1	1	11%		1	1	11%		1	2	22%		1	1	11%	14%
8		1	1	11%		1	1	11%		2	1	22%		1	1	11%	14%
9		0	1	0%		1	1	11%		1	1	11%		0	0	0%	6%
10		0	0	0%		1	2	22%		1	2	22%		0	1	0%	11%

11	0	0	0%	1	2	22%	2	2	44%	1	1	11%	19%
12	3	2	66%	3	2	66%	0	0	0%	0	0	0%	33%
13	3	3	100%	3	3	100%	3	3	100%	3	3	100%	100%
14	2	2	44%	2	2	44%	2	3	66%	2	2	44%	50%
15	2	3	66%	2	2	44%	2	3	66%	2	3	66%	61%
16	2	1	22%	2	3	66%	2	2	44%	1	2	22%	39%
17	2	3	66%	1	2	22%	2	3	66%	1	1	11%	41%
18	1	0	0%	0	0	0%	0	0	0%	0	0	0%	0%
19	2	1	22%	2	3	66%	3	2	66%	2	1	22%	44%
20	1	0	0%	0	0	0%	2	2	44%	0	0	0%	11%
21	3	2	66%	3	3	100%	2	3	66%	3	2	66%	75%
22	2	3	66%	3	3	100%	3	3	100%	2	1	22%	72%
23	2	3	66%	3	3	100%	3	3	100%	2		22%	72%
24	2	2	44%	3	3	100%	2	3	66%	0	0	0%	53%
25	2	3	66%	3	3	100%	2	3	66%	3	2	66%	75%

Al 1 Base de conocimiento o "ground truth"

ANEXO II

Sistema de captura de imágenes.

Una de las partes más complicadas de llevar a cabo del proyecto, fue la instalación del sistema de captura de imágenes. En la instalación se utilizaron cuatro cámaras del modelo Axis M2014-E (Ref. Fabricante: 0549-001) cuyas especificaciones más relevantes se sintetizan como sigue: protección IP 66, resolución máxima de 1280x720 píxeles, entradas externas, programación de eventos sobre entradas, configuración de máscaras sobre las imágenes, conexión a red y otros detalles que se pueden ver en su ficha técnica.

Para suministrar iluminación correcta a la superficie de estudio, se tuvieron que instalar distintos puntos de luz externos, dado que la sala en cuestión no posee suficiente iluminación. Debido a la importancia de este punto, para alcanzar los resultados esperados, se realizaron pruebas con distintos focos de luz variando, tipo de iluminación, potencia de luz, número de puntos de luz y posición de los puntos de luz.

A continuación se muestran las distintas configuraciones probadas antes de alcanzar la más óptima:

1º Prueba: Foco alógeno de 500 vatios y brazo extensible. Debido a la lejanía de la cámara a la palé se pensó en un brazo extensor de 125 cm, que aproximara la cámara al objeto y permitiera además situar el punto de luz.



All 1 Instalación con un punto de luz y brazo extensible”

El resultado no fue positivo, ya que el sistema era inestable y la proximidad del foco junto a la potencia del mismo generaba una gran cantidad de sobras.

2º Prueba: 2 focos alógenos de 500 vatios. Se separan las cámaras del objeto llevándolo hasta el vallado exterior, dando robustez a la instalación. Se utiliza la característica de las cámaras de enmascarado, para eliminar la información que no pertenece al objeto de estudio. Se instalan dos puntos de luz para intentar eliminar el máximo de sobras.



All 2 Instalación con dos puntos de luz sobre el vallado exterior

3º Prueba: Pantalla fluorescente de 50 watos (2x 25w). Habiendo ganado robustez con la instalación del vallado se prueba con iluminación menos focalizada, es decir que no ofrezca tanta potencia de luz en una superficie pequeña, se usará para ello tubos de luz fluorescente.



All 3 Instalación con un punto de luz fluorescente sobre el vallado exterior

Se consigue mejorar la presencia de sombras, aunque este montaje ofrece poca iluminación para que la calidad de la imagen sea válida. Destacar como punto positivo que con este tipo de luz se reduce de forma considerable, 20 veces menos, el consumo energético de la instalación.

4º Prueba: 2 pantallas fluorescentes de 160 watos (2x80w). Se aumenta la potencia luminosa colocando dos pantallas de mayor potencia que las utilizadas en la prueba anterior.



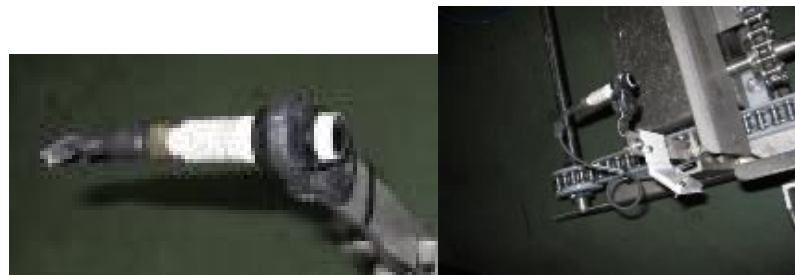
All 4 Instalación con un punto de luz fluorescente sobre el vallado exterior

La distribución con dos puntos de luz fluorescente elimina en gran parte las sombras permitiendo obtener mejores resultados que con las anteriores distribuciones, aunque es probable que todavía exista la posibilidad de mejorar. En la siguiente figura podemos ver la instalación final.



All 5 Instalación final con dos puntos de luz fluorescente sobre el vallado exterior

Además de las cámaras, se utilizaron 4 fotocélulas Omron PNP (Ref. Fabricante: E3f-ds30b1) con espejo, que se encargan de activar la iluminación cuando el palé se encuentra próximo a la posición de captura de la imagen y de activar el disparador de la cámara de fotos cuando el palé alcanza la posición deseada. En la figura All-5 podemos ver la instalación de las fotocélulas mediante soportes metálicos, con las que se detecta la presencia del palé en la posición de captura de imagen.



All 6 Instalación fotocélulas soporte metálico