

TRABAJO FIN DE MÁSTER
Máster en Ingeniería de Sistemas y Control

**DESARROLLO DE TÉCNICAS PARA LA
IDENTIFICACIÓN DE FRUTAS EN
IMÁGENES DEL ESPECTRO VISIBLE,
ULTRAVIOLETA Y TÉRMICO**

Ricardo Gómez Ribao

Directores:

Matilde Santos Peñas

José Miguel Guerrero Hernández



*Departamento de Informática y Automática
Madrid, 2018*

Máster en Ingeniería de Sistemas y Control

**DESARROLLO DE TÉCNICAS PARA LA
IDENTIFICACIÓN DE FRUTAS EN
IMÁGENES DEL ESPECTRO VISIBLE,
ULTRAVIOLETA Y TÉRMICO**

Proyecto Tipo A:

Proyecto específico propuesto por un profesor

Ricardo Gómez Ribao

Directores:

Matilde Santos Peñas

José Miguel Guerrero Hernández



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 blue ink, appearing to read 'R. Gómez Ribao', with a long horizontal stroke extending to the left.

Ricardo Gómez Ribao

Desarrollo de técnicas para la identificación de frutas en imágenes del espectro visible, ultravioleta y térmico

RESUMEN

En este trabajo se propone el estudio, análisis y aplicación de técnicas de visión por computador orientadas a la identificación de elementos en imágenes, en concreto tres tipos de fruta (naranja, manzana y plátano). El propósito final es identificar objetos en la imagen según sus diferentes características geométricas y su textura.

En primer lugar, se ha explorado una imagen en el rango visible, utilizando varios tratamientos del color para determinar nuestra capacidad de detectar y aislar los objetos buscados mediante sus características geométricas. El procesamiento de la imagen en el espectro visible se ha basado en el uso de diferentes metodologías de descomposición de la imagen en sus componentes de color. En función del objeto de la búsqueda se ha aislado la componente de color predominante en cada uno de ellos para extraer la información mediante diferentes técnicas de procesamiento, como la cuantización de los niveles de gris, la posterior binarización selectiva de las imágenes, y las transformaciones morfológicas para mejorar la calidad de las mismas. La identificación y estudio de las regiones de la imagen binaria es otra técnica que aporta información sobre ellas para poder tomar decisiones.

Posteriormente se ha explorado el tratamiento de imágenes con fuente de luz ultravioleta. En este rango de radiación, aun cuando la fuente sigue siendo indirecta, se ha observado otras propiedades de la imagen, y más concretamente, la reflexión de la luz en las superficies de los cuerpos a identificar.

Como tercer paso de este estudio se ha procedido a evaluar parte de los métodos utilizados en las imágenes anteriores con una imagen térmica, en la que la fuente de luz es la radiación infrarroja emitida por los cuerpos de la escena. Basándonos en la homogeneidad térmica de los mismos, esto ha permitido aislarlos mejor en esta fase de procesamiento.

Aunque esta última imagen no proporciona información de color y por ello la identificación del tipo de objeto no es concluyente, en cambio la discriminación por sus propiedades geométricas es muy eficaz.

Por ello, y como última parte del análisis, se ha combinado el procesamiento de la imagen térmica con la información de color que se puede extraer de esa misma región de la escena en el rango visible, de forma que a través de la combinación de la información de ambas imágenes se realiza un filtrado más efectivo. Junto al estudio previo, esta fusión de imágenes y características constituye la principal contribución de este trabajo.

Aunque en este Trabajo Fin de Máster se busca exclusivamente la identificación de tres frutas (manzana, naranja y plátano), el objetivo último es desarrollar una herramienta automática que sea capaz de extraer la información de tres imágenes, térmica, ultravioleta y en el rango visible, para detectar objetos e identificarlos.

Lista de palabras clave

Tratamiento de imágenes, imágenes térmicas, reconocimiento de patrones, texturas, imágenes ultravioletas, rango visible.

ABSTRACT

This paper proposes the study, analysis and application of computer vision techniques targeting object identification in images, in particular, fruits identification (oranges, apples and bananas). The final aim is to identify objects in the images through their different geometrical characteristics and textures.

First, we have studied the visible spectrum images using different color processing techniques to detect and isolate objects searched by their geometrical characteristics. The image processing in the visible spectrum relies upon different methodologies separating the channels.

Depending on the object, we have isolated the principal color component of each one, trying to get information through different color analysis, like quantization of gray levels, selective binarization of the images and morphological transformations to improve the quality of them. The identification and study of regions in the binary image is another technique that provides information for decision-making.

Later, we have explored the ultraviolet image processing. In this spectrum, even with indirect light to the object, we have observed other properties, more precisely, the reflection of light in the surfaces of the bodies to be identified.

The third stage of this paper consists of several methods used to process thermal images, in which the source of light is the infrared radiation emitted by the bodies. Based on the thermal homogeneity of such bodies, we have been able to isolate them better than in previous types of image.

However, this last image does not provide color information, such that the type of object is not uniquely defined and might yield to false detections.

Finally, and as a consequence of the study, we have combined the processing of the thermal image with the color information that can be extracted from the same region in the image of the visible spectrum, making the main contribution.

Although in this work we have searched exclusively the identification of three fruits (apple, orange and banana), the ultimate goal is to develop an automatic tool which can be extended to other fruits or even different objects.

Keywords:

Image processing, thermal images, pattern recognition, textures, ultraviolet images, visible spectrum images.

Índice

1.- Introducción	1
1.1.- Descripción del problema	1
1.2.- Motivación	2
1.3.- Objetivo general	3
1.3.1.- Objetivos específicos	3
1.4.- Organización de la memoria	4
2.- Materiales y métodos	5
2.1.- Sistemas de visión artificial	5
2.2.- Reconocimiento de patrones y descripción de regiones	6
2.3.- Procesamiento de imágenes	8
2.3.1.- Representaciones del color	8
2.3.2.- Modificación del histograma	10
2.3.2.1.- Ecuilización del histograma	11
2.3.2.2.- Expansión del histograma	12
2.3.2.3.- Observaciones en la modificación del histograma	12
2.3.3.- Cuantificación del número de niveles de gris y binarización de la imagen	13
2.3.4.- Transformaciones morfológicas	14
2.3.5.- Extracción de regiones	15
2.4.- Cámaras utilizadas	15
2.4.1.- Cámara del espectro visible	16
2.4.2.- Cámara del espectro infrarrojo	17
2.4.3.- Cámara del espectro ultravioleta	18
2.5.- Descripción de las imágenes	19
2.6.- Software	22
3.- Reconocimiento de objetos en imágenes reales	23

3.1.- Características de los patrones	23
3.2.- Características de las imágenes	27
3.3.- Modificación del histograma, binarización y análisis de las regiones	29
3.3.1.- Modificación del histograma	29
3.3.1.1.- Modificación del histograma de las imágenes del rango visible	29
3.3.1.2.- Modificación del histograma de las imágenes del rango ultravioleta	32
3.3.1.3.- Modificación del histograma de las imágenes térmicas	33
3.3.2.- Binarización	34
3.3.3.- Análisis de regiones	35
3.4.- Resultados del análisis cuantitativo	36
4.- Análisis y discusión de resultados	39
4.1.- Resultados sobre la imagen visible	39
4.1.1.- Obtención de las componentes de color	41
4.1.2.- Modificación del histograma	42
4.1.3.- Binarización de la imagen	46
4.1.4.- Detección de regiones y filtrado	47
4.2.- Resultados sobre la imagen UV	49
4.3.- Resultados sobre la imagen infrarroja.	52
4.4.- Resultados de la combinación de las imágenes infrarroja y visible	55
4.5.- Análisis de eficiencia de los métodos utilizados	57
5.- Conclusiones y futuras líneas de investigación	61
6.- Bibliografía	63
Apéndice 1- Código MATLAB del análisis y filtrado de regiones	67
Apéndice 2- Listado de funciones utilizadas	71
Apéndice 3- Tabla de colores según representación HSV	73

Índice de Figuras

Figura 1	Esquema del proceso general a seguir	4
Figura 2	Ejemplo de imagen multispectral, rango visible (izquierda) e infrarrojo cercano (derecha). En la manzana se aprecia un defecto en el IR que no se ve en el visible.	5
Figura 3	Ejemplo de propiedades de una región de píxeles: perímetro, área y excentricidad.	7
Figura 4	Ejemplo de eualización del histograma de una imagen en escala de grises. Arriba a la izquierda la imagen original, a la derecha arriba la imagen transformada. Abajo, el histograma de la izquierda es el original, el de la derecha es el transformado	11
Figura 5	Ejemplo de expansión del histograma de una imagen en escala de grises. Arriba a la izquierda la imagen original, a la derecha arriba la imagen transformada. Abajo, el histograma de la izquierda es el original, el de la derecha es el transformado.	12
Figura 6	Resultado de aplicar la operación morfológica erosión a una imagen de este estudio	14
Figura 7	Resultado de aplicar la operación morfológica dilatación a una imagen de este estudio	14
Figura 8	Operaciones de apertura (central) y cierre (derecha) a la imagen de la izquierda	15
Figura 9	Cámara del espectro visible	16
Figura 10	Rango espectral de funcionamiento de la cámara del rango visible	17
Figura 11	Cámara del espectro infrarrojo	17
Figura 12	Rango espectral de funcionamiento de la cámara IR.	18
Figura 13	Cámara del espectro UV.	18
Figura 14	Rango espectral de funcionamiento de la cámara UV, indicando la zona de mayor eficiencia del sensor en la captación de luz.	19
Figura 15	Termografía infrarroja en una escena con la fuente de luz detrás de los objetos.	20
Figura 16	Termografía infrarroja en una imagen con iluminación natural por delante de los objetos	20
Figura 17	Imagen UV de la misma escena con fuente de luz natural delante de los objetos.	20
Figura 18	Imagen UV de una escena con la fuente de luz por detrás de los objetos.	20
Figura 19	Imagen en el rango visible, fuente de luz por detrás de los objetos.	20
Figura 20	Imagen en el rango visible de una escena con fuente de luz natural por delante de los objetos.	20
Figura 21	Imagen en el rango visible de una escena con fuente de luz artificial fluorescente por encima de los objetos.	21
Figura 22	Imagen en el rango visible de una escena con fuente de luz natural por delante de los objetos y artificial fluorescente por encima de ellos.	21
Figura 23	Histograma de la imagen termográfica de la Figura 16.	21
Figura 24	Componente V de la representación HSV de la Figura 15, y su histograma adjunto.	22
Figura 25	Componente V de la representación HSV de la Figura 20 y su histograma adjunto.	22
Figura 26	Etapas de procesado de una imagen hasta la clasificación de los objetos.	23
Figura 27	Región detectada y representación del envoltorio convexo de mínima superficie que la engloba	24
Figura 28	Imágenes de las tres frutas, objeto de este estudio, y los patrones binarios obtenidos a partir de ellas.	26

Figura 29	Ejemplo de una manzana iluminada con luz natural y los puntos de medición de color.	28
Figura 30	Ejemplo de transformación de la componente Y de CMY y la modificación del histograma. En este caso se expande el histograma para aumentar el contraste de los objetos	30
Figura 31	Ejemplo de transformación de la componente V de HSV y la modificación del histograma. En este caso se desplaza y contrae el histograma para mejorar la distinción de los objetos	30
Figura 32	Ejemplos de tratamiento de algunas componentes de color con la detección de contornos por gradientes.	32
Figura 33	Ejemplo de histograma modificado para una imagen UV en una escena donde los objetos están iluminados por luz natural frontal	33
Figura 34	Imagen térmica cuyo histograma ha sido modificado, escena en la que los objetos están iluminados frontalmente por luz natural	33
Figura 35	Ejemplo de las componentes de color de varias imágenes según el tipo de iluminación utilizada	40
Figura 36	Ejemplo del tratamiento sobre una manzana considerando su forma tridimensional y los efectos de la reducción del espacio de grises.	41
Figura 37	Histograma de la componente V de HSV de la imagen con luz natural por detrás de los objetos (A1).	42
Figura 38	Histograma de la componente V de HSV de la imagen con luz natural por delante de los objetos (A2).	42
Figura 39	Histograma de la componente V de HSV de la imagen con luz artificial por encima de los objetos (A3).	42
Figura 40	Histograma de la componente V de HSV de la imagen con luz natural frontal a los objetos y luz artificial por encima de ellos (A4).	42
Figura 41	Histograma correspondiente a la componente V de una imagen tomada con luz artificial por encima de los objetos.	43
Figura 42	Histograma modificado de una imagen con luz natural por detrás de los objetos para la componente V de HSV	44
Figura 43	Histograma modificado de una imagen con luz natural por delante de los objetos para la componente V de HSV	44
Figura 44	Histograma modificado de una imagen con luz artificial por encima de los objetos para la componente V de HSV.	44
Figura 45	Histograma modificado de una imagen con luz natural por delante de los objetos y luz artificial por encima para la componente V de HSV.	44
Figura 46	Imágenes resultantes de modificar el histograma en la componente V de HSV. En azul los bordes detectados para facilitar la comprensión de la máscara superpuesta	45
Figura 47	Imagen resultante de modificar el histograma (ecuación 31) en la componente V de HSV para una escena iluminada con luz natural por delante de los objetos.	45
Figura 48	Imágenes resultantes de modificar el histograma en la componente V de HSV. A la izquierda la imagen sin combinar con la detección de contornos por gradientes.	45
Figura 49	Imagen resultante NAND de modificar el histograma en la componente V de HSV para una escena iluminada con luz natural y artificial.	46
Figura 50	Imagen binarizada de una escena iluminada con luz natural por detrás de los objetos. Componente V de HSV.	46
Figura 51	Imagen binarizada de una escena iluminada con luz natural por delante de los objetos. Componente V de HSV	46
Figura 52	Imagen binarizada de una escena iluminada con luz artificial por encima de los objetos. Componente V de HSV	46
Figura 53	Imagen binarizada de una escena iluminada con luz natural por delante de los objetos y luz artificial por encima de ellos. Componente V de HSV	46

Figura 54	Objetos identificados con un filtrado simple (izquierda) y complejo (derecha) en una imagen con luz natural por detrás de los objetos. Componente V de HSV	47
Figura 55	Detección de objetos con filtrado simple (izquierda) y complejo (derecha) de una imagen con luz natural frontal. Componente V de HSV	47
Figura 56	Detección de objetos con filtrado simple (izquierda) y complejo (derecha) en una imagen con luz artificial por encima de los objetos. Componente V de HSV	48
Figura 57	Detección de objetos con filtrado simple (izquierda) y complejo (derecha) de una imagen con luz natural frontal y luz artificial por encima de los objetos. Componente V de HSV	48
Figura 58	Ecuilibración del histograma y binarización para una imagen ultravioleta con los objetos iluminados por detrás	49
Figura 59	Filtrado simple y complejo de una imagen ultravioleta con iluminación detrás de los objetos.	50
Figura 60	Ecuilibración del histograma y binarización de una imagen ultravioleta con iluminación natural por delante de los objetos	50
Figura 61	Filtrado simple y complejo para una imagen ultravioleta con iluminación natural por delante de los objetos.	50
Figura 62	Ecuilibración del histograma en una imagen ultravioleta cuando la iluminación es artificial. Aunque los objetos en la imagen inicial no son claramente visibles, la ecualización del histograma los muestra en la imagen de la derecha	51
Figura 63	Ecuilibración del histograma y binarización de una imagen ultravioleta iluminada con luz natural y artificial	51
Figura 64	Filtrado simple y complejo de una imagen ultravioleta con iluminación natural y artificial	51
Figura 65	Expansión del histograma (arriba izquierda), binarización (arriba derecha), filtrado simple (abajo izquierda) y filtrado complejo de las regiones (abajo derecha) de la imagen térmica de la escena iluminada con luz natural por detrás de los objetos.	52
Figura 66	Expansión del histograma (arriba izquierda), binarización (arriba derecha), filtrado simple (abajo izquierda) y filtrado complejo de las regiones (abajo derecha) de la imagen térmica de la escena iluminada con luz natural frontal	53
Figura 67	Expansión del histograma (arriba izquierda), binarización (arriba derecha), filtrado simple (abajo izquierda) y filtrado complejo de las regiones (abajo derecha) de la imagen térmica de la escena iluminada con luz artificial por encima de los objetos.	53
Figura 68	Expansión del histograma (arriba izquierda), binarización (arriba derecha), filtrado simple (abajo izquierda) y filtrado complejo de las regiones (abajo derecha) de la imagen térmica de la escena con iluminación natural frontal y artificial superior	54
Figura 69	Expansión del histograma (arriba izquierda), binarización (arriba derecha), filtrado simple (abajo izquierda) y filtrado complejo de las regiones (abajo derecha) de la imagen térmica de la escena sin iluminación en el rango visible	54
Figura 70	Termografía con detección de colores en una imagen con iluminación natural por detrás de los objetos.	56
Figura 71	Termografía con detección de colores en una imagen con iluminación natural por delante de los objetos	56
Figura 72	Termografía con detección de colores en una imagen con iluminación artificial por encima de los objetos	56
Figura 73	Termografía con detección de colores en una imagen con iluminación natural por delante de los objetos y luz artificial por encima de los mismos.	56

Figura 74	Superficie de evolución de la eficiencia general de la detección en función del tipo de escena y del método utilizado.	57
Figura 75	Tiempo máximo de cálculo por operación en segundos. La detección de regiones está especificada, además, por tipo de imagen.	59
Figura 76	Tabla de colores según los valores HSV	73

Índice de Tablas

Tabla 1	Hoja de características de la cámara SVS-VISTEK	16
Tabla 2	Características técnicas de la cámara IR	17
Tabla 3	Características técnicas del sensor IR, Gobi-640-GigE.	18
Tabla 4	Características técnicas cámara UV.	19
Tabla 5	Valores de los parámetros de los patrones	27
Tabla 6	Para los puntos marcados en la Figura 29 se indica su posición relativa en la imagen y los valores en las representaciones RGB, HSV, CMY y CIELab, así como el nombre del color según tablas estándares.	28
Tabla 7	Correspondencia de literales de color con los valores de las componentes de la representación de color HSV	55
Tabla 8	Resultados generales según el tipo de imagen y el tipo de iluminación	57
Tabla 9	Estadísticas de las imágenes estudiadas con el análisis de la componente V de HSV.	58
Tabla 10	Estadísticas de las imágenes estudiadas con el análisis de la componente Y de CMY	58
Tabla 11	Estadísticas de las imágenes estudiadas con el análisis de la componente L de Lab	58
Tabla 12	Estadísticas de las imágenes UV estudiadas	58
Tabla 13	Estadísticas de las imágenes térmicas estudiadas	59
Tabla 14	Estadísticas de las imágenes térmicas en combinación con detección de color.	59

1.- Introducción

1.1.- Descripción del problema

En la actualidad los sistemas de visión y procesado automático de información están presentes en múltiples áreas de la industria y de otros entornos comerciales [1-2]. Los sistemas más antiguos y sencillos son capaces de detectar características básicas, principalmente geométricas, de objetos con una forma poco variable y repetitiva, como por ejemplo las piezas fabricadas en una cadena de producción. Estos sistemas devuelven una respuesta binaria y raramente son capaces de discriminar por sí solos qué tipo de objeto evalúan y la bondad de la tendencia del proceso productivo.

En las últimas décadas, debido a la aparición de sistemas de procesado más ágiles y potentes y a una reducción significativa de los costes de los dispositivos, las líneas de producción han integrado sistemas de visión que son capaces de discriminar piezas diferentes e informar a los manipuladores del sistema para que ejecuten tareas más específicas según el elemento. Incluso son capaces de evaluar características más variables y menos precisas matemáticamente como, por ejemplo, la forma de gajos de mandarina [3].

Todos estos sistemas actúan bajo la premisa de unas situaciones y cualidades bastante repetitivas y restringidas. Esto quiere decir que las condiciones de iluminación y la posición y orientación en que se presenta el objeto están controladas y se repiten en un rango de valores muy delimitado.

Por ejemplo, los sistemas de visión artificial que se utilizan en los entornos de alimentación están dispuestos en una zona de la instalación donde capturan las imágenes de las frutas que se desplazan sobre la cinta transportadora [4]. El sistema de visión es capaz de clasificar la fruta según su forma o tamaño, y determinar su grado de maduración a través del color [5].

La inspección en la industria alimentaria se lleva a cabo tanto con cámaras lineales como con cámaras matriciales, dependiendo del tipo de fabricante y del tipo de clasificación. Algunos fabricantes de maquinaria para la industria alimenticia ofrecen también versiones basadas en cámaras infrarrojas [6], que pueden detectar la presencia de hueso en las frutas o determinar características no visibles en su superficie [7-14].

Normalmente estos sistemas incluyen múltiples cámaras para capturar las imágenes de toda la superficie del producto. Cuando los productos son más o menos redondeados, se usan mecanismos para hacer rotar el producto bajo el sistema automático de inspección visual. La clasificación por forma se establece a partir de los diámetros máximo y mínimo, perímetro, área, así como a partir de la relación entre los diferentes parámetros medidos. Las decisiones según el color se basan en la superficie total escaneada. Para determinar las propiedades del color se utilizan parámetros como: relación simple de porcentaje, histograma de valor de intensidad, definición de áreas máximas o mínimas, etc.

La utilización de estos sistemas ha ayudado a mejorar en gran manera la automatización de estas tareas. Uno de los grandes inconvenientes de la inspección manual realizada por personas es la ineficiencia en las tareas que requieren examinar productos aparentemente

muy similares. Después de un periodo de tiempo relativamente corto la agudeza visual se deteriora, incluso en el caso de que la clasificación sea relativamente simple. Aunque no son perfectos, los sistemas de clasificación de alimentos basados en visión artificial son más coherentes en la clasificación de productos buenos y malos, en la separación de productos correctos de extraños, y en clasificación por tamaños.

Las tecnologías de visión artificial se han visto beneficiadas por los avances en los microprocesadores y elementos de procesado digital de la señal, la alta resolución de las cámaras tanto lineales como matriciales, el desarrollo de sistemas de iluminación especializados, así como por las investigaciones realizadas por especialistas en agricultura, conjuntamente con investigadores de automática y visión. El resultado es que los sistemas de visión actuales cumplen los requisitos que demanda la industria alimentaria.

Estos sistemas son capaces de clasificar los alimentos según su apariencia [15], y por tanto, pueden identificar defectos relacionados con plagas, picaduras de animales, efecto de las heladas, oxidación, o cualquier otro tipo de imperfecciones o enfermedades que pueden detectarse a través de la superficie. Algunos fabricantes proporcionan sistemas de visión basados en rayos X, que son capaces de detectar defectos internos o condiciones que no podrían reconocerse tomando sólo la imagen de la superficie del producto.

Recientemente la incorporación al mercado de los sistemas de medición láser 3D ha introducido potentes mejoras en el campo del análisis dimensional de los productos, y la inclusión de las cámaras multiespectrales (registro de imágenes en diferentes frecuencias de onda) ha permitido rebajar los costes y aumentar la polivalencia de los sistemas de detección e identificación.

Pese a todo lo mencionado anteriormente, la mayoría de los sistemas se hacen prácticamente a medida. Esto significa que, a menudo, incorporan iluminación cuyo espectro es específico para el tipo de producto a inspeccionar, de forma que puedan mejorar la selección.

Estos sistemas no se suelen enfrentar al problema de visualizar un objeto en un entorno real, con las dificultades propias de los exteriores.

Se necesitan sistemas inteligentes, capaces de tomar decisiones autónomamente en entornos reales, capaces de identificar objetos y determinar sus propiedades sin las condiciones y restricciones antes comentadas [1]. Un robot aspirador o un dron nunca podrán contar con unas condiciones de iluminación controlada, o con un posicionamiento u orientación de los objetos determinado y, sin embargo, son capaces de identificar de forma eficiente los objetos y reaccionar ante ellos.

La identificación de frutas en imágenes para entornos no industriales es un buen ejemplo de estudio de detección de objetos con geometría irregular, con presencia de brillos y sombras, en unas condiciones no controladas.

1.2.- Motivación

Como se ha comentado anteriormente, abordar el problema de la identificación de objetos en imágenes en exteriores no controlados constituye la puerta de entrada para desarrollar

sistemas autónomos capaces de comportarse de forma flexible ante elementos de geometría y color variables.

Para ello, construir una herramienta automática para la identificación de objetos en imágenes, que a su vez sea simple, es el primer paso para, posteriormente, obtener un sistema que sea capaz de aprender y evolucionar de forma autónoma.

1.3.- Objetivo general

Con este trabajo pretendemos desarrollar unas herramientas básicas que permitan tratar de forma automática imágenes de escenas con objetos con unas características previamente estudiadas, y ser capaces de identificarlos.

Estas herramientas deberán estar adaptadas a las diferentes tecnologías de captación de imágenes que utilicemos, siendo capaces de adquirir información de más de uno de los formatos de imágenes utilizados y fusionarla para tomar una decisión adecuada y fiable.

Las pruebas derivadas de este estudio están encaminadas a resolver la complejidad de la selección de parámetros en los distintos pasos de ajuste, y estimar la bondad de los métodos utilizados para el tipo de imágenes proporcionadas.

El sistema de procesamiento desarrollado debe ser aplicable a diferentes tipos de imágenes de manera efectiva, teniendo en cuenta que los parámetros son adaptables al formato de las mismas. Así mismo proporcionaremos criterios de decisión en tiempo real en función de las características de la imagen.

1.3.1.- Objetivos específicos

Dentro de este trabajo nos hemos propuesto analizar los siguientes aspectos, teniendo en cuenta que se dispone de imágenes adquiridas en distintas frecuencias del espectro.

- i. Procesado de imagen que es el más adecuado para cada tipo de imagen.
- ii. Información a extraer en cada imagen para evaluar los objetos: geométrica, correlación con patrones, color, etc.
- iii. Determinar qué información es la más relevante para los distintos tipos de imágenes.
- iv. Determinar qué características y/o combinación de éstas permite identificar los tres tipos de frutas: manzana, naranja y plátano (objeto del estudio de este trabajo).
- v. Medir el grado de acierto en la selección y decisión.
- vi. Cuál es la combinación más adecuada de operaciones sobre varias imágenes de forma que se pueda mejorar la certeza de la decisión.
- vii. Clasificación de objetos en imágenes de distinto espectro.
- viii. Combinar la información obtenida de distintas imágenes.

Se pretende que el procesado de la imagen sea genérico, para todos los tipos de imágenes, y que la aplicación a cada imagen dependa de sus características de intensidad y distribución de color. En la Figura 1 se representan los principales pasos del proceso.

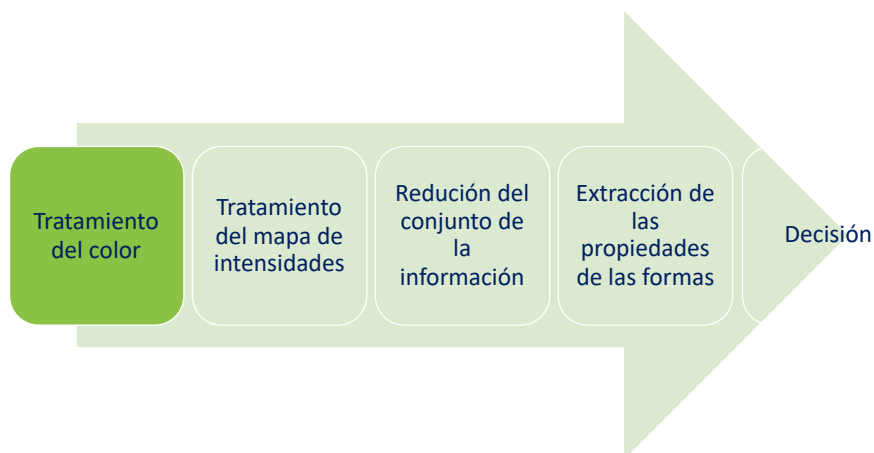


Figura 1. Esquema del proceso general a seguir

Este proceso debe ser capaz de adaptarse a la imagen a estudiar, así como sus parámetros a las propiedades de la misma.

La herramienta a desarrollar se va a aplicar a imágenes de frutas, porque es un tipo de objeto que al utilizar imágenes en diferentes puntos del espectro electromagnético permite detectar muchas propiedades de la misma. Además su morfología presenta la suficiente variabilidad como para ser representativa en el caso del problema estudiado.

Se quiere que la herramienta sea lo suficientemente robusta y fiable como para poder extender su uso a otras aplicaciones y objetos.

1.4.- Organización de la memoria

En el capítulo uno se presenta el problema que se va a analizar en este trabajo fin de máster.

En el capítulo dos se detallan los métodos de reconocimiento de patrones y objetos en imágenes, así como los materiales y métodos empleados en el presente estudio. Se analizan las características que se utilizarán en el procesado y se describen las imágenes objeto de estudio.

En el capítulo tres se exponen con detalle los métodos aplicados a cada tipo de imagen, así como en la combinación de los mismos, justificando su uso.

En el capítulo cuatro se analizan y comparan los resultados de los métodos estudiados.

En el capítulo cinco se presentan las conclusiones obtenidas y se exponen las posibles líneas de trabajo futuro.

2.- Materiales y métodos

2.1.- Sistemas de visión artificial

En los últimos años la evolución tecnológica producida en los diferentes dispositivos electrónicos ha abierto un abanico de aplicaciones, disminuyendo el coste de inversión así como su carga computacional.

En concreto, los sistemas de captación de imágenes, las cámaras, han evolucionado notablemente. En primer lugar, por el gran avance que se ha producido en la integración de los semiconductores, que ha permitido fabricar sensores con mayor resolución, con mayor sensibilidad (capacidad para registrar un rango más amplio de intensidades, especialmente las bajas) y velocidad. Estos sensores también han sido desarrollados para la captación de diferentes rangos del espectro electromagnético. Si inicialmente las cámaras digitales (industriales) eran exclusivamente monocromas (únicamente medían la intensidad general de la escena), en la actualidad son capaces de procesar el color (luz en el rango visible del espectro), de capturar imágenes en el rango cercano de la luz infrarroja (IR), e incluso en el rango del ultravioleta (UV) [14].

En segundo lugar, los dispositivos que contienen el sensor y la óptica asociada a las propias cámaras han seguido esta evolución, añadiendo unidades de procesado más rápidas y potentes. En la actualidad hay muchas cámaras industriales que incorporan ya las herramientas matemáticas capaces de procesar la imagen y devolver el resultado del análisis en tiempo real.

Dentro de los últimos avances tecnológicos producidos en la fabricación de las cámaras digitales se encuentran las cámaras multispectrales. Estos dispositivos se caracterizan por la captación simultánea de varias frecuencias del espectro electromagnético, de forma que retornan al sistema de procesado tanto imágenes a color como en los rangos infrarrojo (IR) y ultravioleta (UV) (Figura 2). Esto proporciona una ventaja en velocidad de procesado importante, porque no necesita capturar la escena con diferentes cámaras ni realizar ajustes de encuadre, sino que en el mismo instante se pueden procesar de forma paralela diferentes tipos de imágenes que proporcionan información que se complementa [5; 15].



Figura 2. Ejemplo de imagen multispectral, rango visible (izquierda) e infrarrojo cercano (derecha). En la manzana se aprecia un defecto en el IR que no se ve en el visible.

En la actualidad también son importantes los avances en el reconocimiento y captación de escenas en tres dimensiones. Estos sistemas se han visto beneficiados por la gran capacidad de computación desarrollada en los nuevos dispositivos tecnológicos.

Desde el punto de vista físico y con respecto a los sistemas de visión, se puede hablar de dos líneas de desarrollo. La primera constaría de un haz láser, que realiza un barrido por la escena, y una cámara capaz de captar dicho haz y reconstruir la escena y sus dimensiones geométricas (en tres dimensiones), a través de la compilación y tratamiento de dichas imágenes. La segunda consta de sistemas multicámara [9; 16-18]. Estos trabajan sobre la combinación matemática de imágenes tomadas desde diferentes puntos de vista (estereoscopia). De esta forma, en cada imagen se procesan los objetos y se combinan las características y coordenadas de cada uno para construir la escena en 3D en el computador.

2.2.- Reconocimiento de patrones y descripción de regiones

El reconocimiento de patrones [19-21] es una disciplina científica cuyo objetivo es la clasificación de objetos en un cierto número de categorías o clases. Nos referiremos a estos objetos de forma genérica llamándolos patrones.

Históricamente los dos enfoques del reconocimiento de patrones han sido el estadístico (o teoría de la decisión) y el sintáctico (o estructural). Con el desarrollo de las redes neuronales ha surgido un nuevo paradigma. Por otro lado, el aumento de la capacidad de almacenamiento de los modernos sistemas de computación 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 propiedades de naturaleza cuantitativa, mientras que el sintáctico se fundamenta en las relaciones geométricas asociadas a la forma de los objetos. El enfoque basado en la apariencia considera distintas formas de vista de los mismos.

Todos los enfoques anteriores están basados en la identificación de unas métricas descriptoras que definan las regiones encontradas en las imágenes, que el sistema pueda utilizar para evaluar y tomar la decisión sobre la naturaleza del patrón.

Una región puede describirse por la forma de su frontera o por sus características internas [22]:

1. Respecto a la forma de su frontera existen diversos métodos aplicables. Por ejemplo, la descripción de segmentos rectos mediante códigos de cadena (donde se relaciona un píxel dado con sus vecinos circundantes), la utilización de ajuste de líneas mediante mínimos cuadrados, la transformada de Hough, etc. El objetivo de estos métodos es describir una frontera y ser capaces de entender sus propiedades, generalmente por asociación a formas geométricas sencillas conocidas, como rectas y circunferencias. Por otro lado, el método de códigos de cadena permite describir una frontera a base de segmentos rectilíneos que conectan los píxeles del contorno. Otro método es la utilización de los descriptores de Fourier, de momentos, o de funciones splines.

- Respecto a las características internas de la región, a la hora de evaluarla como un patrón conocido, se habla de propiedades topológicas (camino para llegar de un píxel de la región a otro sin abandonarla, número de huecos, etc.) o de propiedades métricas (área, perímetro, centroide o centro de gravedad, ejes máximo y mínimo, etc.).

Una propiedad topológica se caracteriza por ser invariante a ciertas deformaciones de las figuras de la imagen.

Entre las propiedades métricas, además de las citadas, cabe destacar la compatibilidad (k), razón entre la longitud de su frontera P (suma de las longitudes de todos los segmentos entre píxeles) y los contenidos en el área A , según las ecuaciones (1) y (2). Su valor es mínimo cuando la forma se asemeja a un disco (Figura 3).

$$P = \sum_i \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \quad (1)$$

$$k = \frac{P^2}{A} \quad (2)$$

El cálculo del centro de gravedad se realiza aplicando la ecuación (3).

$$\bar{x} = \frac{1}{A} \sum_i x_i \quad \bar{y} = \frac{1}{A} \sum_i y_i \quad (3)$$

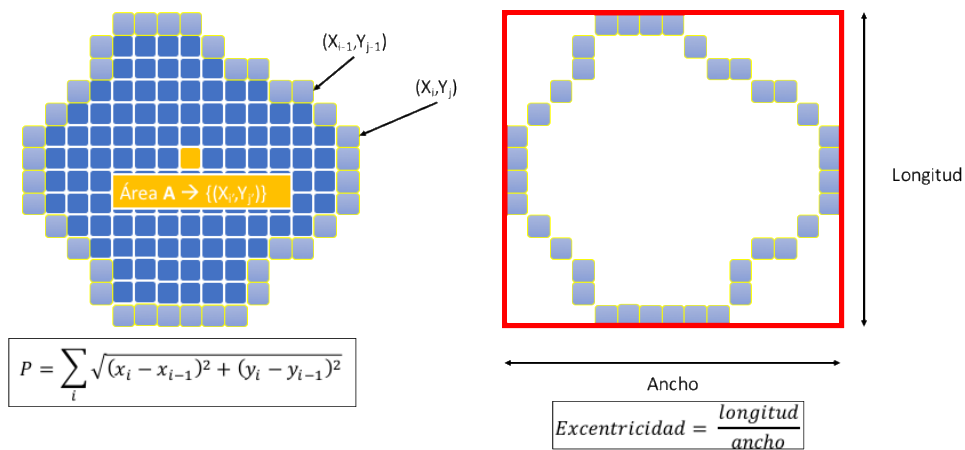


Figura 3. Ejemplo de propiedades de una región de píxeles: perímetro, área y excentricidad.

Otra propiedad de interés es la excentricidad o razón de aspecto, que en el caso de un rectángulo es la razón de su longitud respecto a su ancho. El valor de un cuadrado sería 1. Cuando la figura se circunscribe en un rectángulo, este valor se utiliza para caracterizarla y compararla con el patrón esperado.

También se puede estudiar la redondez de la forma como la dispersión de la medida de la distancia de cada píxel del perímetro al centroide.

Una última aproximación dentro de los métodos métricos para identificar patrones son los momentos invariantes. Los descriptores de Fourier utilizados para describir las fronteras son invariantes a traslación, rotación y cambios de escala. Para la descripción de las regiones se puede utilizar diversos tipos de momentos, igualmente invariantes. Entre ellos destacan los 7 momentos de Hu y los momentos invariantes afines. La utilización de estos métodos de cálculo requiere disponer de unos patrones previamente estudiados con los que se puedan comparar los momentos de las regiones encontradas.

2.3.- Procesamiento de imágenes

En este apartado se presenta una descripción de las diferentes etapas de procesamiento de las imágenes para obtener la información con la que discriminar los objetos.

2.3.1.- Representaciones del color

Para una imagen a color, existen diferentes formas de representar éste, generalmente mediante la combinación de varios componentes o matrices [23-24].

La representación **RGB** corresponde a la descomposición del color en tres componentes: R-rojo/red, G-verde/Green, B-azul/Blue. La combinación en determinadas medidas es la que proporciona el color de cada píxel. Esta representación, para cada imagen, implica la existencia de tres matrices de intensidad de iguales dimensiones.

La determinación del color de cada pixel, y su identificación con ciertos descriptores básicos como “amarillo” o “naranja”, requiere de una formulación matemática con mayor coste computacional que otras representaciones. Por esta razón no se utiliza en este análisis para extraer la información de color de cada región.

El modelo **HSV** responde a las siglas en inglés Hue (Matiz), Saturation (Saturación) y Value (Valor), siendo este último valor el que se puede identificar como brillo. Es una transformación no lineal del modelo RGB que permite identificar la tercera componente como la de intensidad lumínica, a partir de la que se puede estudiar y procesar la imagen.

La conversión entre el modelo RGB y el HSV se realiza mediante las ecuaciones (4-8). Si R, G y B tienen valores entre 0 y 255, primero se normaliza.

$$R'=R/255 \quad G'=G/255 \quad B'=B/255 \quad (4)$$

$$C_{max}=\max(R',G',B')$$

$$C_{min}=\min(R',G',B')$$

$$\Delta=C_{max}-C_{min} \quad (5)$$

Y finalmente se obtienen las distintas matrices.

$$H = \begin{cases} 0^\circ & \text{si } \Delta = 0 \\ 60^\circ \times \left(\frac{G' - B'}{\Delta} \bmod 6 \right) & \text{si } C_{max} = R' \\ 60^\circ \times \left(\frac{B' - R'}{\Delta} + 2 \right) & \text{si } C_{max} = G' \\ 60^\circ \times \left(\frac{R' - G'}{\Delta} + 4 \right) & \text{si } C_{max} = B' \end{cases} \quad (6)$$

$$S = \begin{cases} 0 & \text{si } C_{max} = 0 \\ \frac{\Delta}{C_{max}} & \text{si } C_{max} \neq 0 \end{cases} \quad (7)$$

$$V = C_{max} \quad (8)$$

Este modelo presenta varias ventajas en el tratamiento de imágenes con color. Por un lado, si se quiere evaluar exclusivamente la intensidad de la luz se puede procesar la tercera componente como si fuese una imagen de escala de grises. Por otro lado, si se quiere conocer la componente principal de color básico, se puede promediar sobre el valor de la primera componente. Cada color se podría considerar tomando dicho valor promedio y comparándolo con el rango abarcado por cada color.

El modelo **CMY** se caracteriza por ser el contrario o sustractivo del RGB, de forma que la tercera componente es el color Amarillo (Yellow). Las siglas se identifican con los siguientes colores: C = Cian (Cyan), M= Magenta, Y=Amarillo (Yellow).

Esto quiere decir que si la representación RGB viene dada en enteros de 8 bits, esto es, su valor máximo es 255, entonces la representación CMY se obtiene restando a ese máximo el valor de cada componente de RGB. Si fuese una representación real de 0 a 1, entonces el máximo sería 1 y no 255, tal y como indica la expresión (9).

$$\left. \begin{aligned} C &= \text{maximo} - R \\ M &= \text{maximo} - G \\ Y &= \text{maximo} - B \end{aligned} \right\} \quad (9)$$

La utilidad principal de este modelo, para este análisis, radica en que la componente principal de color de las frutas/objetos buscados es el amarillo. De esta forma el cálculo o procesado hecho sobre la tercera componente de este modelo debiera anticipar los resultados de forma más sencilla, computacionalmente hablando, que con otros modelos de color.

La representación de color **YCbCr** se inventó para su utilización en los sistemas de televisión y vídeo de los años 1940's, con el objetivo de separar las componentes de color de la imagen en diferentes rangos de frecuencia de emisión. La componente Y lleva la información de color de la imagen, algo análogo a la saturación de otros modelos. Las otras dos componentes representan las crominancias o componentes de diferencia de color, una al azul (Cb) y otra al rojo (Cr).

La expresión (10) muestra la conversión desde el espacio RGB al YCbCr.

$$\begin{aligned}
Y &= 0.299*R + 0.587*G + 0.114*B \\
Cb &= (B-Y)*0.564 + 0.5 \\
Cr &= (R-Y)*0.713 + 0.5
\end{aligned}
\tag{10}$$

Por último, se ha utilizado la representación **CIE Lab** para buscar los objetos en las imágenes a color. Este modelo, muy utilizado en edición de fotografía digital, permite reducir el análisis y tratamiento de la imagen a su componente L o Luminosidad. La componente 'a' recorre el rango de rojo a verde en el espacio tridimensional de color y la 'b' recorre el eje azul.

La conversión desde el espacio RGB se explica en [23], aunque se pueden utilizar funciones ya creadas en algunos paquetes de software como MATLAB.

El modelo Lab, según [24], es independiente del dispositivo utilizado y proporciona mejor comportamiento en la medición de colores de objetos con volumen. El modelo RGB es el que más distorsión crea frente a volúmenes y superficies curvas, mientras que el más ajustado al color real percibido por el ojo humano es el Lab. Entre medio se sitúa el modelo HSV, computacionalmente más rápido. Por otro lado, el modelo CMY es análogo al RGB al ser una mera modificación lineal del mismo.

También en [25] se explica cómo aquellos modelos en los que se puede separar la componente de intensidad o luminancia de las componentes cromáticas son más eficaces que aquellos que no lo hacen. De entre todos los modelos, Lab y HSV resultan ser los que mejor lo hacen.

2.3.2.- Modificación del histograma

La ecualización o realzado del contraste de una imagen es una manipulación de la misma orientada a mejorar su calidad. Trata de eliminar efectos no deseados tales como sombras, reflejos y ruido. En este trabajo se ha optado por realizar esta operación a través del histograma.

El histograma es una función discreta que representa el número de píxeles de la imagen en función de los niveles de intensidad, g. La probabilidad P(g) de ocurrencia de un determinado nivel g se define como:

$$P(g) = \frac{N(g)}{M}
\tag{11}$$

Donde M es el número de píxeles de la imagen y N(g) es el número de píxeles en el nivel de intensidad g. Como cualquier distribución de probabilidad todos los valores de P(g) están en el rango [0,1], y su suma total debe ser igual a 1.

En general, un histograma con una distribución de los niveles de gris concentrada en una determinada zona presenta un contraste muy bajo. Así mismo, cuando un histograma de bajo contraste tiene sus valores concentrados en la parte baja del mismo, se corresponde con una imagen oscura, mientras que si están en la parte alta, se corresponde con una imagen brillante.

Cuando el rango de niveles de gris que toma la imagen se encuentra concentrado en una zona del intervalo, la imagen posee poco contraste. Para aumentar el contraste, se puede expandir el histograma o bien realizar una ecualización del mismo.

A continuación se detallan dos métodos de modificación del histograma, por un lado la ecualización y por otro la expansión.

2.3.2.1.- Ecualización del histograma

La ecualización del histograma pretende mejorar el contraste de la imagen, es decir, reparte de forma más o menos uniforme los valores del histograma original.

Para realizar esta transformación, dada una imagen $M \times N$, con n píxeles para cada nivel g , la ecualización del histograma consiste en realizar la siguiente transformación sobre los niveles de intensidad de la imagen (12):

$$E(g) = \frac{L-1}{MN} \sum_{i=0}^g n(i) \quad (12)$$

Siendo L el nivel máximo de gris.

La Figura 4 muestra el resultado de aplicar esta operación sobre una imagen del conjunto de pruebas.

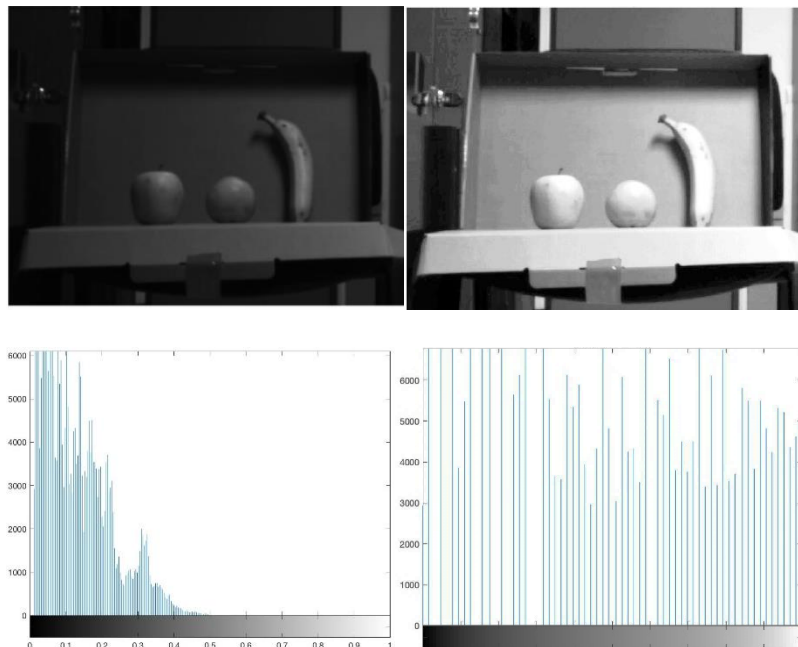


Figura 4. Ejemplo de ecualización del histograma de una imagen en escala de grises. Arriba a la izquierda la imagen original, a la derecha arriba la imagen transformada. Abajo, el histograma de la izquierda es el original, el de la derecha es el transformado.

2.3.2.2.- Expansión del histograma

Otra técnica muy utilizada es la expansión del histograma, que consiste en aumentar el rango de niveles de gris de la imagen aplicando una regla de proporcionalidad a los límites (anteriores y nuevos) de intensidad (13).

$$g(i, j) = \left[\frac{f(i, j) - f(i, j)_{min}}{f(i, j)_{max} - f(i, j)_{min}} \right] \cdot (MAX - MIN) + MIN \quad (13)$$

A diferencia de la ecualización (que es una redistribución de la población de píxeles según otra función estadística de la original), la expansión es una redistribución de la población de píxeles e intensidades atendiendo a una regla simple de proporcionalidad de acuerdo con el rango original y con el deseado a la salida.

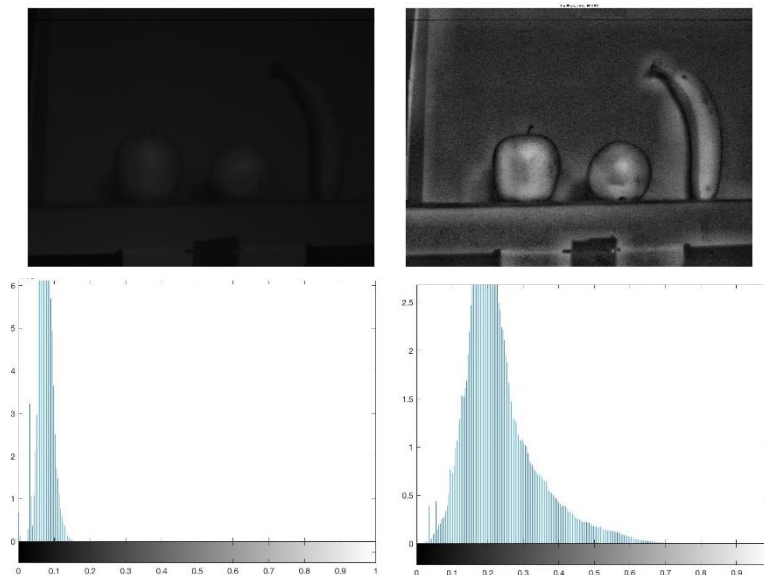


Figura 5. Ejemplo de expansión del histograma de una imagen en escala de grises. Arriba a la izquierda la imagen original, a la derecha arriba la imagen transformada. Abajo, el histograma de la izquierda es el original, el de la derecha es el transformado.

Como se puede comprobar en la Figura 5, la expansión no necesariamente mejora el contraste; sin embargo, permite cambiar las intensidades no saturadas a rangos donde las técnicas de binarización puedan ser más eficaces.

2.3.2.3.- Observaciones a la modificación del histograma

Existen funciones de MATLAB [26] que permiten realizar a la vez la ecualización y expansión del histograma. La utilización de este tipo de funciones se deberá ajustar a las características de la imagen. Así cuando la imagen presenta tanto brillos como sombras,

las funciones que detectan los extremos del histograma no permitirán una buena expansión, al encontrarse estos ya de facto en los extremos del rango.

El uso de un factor gamma de corrección puede mitigar el efecto de los brillos o de las sombras, pero no corrige ambos a la vez.

La mención a los brillos y las sombras se debe a que, en el contexto de la detección de objetos, esto complica notablemente la identificación de los mismos.

Es precisamente la elección que se haga de incorporar o no estos extremos en el tratamiento del histograma la que permitirá extraer más información referente a los contornos de los objetos.

En el caso de que la imagen presente diferencias notables a lo largo de la misma respecto a la iluminación, puede ser muy útil dividir la imagen en una matriz de celdas y aplicar estrategias de corrección del histograma locales a cada una de ellas.

2.3.3.- Cuantificación del número de niveles de gris y binarización de la imagen

La binarización de la imagen es la operación por la cual se convierte una imagen en escala de grises a una imagen lógica, en la que cada píxel puede tomar sólo dos valores, 0 ó 1. Este paso del proceso puede realizarse de dos maneras.

1. Binarización automática aplicando el método de Otsu.

El método de Otsu [27] asume que los píxeles de una imagen se agrupan en dos clases siguiendo un histograma bi-modal, esto es, los píxeles de fondo y los que contienen la información relevante. El método busca el umbral óptimo que minimiza la varianza dentro de las clases y maximiza la separación entre ellas.

2. Cuantificación de los niveles de gris.

Otra forma de proceder es la de realizar una cuantización de los niveles de gris [28-30] y posteriormente aplicar una regla lógica (consistente en una función de paso banda de niveles de intensidad) a los valores obtenidos en cada píxel.

Cuando el histograma de la imagen no está claramente dividido en clases, la aplicación del método de Otsu hace que se pierda información, por lo que no es tan útil pese a ser muy automatizable.

La aplicación del segundo método debe realizarse cuando en la imagen se puede apreciar que localmente se dan las condiciones necesarias para generar una clara clasificación de los píxeles.

Por ejemplo, en una imagen con muchos reflejos y muchas sombras en las proximidades de los contornos de los objetos, habiendo realizado la separación en un conjunto limitado de niveles de gris, se pueden aplicar reglas lógicas de exclusión o inclusión a cada píxel,

de forma que el resultado permitirá elegir como pertenecientes al objeto tanto los píxeles de intensidad alta como los de intensidad baja. O, al contrario, descartar tanto los bordes brillantes como las sombras oscuras.

2.3.4.- Transformaciones morfológicas

Debido a la presencia de los brillos y sombras, las operaciones de binarización frecuentemente dan lugar a píxeles aislados, a aristas separadas o, simplemente, a pseudo-contornos que pueden confundirse en el siguiente paso del procesamiento como falsos objetos. Básicamente se puede decir que los brillos adelgazan la forma y las sombras las engrosan, produciendo en ambos casos deformaciones que dificultan la interpretación.

Para tratar de mitigar estos efectos se recurre a las transformaciones morfológicas, cuyo principal cometido es simplificar la imagen y preservar las formas principales de los objetos.

Para ello se puede utilizar la erosión (Figura 6), una transformación morfológica cuantitativa. Se define como:

$$A \ominus B = \{z \in E^2 | B_z \subseteq A\} \text{ con } B_z = \{b + z | b \in B\} \forall z \in E^2 \quad (14)$$

Siendo A la imagen y B un elemento estructural o subconjunto de E^2 , espacio binario en 2D.

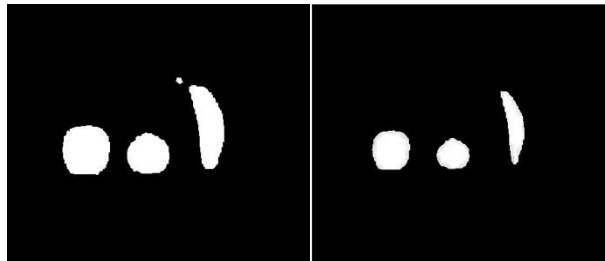


Figura 6. Resultado de aplicar la operación morfológica erosión a una imagen de este estudio

Otras operaciones morfológicas útiles, en función de la información que se quiera eliminar o modificar, son la dilatación, apertura y cierre. Así como la erosión elimina píxeles y ruido de un tamaño “reducido”, la dilatación (Figura 7) incrementa su tamaño de forma que fomenta la fusión de elementos.

$$A \oplus B = \{z \in E | (B^s)_z \cap A \neq \emptyset\} \text{ con } B^s = \{x \in E | -b \in B\} \forall z \in E \quad (15)$$

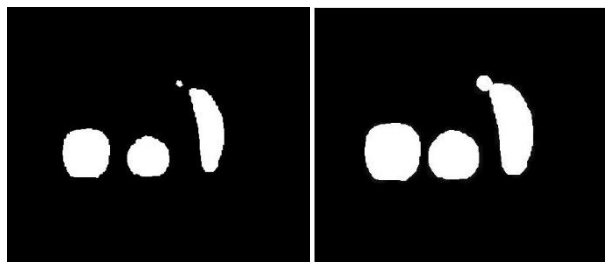


Figura 7. Resultado de aplicar la operación morfológica dilatación a una imagen de este estudio

La apertura y cierre (Figura 8) son dos operaciones en las que la erosión y la dilatación se suceden una a otra, pero en orden inverso. La apertura aplica primero una erosión y luego una dilatación, mientras que el cierre es al contrario.

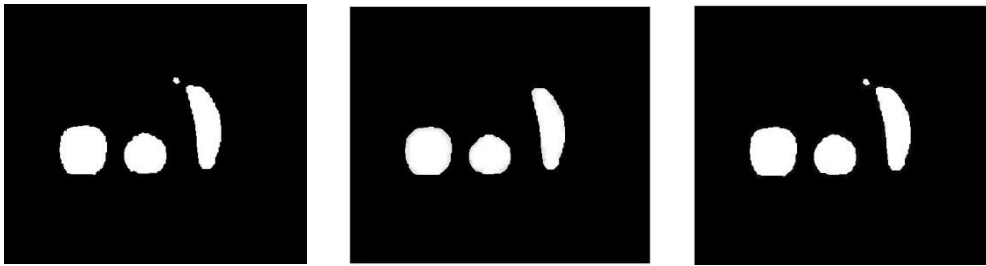


Figura 8. Imagen original (izquierda) tras aplicar las operaciones de apertura (central) y cierre (derecha).

La elección del elemento estructural es importante pues determina la calidad del resultado. Si se elige un elemento estructural como una línea vertical (vector $n \times 1$) el resultado será un adelgazamiento de los píxeles en dirección vertical de la imagen. Si por el contrario se selecciona una línea horizontal (vector $1 \times n$), el adelgazamiento se producirá en la dirección horizontal.

Existen otras formas para el elemento estructural, como el disco, que produce un adelgazamiento en las dos dimensiones, y según su tamaño, tiende a rellenar objetos.

2.3.5.- Extracción de regiones

La extracción de regiones [31-32] se realiza partiendo de una imagen binaria. En primer lugar se busca en la vecindad de cada píxel el valor de sus vecinos. Si el valor del píxel es cero, se ignora. Si es uno, comprueba qué vecinos tienen igualmente valor 1. Por defecto se utiliza conectividad 8, lo que quiere decir que mira en sus 8 vecinos adyacentes, y si varios de ellos tienen valor uno, considera que todos ellos forman parte del mismo objeto o región.

De esta forma, iterando sobre todos los píxeles, encuentra las regiones conectadas y las devuelve como conjuntos de píxeles.

Posteriormente se aplican funciones aritméticas sobre el número de píxeles de cada región, devolviendo unas métricas o propiedades de la región. De entre las múltiples propiedades se hará uso de las siguientes: Centroide, Perímetro, Área, Convex Hull (lista de píxeles del polígono convexo en que se inscribe la forma), y Bounding Box (rectángulo en que se inscribe la forma).

2.4.- Cámaras utilizadas

Las tres cámaras que se han utilizado para obtener las imágenes son las siguientes.

2.4.1.- Cámara del espectro visible: SVS-VISTEK SVS4050CFLGEA

Es una cámara (Figura 99) que se caracteriza por tener una resolución de 2336x1752 píxeles y un sensor de formato 1 pulgada tipo CCD. La velocidad de disparo es de 41.6 fotogramas por segundo (fps). En la Tabla 1 se muestran sus características técnicas.

El rango de absorción para cada una de las componentes de color RGB de una imagen se muestra en la Figura 10. De esta forma el color azul (B) tendría la mayor absorción entre 400 y 500 nm, mientras que la detección del verde (G) sería menos efectiva en ese rango, pero mejor en torno a 500 nm. Por último, el rojo (R) tiene menor eficacia que las otras dos, pero su curva es más extendida, lo que le da más capacidad de detección en un rango del espectro más amplio. La lente utilizada es de 10.4mm de distancia focal fija.



Figura 9. Cámara del espectro visible.

Tabla 1. Hoja de características de la cámara SVS-VISTEK

Resolución	2336 x 1752 px
Velocidad de disparo (max.)	41.6 fps
Interface	Dual GigE Visión
Sensor	KAI-04050-C
Fabricante del Sensor	ON Semiconductor
Tipo de Sensor	Área CCD
Tamaño del sensor(h x v)	12.85 x 9.64 mm
Formato del sensor	1 "
Pixel tamaño (h x v)	5.5 x 5.5 μ m
Tiempo exposición (min)	6 μ s
Tiempo exposición (max)	60 seg (externo ∞)
Rango dinámico	48 dB / 8 bit
Memoria interna	128 MB SDRAM, 136 MB Flash
Montura lente	C-Mount
Dimensiones (ancho x alto x fondo)	50 x 50 x 46.7 mm
Peso	160 g
Temperatura ambiente	-10 to 45 °C
Humedad ambiente	10 to 90 %
Protección clase	IP40
Alimentación	10 to 25 V (DC)
Consumo	7 W

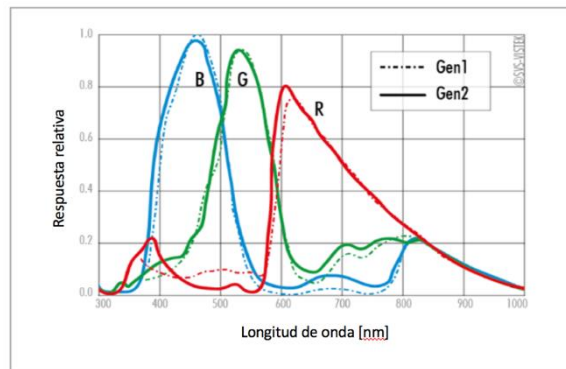


Figura 10. Rango espectral de funcionamiento de la cámara del rango visible.

2.4.2.- Cámara del espectro infrarrojo: Gobi-640GigE de Xenics

Cámara termográfica (Figura 1111) con una resolución de 640x480 píxeles, con una sensibilidad térmica de 55 mK a temperatura ambiente. La velocidad de disparo es de 25 fps. En la Tabla 2 se muestran sus características técnicas, y en la Tabla 3 las características del sensor de la cámara.

Es un dispositivo fiable para la captura de imágenes térmicas, muy utilizada en mediciones de temperatura sin contacto en visión por computador en entornos industriales y control de procesos. Combina alta calidad de las imágenes con alta resolución térmica (0.05°C), así como precisas capacidades de análisis térmico. La lente utilizada es de 25mm de distancia focal fija.



Figura 11. Cámara del espectro infrarrojo.

Tabla 2. Características técnicas de la cámara IR.

Longitud Focal de la lente	18 mm f/1, HFOV 42.6°,
Maximum framerate (full frame)	50 Hz
Rango de tiempo de integración	1 μ s - 80 μ s
Temperatura de estabilización	No requiere refrigeración termoeléctrica
Consumo	< 4.5 W
Voltaje de alimentación	12 V
Resistencia al Impacto	40 G, 11 ms conforme a MIL-STD810G
Vibración	5 G (20 Hz to 2000 Hz) según MIL-STD883J
Rango de temperatura ambiental	-40°C to 60°C
Rango de temperatura de almacenamiento	-40°C to 85°C
Dimensiones (A x Alto x Largo mm³)	49 x 49 x 79
Peso	263 g

Tabla 3. Características técnicas del sensor IR, Gobi-640-GigE.

Tipo de Matriz	Micro bolómetro no refrigerado (a-Si)
Banda espectral	8 μm to 14 μm
Resolución	640 x 480
Pixel pitch	17 μm
Sensibilidad térmica	55 mK @ 30°C con f/1 lente
Refrigeración de la matriz	No refrigerado
Operatividad de píxeles	> 99 %

En la Figura 12 se muestra la curva de eficiencia de la absorción de esta cámara.

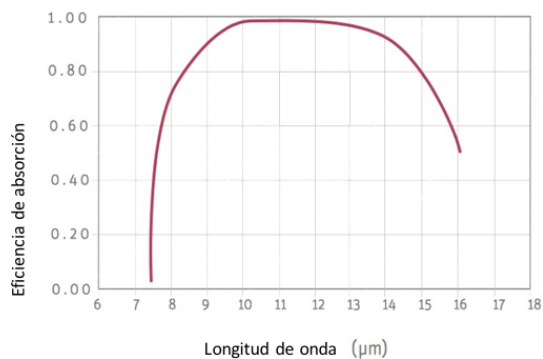


Figura 12. Rango espectral de funcionamiento de la cámara IR.

Como se puede observar, su eficiencia de absorción es máxima entre 9 μm y 15 μm , que es el rango óptimo del infrarrojo lejano, donde se genera la emisión térmica de los cuerpos.

2.4.3.- Cámara del espectro ultravioleta: CM-140GE-UV de GIGE Vision

Esta cámara (Figura 13) es una variante de la versión monocroma, en la que el sensor monocromo se ha sustituido por el modelo Sony ICX407BLA, el cual es capaz de captar más allá del rango visible, permitiendo pasar al espectro ultravioleta. Dado que la cámara también es capaz de captar en el rango visible, se ha instalado un filtro de corte para impedir el paso de la radiación no ultravioleta. Tiene un sensor de formato $\frac{1}{2}$ pulgada, con una resolución de 1392x1040 píxeles y una velocidad de disparo de 16.14 fps. La lente utilizada es de 25mm de distancia focal fija.



Figura 13. Cámara del espectro UV.

En la Tabla 4 se pueden apreciar sus características técnicas.

Tabla 4. Características técnicas cámara UV.

Especificaciones	CM-140GE-UV
Sensor	1/2" escaneado progresivo
Reloj	33.75 MHz
Frame rate	16.14 frames/segundo
Área activa	6.4(h) x 4.8(v) mm
Tamaño de celda	4.65 x 4.65 pixeles
Resolución	1392 x 1040
Ganancia	-3 a +12 dB
Temperatura de operación	-5°C a +45°C
Humedad	20-90% no-condensante
Temperatura almacenaje/humedad	-25°C a 60°C / 20 a 90%
Vibración	10G (20Hz a 200Hz XYZ)
Impacto	70G
Alimentación	12V-24V DC +/- 10% 4.1w
Montura Lente	Tipo C
Dimensiones (alto x ancho x largo)	29 x 44 x 75 mm
Peso	125 g

En la Figura 14 se muestra la zona de mayor eficiencia de absorción de esta cámara, la cual puede operar eficientemente hasta el rango de los 200 nm.

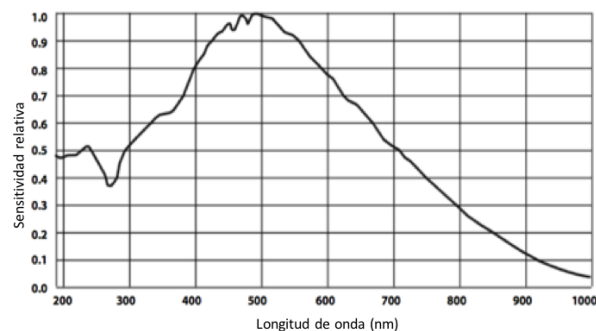


Figura 14. Rango espectral de funcionamiento de la cámara UV, indicando la zona de mayor eficiencia del sensor en la captación de luz.

2.5.- Descripción de las imágenes

A continuación se muestra un ejemplo de las imágenes captadas por este tipo de cámaras, donde pueden verse los tres tipos de frutas a identificar.

Se han considerado diferentes situaciones de iluminación de los objetos. Por ejemplo, escenas en las que la fuente de luz natural se encuentra detrás de los objetos, creando un conjunto de brillos y sombras en los mismos y su entorno que obligan a unas técnicas de procesamiento determinadas. Otras imágenes, tomadas en un entorno cerrado, con menos elementos físicamente próximos a los objetos a identificar, muestran iluminación “frontal” directa sobre los objetos, y por tanto requieren otras técnicas de procesamiento.

También, se ha experimentado con varias condiciones de iluminación: luz natural, luz interior mediante fluorescencia, ambas, y sin iluminación. Esta última ha sido descartada debido a que, como se verá en el capítulo 4 sección 3, la cámara térmica reconoce perfectamente el objeto, pero no es posible utilizar la información de color de la cámara de espectro visible al necesitar ésta de una fuente de luz que ilumine la escena.

El conjunto total de imágenes han sido 480, que se corresponden a 40 imágenes de cada distinto tipo de iluminación (iluminación detrás de los objetos, y frontal; luz natural, artificial, y ambas) para cada una de las tres cámaras disponibles. En las siguientes figuras se muestra un subconjunto de estas imágenes.



Figura 15. Termografía infrarroja en una escena con la fuente de luz detrás de los objetos.



Figura 16. Termografía infrarroja en una imagen con iluminación natural por delante de los objetos

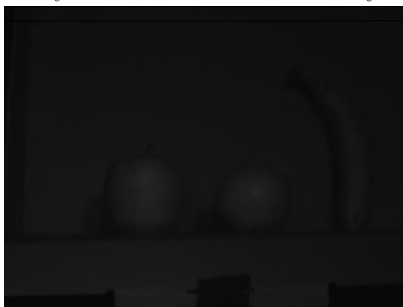


Figura 17. Imagen UV de la misma escena con fuente de luz natural delante de los objetos.



Figura 18. Imagen UV de una escena con la fuente de luz por detrás de los objetos.

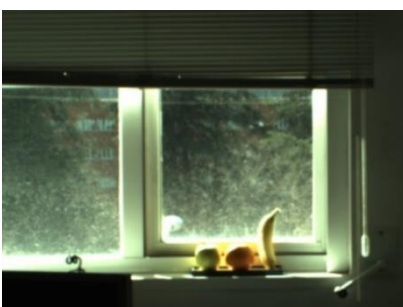


Figura 19. Imagen en el rango visible, fuente de luz por detrás de los objetos.



Figura 20. Imagen en el rango visible de una escena con fuente de luz natural por delante de los objetos.

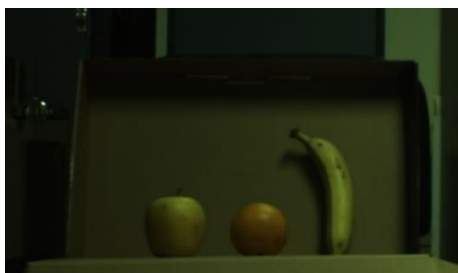


Figura 21. Imagen en el rango visible de una escena con fuente de luz artificial fluorescente por encima de los objetos.



Figura 22 Imagen en el rango visible de una escena con fuente de luz natural por delante de los objetos y artificial fluorescente por encima de ellos.

En las imágenes térmicas no se aprecia gran diferencia por el escaso contraste de las mismas (Figuras 15 y 16), aunque cambie la posición e intensidad de la fuente de luz visible.

En las imágenes ultravioleta existe diferencia dependiendo de la intensidad de la fuente de luz (Figuras 17 y 18).

Como se puede comprobar, cada escena aporta complejidades diferentes, dependiendo de la situación de saturación lumínica de cada zona. Por ejemplo, la combinación de ambas fuentes de iluminación, natural y artificial, o la posición de la fuente, detrás o delante de los objetos, genera dificultades en la detección de contornos. De la misma manera se observa que la luz artificial o eléctrica son fuentes de radiación ultravioleta de menor intensidad que la luz natural, por lo que la capacidad de detección de los objetos en dichas condiciones será menos favorable.

Las imágenes termográficas son en apariencia imágenes con poca información. No obstante, como se puede comprobar en sus histogramas (Figura 23), muy estrechos y centrados, se pueden realizar operaciones de ecuilización y expansión de los mismos, y por lo tanto, encontrar más información y separar los objetos diferenciando sus contornos.

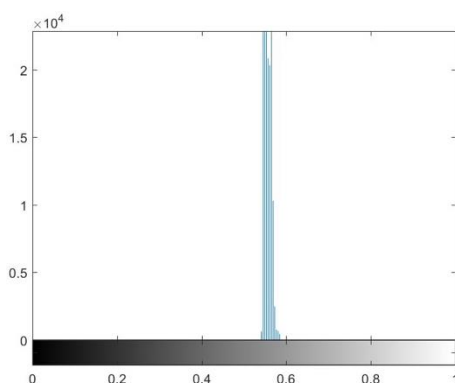


Figura 23. Histograma de la imagen termográfica de la Figura 16.

Sin embargo, los histogramas que corresponden a las imágenes a color (en particular los de sus respectivas componentes de color o intensidad) aparecen muy distribuidos, con altas densidades en los extremos, revelando así saturación de los píxeles de la imagen y por lo tanto pérdida de información (lo que conlleva mayor dificultad en la identificación de los objetos por separación de contornos). Esto se aprecia en las Figuras 24 y 25.

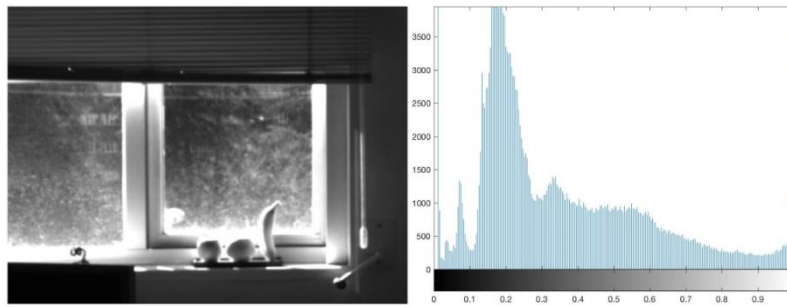


Figura 24. Componente V de la representación HSV de la Figura 19 y su histograma adjunto.

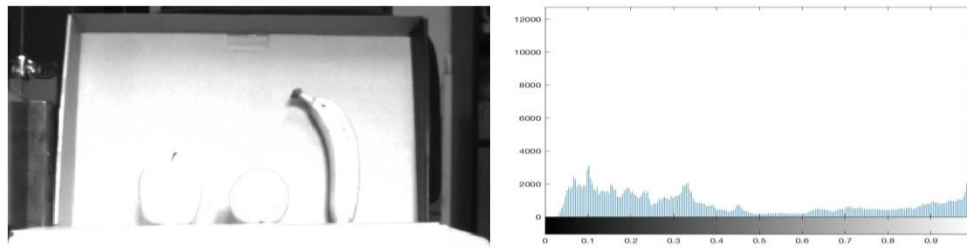


Figura 25. Componente V de la representación HSV de la Figura 20 y su histograma adjunto.

Por otro lado, de los histogramas se deduce que cuando la fuente de iluminación se encuentra por detrás de los objetos, el histograma de las imágenes a color se encuentra desplazado hacia los valores más oscuros, y en las imágenes con iluminación por delante ocurre lo contrario. Este hecho influirá en el algoritmo de procesamiento de las imágenes, pudiendo solventarse con una inversión de la imagen.

También es importante resaltar que en las imágenes donde la saturación se produce en los niveles más altos de luz hay más información por las características físicas de los sensores fotográficos modernos. Esto ya dará indicaciones de qué tipo de iluminación será la más adecuada para una aplicación práctica de este estudio.

2.6.- Software

Para el procesamiento de las imágenes se ha utilizado Matlab R2017a [26,36] junto con algunas de las funciones de la “Image Processing Toolbox” (Apéndice 2). Todo ello ha sido implementado y desarrollado en un PC Acer Aspire Intel i3 con 8Gb de RAM.

3.- Reconocimiento de objetos en imágenes reales

La clasificación de objetos es una forma de selección e identificación de los mismos. De los objetos de una imagen se pueden extraer unas propiedades medibles, las cuales se pueden asociar a las que tienen las clases generales que representan un tipo determinado de objeto.

Según el tipo de imagen, el procesado deberá extraer las componentes de color o no, siendo el resto de las etapas, conceptualmente, las mismas (Figura 26). Las etapas interiores son comunes a las tres tecnologías utilizadas, y son las que se aplican a una imagen de grises. En el caso de una imagen a color, primero se separa la componente que se va a utilizar y se transforma en una matriz de niveles de gris. De la extracción de componentes de color se obtiene un mapa de intensidades y una métrica de color sobre cada región de la imagen.

Una vez obtenido el mapa de intensidades, se practican operaciones de reducción de los niveles de gris hasta concluir en una imagen binaria sobre la que se trabajará para obtener las métricas. La reducción del número de niveles de gris se realiza mediante las técnicas de ecualización del histograma anteriormente descritas. La binarización puede llevarse a cabo mediante métodos de umbral o mediante reglas lógicas sobre los valores de gris de cada píxel.

La parte final del procesado puede hacerse individualmente sobre cada imagen, si bien la clasificación y las métricas de las regiones pueden ser el resultado de la combinación de decisiones sobre varios tipos de imágenes procesadas.



Figura 26. Etapas de procesado de una imagen hasta la clasificación de los objetos.

3.1.- Características de los patrones

En este estudio se realiza un enfoque sintáctico de las formas, para lo que se calculan las siguientes propiedades métricas de las regiones encontradas.

- Compatibilidad: de acuerdo a la expresión (2), se calcula este valor para las figuras patrón previamente obtenidas y posteriormente se compara con el de las regiones encontradas.

- **Solidez:** ratio entre la superficie de la región y la superficie encerrada por el envoltorio convexo (de mínima superficie) que rodea a la región (expresión 16)

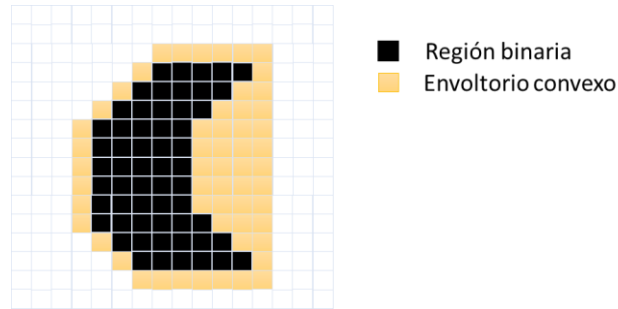


Figura 27. Región detectada y representación del envoltorio convexo de mínima superficie que la engloba.

$$\text{Solidez} = \frac{\text{Superficie de la región}}{\text{Superficie del envoltorio convexo}} \quad (16)$$

- **Excentricidad:** relación entre el ancho y el alto de la imagen o a la inversa, como indica la expresión (17).

$$e = \frac{\text{alto}}{\text{ancho}} \quad (17)$$

Figuras alargadas presentarán valores más altos de excentricidad que aquellas que son cuadradas o redondas, cuyos valores serán próximos a la unidad.

- **Redondez:** dispersión promedio de la distancia medida desde cada píxel perteneciente al perímetro de la región hasta el centroide de la misma. Si la dispersión es “pequeña” la figura podrá ser considerada redonda.

Si se denomina r_i a la distancia del píxel perimetral hasta el centroide, las ecuaciones utilizadas son las siguientes.

$$r_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (18)$$

$$R = \frac{1}{N} \sum_i r_i \quad (19)$$

$$\text{redondez} = \sqrt{\frac{1}{N-1} \sum_i |R - r_i|^2} \quad (20)$$

Cuanto más próximo sea este valor a cero, más redonda es la figura.

- **Color:** se calculan las componentes predominantes en cada región encontrada para determinar el color promedio de la región.
- **Momentos de Hu** [33, 34]: al ser invariantes a traslación, rotación y escala, se podrán comparar con los de los patrones.

El momento de orden (p+q) para una región viene dado por las ecuaciones (21).

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (21)$$

Donde x e y recorren todos los puntos de la región.

El momento central de orden (p+q) viene dado por la ecuación (22), siendo sus elementos los de la expresión (23).

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (22)$$

$$\begin{aligned} \bar{x} &= \frac{m_{10}}{m_{00}} \\ \bar{y} &= \frac{m_{01}}{m_{00}} \end{aligned} \quad (23)$$

Y los momentos centrales normalizados de orden (p+q) se definen de acuerdo a (24)

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \rightarrow \gamma = \frac{p+q}{2} + 1 \quad (24)$$

Para $p+q=2,3,\dots$

En base a estas definiciones se establecen los momentos invariantes de Hu que toman las expresiones (25-31).

$$\phi_1 = \eta_{20} + \eta_{02} \quad (25)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (26)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (27)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (28)$$

$$\begin{aligned} \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ &+ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (29)$$

$$\begin{aligned} \phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ &+ 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \end{aligned} \quad (30)$$

$$\begin{aligned} \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ &(3\eta_{12} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (31)$$

Para que todos los momentos contribuyan por igual en la función de discriminación, y por tanto estén dentro del mismo orden de magnitud, en ocasiones es necesario proceder a una normalización como la que se muestra en (32).

$$\phi'_n = \text{abs}(\ln(\phi_n)) \quad (32)$$

Todas estas características métricas se calculan cuando la imagen ha sido binarizada, esto es, cuando todos sus píxeles han sido convertidos a 0's ó 1's.

Por otro lado, para la caracterización de las regiones encontradas y su identificación como un determinado tipo de fruta, se describen a continuación los patrones utilizados para determinar las características geométricas y matemáticas con las que se comparan en cada caso. Estos patrones son el resultado de aplicar las técnicas de separación de grises, binarización y reconocimiento de regiones que se aplican en este trabajo (apartado 2 sección 3).

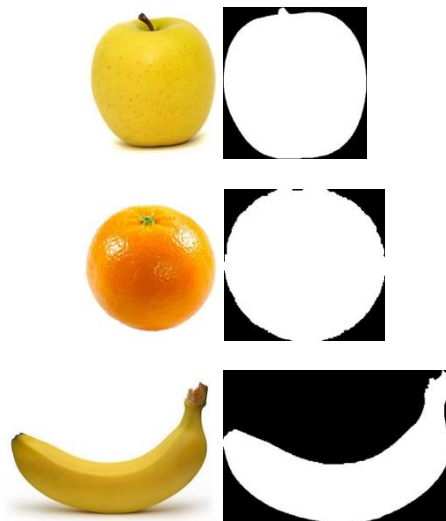


Figura 28. Imágenes de las tres frutas, objeto de este estudio, y los patrones binarios obtenidos a partir de ellas.

Tal y como se muestra en la Figura 28, se comprueba que las siluetas de la naranja y la manzana, aun siendo diferentes, son aproximadamente redondas, por lo que la componente de color será discriminante para su diferenciación. Respecto al patrón que representa al plátano, aunque está en posición horizontal, será totalmente válido para reconocer e identificar objetos en las imágenes ya que los momentos H_u son invariantes a rotación,

La tabla 5 representa los valores de las características arriba descritas de los patrones con los que se comparan las regiones encontradas.

De los valores presentes en la tabla se pueden deducir características cualitativas sobre la forma de las figuras encontradas. Por ejemplo, la manzana y la naranja presentan valores de redondez (expresión 20) muy bajos, y por ello se puede deducir que son formas muy próximas al círculo. No así el plátano, que tendrá una mayor excentricidad (expresión 17) al estar inscrito en un rectángulo alargado.

Tabla 5. Valores de los parámetros de los patrones.

	MANZANA	NARANJA	PLATANO
Redondez	0,030	0,021	0,335
Compatibilidad	15,827	13,743	32,084
Excentricidad	1,063	1,054	1,509
Solidez	0,98	0,985	0,76
Momentos Hu	0,796	0,798	0,445
	4,158	4,343	1,200
	4,951	5,683	1,780
	8,458	8,504	2,609
	15,160	15,810	4,833
	10,940	10,920	3,213
	16,350	15,700	5,250

También se puede observar que la compatibilidad (expresión 2) es menor en la manzana y la naranja, por ser formas más próximas a la circunferencia.

La solidez (expresión 16), al medir la convexidad de la figura, permite determinar cómo una figura puede parecerse a un plátano y no a las otras dos frutas.

Estos mínimos criterios geométricos permiten por ejemplo realizar un primer filtrado sobre las figuras encontradas, de forma que una figura con alta excentricidad y alta solidez no será considerada como plátano ni como las otras frutas. Por el contrario, figuras donde la redondez sea baja y la solidez muy alta, sí podrán ser filtradas como identificación positiva, a la espera de que filtrados más complejos permitan tomar mejores decisiones.

Los momentos de Hu de la manzana y de la naranja son muy parecidos, como se deduce de los propios patrones utilizados. Por ello se hace necesario el uso del color para poder discriminar entre ambas frutas.

3.2.- Características de las imágenes

En el capítulo 2, sección 3, se han comentado las diferentes representaciones de color que se pueden utilizar para medir el color en las imágenes de estudio. Tal y como explica [25], aquellas representaciones en las que se separa la cantidad de luz (luminancia o intensidad) de los valores cromáticos son las más eficaces para ignorar los cambios de tonalidad producidos por la luz reflejada en las superficies tridimensionales (Figura 29). Esto implica que tanto la representación CIELab como la HSV serán las más adecuadas.

En la Figura 29 se muestra una manzana con cinco puntos marcados, en lo que se realizará la medición del color según diversas representaciones. El objetivo de esta imagen es mostrar cómo el propio volumen de los objetos (que determina la intensidad lumínica en las distintas zonas del cuerpo del mismo) influye más o menos según la representación elegida.

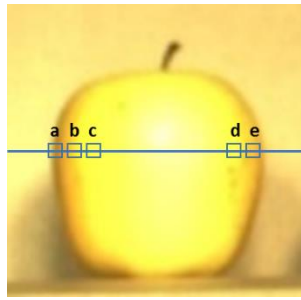


Figura 29. Ejemplo de una manzana iluminada con luz natural y los puntos de medición de color.

Como se puede apreciar en la Tabla 6, los valores de la tonalidad (H) en la representación HSV son bastante estables, percibiendo de forma repetida el color amarillo. El resto de las representaciones también detectan tonalidades cercanas al amarillo, si bien en un rango más amplio. De la representación CMY y de la RGB, si se consideran todas las componentes, el color resultante se muestra en la Tabla 6. De la representación CIELab se considera la combinación de las componentes 'a' y 'b'.

Tabla 6. Para los puntos marcados en la Figura 29 se indica su posición relativa en la imagen y los valores en las representaciones RGB, HSV, CMY y CIELab, así como el nombre del color según tablas estándares.

	X	Y	RGB	Color medido	HSV	Color medido	CMY	Color medido	Lab	Color medido
a	18%	50%	(208,8192 145,5616 42,1632)	verde oliva	(0,1034 0,7981 0,8157)	Amarillo suave	(0,1843 0,4314 0,8353)	verde oliva	(0,6483 15,227 60,0299)	naranja
b	27%	50%	(255,0016 255,0016 70,272)	amarillo intenso	(0,1667 0,7244 0,9961)	Amarillo	(0,0039 0,0039 0,7255)	amarillo intenso	(0,9698 -20,0351 81,9691)	Amarillo / naranja
c	36%	50%	(255,0016 256,0000 87,3472)	amarillo claro	(0,1677 0,6588 1,0000)	Amarillo	(0,0039 0,0000 0,6588)	amarillo claro	(0,9734 -19,7122 76,4912)	Amarillo / naranja
d	75%	50%	(254,0032 254,0032 77,312)	amarillo intenso	(0,1667 0,6957 0,9922)	Amarillo	(0,0078 0,0078 0,698)	amarillo intenso	(0,9669 -19,6549 79,3607)	Amarillo / naranja
e	83%	50%	(249,984 190,7456 60,2368)	naranja suave	(0,1146 0,759 0,9765)	Amarillo suave	(0,0235 0,2549 0,7647)	naranja suave	(0,8035 9,5413 69,5298)	Naranja

En este estudio se ha probado a utilizar el modelo CMY, considerando exclusivamente el peso relativo de la componente Y. En principio, por ser frutas con color predominante amarillo, parece una solución factible y de bajo coste computacional. Pero esta representación es muy sensible a las variaciones de tonalidad debidas a las diferentes condiciones de iluminación creadas por la curvatura de las superficies de los objetos. Se ha utilizado la expresión (33) pero los resultados no han sido muy efectivos.

$$ratio Y = \frac{Y}{C+M+Y} \quad (33)$$

Por esta razón se ha preferido el modelo HSV, donde se ha calculado el valor promedio de la componente H (matiz) de los píxeles de la región encontrada y se ha contrastado con una tabla de valores estándar, adjudicando un nombre de color al objeto encontrado.

En el apéndice 3 se muestra una tabla de colores correspondiente a los valores de las componentes del modelo HSV.

3.3.- Modificación del histograma, binarización y análisis de las regiones

Se parte de una imagen en escala de grises, bien porque se han separado las componentes de color, o bien porque la naturaleza de la imagen es así en su origen.

Por tanto, se procede a estudiar el histograma y a modificarlo de tal manera que se pueda extraer la máxima información del contenido de la imagen.

3.3.1.- Modificación del histograma

Como se ha visto en el capítulo 2.3.2., hay dos operaciones que se utilizan habitualmente en función de las características del histograma. Por un lado, la ecualización es una transformación de los niveles de intensidad según una nueva distribución estadística (uniforme en el caso de este estudio). Por otro lado, está la expansión del histograma, que consiste en extender los niveles de gris de un rango a otro de valores.

Como se verá en los apartados posteriores, en el caso de las imágenes del rango visible y del ultravioleta los histogramas están muy expandidos y presentan mucho contraste, por lo que la mejor opción es la de la expansión del histograma.

En las imágenes térmicas, sin embargo, el histograma es estrecho y con muy poco contraste, por lo que las operaciones aconsejadas son la ecualización y expansión del mismo.

3.3.1.1.- Modificación del histograma de las imágenes del rango visible

Estas imágenes presentan histogramas amplios, con alto contraste y valores extremos altos, indicando la presencia de fuertes luces y sombras. Por lo tanto, se ha optado por una expansión del histograma entre rangos a definir en función de la imagen (Figuras 30 y 31).

Se utilizará la función *imadjust* de Matlab descrita en [26] que redistribuye y transforma una porción del histograma de la imagen original a otro rango con otros extremos.

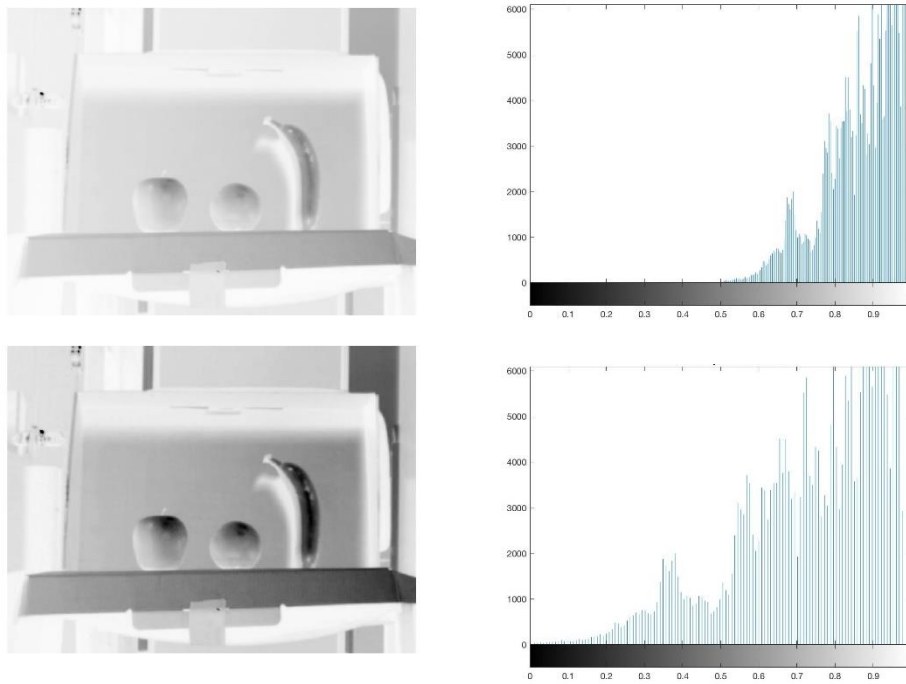


Figura 30. Ejemplo de transformación de la componente Y de CMY y la modificación del histograma. En este caso se expande el histograma para aumentar el contraste de los objetos. Arriba, la imagen original y su histograma; abajo, la imagen modificada junto a su histograma.

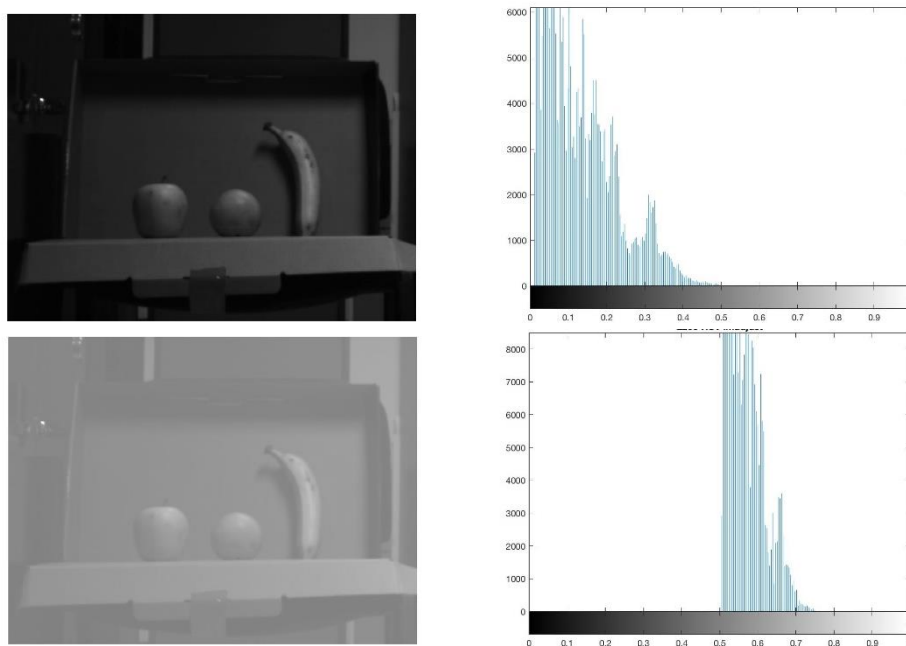


Figura 31. Ejemplo de transformación de la componente V de HSV y la modificación del histograma. En este caso se desplaza y contrae el histograma para mejorar la distinción de los objetos. Arriba, la imagen original y su histograma; abajo, la imagen modificada junto a su histograma.

En el caso de las componentes de color de la imagen en el espectro visible, o de la imagen en el rango ultravioleta, esta función permite seleccionar un rango determinado de intensidades, donde se encuentra la información relevante de los objetos a detectar, y

genera una salida con esos niveles de gris. Esto permite redistribuir el histograma y al mismo tiempo permite eliminar los extremos de saturación, bien sean blancos o negros.

También es importante estudiar el propio gradiente de intensidad que se da en los cuerpos tridimensionales de los objetos a buscar. En este sentido, si la posición relativa de la fuente de luz es tal que satura la superficie de los objetos percibida por el sensor de la cámara, haciendo de ésta una superficie única de alta intensidad rodeada por bordes con tonos de gris más oscuros, será más fácil separar el objeto. En ese caso la función debe contemplar todo el rango de entrada y transformarlo en el rango deseado.

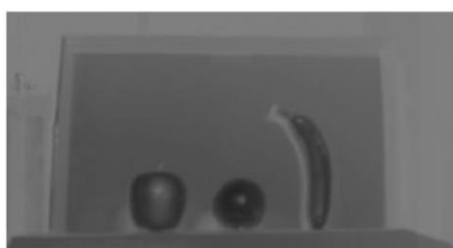
No obstante, si el efecto de intensidad en las superficies de los objetos no es de saturación, entonces es mejor optar por eliminar los extremos del histograma y reducir posteriormente el número de niveles de gris, ya que de esta forma se eliminan los brillos o las sombras colindantes a los contornos de los objetos buscados.

En las imágenes donde la superficie del objeto está saturada por un exceso de iluminación, se mejora el resultado final de la detección amplificando el contraste entre la superficie del objeto y los bordes, los cuales normalmente se encuentran en un rango de intensidades no necesariamente alejado del nivel que satura el objeto.

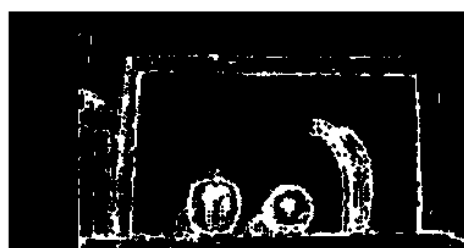
Para conseguir este efecto se utiliza la función *imgradient* de Matlab, descrita en [26], la cual devuelve para cada pixel de la imagen la magnitud y orientación del gradiente de intensidad. Se ha utilizado un operador Roberts para el cálculo del gradiente [35].

La mejora del contraste se realiza de forma directa buscando los gradientes de la imagen. Puesto que las frutas tienen una componente amarilla predominante, se ha observado empíricamente que la superposición de la componente de intensidades elegida (bien sea V de HSV, L de LAB o Y de CMY) con la componente Cb de YCbCr, da una solución óptima en la búsqueda de objetos.

En la Figura 32 se pueden observar varios ejemplos de detección de bordes en componentes de color. De las imágenes mostradas, la que mejores resultados presenta es la componente Cb de YCbCr. Por ello se ha decidido utilizar dicha componente en lugar de otras.



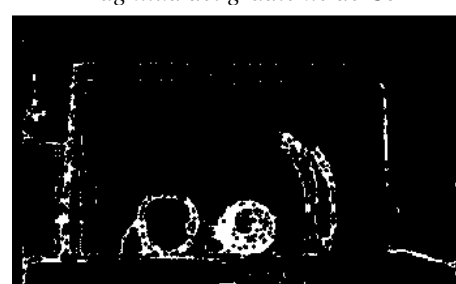
Componente Cb de YCbCr



Magnitud del gradiente de Cb



Componente Cr de YCbCr



Magnitud del gradiente de Cr

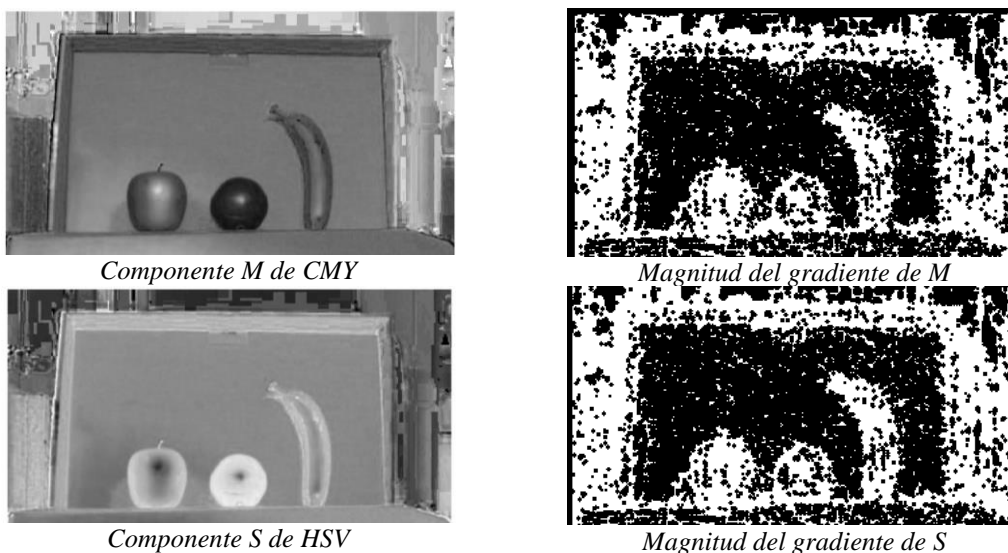


Figura 32. Ejemplos de tratamiento de algunas componentes de color con la detección de contornos por gradientes.

Una vez obtenida la máscara binaria que contiene los bordes, se realiza una operación de enmascaramiento binario tipo NAND para obtener la imagen sobre la que se buscan las regiones (expresión 34).

$$NAND(x, y) = ImagenVisible(x, y) * (max(Bordes) - Bordes(x, y)) \quad (34)$$

Como se verá en el capítulo 4, esta forma de proceder es muy eficiente cuando los objetos buscados presentan iluminación directa que los satura de forma que las superficies tridimensionales parecen planas por el efecto de la intensidad saturada uniforme.

3.3.1.2.- Modificación del histograma de las imágenes del rango ultravioleta

En términos generales estas imágenes presentan un comportamiento similar al de las imágenes del espectro visible, por lo que se aplicará la técnica descrita en el apartado anterior (Figura 33).

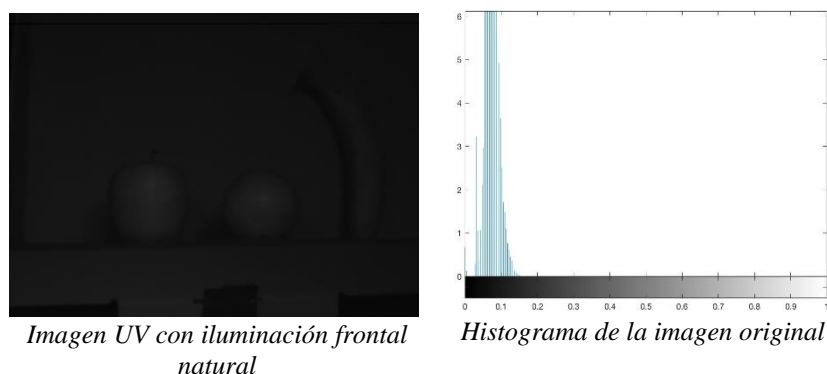
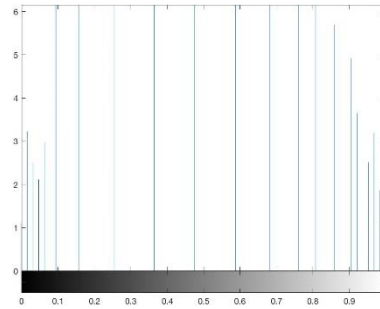




Imagen UV con iluminación frontal natural ecualizada



Histograma ecualizado de la imagen original

Figura 33. Ejemplo de histograma modificado para una imagen UV en una escena donde los objetos están iluminados por luz natural frontal.

No obstante, las imágenes en el rango ultravioleta son matrices en escala de grises, por lo que no se puede extraer información adicional que mejore la detección de ninguna hipotética componente de color. Por ello la función para la detección de bordes y el enmascaramiento con la imagen original no resulta útil ni efectivo.

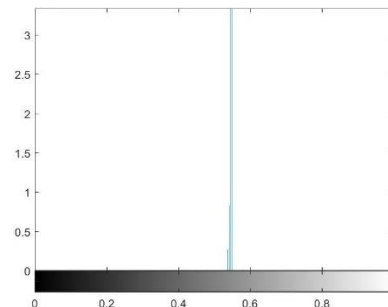
3.3.1.3.- Modificación del histograma de las imágenes térmicas

Las imágenes térmicas presentan poco contraste y sus histogramas son estrechos, como se aprecia en la Figura 23 (histograma de la Figura 16).

Para poder extraer más información de estas imágenes es conveniente ecualizar el histograma (modificar la distribución de intensidades) y expandirlo, esto es, mejorar el contraste (Figura 34).



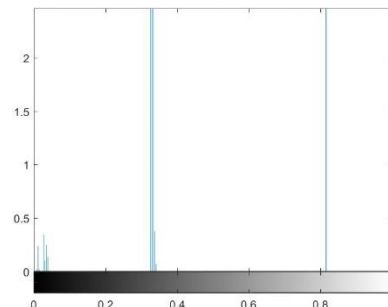
Imagen térmica original



Histograma de la imagen térmica



Imagen térmica ecualizada



Histograma ecualizado de la imagen térmica

Figura 34. Imagen térmica cuyo histograma ha sido modificado, escena en la que los objetos están iluminados frontalmente por luz natural.

En este caso se ha procedido a utilizar la función de Matlab *adapthisteq*, descrita en [26], que realiza una ecualización del histograma junto con una mejora del contraste y que, además, tiene la ventaja de que puede trabajar por zonas. Esta división de la imagen en una matriz de celdas permite mejorar localmente el contraste y aumentar la definición de los contornos, especialmente en las imágenes en el rango visible.

La mejora del contraste tras aplicar este procedimiento es suficientemente buena como para no aplicar el método de resalte de contornos explicado anteriormente.

3.3.2.- Binarización

La fase siguiente es la de binarización. Tal y como se explicó en el capítulo 2 sección 2.3.3, se han considerado dos estrategias: una con funciones de binarización directas siguiendo el método de Otsu [27], y otra con una cuantización previa de los niveles de gris y la selección individual de cada píxel en función de su intensidad.

En las imágenes a color el histograma se ha modificado en el paso anterior para facilitar la separación en términos de intensidad de los objetos y sus contornos. No obstante, estas imágenes siguen presentando una alta variación en la distribución de intensidades en toda su superficie. Si a esto se suma que en estos casos la superposición de los bordes mejora notablemente la diferenciación de los objetos con su entorno, la conclusión es que el método de Otsu [27] es el más eficaz.

En particular este método se ha aplicado con la función *imbinarize* de Matlab, parametrizada para que el proceso de binarización no sea global sino local.

En el caso de las imágenes UV, la eficacia de la modificación del histograma no es grande debido al alto contraste que presentan y a no poder utilizar las componentes de color para mejorar la discriminación de las regiones. Por ello, la binarización no es realmente efectiva.

En el caso de las imágenes termográficas, por el contrario, el método más eficaz ha sido la cuantización y selección de niveles de intensidad píxel a píxel. La ecualización del histograma permite separar los niveles de gris, pero el hecho físico de que los propios niveles de intensidad se deban a la emisividad térmica de los objetos, que a su vez progresa espacialmente de una forma homogénea y gradual, implica que entre dos puntos próximos de la imagen existirá una escala progresiva de intensidades. Por lo tanto, el uso del método de Otsu (que busca distribuciones bimodales en la función de densidad de intensidades) conlleva peores resultados que la selección de umbrales de intensidad a partir de los cuales clasificarlos de forma binaria. De esta forma, si se escoge el rango de intensidades a considerar como aquellas que incluyan la intensidad mayoritaria del objeto, térmicamente homogéneo, el resultado de toda la imagen será el propio objeto.

Esta última operación es básicamente una función lógica o función de paso banda como la presentada en la expresión (35), donde IT corresponde al valor de la Imagen Térmica.

$$PASO(x, y) = Si(IT(x, y) > Limite INF) * Si(IT(x, y) < Limite SUP) \quad (35)$$

3.3.3.- Análisis de regiones

Partiendo de una imagen binaria, el análisis de regiones hace referencia al proceso de decisión sobre si una región, en base al cálculo de las características seleccionadas, puede ser clasificada como uno de los objetos buscados u otro distinto.

En este estudio se comienza con una descripción básica inicial, que clasifica los objetos como manzanas o naranjas si tienen una redondez inferior al 15%, y como plátanos si el máximo de los valores de su excentricidad o su inverso son superiores a 1,5 (así se contempla la posibilidad de que la orientación del plátano sea vertical u horizontal). Estos valores han sido calculados a partir de los valores medios de los patrones estudiados del conjunto de imágenes disponibles. Para el valor de la redondez se ha tomado como valores de partida los presentes en la Tabla 5. Por ello, en la discriminación de las frutas se ha otorgado a las frutas más “redondas” un valor máximo admisible entre 0 (circunferencia perfecta) y el valor medio del plátano. Respecto a la excentricidad, se toman valores superiores al de referencia del plátano (1.5 en la Tabla 5) como el límite mínimo a partir del cual se considera que es plátano y no manzana o naranja.

Este filtrado inicial da como resultado un gran número de falsas detecciones por su simplicidad. No obstante, se mantiene como base para un segundo filtrado en el que se evalúan más propiedades, comparándolas con las de los patrones.

Por un lado, se mantiene el criterio de redondez con el valor comentado anteriormente, y además el de excentricidad se restringe a un rango de hasta 3.2 ya que es el valor máximo obtenido a partir de las muestras estudiadas y excluye, por ejemplo, los casos de rectas que cumplan con los criterios iniciales y se deban a sombras o filtrados inadecuados. Este último criterio lo puede cumplir la silueta de un plátano y un rectángulo perfectamente sólido. Aun así, también se añade un criterio de solidez, restringido al rango [0.72,0.92]. Los valores tomados vuelven a surgir de la Tabla 5, donde se aprecia que una manzana o una naranja tienen valores muy altos de solidez, ya que son figuras convexas en las que la propia región y su envoltorio convexo coinciden. No así en el plátano, donde la parte cóncava de esta fruta establece que dichas superficies no coincidan y que, por tanto, los valores de solidez bajen. El rango se ha establecido tomando como base el valor del patrón plátano y como límite superior el menor valor de solidez obtenido con las muestras de las imágenes tomadas para naranjas y manzanas.

Por último, se exigirá un valor máximo para cada uno de los patrones de un 40% como mínimo, obtenido a partir de los momentos de Hu (expresión 36).

$$Distancia\ normalizada\ a\ los\ momentos = \frac{\sqrt{\sum_1^7 (\mu_i^* - \mu_i)^2}}{\sqrt{\sum_1^7 (\mu_i)^2}} \quad (36)$$

μ_i^* = momento Hu i de la región

μ_i = momento Hu i del patrón

Se calcula la distancia normalizada a los momentos con cada uno de los tres patrones, y si el valor mínimo de los tres obtenidos es inferior al 40%, la clasificación se considera correcta y se clasifica como la fruta con cuyo patrón la distancia es máxima. Se ha elegido este valor porque los momentos H_u son valores bastante explícitos respecto a la figura analizada y se ha considerado como región correcta aquella que está por debajo del 50% de distancia.

3.4.- Resultados del análisis cuantitativo

En este estudio se ha analizado cuantitativamente cada zona del espectro de forma separada para encontrar el algoritmo que resulta más óptimo en cada caso.

Para las imágenes a color se ha estudiado la detección de los objetos considerando una sola de sus componentes según los distintos modelos de color, de manera que se realiza sobre una imagen de escala de grises, y la propiedad del color se selecciona posteriormente. Se ha podido observar que las imágenes en el rango del visible presentan el problema de la iluminación, ya que los brillos y las sombras hacen difícil la distinción de los objetos. De este modo, se ha seleccionado el modelo de color HSV, tomando la componente V, para posteriormente aplicar las técnicas anteriormente descritas.

En cuanto a las imágenes infrarrojas, éstas presentan la ventaja de que registran las emisividades de los objetos en la escena. Físicamente, la emisividad de un cuerpo depende de la temperatura a la que éste se encuentra. Puesto que los objetos están sometidos a la misma temperatura, presentan una homogeneidad y su representación en la escena también es uniforme en intensidad de los píxeles que la componen. Así mismo, no hay fuentes de calor relevantes, por lo que no hay puntos que acumulen una emisividad muy diferente respecto al entorno. Por todo ello, todos los objetos que forman parte de la escena presentan una variación pequeña de temperaturas, y por lo tanto de emisividades. Esto conlleva que el histograma de la imagen es muy estrecho, teniendo muy poco contraste.

También hay que reseñar que la extracción de contornos es muy eficiente, puesto que ninguno de los objetos presenta partes con diferentes emisividades, lo que implica que no aparecerán brillos ni sombras en sus contornos. Por esta misma razón se pueden omitir, en el caso de la imagen tomada en el rango IR, las transformaciones morfológicas que en otros casos puedan ser necesarias en la imagen binarizada. Así pues, en este caso se llevan a cabo las técnicas descritas en este capítulo 3 sección 3, de manera que los objetos puedan ser extraídos con mayor facilidad.

Por último, en el caso de las imágenes UV, donde no se proporciona información de color, se limita el análisis exclusivamente a factores geométricos, como ocurre en el caso de las imágenes térmicas. Aunque en este caso se trata de una imagen de intensidades, sin color, presenta las mismas dificultades que la imagen en el rango visible respecto a los brillos y las sombras.

De todo lo anterior se puede concluir que en las imágenes del rango visible y UV, la presencia de brillos y sombras no permite extraer unos contornos bien definidos para las regiones o formas a estudiar, por lo que su comparación geométrica con los patrones seleccionados es pobre y de poca eficacia. Por el contrario, las imágenes en UV e IR no

transmiten información del color, con lo que algunas de las regiones extraídas pueden corresponder geoméricamente con los objetos buscados, pero no coincidir en color, proporcionando falsas identificaciones.

Por todo lo anterior, se propone realizar una combinación de las imágenes térmicas y de espectro visible, de modo que las características geométricas serán extraídas de la imagen IR, y la información de color será obtenida de la imagen a color.

4.- Análisis y discusión de resultados

La estructura de este capítulo se va a orientar a la naturaleza de las imágenes en primer lugar. En cada apartado se comentarán los resultados de las diferentes técnicas utilizadas según el tipo de imagen. En el caso de las imágenes del rango visible, para las que se han investigado diferentes tratamientos en función de la representación de color utilizada, se presentarán los resultados más óptimos para cada caso, indicando cuál de todas las componentes de color es la más eficiente.

Para las imágenes UV y térmica se presentarán los resultados del procesado de imágenes explicado, indicando la eficiencia de cada una de las decisiones tomadas.

Para este estudio se han tomado 25 fotografías de cada una de las tres cámaras (75 imágenes del conjunto total de 480), con diferentes iluminaciones: sol directo, iluminación artificial e iluminación natural junto con la artificial. También se ha utilizado iluminación frontal a los objetos e iluminación lateral.

4.1.- Resultados cualitativos en imagen visible

El primero de los casos que se presenta son los resultados cualitativos del tratamiento hecho sobre una imagen en el espectro visible. Se expondrán los resultados obtenidos con la imagen descomponiéndola en diferentes componentes de color.

Como ya se ha indicado en apartados anteriores, las imágenes del rango visible se han estudiado según varias representaciones de color:

- HSV, de la que se ha tomado la componente V o luminancia
- CMY, de la que se ha tomado la componente Y o amarillo
- YCbCr, de la que se ha tomado la Cb o cromaticidad azul
- Lab, de la que se ha tomado la componente L o intensidad lumínica

Cada una de las componentes analizadas es una imagen en escala de grises y, por lo tanto, se procesa como se ha explicado en el capítulo 3 sección 3 (Figura 26).

En las siguientes imágenes de la Figura 35 se pueden observar las componentes de partida de algunas de las imágenes estudiadas.



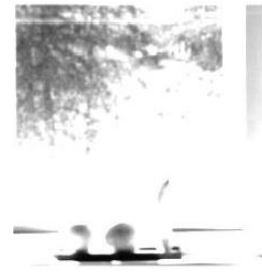
A1.- Componente V de HSV de una imagen con luz natural por detrás de los objetos.



A2.- Componente Y de CMY de una imagen con luz natural por detrás de los objetos.



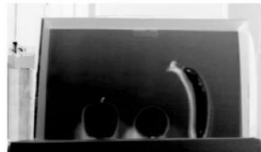
A3.- Componente Cb de YCbCr de una imagen con luz natural por detrás de los objetos.



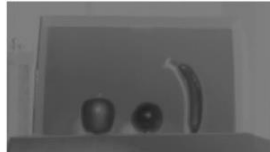
A4.- Componente L de Lab de una imagen con luz natural por detrás de los objetos.



B1.- Componente V de HSV de una imagen con luz natural frontal a los objetos.



B2.- Componente Y de CMY de una imagen con luz natural frontal a los objetos.



B3.- Componente Cb de YCbCr de una imagen con luz natural frontal a los objetos.



B4.- Componente L de Lab de una imagen con luz natural frontal a los objetos.



C1.- Componente V de HSV de una imagen con luz artificial frontal a los objetos.



C2.- Componente Y de CMY de una imagen con luz artificial frontal a los objetos.



C3.- Componente Cb de YCbCr de una imagen con luz artificial frontal a los objetos.



C4.- Componente L de Lab de una imagen con luz artificial frontal a los objetos.



D1.- Componente V de HSV de una imagen con luz natural y artificial frontal a los objetos.



D2.- Componente Y de CMY de una imagen con luz natural y artificial frontal a los objetos.



D3.- Componente Cb de YCbCr de una imagen con luz natural y artificial frontal a los objetos.



D4.- Componente L de Lab de una imagen con luz natural y artificial frontal a los objetos.

Figura 35. Ejemplo de las componentes de color de varias imágenes según el tipo de iluminación utilizada.

De las imágenes mostradas en la Figura 35 se pueden destacar varias consideraciones, que han determinado la eficacia de los resultados, y que son explicadas a continuación.

4.1.1.- Obtención de las componentes de color

En primer lugar, se debe destacar que la ubicación de la fuente de luz es determinante en la capacidad del algoritmo propuesto para separar los objetos de su entorno. En concreto, las sombras en los propios objetos y en los elementos de apoyo van a limitar notablemente la detección de los contornos y, por lo tanto, la correlación de las formas en la imagen binaria.

Esto es particularmente cierto cuando la fuente de luz se encuentra detrás de los objetos, como ocurre en la Figura 17. Igualmente ocurre cuando se encuentra por encima de los objetos, Figura 21. En estos casos, con las sombras bien en la peana o bien en las propias frutas, los resultados visibles del algoritmo han sido pobres, como se explicará más adelante. No obstante, cuando la fuente de luz se encuentra delante de los objetos, los resultados que se aprecian son mejores. En particular con la manzana y la naranja, donde su forma quasi-esférica permite que la iluminación frontal no cree sombras, sino patrones concéntricos (Figura 36) lo que, con la adecuada manipulación del histograma, da lugar a patrones binarios circulares. El plátano, por el contrario, es más complicado ya que es más parecido a un cuerpo poliédrico, y cualquiera que sea la fuente de luz generará sombras.

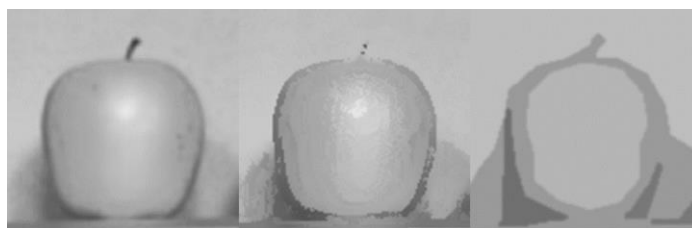


Figura 36. Ejemplo del tratamiento sobre una manzana considerando su forma tridimensional y los efectos de la reducción del espacio de grises.

Por otro lado, en las escenas presentadas en la Figura 35, las que utilizan luz natural son las que dan visualmente mejores resultados de detección en las imágenes finales, ya que la intensidad lumínica ha sido mayor y, por lo tanto, sus histogramas estaban más saturados en las intensidades altas. Esto ha eliminado las zonas de intensidad progresiva o degradada en el propio cuerpo de las frutas, mejorando la detección de los contornos.

Respecto a las componentes a utilizar, cada una de ellas presenta unas necesidades particulares de computación, si bien el conjunto de pasos de procesado del capítulo 3, sección 3, es el que mejores resultados da.

Se puede apreciar en la Figura 35 que las componentes V de HSV e Y de CMY son equivalentes en todos los casos, obteniendo la misma eficacia de detección. Se puede afirmar que cuando la imagen es iluminada por fuente de luz natural, la componente Y es más similar a la de otras fuentes de luz. Y si se obtiene el “negativo” o complementario de la misma (máximo de intensidad – valor del píxel), el resultado es equivalente a la componente de intensidad de HSV.

Por sí misma, la componente Cb, tras el proceso descrito, no permite detectar regiones correctamente. No obstante, cuando la luz es frontal a los objetos y satura su superficie se podrá realizar la combinación de la componente V de HSV en combinación con la

componente Cb a la que se le aplique un método de detección de bordes, tal y como se ha descrito en el capítulo 3, sección 3, subsección 2.

La componente L de Lab no resulta eficaz en absoluto, ya que satura la imagen en valores máximos y la pérdida de información es notable en todos los casos, impidiendo la detección de los objetos.

4.1.2.- Modificación del histograma

La primera de las operaciones que sigue a la obtención de la componente de color es la ecualización del histograma. Las Figuras 37-40 muestran los histogramas de las imágenes ejemplo que se han explicado previamente en la Figura 35, en particular los correspondientes a las figuras A1, A2, A3 y A4.

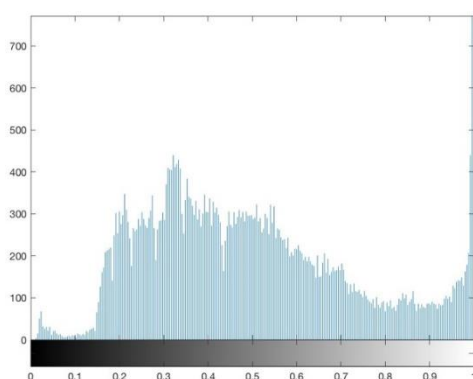


Figura 37. Histograma de la componente V de HSV de la imagen con luz natural por detrás de los objetos (A1).

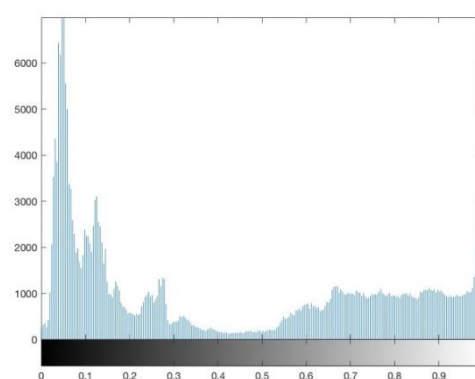


Figura 38. Histograma de la componente V de HSV de la imagen con luz natural por delante de los objetos (A2).

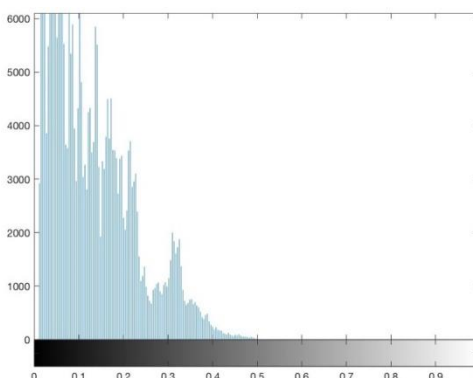


Figura 39. Histograma de la componente V de HSV de la imagen con luz artificial por encima de los objetos (A3).

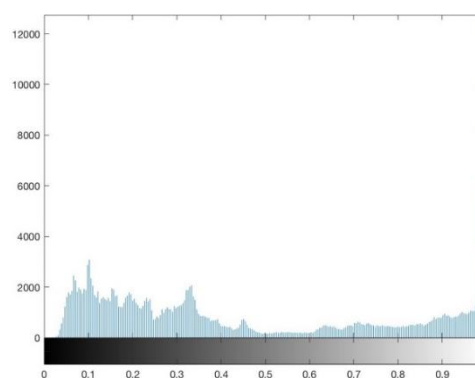


Figura 40. Histograma de la componente V de HSV de la imagen con luz natural frontal a los objetos y luz artificial por encima de ellos (A4).

Aunque sólo se muestran los histogramas de la componente V de HSV (ya se ha explicado que con Y de CMY se obtienen resultados similares), se aprecian las diferencias antes citadas.

En las imágenes con luz natural por detrás de los objetos (Figura 35, A1 y A2), aun cuando hay saturación en altas intensidades, la amplia distribución del histograma anticipa la dificultad que existirá en resolver los contornos de los objetos buscados.

También se comprueba algo similar en el histograma que corresponde a la iluminación con luz artificial (Figura 35, C1), donde no hay saturación en los niveles altos de intensidad, y el histograma en las intensidades bajas es muy uniforme, lo que indica que los gradientes de intensidad al ser procesados darán lugar a falsos contornos, y por tanto falsas detecciones (Figura 41).

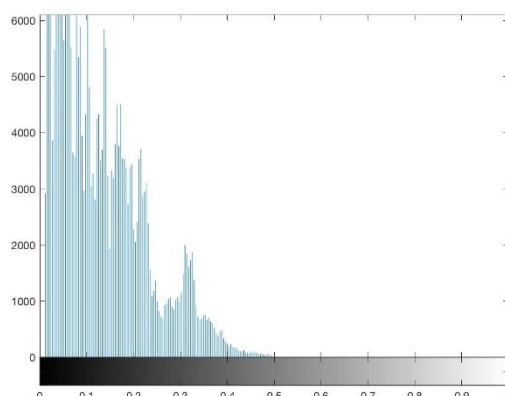


Figura 41. Histograma correspondiente a la componente V de una imagen tomada con luz artificial por encima de los objetos.

Como se ha comentado anteriormente (Figura 36), la naturaleza esférica (tridimensional) de los objetos es la razón por la cual, en ausencia de altos contrastes de luz en los mismos objetos, esto es, ausencia de fuerte iluminación directa y frontal, el volumen de los mismos aparece como una gradación de sombras, siendo el algoritmo (capítulo 3, sección 3, subsección 3) incapaz de definir correctamente los contornos.

Viendo los histogramas de las Figura 37 y 39 se puede deducir que los píxeles correspondientes a altas intensidades mejorarán la detección de los objetos, ya que éstos corresponden a las superficies de los objetos debido a que la iluminación es frontal a los mismos.

El procesado de los histogramas se ha realizado según lo explicado en el apartado 3.3.3.

Para facilitar la detección de los objetos y sus contornos, los histogramas (Figuras 42 a 45) han sido modificados (capítulo 3, sección 3) para incrementar los rangos de alta intensidad y saturar los objetos, de forma que su volumen no afecte a la detección.

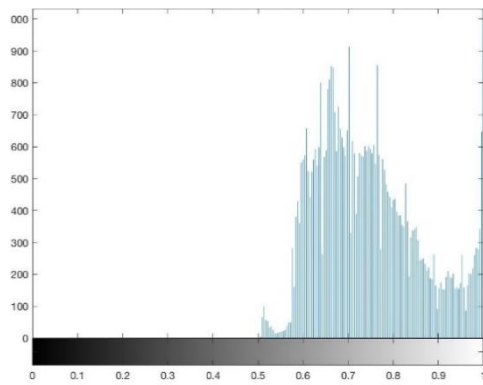


Figura 42. Histograma modificado de una imagen con luz natural por detrás de los objetos para la componente V de HSV

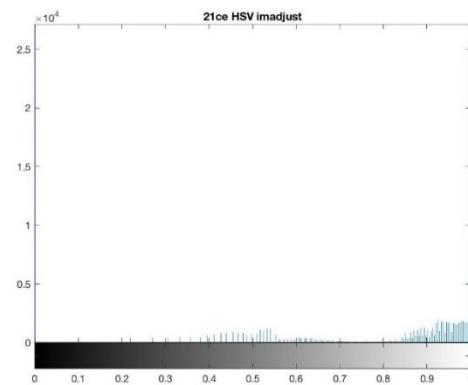


Figura 43. Histograma modificado de una imagen con luz natural por delante de los objetos para la componente V de HSV

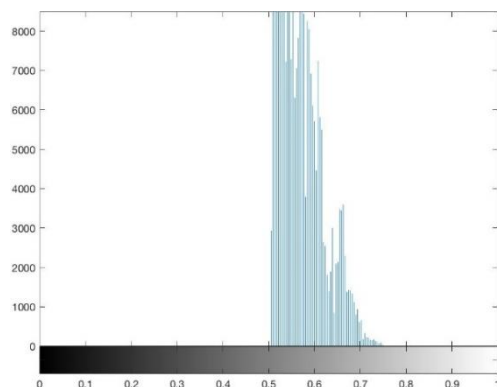


Figura 44. Histograma modificado de una imagen con luz artificial por encima de los objetos para la componente V de HSV.

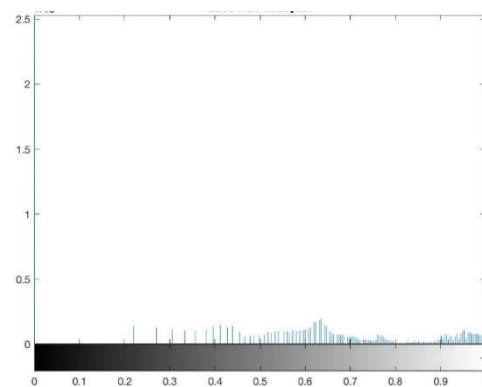


Figura 45. Histograma modificado de una imagen con luz natural por delante de los objetos y luz artificial por encima para la componente V de HSV.

En las imágenes que corresponden a estos histogramas modificados (Figuras 46 a 49) se aprecia que cuando la luz satura la superficie de los objetos, estos son más difíciles de ver a simple vista, pero son más homogéneos en intensidades ya que hay menos gradientes de grises.

La utilización de funciones gradiente [35] para detectar contornos en las componentes Cb de YCbCr y su combinación con las imágenes anteriores resulta de especial utilidad en los casos en los que la gradación de grises en los objetos es mínima, esto es, las que tienen saturación en altas intensidades.

La operación realizada para obtener la imagen sobre la que buscar las regiones consiste en una operación de enmascaramiento binario tipo NAND según la ecuación (34).

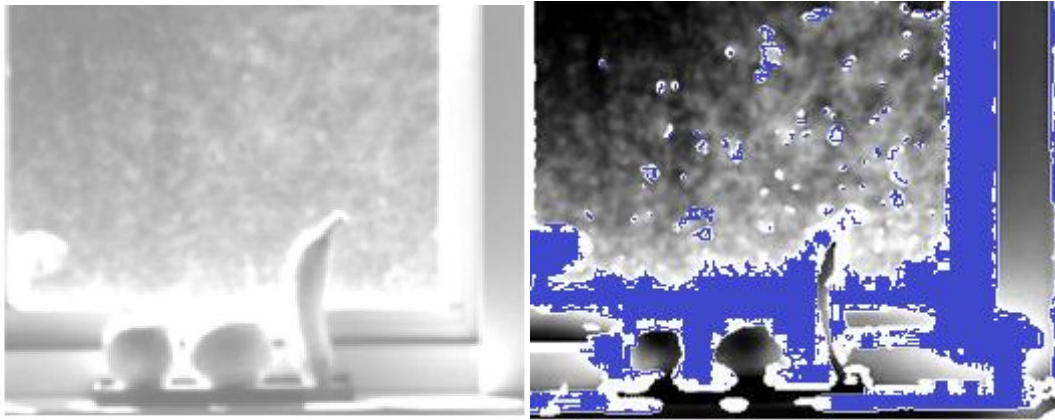


Figura 46. Imágenes resultantes de modificar el histograma en la componente V de HSV. En azul los bordes detectados para facilitar la comprensión de la máscara superpuesta.

Observando la Figura 46, a la izquierda está la imagen sin combinar con la detección de contornos por gradientes. A la derecha se ve la utilización de gradientes. La detección de contornos con el uso de los gradientes reduce considerablemente los objetos impidiendo la correcta identificación de los mismos.

Por otro lado, en las Figuras 47 y 49 se observa que la saturación en altas intensidades, junto con la detección de bordes, permite una detección bastante precisa.

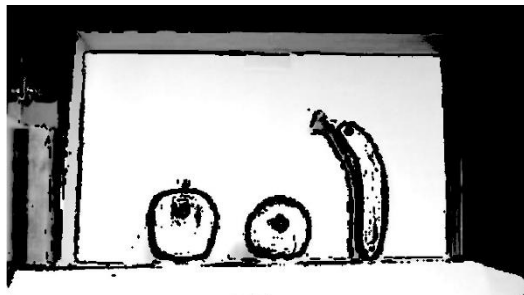


Figura 47. Imagen resultante de modificar el histograma (ecuación 31) en la componente V de HSV para una escena iluminada con luz natural por delante de los objetos.

Pero cuando las imágenes se han obtenido con luz no frontal a los objetos y sin saturar sus superficies, el resultado de aplicar esta función no es tan efectivo, tal y como se aprecia en la Figura 48.

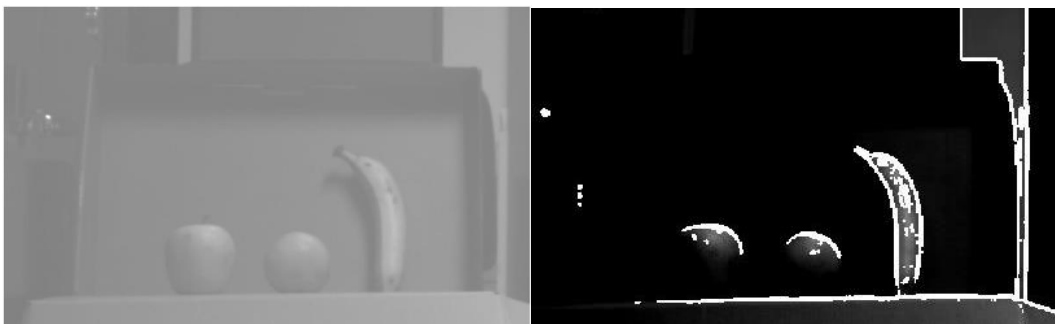


Figura 48. Imágenes resultantes de modificar el histograma en la componente V de HSV. A la izquierda la imagen sin combinar con la detección de contornos por gradientes.

La Figura 49 representa una imagen tomada con luz natural frontal a los objetos y artificial por encima de ellos. El hecho de que la luz natural incida directamente sobre los mismos de forma frontal saturando sus superficies mejora el resultado de las funciones gradiente.

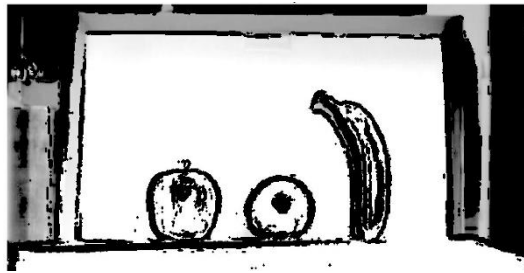


Figura 49. Imagen resultante NAND de modificar el histograma en la componente V de HSV para una escena iluminada con luz natural y artificial.

4.1.3.- Binarización de la imagen

El proceso de binarización será más sencillo en los casos de luz natural frontal a los objetos, tal y como se puede observar en las Figuras 50 a 53.



Figura 50. Imagen binarizada de una escena iluminada con luz natural por detrás de los objetos. Componente V de HSV.

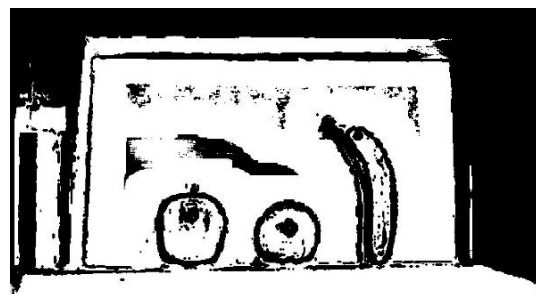


Figura 51. Imagen binarizada de una escena iluminada con luz natural por delante de los objetos. Componente V de HSV.



Figura 52. Imagen binarizada de una escena iluminada con luz artificial por encima de los objetos. Componente V de HSV.

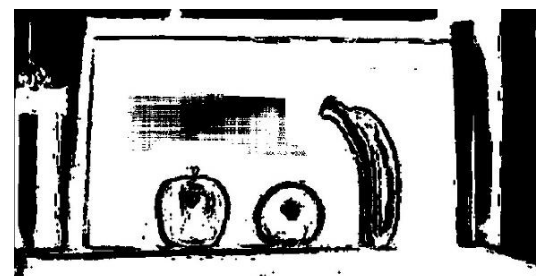


Figura 53. Imagen binarizada de una escena iluminada con luz natural por delante de los objetos y luz artificial por encima de ellos. Componente V de HSV

4.1.4.- Detección de regiones y filtrado

Una vez binarizadas las imágenes, se procede a realizar un proceso de etiquetado [22] de las regiones blancas, de manera que se asigna la misma etiqueta a cada uno de los píxeles de una región conexa. Este procedimiento no es más que agrupar los píxeles de una imagen basándose en la conectividad, de manera que permita extraer las características de cada región que posteriormente serán evaluadas.

Una vez etiquetadas las diferentes regiones, se procede a realizar dos filtrados con el fin de eliminar aquellos objetos que no tenga similitud con las frutas. El primer filtrado sencillo se basa en características geométricas de forma, mientras que el segundo es más complejo e integra correlaciones con momentos invariantes y otras ratios de forma (ver capítulo 3, sección 3, subsección 3). Como se puede apreciar en las Figuras 55 y 57, los objetos serán detectados con mayor eficacia en estas dos imágenes que en las otras (Figuras 54 y 56).

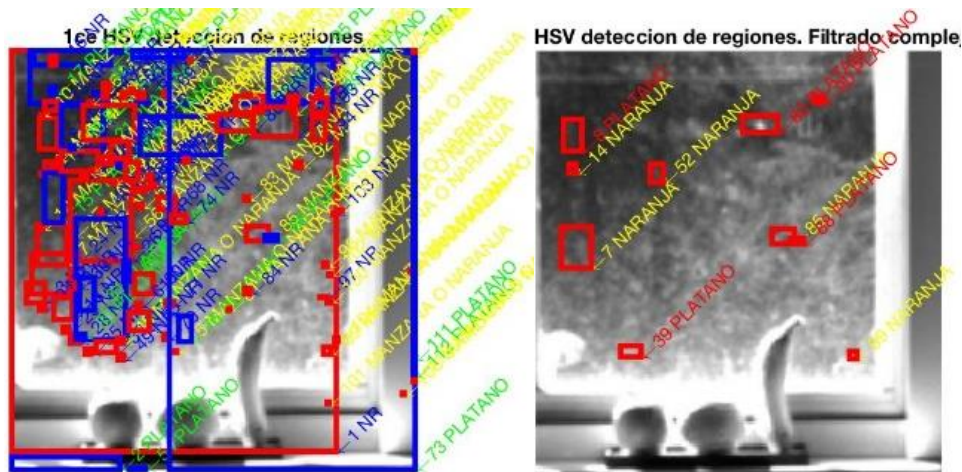


Figura 54. Objetos identificados con un filtrado simple (izquierda) y complejo (derecha) en una imagen con luz natural por detrás de los objetos. Componente V de HSV

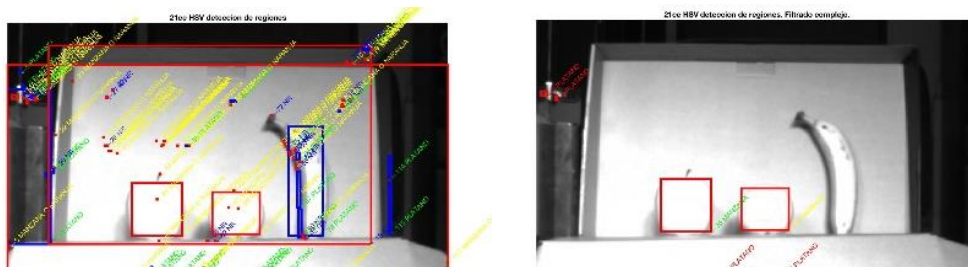


Figura 55. Detección de objetos con filtrado simple (izquierda) y complejo (derecha) de una imagen con luz natural frontal. Componente V de HSV

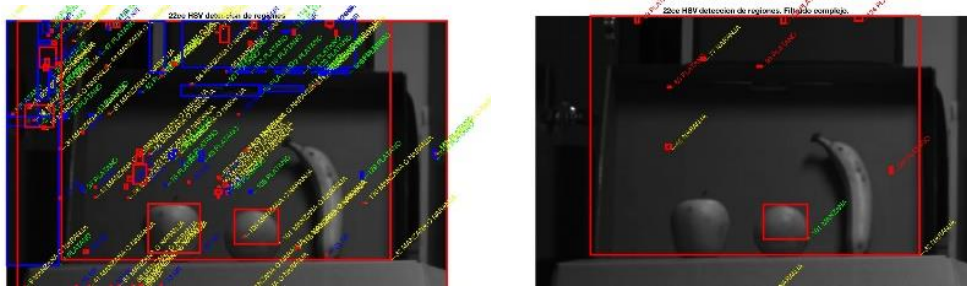


Figura 56. Detección de objetos con filtrado simple (izquierda) y complejo (derecha) en una imagen con luz artificial por encima de los objetos. Componente V de HSV.

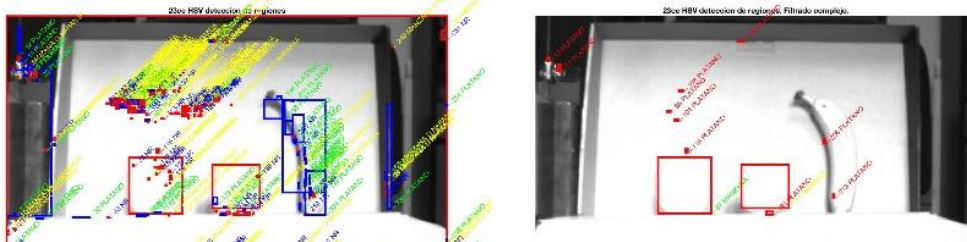


Figura 57. Detección de objetos con filtrado simple (izquierda) y complejo (derecha) de una imagen con luz natural frontal y luz artificial por encima de los objetos. Componente V de HSV

Como puede observarse en las Figuras 54 a 57, los primeros filtrados simples dan lugar a múltiples detecciones que no se corresponden con los objetos buscados. No obstante, dichas detecciones son eliminadas en su mayor parte con el segundo filtrado complejo.

También cabe destacar que el filtrado simple falla en mayor medida cuando los objetos están iluminados por detrás, y se debe a que los objetos aparecen conectados en la imagen binaria, de manera que no puede hacer una división de forma eficaz de los propios objetos y los otros elementos circundantes, tales como la peana o los contornos (Figura 54). En el resto de casos con iluminación directa (natural, artificial, o ambas), el resultado del filtrado simple es efectivo en cuanto que detecta casi todos los objetos buscados.

Cuando se aplica el filtrado complejo, muchas de las detecciones creadas en el proceso de binarización, y cuyo origen viene de zonas con degradados de intensidades propios de zonas de brillos o sombras puntuales, desaparecen. Aunque inicialmente hubieran superado los criterios de redondez o de excentricidad, la búsqueda de correlación con los patrones elegidos, así como la solidez de las regiones, determina su exclusión.

Desafortunadamente, en muchos de estos casos, aunque las detecciones iniciales sí mostraban la presencia del objeto “plátano”, el filtrado final lo desecha. La razón está en la propia forma pseudo-poligonal de esta fruta, que crea sombras que desvirtúan la comparación con el patrón utilizado. Esto se puede observar en la diferencia entre la Figura 55 y la 57.

Cuando la saturación en altas intensidades es menor, como en el caso de la iluminación con luz artificial de la Figura 56, se puede comprobar que el número de regiones detectadas es mucho más alto debido a las zonas de grises incorrectamente definidas en la binarización. Se puede comprobar que estas regiones son en su mayoría mucho más pequeñas que los objetos buscados.

No se ha añadido ningún filtro de tamaño relativo para eliminarlas directamente en el filtrado simple, ya que el factor de escala es desconocido cuando se aplica el algoritmo, y se ha intentado que sea lo más generalista posible. En caso de conocer o disponer de un factor de escala o de distancia a los objetos, se podría eliminar estas regiones basándonos en su tamaño reducido.

En las Figuras 55 y 57 se comprueba la eficacia de la detección e identificación en situaciones de iluminación frontal con luz (natural) que satura los niveles altos de intensidad. Por el contrario, cuando la iluminación es indirecta, como en la Figura 56, la detección está muy deformada y la identificación no es correcta.

4.2.- Resultados sobre la imagen UV

La imagen tomada en el rango de los UV comparte muchas de las características que presenta la imagen capturada en el rango visible, pero con la dificultad de que no existe información de color que permita mejorar la detección en condiciones de brillo y sombras.

Los resultados obtenidos en este estudio son, por tanto, similares a los que se obtendrían en el caso anterior.

Sin embargo, es importante resaltar que cuando la luz es artificial fluorescente, aunque ésta sí emite en la zona del espectro del UV, lo hace en menor medida que la luz solar. Por ello las imágenes tomadas con luz de interior son más oscuras, dificultando en gran medida la detección de los elementos de la escena.

Si se aplican los pasos de procesado anteriormente descritos a las imágenes en el rango ultravioleta se obtienen los resultados que se muestran a continuación.

En las Figuras 58 y 59 se muestra el procedimiento descrito aplicado a la imagen ultravioleta de la escena en la que los objetos están iluminados con luz natural por detrás de los mismos.



Figura 58. Ecuilización del histograma y binarización para una imagen ultravioleta con los objetos iluminados por detrás.

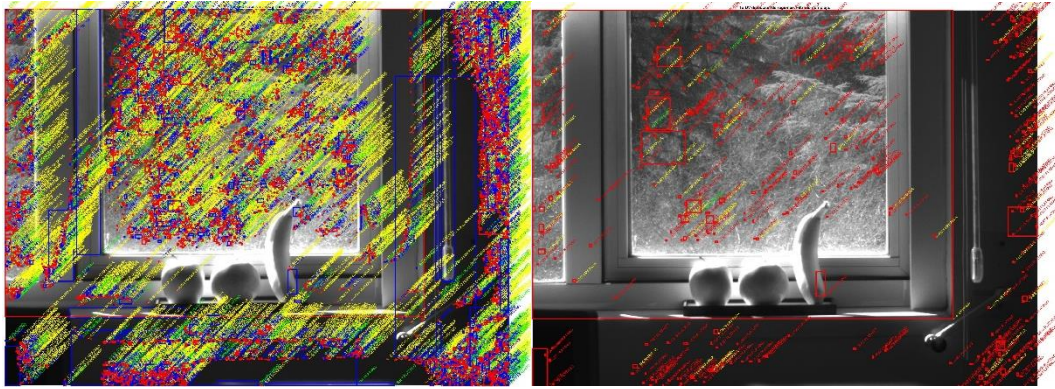


Figura 59. Filtrado simple y complejo de una imagen ultravioleta con iluminación detrás de los objetos.

Como se aprecia en ellas, el contraste de la imagen es elevado y, por tanto, se va a encontrar la misma problemática que ofrecía este tipo de imágenes en el espectro visible. El resultado de las operaciones realizadas muestra multitud de falsas detecciones tanto en el filtrado simple como en el complejo, sin que realmente se encuentren los objetos buscados.

Cuando la iluminación pasa a ser frontal a los objetos, Figuras 60 y 61, el algoritmo mejora al detectar los objetos naranja y manzana, pero el elevado contraste impide reducir el conjunto de falsas detecciones. Por otro lado, la figura del plátano no es detectada. El ruido presente en las imágenes no es filtrado y es el causante de la detección de tantas falsas regiones.

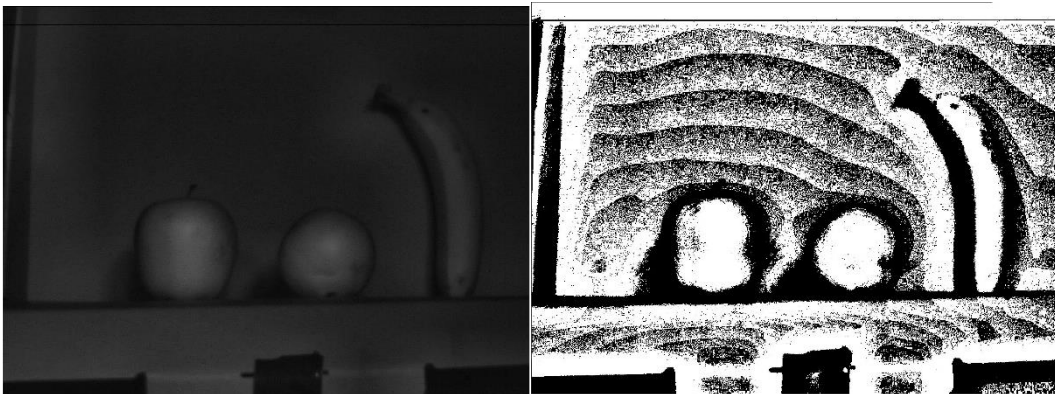


Figura 60. Ecuación del histograma y binarización de una imagen ultravioleta con iluminación natural por delante de los objetos.

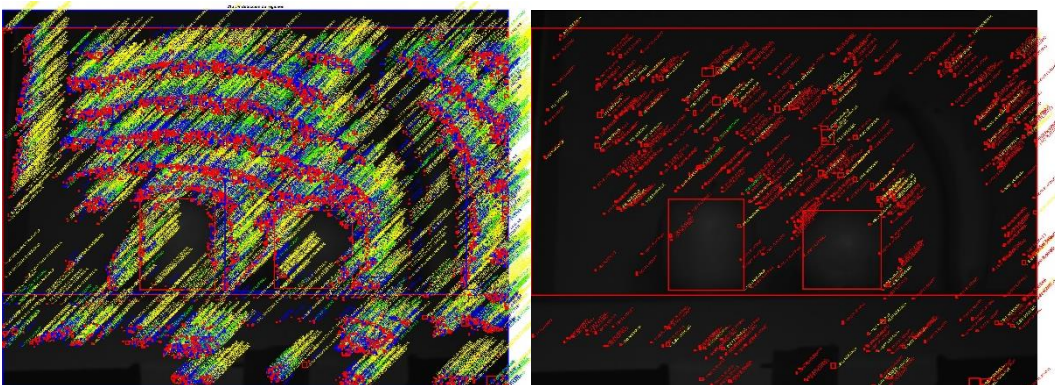


Figura 61. Filtrado simple y complejo para una imagen ultravioleta con iluminación natural por delante de los objetos.

Cuando la iluminación se sitúa por encima de los objetos, la luminosidad se atenúa mucho y el ruido es dominante, lo que conduce a muchas falsas detecciones y ninguna de los objetos buscados, como se muestra en la Figura 62.

Cuando se da la combinación de ambos tipos de iluminación, como se muestra en las Figuras 58 a 59, el resultado es similar a las Figuras 55 y 56, muchas falsas detecciones y únicamente identifica correctamente la manzana y la naranja.



Figura 62. Ecuilización del histograma en una imagen ultravioleta cuando la iluminación es artificial. Aunque los objetos en la imagen inicial no son claramente visibles, la ecualización del histograma los muestra en la imagen de la derecha.

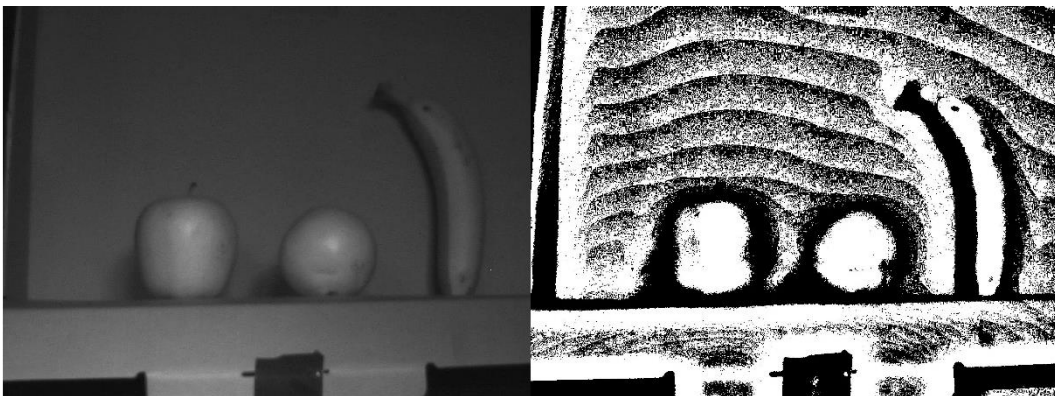


Figura 63. Ecuilización del histograma y binarización de una imagen ultravioleta iluminada con luz natural y artificial.

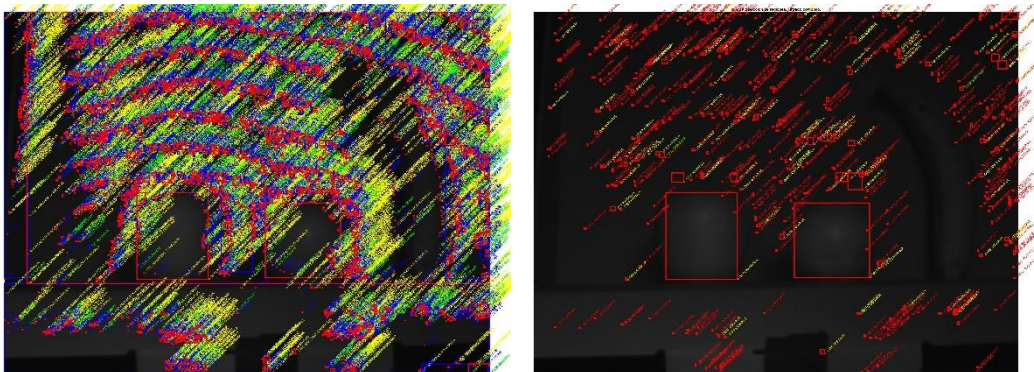


Figura 64. Filtrado simple y complejo de una imagen ultravioleta con iluminación natural y artificial.

La calidad de la imagen ultravioleta proporciona más ruido, por lo que en la ecualización del histograma genera más falsas detecciones. No obstante, cuando la iluminación es frontal a los objetos y de mayor intensidad, al menos se detectan la manzana y la naranja.

4.3.- Resultados sobre la imagen infrarroja.

La homogeneidad de temperaturas de cada objeto hace que estos aparezcan como una región muy uniforme en términos de intensidad y, por lo tanto, que puedan ser diferenciados de su entorno de forma más eficaz.

El resultado al procesar de la forma descrita hasta el momento las imágenes térmicas se presenta a continuación, pudiendo apreciar la bondad del método y su eficacia, tanto en registro geométrico de las regiones (ajuste del cuadrado al contorno del objeto detectado) como en la propia identificación de los mismos.

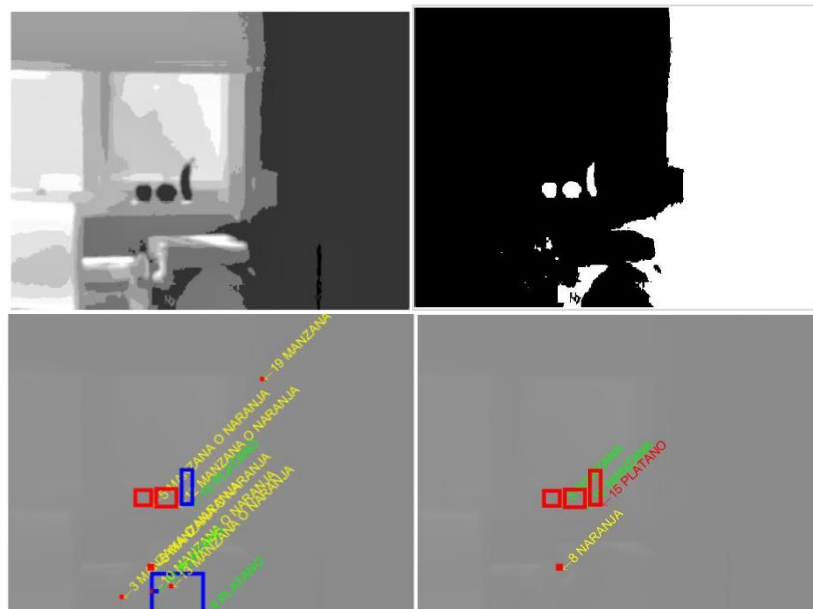


Figura 65. Expansión del histograma (arriba izquierda), binarización (arriba derecha), filtrado simple (abajo izquierda) y filtrado complejo de las regiones (abajo derecha) de la imagen térmica de la escena iluminada con luz natural por detrás de los objetos.

En la Figura 65, fila superior, las frutas presentan en todos los casos una temperatura muy diferente de su entorno, en particular más fría, lo que permite diferenciar los objetos claramente.

Las imágenes proporcionadas por la cámara permiten realizar una expansión del histograma muy eficiente, que lleva a separar notablemente los niveles de intensidad y diferenciar los contornos de los elementos de la escena, aunque en las imágenes originales no fuese evidente tal diferenciación.

Respecto al filtrado de las regiones (Figura 65, fila inferior), tanto el filtrado simple como el complejo identifican perfectamente los objetos buscados. La homogeneidad de las imágenes creada por la propia homogeneidad de las temperaturas de los elementos de la

escena hace que el número de falsas detecciones sea muy inferior a la de la imagen visible o del rango UV. De ahí la bondad de ambos filtrados.

La fuente de luz resulta irrelevante en estos casos (Figuras 66 a 69), ya que la emisividad de los objetos es la que determina, en el caso de esta cámara, el contenido de la imagen. La radiación emitida, tanto por la luz natural como por la artificial, introducen distorsión con respecto a la propia del objeto como fuente térmica.

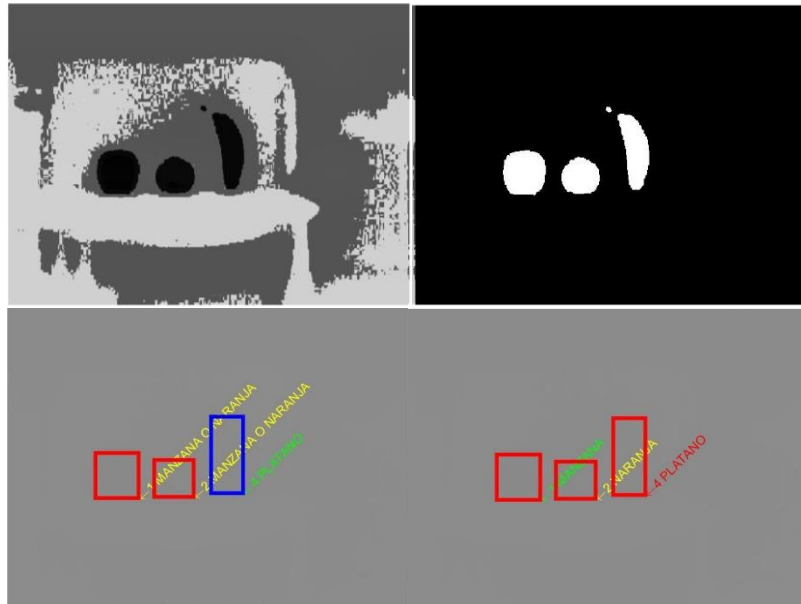


Figura 66. Imagen resultante tras la expansión del histograma (arriba izquierda), binarización (arriba derecha), filtrado simple (abajo izquierda) y filtrado complejo de las regiones (abajo derecha) de la imagen térmica de la escena iluminada con luz natural frontal.

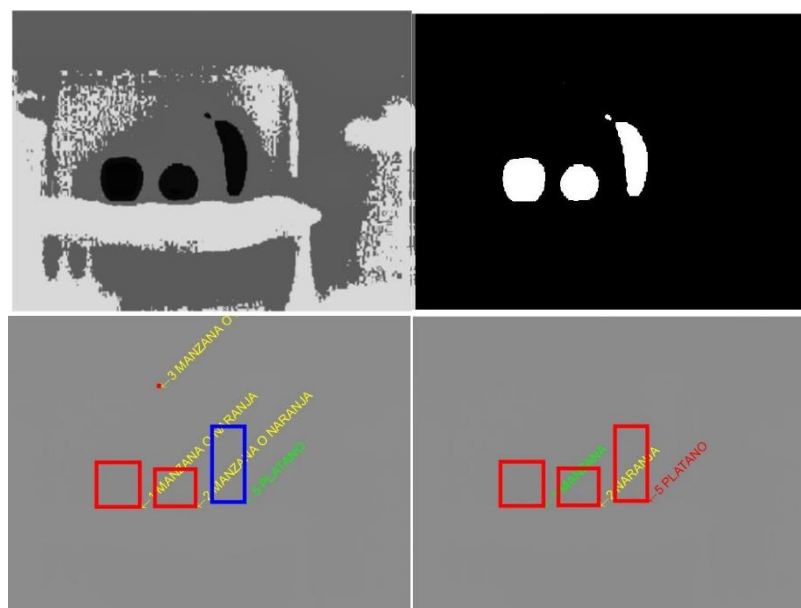


Figura 67. Imagen resultante tras la expansión del histograma (arriba izquierda), binarización (arriba derecha), filtrado simple (abajo izquierda) y filtrado complejo de las regiones (abajo derecha) de la imagen térmica de la escena iluminada con luz artificial por encima de los objetos.

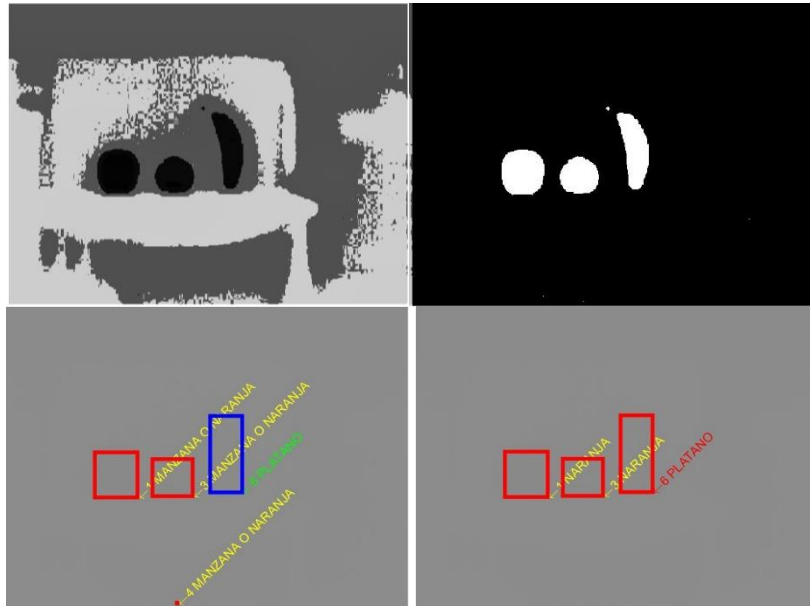


Figura 68. Imagen resultante tras la expansión del histograma (arriba izquierda), binarización (arriba derecha), filtrado simple (abajo izquierda) y filtrado complejo de las regiones (abajo derecha) de la imagen térmica de la escena con iluminación natural frontal y artificial superior.

De la misma manera, la cantidad de iluminación recibida por los objetos no es determinante para la detección, ya que ésta no aporta energía suficiente como para calentar los objetos y alterar su emisividad térmica. En este sentido, las imágenes sin fuente alguna de iluminación mostrarán resultados igualmente efectivos que en los casos con iluminación adicional, tal y como se observa en la Figura 69, cuya captura se realizó completamente a oscuras.

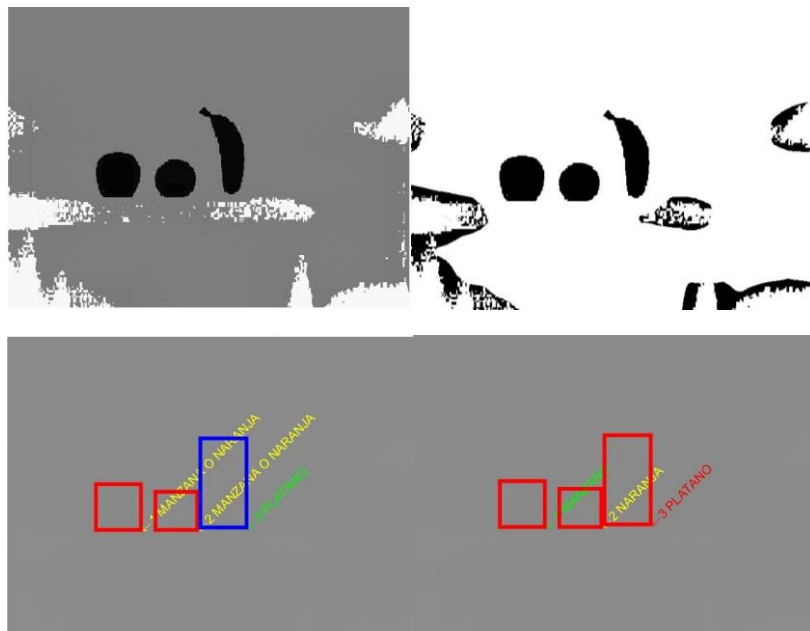


Figura 69. Imagen resultante tras la expansión del histograma (arriba izquierda), binarización (arriba derecha), filtrado simple (abajo izquierda) y filtrado complejo de las regiones (abajo derecha) de la imagen térmica de la escena sin iluminación en el rango visible.

4.4.- Resultados de la combinación de las imágenes infrarroja y visible

El procesado que se ha seguido en este caso se ha vertebrado sobre el que se utilizó para las imágenes térmicas o infrarrojas. Sobre este esquema de detección se ha medido el color de las mismas regiones yuxtapuestas sobre la correspondiente imagen a color.

Las dos cámaras se han alineado en el eje horizontal, de forma que en la etapa anterior al procesado de ambas imágenes se garantiza la correspondencia de coordenadas entre ellas, siendo necesario realizar un recorte de la imagen a color de manera que exista una correspondencia directa de dicha imagen con la imagen térmica. El método de combinación consiste en realizar el procesado anteriormente descrito a la imagen térmica, identificando las regiones filtradas y, posteriormente, medir el color en la correspondiente imagen del espectro visible, contabilizando los píxeles de la región equivalente en ella.

En las Figuras 70 a 73 se muestran las regiones detectadas con el color medido. Para la medición del color se ha tomado el promedio de los valores H y V de la representación HSV de los píxeles contenidos en un cuadrado de 20 píxeles de lado alrededor del centroide de cada región.

En la Tabla 7 se muestra el valor asociado a cada uno de los colores.

Tabla 7. Correspondencia de literales de color con los valores de las componentes de la representación de color HSV.

Texto de color	Valores porcentuales
Rojo	HSV(0°,100% ,50%)
Amarillo	HSV(60°,100% ,50%)
Verde	HSV(120°,100% ,50%)
Cian	HSV(180°,100% ,50%)
Azul	HSV(240°,100% ,50%)
Magenta	HSV(300°,100% ,50%)

En las Figuras 70 a 73 se muestran las etiquetas de los colores detectados (ver Apéndice 3). Aunque el método elegido para la medición de color se basa en la componente H (Capítulo 3, sección 3), porque se considera que este método es más independiente de la intensidad lumínica que reciben los objetos en la escena, se puede observar cómo en la imagen los tonos detectados para los objetos (mayormente amarillos según la percepción visual humana) oscilan en torno al amarillo y al verde claro o color lima.

Cuando la iluminación es frontal (Figura 71) a los objetos se aprecian mejor los tonos de los objetos y la detección es más fiel a lo percibido por el ojo humano básicamente porque la cantidad lumínica sí termina teniendo impacto en las medidas. Cuando se observan penumbras en los objetos los tonos tienden al verde claro (Figuras 70 y 72).



Figura 70. Termografía con detección de colores en una imagen con iluminación natural por detrás de los objetos.

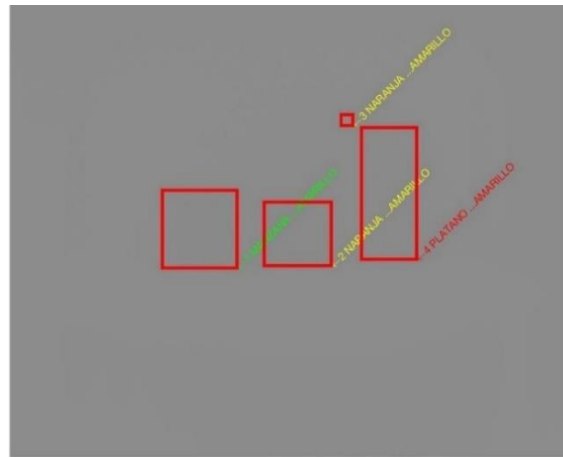


Figura 71. Termografía con detección de colores en una imagen con iluminación natural por delante de los objetos

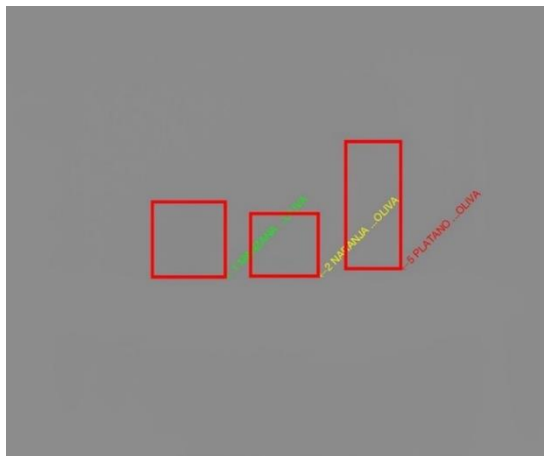


Figura 72. Termografía con detección de colores en una imagen con iluminación artificial por encima de los objetos.

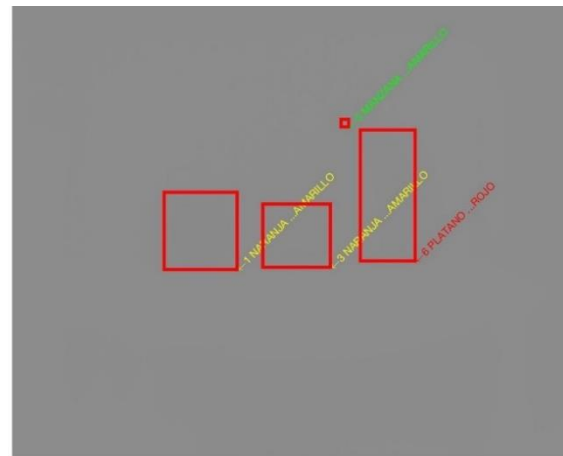


Figura 73. Termografía con detección de colores en una imagen con iluminación natural por delante de los objetos y luz artificial por encima de los mismos.

La técnica utilizada de promediado en torno al centroide presenta problemas de eficacia en el caso del plátano, donde dicho punto puede caer fuera de la silueta del objeto. También es importante resaltar que la tonalidad de la iluminación determina fuertemente los colores detectados, así como su intensidad.

No obstante, este método es complementario a lo estudiado en el apartado de las imágenes térmicas y, por tanto, es útil al añadir más información para el filtrado. Se debe tener en cuenta que la tonalidad amarilla se ha considerado en un rango amplio que cubre también los verdes (color claro o verde lima) y hasta los rojos por el otro extremo.

En las imágenes estudiadas no había grandes componentes azules, pero este sistema permite determinar cuáles de ellas no son regiones válidas.

4.5.- Análisis de eficiencia de los métodos utilizados

A continuación, se presentan diversas estadísticas resultantes de la aplicación de los métodos al conjunto de 75 imágenes utilizadas. También se muestra el número de detecciones en las pruebas realizadas con los diferentes tipos de imagen.

La eficacia del filtrado se calcula como la ratio entre las detecciones correctas (es una fruta y está bien identificada) y las detecciones del último filtrado. Este indicador informa sobre si los objetos encontrados son frutas, pero no evalúa cuántas frutas se han encontrado. Un segundo indicador mide si se han encontrado las tres frutas buscadas en las imágenes de este estudio.

Por último, el producto de los dos es lo que se considera eficacia general del método, ya que indica la bondad en la detección de exclusivamente tres frutas y en su identificación correcta. La Figura 74 muestra la evolución de este último indicador en función del método y la iluminación, siendo 0 la mínima eficiencia y 1 la máxima. Como se ha comentado anteriormente, tanto el uso de la representación de color Lab como el uso de imágenes en el rango ultravioleta, no presentan buenos resultados.

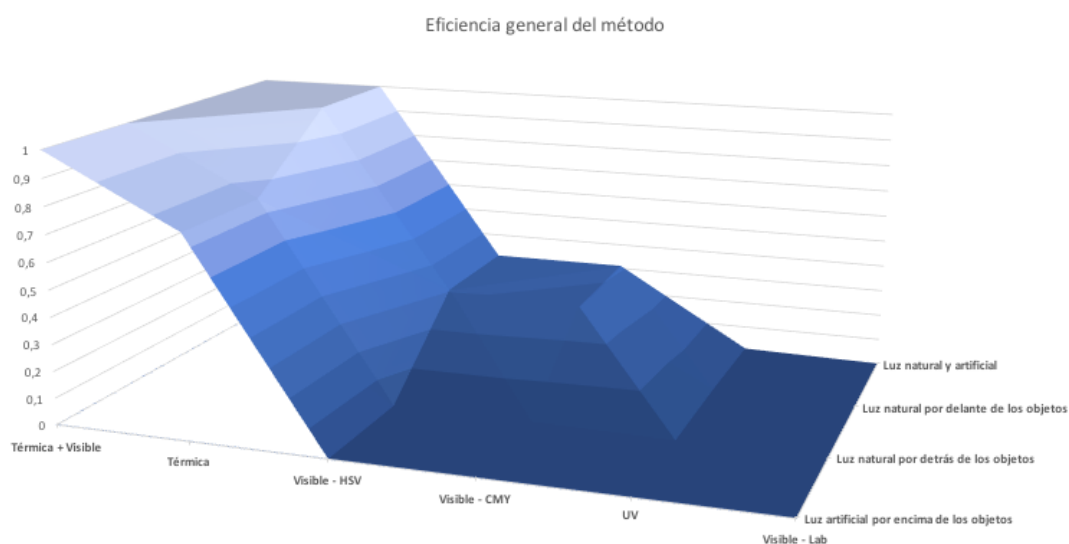


Figura 74. Superficie de evolución de la eficiencia general de la detección en función del tipo de escena y del método utilizado.

En todos los casos la eficacia mejora cuando se utilizan imágenes térmicas. En el caso de los métodos del espectro visible, son idénticos en cuanto a los resultados el método con HSV y el CMY, mejorando cuando la luz natural se aplica por delante de los objetos.

En la Tabla 8 se muestra la eficiencia de cada método en tanto por ciento. Se puede observar cómo el método más eficiente es el resultante de analizar la imagen térmica junto con la imagen a color, añadiendo este último como criterio adicional al filtrado complejo.

Tabla 8. Resultados generales según el tipo de imagen y el tipo de iluminación.

	Térmica + Visible	Térmica	Visible - HSV	Visible - CMY	UV	Visible - Lab
Luz artificial por encima de los objetos	100%	75%	0%	0%	0%	0%
Luz natural por detrás de los objetos	100%	75%	1%	1%	0%	0%
Luz natural por delante de los objetos	100%	100%	31%	29%	1%	0%
Luz natural y artificial	100%	100%	31%	31%	1%	0%

En las Tablas 9 a 14 se pueden comprobar los resultados de cada una de las técnicas aplicadas, basadas en el promedio de los casos y simulaciones estudiados. Un detalle a resaltar es el número de detecciones que se producen en las imágenes UV, donde el alto contraste local de imágenes con tanta resolución hace que el proceso de ecualización del histograma y la posterior binarización no sean tan eficaces, creando multitud de pequeñas islas de intensidad alta que acaban siendo detectadas como regiones.

Tabla 9. Estadísticas de las imágenes estudiadas con el análisis de la componente V de HSV.

Visible - HSV	Número de detecciones iniciales	Número de detecciones tras filtrado simple	Número de detecciones tras filtrado complejo	Falsas detecciones	Falsas identificaciones	Detecciones correctas	Eficacia del filtrado	Eficacia de la detección sobre 3 objetos	Eficacia general
Luz natural por detrás de los objetos	199,5	40,1	10,1	9,5	0,3	0,5	5,0%	16,7%	1%
Luz natural por delante de los objetos	128,6	44,7	6,8	4,2	0,3	2,5	36,8%	83,3%	31%
Luz artificial por encima de los objetos	99,3	34,3	9,9	8,7	1,0	0,1	1,0%	3,3%	0%
Luz natural y artificial	126,4	44,2	6,7	4,2	0,3	2,5	37,3%	83,3%	31%

Tabla 10. Estadísticas de las imágenes estudiadas con el análisis de la componente Y de CMY

Visible - CMY	Número de detecciones iniciales	Número de detecciones tras filtrado simple	Número de detecciones tras filtrado complejo	Falsas detecciones	Falsas identificaciones	Detecciones correctas	Eficacia del filtrado	Eficacia de la detección sobre 3 objetos	Eficacia general
Luz natural por detrás de los objetos	183,5	39,3	9,5	8,7	0,2	0,5	5,3%	16,7%	1%
Luz natural por delante de los objetos	118,2	40,7	7,3	4,6	0,2	2,5	34,2%	83,3%	29%
Luz artificial por encima de los objetos	105,4	33,6	9,2	8,1	1	0,1	1,1%	3,3%	0%
Luz natural y artificial	112,3	41,5	6,7	4	0,2	2,5	37,3%	83,3%	31%

Tabla 11. Estadísticas de las imágenes estudiadas con el análisis de la componente L de Lab

Visible - Lab	Número de detecciones iniciales	Número de detecciones tras filtrado simple	Número de detecciones tras filtrado complejo	Falsas detecciones	Falsas identificaciones	Detecciones correctas	Eficacia del filtrado	Eficacia de la detección sobre 3 objetos	Eficacia general
Luz natural por detrás de los objetos	62	25,8	7,2	6,7	0,1	0,1	1,4%	3,3%	0%
Luz natural por delante de los objetos	39,8	10	3	2,8	0	0,1	3,3%	3,3%	0%
Luz artificial por encima de los objetos	105,6	42,2	6,3	10	1,1	0,1	1,6%	3,3%	0%
Luz natural y artificial	19,8	7,9	2	1,5	0	0	0,0%	0,0%	0%

Tabla 12. Estadísticas de las imágenes UV estudiadas

UV	Número de detecciones iniciales	Número de detecciones tras filtrado simple	Número de detecciones tras filtrado complejo	Falsas detecciones	Falsas identificaciones	Detecciones correctas	Eficacia del filtrado	Eficacia de la detección sobre 3 objetos	Eficacia general
Luz natural por detrás de los objetos	1200,0	723,6	182,3	180,7	1,8	0,1	0,1%	3,3%	0%
Luz natural por delante de los objetos	950,0	990,6	150,7	145,7	2,2	2,1	1,4%	70,0%	1%
Luz artificial por encima de los objetos	450,0	466,3	134,6	129,6	4,5	0,5	0,4%	16,7%	0%
Luz natural y artificial	1050,0	1101,5	167	163,4	2,5	2,1	1,3%	70,0%	1%

Tabla 13. Estadísticas de las imágenes térmicas estudiadas

Térmica	Número de detecciones iniciales	Número de detecciones tras filtrado simple	Número de detecciones tras filtrado complejo	Falsas detecciones	Falsas identificaciones	Detecciones correctas	Eficacia del filtrado	Eficacia de la detección sobre 3 objetos	Eficacia general
Luz natural por detrás de los objetos	20,3	10,0	4,0	1,0	0,0	3,0	75,0%	100,0%	75%
Luz natural por delante de los objetos	10,5	6,0	3,0	0,0	0,0	3,0	100,0%	100,0%	100%
Luz artificial por encima de los objetos	11,6	7,0	4,0	1,0	0,0	3,0	75,0%	100,0%	75%
Luz natural y artificial	9,4	5,0	3,0	0,0	0,0	3,0	100,0%	100,0%	100%

Tabla 14. Estadísticas de las imágenes térmicas en combinación con detección de color.

Térmica + Visible	Número de detecciones iniciales	Número de detecciones tras filtrado simple	Número de detecciones tras filtrado complejo	Falsas detecciones	Falsas identificaciones	Detecciones correctas	Eficacia del filtrado	Eficacia de la detección sobre 3 objetos	Eficacia general
Luz natural por detrás de los objetos	19,8	10,0	3,0	0,0	0,0	3,0	100,0%	100,0%	100%
Luz natural por delante de los objetos	10,3	6,0	3,0	0,0	0,0	3,0	100,0%	100,0%	100%
Luz artificial por encima de los objetos	11,0	7,0	3,0	0,0	0,0	3,0	100,0%	100,0%	100%
Luz natural y artificial	8,8	5,0	3,0	0,0	0,0	3,0	100,0%	100,0%	100%

Por último, es importante estimar el coste computacional del procesamiento propuesto. Las operaciones de ecualización y binarización de las imágenes (incluyendo la separación en componentes de color cuando sea necesario) no superan los 3 segundos de cálculo cada una, considerando las imágenes con mayor resolución (en este estudio las imágenes a color). No obstante, la detección de regiones eleva notablemente el tiempo de computación, pudiendo llegar a realizar los dos filtrados en una media de 0.5 segundos por región encontrada. Dado que el mejor de los casos es con la combinación de las imágenes térmicas y visibles, para este método el tiempo promedio de ejecución es de 14 segundos (suma de los tiempos de carga de imagen, separación de componentes, ecualización del histograma, binarización y detección de regiones de la imagen térmica con medición de color en su correspondiente imagen a color).

En la Figura 75 se muestra la gráfica de los valores máximos del tiempo de cálculo de las operaciones en función del tipo de imagen estudiada (que determina el número de detecciones y por tanto el tiempo de computación), lo que indica el tiempo en el peor de los casos.



Figura 75. Tiempo máximo de cálculo por operación en segundos. La detección de regiones está especificada, además, por tipo de imagen.

5.- Conclusiones y futuras líneas de investigación

En el desarrollo de este Proyecto Fin de Máster se ha trabajado con tres tipos de cámaras y las correspondientes imágenes obtenidas en diferentes condiciones de iluminación.

Respecto a las imágenes en el rango visible se pueden extraer varias conclusiones.

La iluminación utilizada en una aplicación para detectar los objetos en una imagen deberá saturar lumínicamente la superficie de los objetos, para evitar que su geometría tridimensional provoque distorsiones en la detección al crear una gradación de grises. Además, la mejor dirección de iluminación es la frontal, más o menos perpendicular al plano del sensor de la cámara.

Por otro lado, la descomposición de la imagen utilizando diferentes representaciones del espacio de color ha demostrado que tanto la representación HSV (utilizando su componente V) como la CMY (utilizando la Y) son prácticamente equivalentes. No obstante, esta equivalencia se debe a que los objetos buscados tenían una componente amarilla predominante. En otros casos, con otras frutas u otras fuentes de luz seguramente será más conveniente ceñirse a la componente V de HSV.

También se puede destacar cómo la utilización de la componente Cb de YCbCr para la detección de contornos y su posterior superposición con la componente V de HSV resulta en una mejora notable de la detección de las regiones al introducir un borde que separa los objetos de los elementos de su entorno. Sin embargo, la utilización de esta técnica ha presentado ciertas deficiencias o inconvenientes, principalmente provocados por la iluminación indirecta o insuficiente de los objetos, ya que los diferentes gradientes de intensidad creados por las sombras y zonas de baja intensidad provocan detecciones de falsos contornos que desvirtúan la detección de los objetos.

También se ha podido demostrar que la representación LAB no es eficaz.

Sobre la imagen UV se puede concluir que produce malos resultados en general, ya que los reflejos dan lugar a informaciones erróneas en el procesado. Además, no aporta información de color, con lo que no se puede mejorar la detección con otras componentes de color como se ha hecho en las imágenes del espectro visible.

Por último, la imagen térmica o infrarroja es con diferencia la que mejores resultados proporciona en cuanto a la identificación de los objetos. La uniformidad del mapa de intensidades obtenido para cada objeto permite una detección correcta de los objetos en su volumen y una buena identificación mediante sus características geométricas. Como objeción al uso de esta técnica, se puede encontrar la ausencia de información de color. Este hecho puede conducir a falsas detecciones de objetos de similar geometría.

La posibilidad de combinar la imagen térmica con la imagen de espectro visible resuelve este problema al aportar información sobre el color de los objetos detectados, permitiendo añadir un filtro adicional a los aplicados inicialmente.

En el análisis realizado no se han realizado medidas de los objetos ya que no se dispone de ninguna medida de profundidad o distancia al sensor. En caso de poder obtener la distancia a los objetos mediante algún sistema de medida (tipo láser o mediante patrones

de dimensión conocida), se podrían obtener por cálculos geométricos los tamaños de los objetos para aplicar un filtro adicional sobre esa propiedad.

Como línea de investigación futura se plantea la construcción de un sistema formado por una cámara multiespectral y un sistema de iluminación frontal con luz blanca visible perpendicular al plano del sensor de la cámara.

Respecto a los objetos a detectar hay que comentar dos aspectos. El primero se refiere a la diversidad de objetos, por ejemplo, vasos, piñas o juguetes. Evidentemente las propiedades geométricas variarán, pero también las componentes de color. Aunque la utilización de la representación HSV es la más flexible ante cambios de color según este estudio, la realidad es que la utilización de la componente Cb de YCbCr puede no ser la más adecuada si el color del nuevo objeto no es de componente predominantemente amarilla.

El segundo aspecto a valorar es la propia diversidad dentro de objetos de la misma naturaleza, por ejemplo, manzanas de distinto tamaño y forma similar. En este caso sería aconsejable crear un sistema inteligente que sea capaz de aprender de la experiencia reconociendo patrones. El uso de técnicas de reconocimiento de patrones mediante inteligencia artificial, redes neuronales, o patrones estadísticos sería aplicable. De esta forma la determinación de las regiones por medio de variables geométricas sería un ejercicio de estimaciones estadísticas con unas probabilidades de certidumbre calculables.

Recientemente se están implementando sistemas comerciales que combinan escaneado láser con captación de imágenes con cámara. La unión de ambos permitiría añadir filtros de geometría tridimensional a los objetos siendo, de esta forma, muy sencillo eliminar las falsas figuras creadas por sombras planas o gradientes de luz.

Para finalizar, se propone realizar una implementación de este sistema en lenguaje C, con la finalidad de reducir el coste computacional que requiere un lenguaje de programación prototipado como Matlab y poder implementarlo en sistemas hardware que permitan su uso en tiempo real.

6.- Bibliografía

- [1]. Guerrero, J.M., Ruz, J.J., Pajares, G. (2017). Crop rows and weeds detection in maize fields applying a computer vision system based on geometry. *Computers and Electronics in Agriculture*. Volume 142, Part A. Pages 461-472. ISSN 0168-1699. <https://doi.org/10.1016/j.compag.2017.09.028>.
- [2]. Gundupalli, S.P., Hait, S., Thakur, A. (2017). A review on automated sorting of source-separated municipal solid waste for recycling. *Waste Management*, Volume 60, Pages 56–74. ISSN 0956-053X. <https://doi.org/10.1016/j.wasman.2017.09.019>.
- [3]. Suardíaz Muro J. , Navarro Lorente P.J., Alcover Garau P., Fernández Andrés C., Al-Hadithi B.M. (2014) Solucion basada en vision artificial para la inspeccion automatizada de gajos de mandarina.
- [4]. Montoya Holguin C., Cortés Osorio J.A., Chaves Osorio J.A. (2014) Sistema automático de reconocimiento de frutas basado en visión por computador. *Ingeniare. Revista Chilena de Ingeniería*, vol. 22 No 4, pp. 504-516.
- [5]. Giró S. (2017) Vision artificial en la industria alimentaria. Grupo Infaimon.
- [6]. Vollmer M. y Möllmann K.P. (2013) Termografía infrarroja. Editorial Universitat Politecnica de Valencia.
- [7]. Arivazhagan S., Newlin Shebiah R., Selva Nidhyanandhan S., Ganesan L. (2009) Fruit Recognition using Color and Texture Features. *Journal of Emerging Trends in Computing and Information Sciences*.
- [8]. Baohua Zhanga, Wenqian Huang, Jiangbo Li, Chunjiang Zhao, Shuxiang Fan, Jitao Wu, Chengliang Liu. (2004) Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. *Food Research International* Volume 62, August 2014, Pages 326-343.
- [9]. INFAIMON Control y optimización de procesos. (2014) Determinación de la calidad y la madurez de la fruta mediante imagen hiperespectral.
- [10]. Kay G. and de Jager G. (1992) A versatile colour system capable of fruit sorting and accurate object classification. *IEEE Conference Publications*, pp. 145-148.
- [11]. Moreda G. P., Muñoz M. A., Barreiro. (2011) Inspeccion de frutas y hortalizas mediante videocamara estado actual de la tecnica. P. LPF-TAGRALIA. Departamento Ingeniería Rural UPM.
- [12]. Rocha A., Hauagge D. C., Wainer J. and Goldenstein S. (2010) Automatic fruit and vegetable classification from images. *ScienceDirect Computers and Electronics in Agriculture*, pp. 96-104.
- [13]. Woo Chaw Seng, Seyed Hadi Mirisae. (2009) A New Method for Fruits Recognition System. *Electrical Engineering and Informatics*.

- [14]. Xu, J., Riccioli, C., Sun, D. (2017). Comparison of hyperspectral imaging and computer vision for automatic differentiation of organically and conventionally farmed salmon. *Journal of Food Engineering*, Volume 196, Pages 170-182.
- [15]. Romero-Acero A. (2014) Sistema de clasificación por visión artificial de mangos tipo Tommy. Grupo IAE- Facultad de Minas Universidad Nacional de Colombia.
- [16]. García-Lamont F., Cervantes J., Ruiz S., and López-Chau A. (2015) Color Characterization Comparison for Machine Vision-Based Fruit Recognition. D.- S. Huang et al. (Eds.): ICIC 2015, Part I, LNCS 9225, pp. 258–270
- [17]. Hyperspectral imaging en Wikipedia
- [18]. Gowen A.A., O'Donnell C.P., Cullen P.J., Downey G. Frias J.M. (2007) Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, Vol 18, Issue 12, pp 590-598
- [19]. Brown M., Godman M. (2011) Simple Pattern Recognition via Image Moments. Electrical Engineering Department New Mexico Institute of Mining and Technology.
- [20]. Theodoridis S. y Koutroumbas K. (2006) *Pattern Recognition (3ed.)*. Associated Press.
- [21]. Theodoridis S. y Koutroumbas K. (2010) *Introduction to Pattern Recognition. A Matlab Approach*. Associated Press.
- [22]. Pajares G. y De la Cruz García J.M. (2007) *Visión por computador (2ed.)*. Ra-Ma.
- [23]. CIE 15.3:2004 COLORIMETRY – Technical Report
- [24]. Mendoza, F., Dejmek, P., Aguilera, J.M. (2006). Calibrated color measurements of agricultural foods using image analysis. *Postharvest Biology and Technology*. 41:285-295.
- [25]. G. García-Mateos, S. Jaén Terrones, D. Escarabajal-Henarejos, J.M. Molina-Martínez, J.L. Hernández-Hernández (2014). Segmentación automática de imágenes de cultivos: estudio comparativo de modelos de color. *Symposium Nacional de Ingeniería Hortícola*
- [26]. Gonzalez R., Wods R.E. y Eddins S.L. (2010) *Digital Image Processing Using Matlab (2ed.)*. Mc Graw Hill
- [27]. Otsu N. (1979) A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics* (Volume: 9, Issue: 1)
- [28]. Chabrier S., Rosenberger C., Laurent H., Emile B. and Marche P. (2003) EVALUATING THE SEGMENTATION RESULT OF A GRAY-LEVEL IMAGE. *Laboratoire Vision et Robotique, UPRES EA 2078 ENSI de Bourges / Universite d'Orleans*.

- [29]. Kavitha A.R., Chellamuthu C. (2010) Implementation of Gray-level Clustering Algorithm for Image Segmentation. *Procedia Computer Science* 2 (2010) 314–320.
- [30]. Sourav Samanta, Nilanjan Dey, Poulami Das, Suvojit Acharjee, Sheli Sinha Chaudhuri. (2012) Multilevel Threshold Based Gray Scale Image Segmentation using Cuckoo Search. In proceedings of ICECIT – 2012, Elsevier 27-34.
- [31]. Randolph Maire M. (2009) Contour Detection and Image Segmentation. Electrical Engineering and Computer Sciences University of California at Berkeley Technical Report No. UCB/EECS-2009-129
<http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-129.html>.
- [32]. Wei-Ying Ma and B. S. Manjunath. (1995) EdgeFlow: A Technique for Boundary Detection and Image Segmentation. Department of Electrical and Computer Engineering University of California.
- [33]. Hu Ming-Kuei. (1962) Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory* 179-187.
- [34]. Zhihu Huang, Jinsong Leng. (2011) Analysis of Hu's Moment Invariants on Image Scaling and Rotation. ECU Publications Pre.
- [35]. LS. Davis, A survey of edge detection techniques, (1975) *Computer Graphics and Image Processing*, 4,3, 248-260
- [36]. Image Processing Toolbox. User's Guide. Matlab R2017a. Mathworks.

Apéndice 1- Código MATLAB del análisis y filtrado de regiones

CARGA DE IMÁGENES Y SEPARACIÓN DE COMPONENTES

```
function [imagen] = cargafigura (nombre,tipo,operacion,invertir)
try
    % mac
    camino = '/Users/rgribao/Documents/Mis Documentos/OneDrive/Privada/UNED/2016_17_18/TFM/imagenes/';
    i0000 = imread(strcat(camino,nombre,'.',tipo),tipo);
catch
    % windows
    camino = 'C:\Users\ricar\OneDrive\Privada\UNED\2016_17_18\TFM\imagenes\';
    i0000 = imread(strcat(camino,nombre,'.',tipo),tipo);
end

informacion = imfinfo(strcat(camino,nombre,'.',tipo));
tipo_leido = informacion.BitDepth;
if max(size(size(i0000))) > 2
    tipo_leido = tipo_leido / 3;
end
max_bytes = 1 / ((2 ^ tipo_leido) - 1);

if strcmp(operacion,'gris')
    i0 = max_bytes * double(rgb2gray(i0000));

elseif strcmp(operacion,'hsv')
    i000 = rgb2hsv(i0000);
    i0 = double(i000(:,:,3));

elseif strcmp(operacion,'cmy')
    i000 = rgb2hsv(i0000);
    i00 = double(i000(:,:,3));
    i0 = double(ones(size(i00))) - i00;

elseif strcmp(operacion,'ycbcr')
    YCBCR = rgb2ycbcr(i0000);
    i000 = YCBCR(:,:,2);
    i0 = max_bytes * double(i000);

elseif strcmp(operacion,'lab')
    LAB = double(rgb2lab(i0000)); % convert to L*a*b*
    max_luminosity = 100.0;
    i0 = LAB(:,:,1)/max_luminosity;

end
%[nn,mm] = size(i0);

% para imagenes donde lo que buscamos está iluminado en vez de estar oscurecido,
% invertimos, ya que nuestros algoritmos invierten los elementos "oscuros"
if invertir == 0
    imagen = i0;
else
    imagen = double(ones(size(i0))) - i0;
end

end
```

ECUALIZACIÓN DE IMAGENES A COLOR Y UV

```
if saturado == 0
    i1 = imadjust(imgIN, [0.0 1.0],[0.5 1.0],1.0);
else
    i1 = imadjust(imgIN, [0.2 0.8],[1],0.3);
end

[imag,idir] = imggradient(imgYCbCr,'Roberts');
se = strel('disk',1);
imag = imdilate(imag,se);
imag = imbinarize(imag);
i1 = miMascara(i1,saturado * imag);

h = figure; imshow(i1);title(strcat(texto,' imadjust')); print(h,'-djpeg',strcat(texto,' imadjust'));clear h;
h = figure; imhist(i1);title(strcat(texto,' imadjust hist')); print(h,'-djpeg',strcat(texto,' imadjust hist'));clear h;

isalida = i1;
```

ECUALIZACIÓN DE IMAGENES TÉRMICAS

```
i1 = adapthisteq(imgIN,'NumTiles',[8 8],'ClipLimit',0.01);
h = figure; imshow(i1);title(strcat(texto,' adapthisteq 8 8 001')); print(h,'-djpeg',strcat(texto,' adapthisteq 8 8 001'));clear h;
h = figure; imhist(i1);title(strcat(texto,' adapthisteq 8 8 001 hist')); print(h,'-djpeg',strcat(texto,' adapthisteq 8 8 001 hist'));clear h;

isalida = i1;
```

BINARIZACIÓN DE IMÁGENES A COLOR

```
% binarizacion con imbinarize con estadísticas locales-adaptive y ajuste
% sensibilidad
i2 = imbinarize(imgIN,'adaptive','ForegroundPolarity','dark','Sensitivity',0.6);
```

```

h = figure; imshowpair(imgIN,i2,'montage');title(strcat(texto,' imbinarize local adaptive sensitivity')); print(h,'-
djpeg',strcat(texto,' imbinarize local adaptive sensitivity')); clear h;

% operacion morfologica
if morfo == 1
    se = strel('disk',1);
    i2 = imerode(i2,se);
elseif morfo == 2
    se = strel('disk',1);
    i2 = imdilate(i2,se);
end

isalida = i2;

```

BINARIZACIÓN DE IMÁGENES TÉRMICAS

```

% binarizacion por umbrales
lvl = multithresh(imgIN,3);
val = [lvl max(imgIN(:))];
i21 = imquantize(imgIN,lvl,val);
i2 = (i21(:,>0) & (i21(:,<((10.0/255.0) + min(lvl)))));
h = figure; imshowpair(imgIN,i2,'montage');title(strcat(texto,' imquantize multithresh 4niv')); print(h,'-
djpeg',strcat(texto,' imquantize multithresh 4niv')); clear h;

% operacion morfologica
if morfo == 1
    se = strel('disk',5);
    i2 = imerode(i2,se);
elseif morfo == 2
    se = strel('disk',5);
    i2 = imdilate(i2,se);
end

isalida = i2;

```

DETECCION DE REGIONES

```

function deteccion_regiones(img_in,img_orig,texto)

cc = bwconncomp(img_in);
datos = regionprops(cc,'all');
[n m]=size(datos);
clear cc;

g = figure;
imshow(img_orig); title(strcat(texto,' deteccion de regiones')); hold on;
h = 0.0;
hh = 0.0;

mHuMan = [0.796300 4.158000 4.951000 8.458000 15.160000 10.940000 16.350000];
normMan = 27.033404;

mHuNar = [0.797600 4.343000 5.683000 8.504000 15.810000 10.920000 15.700000];
normNar = 27.199353;

mHuPla = [0.444600 1.200000 1.780000 2.609000 4.833000 3.213000 5.250000];
normPla = 8.535614;

fileID = fopen(strcat(texto,' analisis.txt'),'w');
fprintf(fileID,strcat('\n',texto,'\n'));
fprintf(fileID,'num:redondez;compatibilidad;excentricidad;solidez;corr2Man;corr2Nar;corr2Pla;corrMan;corrNar;corrPla;m
Hu1;mHu2;mHu3;mHu4;mHu5;mHu6;mHu7\n');
fprintf(fileID,'-----\n');

for i=1:n
    xc = datos(i).Centroid(1);
    yc = datos(i).Centroid(2);
    [m,p] = size(datos(i).ConvexHull);
    clear r; r = zeros(m,1);
    for j=1:m
        x = datos(i).ConvexHull(j,1);
        y = datos(i).ConvexHull(j,2);
        r(j) = sqrt((x-xc)^2+(y-yc)^2);
    end
    R = mean(r);
    redondez = std(r)/R; compatibilidad = ((datos(i).Perimeter)^2)/datos(i).Area;
    excentricidad = datos(i).MajorAxisLength/datos(i).MinorAxisLength;
    solidez = datos(i).Solidity;
    mHu = Momentos_Hu(datos(i).FilledImage);

    corrMan = (mHu(1)-mHuMan(1))^2 + (mHu(2)-mHuMan(2))^2 + (mHu(3)-mHuMan(3))^2 + (mHu(4)-mHuMan(4))^2 + (mHu(5)-
mHuMan(5))^2 + (mHu(6)-mHuMan(6))^2 + (mHu(7)-mHuMan(7))^2;
    corrMan = sqrt(corrMan) / normMan;

    corrNar = (mHu(1)-mHuNar(1))^2 + (mHu(2)-mHuNar(2))^2 + (mHu(3)-mHuNar(3))^2 + (mHu(4)-mHuNar(4))^2 + (mHu(5)-
mHuNar(5))^2 + (mHu(6)-mHuNar(6))^2 + (mHu(7)-mHuNar(7))^2;
    corrNar = sqrt(corrNar) / normNar;

    corrPla = (mHu(1)-mHuPla(1))^2 + (mHu(2)-mHuPla(2))^2 + (mHu(3)-mHuPla(3))^2 + (mHu(4)-mHuPla(4))^2 + (mHu(5)-
mHuPla(5))^2 + (mHu(6)-mHuPla(6))^2 + (mHu(7)-mHuPla(7))^2;
    corrPla = sqrt(corrPla) / normPla;

```

```

fprintf(fileID, '%3d;%4.3f;%4.3f;%4.3f;%4.3f;%4.3f;%4.3f;%4.3f;%4.3f;%4.3e;%4.3e;%4.3e;%4.3e;%4.3e;%4.3e;%4.3e;%4.3e\n', i, redondez, compatibilidad, excentricidad, solidez, corr2(mHu, mHuMan), corr2(mHu, mHuNar), corr2(mHu, mHuPla), corrMan, corrNar, corrPla, mHu);

% la variaci3n en radios es menor de un 10%
if redondez <= 0.13
    text(datos(i).BoundingBox(1)+datos(i).BoundingBox(3), datos(i).BoundingBox(2)+datos(i).BoundingBox(4), strcat('\leftarrow', num2str(i), ' PLATANO O NARANJA'), 'Color', 'yellow', 'Rotation', 45);
    rectangle('Position', datos(i).BoundingBox, 'EdgeColor', 'r', 'LineWidth', 3);

elseif excentricidad >= 2.5
    text(datos(i).BoundingBox(1)+datos(i).BoundingBox(3), datos(i).BoundingBox(2)+datos(i).BoundingBox(4), strcat('\leftarrow', num2str(i), ' PLATANO'), 'Color', 'green', 'Rotation', 45);
    rectangle('Position', datos(i).BoundingBox, 'EdgeColor', 'b', 'LineWidth', 3);

else
    text(datos(i).BoundingBox(1)+datos(i).BoundingBox(3), datos(i).BoundingBox(2)+datos(i).BoundingBox(4), strcat('\leftarrow', num2str(i), ' NR'), 'Color', 'blue', 'Rotation', 45);
    rectangle('Position', datos(i).BoundingBox, 'EdgeColor', 'r', 'LineWidth', 3);
end
end
hold off;

print(g, '-djpeg', strcat(texto, '_imagfinal1'));
clear g;
fclose(fileID);

g = figure;
imshow(img_orig); title(strcat(texto, ' deteccion de regiones. Filtrado complejo.)); hold on;

for i=1:n
    xc = datos(i).Centroid(1);
    yc = datos(i).Centroid(2);
    [m,p] = size(datos(i).ConvexHull);
    clear r; r = zeros(m,1);
    for j=1:m
        x = datos(i).ConvexHull(j,1);
        y = datos(i).ConvexHull(j,2);
        r(j) = sqrt((x-xc)^2+(y-yc)^2);
    end
    R = mean(r);
    redondez = std(r)/R; %compatibilidad = ((datos(i).Perimeter)^2)/datos(i).Area;
    excentricidad = max([datos(i).BoundingBox(3)/datos(i).BoundingBox(4) ;
    datos(i).BoundingBox(4)/datos(i).BoundingBox(3)]);
    solidez = datos(i).Solidity;
    mHu = Momentos_Hu(datos(i).FilledImage);

    corrMan = (mHu(1)-mHuMan(1))^2 + (mHu(2)-mHuMan(2))^2 + (mHu(3)-mHuMan(3))^2 + (mHu(4)-mHuMan(4))^2 + (mHu(5)-
    mHuMan(5))^2 + (mHu(6)-mHuMan(6))^2 + (mHu(7)-mHuMan(7))^2;
    corrMan = sqrt(corrMan) / normMan;

    corrNar = (mHu(1)-mHuNar(1))^2 + (mHu(2)-mHuNar(2))^2 + (mHu(3)-mHuNar(3))^2 + (mHu(4)-mHuNar(4))^2 + (mHu(5)-
    mHuNar(5))^2 + (mHu(6)-mHuNar(6))^2 + (mHu(7)-mHuNar(7))^2;
    corrNar = sqrt(corrNar) / normNar;

    corrPla = (mHu(1)-mHuPla(1))^2 + (mHu(2)-mHuPla(2))^2 + (mHu(3)-mHuPla(3))^2 + (mHu(4)-mHuPla(4))^2 + (mHu(5)-
    mHuPla(5))^2 + (mHu(6)-mHuPla(6))^2 + (mHu(7)-mHuPla(7))^2;
    corrPla = sqrt(corrPla) / normPla;

    minCorr = min([corrMan corrNar corrPla]);

    critRedondez = (redondez > 0.0) * (redondez < 0.13);
    critExcentricidad = (excentricidad > 1.5) * (excentricidad < 3.2) * (solidez > 0.72) * (solidez < 0.92);
    critCorrHu = (minCorr < 0.52);

    criterio = (critRedondez + critExcentricidad) * critCorrHu;

    if criterio > 0
        if critExcentricidad > 0
            untexto = ' PLATANO';
            color = 'red';
        else
            if corrNar > corrMan
                untexto = ' NARANJA';
                color = 'yellow';
            else
                untexto = ' MANZANA';
                color = 'green';
            end
        end
        end

        text(datos(i).BoundingBox(1)+datos(i).BoundingBox(3), datos(i).BoundingBox(2)+datos(i).BoundingBox(4), strcat('\leftarrow', num2str(i), untexto), 'Color', color, 'Rotation', 45);
        rectangle('Position', datos(i).BoundingBox, 'EdgeColor', 'r', 'LineWidth', 3);
    end
end
hold off;

print(g, '-djpeg', strcat(texto, '_imagfinal1_complex'));
clear g;

end

```


Apéndice 2- Listado de funciones utilizadas

Las funciones principales utilizadas de Matlab Image Processing Toolbox son:

adapthisteq

Realiza una ecualización adaptativa del histograma de una imagen en niveles de gris. Lo hace de forma local dividiendo la imagen original en una cuadrícula cuyo tamaño se especifica a la entrada (por defecto divide la imagen en 8x8 cuadros).

bwconncomp

Devuelve los componentes conectados encontrados en una imagen binaria (variable de entrada a la función). Por defecto utiliza una conectividad 8 de vecinos en 2D.

imadjust

Redistribuye los niveles de gris de una imagen a nuevos valores fijados en el vector de entrada. Normalmente se utiliza en conjunción con la función stretchlim.

imbinarize

Binariza una imagen (normalmente en escala de grises) tomando un valor umbral. Puede calcular el valor umbral por un método Otsu sobre toda la imagen de entrada, método global, o por un método local, utilizando estadísticas de primer orden en la vecindad de cada píxel. El umbral puede ser especificado como un valor límite de luminancia (utilizado en el método global) o como una matriz (del mismo tamaño que la imagen de entrada), en cuyo caso es interpretado como un umbral local adaptado. Los umbrales son en todo caso valores entre 0 y 1.

imdilate

Función que realiza la transformación morfológica de dilatación en una imagen.

imerode

Función que realiza la transformación morfológica de erosión en una imagen.

imgradient

Detecta la magnitud y dirección del gradiente en una imagen. Nos permite detectar contornos al tomar estos valores máximos del gradiente.

imhist

Muestra un objeto gráfico con el histograma de una imagen en escala de grises.

imquantize

En combinación con la función anterior redistribuye los niveles de gris a los especificados como parámetros de entrada. Se dice que cuantiza los niveles en tanto que traduce una escala continua de niveles de intensidad a una serie de niveles discretos.

imshow

Muestra una imagen en pantalla.

imshowpair

Muestra dos imágenes pegadas en pantalla.

imread

Lee una imagen de un archivo del sistema.

multithresh

Devuelve un valor umbral calculado para una imagen de entrada utilizando un método de Otsu.

regionprops

Con los componentes conectados encontrados en la función `bwconncomp` utilizados como variable de entrada, esta función devuelve unas medidas de cada una de las componentes, según se especifique en los parámetros de entrada. En el capítulo 3 se detallan las utilizadas.

rgb2gray

Convierte una imagen color RGB a una imagen de grises representando intensidades. Directamente elimina las componentes de saturación y tonalidad de la representación HSV de color, reteniendo exclusivamente la intensidad o luminancia.

rgb2hsv

Devuelve las componentes del espacio de color HSV a partir de las tres componentes del espacio de color RGB.

rgb2lab

Devuelve las componentes del espacio de color LAB a partir de las tres componentes del espacio de color RGB.

rgb2ycbcr

Devuelve las componentes del espacio de color YCbCr a partir de las tres componentes del espacio de color RGB.

stretchlim

Para una imagen *I* (en nuestro caso imagen de grises) devuelve un vector con los límites inferior y superior que se pueden utilizar para estirar el contraste de la imagen.

strel

Devuelve un elemento estructural para transformaciones morfológicas en imágenes binarias. Normalmente devuelve una matriz bi-dimensional de ceros y unos que indica la relación del pixel de la imagen original a transformar morfológicamente con sus vecinos en la misma imagen.

Apéndice 3- Tabla de colores según representación HSV

Fuente: Wikipedia

HSV [\[edit \]](#)

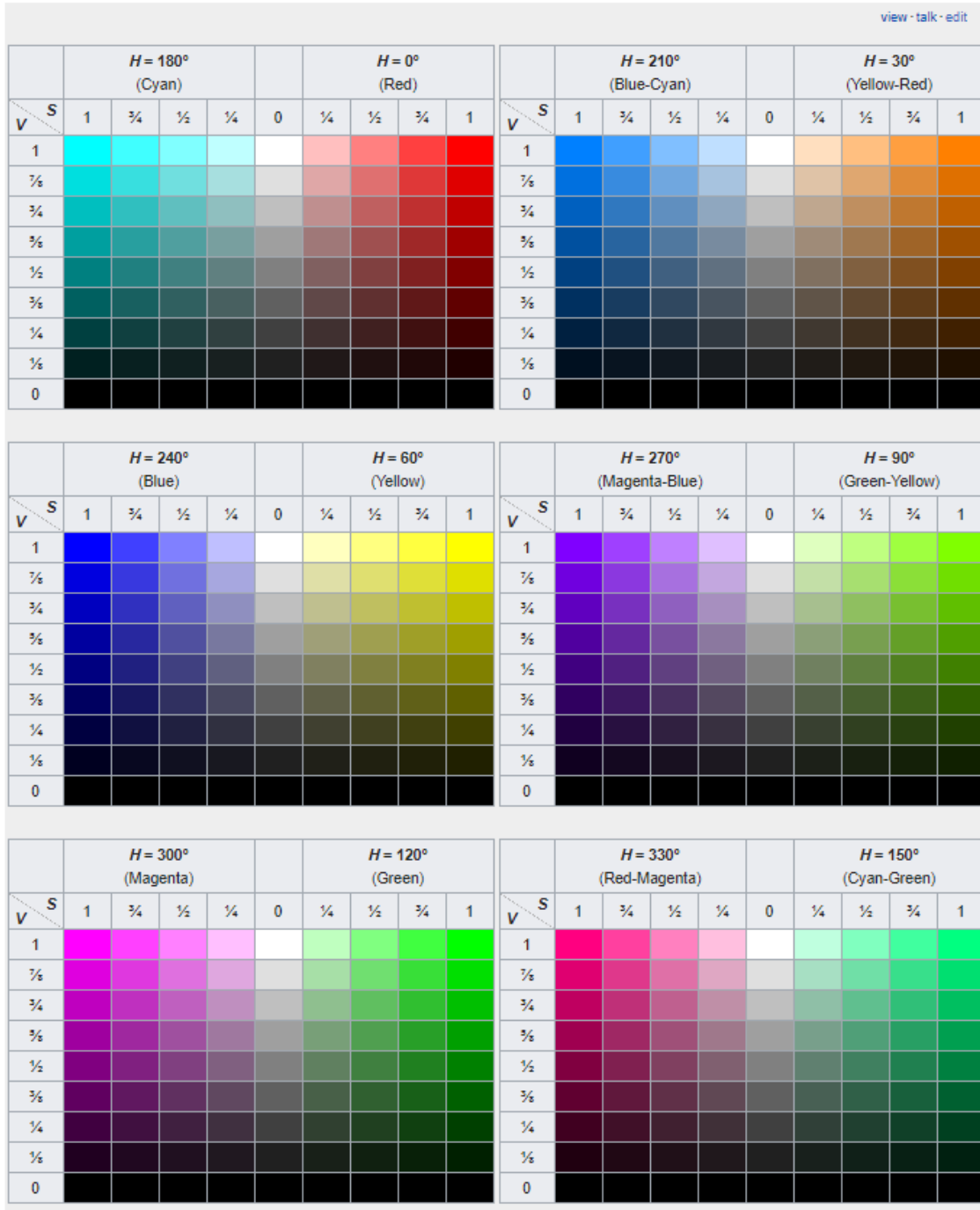


Figura 32. Tabla de colores según los valores HSV.