



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
Posgrado en Ciencia e Ingeniería de la Computación  
Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas  
Señales, imágenes y ambientes virtuales

**Sistema de detección de productos cilíndricos con descriptores de  
imágenes basados en puntos característicos**

TESIS  
QUE PARA OPTAR POR EL GRADO DE  
MAESTRO EN CIENCIAS E INGENIERÍA DE LA COMPUTACIÓN

PRESENTA:  
PEDRO PABLO BECERRIL CALDERÓN

Director de Tesis:  
Dr. Boris Escalante Ramírez  
Facultad de Ingeniería

Ciudad Universitaria, Cd. De México

noviembre 2017



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## **Dedicatoria**

ii

Dedico esta tesis con todo cariño a mi madre, a mi tía-hermana y a mis abuelos, las personas más trabajadoras que conozco y las mejores influencias que me han impulsado para llegar más lejos de lo que esperaba. Este es el resultado del esfuerzo conjunto, de sus consejos, su amor y su apoyo.

## Agradecimientos

iii

Quiero expresar a través de las siguientes líneas mi más sincero agradecimiento a todas las personas e instituciones que formaron parte de la realización de este trabajo de tesis.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la ayuda económica brindada durante mi estancia en el posgrado.

Muy especialmente a mi director de tesis, el Dr. Boris Escalante Ramírez, por el soporte, la orientación y sus invaluable conocimientos que me permitieron completar de manera satisfactoria el proyecto de esta tesis.

Gracias al Dr. Mario Peña, al Maestro Humberto Gómez y los miembros del jurado por las contribuciones al proyecto y por el tiempo invertido en la revisión de mi trabajo.

Agradezco al Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS), al actual coordinador del Posgrado en Ciencia e Ingeniería de la Computación (PCIC), el Dr. Javier Gómez Castellanos, al ex coordinador, el Dr. Jorge Luis Ortega Arjona y en general a todo el personal que conforma el PCIC por estar al pendiente del alumnado.

Gracias especialmente a los profesores del posgrado por su dedicación y por compartir sus conocimientos, especialmente a aquellos que dan el extra dentro y fuera de las aulas sin pedir nada a cambio.

Agradezco a la Universidad Nacional Autónoma de México por permitirme seguir preparándome profesionalmente y por mantener el conocimiento y las oportunidades al alcance de aquellos que tengan la voluntad de superarse.

Finalmente, agradezco a mi familia y amigos de Hiperreality y Realidad Virtual MX por apoyarme en momentos de necesidad y por ayudarme a pensar fuera de la caja.

En las siguientes páginas se describe el desarrollo de un sistema de reconocimiento que hace uso de visión por computadora, técnicas de procesamiento de imágenes y fotografías de anaqueles de supermercado, con el fin de agilizar el proceso de auditoría en los centros comerciales al identificar productos con envases cilíndricos a partir de proyecciones planares de éstos.

En el capítulo 1 se presenta el estado del arte al momento del desarrollo de este proyecto y una comparativa entre el trabajo realizado y el existente en el estado del arte. El capítulo 2 describe de manera breve y sencilla el funcionamiento completo del sistema de reconocimiento. Los capítulos 3, 4 y 5 describen respectivamente los tres sub sistemas principales del sistema de reconocimiento: i) De adquisición de imágenes, ii) De detección y calificación de productos y iii) De reconocimiento por probabilidad. El capítulo 6 muestra los resultados de las pruebas de desempeño a las que fue sometido el algoritmo y finalmente el capítulo 7 plantea el posible trabajo que se podría implementar a futuro para mejorar el trabajo realizado en esta tesis.

## Tabla de Contenidos

v

Capítulo 1: Planteamiento del problema.....	1
1.1 Estado del arte.....	2
1.1.1 Aplicaciones accesibles al público.....	2
1.1.2 Aplicaciones en desarrollo.....	3
1.2 Precedente para el reconocimiento de cajas.....	8
1.3 Comparativa del proyecto elaborado contra el estado del arte.....	9
Capítulo 2: Descripción general del sistema de reconocimiento de productos.....	11
2.1 Ingreso de información.....	12
2.2 Sistema de adquisición de imágenes.....	13
2.3 Sistema de detección y calificación de productos.....	14
2.4 Sistema de reconocimiento por probabilidad.....	15
2.4 Salida del sistema.....	15
Capítulo 3: Sistema de adquisición de imágenes.....	16
3.1 Módulo de adquisición de base de datos.....	17
3.1.1 Descripción de la cámara de línea y lentes.....	18
3.1.2 Descripción de la plataforma giratoria.....	19
3.1.3 Resultados del módulo de adquisición de base de datos.....	20
3.2 Hardware de fotografiado de estantes.....	22
3.2.1 Descripción de la aplicación de fotografiado.....	23
Capítulo 4: Sistema de detección y calificación de productos.....	26
4.1 Separación de repisas y productos.....	26
4.1.1 Segmentación de repisas.....	26

4.1.2 Separación de productos .....	30vi
4.2 Metodología de selección de SURF como parámetro de reconocimiento .....	34
4.3 Calificación de productos .....	35
4.3.1 Prueba de tamaño .....	36
4.3.2 Calificación por puntos SURF .....	37
4.3.3 Calificación por color .....	40
4.3.4 Salida del sistema de calificación .....	45
Capítulo 5: Sistema de reconocimiento por probabilidad.....	48
5.1 Funcionamiento del sistema.....	48
5.2 Salida del sistema de evaluación de probabilidad.....	50
Capítulo 6: Pruebas de desempeño y resultados.....	52
6.1 Observaciones finales .....	58
Capítulo 7: Trabajo a futuro.....	60
Lista de referencias .....	61
Apéndice .....	A

## Lista de tablas

vii

Tabla 1 Productos seleccionados para pruebas de desempeño .....	53
Tabla 2 Tabla de resultados para matriz de confusión de ejecución a 0° .....	54
Tabla 3 Tabla de resultados para matriz de confusión de ejecución a 30° .....	55
Tabla 4 Tabla de resultados para matriz de confusión de ejecución a 60° .....	56
Tabla 5 Tabla de resultados para matriz de confusión de ejecución a 90° .....	57
Tabla 6 Desempeño generalizado de cada ejecución.....	58



Figura 1 Configuración principal de la patente.....	4
Figura 2 Variante de la configuración principal .....	5
Figura 3 Ejecución de Smart Shelfie .....	7
Figura 6 Diagrama de bloques del sistema de adquisición de imágenes .....	16
Figura 7 Funcionamiento del módulo de adquisición de base de datos.....	18
Figura 8 Cámara de línea .....	18
Figura 9 Módulo de escaneo completamente ensamblado y funcional .....	20
Figura 10 Escaneo crudo.....	21
Figura 11 Escaneo con rotación y recorte.....	21
Figura 12 Dispositivo de fotografiado descartado .....	22
Figura 13 Estante experimental .....	23
Figura 14 Corrección de fotografía de estante.....	25
Figura 15 Generación del vector de saturación de puntos SURF .....	27
Figura 16 Proceso de suavizado del vector renglón.....	28
Figura 17 Segmentos de imagen generados por el módulo de segmentación de repisas.....	29
Figura 18 Imagen de repisa tras aplicar apertura por reconstrucción .....	30
Figura 19 Imagen de cierre por reconstrucción .....	30
Figura 20 Marcadores encontrados por máxima regional.....	31
Figura 21 Imágenes binarias generadas por cada kernel y suma de las mismas.....	32
Figura 22 Vector renglón generado por la máscara binaria.....	33
Figura 23 Resultado final de segmentación de productos por repisa.....	33
Figura 24 Imagen de mayor número de coincidencias .....	39
Figura 25 Imagen de match de un producto no concordante con el buscado .....	39
Figura 26 Área relevante para comparación de color .....	41
Figura 27 Área relevante para comparación de color .....	42
Figura 28 Plano polar para representación de gráfica de color.....	43
Figura 29 Imágenes con alta coincidencia de color 78%.....	44
Figura 30 Imagen con menor coincidencia de color (49%).....	44
Figura 31 Ejemplos de imágenes de probabilidad .....	46
Figura 32 Salida gráfica del sistema de calificación.....	47
Figura 33 Interfaz del sistema de evaluación de probabilidad.....	48
Figura 34 Imágenes de probabilidad de productos clase A y B.....	49
Figura 36 Estante experimental en 4 orientaciones .....	52

## Capítulo 1:

### Planteamiento del problema

La auditoría en los centros comerciales es vital para el eficiente reabastecimiento de los productos que éste ofrece. Un buen supermercado debe poder llevar un control de inventario adecuado en su administración de manera tal que le permita discernir si necesita reabastecer sus anaqueles, adquirir más cantidad de algún producto, o bien, verificar que aún cuenta con suficiente producto almacenado en bodega para rellenar sus estantes y poder seguir ofreciéndolo a su clientela.

De acuerdo al artículo *Smart Shelfie*<sup>1</sup>, 31% de los clientes en un supermercado terminan comprando en otro lugar si no encuentran un producto disponible y 26% compran una marca distinta del producto que buscan si no lo encuentran a primera vista. Lo anterior significa que, para un centro comercial, una mala auditoría de sus anaqueles implica pérdidas tanto para el supermercado como para las marcas que lo abastecen.

Actualmente, la auditoría de los anaqueles se realiza a mano por personal del supermercado mediante la lectura de los códigos de barras de cada uno de los productos de un anaquel. El algoritmo presentado en esta tesis propone un método de identificación de productos cilíndricos mediante análisis de imágenes, más específicamente, la comparación de puntos característicos SURF, análisis de color de la imagen y análisis de probabilidad.

---

<sup>1</sup> Kaulgud, P., Prahlad, S., Satapathy, R. (2015). IEEE Region 10 Symposium. *Smart Shelfie - Internet of Shelves: For Higher On-Shelf Availability*. [En línea] <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7166240> [Accedido por última vez en marzo del 2016]

El algoritmo tiene como finalidad ayudar al auditor a determinar la existencia o falta de productos y así agilizar el tiempo que conlleva el proceso de auditoría.

## **1.1 Estado del arte**

En esta sección, se resumen algunas aplicaciones existentes cuya finalidad es el reconocimiento de productos de supermercado, a partir de las cuales se establecieron las bases del trabajo realizado en esta tesis. El contenido se encuentra dividido en: Aplicaciones accesibles al público y Aplicaciones en desarrollo.

### **1.1.1 Aplicaciones accesibles al público**

A continuación, se describen algunas aplicaciones comerciales que realizan reconocimiento de productos a partir de imágenes digitales. Cabe mencionar que, a pesar de ser públicas, las aplicaciones mencionadas en esta sección son de paga por lo que su metodología de análisis y procesamiento no está disponible de manera abierta.

#### **1.1.1.1 Plano Mobile**

*Plano Mobile* es una aplicación móvil de la empresa *Planorama*. Esta empresa ofrece soluciones de administración de inventarios y análisis estadístico mediante el procesamiento de planos digitales de los estantes, más conocidos como planogramas. *Plano Mobile* se encarga de tomar fotografías digitales a través de la cámara de un dispositivo móvil y subirlas a los servidores de *Planorama*, los servidores procesan la imagen y generan estadísticas y reportes de calidad.

*Planorama* tiene convenios con diferentes marcas comerciales por lo que su análisis depende en gran medida de las empresas con las que está asociada.

### **1.1.1.2 Symphony Gold**

Symphony Gold es un servicio que ofrece soluciones de análisis de planogramas a través de dispositivos móviles. No cuenta con una aplicación móvil para el público general, sino que ofrece la creación de una para uso exclusivo la empresa que lo solicite.

Las aplicaciones ofrecidas por Symphony Gold ofrecen un reconocimiento basado en un servicio en la nube para el análisis de planogramas mediante el cual realizan análisis de inventariado, promoción, introducción de productos a inventario y generación de realogramas.

### **1.1.2 Aplicaciones en desarrollo**

En esta sección se describen algunos de los enfoques más recientes que han tratado de dar solución al problema de detección de productos y administración de anaqueles en supermercados mediante técnicas de procesamiento de imágenes y que han sido expuestos en patentes, artículos y conferencias de carácter científico.

#### **1.1.2.1 Planogram extraction based on image processing (Patente 8189855B2)**

Esta patente [2] describe varias técnicas que hacen uso de visión por computadora y análisis de imágenes para resolver diferentes problemáticas que se presentan en la auditoría de anaqueles, por ejemplo, el desabasto, la mala presentación de un producto y en general la auditoría de anaqueles. La patente presenta diferentes configuraciones en las que se implementan dispositivos de captura de imágenes acomodados de distintas maneras a fin de capturar las imágenes de los estantes que posteriormente serán procesadas mediante técnicas de análisis de imágenes y reconocimiento de objetos.

La configuración principal descrita en la patente se muestra en la siguiente figura:

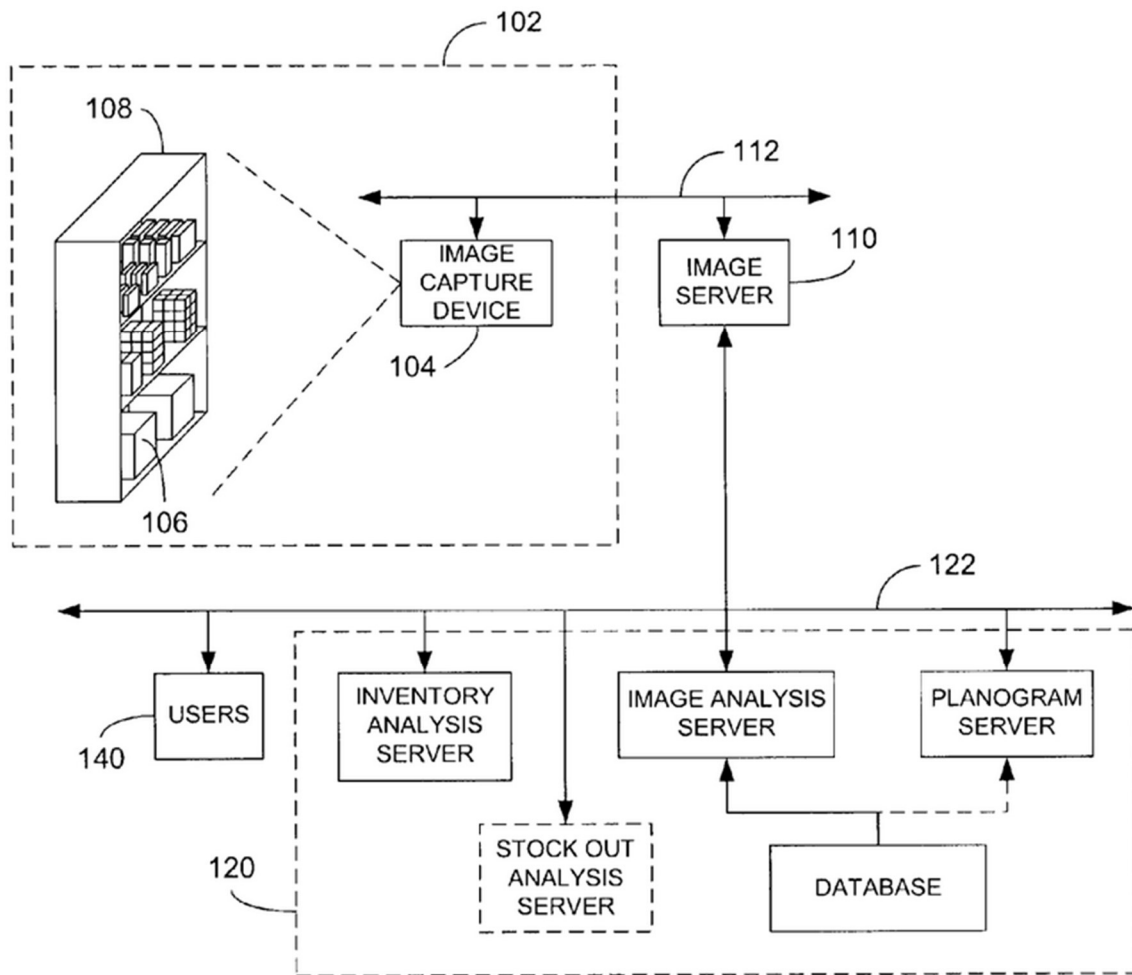


Figura 1 Configuración principal de la patente [2]

En la figura se pueden apreciar el estante (108) con productos (106) el cual es fotografiado por los dispositivos de captura de imágenes (104). Éstos envían las imágenes a un servidor (110) a través de una red (112) para su análisis. El servidor está conectado a través de una segunda red (122) con un sistema de administración de inventario (120) que se encarga de generar un planograma a partir de las imágenes recibidas para finalmente comparar el planograma extraído contra un planograma de referencia. El resultado del análisis es enviado a los auditores (140) para indicarles en qué momento un anaquel requiere atención.

Otras variantes de esta configuración permiten detectar la carencia de productos mediante la detección de patrones adheridos al estante (figura 2 izquierda) o mediante la detección de alguna característica propia del estante como colores, texturas, etc. (figura 2 derecha).

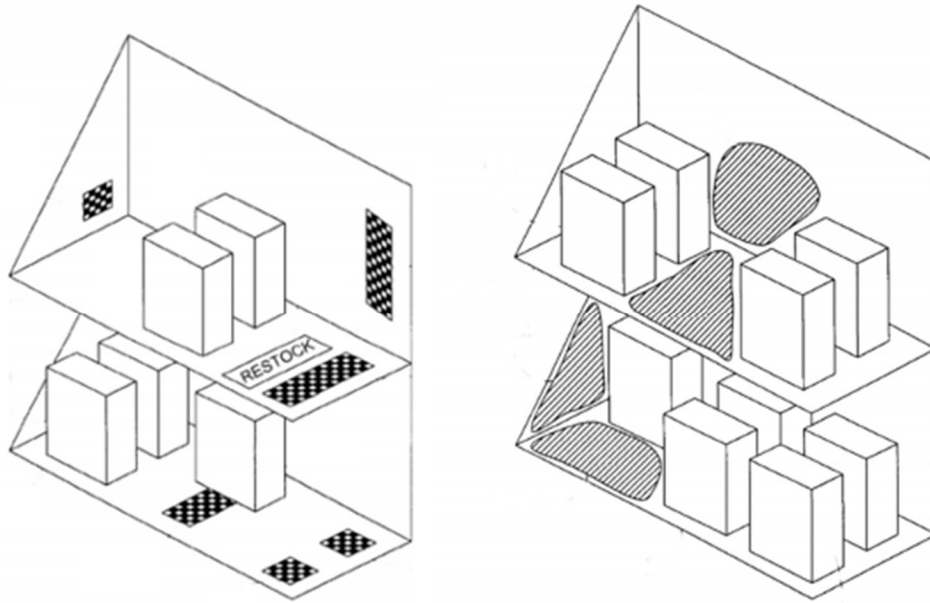


Figura 2 Variante de la configuración principal [2]

### 1.1.2.2 Smart Shelfie – Internet of Shelves

El artículo<sup>2</sup>[1] fue publicado en Mayo del 2015 y describe un sistema de alertas automáticas por internet que hace uso del procesamiento de imágenes para detectar la carencia de algún producto en estantes de supermercado y dar aviso a los trabajadores para

---

<sup>2</sup> Kaulgud, P., Prahlad, S., Satapathy, R. (2015). IEEE Region 10 Symposium. *Smart Shelfie - Internet of Shelves: For Higher On-Shelf Availability*. [En línea] <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7166240> [Accedido por última vez en marzo del 2016]

que éstos atiendan únicamente los estantes que lo requieran, optimizando así el tiempo invertido en la auditoría de anaqueles.

El sistema funciona de la siguiente manera: Cada estante cuenta con una cámara instalada frente a él con la que se realiza la toma eventual de una fotografía (toma una *selfie*, de ahí el nombre *Shelfie*). Esta fotografía es comparada contra la toma del estante en un escenario ideal de dos formas distintas: Mediante procesamiento de manchas de color (blobs) o mediante emparejamiento exhaustivo de modelo (template matching). El primer método toma manchas de color características un producto, por ejemplo, manchas rojas para una Coca-Cola o manchas azules de una Pepsi y basa sus resultados en el color de dichas manchas. El segundo proceso realiza la comparación de un producto contra distintos “modelos”, los cuales consisten en imágenes del producto expuestas a diferentes condiciones de luz y tomadas en diferentes orientaciones y distancias. El método de blobs es muy rápido pero no tiene buen nivel de reconocimiento mientras que el método de template matching tiene mejor reconocimiento pero es mucho más lento. Cabe mencionar que template matching permite reconocer aquellos productos que están mal orientados, es decir, aquellos cuyo logo no está hacia el frente, sin embargo el reconocimiento es por omisión del mismo, es decir, no es marcado como un producto mal orientado, simplemente es omitido. Esto se aprecia en la figura 3 en el estante inferior de la imagen.

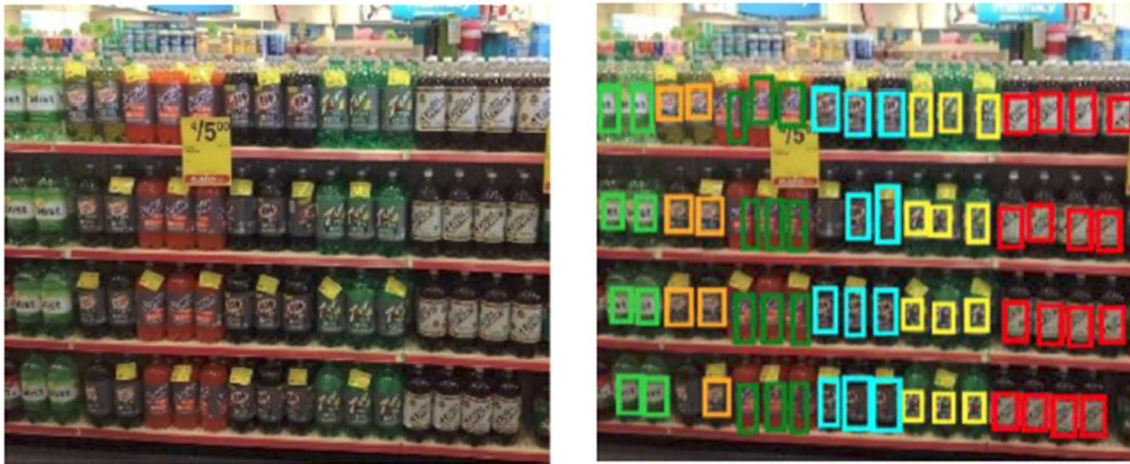


Figura 3 Ejecución de Smart Shelfie [1]

De acuerdo con el artículo, *Smart Shelfie* tiene un tiempo de procesamiento que varía de 2 a 6 segundos por imagen, sin embargo es susceptible a cambios de iluminación o a problemas en las imágenes de entrada como por ejemplo tomas borrosas o de baja resolución.

### 1.1.2.3 Identification for refilling of shelves using image processing

El documento<sup>3</sup>[3] describe un método teórico para la detección de productos faltantes en estantes de supermercado a través de técnicas de visión por computadora, hardware de electrónica básica con una arquitectura maestro-esclavo y procesamiento de imágenes mediante SURF. Su funcionamiento se basa en la obtención de una imagen digital a través de una cámara CMOS TTL JPEG conectada a un microcontrolador Arduino. Este micro controlador y la cámara conforman la sección esclavo de la arquitectura. Un segundo Arduino conforma la parte maestro y se encarga de recibir la

---

<sup>3</sup> *Identification for refilling of shelves using image processing*  
 [En línea] [http://www.superior.edu.pk/ICEET/pdf/research2015/submission\\_69.pdf](http://www.superior.edu.pk/ICEET/pdf/research2015/submission_69.pdf) [Accedido por última vez en marzo del 2016]



imagen desde el primer microcontrolador y enviarla al módulo de procesamiento de imágenes que consiste en una computadora con Matlab. Durante el procesamiento, se compara la imagen recibida contra una imagen del escenario ideal, detectando así los espacios vacíos y por consiguiente la carencia de determinados productos, así como su ubicación.

Finalmente, para saber los nombres de los productos faltantes se toman las ubicaciones donde hubo espacios vacíos y se segmenta la imagen del escenario ideal en dicha sección. Del segmento resultante se calculan los puntos SURF y se comparan contra los puntos SURF de imágenes individuales de cada producto contenidas en una base de datos. De las coincidencias más altas se toma el nombre de la imagen y dicho nombre se considera como el del producto faltante.

Éste método en específico tiene muchas desventajas, la principal de ellas la de ser puramente teórico, los resultados presentados se basan en pruebas con imágenes generadas por computadora que están muy lejanas de un escenario realista, además de que la implementación de esta metodología requiere de mucho hardware, sin embargo, fue de ayuda para la planeación de formas de llevar a cabo la segmentación de los estantes y productos en este trabajo de tesis.

## **1.2 Precedente para el reconocimiento de cajas**

Existe un proyecto precedente al trabajo de esta tesis cuyo objetivo es el reconocimiento de productos rectangulares (cajas) de un estante. El proyecto funciona mediante la comparación de la fotografía de un estante contra la imagen de la clase buscada, es decir, la imagen de la caja del producto buscado. La imagen clase es primero escalada

al mismo tamaño que tendría dentro de la imagen del estante y después es correlacionada en toda la extensión de la imagen del estante. Este proceso arroja diferentes ubicaciones donde la correlación obtuvo mayor coincidencia. De todas las ubicaciones obtenidas se toma la más fuerte y se segmenta la imagen del estante en dicha ubicación. Finalmente, con el segmento obtenido se repite el proceso de correlación sobre la imagen del estante para obtener todas las ubicaciones donde se tiene alto nivel de coincidencia, las cuales son consideradas como las ubicaciones finales de los productos buscados.

Este proyecto trabaja únicamente mediante correlación y con imágenes de intensidad, es decir, en escala de grises.

### **1.3 Comparativa del proyecto elaborado contra el estado del arte**

Como se puede apreciar en la sección 1.1, muchas de las aplicaciones descritas basan sus resultados en los siguientes puntos:

- i. Comparación del estado de un anaquel contra el escenario ideal.

La inclusión de un escenario ideal implica la captura de dicho escenario. Esto puede representar una desventaja debido a que se requiere personal que haga dichas capturas y contradice el objetivo de facilitar el trabajo de los auditores.

- ii. No se hace análisis exhaustivo de los canales de color de la imagen.

En pocos artículos se menciona un análisis de color que vaya más allá de un análisis de manchas por lo que el análisis de color es un área aprovechable que puede mejorar el reconocimiento.

- iii. Análisis limitado a una sola metodología.

En los artículos se menciona la detección mediante el uso de puntos SURF, patrones, marcadores o correlación y sus resultados se basan únicamente en el respectivo método seleccionado, por lo que la inclusión de uno o más niveles de análisis podría ayudar a refinar los resultados.

El trabajo descrito en esta tesis describe un sistema de captura y reconocimiento de productos cilíndricos en el cual se contempla un análisis robusto integrado por 3 etapas de calificación que permiten obtener resultados mucho más exactos: una etapa de calificación basada en puntos característicos SURF, una etapa de calificación basada en las propiedades de color de la imagen y un análisis de probabilidad basado en las primeras dos calificaciones, todo lo anterior sin necesidad de tener un escenario ideal del estante analizado.

Cabe mencionar además que este proyecto surge ante la necesidad de una aplicación que permitiera facilitar el trabajo a elementos de una empresa de auditoría.

## Capítulo 2:

### Descripción general del sistema de reconocimiento de productos

El sistema de reconocimiento de productos es un programa elaborado en Matlab que hace uso de sus herramientas gráficas y librerías de procesamiento de imágenes para ejecutar un análisis sobre una imagen de un estante buscando uno o más productos.

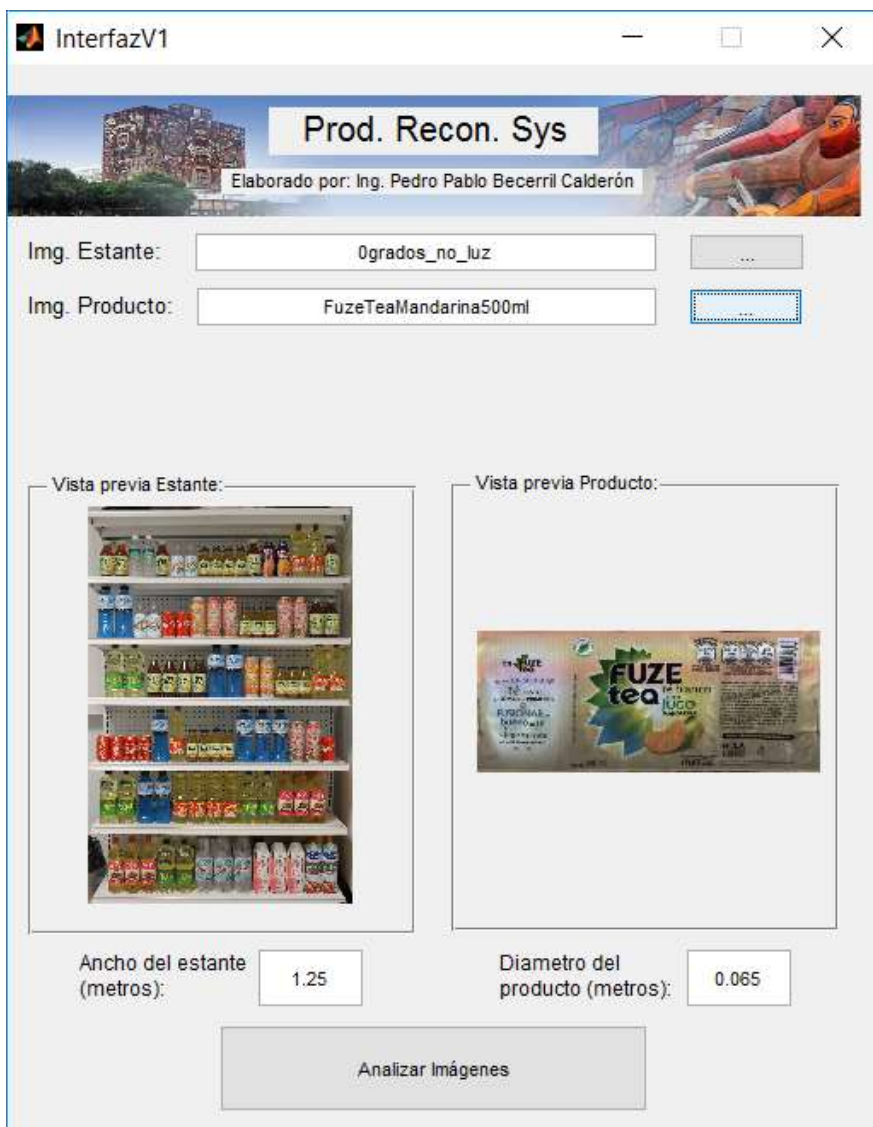


Figura 4 Interfaz principal del sistema de calificación de productos

El siguiente diagrama de bloques ilustra cada una de las etapas y sub sistemas que conforman la ejecución del sistema reconocimiento de productos cilíndricos. Cada sistema se explica más ampliamente en los capítulos subsecuentes, éste capítulo se limita a describir a grandes rasgos el funcionamiento de cada sistema.



*Figura 5 Diagrama de bloques del sistema de calificación*

## 2.1 Ingreso de información

Durante esta etapa se carga la información necesaria para el sistema. La información requerida es la siguiente:

- Imagen del estante
- Imagen del producto buscado
- Medida del largo del estante en metros
- Medida del diámetro del producto en metros.

La interfaz gráfica está diseñada para facilitar el ingreso de la información solicitada y además avisará al usuario en caso de que se omita algún dato o que la información sea incorrecta, sin embargo, es importante cuidar que las imágenes de entrada cumplan con los siguientes puntos:

**Imagen del estante:** Los bordes del estante en la fotografía deben estar lo mejor alineados posible con respecto a los bordes de la imagen. Esto se debe a que, durante el

proceso de separación de productos del sistema de detección, se realiza un proceso de descarte por tamaño, dicho proceso requiere de la transformación del diámetro del producto en metros a su equivalente en pixeles y se calcula bajo la premisa de que el ancho del estante es igual al ancho de la imagen (ver capítulo 4.1).

**Imagen de producto:** Tal como se mencionó en la descripción del “módulo de adquisición de base de datos”, las proyecciones planares que representan a los productos de la base de datos, deben estar correctamente orientadas y recortadas al área de información relevante. De estos dos aspectos, la rotación puede llegar a omitirse ya que el algoritmo cuenta con un proceso automático de detección de orientación y realiza la corrección de manera automática. El recorte del área relevante debe tomarse en cuenta ya que puede llegar a alterar los resultados de la identificación por color (ver sección Análisis de coincidencias).

## **2.2 Sistema de adquisición de imágenes**

El sistema de adquisición de imágenes se encarga de realizar la captura de las imágenes que conformarán la base de datos del sistema de reconocimiento y la captura de las imágenes de los estantes por auditar. Las imágenes de la base de datos son imágenes de cada producto que se puede llegar a reconocer, es decir, cada una de las clases reconocibles, y se obtienen a través de un prototipo electro-mecánico y una cámara de línea. Las imágenes de cada estante son obtenidas a partir de una cámara de fotografía profesional y se someten a un pre procesamiento a fin de ajustar sus bordes a los bordes del estante. Se puede revisar información más detallada de éste sistema en el capítulo 3 de esta tesis.

### **2.3 Sistema de detección y calificación de productos**

Para poder realizar la identificación de cada producto que aparece en la imagen de un estante se deben detectar individualmente cada uno de los productos que hay sobre dicha imagen y calificarlos mediante una métrica que defina la similitud con alguna clase. El sistema de detección de productos se encarga de realizar dichos procesos: segmenta y califica cada instancia detectada sobre la imagen de un estante.

La segmentación de cada producto se realiza en dos fases, primero se realiza una segmentación de repisas mediante la detección de zonas de concentración promedio de puntos SURF. Una vez segmentada cada repisa la segunda fase se encarga de segmentar cada producto habido en cada repisa mediante el cálculo de una máscara binaria. La máscara binaria define en qué posiciones sobre la repisa se encuentra cada instancia de un producto y es calculada de manera individual para cada repisa mediante un proceso de apertura y cierre por reconstrucción, el cual permite detectar marcadores (blobs) de objetos en primer plano de una imagen.

Una vez que se tienen todas las instancias de los productos de cada repisa del estante se procede a calificar la similitud de cada instancia con las imágenes de la base de datos, es decir, el parecido de cada instancia contra cada clase. Para determinar el parecido se realizan 3 pruebas: La primera prueba verifica que el diámetro de la instancia segmentada sea similar al diámetro de la clase contra la que se está comparando; si el tamaño es similar se procede a la segunda y tercera prueba: un reconocimiento por puntos SURF y por canales de color.

Al final del proceso cada instancia de producto tendrá asignado un par de valores numéricos que definen su similitud con respecto a una clase: una calificación de similitud por puntos SURF y otra por similitud de paleta de color. Con estos dos valores numéricos se calcula una tercera métrica de similitud llamada probabilidad conjunta. Al final de la ejecución de este sistema, cada instancia tendrá múltiples de valores, uno por cada clase contra la que fue comparada.

#### **2.4 Sistema de reconocimiento por probabilidad**

El valor de probabilidad conjunta define el porcentaje de similitud que tiene una instancia determinada para una clase determinada. Dado que cada instancia llega a tener múltiples valores de probabilidad conjunta (uno por cada clase en la base de datos), el sistema de reconocimiento por probabilidad determina, con base en dichos valores, a qué clase se parece más una instancia. El sistema se ejecuta una vez que todas las instancias de un estante han sido comparadas contra todas las clases de la base de datos, de ésta manera cada instancia terminará perteneciendo exclusivamente a una sola clase, eliminando así ambigüedades que pudiera haber entre productos muy similares.

#### **2.4 Salida del sistema**

Como salida final, el sistema genera un archivo de texto en el que se describe las ubicaciones de los productos coincidentes por cada clase. Adicionalmente se generan archivos de imágenes con las mismas ubicaciones contenidas en el archivo de texto, a fin de tener una forma de visualización gráfica y no solamente las coordenadas numéricas.



### Capítulo 3:

#### Sistema de adquisición de imágenes.

Para poder realizar el reconocimiento de productos se requiere de una base de datos que contenga las clases buscadas, es decir, las imágenes de las etiquetas<sup>4</sup> de cada uno de los productos buscados, y fotografías de muy buena calidad de los estantes auditados.

El sistema de adquisición de imágenes se encarga de las dos tareas descritas en el párrafo anterior: Realiza la adquisición de las imágenes que conforman la base de datos del sistema de reconocimiento y la adquisición de las fotografías del estante. La primera tarea corre a cargo *del módulo de adquisición de base de datos* y la segunda corresponde al *hardware de fotografiado de estantes*.

Visto como un diagrama de bloques, el sistema de adquisición de imágenes puede interpretarse como se muestra en la figura 6.

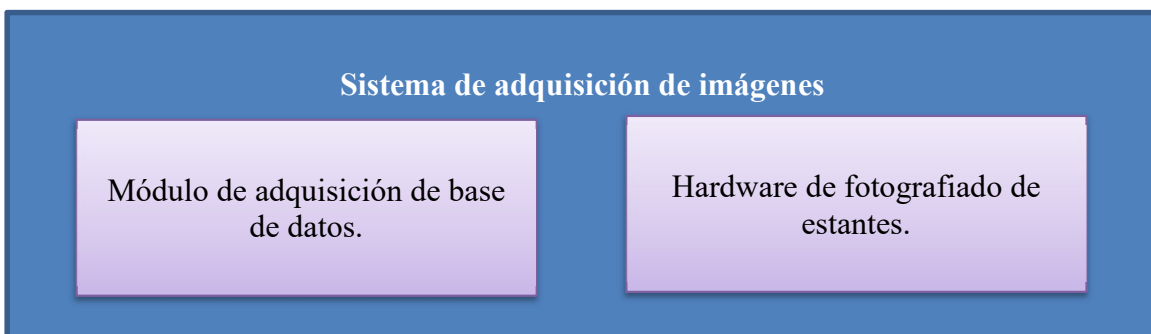


Figura 6 Diagrama de bloques del sistema de adquisición de imágenes

---

<sup>4</sup> Entiéndase en este contexto “etiqueta” como el logotipo gráfico o la presentación que ofrece el producto al consumidor.

### 3.1 Módulo de adquisición de base de datos

Debido a que los productos cilíndricos pueden presentar distintas orientaciones, las tomas contenidas en la base de datos deben ser imágenes que cubran todos los ángulos del producto y que a su vez conserven una buena calidad y eviten factores como ruido, desenfoque o deformaciones. Además, debido a que las presentaciones de los productos cambian a un ritmo acelerado, ya sea por temporada, festividad o debido a promociones, es necesario que las imágenes puedan actualizarse fácilmente.

En un primer enfoque se optó por desprender las etiquetas de los productos y escanearlas tal como se haría con un documento, sin embargo, esto no era una solución universal considerando que algunos productos tenían etiquetas de aluminio altamente propensas al ruido o simplemente porque la etiqueta estaba impresa en el mismo envase.

El módulo de adquisición de base de datos solventa este problema al realizar el escaneo a través de una cámara de línea, un lente de 50mm y una base giratoria controlada electrónicamente. El funcionamiento es el siguiente y se ilustra en la figura 7:

- 1) Se coloca el producto sobre la base giratoria procurando alinear el centro del producto con el eje de giro de la base y se activa la plataforma para hacer girar el producto.
- 2) Con ayuda de un tripié, se posiciona la cámara a cierta distancia y altura con respecto a la plataforma, de manera que el campo visual de ésta cubra todo el producto a escanear, y se enfoca el lente hasta lograr la mejor nitidez posible.

- 3) Se selecciona en la plataforma una velocidad de giro y se oprime la señal de adquisición a fin de lograr la captura de varias líneas que en conjunto formarán una imagen plana de 360 grados del producto.

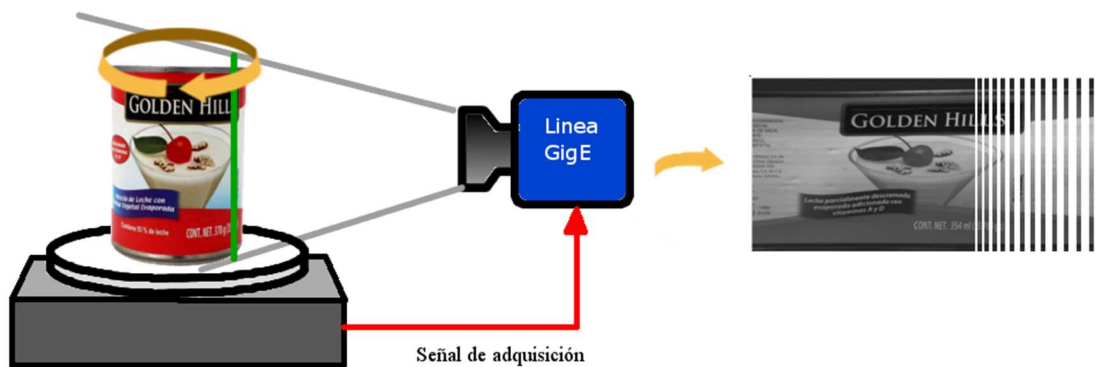


Figura 7 Funcionamiento del módulo de adquisición de base de datos

Mediante esta forma de obtención de imágenes se generan tomas de planares, sin deformaciones, de 360 grados, fácilmente actualizables de cualquier producto cilíndrico. La iluminación también es cuidada a través del uso de un estudio de fotografiado portátil.

### 3.1.1 Descripción de la cámara de línea y lentes

La cámara de línea es un modelo *Linea GigE* de Teledyne. En conjunto con el software *Sapera Cam Expert*, permite capturar imágenes de alta resolución a color. En las primeras etapas de desarrollo se contaba con una cámara de las mismas características pero con salida monocromática, ésta permitía hacer reconocimiento por puntos característicos, sin embargo más adelante se determinó que el algoritmo debía incluir un módulo de reconocimiento por color para hacerlo más robusto, por lo cual fue necesaria la versión de la cámara con salida a color.

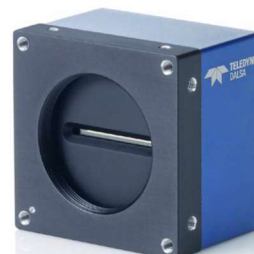


Figura 8 Cámara de línea

Los lentes de prueba para la cámara se adquirieron por separado en 3 diferentes medidas de distancia focal: 28, 32 y 50mm. Para la selección de lentes útiles, se realizaron pruebas basadas la distancia focal y la altura del producto escaneado, tomando como premisa que la distancia focal de un lente fuera suficiente para abarcar la totalidad de la altura del producto y permitiera además enfocar nítidamente a éste. Se concluyó que el lente de 50mm es ideal para el escaneo de todos los productos, ya que con solo alejar o acercar la cámara y enfocar de acuerdo a la distancia, se podía lograr una toma completa y nítida de todos los productos, los cuales variaban en alturas desde los 8cm y hasta los 35cm.

### **3.1.2 Descripción de la plataforma giratoria**

La plataforma giratoria es un prototipo con las siguientes características:

- Base giratoria circular de 30 cm de diámetro con superficie antiderrapante y marcas radiales de referencia para alineación del producto.
- Torque alto para soportar productos de 8 a 10kg de peso.
- Velocidad de giro gradual controlable.
- Salida de señales lógicas para la activación de la cámara mediante electrónica (señal de adquisición).

Además de la plataforma se adquirió un estudio de fotografiado portátil, el cual permitió obtener escaneos de productos con un buen nivel de iluminación y con poca cantidad de ruido causado por brillos o sombras en la imagen. La figura 9 muestra el módulo de escaneo completamente ensamblado y funcional con los componentes anteriormente descritos.



*Figura 9 Módulo de escaneo completamente ensamblado y funcional*

### **3.1.3 Resultados del módulo de adquisición de base de datos**

Los resultados de este módulo consisten en colecciones de imágenes correspondientes a las proyecciones planares de cada uno de los productos cilíndricos. Éstos conjuntos de imágenes conforman la base de datos del sistema.

Cabe mencionar que los escaneos por default tienen una orientación (rotación) vertical e información irrelevante para el reconocimiento, como el fondo o el contenido de las botellas, por lo que se deben rotar y recortar manualmente al área de interés (la etiqueta)

mediante un programa de edición de imágenes para poder utilizarlas en el sistema de reconocimiento. La figura 10 ejemplifica el resultado del escaneo recién capturado y la figura 11 el escaneo con reorientación y recorte del área relevante.

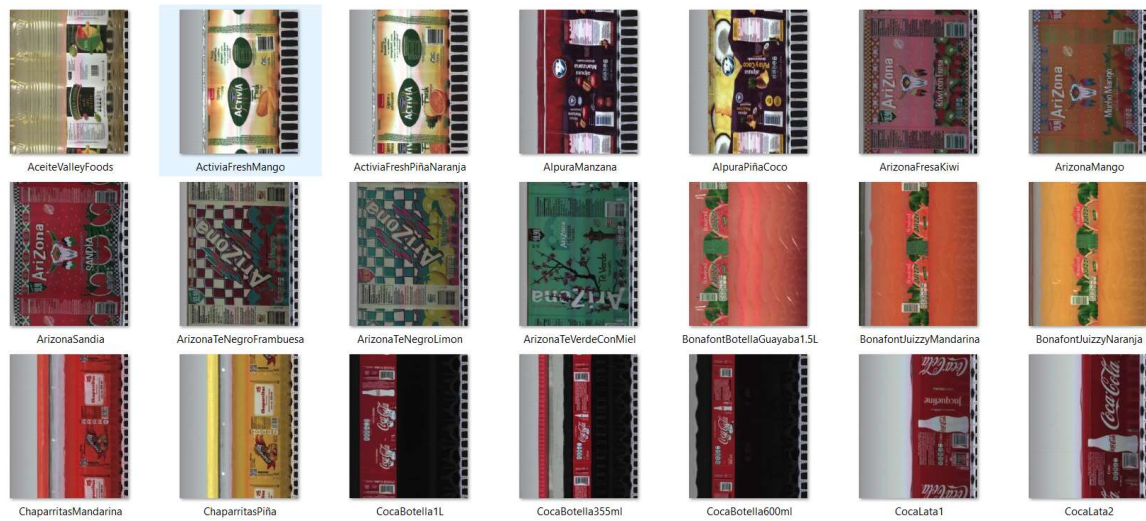


Figura 10 Escaneo crudo

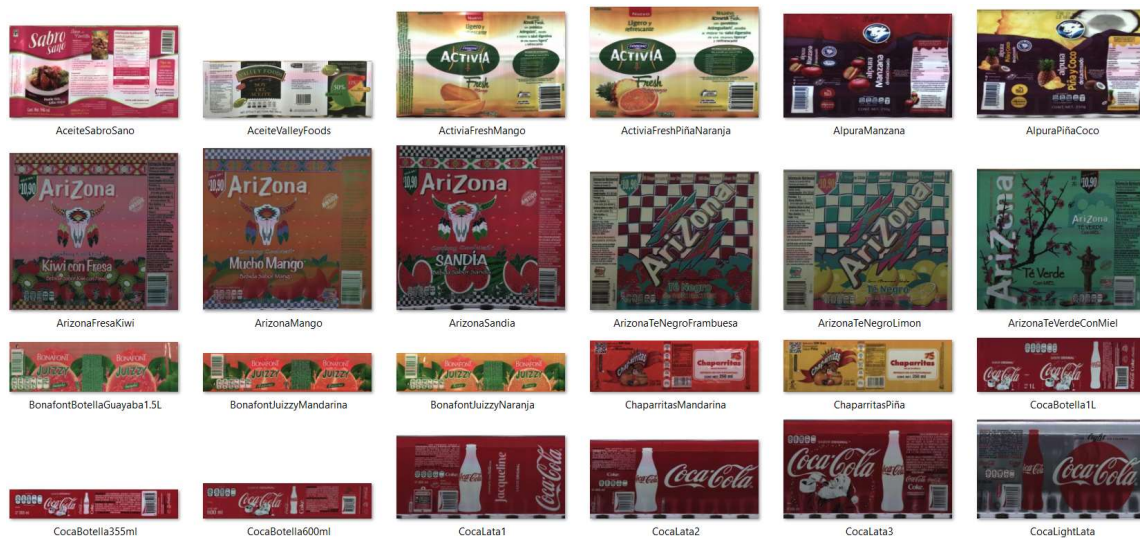


Figura 11 Escaneo con rotación y recorte

### 3.2 Hardware de fotografiado de estantes

Como parte de los requerimientos del proyecto por parte de la empresa auditora, el hardware de fotografiado de estantes debía consistir en un dispositivo portátil que el personal pudiera llevar consigo y que fuera manipulable in situ. Bajo estos criterios se evaluaron dos opciones: un teléfono celular marca Samsung, modelo GT-S6810L y una cámara de fotografía profesional marca Canon, modelo EOS Rebel T3.



*Figura 12*

*Dispositivo de  
fotografiado*

Para determinar la usabilidad de los dispositivos se realizaron pruebas experimentales en las que se comprobó el desempeño del algoritmo con las imágenes de ambas cámaras. A pesar de que el celular era capaz de tomar fotografías de alta resolución<sup>5</sup>, los resultados de los experimentos determinaron que las imágenes de la cámara Canon permitían al algoritmo tener un mejor desempeño en cuanto a la correcta segmentación de productos y la generación de una cantidad más alta de puntos característicos debido a su resolución<sup>6</sup>. Con este dispositivo se tomaron las fotografías del estante experimental, las cuales fueron utilizadas en la mayoría de las pruebas de desempeño.

Las imágenes capturadas por la cámara Canon son imágenes de un anaquel experimental, el cual consiste en un estante de supermercado de 2 metros de alto por 1.22 metros de ancho, armado en el sitio de pruebas en un entorno controlado, con productos de

---

<sup>5</sup> 2560 x 1920 pixeles, más alta que el estándar HD: 1920x1080

<sup>6</sup> Resolución de cámara Canon: 4272 x 2848 pixeles

prueba adquiridos para el proyecto. Un ejemplo fotografía del anaquel experimental capturado con este dispositivo se puede apreciar en la siguiente figura.



*Figura 13 Estante experimental*

### **3.2.1 Descripción de la aplicación de fotografiado**

Como se verá más a detalle en la sección de “Funcionamiento del sistema de reconocimiento”, el algoritmo requiere que la imagen de entrada, correspondiente al estante, sea una toma lo más perpendicular posible y lo mejor alineada posible en sus bordes. Debido a que puede haber limitantes físicas que impidan la adquisición de una imagen correctamente alineada, se desarrolló una aplicación móvil para dispositivos con sistema operativo Android que permite corregir las tomas mediante una transformación afin.



La aplicación corrige dos aspectos de la fotografía: el área correspondiente al estante y el ángulo desde el cual se realizó la toma. Estas dos correcciones se realizan a través de la selección, en pantalla del dispositivo, de las 4 esquinas del estante. Con los 4 puntos seleccionados se lleva a cabo un proceso de corrección de perspectiva que genera una imagen que simula una toma perpendicular al estante.

A continuación, se muestra un ejemplo de cómo trabaja la aplicación. La imagen de la izquierda en la figura 14 muestra como la toma inicial se extiende más allá del estante de interés, el cual se encuentra al centro de la imagen. Mediante la selección de las esquinas del estante se elimina la información innecesaria (área roja) y se realiza una transformación afin sobre la imagen. El resultado es una proyección perpendicular al estante de interés (figura 14 derecha).



*Figura 14 Corrección de fotografía de estante*

## **Capítulo 4:**

### **Sistema de detección y calificación de productos.**

El sistema de detección y calificación de productos lleva a cabo dos tareas primordiales del sistema de reconocimiento: la segmentación automática de los productos visibles sobre la imagen del estante y la asignación de una calificación de similitud a cada uno de ellos de acuerdo a su semejanza con respecto a una clase.

#### **4.1 Separación de repisas y productos**

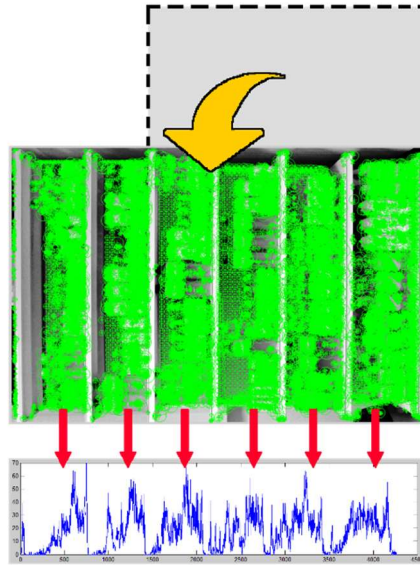
La segmentación automática de productos se realiza en dos etapas, primero se segmentan las áreas de la imagen del estante que corresponden a las repisas y finalmente se segmentan los productos sobre cada una de las repisas segmentadas.

##### **4.1.1 Segmentación de repisas**

Durante este paso, se toma la imagen del estante y se segmentan las repisas que hay en ella a través de la detección del área de saturación de puntos característicos. Éste proceso consiste en obtener todos los puntos SURF que hay sobre la imagen del estante y detectar las áreas de mayor concentración de puntos. Dichas áreas representan la ubicación de las repisas.

El primer paso del proceso consiste en cargar la imagen y convertirla al espacio de color HSV. De la imagen resultante se separa el canal de intensidad (Value) y se obtienen los puntos SURF sobre la imagen. Dado que la imagen de intensidad es una matriz bidimensional, se puede calcular fácilmente la saturación de puntos mediante conteo de puntos característicos por renglón de la matriz. Visto de otra manera, si rotamos la imagen del estante y hacemos que todos los puntos “caigan” sobre el costado, podemos obtener un

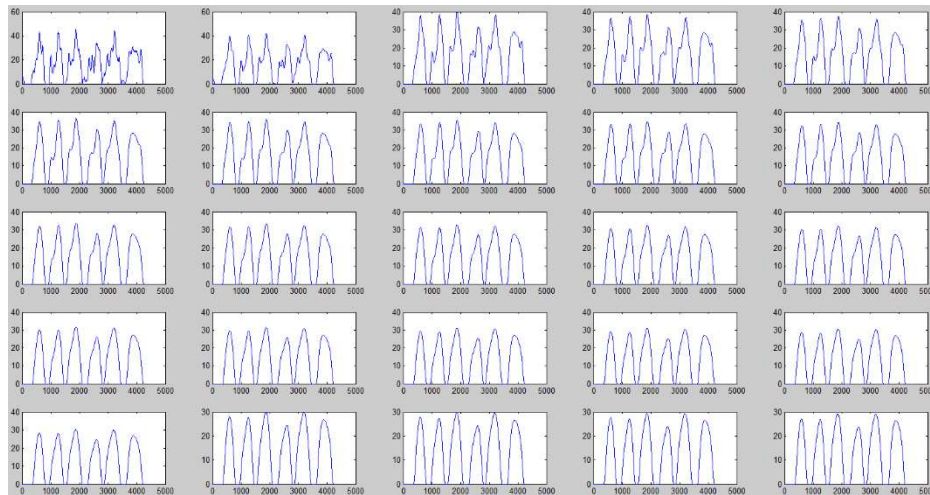
vector renglón que representa la saturación de puntos en cada renglón de la imagen de intensidad (ver figura15).



*Figura 15 Generación del vector de saturación de puntos SURF*

Las crestas del vector renglón obtenido en el paso anterior nos permiten definir de qué punto a qué punto se extiende una repisa, sin embargo, para acotar de mejor manera la extensión de cada región, el vector se somete a un proceso de suavizado.

El proceso de suavizado se realiza en varias iteraciones, cada una de las cuales descarta valores menores al 10% del pico más alto para eliminar outliers y se filtra el vector con un kernel promediador de tamaño igual al 1.6% del total de renglones de la imagen, es decir, 1.6% de la altura de la imagen del estante. El resultado del suavizado se muestra en la siguiente figura.



*Figura 16 Proceso de suavizado del vector renglón*

Como último paso, se recortan de la imagen del estante los segmentos que corresponden a cada repisa tomando como referencia los puntos de inicio y termino de cada cresta, los cuales representan los renglones de la matriz correspondiente a la imagen de intensidad. El resultado final del proceso de segmentación se muestra en la figura 17.

Este proceso de segmentación fue puesto a prueba no solamente con las imágenes del estante experimental y los productos en distintas orientaciones, sino también sobre imágenes de estantes de un anaquel de supermercado real. En ambos casos los resultados fueron positivos, se segmentaron correctamente cada una de las repisas en todas las fotografías.



Figura 17 Segmentos de imagen generados por el módulo de segmentación de repisas

### 4.1.2 Separación de productos

Una vez que se cuenta con imágenes de las repisas separadas se procede a segmentar cada producto que se encuentre en ellas. El proceso que permite llevar a cabo la separación de productos es la creación de una máscara binaria (figura 21) que define las zonas de aparición de un producto.

La creación de la máscara binaria depende de una función llamada *ForegroundMask* que permite encontrar los elementos en primer plano de las imágenes de las repisas. Los elementos en primer plano son aquellos que visiblemente se pueden apreciar por encima del fondo de la imagen y suelen identificarse por ser regiones de pixeles conexos, también conocidas como marcadores. Los marcadores en este caso representan a cada uno de los productos de la repisa.

La detección de marcadores de *ForegroundMask*, utiliza dos procesos sobre la imagen de intensidad, primero, apertura por reconstrucción para crear regiones de áreas conexas (figura 18) y segundo, cierre por reconstrucción para eliminar impurezas de dichas áreas (figura 19).

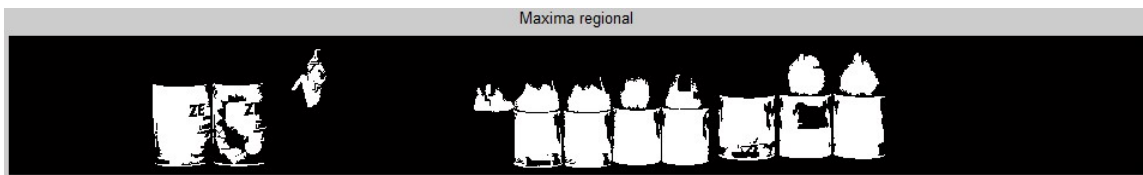


Figura 18 Imagen de repisa tras aplicar apertura por reconstrucción



Figura 19 Imagen de cierre por reconstrucción

Estos dos procesos dan como resultado una imagen de intensidad con diversos marcadores áreas determinadas, sin embargo, aún hay una gran variedad de valores sobre la imagen como para definir correctamente qué marcadores corresponden a cada producto. Para solucionar esto, en un último proceso se obtiene la máxima regional de la imagen de marcadores para obtener una imagen binaria como la que se muestra a continuación.



*Figura 20 Marcadores encontrados por máxima regional*

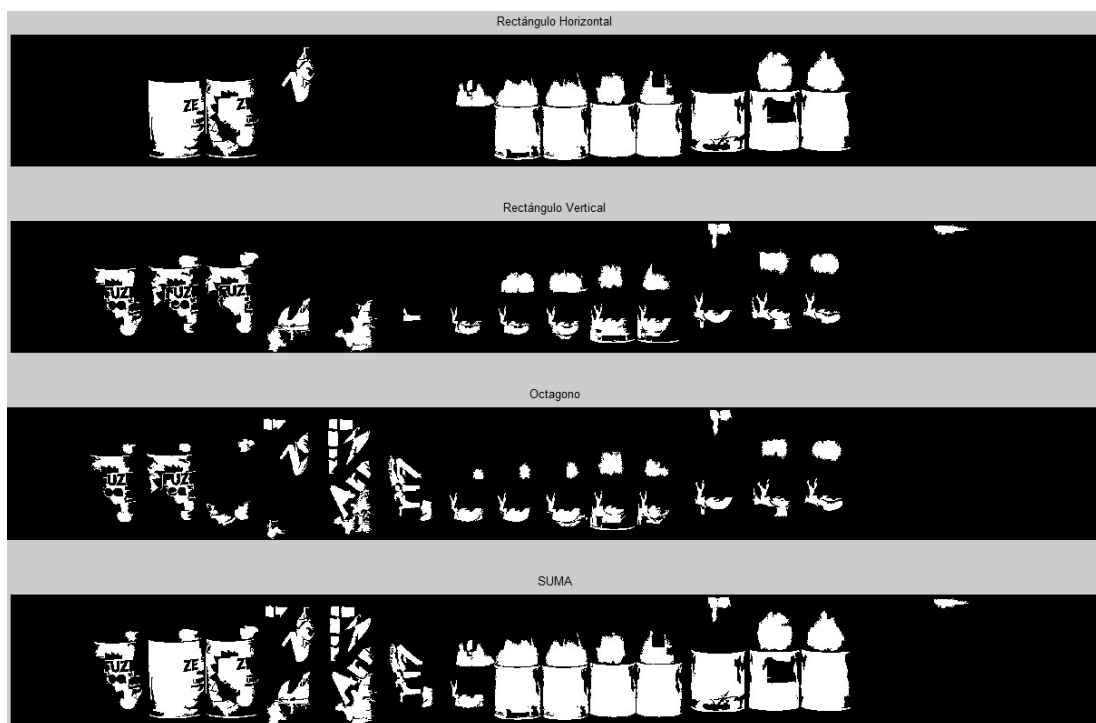
Como se puede apreciar en la figura 20, no todos los productos fueron correctamente localizados; esto se debe a que los procesos de apertura y cierre por reconstrucción realizan un proceso de erosión y dilatación respectivamente. Éstos dos últimos procesos requieren de un kernel que tiene una forma específica (cuadrado, rectangular, circular, etc). Dependiendo de la forma del kernel la erosión y dilatación arrojan resultados distintos y en consecuencia alteran el resultado de la apertura y cierre por reconstrucción. Por esta razón, el algoritmo ejecuta 3 veces la función *ForegroundMask* pero con 3 kernels distintos: un rectángulo horizontal de 60 x 10 píxeles, un rectángulo vertical de 10 x 40 píxeles y un octágono de 9 píxeles de radio.

Las formas de los kernels fueron seleccionadas con base en el análisis de los resultados de diferentes experimentos en los que se filtraron las imágenes de las repisas con distintas formas de kernel. Los resultados mostraron que cada forma genera diferentes blobs al obtener la máxima regional. La forma rectangular vertical genera blobs correspondientes a la forma cilíndrica de la mayoría de los productos, la forma rectangular



horizontal resalta las áreas de productos con etiquetas pequeñas, las cuales son más amplias a lo largo que a lo alto y finalmente la forma octagonal resalta el cuerpo de los envases, especialmente en el área del cuello y base de las botellas.

El resultado de ejecutar 3 veces la función *ForegroundMask* con las 3 formas de kernel son 3 imágenes binarias que, al sumarlas mediante una operación OR lógica, generan una cuarta imagen con una mayor cantidad de marcadores representativos de los productos de la repisa.



*Figura 21 Imágenes binarias generadas por cada kernel y suma de las mismas*

Una vez generada la imagen de marcadores, se realiza una suma por columna de los píxeles de la imagen binaria (figura 21) para generar un vector renglón cuyas crestas indican posición de inicio y fin de un producto. A partir de las crestas obtenidas se puede segmentar el área correspondiente a cada envase.

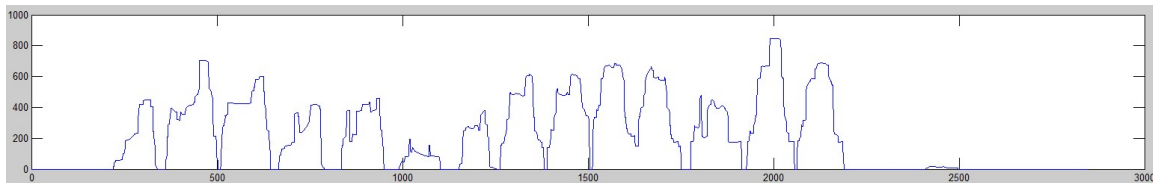


Figura 22 Vector renglón generado por la máscara binaria

El resultado final de la segmentación es un conjunto de segmentos de imagen del estante que corresponden a cada producto de cada repisa. Para ilustrar el resultado, se muestran señalados a continuación los productos encontrados por cada repisa.



Figura 23 Resultado final de segmentación de productos por repisa

## **4.2 Metodología de selección de SURF como parámetro de reconocimiento**

Una de las partes más importantes de este trabajo de tesis es la implementación de SURF como evaluador para el reconocimiento de productos cilíndricos, por lo que antes de dar paso a la descripción del funcionamiento del sistema de calificación, se describe brevemente las bases que dieron pie a la selección de SURF como medida de reconocimiento.

En trabajos similares a éste, el parámetro de evaluación se basa en la mayoría de los casos en la comparación del anaquel en un escenario ideal contra el anaquel en un instante de tiempo. Dicha comparación se lleva a cabo mediante correlación, sin embargo, para que la correlación tenga un buen resultado se necesita que las imágenes a comparar sean coincidentes en escala y ubicación, por ejemplo, en el trabajo precedente que se describe en el estado del arte, la correlación se realiza una vez que la imagen del producto buscado ha sido reducida (escalada) al mismo tamaño que presenta el producto en la imagen del anaquel. Además de esto, los productos tenían la propiedad de ser cajas rectangulares que difícilmente variarían en rotación o en deformación.

Experimentos elaborados durante las primeras etapas del trabajo descrito en esta tesis demostraron que la correlación en el caso de productos cilíndricos no es viable debido a que el área detectable por producto es mucho más pequeña y a que cada producto es propenso a una gran cantidad de rotaciones y deformaciones por la forma cilíndrica del envase. Para sustituir a la correlación, se decidió llevar a cabo la comparación mediante puntos característicos.

La comparación de puntos característicos se realiza de la siguiente manera: Se segmenta cada producto del estante (instancias) y se compara contra el producto buscado (clase) mediante el match de sus respectivas nubes de puntos característicos. Para que la comparación se más precisa se aproxima el tamaño de la imagen de la clase al tamaño de la instancia.

Se probaron mediante experimentación 3 diferentes algoritmos de detección de puntos característicos: Harris, SURF y BRISK. Para seleccionar el mejor candidato se tomaron en cuenta 3 factores importantes:

- Invariación a la rotación: Para solventar el hecho de que los productos pueden presentar diferentes orientaciones.
- Invariación a la escala: Para solventar el hecho de que el tamaño de la imagen de clase e instancias llegan a variar levemente.
- Mejor desempeño: Para tener un mejor margen sobre el cual basar la calificación numérica asignada a cada instancia.

El detector de Harris es invariante a la rotación, pero no lo es a la escala por lo que se descartó. SURF y BRISK por otro lado, son invariantes tanto a rotación como a escala. De entre estos dos algoritmos, SURF logró obtener una mayor cantidad de puntos detectados por instancia, lo que da lugar a un margen más amplio al momento de calcular la calificación numérica por comparación de puntos.

### **4.3 Calificación de productos**

La calificación de coincidencias determina el nivel de similitud entre la imagen del producto buscado (clase) y cada uno de los productos segmentados del estante (instancias)

mediante la asignación de dos calificaciones numéricas a cada instancia: Calificación de puntos SURF y calificación de color. Para reducir el número de comparaciones entre clase e instancias, antes de realizar el proceso de calificación por SURF y color, se realiza una prueba de tamaño para comparar únicamente aquellas instancias que se aproximan a la medida del producto buscado.

El reconocimiento de los mejores candidatos se lleva a cabo una vez que todas las instancias obtienen una calificación de puntos SURF y color. Todo producto que tenga un nivel de coincidencia mayor o igual al 75% en puntos SURF y un nivel de coincidencia mayor al 40% en color es considerado como un candidato fuerte y su definitividad dependerá del análisis final ejecutado por el *Sistema de evaluación de probabilidad*.

#### **4.3.1 Prueba de tamaño**

La prueba de tamaño no asigna una calificación numérica al producto, simplemente determina si el tamaño del producto segmentado se aproxima lo suficientemente al tamaño de la clase buscada como para proceder a su calificación por puntos SURF y color. Si el tamaño del segmento que representa la instancia de un producto es demasiado pequeño con respecto al tamaño de la clase buscada, se descarta la instancia y las calificaciones de SURF y color se consideran cero. Este proceso ayuda a reducir la cantidad de productos analizados y así aumentar la eficiencia de la ejecución.

La prueba de tamaño consiste en la comparación del ancho en píxeles de un producto segmentado (una instancia), contra el ancho en píxeles de un *bounding box*.

Tal como se menciona en la sección 2.2.2 de este capítulo, dos de los parámetros de entrada del algoritmo corresponden al ancho del estante y al diámetro del producto,

ambos en metros. Con esta información se lleva a cabo el cálculo del ancho del *bounding box* en pixeles utilizando una regla de 3 como se muestra a continuación.

$$\text{Ancho bounding box} = \frac{\text{Diámetro del producto (m)} * \text{Ancho de la imagen del estante (pix)}}{\text{Ancho del estante (m)}}$$

*Ecuación 1 Cálculo del ancho de bounding box, equivale a diámetro del producto en pixeles*

La ecuación anterior supone que el ancho del estante en metros es equivalente al ancho en pixeles de la imagen del estante, esto puede darse por hecho ya que como se menciona, en el sub tema *2.1.2.1 Descripción la aplicación de fotografiado*, la imagen de entrada del estante fue capturada de manera tal que los bordes del anaquel coinciden con los bordes de la imagen. Si la imagen de entrada del estante no cumple con esta característica, el cálculo del tamaño esperado puede fallar y en el peor de los casos, descartar a todos los productos por no pasar la prueba de tamaño.

### **4.3.2 Calificación por puntos SURF**

La calificación por puntos SURF (Speed Up Robust Features) consiste en la comparación del conjunto de puntos SURF del producto segmentado contra el conjunto de puntos SURF del producto buscado. Al final del análisis, cada producto contará con una calificación numérica que representa la cantidad de coincidencias que se obtuvieron al comparar ambos conjuntos.

A pesar de que SURF es un algoritmo invariante a la escala, pruebas experimentales demostraron que la reducción del producto buscado a una medida aproximada a la etiqueta del producto segmentado ayuda a incrementar la precisión de coincidencia, es por esto que

antes de obtener los conjuntos de puntos característicos de ambas imágenes se lleva a cabo una reducción de la imagen de la clase buscada (imagen de cámara de línea del producto buscado) a un tamaño aproximado al que tendría la etiqueta dentro de la fotografía del estante. Haciendo uso del diámetro del producto en píxeles, calculado en la ecuación 1, se calcula el factor de escala como se muestra en la ecuación 2.

$$factor\ escala = \frac{perimetro\ esperado\ (pix)}{ancho\ de\ objeto\ buscado\ (pix)} \qquad perimetro\ esperado = diametro\ (pix) * \pi$$

*Ecuación 2 Cálculo del factor de escala (reducción) de imagen*

Con el factor de escala calculado se escala la imagen de la clase buscada. El resultado es una imagen del producto buscado de dimensiones similares a las de la instancia contra la que será comparada.

Una vez aproximados los tamaños de imagen de producto buscado y producto segmentado, se obtienen los conjuntos de puntos característicos de ambas imágenes y se realiza la comparación de los mismos. Las figuras 24 y 25 muestran el proceso gráfico de matching entre los dos conjuntos de puntos; la primera imagen ilustra el mejor match encontrado (130 puntos coincidentes) y la segunda a otro producto con menor cantidad de puntos coincidentes (33 coincidencias). En ambas imágenes se aprecia a la izquierda el producto buscado y a la derecha uno de los productos segmentados.

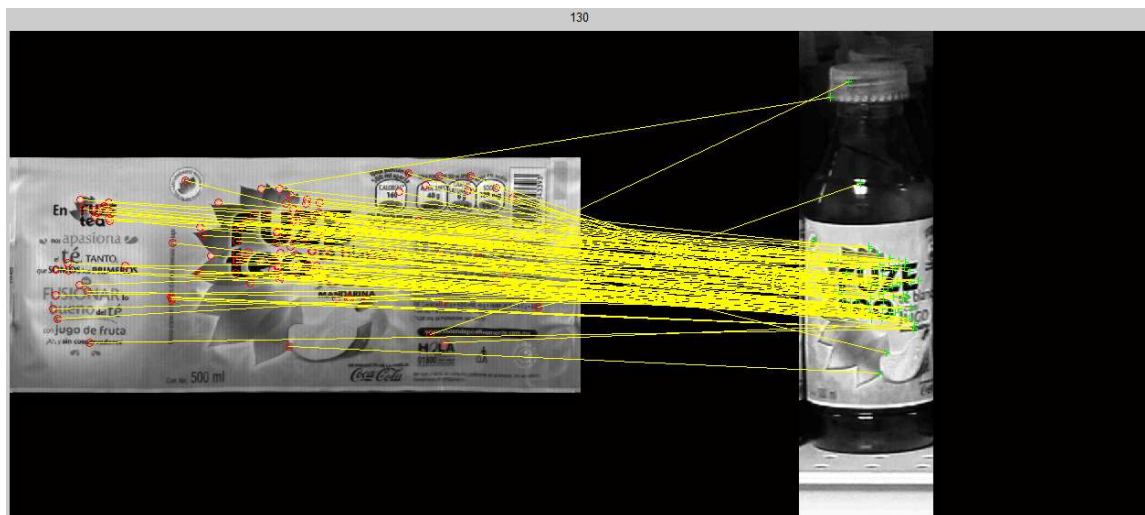


Figura 24 Imagen de mayor número de coincidencias

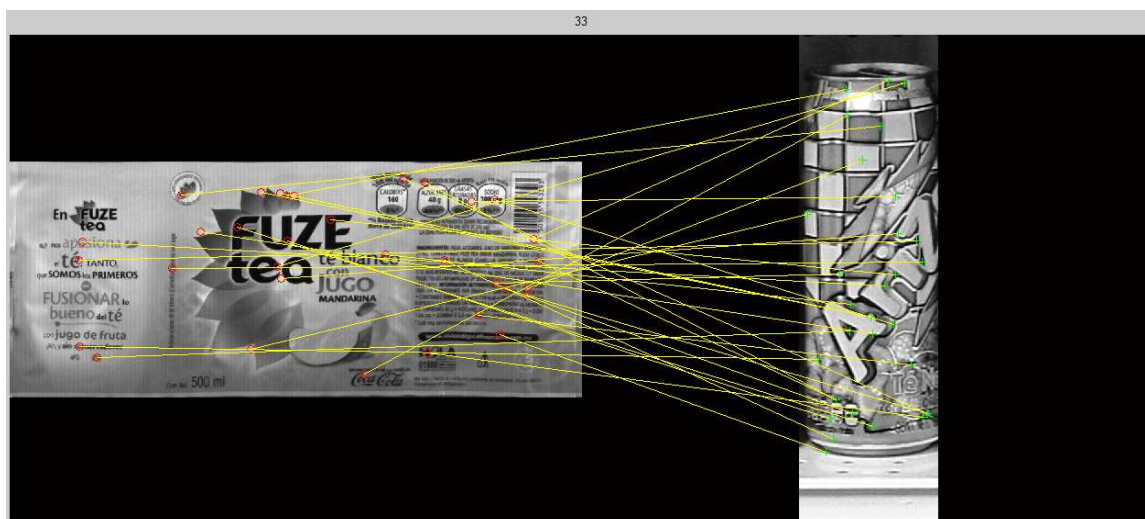


Figura 25 Imagen de match de un producto no concordante con el buscado

Para determinar los candidatos fuertes de entre todas las instancias existentes, se toma al producto con el número de coincidencias más alto y su calificación se utiliza como



base para determinar el porcentaje de similitud de cada producto. Por ejemplo, si después de calificar a todos los productos, la calificación más alta es igual a 120 puntos coincidentes, 120 se considerará el 100% de similitud. De esta manera, aquellos productos con 30, 60 y 90 puntos coincidentes equivaldrían al 25%, 50% y 75% de similitud respectivamente. Los candidatos fuertes son todos aquellos cuyo porcentaje de similitud es mayor o igual al 75% y su definitividad dependerá de la calificación de color.

El algoritmo SURF es una versión computacionalmente más rápida del algoritmo SIFT (Scale Invariant Feature Transform), y basa su comportamiento en el mismo principio de funcionamiento que SIFT, el cual se realiza mediante la ejecución de 4 pasos: Detección de máximos en espacio de escala, localización de puntos característicos, asignación de orientación y creación de descriptor de puntos característicos.

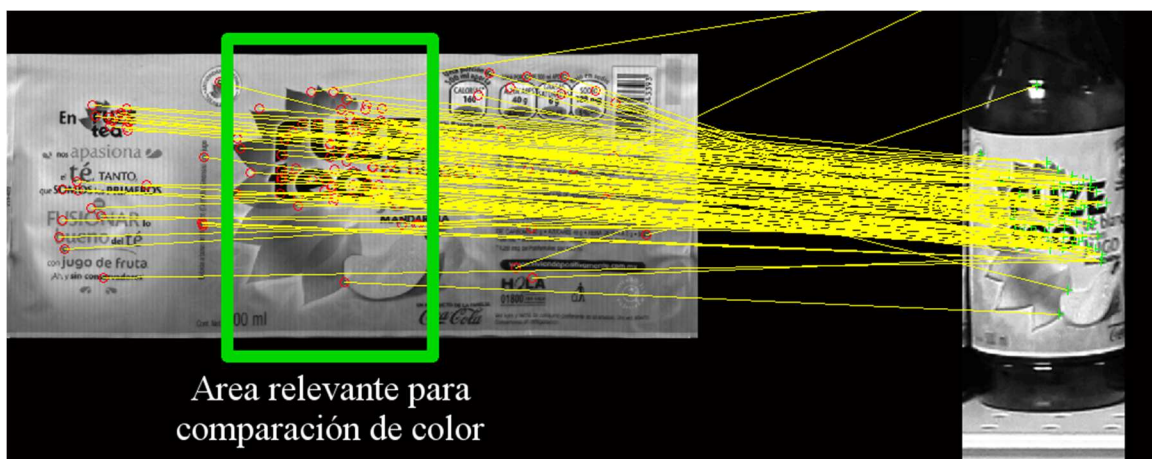
### **4.3.3 Calificación por color**

La calificación por color compara los valores de tonalidad, saturación e intensidad del producto buscado (clase) contra los mismos valores del producto segmentado (instancia) haciendo uso de la información contenida en los canales HSV de ambas imágenes. El análisis conlleva un pre procesamiento a fin de evaluar únicamente los colores en las áreas más concordantes de cada par de imágenes y de esta manera obtener una mayor precisión en la comparación de color.

#### **4.3.3.1 Pre procesamiento**

El producto buscado, al ser una proyección plana, muestra todos los ángulos del envase, mientras que el producto segmentado del estante únicamente muestra la sección que queda de frente a la cámara. Esto significa que habrá colores en el producto buscado

que no estarán presentes en el producto segmentado cuando se realice la comparación de color, causando un resultado impreciso. Para resolver este problema, el preprocesamiento calcula el área de relevancia de la siguiente manera: sobre la imagen del producto buscado se calcula la ubicación promedio<sup>7</sup> de todos los puntos coincidentes, considerando únicamente sus posiciones a lo largo de la imagen, es decir, horizontalmente. A partir de la ubicación calculada se recorta un área de ancho igual al diámetro del producto, el cual fue calculado durante la prueba de tamaño. El resultado es un segmento de la imagen planar que corresponde a la sección visible del producto segmentado (figura 26).



*Figura 26 Área relevante para comparación de color*

El producto segmentado también incluye información de color irrelevante en las áreas superiores e inferiores a la etiqueta, por ejemplo, el color del contenido del envase y el fondo del estante. Para segmentar el área de la etiqueta, el pre procesamiento calcula la

<sup>7</sup> Se calcula un promedio acotado para descartar outliers.

ubicación promedio<sup>8</sup> de la concentración de los puntos coincidentes sobre la altura de la imagen, es decir, verticalmente. A partir de la ubicación obtenida se recorta un área de altura igual a la altura de la imagen del producto buscado, la cual es una medida aproximada de la altura de la etiqueta y fue calculada durante la prueba de tamaño. El resultado es una imagen que muestra únicamente el área que corresponde a la etiqueta del producto (figura 27).

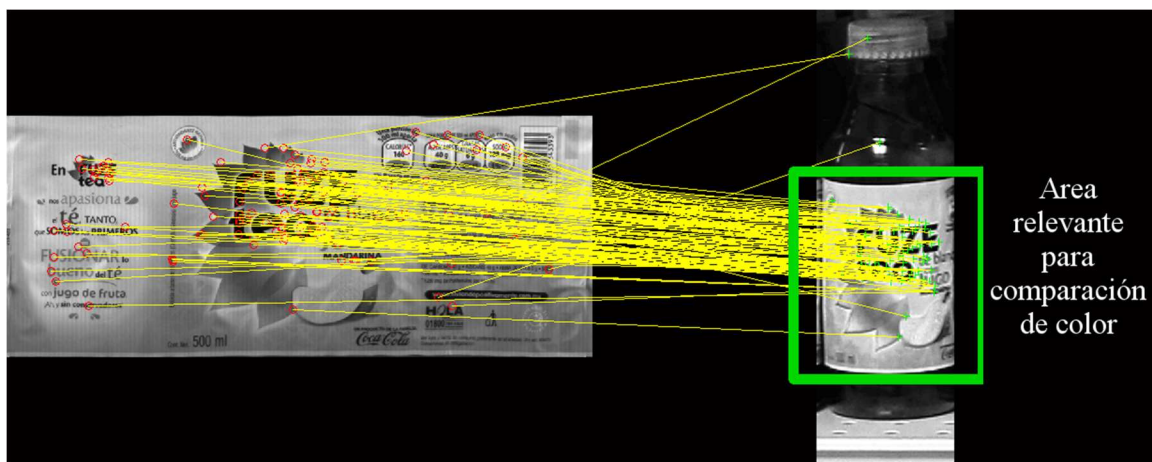
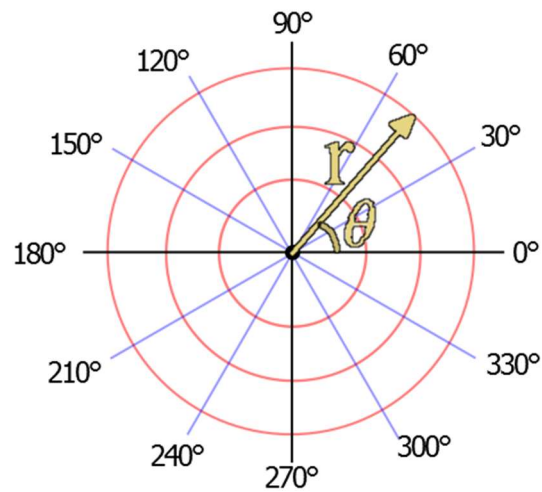


Figura 27 Área relevante para comparación de color

#### 4.3.3.2 Comparación de canales de color

Una vez segmentadas las áreas relevantes del producto buscado y el segmentado, se grafica la distribución de color en un plano polar como el de la figura 28, en donde  $r$  es el valor de saturación contenido en el canal S de la imagen y  $\theta$  es la tonalidad contenida en el canal H.

<sup>8</sup> Se calcula un promedio acotado para descartar outliers.



*Figura 28 Plano polar para representación de gráfica de color*

El canal de intensidad ( $V$ ) no tiene representación gráfica sobre el plano polar, sin embargo, se utiliza para filtrar los valores  $H$  y  $S$  que deben aparecer sobre la gráfica. Aquellos píxeles con intensidad ( $V$ ) muy baja son descartados para eliminar zonas muy oscuras. De manera similar, el canal de saturación ( $S$ ) filtra a aquellos píxeles cuya saturación se encuentre entre el 0.4 y 0.6 (valores normalizados) para descartar colores muy pálidos (saturación tiende a 0), o colores muy puros (saturación tiende a 1); éstos últimos son muy difícilmente encontrados en imágenes reales.

El resultado final de la graficación de color se muestra al centro de las figuras 29 y 30. En ellas se pueden apreciar las manchas de colores correspondientes al producto buscado y al producto segmentado del estante. Cada mancha es como una huella característica del producto y la calificación de color depende de la coincidencia que haya al traslapar dichas manchas. A la derecha de las figuras 29 y 30 se muestra una tercera gráfica que representa el traslape mencionado.

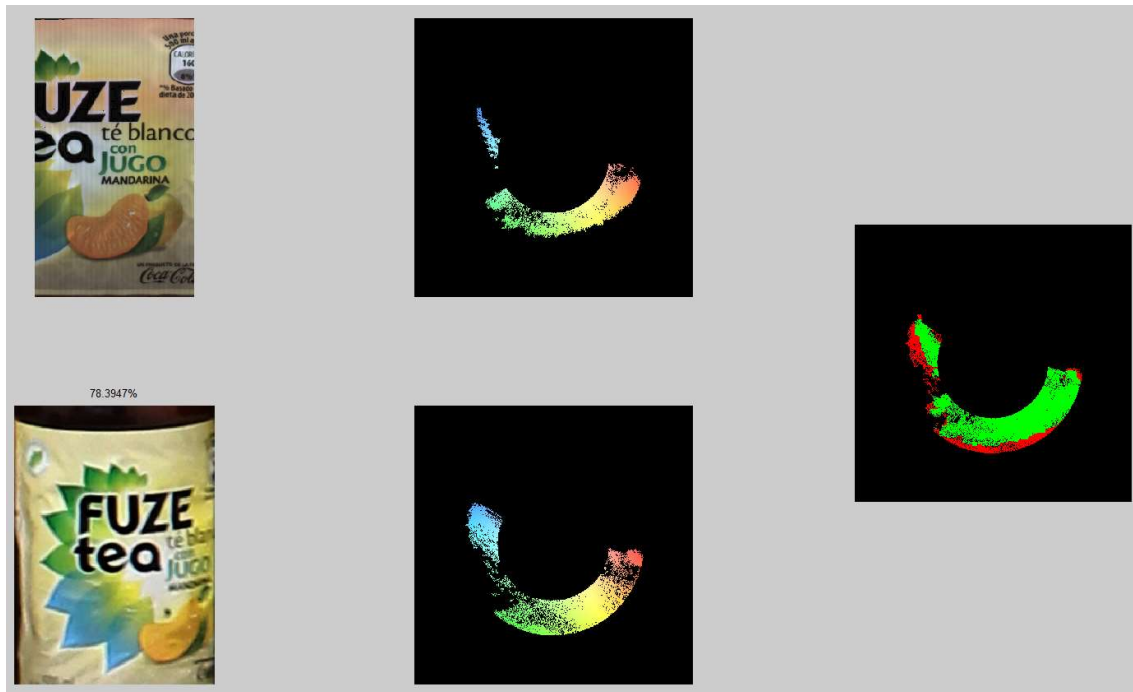


Figura 29 Imágenes con alta coincidencia de color 78%

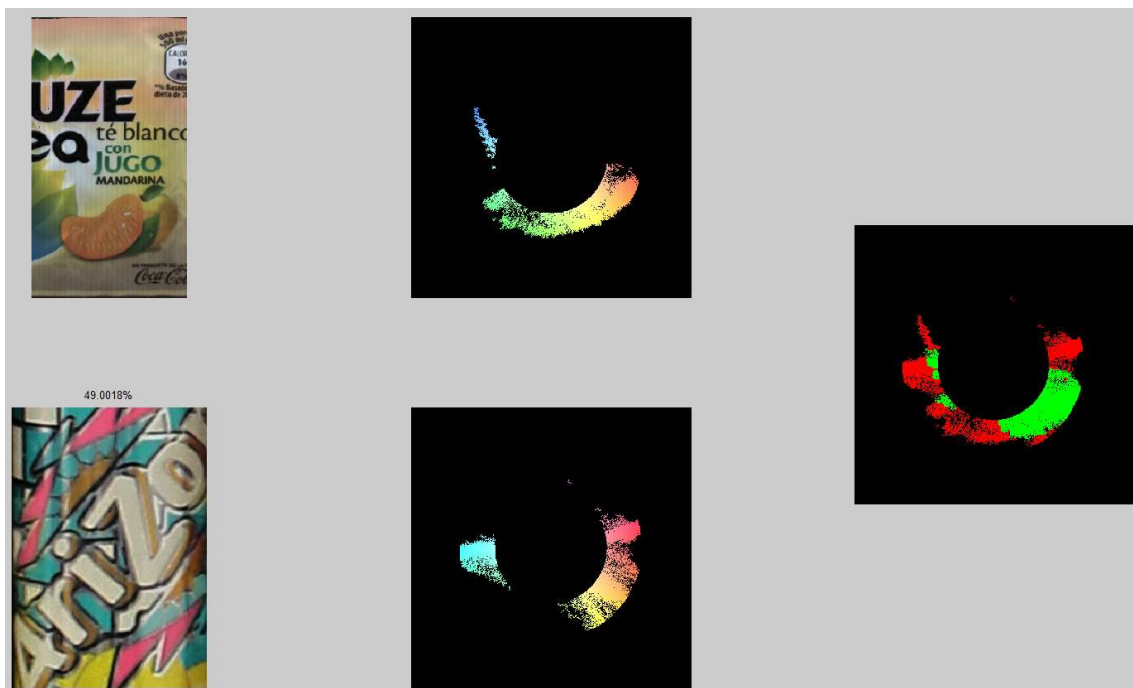


Figura 30 Imagen con menor coincidencia de color (49%)

El área verde representa a los colores coincidentes mientras que el área roja representa las no coincidencias. A partir de ésta última gráfica se puede obtener la calificación de color simplemente con calcular el porcentaje de área verde con respecto al área roja como se muestra en la ecuación 3.

$$\text{porcentaje de coincidencia} = \frac{\text{píxeles en verde}}{\text{píxeles en verde} + \text{píxeles en rojo}}$$

*Ecuación 3* Cálculo de porcentaje de coincidencia de color

Los candidatos fuertes de la calificación de color son los mismos obtenidos durante el proceso de calificación por puntos SURF, descartando únicamente aquellos que obtuvieron un porcentaje de coincidencia de color menor al 40%.

#### **4.3.4 Salida del sistema de calificación**

Por cada búsqueda de una clase, el sistema de calificación genera un archivo que almacena una *matriz de ubicaciones* de los candidatos más fuertes y una *imagen de probabilidades* que muestra la probabilidad conjunta de cada candidato fuerte mediante un recuadro en un tono de gris. Cada archivo es guardado con el nombre de la clase buscada y a su vez, todos los archivos generados son almacenados en una misma carpeta, la cual lleva el nombre del estante analizado.

La matriz de ubicaciones es una matriz de  $n$  renglones por 5 columnas, donde  $n$  es el número de candidatos fuertes del producto buscado y las columnas corresponden a las coordenadas de inicio y fin de cada una de esas apariciones  $(x_i, y_i, x_f, y_f)$ . La quinta columna de cada renglón contiene la probabilidad conjunta del producto, calculada a partir de la ecuación 4.

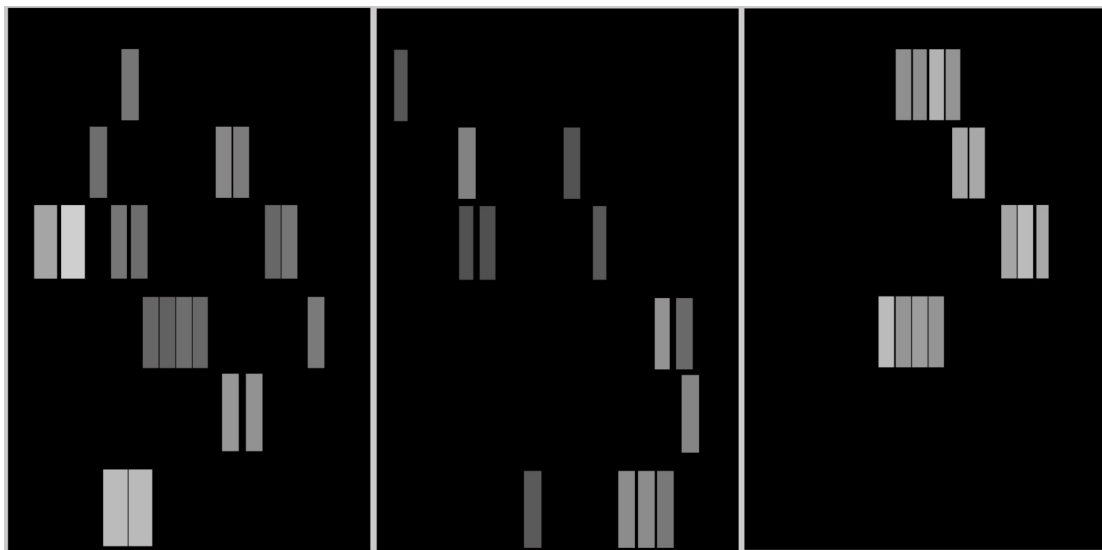
$$P_{color} = \text{porcentaje de coincidencia de color}$$

$$P_{SURF} = \text{porcentaje de coincidencia de puntos SURF}$$

$$\text{Probabilidad conjunta} = P_{SURF} * P_{color}$$

*Ecuación 4 Cálculo de la probabilidad conjunta*

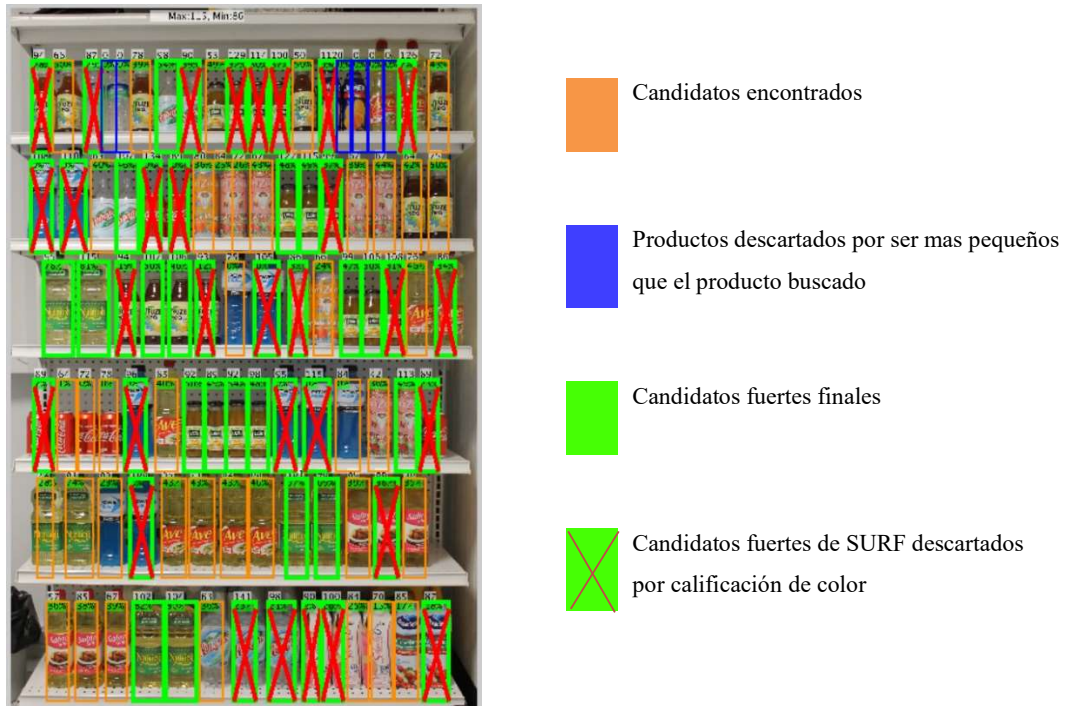
La imagen de probabilidad es una imagen de intensidad en la que cada ubicación de cada producto candidato fuerte es rellenada con el valor de su probabilidad conjunta. Esta imagen es utilizada por el sistema de evaluación de probabilidad para discernir a qué clase pertenece cada instancia en caso de haber conflicto entre dos posibles clases.



*Figura 31 Ejemplos de imágenes de probabilidad*

Como salida extra del sistema de calificación se genera una imagen del estante completo con recuadros de colores superpuestos indicando los productos encontrados (ámbar), los productos descartados por prueba de tamaño (azul), las mejores coincidencias por puntos característicos (verde) y los productos descartados por prueba de color (verde con cruz roja).

La imagen incluye información adicional en recuadros blancos que indican el número de coincidencias de puntos característicos y anotaciones en recuadros verdes con el porcentaje de similitud de color. Ésta última salida del sistema es utilizada únicamente con fines ilustrativos y de depuración.



*Figura 32 Salida gráfica del sistema de calificación*



## Capítulo 5:

### Sistema de reconocimiento por probabilidad.

El sistema de evaluación de probabilidad utiliza los archivos de salida generados por el sistema de detección y calificación de productos (matriz de ubicaciones e imagen de probabilidad) para filtrar los mejores resultados. El sistema es controlado mediante una interfaz elaborada en Matlab (figura 33) y se ejecuta una vez que todas las clases de la base de datos han sido analizadas por el sistema de calificación.

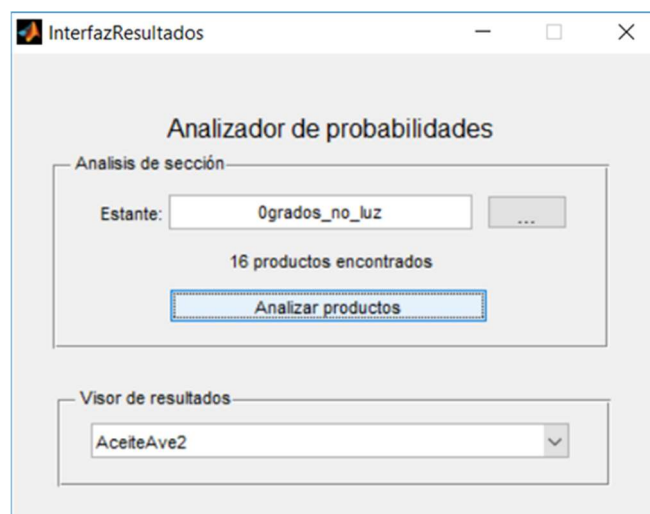


Figura 33 Interfaz del sistema de evaluación de probabilidad

#### 5.1 Funcionamiento del sistema

Por cada producto (clase), el sistema de evaluación de probabilidad lee la correspondiente matriz de ubicaciones y compara todas las ubicaciones registradas en ella contra todas las demás ubicaciones almacenadas en las matrices del resto de las clases. Si la ubicación de una instancia es compartida por dos o más clases, dicha ubicación prevalecerá únicamente para aquella clase con valor de probabilidad conjunta más alto.

Debido a que las ubicaciones no llegan a ser exactamente las mismas a causa del diámetro de cada producto, el sistema utiliza las imágenes de probabilidad para determinar la prevalencia de una ubicación entre distintos productos de la siguiente manera: Supongamos un producto clase *A*, tiene 4 instancias fuertes ubicadas en  $u_1, u_2, u_3$  y  $u_4$  y su imagen de probabilidad corresponde a la figura 34 (a). Supongamos también un producto B con 4 instancias fuertes ubicados en  $v_1, v_2, v_3$  y  $v_4$  e imagen de probabilidad mostrada en la figura 34(b).

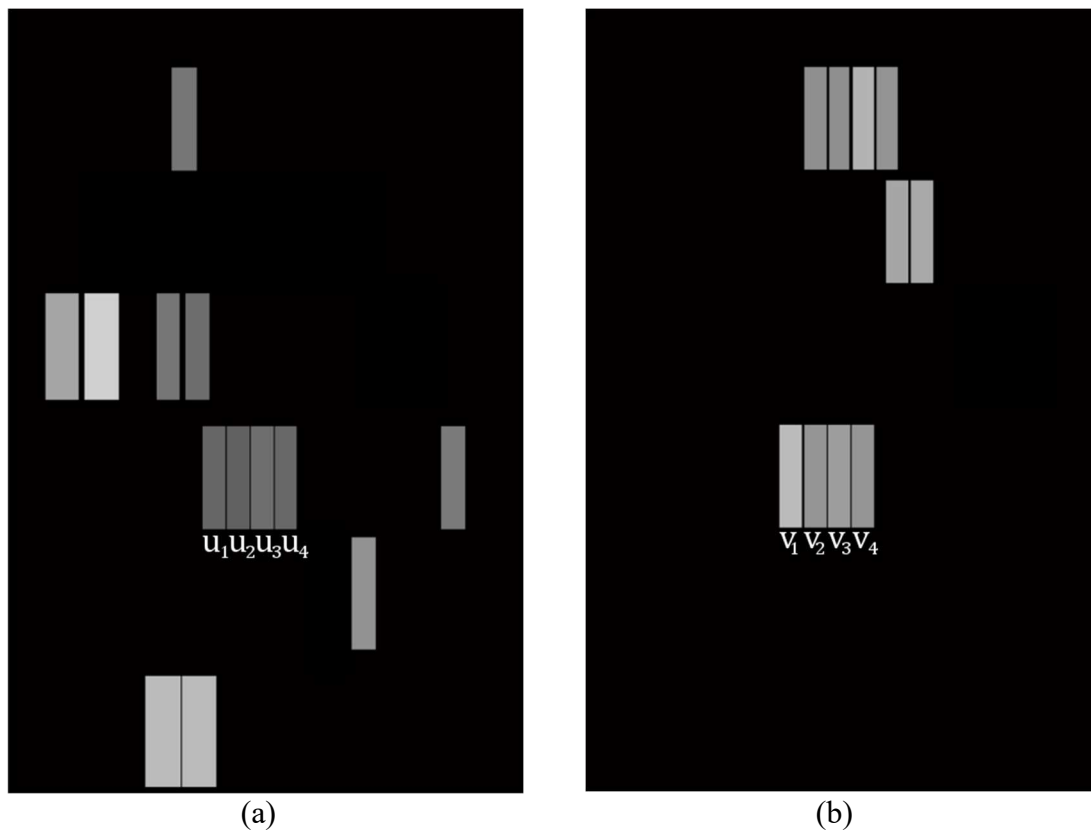


Figura 34 Imágenes de probabilidad de productos clase *A* y *B*

Cada recuadro de las imágenes de probabilidad está relleno homogéneamente con un valor igual al de la probabilidad conjunta del producto que representa. Las ubicaciones correspondientes a las instancias  $u_1, u_2, u_3$  y  $u_4$  traslapan respectivamente con

las ubicaciones de las instancias  $v_1, v_2, v_3$  y  $v_4$ . Dado que cada ubicación no es exactamente concordante, el cálculo de la moda de intensidades dentro de las áreas a comparar determina qué ubicación contiene la probabilidad conjunta más alta. Para nuestro ejemplo, es claramente apreciable que la tonalidad de los recuadros que los recuadros  $v$  tienen una intensidad más cercana a blanco, lo que indica un valor de probabilidad conjunta más alta, por lo que dichas ubicaciones prevalecerán como candidatos finales para la clase B y no para la clase A.

## **5.2 Salida del sistema de evaluación de probabilidad**

El sistema de evaluación de probabilidad genera un archivo con las ubicaciones de los productos seleccionados como coincidentes finales de cada clase. Dichas ubicaciones pueden ser utilizadas posteriormente para su comparación contra un planograma.

Además del archivo de ubicaciones, el sistema cuenta con una salida gráfica que permite visualizar las ubicaciones finales antes y después del análisis de probabilidad. La figura 35 muestra una salida gráfica de ejemplo para la clase *Aceite Nutrioli*.



*Antes de análisis de probabilidad  
(Candidatos fuertes en verde)*

*Después de análisis de probabilidad  
(Candidatos finales en verde)*

*Figura 35 Salida gráfica del sistema*

## Capítulo 6:

### Pruebas de desempeño y resultados.

Para las pruebas de desempeño se utilizó el estante experimental y 16 productos acomodados en 4 orientaciones: frontal, 30, 60 y 90 grados. Los productos fueron rotados con el fin de verificar la efectividad del algoritmo ante posibles orientaciones que se podrían encontrar en la tienda.

La tabla 1 muestra los 16 productos seleccionados y la cantidad de éstos mismos sobre el estante experimental. La figura 36 muestra las fotografías del estante experimental.



*Figura 36 Estante experimental en 4 orientaciones*

*Tabla 1 Productos seleccionados para pruebas de desempeño*

<i>Nombre del producto</i>	<i>Cantidad en el estante</i>
1. Aceite Ave	8
2. Aceite Nutrioli	6
3. Aceite Sabrosano	4
4. Arizona Fresa Kiwi	6
5. Arizona Mango	3
6. Coca lata logo horizontal	5
7. Coca lata logo vertical	2
8. Del valle durazno 250ml	14
9. Fuze tea mandarina 500ml	11
10. Jumex mango 500ml	2
11. Leche Siluette	3
12. Ocean Spray	2
13. Paladium	2
14. Peñafiel grande (botella)	3
15. Peñafiel vidrio (chico)	4
16. Powerade azul 1L	10

Se realizó la ejecución del algoritmo con los 16 productos (clases) de la tabla 1 y las 4 imágenes de los anaqueles. Los resultados se registraron en 4 matrices de confusión, disponibles en el apéndice de esta tesis. En las matrices, los renglones representan los productos buscados (clases) y las columnas las coincidencias (instancias) de un producto. Tanto renglones como columnas comparten la numeración de productos mostrada en la tabla 1.

Las tablas 2 a 5 muestran los resultados registrados en las matrices de confusión para cada una de las pruebas a 0°, 30°, 60° y 90°. La nomenclatura de las columnas se muestra a continuación:

- $TP$  = Verdaderos positivos (True Positives)
- $TN$  = Verdaderos negativos (True Negatives)
- $FP$  = Falsos positivos (False Positives)
- $FN$  = Falsos negativos (False Negatives)
- $TPR$  = Sensibilidad (True Positive Ratio)
- $TNR$  = Especificidad (True Negative Ratio)
- $FPR$  = Probabilidad de falsa alarma (FPR)

*Tabla 2 Tabla de resultados para matriz de confusión de ejecución a 0°*

<b>Producto</b>	<b>Existencias</b>	<b>TP</b>	<b>TN</b>	<b>FP</b>	<b>FN</b>	<b>TPR</b>	<b>TNR</b>	<b>FPR</b>
1	8	5	66	1	2	0.7143	0.9851	0.0149
2	6	6	66	0	2	0.7500	1.0000	0.0000
3	4	5	69	0	0	1.0000	1.0000	0.0000
4	6	2	67	2	3	0.4000	0.9710	0.0290
5	3	1	71	1	1	0.5000	0.9861	0.0139
6	5	1	72	1	0	1.0000	0.9863	0.0137
7	2	1	72	1	0	1.0000	0.9863	0.0137
8	14	13	60	1	0	1.0000	0.9836	0.0164
9	11	9	64	1	0	1.0000	0.9846	0.0154
10	2	1	72	1	0	1.0000	0.9863	0.0137
11	3	3	71	0	0	1.0000	1.0000	0.0000
12	2	2	72	0	0	1.0000	1.0000	0.0000
13	2	0	74	0	0	0.0000	1.0000	0.0000
14	3	2	71	2	4	0.3333	0.9726	0.0274
15	4	2	71	1	0	1.0000	0.9861	0.0139
16	10	9	65	0	0	1.0000	1.0000	0.0000

*Tabla 3 Tabla de resultados para matriz de confusión de ejecución a 30°*

<i>Producto</i>	<i>Existencias</i>	<i>TP</i>	<i>TN</i>	<i>FP</i>	<i>FN</i>	<i>TPR</i>	<i>TNR</i>	<i>FPR</i>
1	8	1	53	2	1	0.5000	0.9636	0.0364
2	6	3	53	1	0	1.0000	0.9815	0.0185
3	4	5	50	1	1	0.8333	0.9804	0.0196
4	6	0	57	0	0	0.0000	1.0000	0.0000
5	3	0	55	2	0	0.0000	0.9649	0.0351
6	5	0	56	1	0	0.0000	0.9825	0.0175
7	2	2	53	1	1	0.6667	0.9815	0.0185
8	14	8	28	2	0	1.0000	0.9333	0.0667
9	11	9	48	0	0	1.0000	1.0000	0.0000
10	2	0	55	2	0	0.0000	0.9649	0.0351
11	3	3	53	0	1	0.7500	1.0000	0.0000
12	2	2	54	0	1	0.6667	1.0000	0.0000
13	2	0	57	0	0	0.0000	1.0000	0.0000
14	3	2	46	1	8	0.2000	0.9787	0.0213
15	4	0	51	0	6	0.0000	1.0000	0.0000
16	10	3	48	6	0	1.0000	0.8889	0.1111



*Tabla 4 Tabla de resultados para matriz de confusión de ejecución a 60°*

<i>Producto</i>	<i>Existencias</i>	<i>TP</i>	<i>TN</i>	<i>FP</i>	<i>FN</i>	<i>TPR</i>	<i>TNR</i>	<i>FPR</i>
1	8	5	49	1	7	0.4167	0.9800	0.0200
2	6	3	59	0	0	1.0000	1.0000	0.0000
3	4	3	56	3	0	1.0000	0.9492	0.0508
4	6	0	60	0	2	0.0000	1.0000	0.0000
5	3	2	59	1	0	1.0000	0.9833	0.0167
6	5	0	58	2	2	0.0000	0.9667	0.0333
7	2	0	59	1	2	0.0000	0.9833	0.0167
8	14	11	50	0	1	0.9167	1.0000	0.0000
9	11	6	54	1	1	0.8571	0.9818	0.0182
10	2	2	60	0	0	1.0000	1.0000	0.0000
11	3	0	56	6	0	0.0000	0.9032	0.0968
12	2	0	62	0	0	0.0000	1.0000	0.0000
13	2	0	60	0	2	0.0000	1.0000	0.0000
14	3	0	54	2	6	0.0000	0.9643	0.0357
15	4	0	59	3	0	0.0000	0.9516	0.0484
16	10	7	52	3	0	1.0000	0.9455	0.0545

*Tabla 5 Tabla de resultados para matriz de confusión de ejecución a 90°*

<i>Producto</i>	<i>Existencias</i>	<i>TP</i>	<i>TN</i>	<i>FP</i>	<i>FN</i>	<i>TPR</i>	<i>TNR</i>	<i>FPR</i>
1	8	0	42	1	8	0.0000	0.9767	0.0233
2	6	2	46	0	3	0.4000	1.0000	0.0000
3	4	3	41	3	4	0.4286	0.9318	0.0682
4	6	0	45	2	4	0.0000	0.9574	0.0426
5	3	0	50	1	0	0.0000	0.9804	0.0196
6	5	0	45	5	1	0.0000	0.9000	0.1000
7	2	1	47	1	2	0.3333	0.9792	0.0208
8	14	3	46	2	0	1.0000	0.9583	0.0417
9	11	1	46	4	0	1.0000	0.9200	0.0800
10	2	2	49	0	0	1.0000	1.0000	0.0000
11	3	0	43	5	3	0.0000	0.8958	0.1042
12	2	1	48	1	1	0.5000	0.9796	0.0204
13	2	0	49	1	1	0.0000	0.9800	0.0200
14	3	1	48	0	2	0.3333	1.0000	0.0000
15	4	0	48	1	2	0.0000	0.9796	0.0204
16	10	6	41	4	0	1.0000	0.9111	0.0889

En las tablas anteriores (tablas 2 a 5), para cada producto, la columna de sensibilidad (TPR) indica la proporción de positivos que fueron correctamente clasificados como positivos, la especificidad (TNR) indica la proporción de negativos que fueron correctamente identificados como negativos y la probabilidad de falsa alarma (FPR) indica la expectativa de aparición de un falso positivo. Si se promedian los resultados obtenidos en cada tabla en las 3 columnas mencionadas podemos obtener una tabla que muestre el desempeño generalizado de cada ejecución (ver tabla 6).

*Tabla 6 Desempeño generalizado de cada ejecución*

<i>Ejecuciones realizadas</i>	<i>Promedio TPR</i>	<i>Promedio TNR</i>	<i>Promedio FPR</i>
Ejecución a 0°	0.7936	0.9893	0.0107
Ejecución a 30°	0.4760	0.9763	0.0237
Ejecución a 60°	0.4494	0.9756	0.0244
Ejecución a 90°	0.3747	0.9594	0.0406

Como se puede ver en la tabla 6, la ejecución a 0° (frontal) contiene los valores más altos promedio de sensibilidad (TPR) y especificidad (TNR) así como el más bajo en falsas alarmas (FPR). Esto indica que el sistema reconoce con mayor precisión a aquellos estantes cuyos productos están mejor alineados.

### **6.1 Observaciones finales**

Con base en los resultados obtenidos en la tabla 6, se puede concluir que el reconocimiento de productos es mucho más efectivo cuando estos se encuentran de frente a la cámara. Esto se debe a que entre mayor es la rotación, la sección de informativa del

producto queda más expuesta y dicha sección suele aportar muy poca información de reconocimiento debido a que los logotipos y letras en esa zona son muy pequeños y la resolución de la cámara no los captura con suficiente claridad. Éste es un factor difícil de mejorar ya que depende enteramente de la resolución disponible en el hardware con el que se capturen las fotografías del estante.

A pesar de que en las pruebas de orientación a 60 y 90 grados, el desempeño fue decayendo, se debe tomar en consideración que en dichos casos, el estante experimental es un diseño en el que se busca el peor caso posible, es decir, representan situaciones que son muy poco probables de encontrar en un estante real de supermercado. Además de esto, se observó que el problema de rotación en supermercados reales tiende a reducirse gracias a que éstos, por cuestiones de mercadotecnia, procuran mantener sus productos lo mejor alineados posible de manera que la clientela los pueda ubicar fácilmente.

## **Capítulo 7:**

### **Trabajo a futuro.**

El trabajo realizado permitió plantear posibles expansiones en diferentes áreas del sistema de reconocimiento que permitan mejorar los resultados obtenidos.

En cuanto a refinación de resultados, la inclusión de un planograma, como información de entrada adicional en el que se incluyan las ubicaciones esperadas de los productos buscados podría ayudar a discernir entre aquellos productos en los que se tiene menor certeza de a qué clase pertenecen.

En cuanto a detección y segmentación de productos, la implementación de hardware adicional como cámaras infrarrojas permitiría detectar la profundidad de los productos o el estante mismo, de manera que se logre obtener información adicional que mejore la segmentación de ubicaciones y medidas de cada instancia de producto.

Finalmente, en cuanto a rendimiento, la ejecución en paralelo de las comparaciones entre clase e instancias podría mejorar el tiempo de ejecución que se tiene entre cada análisis.

### Lista de referencias

- [1] Kaulgud, P., Prahlad, S., Satapathy, R. (2015). IEEE Region 10 Symposium. *Smart Shelfie - Internet of Shelves: For Higher On-Shelf Availability*.  
 [En línea] <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7166240>  
 [Accedido por última vez en marzo del 2016]
- [2] Patent US8189855 - *Planogram extraction based on image processing*. (2016).  
 Google Books. [En línea] <http://www.google.com.mx/patents/US8189855>  
 [Accedido por última vez el 20 de Abril del 2016]
- [3] Fareed, M., Hussain, S., Meraj, M., Wagas, M., Wasi, K. (2015). Proceedings of the 2nd International Conference on Engineering & Emerging Technologies (ICEET), Superior University, Lahore.  
*Identification for refilling of shelves using image processing*  
 [En línea] [http://www.superior.edu.pk/ICEET/pdf/research2015/submission\\_69.pdf](http://www.superior.edu.pk/ICEET/pdf/research2015/submission_69.pdf)  
 [Accedido por última vez en marzo del 2016.]
- [4] Fan, J., Zhang, T. Hewlett-Packard Laboratories, Palo Alto, California, USA  
*Shelf detection via vanishing point and radial projection*  
 [En línea]  
<http://ieeexplore.ieee.org/pbidi.unam.mx:8080/stamp/stamp.jsp?tp=&arnumber=7025315>  
 [Accedido por última vez en marzo del 2016]
- [5] Department of Computer Science, University of British Columbia, Vancouver, Canada (2006),  
 Matthew Brown\* and David G. Lowe  
*Automatic Panoramic Image Stitching using Invariant Features*
- [6] Department of Computer Science, University of British Columbia, Vancouver, Canada,  
 David G. Lowe  
*Object Recognition from Local Scale-Invariant Features*
- [7] School of Computer Science, Carnegie Mellon University; Intel Research Pittsburgh,  
 Yan Ke, Rahul Sukthankar  
*PCA-SIFT: A More Distinctive Representation for Local Image Descriptors*
- [8] Dept. of Electronics Engineering, Myong Ji University, Korea, SOC Platform Research Division, Korea Electronics Tech. Inst., Korea,  
 Thao Nguyen, Eun-Ae Park, Jiho Han, Dong-Chul Park, and Soo-Young Min  
*Object Detection Using Scale Invariant Feature Transform*

## Apéndice

Matrices de confusión en las que se describen las clases buscadas (renglones) y el número de instancias encontradas por cada clase (columnas).

		Instancias															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Clases	1	5				1	1										
	2		6	1											1		
	3			5													
	4				2	1									1	1	
	5					1				1							
	6						1										
	7							1									
	8								13								
	9									9							
	10										1						
	11											3					
	12												2				
	13																
	14	1			1			1		1					2		
	15															2	
	16																9

No.	Nombre	Total en estante	No encontrados
1	Aceite Ave	8	3
2	Aceite Nutrioli	8	2
3	Aceite Sabrosano	6	1
4	Arizona Fresa Kiwi	6	4
5	Arizona Mango	3	2
6	Coca lata logo horizontal	5	4
7	Coca lata logo vertical	2	1
8	Del valle durazno 250ml	14	1
9	Fuze tea mandarina 500ml	11	2
10	Jumex mango 500ml	2	1
11	Leche Siluette	3	0
12	Ocean Spray	2	0
13	Paladium	2	2
14	Peñafiel 355ml	4	2
15	Peñafiel botella	3	1
16	Powerade azul 1L	10	1

1 Matriz de confusión de prueba frontal

		Instancias															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Clases	1	1									1						
	2		3														
	3	1		5													
	4																
	5																
	6																
	7							1	2								
	8									8							
	9										9						
	10																
	11								1			3					
	12												2		1		
	13																
	14	1	1	1		2			2		1				2		
	15																6
	16																3

No.	Nombre	Total en estante	No encontrados
1	Aceite Ave	8	7
2	Aceite Nutrioli	8	5
3	Aceite Sabrosano	6	1
4	Arizona Fresa Kiwi	6	6
5	Arizona Mango	3	3
6	Coca lata logo horizontal	5	5
7	Coca lata logo vertical	2	0
8	Del valle durazno 250ml	14	6
9	Fuze tea mandarina 500ml	11	2
10	Jumex mango 500ml	2	2
11	Leche Siluette	3	0
12	Ocean Spray	2	0
13	Paladium	2	2
14	Peñafiel 355ml	4	2
15	Peñafiel botella	3	3
16	Powerade azul 1L	10	7

2 Matriz de confusión de prueba a 30 grados



		Instancias															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Clases	1	5		2			1	1		1					2		
	2		3														
	3			3													
	4											1					1
	5					2											
	6												2				
	7						1						1				
	8			1						11							
	9	1									6						
	10											2					
	11																
	12																
	13																2
	14					1							2			3	
	15																
	16																7

No.	Nombre	Total en estante	No encontrados
1	Aceite Ave	8	3
2	Aceite Nutrioli	8	5
3	Aceite Sabrosano	6	3
4	Arizona Fresa Kiwi	6	6
5	Arizona Mango	3	1
6	Coca lata logo horizontal	5	5
7	Coca lata logo vertical	2	2
8	Del valle durazno 250ml	14	3
9	Fuze tea mandarina 500ml	11	5
10	Jumex mango 500ml	2	0
11	Leche Siluette	3	3
12	Ocean Spray	2	2
13	Paladium	2	2
14	Peñafiel 355ml	4	4
15	Peñafiel botella	3	3
16	Powerade azul 1L	10	3

3 Matriz de confusión de prueba a 60 grados

		Instancias															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Clases	1				1	1	3			3							
	2		2						2					1			
	3			3	1					1		1				1	
	4											3	1				
	5																
	6											1					
	7			1			1	1									
	8								3								
	9									1							
	10										2						
	11			1			1	1									
	12												1				1
	13																1
	14	1		1												1	
	15																2
	16																6

No.	Nombre	Total en estante	No encontrados
1	Aceite Ave	8	8
2	Aceite Nutrioli	8	6
3	Aceite Sabrosano	6	3
4	Arizona Fresa Kiwi	6	6
5	Arizona Mango	3	3
6	Coca lata logo horizontal	5	5
7	Coca lata logo vertical	2	1
8	Del valle durazno 250ml	14	11
9	Fuze tea mandarina 500ml	11	10
10	Jumex mango 500ml	2	0
11	Leche Siluette	3	3
12	Ocean Spray	2	1
13	Paladium	2	2
14	Peñafiel 355ml	4	3
15	Peñafiel botella	3	3
16	Powerade azul 1L	10	4

4 Matriz de confusión de prueba a 90 grados