

**DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA PARA EL RECONOCIMIENTO DE
PATRONES UTILIZANDO VÍDEOCÁMARAS**

Juan Guillermo Aguilar Gaviria

Viviana Andrea Castaño Soto

Proyecto de grado presentado como requisito parcial
para aspirar al título de Ingeniero Electrónico

Director

Andrés Felipe Calvo, Ing. M.Sc.

UNIVERSIDAD TECNOLÓGICA DE PEREIRA
PROGRAMA DE INGENIERÍA ELECTRÓNICA
PEREIRA

2016

Nota de aceptación:

Firma Director de Trabajo de Grado

MSc. ANDRÉS FELIPE CALVO S.

Firma del Director de Programa

Msc. EDWIN ANDRÉS QUINTERO

Pereira, Junio 20 de 2016

DEDICATORIA

A Dios por sostenerme en los momentos difíciles y darme salud y sabiduría para culminar con este sueño.

A mis padres por darme la vida, creer en mi y darme el apoyo necesario en todas las etapas de mi vida.

A mis hijos por quienes me levanto cada día tratando de darles el mejor ejemplo.

JUAN GUILLERMO AGUILAR GAVIRIA

A Dios por darme salud y guiarme en cada paso que doy para poder cumplir con cada de mis metas.

A mi familia y todas aquellas personas que directa o indirectamente me apoyaron en este proceso y gracias a ellos hoy puedo culminar esta etapa de mi vida

A mi hija quien ha sido una gran motivación para levantarme cada día y luchar por mis sueños.

VIVIANA ANDREA CASTAÑO SOTO

AGRADECIMIENTOS

Agradecemos a nuestro Director Andrés Felipe Calvo por su acompañamiento y el gran apoyo que nos ha brindado en este proceso y gracias a el pudimos culminar este gran sueño.

CONTENIDO

	pág.
1. INTRODUCCIÓN	3
2. DEFINICIÓN DEL PROBLEMA	4
3. OBJETIVOS	5
3.1. OBJETIVO GENERAL	5
3.2. OBJETIVOS ESPECÍFICOS	5
4. METODOLOGÍA	6
4.1. CONSTRUCCIÓN DE LA BASE DE DATOS	7
4.2. EXTRACCIÓN DE CARACTERÍSTICAS	9
4.2.1. Pre-Procesamiento de las Imágenes	10
4.3. DESCRIPTORES MODELOS DE COLOR	11
4.3.1. Descriptor RGB	11
4.3.2. Descriptor CMYK	12
4.3.3. Modelo TSL	12
4.4. DESCRIPTORES MODELOS DE TEXTURA	13
4.4.1. Descriptor SOBEL	14
4.4.2. Descriptor CANNY	14
4.5. DESCRIPTORES MODELOS DE FORMA	15
4.5.1. Descriptor HOG	16
4.5.2. Descriptor SURF	17
4.5.3. MSER	18

4.6.	FUSIÓN DE DATOS	19
4.7.	ENTRENAMIENTO Y VALIDACIÓN	19
4.7.1.	Maquinas de soporte vectorial	19
4.7.2.	Máquinas de Soporte Vectorial Multiclase	21
4.8.	VALIDACIÓN DEL CLASIFICADOR UTILIZANDO MONTECARLO	21
5.	ANÁLISIS Y RESULTADOS	22
5.1.	FILTRADO Y SEGMENTACIÓN	22
5.2.	DESCRIPTORES DE COLOR	23
5.2.1.	Descriptor RGB	24
5.2.2.	Descriptor CMYK	24
5.2.3.	Descriptor TSL	25
5.3.	DESCRIPTORES SOBEL Y CANNY	26
5.4.	DESCRIPTOR HISTOGRAMAS DE GRADIENTES ORIENTADOS	27
5.5.	DESCRIPTOR SURF	28
5.6.	DESCRIPTOR MSER <i>Maximally Stable Extremal Regions</i>	29
5.7.	VALIDACIÓN Y CLASIFICACIÓN	29
6.	CONCLUSIONES Y RECOMENDACIONES	34
7.	ESTUDIANTES	36
8.	DIRECTOR	36
	BIBLIOGRAFÍA	37

Índice de cuadros

1.	Descripción de cada clase	8
2.	Eficiencia de los descriptores RGB, CMYK, TSL y SOBEL	30
3.	Eficiencia de los descriptores CANNY, HOG, SURF y MSER	30
4.	Eficiencia y desviación estándar promedio.	31
5.	Eficiencias de las clases	32
6.	Eficiencia y desviación con fusión de datos	33
7.	Eficiencia y desviación promedio mezcla de descriptores	34
8.	Eficiencias fusión de datos sin desviación estándar	34

Índice de figuras

1.	Esquema general de implementación.	7
2.	Captura de la base de datos	8
3.	Ejemplo de cada una de las clases de la base de datos	9
4.	Filtrado de la imagen.	23
5.	Detección de bordes y segmentación	23
6.	Imagen de botella con tapa defectuosa y su descriptor correspondiente	24
7.	Ejemplo de representación en el modelo de color CMYK	25
8.	Descriptor CMYK	25
9.	Resultados modelo TSL	26
10.	Resultados operadores Canny y Sobel	26
11.	Descriptor Canny y Sobel	27
12.	Ejemplo descriptor HOG	28
13.	Ejemplo Descriptor SURF	28
14.	Ejemplo resultado MSER y descriptor del vector de características	29
15.	Gráfico de barras eficiencia y desviación estándar promedio	31

1. INTRODUCCIÓN

El continuo desarrollo de la sociedad y en especial el sector industrial cada vez es más competitivo y exigente. La demanda de productos en todos los sectores exige que los productos adquiridos cumplan siempre con estándares de calidad que garanticen el mejor desempeño, funcionalidad y satisfacción. La gran demanda de productos, hace que sea común encontrar en las líneas de producción el manejo de diferentes productos, lo que genera que las metodologías de reconocimiento aplicadas a los sistemas de identificación, no funcionen en algunos de estos ambientes. Para cumplir con el estándar anterior, la industria implementa metodologías que permiten identificar cualquier irregularidad en los productos fabricados u ofrecidos.

La etapa del reconocimiento de patrones es realizado en ocasiones por personal calificado mediante la inspección manual y visual a cada producto, siendo este un método no eficiente y de baja rigurosidad que puede presentar errores y disminuir la efectividad en los procesos debido al agotamiento físico, pérdida de atención y criterios subjetivos del ser humano [1, 2, 3]. Con el objetivo de mejorar este proceso, se ha implementado sistemas de reconocimiento automático mediante el uso de vídeo cámaras y técnicas de visión por computadora, el cual permita combinar y realizar una fusión de datos entregados por diferentes descriptores, los cuales permiten reconocer cambios de color, forma y textura presentes en un objeto. Tales cambios son de vital importancia y permiten definir modelos matemáticos para detección de patrones, como hendiduras, rayones, golpes, cambios de forma, etc [2, 4, 5, 6, 7].

Debido a la gran demanda de productos, en las líneas de producción es común encontrar que estas no manejan un solo producto, lo que conlleva a que aplicar una metodología para el reconocimiento de patrones deba cumplir con características específicas. Sumado a esto es importante tener en cuenta que no siempre se cuenta con la posibilidad de garantizar unas condiciones ideales bajo las cuales pueda operar el sistema de identificación [4]. Además, se suma que cada tipo de descriptor no se adapta a diferentes entornos, ejemplo de ello se evidencia con los descriptores de color, los cuales no son robustos a cambios de iluminación, y los descriptores morfológicos, los cuales no son robustos a oclusiones parciales, totales del objetivo o al ángulo con que incide la cámara sobre este[8], haciendo

que muchas de la metodologías descritas en el estado del arte no sean adecuadas para diferentes ambientes y que deban ser ajustadas según las condiciones del proceso y del patrón que se desea identificar [4, 6].

Lo anterior evidencia que un solo tipo de descriptor no es suficiente para caracterizar una amplia variedad de patrones en ambiente no estructurado, por lo que la fusión de datos es un enfoque que permite mejorar el factor discriminante de una observación aprovechando las ventajas individuales de cada descriptor. Esto hace pensar en el desarrollo de una metodología que utilice las ventajas de cada descriptor para obtener un mayor rendimiento en el reconocimiento de patrones.

2. DEFINICIÓN DEL PROBLEMA

El problema fundamental de la visión por computador es la dificultad de diferenciar un objeto de otro en una línea de producción. El reconocimiento de patrones en las aplicaciones industriales y comerciales ha despertado la necesidad de generar aplicaciones de diferentes niveles de complejidad ya que son muy útiles en diferentes áreas de control de calidad, robótica, seguridad, etc. Los protocolos de calidad han generado la implementación de diferentes modelos o métodos que permitan identificar las fallas en los productos ofrecidos sin cambiar los tiempos en la producción. Los problemas más frecuentes en el control de calidad son las variaciones de color, problemas con la etiqueta, rayones o golpes. Hoy en día la forma como se determinan estas fallas es mediante una revisión manual por personal calificado mediante la inspección visual de una pequeña muestra de un lote específico para determinar el resultado general, este proceso es ineficiente debido a que se incurre en criterios subjetivos, agotamiento físico y pérdida de atención[2, 3]. Esto genera errores e incertidumbre en la revisión del producto causando pérdidas económicas a la empresa, por tal razón, la pequeña muestra que se toma del lote de producción se analiza estadísticamente representando toda la producción [1]. Con el creciente aumento de la tecnología, la electrónica y los dispositivos de cómputo, se han diseñado métodos para realizar estos análisis, con el fin de mejorar el reconocimiento automático mediante el uso de vídeo cámaras y técnicas de visión por computador.

Debido a la gran demanda del mercado, es normal encontrar que se manejan diferentes productos

en una línea de producción, por tal razón, se deben implementar metodologías que cumplan con características como: adaptabilidad, precisión, rapidez, etc. Además, es importante tener en cuenta que las condiciones para trabajar estos sistemas no siempre son óptimas [4]. Se debe tener en cuenta que cada descriptor no se adapta a cualquier entorno, haciendo que unos sean más robustos a ruido, otros a cambios de iluminación entre otros; tal vez las metodologías descritas en el estado del arte no sean las más adecuadas a los ambientes analizados por ello deben ser ajustadas dependiendo del proceso o las condiciones externas, además, del patrón a analizar [4, 6].

3. OBJETIVOS

3.1. OBJETIVO GENERAL

Implementar una metodología para el reconocimiento automático de patrones (hendiduras, rayones y fracturas) en una botella plástica utilizando vídeo cámaras.

3.2. OBJETIVOS ESPECÍFICOS

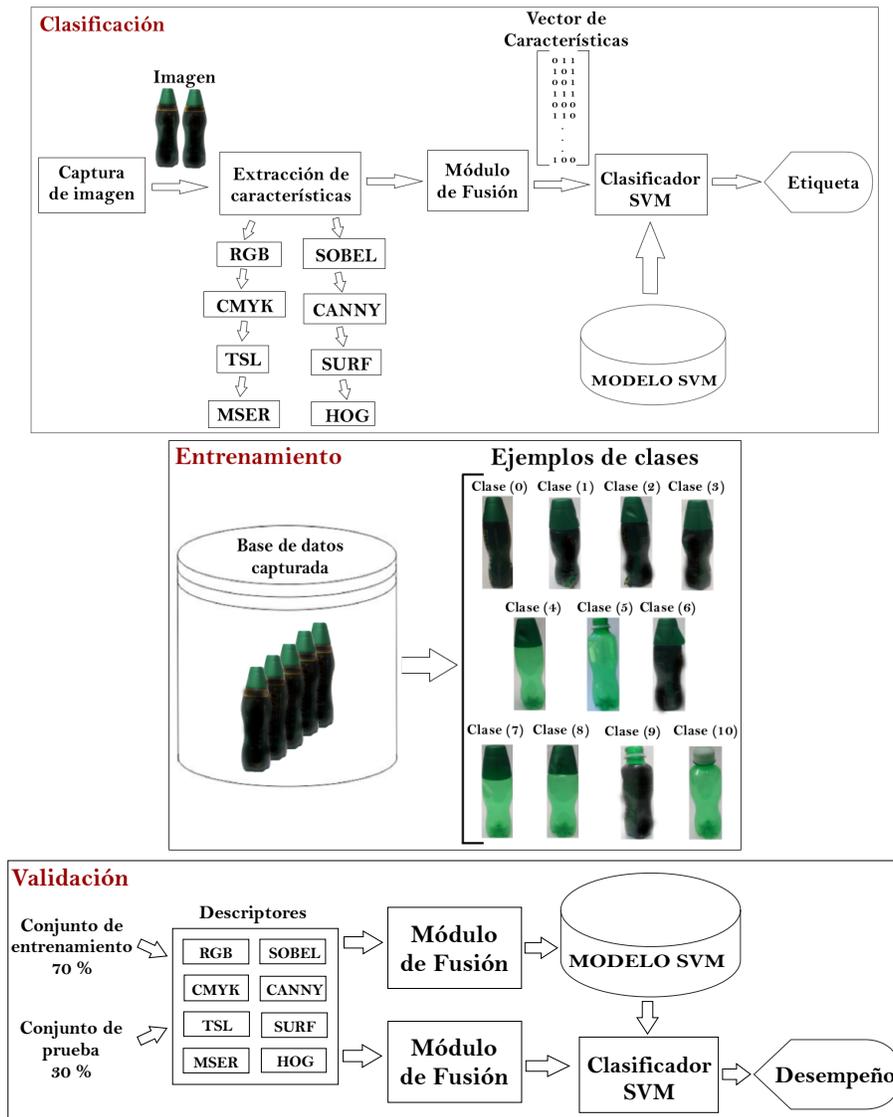
1. Definir e implementar los descriptores que serán usados para identificar color, forma y textura tales como RGB, HOG, SURF, Harris, Sobel y Canny.
2. Diseñar la metodología de fusión de datos de los descriptores definidos.
3. Diseñar la metodología de clasificación de patrones.
4. Validar estadísticamente el método de reconocimiento.

4. METODOLOGÍA

Siendo el objetivo principal de este trabajo analizar modelos basados en histogramas para la detección de patrones, como hendiduras, rayones, golpes, cambios de forma de un objeto para el diseño de aplicaciones en control de calidad utilizando descriptores basados en histogramas. Los métodos aquí descritos, poseen cualidades de gran interés ya que utilizan patrones de color, forma y textura los cuales se adecuan a las necesidades del entorno, proporcionando eficientemente un análisis sobre objetos. En posteriores secciones se explicara los conceptos, pasos y consideraciones de los métodos utilizados. Además, partiendo de los resultados de los métodos implementados se pretende realizar una combinación entre ellos para evaluar su rendimiento y analizar una a una las 11 clases que se tienen en la base de datos. En esta sección se realiza un análisis multimodal para analizar los diferentes patrones que entregan los descriptores implementados (RGB, CMYK, TSL, SOBEL, CANNY, SURF, MSER Y HOG) utilizando un método de validación adaptativo.

La metodología propuesta en este trabajo consta de tres etapas fundamentales como se puede observar en la en la figura 1. En la etapa de clasificación, una base de datos creada a partir de las imágenes del producto a evaluar, sirve de fuente para la extracción de características y posteriormente realizar una fusión de los datos, los cuales servirán para la etapa de entrenamiento y validación en la cuales mediante el uso de una SVM, sea posible establecer a que clase pertenece un conjunto de características.

Figura 1. Esquema general de implementación.



4.1. CONSTRUCCIÓN DE LA BASE DE DATOS

En un entorno industrial y más específicamente en los procesos de control de calidad, generalmente se usan condiciones ideales de iluminación y cámaras digitales tipo CCD (charge-coupled device) para satisfacer las necesidades que se requieren en el procesamiento de imágenes, tales como, eliminación de sombras, reflejos y buena definición de la imagen [2]. En el mercado es posible encontrar una gran variedad de dispositivos de iluminación generalmente tipo LED y cámaras digitales con diferentes protocolos de comunicación que satisfacen las necesidades requeridas.

La base de datos construida se creó con botellas de un producto comercial el cual fue elegido por su morfología y color, características ideales para el objeto de estudio de este trabajo. La base de datos adquirida presenta secciones difusas de color negro que tienen como objetivo no revelar la marca del producto debido a derechos de *copyright for trade mark*. Cabe resaltar que el trabajo aquí realizado y los resultados obtenidos se obtuvieron con base en las imágenes originales sin difuminación. Con el fin de reducir efectos de sombras no deseadas y resaltar el objeto del fondo, se propuso un ambiente no controlado que consta de una base fija con una tela de color blanco que posee una curvatura en su parte inferior como se puede apreciar en la figura 2.

Figura 2. Captura de la base de datos

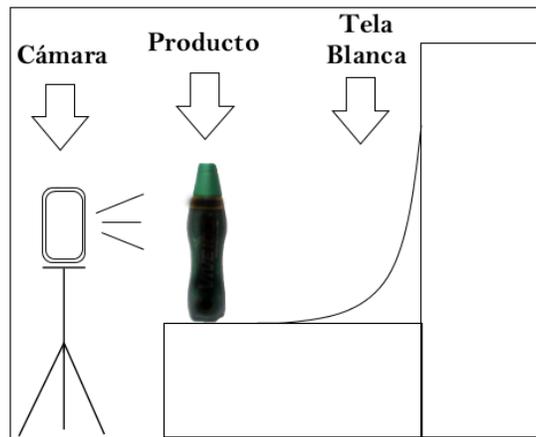
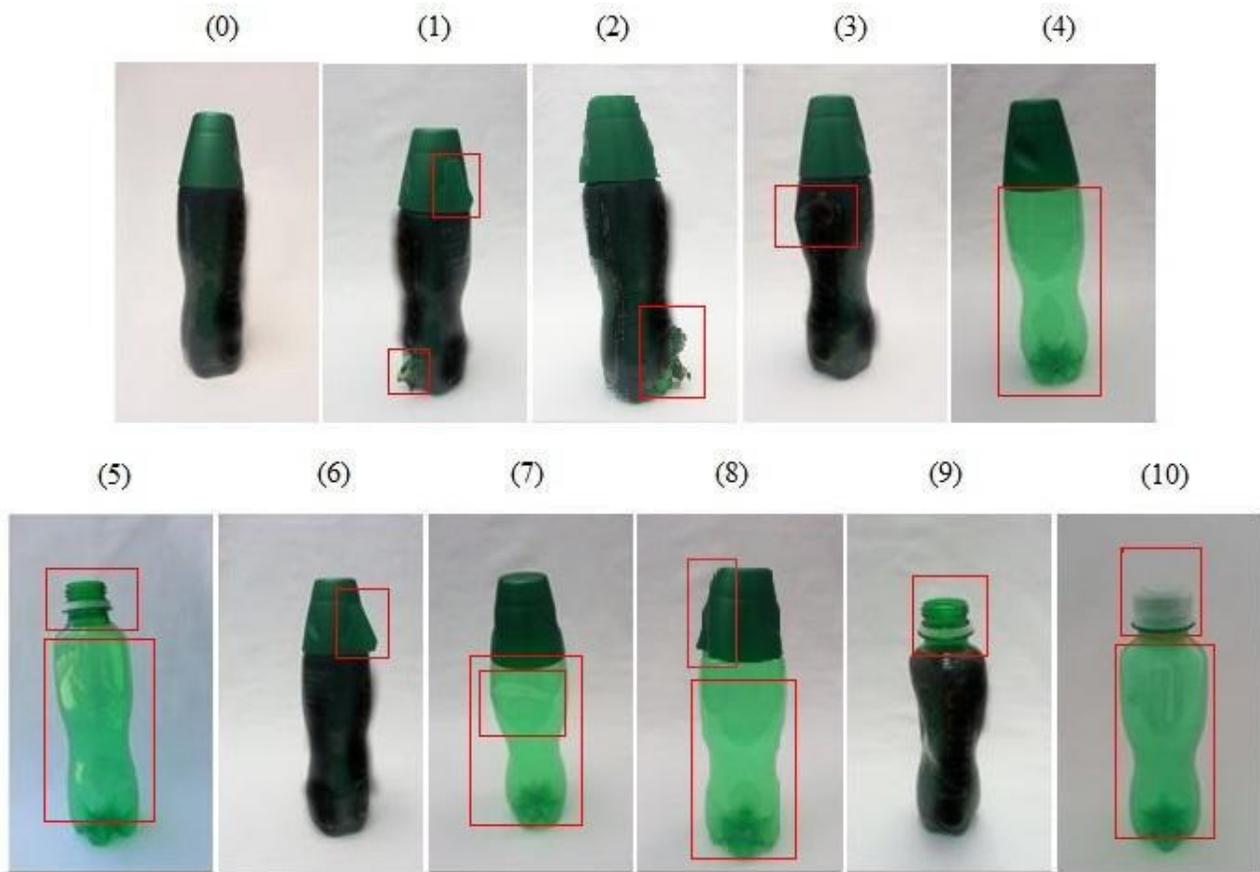


Tabla 1. Descripción de cada clase

0	Botellas buenas
1	Botella con tapa y envoltura imperfecta
2	Botella con envoltura imperfecta
3	Botella imperfecta
4	Botella sin envoltura
5	Botella sin tapa y sin envoltura
6	Botella con tapa imperfecta
7	Botella imperfecta y sin envoltura
8	Botella con tapa dañada y sin envoltura
9	Botella imperfecta y sin tapa
10	Botella sin envoltura y con tapa imperfecta

Figura 3. Ejemplo de cada una de las clases de la base de datos



La base de datos creada según el ambiente propuesto en la imagen anterior, fue adquirida mediante una cámara de un celular Samsung S3 modelo SGH I-747 a una resolución de 720×1280 , en condiciones no controladas de iluminación. Se crea entonces una base datos multiclase con 50 fotos por cada clase y en la cual existen diversas situaciones que involucran hendiduras, rayones, falta de etiquetas, botellas incompletas entre otros, como se puede apreciar en la tabla 1, cabe resaltar que dentro de cada clase se proponen múltiples imperfectos en el producto final, es decir, la clase botella con tapa imperfecta consta de tapas con rayones, tapas con hendiduras, tapas rotas o combinaciones de estas. En la figura 3, se presenta un ejemplo de cada una de las clases propuestas, donde los recuadros de color rojo destacan la característica que define cada clase.

4.2. EXTRACCIÓN DE CARACTERÍSTICAS

En esta etapa el objetivo principal es obtener de cada imagen contenida en la base de datos, la información de los datos que la representan en cada espacio de color o modelo propuesto mediante el uso de histogramas, lo cual permite la clasificación de los datos contenidos en cada imagen. Mediante los histogramas se forma lo que denominaremos el descriptor o vector de características, sin embargo,

antes de realizar este proceso, se debe realizar una etapa de pre-procesamiento, la cual se explica a continuación.

4.2.1. Pre-Procesamiento de las Imágenes

Todas las imágenes de la base de datos, cuentan con un área de fondo de color blanco que debe ser eliminada debido a que esta no pertenece al objeto de interés. Por esta razón se debe realizar una serie de procesos mediante los cuales sea posible eliminar esta zona y destacar el objeto de interés. Los procesos involucrados en esta etapa incluyen conversión a escala de grises, etapa de filtrado y finalmente se aplica el descriptor Canny el cual mediante la aplicación del gradiente a la imagen permite la detección de los bordes que pertenecen al objeto.

Conversión a escala de grises: Llevar las imágenes a este tipo de representación permite dar a cada *pixel* un valor de luminancia que varía en una escala comprendida entre el negro y blanco asignando un valor numérico a partir de los valores de los pixeles de cada una de las capas del modelo RGB. En este tipo de representación la imagen pasa de un modelo de tres capas a una mediante la siguiente ecuación:

$$I = 0,21R + 0,72G + 0,07B \quad (1)$$

Es posible observar como se involucra cada una de las capas del modelo RGB y el factor por el cual debe ser ponderado cada uno de los pixeles allí presentes.

Filtrado: Esta etapa se realiza mediante el uso de una máscara o núcleo integral tipo Gaussiano con el fin de suavizar las imágenes mediante la reducción de la intensidad entre pixeles vecinos, de esta manera resulta más efectiva la distinción entre los pixeles que pertenecen al objeto o al fondo. El filtro aplicado en este trabajo se propone como un pasa bajo tipo Gaussiano con un tamaño de mascara de 5x5 y una desviación estándar positiva de 0.9.

Segmentación por detección de bordes: En esta etapa se implementó un detector de bordes discreto de tipo diferencial mediante el operador Canny el cual calcula el gradiente de la intensidad de la imagen en cada par de coordenadas de todos los píxeles de la imagen. Con el resultado que entrega este operador es posible recortar la imagen mediante el uso de las coordenadas máximas y mínimas que fueron identificadas como un cambio de borde.

4.3. DESCRIPTORES MODELOS DE COLOR

Los modelos de color se pueden entender como una representación matemática que permite asignar un valor numérico a un color que proviene de un espacio conformado por múltiples componentes al cual se le denomina espacio de color. Estos modelos son muy utilizados en las prácticas relacionadas con el procesamiento de imágenes debido a que mediante estos es posible obtener información detallada de las características presentes en una imagen [9, 10].

4.3.1. Descriptor RGB

El espacio de color RGB se basa en el control de la intensidad de las tres componentes de luz principales en una imagen, en la cual la combinación de estas da como resultado un color deseado dependiendo de la intensidad de cada componente. Para este espacio de color las componentes RGB (R = Red, G = Green, B = Blue) se almacenan de forma independiente y su rango puede variar desde el color negro que representa ausencia de intensidades, hasta el color blanco que representa máxima intensidad en una resolución de 8 *bits*[9].

Para el cálculo del descriptor se utiliza un histograma que mide la distribución de los niveles de intensidad de la imagen, realizando uno por cada canal y finalmente concatenándolos en uno solo de tal manera que el descriptor queda de la forma $D_{rgb} = [D_r, D_g, D_b]$ donde D_r , D_g , y D_b representan el histograma de tamaño $1 \times n$ de cada capa, siendo n la cantidad de *bins*. Se eligen de manera experimental una cantidad de 20 *bins* por canal, buscando el mejor desempeño.

4.3.2. Descriptor CMYK

Este espacio de color de tipo sustractivo, es generalmente usado para los sistemas de impresión realizando una conversión a partir del modelo RGB [9], sin embargo, se ha evidenciado buenos resultados cuando es utilizado en la identificación de cambios de intensidades de color [11], lo cual se puede asociar en este trabajo como un cambio de intensidad debido a una zona con alguna irregularidad.

Para obtener el modelo CMYK a partir de los valores del RGB se tiene que:

$$C = 1.0 - R - uK_b \quad (2)$$

$$M = 1.0 - G - uK_b \quad (3)$$

$$Y = 1.0 - B - uK_b \quad (4)$$

$$K = bK_b \quad (5)$$

donde

$$K_b = \text{MIN} \{1.0 - R, 1.0 - G, 1.0 - B\} \quad (6)$$

Y $0.0 \leq u \leq 1.0$ es el factor de eliminación de color y $0.0 \leq b \leq 1.0$ es el factor de oscuridad .

Para obtener el descriptor correspondiente a este espacio de color, se debe realizar un histograma de 15 *bins* por cada capa, y posteriormente concatenar los resultados para obtener un descriptor de la forma $D_{cmyk} = [D_c, D_m, D_y, D_k]$ de tamaño 1×60 .

4.3.3. Modelo TSL

Es un espacio de color perceptual que define el color como tinte (grado en el cual un estímulo puede ser igual o diferente a otro estímulo, podría compararse con una tonalidad de blanco), saturación (colorido de un estímulo con relación a su propio brillo) y luminosidad (la luminosidad de un estímulo en relación a otro que aparece en blanco) . Aunque fue desarrollado principalmente para la detección de rostros y piel humana [12, 13], resulta muy eficiente en la segmentación del color en general [14]

y con base en que este modelo es robusto a cambios de iluminación, se propone su implementación en este trabajo.

$$T = \begin{cases} \frac{1}{2\pi} \arctan \frac{r'}{g'} + \frac{1}{4}, & \text{if } g' > 0 \\ \frac{1}{2\pi} \arctan \frac{r'}{g'} + \frac{3}{4}, & \text{if } g' < 0 \\ 0, & \text{if } g' = 0 \end{cases} \quad (7)$$

$$S = \sqrt{\frac{9}{5} (r'^2 + g'^2)} \quad (8)$$

$$L = 0.299R + 0.587G + 0.114B \quad (9)$$

donde

$$r' = r - \frac{1}{3} \quad (10)$$

$$g' = g - \frac{1}{3} \quad (11)$$

$$r = \frac{R}{R + G + B} \quad (12)$$

$$g = \frac{G}{R + G + B} \quad (13)$$

Para obtener el descriptor de este espacio de color, se parte del hecho de que la información del canal L contiene la variación de iluminación presente en la imagen, razón por la cual esta capa no es tenida en cuenta, quedando entonces los canales T y S con la información que corresponden al tono y saturación. Con base en lo anterior se define el descriptor $DTSL = [Dt, Ds]$ de dimensiones 1×60 correspondiente a los histogramas de los canales T y S con una cantidad de 30 *bins* por canal.

4.4. DESCRIPTORES MODELOS DE TEXTURA

Los operadores Canny y Sobel permiten la identificación de puntos de interés, para el caso de estos cálculos los cambios bruscos de intensidad como bordes en la escena, las texturas presentes en objetos entre otros [15, 16, 17]. Por lo tanto, estos métodos se proponen como una herramienta útil para la detección de texturas que sirven como patrón para identificación de imperfección en un producto. En esta sección se hará uso *Image Processing toolbox* de *Matlab*[®].

4.4.1. Descriptor SOBEL

Este método que mediante la convolución de dos matrices o *Kernel* con una imagen I en escala de grises, evalúa los cambios de intensidad entre los vecinos de un *pixel* de forma horizontal y vertical. El resultado de realizar la convolución es almacenado en dos vectores de manera independiente, de tal forma que, uno almacena las variaciones en dirección x y otro las variaciones en dirección y [18, 2], como se muestra a continuación.

$$G_x = \begin{vmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{vmatrix} * I \quad (14)$$

$$G_y = \begin{vmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{vmatrix} * I \quad (15)$$

$$G = \sqrt{G_x^2 + G_y^2} \quad (16)$$

$$\Theta = \arctan\left(\frac{G_x}{G_y}\right) \quad (17)$$

Para el caso del descriptor, este se define a partir de las coordenadas (x, y) en pixeles, que se hallan identificado como un borde, realizando una concatenación de los histogramas para cada coordenada. De tal manera que de una imagen resulta un vector que contiene un conjunto de pares ordenados de la forma $Vp[(x_1, y_1), \dots, (x_i, y_i)]$ con $i = 1, 2, 3, 4, \dots, n$, representando n la cantidad máxima de coordenadas. Esto genera dos histograma, denominados Hx y Hy de 50 *bins* cada uno que finalmente forman el descriptor de la forma $Dcs = [Hx, Hy]$, de tamaño 1×100 .

4.4.2. Descriptor CANNY

Asumiendo una imagen en escala de grises, el método de CANNY permite detectar las discontinuidades debidas a la diferencia en la magnitud entre dos pixeles. En este método se aplica filtrado, decisión inicial e histéresis [19]. Este operador está basado en un modelo de dominio continuo de un borde de paso de la amplitud h_E más ruido Gaussiano y desviación estándar σ_n . La detección de los bordes se realiza entre la convolución de una función de flanco $f(x)$ y una respuesta al impulso

$h(x)$, con límites $[-W, W]$. Un borde está marcado por la convolución del gradiente $f(x)*h(x)$. La respuesta al impulso de Canny debe satisfacer 3 criterios que son [9]:

- **Buena detección:** la amplitud de la señal a ruido de la pendiente se maximiza para disminuir la probabilidad de error y marcar puntos de bordes reales y disminuir el marcado de puntos de borde falsos. La relación SNR para este modelo es:

$$SNR = \frac{h_E S(h)}{\sigma_n} \quad (18)$$

con

$$S(h) = \frac{\int_{-w}^0 h(x) dx}{\int_{-w}^w [h(x)]^2 dx} \quad (19)$$

- **Buena Localización:** los puntos de borde que son marcados por el operador deben estar lo más cerca posible del centro del borde. El factor de localización está definido por:

$$LOC = \frac{h_E L(h)}{\sigma_n} \quad (20)$$

con

$$L(h) = \frac{h'(0)}{\int_{-w}^w [h'(x)]^2 dx} \quad (21)$$

donde $h'(x)$ es la derivada de $h(x)$.

- **Única respuesta:** solo debe existir una respuesta a un verdadero borde. La distancia entre los picos del gradiente se denota x_m solo cuando se presenta ruido, y se toma una fracción k del factor de anchura W . Para ello se obtiene:

$$x_m = kW \quad (22)$$

El descriptor para este operador se implementa de la misma manera que para Sobel, debido a que estos modelos finalmente entregan el resultado de la convolución de la misma forma.

4.5. DESCRIPTORES MODELOS DE FORMA

Los métodos propuestos a continuación identifican las zonas de interés en las cuales se presentan cambios de intensidad, HOG lo hace mediante la aplicación del gradiente en pequeñas subregiones de la imagen, SURF utiliza réplicas de la imagen para identificar regiones invariantes a la escala y

MSER se basa en las regiones en las cuales identifica estabilidad de intensidad entre los *pixel* vecinos. En la implementación del descriptor se utilizó *Computer Vision System Toolbox* de *Matlab*[®] para obtener el vector de características.

4.5.1. Descriptor HOG

Los HOG (histogramas de gradientes orientados) son comúnmente usados en la identificación de objetos mediante la extracción de los bordes presentes en una imagen. Los HOGs dividen la imagen en secciones o celdas y en cada una de estas se hace uso del histograma proveniente de los gradientes orientados los cuales identifican en qué dirección y con qué magnitud varían los pixeles en una imagen, de esta forma se crean vectores que marcan la variación de la intensidad de color. Generalmente este descriptor es aplicado en una imagen en escala de grises con lo cual se logra disminuir el ruido debido a una deficiente iluminación [20].

El histograma de gradientes orientados es un histograma de la orientación del gradiente presente en la imagen $g(x)$. Por lo tanto, la orientación está estimada para cada *pixel*. El gradiente $g(x) = (g_x(x), g_y(x))^T$ puede estar estimado por diferentes filtros lineales. Para algunas derivadas, la orientación $\theta(x)$ y la norma $u(x)$ del gradiente de la imagen de *pixel* puede estimarse de la siguiente manera: Solo los *pixel* que son considerados en el cálculo del gradiente son aquellos para los que $u(x)$ tiene un umbral fijo u_{min} , ya que la dirección de gradientes bajos no se puede estimar de forma confiable debido al ruido.

$$\theta(x) := \arctan \frac{g_y(x)}{g_x(x)} \in [0, \pi) \quad (23)$$

$$u(x) := \|\nabla g(x)\| = \sqrt{(g_x(x))^2 + (g_y(x))^2} \quad (24)$$

En el cálculo del descriptor HOG, se utiliza la configuración de 0 a 360° con un tamaño de celda de 32×32 que se propone para detección de objetos que no sean humanos, que entrega la concatenación de los diferentes histogramas de la escena. Para la configuración de la función se utiliza un bloque de y sin traslape y se genera un histograma para el descriptor de 100 *bins*.

4.5.2. Descriptor SURF

El método SURF (Speedeered Up Robust Features) es un identificador de esquinas de alta eficiencia, basado en la asignación de un factor de escala y de orientación asociado a cada punto característico y mediante el cual es posible tener una buena correspondencia entre dos o más imágenes de un mismo objeto tomadas a una diferente escala, luminosidad y presencia de ruido. Este algoritmo funciona principalmente desarrollando dos etapas, una es la detección de los puntos de interés y la otra es la descripción de los puntos de interés. Cada uno de los puntos anteriores consta de unos pasos para su desarrollo. En la detección de los puntos de interés lo primero es integrar la imagen, la imagen que se obtiene como resultado se utiliza para el análisis del espacio de escala, y a partir de esta imagen se puede hallar el punto de interés. En la descripción del punto de interés se le asignan orientaciones a los puntos localizados y luego se describen las características.

Una imagen integral $I_{\Sigma}(x)$ en una locación $x = (x, y)$ se utiliza para resumir todos los *pixel* dentro de una región rectangular formada por el origen y x . Su fórmula se define así:

$$I_{\Sigma}(x) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j) \quad (25)$$

Una vez calculada la integral de la imagen, es necesario realizar cuatro pasos para calcular la suma de las intensidades sobre cualquier área. SURF es un detector basado en la Matriz Hessiana, ya que el determinante de esta matriz es usado para obtener las medidas de ubicación y escala. Esta matriz es aproximada, usando un conjunto de filtros y sin suavizado al pasar de una escala a otra, para hallar la matriz Hessiana $\mathcal{H}(x, \sigma)$ en x a escala σ se define así:

$$\mathcal{H}(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (26)$$

donde $L_{xx}(x, \sigma)$ es la segunda derivada de la Gaussiana representada por $\frac{\partial^2}{\partial x^2} g(\sigma)$ con la imagen I en el punto x , y es similar para $L_{xy}(x, \sigma)$ y $L_{yy}(x, \sigma)$. La máscara de 9x9 son aproximaciones de un Gaussiano con $\sigma = 1.2$ la cual representa la escala mas baja.

EL descriptor se realiza para una cantidad de 100 *bins* con base en el vector de características que entrega la aplicación de este modelo.

4.5.3. MSER

Es utilizado como un método de reconocimiento de regiones especiales en las imágenes analizando la intensidad entre *pixels*. Esta técnica se propuso para encontrar la correspondencia entre los elementos de dos imagen con diferente punto de vista. *Maximally Stable Extremal Regions* es una componente que conecta aproximadamente el umbral de la imagen. La palabra “*Extremal*” se refiere a la propiedad que todo los *pixels* tienen entre su mayor y menor intensidad con los pixels del límite exterior. El “*Maximally Stable*” describe la propiedad que optimiza el umbral en la selección del proceso. El conjunto de todos los componentes obtenidos por umbralización, tiene un número de propiedades ideales. Primero, un cambio repetitivo en la imagen no permite cambios. Segundo, las transformaciones continuas de la geometría conservan la topología. Por último, no hay más regiones extremas que los *pixels* que hay en la imagen. Por esta razón se crea un conjunto de regiones que conserva bajo la misma geometría los cambios y aún conserva la misma importancia, el cual está descrito por:

Una región r_t está conectada por un conjunto de elementos de la escena s con una intensidad $f(s) > t$, donde t es un umbral arbitrario.

Una medida de inestabilidad $s_\Delta(r_t, t)$ mide los cambios de tamaño de r_t cuando el umbral es cambiado por algún valor de Δ , como se muestra en la siguiente ecuación:

$$s_\Delta(r_t, t) = |r_{t+\Delta} - r_{t-\Delta}| / |r_t| \quad (27)$$

donde $|\cdot|$ indica la computación del volumen de la región respectiva.

Una región r_t es máximamente estable si $s_\Delta(r_t, t)$ es mínimo para t [21].

EL descriptor se realiza para una cantidad de 100 *bins* con base en el vector de características que entrega la aplicación de este modelo.

4.6. FUSIÓN DE DATOS

La fusión de datos es fundamental en cualquier proceso o tarea en la cual se hace necesario combinar, fusionar y agrupar gran cantidad de datos para obtener la mejor calidad con respecto al punto de interés. Es importante resaltar que este proceso se encuentra presente en aplicaciones de robótica, control de procesos, equipos de monitoreo, etc[22]. Básicamente en este trabajo la fusión de datos consiste en la combinación lineal de los descriptores calculados en la etapa de extracción de características. Esta combinación lineal consisten en la concatenación de los descriptores para formar un solo vector de características.

4.7. ENTRENAMIENTO Y VALIDACIÓN

En esta etapa se pretende mediante el uso de los descriptores obtenidos, la identificación de las clases propuestas en la máquina de soporte mediante el método de validación cruzada Montecarlo. Es importante resaltar que antes de ingresar los descriptores en la SVM, estos son introducidos en matrices, una por cada descriptor, de manera que existen tantas matrices como modelos implementados.

4.7.1. Maquinas de soporte vectorial

Las máquinas de soporte vectorial han resultado ser más eficientes en comparación con otras técnicas para la clasificación de imágenes. Están basadas en un método de clasificación de un hiperplano que separa dos conjuntos de datos pertenecientes a clases distintas, haciendo énfasis en la frontera de separación y utilizando técnicas de programación cuadrática. Después de la fase de aprendizaje se comprueba el error y se toma otra muestra de datos para ser comparada con la clase real.

En estos procesos de clasificación se consideran solo dos clases: Positiva ($y = 1$) y Negativa ($y = -1$). Algunas veces sucede que los datos no sean linealmente separables o presentan cierto nivel de ruido. Por ende se utilizan diferentes tipos de SVM para implementarlos según se requiera. Los tipos de SVM son:

- SVM lineal con margen máximo.
- SVM para la clasificación no lineal
- SVM lineal con margen blando

La SVM se considera un clasificador binario, el cual calcula una etiqueta $y \in \{-1, 1\}$ utilizando el siguiente modelo:

$$y = w^T \phi(x) + b \quad (28)$$

Donde, $\phi(x) : \mathbb{R} \rightarrow H$ es una transformación del conjunto de entrada a un conjunto con igual o mayor dimensión, lo que se supone que facilita la separación de las clases. El vector w define el hiperplano de separación en espacio $\phi(X)$. b representa el *bias* o sesgo con respecto al origen de coordenadas. La SMV aborda el problema de clasificación relajando el concepto de margen, para lo que se introduce variables de sesgo $\xi_i \geq 0$, donde la función de costo o función *fitness* se define como:

$$c(y, f(x, w)) = \begin{cases} 0 & \text{si } y = \text{sgn}(w^T \phi(x) + b) \\ 1 & \text{si } y \neq \text{sgn}(w^T \phi(x) + b) \end{cases} \quad (29)$$

donde existe el riesgo empírico $R_{emp} = \sum_{i=1}^n \xi_i^\sigma$, siendo generalmente $\sigma = 1$ en problemas de clasificación. Por lo tanto, el problema de estimación del hiperplano de separación, se define como la solución al siguiente problema de optimización cuadrático (QP-*Quadratic Problem*):

$$\text{mín } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad (30)$$

Sujeto a

$$y_i (w^T \phi(x_i) + b) \geq 1 - \xi_i \forall i = 1, \dots, n \quad (31)$$

$$\xi_i \geq 0 \forall i = 1, \dots, n \quad (32)$$

Donde:

- x_i es una muestra del conjunto de entrenamiento con su respectiva etiqueta.
- C se define como una constante de regularización.

4.7.2. Máquinas de Soporte Vectorial Multiclase

Las SVM son consideradas clasificadores para resolver problemas binarios, de solo dos clases, sin embargo, es común encontrar problemas en los que se debe clasificar $k > 2$ distintas, donde k es el número de clases que se deben separar. Para ello existen diferentes metodologías con las cuales a partir de las SVM se puede llegar a la clasificación multiclase, estas metodologías son:

Uno contra todos: Consiste en la estimación de una SVM por clase, la cual es capaz de distinguir las características propias de la clase seleccionada de las características del resto de clases. Generalmente, cuando hay una clase en particular, la SVM detecta positivo un ejemplo y se le asigna la respectiva etiqueta. Cuando más de un modelo SVM separa los datos positivos, se calcula la probabilidad de pertenecer a la clase y se escoge la que tenga máxima probabilidad.

Uno contra uno: Consiste en construir un clasificador por par de clases, la pertenencia a la clase se obtiene observando la que halla recibido más votos, si existe un empate entre dos clases, se calcula la probabilidad de pertenecer a la clase y se escoge la que tenga máxima probabilidad.

Para implementar esta metodología se requiere de m clasificadores, este método es más lento que uno contra todos. Para calcular m se utiliza la siguiente fórmula:

$$m = \frac{n_{clases}(n_{clases} - 1)}{2} \quad (33)$$

4.8. VALIDACIÓN DEL CLASIFICADOR UTILIZANDO MONTECARLO

En esta sección se verifica el desempeño de los clasificadores en la detección de fallas en el objeto analizado (botella) utilizando la información obtenida en la captura de las imágenes. Además, se desarrollan dos etapas muy importantes y decisivas: la primera es la clasificación, la cual consiste

en analizar los modelos para encontrar el método que funciona mejor, y la validación que es donde se aplica el método de MonteCarlo para encontrar la convergencia del desempeño garantizando una buena respuesta de este análisis.

En la etapa de clasificación, se utilizó una máquina de soporte vectorial (SVM) con estrategia «*One-vs-All*», este método de aprendizaje utiliza un kernel Gaussiano con radio adaptativo y la estimación de los vectores de soporte lo calcula con el método de minimización «*sequential minimal optimization*» (SMO).

Para realizar el procedimiento de evaluación del modelo entrenado se utilizó una estrategia de validación cruzada, separando la base de datos 70% y 30%. Con el primer porcentaje se realiza el entrenamiento y con el otro porcentaje se realiza la evaluación en cada iteración de MonteCarlo; este proceso realiza las iteraciones con criterio de paro $\|diag(M_k) - diag(M_{k-1})\| < 0.001$, donde $diag(M_k)$ es el vector generado por la diagonal de la matriz de confusión y k es la iteración promedio actual de MonteCarlo. Las estadísticas obtenidas en el experimento consisten en el cálculo de la matriz de confusión y el comportamiento promedio del porcentaje de acierto y su desviación estándar σ .

5. ANÁLISIS Y RESULTADOS

5.1. FILTRADO Y SEGMENTACIÓN

Como se explicó en la sección 4.2.1, la etapa de filtrado se realiza con el fin de suavizar la imagen y disponerla para la identificación de sus bordes. En la figura 4 se muestra la función del filtro aplicado y el resultado de su aplicación en una imagen. Es posible observar en la parte superior de la botella, que la línea del contorno adquiere una mejor definición. Ya realizado el paso anterior la imagen debe ser segmentada mediante la identificación de los bordes y posterior elección de las coordenadas máximas y mínimas, resultado que se evidencia en la figura 5

Figura 4. Filtrado de la imagen.

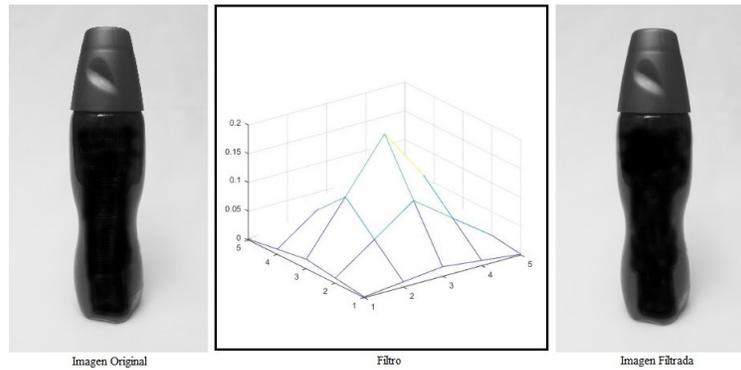
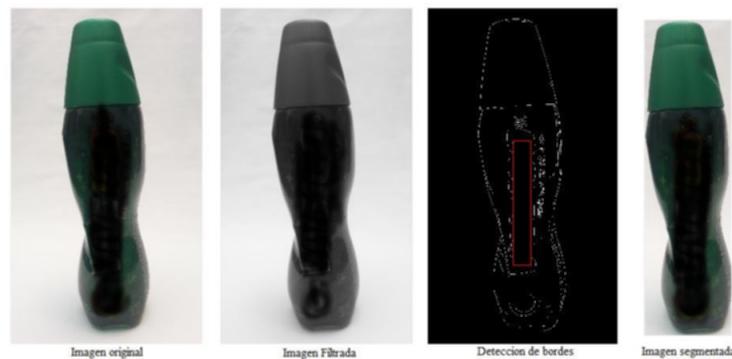


Figura 5. Detección de bordes y segmentación



Los resultados mostrados en las dos figuras anteriores evidencian que una buena segmentación con base en una imagen a escala de grises y que ha pasado por un proceso de filtrado, permite eliminar el área de fondo presente en la imagen, recortándola al tamaño que ocupa el objeto.

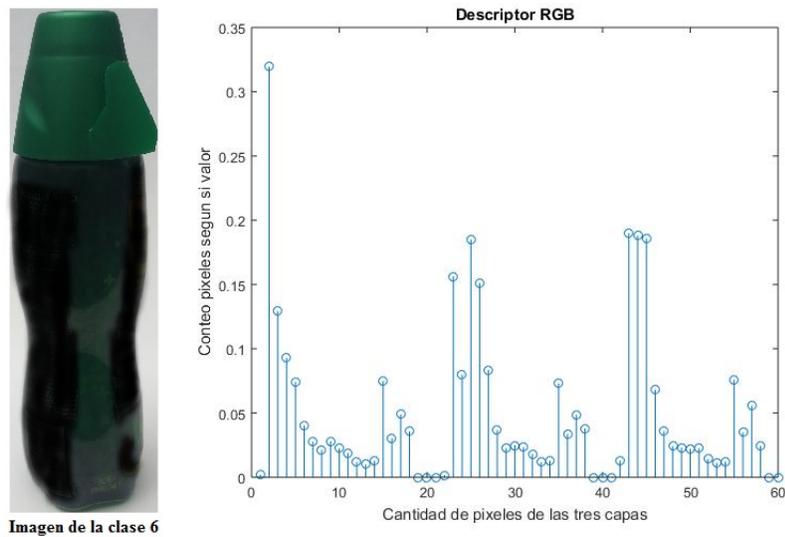
5.2. DESCRIPTORES DE COLOR

Los modelos de color no solo permiten representar una imagen de forma numérica, sino que también son prácticos y útiles cuando se desea resaltar ciertos aspectos de la imagen o simplemente por conveniencia específica para alguna aplicación. En este trabajo los modelos de color implementados se describen con detalle en la sección 4.2. A continuación se deja evidencia de los resultados de cada modelo implementado y finalmente se presenta un ejemplo del descriptor basado en el histograma que se crea a partir de cada imagen.

5.2.1. Descriptor RGB

En la figura 6, se presenta el descriptor obtenido a partir de una de las imágenes de la base de datos con base en lo explicado en la sección 4.3.1, allí es posible observar que debido a el aporte de cada capa se presentan 3 secciones notoriamente marcadas aportando información para el descriptor de manera independiente.

Figura 6. Imagen de botella con tapa defectuosa y su descriptor correspondiente



5.2.2. Descriptor CMYK

En la figura 7 se puede apreciar el resultado de la aplicación de este modelo, allí se pueden observar las cuatro capas que conforman la imagen en este tipo de representación. Este modelo se obtuvo a partir del modelo RGB como se mostró en la sección 4.3.1 y el descriptor que contiene la información de cada uno de los histogramas de las cuatro capas se puede observar en la figura 8.

Figura 7. Ejemplo de representación en el modelo de color CMYK

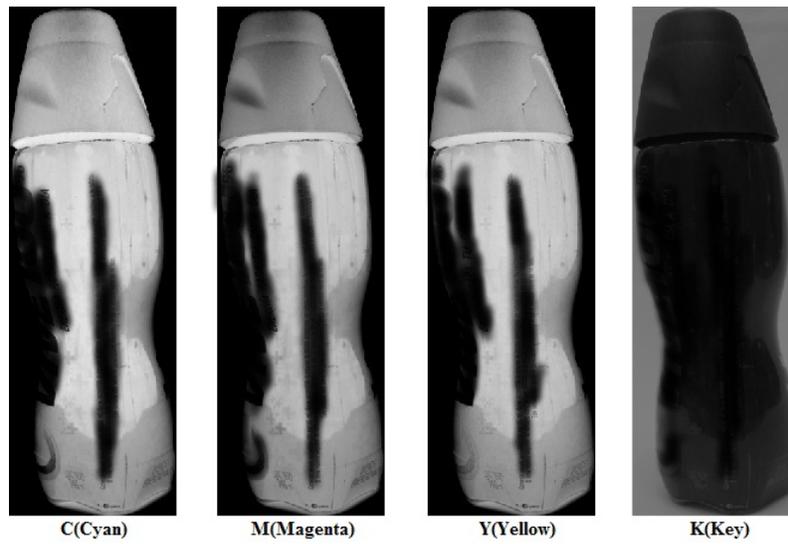
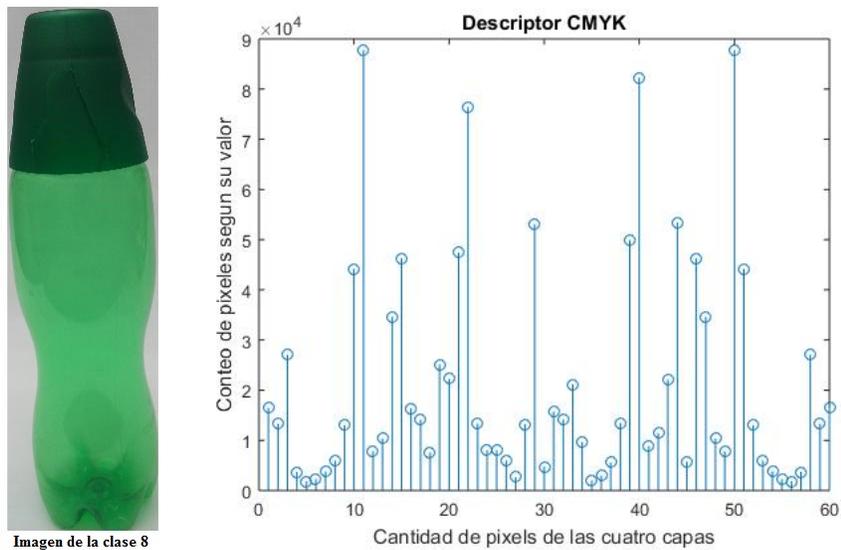


Figura 8. Descriptor CMYK

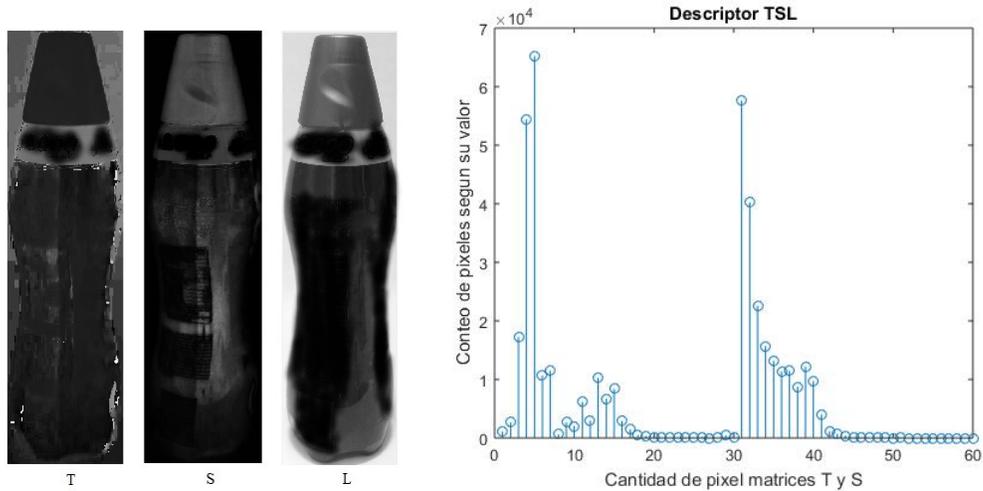


5.2.3. Descriptor TSL

Como se mencionó en la sección 4.3.3 el descriptor de este modelo no contendría la información aportada por la capa de luminancia L , el resultado de esta condición se puede evidenciar en la figura 9 donde se observa dos regiones las cuales pertenecen al aporte de la capa T y la capa S . Además,

en esta misma figura se puede observar una imagen de la base de datos representada en el modelo TSL mostrando cada una de las capas.

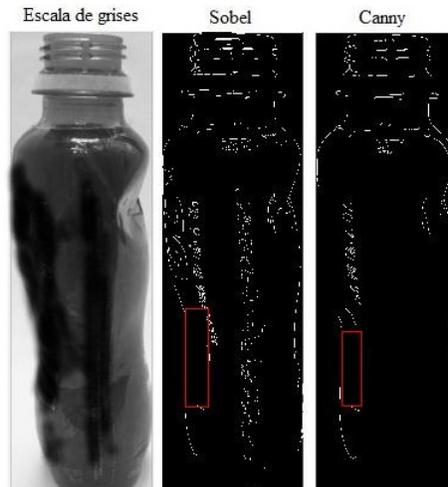
Figura 9. Resultados modelo TSL



5.3. DESCRIPTORES SOBEL Y CANNY

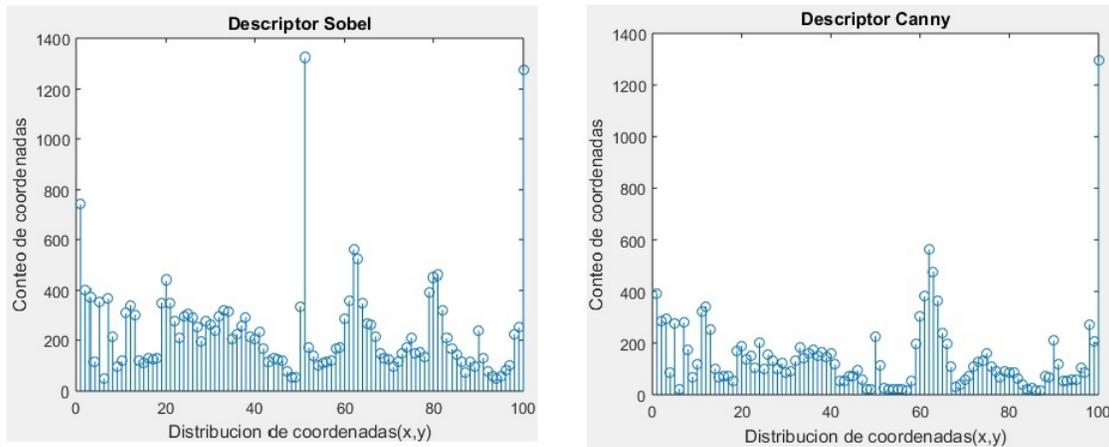
Los resultados obtenidos para Sobel y Canny basados en lo expuesto en la sección 4.4.1 y 4.4.2 se presentan en la figura 10, allí es posible observar que ambos operadores logran identificar puntos que pertenecen a una zona con imperfección, la cual se resalta en el área señalizada.

Figura 10. Resultados operadores Canny y Sobel



Los descriptores generados a partir de el resultado del histograma de ambos operadores, se presenta en la siguiente figura, donde es posible observar que existe cierta similitud entre ellos, la cual es debida a que ambos métodos utilizan la convolución como operador principal para obtener lo descrito en las secciones 4.4.1 y 4.4.2.

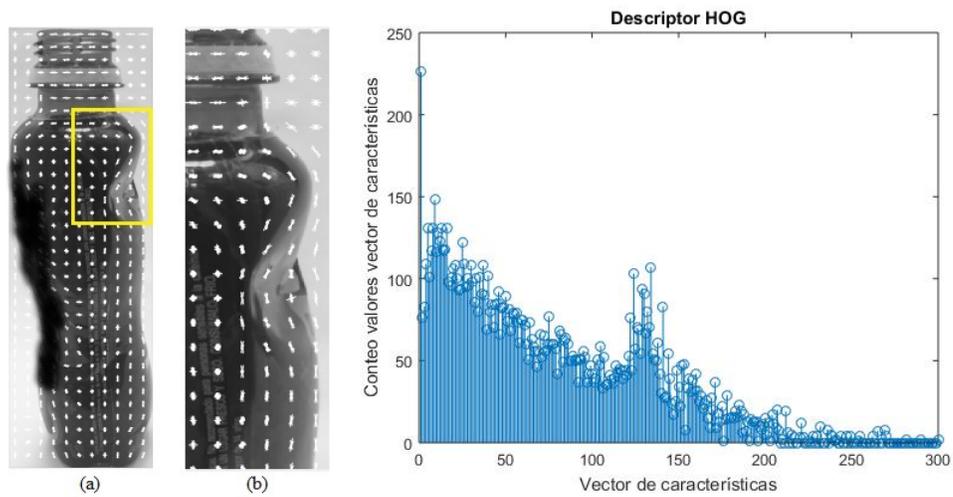
Figura 11. Descriptor Canny y Sobel



5.4. DESCRIPTOR HISTOGRAMAS DE GRADIENTES ORIENTADOS

Los HOG suelen ser de gran utilidad cuando se desea detectar regiones de interés razón por la cual se realizó su implementación, con base en lo mencionado en la sección 4.5.1. El vector de características de los resultados obtenidos es de una tamaño $1 \times n$ y allí se encuentra contenida toda la información de las regiones encontradas en las cuales se presenta un resultado positivo para el cálculo del gradiente y con esta información es con la cual se realizan las labores de clasificación, detección, o seguimiento según sea el caso, siendo la primera la razón de este trabajo. El descriptor implementado con base en la información recibida del vector de características se muestra en la figura 12, allí también es posible observar la orientación de los gradientes en la sección que tiene un imperfecto.

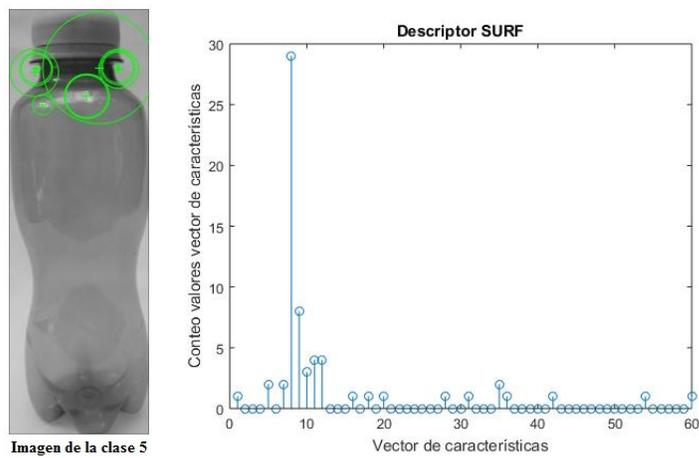
Figura 12. Ejemplo descriptor HOG



5.5. DESCRIPTOR SURF

En la siguiente figura se muestra el descriptor generado a partir del histograma del vector de características para una configuración de 4 octavas.

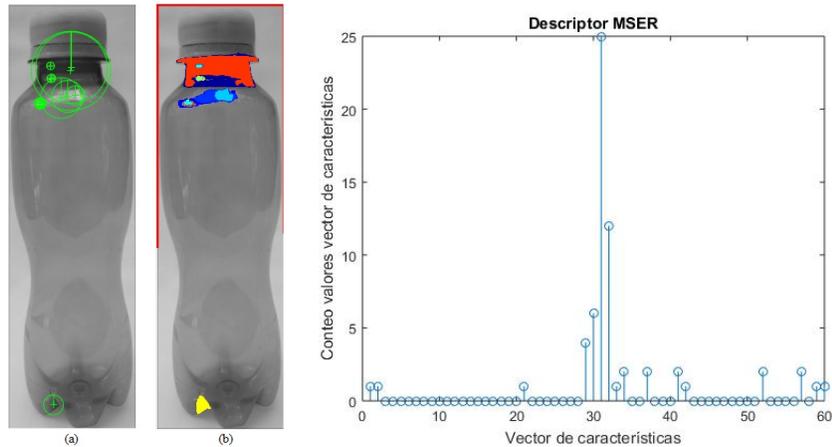
Figura 13. Ejemplo Descriptor SURF



5.6. DESCRIPTOR MSER *Maximally Stable Extremal Regions*

En la siguiente figura se muestran las regiones identificadas en las cuales no se presentan cambios bruscos de intensidad entre píxeles.

Figura 14. Ejemplo resultado MSER y descriptor del vector de características



5.7. VALIDACIÓN Y CLASIFICACIÓN

Luego de haber usado los descriptores en la etapa de validación y clasificación, los resultados a continuación presentados corresponden a los entregados por el método de validación. En los resultados solo se presenta la diagonal principal de la matriz de confusión ya que estos valores corresponden a el porcentaje de acierto de clasificación, los demás valores de esta matriz indican el error en la clasificación. Las tablas 2 y 3, contienen los valores correspondientes al porcentaje de acierto para cada uno de los descriptores.

Tabla 2. Eficiencia de los descriptores RGB, CMYK, TSL y SOBEL

CLASE	RGB	CMYK	TSL	SOBEL
0	89,04 ± 8,11	97,44 ± 4,87	73,19 ± 11,48	79,47 ± 12,2
1	82,13 ± 10,83	59,99 ± 13,63	70,8 ± 13,72	58,15 ± 15,6
2	66,28 ± 14,55	67,43 ± 13,13	72,74 ± 13,9	73,52 ± 15,57
3	56,44 ± 14,93	70,81 ± 15,25	78,35 ± 12,32	73,46 ± 15,14
4	99,96 ± 0,67	99,9 ± 0,83	98,49 ± 4,8	86,6 ± 10,57
5	98,43 ± 4,84	98,06 ± 4,37	98,46 ± 5,06	99,63 ± 1,82
6	60,34 ± 14,98	69,49 ± 15,72	77,52 ± 11,66	50,36 ± 12,5
7	86,99 ± 9,52	87,32 ± 8,86	86,79 ± 9,95	74,42 ± 9,96
8	88,55 ± 9,25	93,5 ± 8,17	89,51 ± 7,77	78,96 ± 12,01
9	98,71 ± 2,74	96,53 ± 5,02	100 ± 0	96,52 ± 5,62
10	93,6 ± 5,47	92,29 ± 7,31	87,35 ± 7,56	88,08 ± 7,63

Tabla 3. Eficiencia de los descriptores CANNY, HOG, SURF y MSER

CLASE	CANNY	HOG	SURF	MSER
0	73,75 ± 14,85 %	91,35 ± 7,84 %	53,5 ± 17,63 %	84,65 ± 13,42 %
1	50,49 ± 14,6 %	67,09 ± 13,62 %	31,19 ± 13,23 %	33,57 ± 11,68 %
2	74,06 ± 14,36 %	79,71 ± 14,22 %	35,8 ± 13,78 %	40,91 ± 15,04 %
3	64,73 ± 13,91 %	93,44 ± 6,99 %	55,29 ± 13,3 %	57,53 ± 15,09 %
4	73,29 ± 12,54 %	95,21 ± 5,27 %	51,39 ± 21,19 %	68,38 ± 16,62 %
5	98,89 ± 3,24 %	95,11 ± 5,62 %	63,7 ± 13,49 %	59,99 ± 14,38 %
6	57,93 ± 13,19 %	82,43 ± 13,29 %	51,12 ± 17,24 %	46,76 ± 15,8 %
7	75,41 ± 12,96 %	87,55 ± 11,56 %	42,84 ± 13,09 %	47,46 ± 16,42 %
8	62,46 ± 10,96 %	90,75 ± 7,85 %	55,43 ± 19,86 %	65,62 ± 23,57 %
9	88,06 ± 8,77 %	100 ± 0 %	54,93 ± 13,91 %	90,29 ± 12,05 %
10	93,1 ± 6,64 %	91,04 ± 5,6 %	73,11 ± 15,73 %	47,38 ± 10,66 %

Los descriptores de color a pesar de tener algunas clases con un rendimiento inferior al 80 %, presentan resultados muy aceptables en sus demás clases que varían entre el 82, 13 ± 10, 83 % y 100 ± 0 %, este último resultado lo presentó TSL en la clase 9. HOG al igual que TSL identifica la clase 9 con un acierto del 100 ± 0 %, sin embargo, los resultados obtenidos en HOG comparados con los de los demás descriptores son los que presentan las eficiencias más altas. Por otro lado los resultados de los descriptores de textura presentan algunas clases con eficiencias que superan el 80 % pero a nivel general su porcentaje de clasificación se puede considerar muy bajo. Finalmente se observa que los descriptores SURF y MSER son los que presentan los resultados más bajos presentando eficiencias de 31, 19 ± 13, 23 % en SURF y de 33, 57 ± 11, 68 % en MSER.

Es importante conocer cual es el promedio de los resultados contenidos en las dos tablas anteriores,

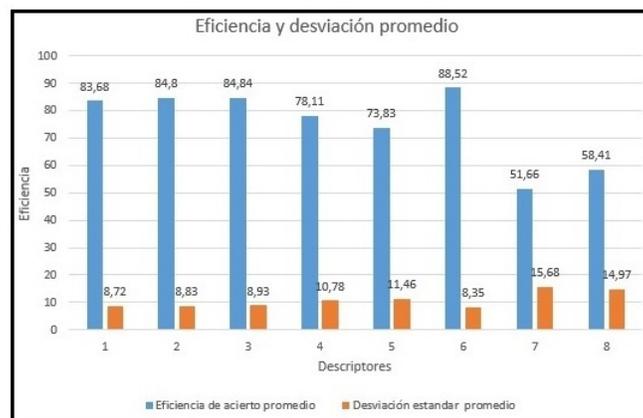
a continuación se presenta este resultado mediante una tabla y un gráfico de barras en el cual se evidencia que:

- De los descriptores de forma HOG es el que presenta la mayor eficiencia promedio con un valor del $88,52 \pm 8,35$ %, siendo este un valor muy aceptable para la clasificación e identificación de las etiquetas propuestas.
- La eficiencia de los modelos de color está por encima del 80 %, lo cual indica que los descriptores basados en estos modelos son aptos para la clasificación, sin embargo es importante tener en cuenta que para las clases 1, 2, 3 y 6 se presenta la eficiencia mas baja en los tres descriptores con un mínimo del $60,34 \pm 14,98$ % y un máximo del $82,13 \pm 10,83$ %.
- Se evidencia que los resultados SURF y MSER presentan la eficiencia promedio mas baja con respecto a los demás, razón por la cual, no seria conveniente su uso en la detección de las clases propuestas.

Tabla 4. Eficiencia y desviación estándar promedio.

Descriptor	Eficiencia promedio	Desviación estándar promedio
RGB	83,68 %	8,72 %
CMYK	84,8 %	8,83 %
TSL	84,84 %	8,93 %
SOBEL	78,11 %	10,78 %
CANNY	73,83 %	11,46 %
HOG	88,52 %	8,35 %
SURF	51,66 %	15,68 %
MSER	58,41 %	14,97 %

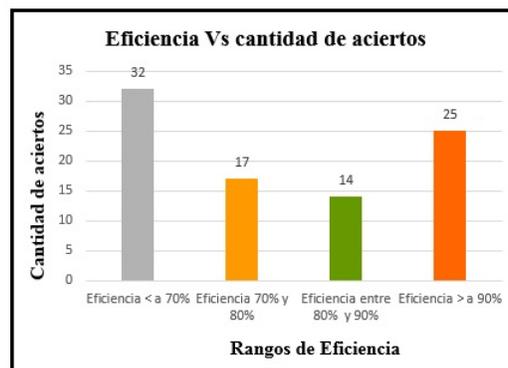
Figura 15. Gráfico de barras eficiencia y desviación estándar promedio



En la siguiente tabla se muestra como fue el resultado general con base en unos rangos definidos los cuales se pueden observar en el gráfico de barras de la misma figura. En la tabla es posible observar que los descriptores de color RGB y CMYK tienden a un porcentaje de acierto superior al 70%, con solo tres clases en cada uno de ellos por debajo de este rango. En el caso de TSL, no presenta clases por debajo de este rango. De los demás descriptores SURF y MSER presentan mas del 50% de las clases por debajo del 70% de eficiencia, Sobel y Canny solo presentan dos eficiencias superiores al 90% y HOG como ya se había indicado es el que presenta las mejores eficiencias con solo una clase por debajo del 70%.

Tabla 5. Eficiencias de las clases

	RGB	CMYK	TSL	SOBEL	CANNY	HOG	SURF	MSER
0	89,04	97,44	73,19	79,47	73,75	91,35	53,50	84,66
1	82,13	59,99	70,80	58,15	50,49	67,09	31,19	33,58
2	66,28	67,43	72,74	73,52	74,06	79,71	35,80	40,92
3	56,44	70,81	78,35	73,46	64,73	93,44	55,29	57,54
4	99,96	99,90	98,49	86,60	73,29	95,21	51,39	68,38
5	98,43	98,06	98,46	99,63	98,89	95,11	63,70	60,00
6	60,34	69,49	77,52	50,36	57,93	82,43	51,12	46,77
7	86,99	87,32	86,79	74,42	75,41	87,55	42,84	47,46
8	88,55	93,50	89,51	78,96	62,46	90,75	55,43	65,62
9	98,71	96,53	100,00	96,52	88,06	100,00	54,93	90,29
10	93,60	92,29	87,35	88,08	93,10	91,04	73,11	47,38



Como se pudo observar, en los resultados obtenidos se tienen descriptores que presentan un desempeño muy aceptable, estos descriptores son: RGB, CMYK, TSL y HOG, siendo el ultimo el que presenta la eficiencia más alta. Con base en lo anterior y con el fin de mejorar los resultados obtenidos, se propone una metodología que mediante la mezcla de dos descriptores o mediante la combinación lineal de estos, sea posible incrementar los resultados obtenidos en la identificación de las clases.

Debido a que HOG es el descriptor con la eficiencia más alta, se propone realizar combinaciones lineales entre los descriptores de color con este descriptor y por último una combinación de los descriptores CMYK y TSL, para evaluar nuevamente la eficiencia. Los resultados obtenidos de las combinaciones propuestas, se muestran en la tabla 6.

Tabla 6. Eficiencia y desviación con fusión de datos

CLASE	RGB+HOG	CMYK+HOG	TSL+HOG	CMYK+TSL
0	95,17 ± 6,36 %	94,13 ± 6,13	92,37 ± 6,68	88,74 ± 8,77
1	66,41 ± 12,95 %	69,69 ± 13,05	68,74 ± 12,27	86,25 ± 10,29
2	81,07 ± 12,25 %	82,69 ± 11,09	84,46 ± 13,55	74,93 ± 12,81
3	91,09 ± 8,74 %	91,87 ± 8,41	91,82 ± 8,53	80,34 ± 12,94
4	97,12 ± 4,31 %	96,99 ± 4,69	97,98 ± 3,53	99,77 ± 1,29
5	95,08 ± 5,60 %	94,34 ± 5,93	96,16 ± 5,51	97,53 ± 5,10
6	82,46 ± 12,32 %	82,58 ± 13,22	83,37 ± 13,30	81,47 ± 11,45
7	86,86 ± 10,12 %	85,77 ± 10,05	88,37 ± 9,61	89,32 ± 8,66
8	91,02 ± 8,16 %	91,84 ± 7,62	92,98 ± 6,20	92,45 ± 5,84
9	100 ± 0 %	100 ± 0	100 ± 0	99,62 ± 1,59
10	88,72 ± 8,69 %	90,02 ± 7,64	90,41 ± 7,28	89,04 ± 8,60

En los resultados de la tabla anterior se puede observar que en las combinaciones de los modelos de color con HOG, solo la clase 1 presentó un desempeño del 66,41 ± 12,45 % en la combinación RGB con HOG, las demás clases todas superan el 81,07 ± 12,25 % presentado en la misma combinación. La clase 2 en todas las combinaciones aumento con respecto a los valores de los descriptores de manera independiente presentando un valor mínimo de 74,93 ± 12,81 % en la combinación CMYK con TSL, este mismo valor es el único por debajo del 80 % para esta combinación. La eficiencia de clases de manera general presentó un aumento significativo, sin embargo, en algunos casos como por ejemplo en la clase 10 la eficiencia de los descriptores RGB, CMYK y HOG era superior antes del cruce.

Todas las combinaciones a excepción de CMYK con TSL presentan más de la mitad de las clases con un desempeño superior al 90 %, presentando eficiencias mínimas del 99,62 ± 1,59 % en la clase 9, 94,34 ± 5,93 % en la clase 5 y 96,99 ± 4,69 % en la clase 4.

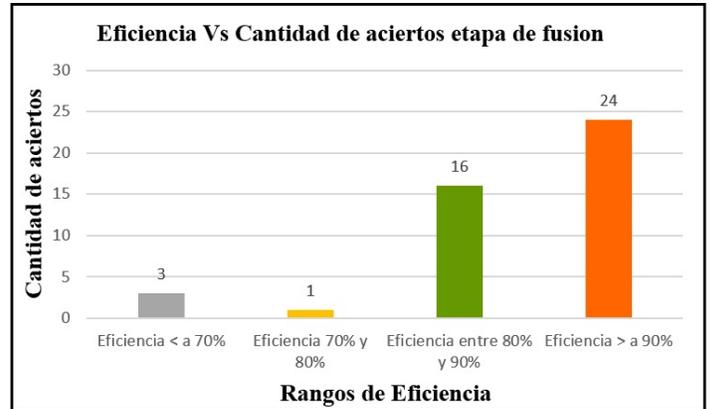
En la siguiente tabla se presentan los valores promedio de eficiencia y desviación de la mezcla de los descriptores, allí se puede observar que sus resultados superan los obtenidos en la tabla 4. Lo anterior indica que la fusión de descriptores brinda mejores resultados obteniendo un promedio mínimo del

Tabla 7. Eficiencia y desviación promedio mezcla de descriptores

Descriptor	Eficiencia Promedio	Desviación estándar promedio
RGB+HOG	88,64 %	8,14 %
CMYK+HOG	89,08 %	7,98 %
TSL+HOG	89,7 %	7,8 %
CMYK+TSL	89,04 %	7,94 %

Tabla 8. Eficiencias fusión de datos sin desviación estándar

CLASES	RGB&HOG	CMYK&HOG	TSL&HOG	CMYK&TSL
0	95,17	94,13	92,37	88,74
1	66,41	69,69	68,74	86,25
2	81,07	82,69	84,46	74,93
3	91,09	91,87	91,82	80,34
4	97,12	96,99	97,98	99,77
5	95,08	94,34	96,16	97,53
6	82,46	82,58	83,37	81,47
7	86,86	85,77	88,37	89,32
8	91,02	91,84	92,98	92,45
9	100,00	100,00	100,00	99,62
10	88,72	90,02	90,41	89,04



88,64 ± 8,14 % y máximo del 89,7 ± 7,8 %

En la siguiente tabla se muestran los resultados generales de la fusión de los datos con base en unos rangos definidos y su respectivo gráfico de barras.

6. CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

Se desarrolló un sistema para el reconocimiento de fallas en un producto terminado, aplicando métodos de aprendizaje supervisado como las SVM obteniendo desempeños promedio superiores al 88,64 ± 8,14 % en todos los métodos analizados.

Se permitió verificar el nivel de descripción de cada metodología propuesta en el trabajo, tales como histogramas de color como RGB, TSL y CMYK, y en descriptores morfológicos como CANNY,

SOBEL, SURF, MSER Y HOG, verificando el desempeño en la etapa de validación con MonteCarlo.

Como se puede observar en la tabla de rendimiento promedio los métodos que presentan mejor rendimiento son HOG con un $88,52 \pm 8,35 \%$, RGB con un $83,68 \% \pm 8,72 \%$, CMYK con un $84,8 \% \pm 8,83 \%$ y TSL con un $84,84 \% \pm 8,93 \%$.

Al realizar la fusión de los métodos se comprueba que el rendimiento mejoró notablemente y las eficiencias en su mayoría superan el 91% . Esto se debe en gran parte a que las fallas presentadas en esta base de datos son ampliamente descritas por espacios de color y sus variaciones en forma son lo suficientemente notables para el clasificador.

Como se logra observar en la tabla 8 los descriptores de color tienen un desempeño más fuerte permitiendo que el resultado de la fusión de los métodos tenga un alto porcentaje de eficiencias.

Se demostró la viabilidad en la implementación de las metodologías mencionadas ya que se utilizan cámaras de bajo costo y condiciones de iluminación no controladas obteniendo resultados que presentan una eficiencia que varía entre $88,64 \pm 8,14 \%$ y el $89,70 \pm 7,8 \%$.

Aunque el modelo TSL fue creado para la detección de piel humana, se evidenció que en este tipo de aplicaciones su rendimiento superó el de los modelos RGB y CMYK presentando una eficiencia promedio del $84,84 \pm 8,93 \%$, por otro lado se verifica que la capa de luminancia L , no aporta información relevante en la identificación de parámetros.

La metodología utilizada para la fusión de datos presentó un alto desempeño a la hora de extraer características de objetos, ya que la combinación de los métodos mejora los resultados de clasificación en el proceso de control de calidad; arrojando que la mejor combinación entre métodos es TSL&HOG presentando una eficiencia del $89,70 \pm 7,8 \%$.

RECOMENDACIONES

El presente trabajo se enfocó a realizar un análisis morfológico de objetos en un producto de línea de producción, teniendo en cuenta que gran cantidad de aplicaciones hacen énfasis en un solo factor de distinción en el objeto, se realizó fusión de datos para optimizar los modelos estudiados e implementados y poder resaltar diferentes patrones como forma, color y textura, para futuros trabajos se recomienda buscar otros métodos de fusión de datos que permitan que la combinación de los descriptores generen los patrones necesarios para aumentar la eficiencia de la clasificación.

7. ESTUDIANTES

- JUAN GUILLERMO AGUILAR GAVIRIA

Cédula 4.517.312 de PEREIRA

Estudiante Programa de Ingeniería Electrónica

Universidad Tecnológica de Pereira.

- VIVIANA ANDREA CASTAÑO SOTO

Cédula 34.066.314 de DOSQUEBRADAS

Estudiante Programa de Ingeniería Electrónica

Universidad Tecnológica de Pereira.

8. DIRECTOR

- ANDRÉS FELIPE CALVO. Ing. M.Sc.

Profesor Transitorio de Ingeniería Electrónica

Grupo de investigación en Ingeniería Electrónica

Universidad Tecnológica de Pereira.

BIBLIOGRAFÍA

- [1] BEJARANO, Arley y CALVO, Àndres. ESCÀNER 3D PARA CONTROL DE CALIDAD DE PIEZAS METALÚRGICAS. Inf. téc., Universidad Tecnològica de Pereira, 2012.
- [2] NEYRA, Ing Fernando Gabino Ramàrez. "Visi3n artificial para detecci3n automàtica de fallas estructurales en botellas de vidrio". Tesis de Maestría, Instituto polotécnico nacional, 2011.
- [3] MANUEL, Prats Montalbán José. Control Estadístico de Procesos mediante Anàlisis Multivariante de Imàgenes. Tesis Doctoral, Universidad Politecnica de Valencia, 2005.
- [4] SPINOLA, C., *et al.* Image processing for surface quality control in stainless steel production lines. En: Imaging Systems and Techniques (IST), 2010 IEEE International Conference on, 2010, págs. 192–197.
- [5] SPINOLA, C.G., *et al.* Real-time image processing for edge inspection and defect detection in stainless steel production lines. En: Imaging Systems and Techniques (IST), 2011 IEEE International Conference on, 2011, págs. 170–175.
- [6] SHARIFZADEH, M., *et al.* Detection of steel defect using the image processing algorithms. En: Multitopic Conference, 2008. INMIC 2008. IEEE International, 2008, págs. 125–127.
- [7] MAHALE, B. y KORDE, S. Rice quality analysis using image processing techniques. En: Convergence of Technology (I2CT), 2014 International Conference for, 2014, págs. 1–5.
- [8] ABDEL-HAKIM, A.E. y FARAG, A.A. CSIFT: A SIFT Descriptor with Color Invariant Characteristics. En: Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, tomo 2, 2006. ISSN 1063-6919, págs. 1978–1983.
- [9] PRATT, William K. Digital Image Processing. JOHN WILEY & SONS, INC, Los Altos, California, 2001.
- [10] STOKMAN, H. y GEVERS, T. Selection and Fusion of Color Models for Image Feature Detection. En: IEEE Transactions on Pattern Analysis and Machine Intelligence, tomo 29, n^o 3, 2007, págs. 371–381. ISSN 0162-8828.

- [11] GHOSH, T., *et al.* A feature extraction scheme from region of interest of wireless capsule endoscopy images for automatic bleeding detection. En: 2014 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2014. ISSN 2162-7843, págs. 000256–000260.
- [12] MO, Shu; CHENG, Shihai y XING, Xiaofen. Hand gesture segmentation based on improved kalman filter and TSL skin color model. En: Multimedia Technology (ICMT), 2011 International Conference on, 2011, págs. 3543–3546.
- [13] RUANGYAM, P. y COVAVISARUCH, N. An efficient region-based skin color model for reliable face localization. En: 2009 24th International Conference Image and Vision Computing New Zealand, 2009. ISSN 2151-2191, págs. 260–265.
- [14] TERRILLON, J. C., *et al.* Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. En: Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on, 2000, págs. 54–61.
- [15] HAO, G.; MIN, L. y FENG, H. Improved Self-Adaptive Edge Detection Method Based on Canny. En: Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2013 5th International Conference on, tomo 2, 2013, págs. 527–530.
- [16] ZHAO, H.; QIN, G. y WANG, X. Improvement of canny algorithm based on pavement edge detection. En: Image and Signal Processing (CISP), 2010 3rd International Congress on, tomo 2, 2010, págs. 964–967.
- [17] GAO, Wenshuo, *et al.* An improved Sobel edge detection. En: Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on, tomo 5, 2010, págs. 67–71.
- [18] CALOT, Enrique. Reconocimiento de patrones en imágenes médicas basado en sistemas inteligentes. Inf. téc., Universidad de Buenos Aires, 2008.
- [19] LAGANIÈRE, Robert. OpenCV2 2 Computer Vision Application Programming Cookbook. Packt Publishing Ltd, Birmingham, 2011.

- [20] Borrego Jiménez María Asunción, Manual UF0242: Calidad en los procesos gráficos, Editorial IC.
- [21] ALYAMMAHI, S., *et al.* A hardware accelerator for real-time extraction of the linear-time MSER algorithm. En: Industrial Electronics Society, IECON 2015 - 41st Annual Conference of the IEEE, 2015, págs. 000065–000069.
- [22] JEON, Byeungwoo y LANDGREBE, D. A. Decision fusion approach for multitemporal classification. En: IEEE Transactions on Geoscience and Remote Sensing, tomo 37, nº 3, 1999, págs. 1227–1233. ISSN 0196-2892.