

Un Enfoque para Aplicar la Visión por Computador en el Control de Calidad de Refrigeradores sobre la Línea de Ensamble

Carlos A. López¹, Carlos D. Ferrin² y Luis F. Castillo^{1,3}

(1) Facultad de Ingeniería y Arquitectura, Universidad Nacional Sede Manizales, Manizales-Colombia.
(e-mail: caralopezcas@unal.edu.co; lfcastilloos@unal.edu.co)

(2) Facultad de Ingeniería, Universidad del Valle, Santiago de Cali-Colombia
(e-mail: carlos.ferrin@correounivalle.edu.co)

(3) Facultad de Ingenierías, Universidad de Caldas, Grupo de Investigación GITIR, Manizales-Colombia
(e-mail: luis.castillo@ucaldas.edu.co)

Recibido Jun. 28, 2018; Aceptado Sep. 10, 2018; Versión final Nov. 1, 2018, Publicado Abr. 2019

Resumen

En este trabajo se propone una metodología de visión por computador para detectar accesorios internos en tiempo real y dictaminar su correcto ensamblado dentro de los refrigeradores. Las devoluciones de refrigeradores por fallas estéticas son de los mayores problemas que enfrenta la industria de manufactura de línea blanca. Esto se debe principalmente a que la evaluación estética es un proceso subjetivo que depende de las capacidades visuales y cognitivas de un inspector de calidad. La metodología consta de un bloque de adquisición de múltiples vistas utilizando sensores RGB-D-NIR, de algoritmos de segmentación, seguimiento, detección de regiones de interés y dictamen basado en puntos característicos. Los resultados obtenidos permiten concluir que el método propuesto presenta buen desempeño y rapidez, no introduce cadencia a la línea de producción y permite que las técnicas de dictamen puedan ser mejorados para lograr la automatización completa de la inspección visual de los accesorios.

Palabras clave: control de calidad; manufactura; refrigeradores; visión por computador

An Approach for Applying Computer Vision in Refrigerator Quality Control on Assembly Lines

Abstract

This paper proposes a computer vision methodology to detect internal accessories in real time and discern their correct assembly inside the refrigerators. Refrigerator returns due to aesthetic failures are one of the biggest problems facing the white goods manufacturing industry. Main cause of this problem is that aesthetic evaluation is a subjective task that depends on the visual and cognitive skills of quality inspectors. The methodology consists of an acquisition block of multiple views using RGB-D-NIR sensors and algorithms such as segmentation, object tracking, regions of interest detection and key point detection based verdict. Experimental results showed that the proposed methodology presents good performance and speed; besides, it does not introduce cadence to the production line and allows that the key point detection based verdict algorithm may be improved in future works in order to achieve a complete automation of the visual inspection process for accessories in the refrigerator.

Keywords: quality assurance; manufacturing; refrigerator; computer vision

INTRODUCCIÓN

La calidad de los refrigeradores es un objetivo clave en la industria de línea blanca ya que contribuye a la reducción de costos por ejemplo en servicio postventa y mejor aún, permite buscar mayor participación de mercado gracias al reconocimiento de los consumidores; para alcanzar este objetivo, se hace necesario ejecutar pruebas funcionales y estéticas, las cuales son realizadas durante el proceso de fabricación (Harrington, 2001). Las pruebas funcionales verifican, entre otras cosas, que el consumo de potencia y el control de temperatura cumpla con las especificaciones de diseño. Estas pruebas se llevan a cabo con dispositivos electrónicos que informan de manera precisa y en tiempo real (aproximadamente 30 segundos) si la unidad de refrigeración cumplió o no con la prueba. Por su parte, las pruebas estéticas aseguran dos cosas: primero que no haya ningún tipo de daño en los accesorios internos y externos en la unidad de refrigeración; y segundo, que para un modelo en particular los accesorios estén presentes (Khangura, 2012). Las pruebas estéticas requieren de la capacidad visual de un experto (llamado también Inspector Visual de Calidad) quien debe conocer los accesorios de cada modelo en forma exacta. En una de las líneas de producción más rápida en MABE, donde pasan 30 a 60 modelos diferentes mensuales, se estima que un inspector visual de calidad tarda entre 45 a 60 segundos aproximadamente en inspeccionar una nevera. De acuerdo a lo anterior, es evidente el alto grado de subjetividad que poseen este tipo de evaluaciones estéticas y en consecuencia la alta susceptibilidad a errores. Para el caso de accesorios internos (en compartimento y puerta del refrigerador), un tipo de error son los accesorios faltantes o equivocados; en este sentido, concebir un sistema capaz de emular la capacidad visual del inspector es un gran reto, teniendo en cuenta además que el sistema debe ser no invasivo y ser capaz de analizar el interior del refrigerador a medida que este se mueve en la línea de producción.

Para comprender la problemática planteada anteriormente, se debe describir el proceso de evaluación estética en la línea de producción: durante la evaluación estética un experto debe concentrar su atención en la unidad de refrigeración que está analizando (separarla del resto de unidades de refrigeración y del fondo), seguirla visualmente y sobre la banda, identificar el modelo de la unidad de refrigeración (esto se hace con la ayuda de un lector portable de código de barras) y analizar cada accesorio para determinar si falta o está equivocado (reemplazado por el accesorio de otro modelo de nevera). El mismo inspector (o un asistente) se encarga de retirar la unidad de refrigeración de la línea, e informar a través de un aplicativo en el lector portable que la nevera no ha pasado la prueba estética. En consecuencia, en un sistema de visión por computador, el reto de identificar accesorios faltantes/equivocados en unidades de refrigeración puede ser abordado dividiéndolo en dos grandes problemas: 1) detección y seguimiento del accesorio; y 2) clasificación del accesorio en correcto, faltante o equivocado. En este trabajo, se abordará una propuesta de solución al primer problema: detección y seguimiento de accesorios en unidades de refrigeración para su posterior análisis (p.e., clasificación).

A la fecha, el único trabajado encontrado relacionado con la inspección visual de accesorios internos en refrigeradores es el reportado en (Sudeep, 2015). En este trabajo se aborda únicamente el problema de accesorios faltantes utilizando técnicas de correspondencia de puntos característicos entre las imágenes (en monocromo) del accesorio de prueba y del modelo de referencia. Sin embargo, en los procesos de ensamblados no solo ocurren problemas de accesorio faltante sino también problemas de accesorios equivocados, esto es, un accesorio de otro modelo de nevera (con diseño totalmente diferente) es ensamblado en una nevera que no corresponde con dicho modelo.

METODOLOGÍA

Para adelantar esta investigación se plantea una metodología para la detección y seguimiento de accesorios, así como la clasificación de cada accesorio en correcto faltante o equivocado en la Fig. 1. Dicha metodología inicia con la etapa de adquisición de imágenes, que permite obtener el conjunto de datos iniciales en forma de imagen de la escena donde se encuentra el objeto de análisis, p.e., la unidad de refrigeración con la puerta abierta. Se continúa con la etapa de segmentación que permite obtener imágenes Regiones de Interés, ROI, que delimitan el gabinete y la puerta en cada uno de los tipos de imágenes en las diferentes vistas. Esta etapa es diferente para gabinete y puerta debido a que para el primer caso el fondo se tiene controlado y en el otro solo parcialmente.

Luego, una etapa de seguimiento de permite centrar la atención del análisis en una de las neveras que pasa en frente del campo de visión del sistema hasta que esta llega a una posición estratégica en la imagen de todos los sensores. Por último, mediante etapas de detección de ROI's y Dictamen basado en puntos característicos, se detecta y clasifica cada accesorio interno de la nevera. Para estas últimas etapas se utiliza información del ROI del gabinete y de la puerta y de los accesorios registrados previamente los cuales se encuentran en una base de datos para cada modelo. A continuación, se describe cada una de las etapas de la metodología propuesta y se finaliza con los detalles de implementación para la realización de pruebas experimentales.

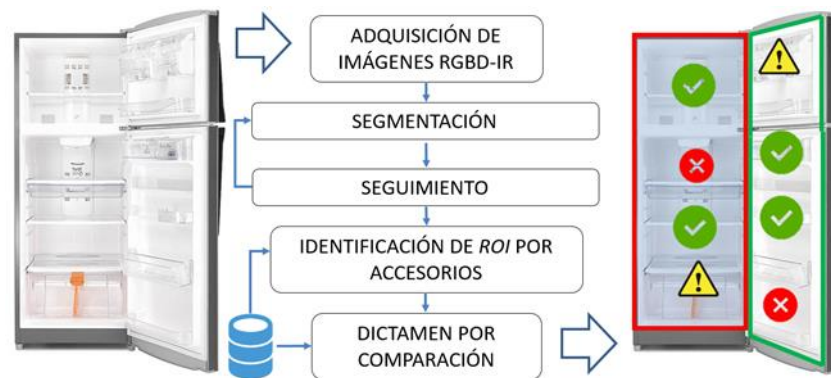


Fig. 1: Metodología para la detección y seguimiento de accesorios en unidades de refrigeración en línea de producción.

Etapa 1: Adquisición de imágenes

Un análisis detallado de la escena y las restricciones propias de un escenario real de producción determina que un conjunto de cuatro Kinect V2 (Hong, Saavedra, y Martínez-Corral, 2017; Kim, Yun, Jung, y Sun, 2015) (dos para gabinete y dos para puerta) son necesarios para poder capturar imágenes adecuadas para la detección, seguimiento y análisis de los accesorios. Es decir, el bloque de adquisición responde a necesidades no solo de la problemática de detección y seguimiento, sino también la de clasificación de accesorios. En la Fig. 2 se observa la ubicación y orientación de los cuatro sensores. La elección de los Kinect V2 se debe a que estos permiten disponer de tres tipos de imágenes (imágenes multimodales) por sensor, con las siguientes características: Imagen RGB de 1920x1080 (24 bit), Imagen PROFUNDIDAD de 512x424 (11 bit) y una Imagen NIR de 512x424 (11 bit). La frecuencia máxima de captura de cada sensor es de 30 Hz. Esta frecuencia disminuye cuando se realiza la interfase entre el sensor y una estación de cómputo.

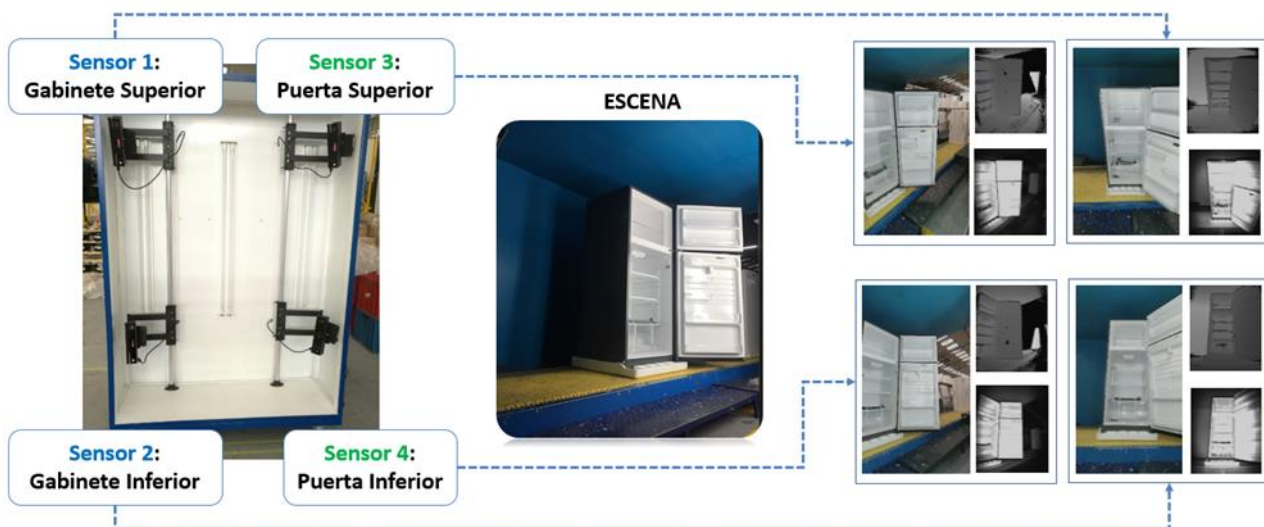


Fig. 2: Sistema de adquisición de 4 sensores Kinect V2 (Izq.) capturando imágenes de la escena (cent.) genera un total 12 imágenes en una sola captura.

Los Kinect V2 superiores (Sensores 1 y 3) se encuentran a 1.6 m de la base de la estructura metálica y los Kinect inferiores (Sensores 2 y 4) a una altura de 0,6 m. Además, nótese en la Fig. 2 que los Kinect se han girado 90° para aprovechar la apertura angular horizontal de los sensores la cual resulta mayor que la vertical, a saber, el FOV horizontal es de 84.70° y el vertical de 54.4° (Hong et al., 2017). Los soportes de los brazos de los Kinect V2 están separados 1.0 m. La configuración anterior obedece a los requerimientos de una línea de producción que trabaja modelos de neveras pequeñas (1.5 m de altura) y grandes (2 m de altura) y una separación vertical de la banda a los Kinect de 1.7 m. Detrás de las neveras se ha ubicado un fondo polietileno opaco (Demant, Streicher-Abel, y Garnica, 2013) para evitar que fuentes parásitas infrarrojas lleguen de forma directa a los sensores de gabinete, sin embargo, este fondo no ayuda a controlar la presencia de fuentes parásitas hacia los sensores de la puerta, esto genera la necesidad de concebir dos técnicas diferentes para segmentación. Con ayuda de lámparas LED blancas se controla la iluminación en el espectro visible, con el fin de evitar que el sensor realice balance de blancos de forma automática en la imagen RGB.

Dado que a la fecha el SDK de Kinect V2 para Windows (Microsoft, 2018) no soporta la captura de simultánea de dos sensores en una misma estación de cómputo, se gestiona entonces la captura de imágenes de estos sensores con ayuda de la biblioteca multiplataforma Libfreenect2 (Xiang et al., 2016) en el lenguaje de programación C++. Esta biblioteca captura la trama de datos enviadas por el sensor a través del protocolo USB 3.0 (utilizando LibUSB 1.0 (Organization, 2016) hacia el computador, y con ayuda de tarjetas gráficas de NVIDIA (CUDA 8.0 (Nickolls, Buck, Garland, y Skadron, 2008)) se decodifica finalmente las imágenes RGB, PROFUNDIDAD y NIR.

Etapa 2: Segmentación

Tal como se mencionó en la sección anterior, se utiliza un fondo de polietileno opaco para controlar la presencia de fuentes parásitas en los sensores de gabinete. Este fondo puede aprovecharse para implementar una técnica de sustracción de fondo (Elgammal et al., 2014; Klette, 2014) sobre la imagen de profundidad (ecuación (1)), utilizando un modelo de media y desviación estándar de la escena sin la presencia de ninguna nevera. Gracias a esto, en la Fig. 3 se propone un conjunto de técnicas para segmentar el gabinete tanto superior e inferior utilizando la ya mencionada técnica de sustracción de fondo.

$$BG(x, y) = \begin{cases} 1, & \text{si } |D(x, y) - MD(x, y)| \geq SD(x, y) \\ 0, & \text{en caso contrario} \end{cases} \quad (1)$$

Donde BG (x, y) es la imagen binaria de salida de gabinete, D(x, y) es la imagen de profundidad, MD(x, y) y SD(x, y) son modelo promedio y desviación estándar del fondo obtenido a partir de 100 imágenes D(x, y). La ecuación (1) se aplica para gabinete superior e inferior. Los niveles de gris en la imagen de profundidad son normalizados con respecto a su valor máximo posible, el cual es obtenido mediante una función que provee el SDK de Microsoft para el Kinect V2.

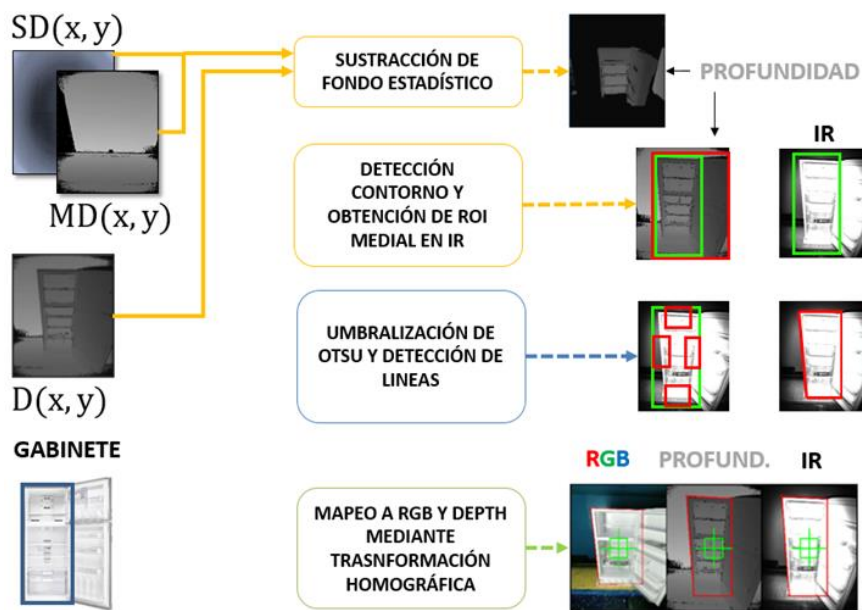


Fig. 3: Flujo secuencial de técnicas utilizadas para la segmentación de Gabinete.

La técnica de sustracción de fondo para gabinete no puede utilizarse para puertas, ya que en este caso no es posible ubicar un fondo sin obstaculizar el paso de las neveras a lo largo de banda. Por tal razón, se utiliza el siguiente criterio de umbralización fija (Szeliski, 2011) tanto en la imagen NIR como en la de profundidad: un pixel es clasificado como perteneciente a la nevera si se encuentra cercano en la imagen de profundidad y es suficientemente brillante dentro de un rango determinado en la imagen infrarroja. Dado que el criterio involucra dos tipos de imágenes, la técnica de umbralización así implementada se denomina umbralización multimodal, ecuación (2). En la Fig. 4 se observa el conjunto de técnicas que complementan el método de segmentación de puertas.

$$BP(x, y) = \begin{cases} 1, & \text{si } D1 \leq D(x, y) \leq D2 \wedge N1 \leq NIR(x, y) \leq N2 \\ 0, & \text{en caso contrario} \end{cases} \quad (2)$$

Donde BP(x, y) es la imagen binaria de salida del gabinete, D(x, y) es la imagen de profundidad, NIR(x, y) es la imagen infrarroja. D1, D2, N1 y N2 son valores fijos en el intervalo [0, 1]. La ecuación (2) se aplica para puerta superior e inferior.

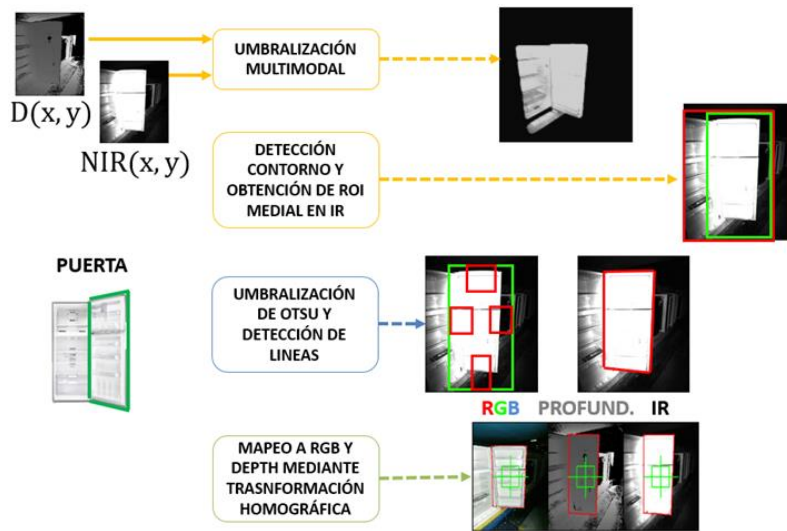


Fig. 4: Flujo secuencial de técnicas utilizadas para la segmentación de Puerta.

Obsérvese en las Fig. 3 y Fig. 4 que los métodos de segmentación comparten un mismo conjunto de técnicas después de: 1) realizar la sustracción de fondo para gabinete y 2) la umbralización multimodal para puertas. En primer lugar, la detección de contornos y aproximación medial, permite obtener una primera aproximación del ROI que contiene gabinete o puerta, utilizando un algoritmo de etiquetado con conectividad 8 (Demant et al., 2013; Szeliski, 2011) y una aproximación medial del ROI. El etiquetado mediante componentes conexos con conectividad 8 tiene como propósito obtener las etiquetas, L , de los n contornos detectados y el conjunto, C , de $m(L)$ puntos 2D, (x_i^L, y_i^L) , por cada etiqueta L . Ver Fig. 5. El ROI por cada etiqueta L , es decir $ROI(L)$, se define mediante un conjunto de 4 puntos 2D obtenidos a través la ecuación (3).

$$ROI(L) = \{([x^*]_k^L, [y^*]_k^L) \mid k = 1, 2, 3, 4\} \quad (3)$$

Donde:

$$([x^*]_1^L, [y^*]_1^L) = (\min\{x_i^L\}, \min\{y_i^L\})$$

$$([x^*]_2^L, [y^*]_2^L) = (\max\{x_i^L\}, \min\{y_i^L\})$$

$$([x^*]_3^L, [y^*]_3^L) = (\max\{x_i^L\}, \max\{y_i^L\})$$

$$([x^*]_4^L, [y^*]_4^L) = (\min\{x_i^L\}, \max\{y_i^L\})$$

Dado que esta primera aproximación del ROI contiene en el caso de gabinete pixeles pertenecientes a la puerta y en el caso de puerta pixeles de gabinete, se procede de manera heurística a redefinir el ROI mediante una aproximación medial, la cual consiste en tomar la mitad del ROI obtenido por cada etiqueta L .

$$ROI_G(L) = \{([x_m^*]_k^L, [y_m^*]_k^L) \mid k = 1, 2, 3, 4\} \quad (4)$$

Donde:

$$([x_m^*]_1^L, [y_m^*]_1^L) = ([x^*]_1^L, [y^*]_1^L)$$

$$([x_m^*]_2^L, [y_m^*]_2^L) = \left(\frac{[x^*]_1^L + [x^*]_3^L}{2}, [y^*]_1^L \right)$$

$$([x_m^*]_3^L, [y_m^*]_3^L) = \left(\frac{[x^*]_1^L + [x^*]_3^L}{2}, [y^*]_3^L \right)$$

$$([x_m^*]_4^L, [y_m^*]_4^L) = ([x^*]_4^L, [y^*]_4^L)$$

$$ROI_P(L) = \{([x_m^*]_k^L, [y_m^*]_k^L) \mid k = 1, 2, 3, 4\} \quad (5)$$

Donde:

$$([x_m^*]_1^L, [y_m^*]_1^L) = \left(\frac{[x^*]_1^L + [x^*]_3^L}{2}, [y^*]_1^L \right)$$

$$([x_m^*]_2^L, [y_m^*]_2^L) = ([x^*]_2^L, [y^*]_2^L)$$

$$([x_m^*]_3^L, [y_m^*]_3^L) = ([x^*]_3^L, [y^*]_3^L)$$

$$([x_m^*]_4^L, [y_m^*]_4^L) = \left(\frac{[x^*]_1^L + [x^*]_3^L}{2}, [y^*]_3^L \right)$$

Este ROI se toma hacia la izquierda en gabinete, ecuación (4), y hacia la derecha en puerta, ecuación (5), si la unidad de refrigeración abre la puerta hacia la derecha en la imagen. Obviamente debe tomarse en sentido contrario en caso que la nevera abra la puerta hacia la izquierda, en la imagen. Luego, dentro de este ROI medial, se obtiene un ROI no necesariamente rectangular y definido por cuatro puntos coordenados 2D en la imagen. Para esto, se identifican las ecuaciones de línea que delimitan el contorno de la nevera por la parte superior, inferior, izquierda y derecha tanto del gabinete como de la puerta. Estas líneas se obtienen utilizando un ajuste lineal mediante mínimos cuadrados (Bertsekas, 2016) sobre el conjunto de puntos frontera, que resulta del borde obtenido mediante un kernel 3x3 de derivada direccional tanto en X como Y, sobre la imagen umbralizada; esta última umbralización se realiza mediante la técnica de Otsu (Liu y Yu, 2009).

Las sub-imágenes son extraídas de la imagen infrarroja y se definen de forma heurística a partir del ROI medial con ayuda de la técnica mostrada en la Fig. 6. Una vez encontrada estas líneas, los cuatro puntos que definen el ROI se obtienen utilizando ecuaciones de punto de intersección entre líneas (Jänich, 1994): líneas superior e izquierda para la esquina superior-izquierda del ROI, líneas superior y derecha para el punto la esquina superior-derecha del ROI, líneas derecha e inferior para el punto inferior-derecho del ROI, y entre las líneas inferior e izquierda para el punto de la esquina inferior-izquierda del ROI.

Para el caso de gabinete no es posible obtener una ecuación de línea ubicada al costado derecho con la técnica mencionada anteriormente (que por cierto separa gabinete de puerta en la imagen). En este caso se toma como línea derecha, aquella obtenida en el ROI medial. Se procede de igual forma para la línea izquierda en puerta. Finalmente, las coordenadas de los puntos obtenidos en la imagen NIR se toman como los mismos en la imagen de PROFUNDIDAD (se considera despreciables las distorsiones en las imágenes del Kinect V2), y mediante técnicas homográficas (Hartley y Zisserman, 2004) se mapean estos puntos a la imagen RGB. La relación homográfica entre puntos coplanares desde dos vistas (este caso NIR y RGB) viene dado por la ecuación (6) utilizando coordenadas homogéneas asociadas a los puntos 2D. La matriz $H = h_{ij}$ ($i, j = 1, 2, 3$) es la homografía. Dado que H contiene 9 incógnitas, se propone el sistema de ecuaciones (6.2) para cuatro pares de puntos correspondientes, y se resuelve el sistema por descomposición de valores singulares (Hartley y Zisserman, 2004; Jänich, 1994) de la matriz A , ecuación (7).

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x'_1 & -y_1y'_1 & -x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y'_1 & -y_1y'_1 & -y'_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2x'_2 & -y_2y'_2 & -x'_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2y'_2 & -y_2y'_2 & -y'_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3x'_3 & -y_3y'_3 & -x'_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3y'_3 & -y_3y'_3 & -y'_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4x'_4 & -y_4y'_4 & -x'_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4y'_4 & -y_4y'_4 & -y'_4 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (7)$$

$\underbrace{\hspace{15em}}_A$

Etapa 3: Seguimiento.

La técnica de segmentación generará tantos ROI's como gabinetes y puertas sean segmentadas mediante la sustracción de fondo o la Umbralización multimodal respectivamente. Por esto, con ayuda de los ROI's, se construyen contornos los cuales son etiquetados nuevamente mediante componentes con conectividad 8 (Szeliski, 2011). Ver Fig. 5.

Se define como posición óptima el centro de la imagen y entonces se elige el contorno de máxima área cuyo centroide está más cercano al centro de la imagen por el lado izquierdo. En caso que las neveras entren por el lado derecho, debe aplicarse el criterio anterior pero ahora por este nuevo lado. Las etapas claves de este bloque se muestran en la Fig. 7.

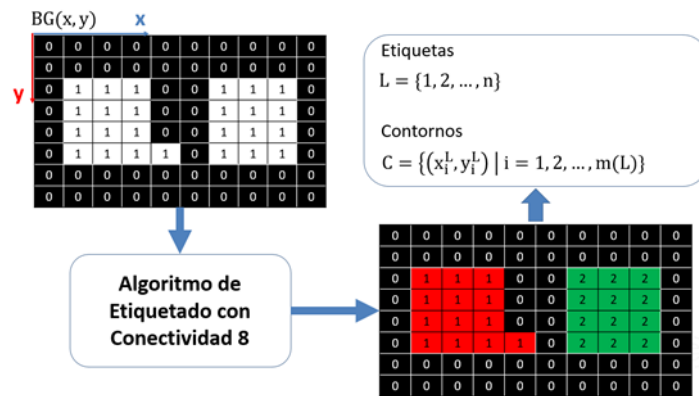


Fig. 5: Resultado aplicación algoritmo de etiquetado sobre una imagen binaria. La salida de este algoritmo son las etiquetas, L, de los n contornos detectados y el conjunto, C, de m(L) puntos 2D, (x_i^L, y_i^L) , por cada etiqueta L.

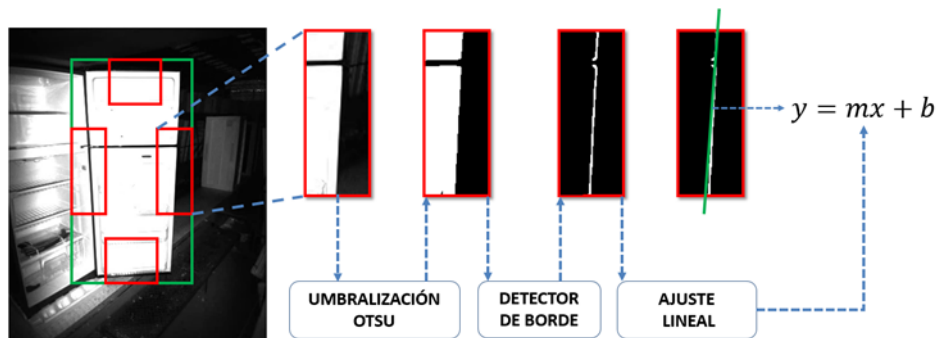


Fig. 6: Técnica de obtención de ecuación de línea sobre borde de nevera en imagen infrarroja.

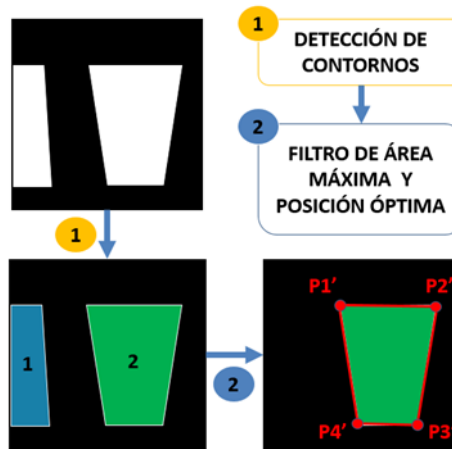


Fig. 7: Etapas claves para la selección del ROI de interés en el bloque de seguimiento.

Etapas clave para la selección del ROI de interés en el bloque de seguimiento.

Hasta ahora, los bloques de adquisición, segmentación y seguimiento se han elaborado independiente del modelo de unidad de refrigeración que se está analizando en el campo de visión de los sensores. Dado que el número de accesorios de un modelo a otro varía y que las ubicaciones y tipos de accesorios es una variable no controlada, se requiere entonces, para la obtención de los ROI's de los accesorios, conocer con anterioridad el modelo de unidad de refrigeración para la prueba. Esto se logra almacenando información de los modelos y accesorios en una base de datos a través de un proceso de registro del modelo. Por cada modelo se requiere el ROI que delimita el gabinete y la puerta, y los ROI's y etiquetas de cada uno de los accesorios que se deseen analizar. En la Fig. 8 se muestra las variables que se dispone para realizar la obtención de los ROI's de los accesorios para la prueba final. Suponiendo N accesorios, la estimación de los ROI's de los accesorios en la unidad de refrigeración en cuestión se obtiene mediante el siguiente algoritmo, ver Tabla 1. basado en transformación homográfica (Hartley y Zisserman, 2004) y el cual se planteó en las ecuaciones (6) y (7).

Tabla 1: Algoritmo de estimación de ROI's para accesorios de prueba a partir de los ROI's en la referencia.

Entrada	$\{P_i\}, \{PC_{ij}\}, \{P'_i\} : i = 1, \dots, 4; j = 1, \dots, N.$ <ul style="list-style-type: none"> - $\{P_i\}$: ROI del gabinete o puerta en la imagen de referencia. - $\{P'_i\}$: ROI del gabinete o puerta en la imagen de prueba. - $\{PC_{ij}\}$: ROI de los accesorios en la imagen de referencia. - N: Número de accesorio registrados en la imagen de referencia.
Salida	$\{PC'_{ij}\} : i = 1, \dots, 4; j = 1, \dots, N.$ <ul style="list-style-type: none"> - $\{PC'_{ij}\}$: ROI de los accesorios en la imagen de prueba. - N: Número de accesorio registrados en la imagen de referencia.
Algoritmo	<ol style="list-style-type: none"> 1: $H =$ homografía ($\{P_i\}, \{P'_i\}$); 2: $\{PC'_{ij}\} =$ inicializa(4, N); 3: Para $j = 1$ hasta N 4: Para $i = 1$ hasta 4 5: $X = [\{PC_{ij}\}.x, \{PC_{ij}\}.y, 1]^T$; 6: $X_{out} = HX$; 7: $\{PC'_{ij}\}.x = X_{out}.x$; 8: $\{PC'_{ij}\}.y = X_{out}.y$; 9: Fin Para 10: Fin Para 11: Retorna $\{PC'_{ij}\}$;

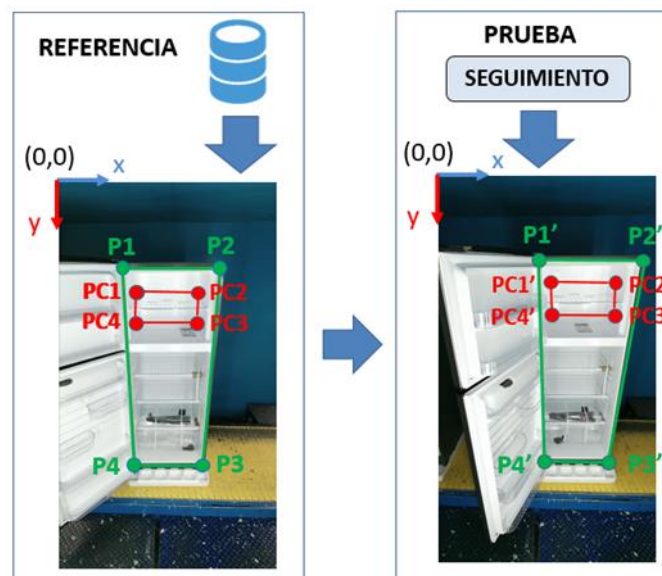


Fig. 8: Los puntos $\{P_i\}$, $\{P'_i\}$ y $\{PC_i\}$ son conjuntos de puntos conocidos: $\{P_i\}$ y $\{PC_i\}$ provienen de la base de datos de la referencia registrada, $\{P'_i\}$ proviene del resultado de la etapa de seguimiento de la nevera que bajo prueba. $\{PC'_i\}$ es el conjunto de puntos desconocidos del accesorio en la nevera de prueba. ($i = 1, 2, 3$ y 4).

Etapa 5: Dictamen basado en Puntos Característicos.

El dictamen de cada accesorio se lleva a cabo mediante la comparación de puntos característicos tipo ASIFT (Yu y Morel, 2011) de los ROI's de los accesorios tanto en el modelo de referencia como en la prueba, ver Fig. 9. A diferencia de los puntos característicos tipo SIFT (Lowe, 1999), los puntos característicos ASIFT además de ser invariantes a escala, son invariantes a transformaciones de perspectiva en la imagen. Este último resulta muy útil en la propuesta metodológica aquí desarrollada ya que no siempre logra obtenerse una imagen del accesorio en la prueba totalmente alineada con la imagen del accesorio en el modelo de referencia. La idea clave para determinar el grado de similitud entre el accesorio de referencia y el de prueba es hallar NR (número de puntos característicos en la referencia) y NP (número de puntos característicos en la prueba). Primero, NR se determina así: 1) se obtiene el centroide de los descriptores asociados a los puntos característicos en el accesorio de referencia, 2) se calcula la mínima y máxima distancia euclidiana de los vectores descriptores asociados a cada punto característico respecto al centroide; 3) aquellos cuya distancia está entre $[\min N, \alpha * (\max N - \min N)]$ (donde $\min N$ y $\max N$ son la mínima y máxima distancia en el espacio de descriptores respecto al centroide) son seleccionados como los puntos de mayor peso.

Es importante mencionar que cada descriptor asociado a cada punto se obtiene utilizando un vecindario de 16x16 píxeles alrededor de cada punto. Este, a su vez, se divide en sub-bloques de tamaño 4x4, y para cada uno se crea un histograma de orientaciones tal como se describe en (Lowe, 1999). La concatenación en un vector de los valores de cada histograma para los 16 sub-bloques del punto de interés constituye su descriptor. Luego, los NP puntos asociados a la prueba se determinan de una manera similar a la anterior: 1) se mide la distancia euclidiana en el espacio de descriptores de cada uno de los NR puntos obtenidos en el accesorio de referencia con cada uno de los puntos característicos tipo ASIFT obtenidos para el accesorio de la prueba. 2) si esa distancia se encuentra en el rango $[\min N, \beta^*(\max N - \min N)]$, entonces es seleccionado como punto correspondiente. En caso de existir más de un punto en el accesorio de la prueba asociado a un punto en el accesorio de referencia, se seleccionará la correspondencia entre puntos de menor distancia. Para las pruebas realizadas se fijó de forma heurística $\alpha = \beta = 0.7$.

De lo descrito anteriormente se deduce que $NP < NR$; en consecuencia se propone el grado de similitud entre los accesorios, MSA, de forma porcentual, ecuación (8).

$$MSA = \frac{NP}{NR} \times 100\% \tag{8}$$

Pruebas preliminares realizadas sobre dos modelos de neveras permiten identificar los siguientes umbrales para dictaminar el accesorio en correcto, faltante o equivocado, Tabla 2.

Tabla 2: Rangos de valores de MSA para dictámenes.

Rango MSA	Dictamen
0 – 10	Faltante
11 – 30	Equivocado
31 -100	Correcto

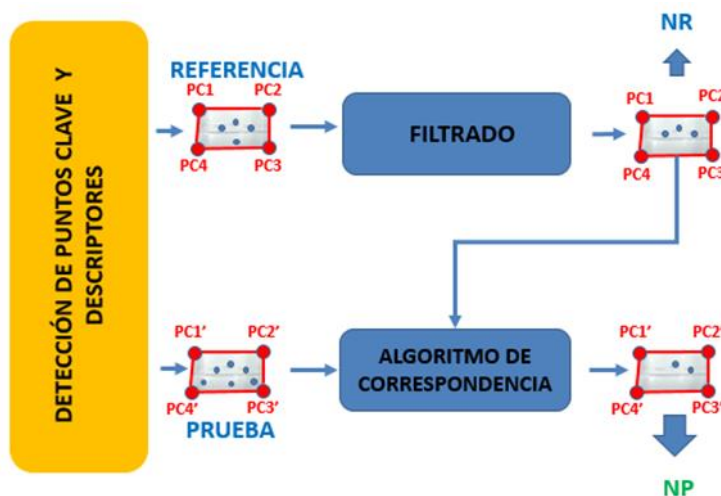


Fig. 9: Algoritmo de dictamen mediante comparación de puntos característicos.

Detalles de Implementación

Para la implementación de los algoritmos involucrados en las diferentes etapas, se identifican bibliotecas y rutinas software con soporte CPU o GPU. Se selecciona la implementación que mejor desempeño ofrezca para cada caso, ver Tabla 3. Se propone la arquitectura de la Fig. 10 para la ejecución de los algoritmos de visión por computador, los cuales fueron implementados en C/C++ y CUDA. La etapa de adquisición, como ya se mencionó, se implementó utilizando la biblioteca Libfreenect 2 la cual trae soporte GPU, y las implementaciones de las otras etapas se hizo utilizando OpenCV 3.2 (Howse, Puttemans, Hua, y Sinha, 2015), para la implementación de ASIFT en C++ se utilizó la reportada en (Yu y Morel, 2011). Cada Workstation (Intel Core i7, 3.40 GHz, 32 GB RAM, NVIDIA Quadro M5000) ejecuta los algoritmos para puerta y gabinete; a su vez, para cada Kinect V2 se ejecuta un proceso diferente en el sistema operativo Windows 10 para que la captura de las imágenes sea cuasi-simultánea. Los resultados son enviados a un equipo de visualización el cual envía los resultados a una base de datos. En esta base de datos se encuentran almacenados los modelos de referencia. Un lector de código de barras permite la captura automática del serial y referencia de una nevera particular.

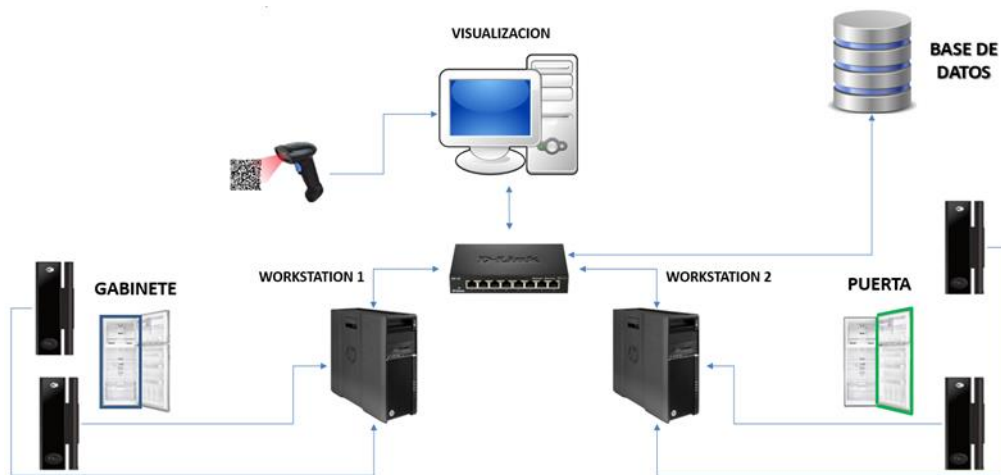


Fig. 10: Plataforma computacional: arquitectura basada en Workstations CPU/GPU.

Tabla 3: Especificación de las implementaciones para cada una de las etapas del sistema de visión por computador.

Etapa	Compartimento	
	Gabinete	Puerta
Adquisición	Basada en Libfreenect 2 (GPU)	
Segmentación	<ul style="list-style-type: none"> Sustracción de fondo estadístico DEPTH (GPU) Segmentación geométrica + mapeo homográfico (CPU) 	<ul style="list-style-type: none"> Umbralización multimodal DEPTH + IR (GPU). Segmentación geométrica + mapeo homográfico (CPU)
Seguimiento	Detección de contornos + criterio de máxima área + posición óptima. (CPU)	
Detección ROI	Mapeo de ROI's mediante transformación homográfica (CPU)	
Dictamen	Comparación de descriptores asociados a puntos claves invariantes a escala transformaciones de perspectiva (CPU/GPU)	

RESULTADOS Y DISCUSIÓN

Para evaluar la propuesta metodológica se registran 3 modelos: MCRMU300WANB (nevera pequeña: 1.5 m), MCRMC320WACX (nevera mediana 1.7 m), MCRMC390WACX (nevera grande 2.0 m). Para el modelo MCRMU300WANB se registran 6 accesorios y 7 para los otros dos modelos. En la parte superior de la Fig. 11, Fig. 12 y Fig. 13 se muestran los accesorios registrados para cada uno de los sensores en los 3 modelos. En la Tabla 3 se recoge el número de pruebas y porcentajes de acierto obtenidos para cada uno de los sensores. Se observa de esta que los algoritmos para gabinete obtuvieron mejor resultado que los de puerta. Esto se debe a que en general el algoritmo de segmentación de gabinete tuvo mejor desempeño que el de puertas durante las pruebas realizadas.

Con los umbrales definidos en la Tabla 2 se construye la matriz de confusión de la Tabla 5 y se obtiene sensibilidades y especificidades para cada caso: equivocado, faltante y correcto. Cabe resaltar que estas pruebas fueron llevadas a cabo en compañía de un experto en calidad de neveras quien aportó su dictamen en cada caso. Se puede ver, en primer lugar, que ningún valor de sensibilidad es superado por 100 – Especificidad, esto indica que el sistema está clasificando (no adivinando) y lo hace en general con un porcentaje del 70 %. En segundo lugar, se observa que las precisiones fueron buena (90 %), regular (73 %), y deficiente (23 %) en la clasificación de accesorio correcto, faltante y equivocado, respectivamente. Esto indica que el sistema tiene mayor precisión en la clasificación de accesorios correctos y faltantes que equivocados. Se pudo identificar además que estos resultados dependen más de la etapa de segmentación que de las otras etapas ya que la principal fuente de error provenía de una segmentación incorrecta de la nevera o de los accesorios en las imágenes RGB-D-NIR. Esto sugiere que debe realizarse un estudio más amplio para identificar técnicas más robustas de segmentación.

Aun cuando en el trabajo mostrado en (Sudeep, 2015) se reporta una precisión del 98 % en el dictamen de accesorios faltantes, no es posible la comparación formal de los resultados aquí obtenidos con estos ya que no se evidencia el método ni el diseño de experimentos llevados a cabo para la obtención de este resultado. Esto por el contrario permite resaltar que el trabajo aquí realizado es de los primeros en visión por computador para la clasificación de accesorios internos en refrigeradores formalmente expuestos ante la comunidad científica.

Por último, con el fin de evaluar la viabilidad del sistema en un escenario de producción real, se midieron los tiempos de realización de prueba sobre una línea rápida de producción (35 s/nevera), además del nivel de memoria RAM ocupada por el algoritmo y los niveles de ocupación de CPU y GPU, en dos arquitecturas diferentes. En la Tabla 6 se recogen los resultados promedios de 120 pruebas realizadas para estas dos arquitecturas. Se observa que ambas arquitecturas tienen tiempos de realización de prueba por debajo de los tiempos necesarios de producción. Esto significa que no es necesario detener la nevera para su análisis, convirtiéndolo en un sistema que no afecta la cadencia de la línea. A su vez nótese que en ambos casos es posible una implementación acelerada por GPU de aquellos algoritmos que se implementaron en CPU, ya que la GPU no está siendo completamente utilizada, además, la ejecución de los aplicativos no supera 2GB RAM.

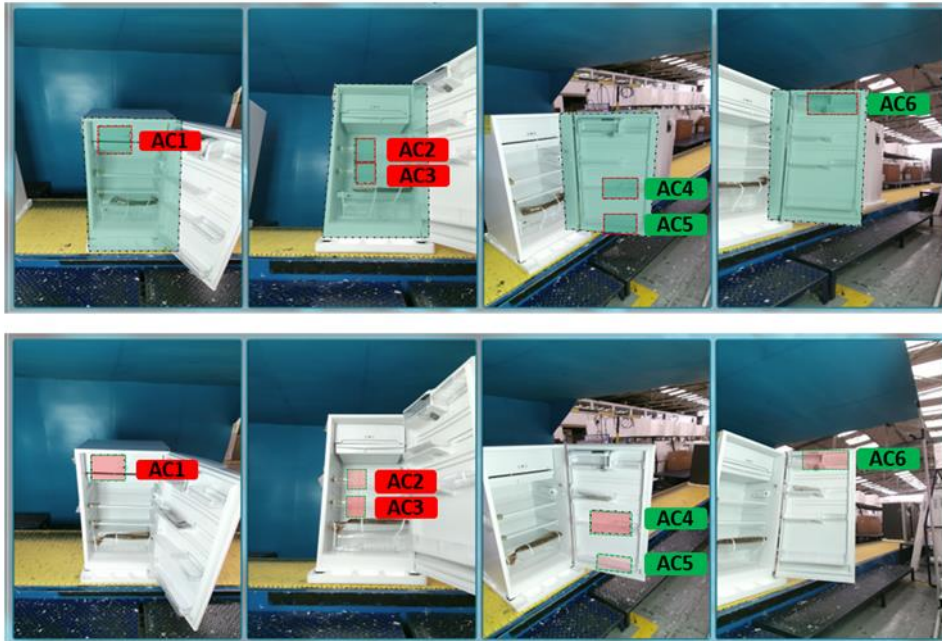


Fig. 11: Resultado de obtención de ROI's de accesorios en posición central de la imagen. Prueba realizada con una nevera perteneciente al modelo MCRMU300WANB. Imágenes del modelo de referencia en la parte superior y de la prueba en la parte inferior.

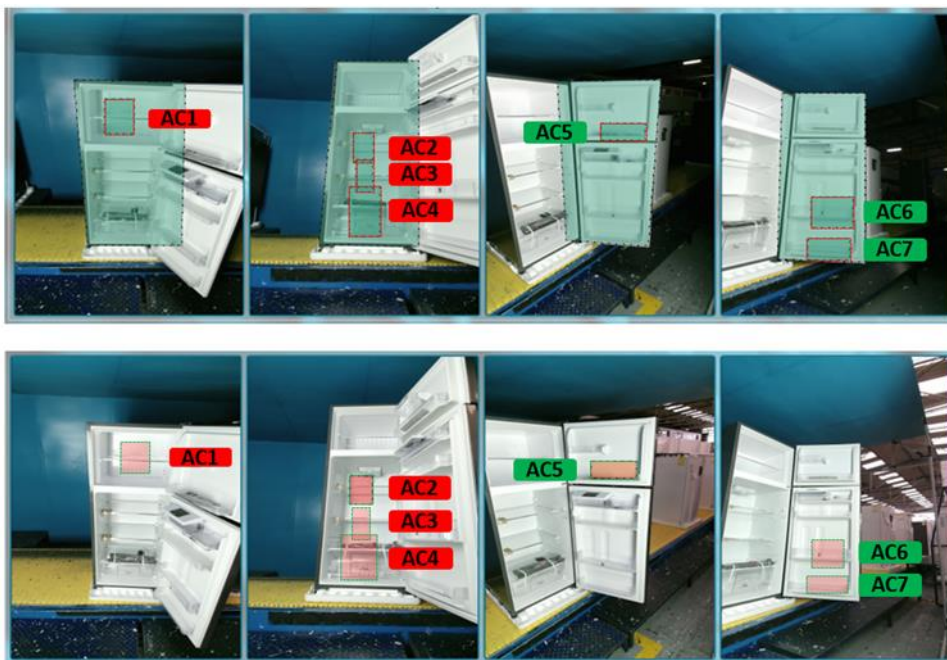


Fig. 12: Resultado de obtención de ROI's de accesorios en posición central de la imagen. Prueba realizada con una nevera perteneciente al modelo MCRM320WACX. Imágenes del modelo de referencia en la parte superior y de la prueba en la parte inferior.



Fig. 13: Resultado de obtención de ROI's de accesorios en posición central de la imagen. Prueba realizada con una nevera perteneciente al modelo MCRMC390WACX. Imágenes del modelo de referencia en la parte superior y de la prueba en la parte inferior.

Tabla 4: Resultados para cada sensor en el proceso de dictamen de accesorio.

<i>Resumen Dictámenes por Sensor Kinect V2</i>	<i>Número de Aciertos</i>	<i>Número de Pruebas</i>	<i>Porcentaje de Acierto</i>
Gabinete Inferior	173	240	72%
Gabinete Superior	70	90	78%
Puerta Inferior	105	150	70%
Puerta Superior	78	120	65%
Total	426	600	71%

Tabla 5: Matriz de confusión y valores de sensibilidad y especificidad para los dictámenes del sistema desarrollado respecto al dictamen de un experto.

<i>Dictamen Sistema</i>	<i>Dictamen Experto</i>			<i>Total Pruebas</i>	<i>Sensibilidad (%)</i>	<i>Especificidad (%)</i>
	<i>Equivocado</i>	<i>Faltante</i>	<i>Correcto</i>			
Equivocado	23	29	5	100	23	93
Faltante	28	146	25	200	73	87
Correcto	49	25	270	300	90	87
Porcentaje Clasificación Total				70		

Tabla 6: Desempeño del sistema en 2 arquitecturas computacionales.

<i>Arquitectura Computacional</i>	<i>Tiempo Promedio Realización Prueba (s)</i>	<i>RAM (GB)</i>	<i>Ocupación CPU (%)</i>	<i>Ocupación GPU (%)</i>
A1: Intel(R) Xeon, 12 Núcleos, 2.40 GHz, 16 GB RAM, Nvidia Quadro K 2200.	29	1.1	60	15
A2: Intel(R) Core I7, 8 Núcleos, 3.40, GHz, 32 GB RAM, Nvidia Quadro M5000	24	1.2	70	10

CONCLUSIONES

Se diseñó e implementó una metodología de visión por computador para la detección, seguimiento y clasificación de accesorios internos en unidades de refrigeración en tiempo real para control de calidad. El sistema propuesto tiene mayor precisión en la clasificación de accesorios correctos y faltantes que equivocados. A pesar de esto, se destaca que la rapidez de la ejecución de los algoritmos en las arquitecturas computacionales fue exitosa ya que no introducen cadencia a la línea de producción. Esto permitirá concebir la utilización de técnicas más robustas de segmentación las cuales seguramente aumentaran el tiempo general de los algoritmos, pero mejorarán la precisión del sistema.

REFERENCIAS

- Bertsekas, D., *Nonlinear Programming*, 3ª Ed., 102-119, Athena Scientific, Massachusetts, United States (2016)
- Bowmans, T., F. Porikli, B. Höferlin y A. Vacavant, *Background Modeling and Foreground Detection for Video Surveillance*, CRC Press (2014)
- Demant, C., B. Streicher-Abel y C. Garnica, *Industrial Image Processing*, 25-63, Springer-Verlag Berlin Heidelberg, Berlin, Germany (2013)
- Harrington, L., *A Comparative Assessment of Refrigerator Test Methods*, Proceedings of the 2001 Summer Study of the European Council for an Energy Efficient Economy (ECEEE), 401-414 (2001)
- Hartley, R. y A. Zisserman, *Multiple View Geometry in Computer Vision*, 2ª Ed., 87-127, Cambridge University Press, Cambridge, United Kingdom (2004)
- Hong, S., G. Saavedra y M. Martínez-Corral, *Full Parallax Three-dimensional Display from Kinect v1 and v2*, doi: 10.1117/1.OE.56.4.041305, *Optical Engineering*, 56(4), 041305 (2017)
- Howse, J., S. Puttemans, Q. Hua y U. Sinha, *OpenCV 3 Blueprints*, PACKT Publishing Ltd (2015)
- Jänich, K., *Linear Algebra*, 1ª Ed., 15-42, New York, Springer New York, United States (1994)
- Khangura, A., S. Gandhi, *Design and Development of the Refrigerator with Quality Function Deployment Concept*, *International Journal on Emerging Technologies*, 3(1), 173-177 (2012)
- Kim, C., S. Yun, J. Seung-Won y C. Sun, *Color and Depth Image Correspondence for Kinect V2*, doi: 10.1007/978-3-662-47487-7, *Lecture Notes in Electrical Engineering*, Springer Berlin Heidelberg, 352, 111-116 (2015)
- Klette, R., *Concise Computer Vision*, 1ª Ed., Springer-Verlag London, United Kingdom (2014)
- Liu, D. y J. Yu, *Otsu Method and K-means*, doi:10.1109/HIS.2009.74, *Ninth IEEE International Conference on Hybrid Intelligent Systems*, 1, 344–349 (2009)
- Lowe, D.G., *Object Recognition from Local Scale-invariant Features*, doi: 10.1109/ICCV.1999.790410, *Proceedings of the Seventh IEEE International Conference on Computer Vision*, 2, 1150–1157 (1999)
- Microsoft C., *Kinect for Windows SDK 2.0*, Microsoft Corporation (2018)
- Nickolls, J., I. Buck, M. Garland y K. Skadron, *Scalable Parallel Programming with CUDA*, doi: 10.1145/1365490.1365500, *Queue*, 6(2), 40 (2008)
- Organization, L., *LibUSB 1.0*, Sourceforge, (2016)
- Sudeep, S., *Refrigerator Inspection: LG Electronics*, Spookfish (2015)
- Szeliski, R., *Computer Vision: Algorithms and Applications*, 1ª Ed., 181-234, Springer-Verlag London, UK (2011)
- Xiang, L., F. Echtler, C. Kerl, T. Wiedemeyer, R. Gordon y F. Facioni, *Libfreenect2: Release 0.2*, doi: 10.5281/zenodo.594510, Zenodo (2016)
- Yu, G. y J. Morel, *ASIFT: An Algorithm for Fully Affine Invariant Comparison*, doi: 10.5201/ipol.2011.my-asift, *Image Processing on Line*, 1, 11–38 (2011)

