

Original research / Artículo original / Pesquisa original - Tipo 1

# 3D reconstruction system for semi-automatic estimation of objects length and area by means of stereo vision

**Bryan García**<sup>1,2</sup> / bryan.garcia@correounivalle.edu.co

**Carlos Diego Ferrin B.**<sup>1</sup> / carlos.ferrin@correounivalle.edu.co

**Jorge Humberto Erazo Aux,**<sup>1,2</sup> / jorge.erazo@correounivalle.edu.co

Institución Universitaria Antonio José Camacho, Cali-Colombia<sup>1</sup>

Universidad del Valle, Cali – Colombia<sup>2</sup>

**ABSTRACT** It is mandatory to characterize dimensionally the manufactured industrial pieces for quality control purposes. As it is not possible to touch some pieces when trying to retrieve dimensional information, then non-invasive techniques are required to do so. Stereo vision is a passive technology which is both robust and accurate for non-invasive applications. For this reason, in this work we describe the design and implementation of a 3D reconstruction system for the estimation of the length and area of certain objects. This tool allows to easily incorporate new image correspondence techniques to its main execution pipeline. We carry some experiments and show certain benefits when selecting an accurate image correspondence technique for the estimation of the length and area.

**KEYWORDS** Stereo vision; 3D reconstruction; digital image processing.

Sistema de reconstrucción 3D para la estimación semi-automática de la longitud y el área de objetos mediante visión estéreo

**RESUMEN** En los entornos de manufactura de piezas industriales es necesaria la obtención de medidas básicas, como son la longitud y el área, en procesos de control de calidad. Este proceso se dificulta cuando no se puede tener contacto con las piezas y se debe recurrir a métodos de no-contacto. La visión estéreo es una técnica pasiva, robusta y confiable para llevar a cabo este tipo de tareas no invasivas. En este trabajo se describe la construcción de un sistema de visión estéreo para reconstrucción 3D y estimación de longitud y área de objetos, esta herramienta permite incorporar fácilmente técnicas del estado del arte en correspondencia de imágenes a su pipeline de ejecución principal. Las pruebas realizadas sobre algunos objetos muestran los beneficios de disponer de técnicas de correspondencia de imágenes cada vez más exactas.

**PALABRAS CLAVE** Visión estéreo; reconstrucción 3D; Procesamiento digital de imágenes.

Sistema de reconstrução 3D para a estimativa semiautomática do comprimento e da área de objetos utilizando visão estéreo

**RESUMO** Em processos de controle de qualidade na fabricação de peças industriais, é necessária a obtenção de medidas básicas, tais como o comprimento e a área. Este procedimento fica mais difícil quando não é possível ter contato com as peças e devem se utilizar métodos sem contato. A visão estéreo é uma técnica passiva, robusta e confiável para executar este tipo de tarefas não invasivas. Este trabalho descreve a construção de um sistema de visão estéreo para reconstrução em 3D e a estimativa do comprimento e da área de objetos; esta ferramenta permite facilmente incorporar técnicas do estado da arte em correspondência de imagens ao seu pipeline de execução principal. Os testes realizados em alguns objetos mostram os benefícios de ter técnicas de correspondência de imagens cada vez mais precisas.

**PALAVRAS-CHAVE** Visão estéreo; reconstrução 3D; Processamento digital de imagens.

## I. Introduction

In the manufacturing of industrial pieces, the obtaining of basic measurement as length and area are required for quality control processes (Demant, Streicher-Abel, & Garnica, 2013; Hagen et al., n.d.; Krig, 2014). In certain pieces, this process is more difficult, since the physical contact with them is not allowed; hence, contactless methods to carry on this task are required and the stereo vision (Wöhler, 2013) is an alternative for this issue.

The stereo vision is a widely used technique in the computer vision field (Szeliski, 2011), punctually, it is used to estimate the position in the three-dimensional space where an object is located (Hamzah & Ibrahim, 2016). A stereo vision system consists of a pair of cameras horizontally located, each next to the other one. The captured images are processed with the purpose of recovering the information of the deepness in the scene (Bodkin, 2012). The challenge is mainly related with the ability to determine the best method to approximate the differences between the two views to map the correspondence (the disparity map) of the scene. Intuitively, a disparity map represents the corresponding pixels that are horizontally rearranged between the two images. Each year, new methods to solve this problem arise, focusing mainly in the accuracy and computing time elapsed (Hamzah & Ibrahim, 2016).

The time-of-flight [ToF] and the structured light based cameras are active alternative technologies in the industry; however, these devices present some problems (Hansard, Lee, Choi, & Horaud, 2013) as that they are exposed to systematic errors such as noise and ambiguity; which are related to the particular sensor used (e.g., Microsoft's Kinect). Second, they also suffer non-systematic errors such as the dispersion and the soft focus caused by movement (i.e., motion blur). According to the comparative studies performed by Foix & Aleny (2011), Kim, Ayaz, Park, & Roh (2014), and Zhang, Wang, & Chan (2013), the ToF and structured light sensors work flawlessly in certain ranges (approximately 5 to 7 meters) and they are very sensible to be used in outdoors, especially in highly illuminated places. Due to these limitations, the stereo vision systems (passive) are more robust and trustworthy, they are able to produce high-resolution disparity maps, and they are also appropriate for its use both in indoors and outdoors. In order to obtain better results, the assistance of a person that solves the disparities estimation is still required.

## I. Introducción

En los entornos de manufactura de piezas industriales es necesaria la obtención de medidas básicas, como longitud y área, para procesos de control de calidad (Demant, Streicher-Abel, & Garnica, 2013; Hagen et al., n.d.; Krig, 2014). En algunas piezas se dificulta este proceso debido a que no se puede tener contacto con ellas, por lo que se necesitan métodos de no-contacto, la visión estéreo (Wöhler, 2013) es una alternativa ante esta dificultad.

La visión estéreo es una técnica muy utilizada dentro del campo de la visión por computador (Szeliski, 2011) para estimar la posición en el espacio tridimensional a la que se encuentra un determinado objeto (Hamzah & Ibrahim, 2016). Un sistema de visión estéreo consiste, básicamente, en un par de cámaras ubicadas horizontalmente, una al lado de la otra. Las dos imágenes capturadas por estas dos cámaras son procesadas para recuperar información de profundidad en la escena (Bodkin, 2012). El gran reto es poder determinar el mejor método para aproximar las diferencias entre las dos vistas para mapear la correspondencia (el mapa de disparidad) de la escena. Intuitivamente, un mapa de disparidad representa los píxeles correspondientes entre sí que se encuentran desplazados horizontalmente entre las dos imágenes. Cada año aparecen nuevos métodos para resolver este problema, principalmente enfocados en exactitud y tiempo de cómputo (Hamzah & Ibrahim, 2016).

Las cámaras de tiempo de vuelo [ToF] y las cámaras basadas en luz estructurada son alternativas de tecnología activa en la industria, sin embargo, estos dispositivos padecen algunos problemas (Hansard, Lee, Choi, & Horaud, 2013): primero, están sujetos a errores sistemáticos, tales como el ruido y la ambigüedad, los cuales se relacionan con el sensor que en particular se utilice, por ejemplo el Kinect de Microsoft; segundo, están sujetos a errores no sistemáticos, tales como la dispersión y el desenfoque por movimiento *—motion blur—*. De acuerdo con los estudios comparativos de Foix & Aleny (2011), Kim, Ayaz, Park, y Roh (2014), y Zhang, Wang, y Chan (2013), los sensores ToF y los de luz estructurada trabajan muy bien dentro de ciertos rangos (aproximadamente 5 a 7 metros) y son muy sensibles para ser usados en exteriores, especialmente en zonas de mucha iluminación. A causa de estas limitaciones, los sistemas de visión estéreo (pasivos) son más robustos y confiables, son capaces de producir mapas de disparidad de alta resolución y son apropiados para su uso, tanto en exteriores, como en interiores. Para obtener mejores resultados, todavía se requiere de la asistencia de una persona que resuelva errores en la estimación de disparidades.

Por lo anterior, la hipótesis en la que se apoya este trabajo se centra en que a medida que las técnicas de correspondencia de imágenes mejoren, se mejorará la obtención de variables dimensionales, tales como la longitud y el área de los objetos, utilizando técnicas de visión estéreo. A continuación se describe la construcción de un sistema de reconstrucción 3D, mediante visión estéreo, para la determi-



Figure 1. Pair of cameras with FireWire connection for stereo vision / Par de cámaras con conexión firewire para visión estereo

nación semiautomática de la longitud y el área de objetos, que permite la fácil incorporación de nuevos módulos de correspondencia de imágenes para evaluar su efecto sobre la obtención de las ya mencionadas variables dimensionales.

## II. Método

El sistema propuesto consta de una etapa de adquisición consistente en dos cámaras BASLER con lentes de 8 mm, cada una (ver FIGURA 1). La separación (*baseline*) entre los centros ópticos de las cámaras es de aproximadamente 48 mm. Estas cámaras se conectan a través de protocolo IEEE 1394 (*Firewire*) hacia el computador, lo que permite obtener un tamaño de imagen de 658x492 (MxN) a color (24 bit, RGB).

Utilizando el proceso convencional de reconstrucción 3D mediante un sistema de visión estereo (Hamzah & Ibrahim, 2016; Szeliski, 2011; Wöhler, 2013), se propone la meto-

For this purpose, the base hypothesis of this work is correlated in the fact that, whilst the techniques for image correspondence improve with time, the gathering of dimensional variables as length and area of the objects using stereo vision techniques will also be improved. In the next section, we describe the making of a 3D reconstruction system using stereo vision for the semiautomatic determination of the length and area of objects; our proposal allows the easy incorporation of new image corresponding modules to assess their effect in the collecting of the already mentioned dimensional variables.

## II. Method

The proposed system has a stage of constant data gathering in two BALSER cameras with 8 mm lens each one (see FIGURE 1). The baseline between the cameras optical centers is approximately 48 mm. These cameras are connected through the IEEE 1394 protocol (FireWire) to the computer, which allows to obtain images with a size up to 658x492 pixels (MxN) with 24 bit RGB color calibration.

Using the 3D conventional process for reconstruction through a stereo vision system (Hamzah & Ibrahim, 2016; Szeliski, 2011; Wöhler, 2013), we propose the methodology exposed in FIGURE 2 to carry on the calibration, rectification, correspondence, and gathering of the point clouds processes. This, in order to obtain—in a semiautomatic way—the distances between points (lengths) and the areas of the surfaces.

By considering this methodology, we built a graphical user interface (see FIGURE 3) in Matlab—2015a or superior—, in order to mainly perform two processes: the cali-

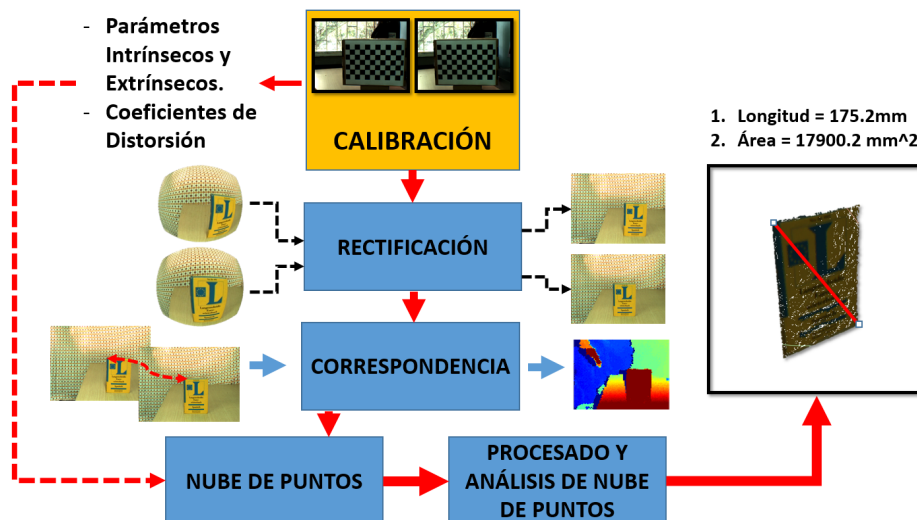


Figure 2. Block diagram of the stages in the stereo vision system for the obtaining of objects length and area / Diagrama de bloques de las etapas del sistema de visión estereo para la obtención de la longitud y el área de objetos

bration of cameras that allow the gathering of the intrinsic and extrinsic parameters plus the radial and tangential distortion coefficients (Szeliski, 2011); besides of the rectification, correspondence, gathering, and analysis of the point clouds from a pair of stereo images of the object (Demant et al., 2013; Krig, 2014). In **FIGURE 3**, section 1 corresponds to the cameras calibration and section 2 is the rectification, correspondence, gathering, and analysis of the point clouds. The latter section corresponds to the set of tools that allow the observation of the applicative credits and the user manual; it also allows the exiting of the main interface.

In **FIGURE 4**, it is possible to observe the interface for the first module (i.e., calibration); the same that uses the Matlab functions for stereo calibration. The interface allows, besides of gathering the intrinsic and extrinsic matrices, the visualization of the average reprojection errors per each pair of images, and the spatial location of the cameras relative to each one of the calibration points in the space.

The second module (**FIGURE 5**) consist of a set of sequential tools that allow the final measurements of length and superficial area of objects. For this, it is necessary to load the pair of images and the calibration matrices (obtained in the first module) of the studied object. Afterwards, the correspondence technique to be used is selected—for now we used the block matching (Hamzah, Hamid, & Salim, 2010) and random walk (Lee, Lee, Lim, & Suh, 2015) ones—. It is important to mention that—in this point—the applica-

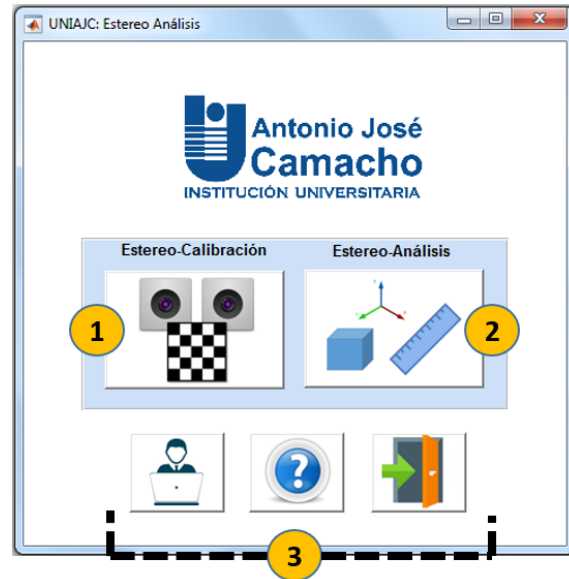


Figure 3. Main graphical user interface / Interfaz gráfica de usuario principal

dología expuesta en la **FIGURA 2**, para realizar el proceso de calibración, rectificación, correspondencia y obtención de la nube de puntos, y el procesado y análisis final de este último, para obtener, de forma semiautomática, distancias entre puntos (longitudes) y área de superficies.

Teniendo presente esta metodología se construye una interfaz gráfica de usuario (ver **FIGURA 3**) en Matlab™–2015a o superior–, para realizar principalmente dos procesos: la calibración de cámaras que permite la obtención de los parámetros intrínsecos y extrínsecos, así como los coefi-

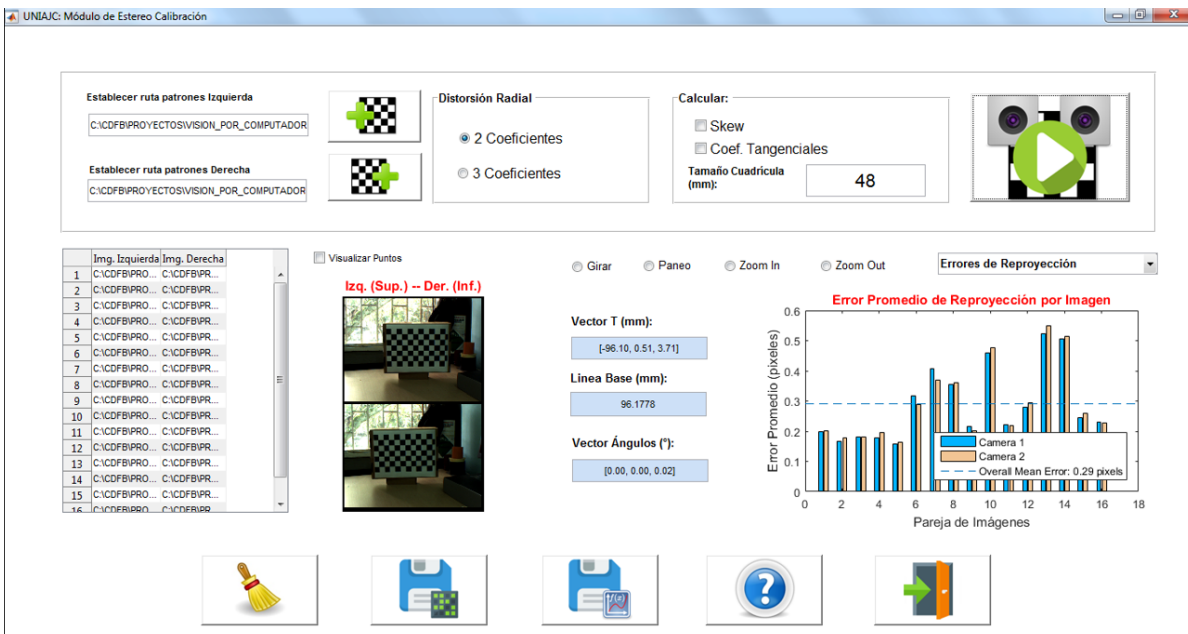


Figure 4. Graphical interface for the stereo calibration process / Interfaz gráfica para proceso de calibración estereo

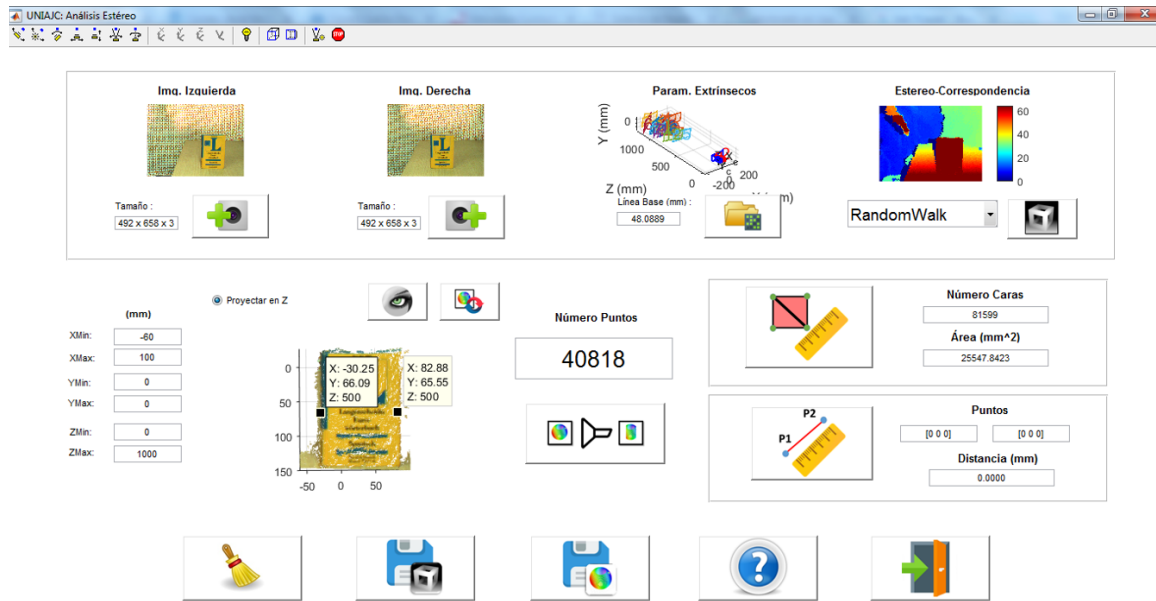


Figure 5. Graphical interface for the gathering of the length and area of the objects measured using stereo vision /  
 Interfaz gráfica para la obtención de la longitud y el área de objetos mediante visión estéreo

cientes de distorsión radial y tangencial (Szeliski, 2011); y la rectificación, correspondencia, obtención y análisis de la nube de puntos a partir de un par estéreo de imágenes de objeto (Demant et al., 2013; Krig, 2014). En la FIGURA 3: la sección 1 corresponde a la calibración de cámaras y la sección 2 a la rectificación, correspondencia y obtención y análisis de nubes de puntos; la sección corresponde al conjunto de herramientas que permite ver los créditos del aplicativo y el manual de usuario, y permite salir de la interfaz principal.

En la FIGURA 4 se observa la interfaz para el primer módulo –calibración–, el mismo que utiliza las funciones de Mat-

tion allows the incorporation of new functions for image corresponding by simply locating the script inside the *AlgoritmosCorrespondencia* folder located next to the *Library* one. The source code of the project can be seen in Ferrin (2017).

The script must have as input parameters two images in gray scale (with  $M \times N$  size and type *uint8*) and as output parameter an image of the same size the input image is, but in the *single* type. Once the image corresponding technique is applied, and using the calibration parameters, the system

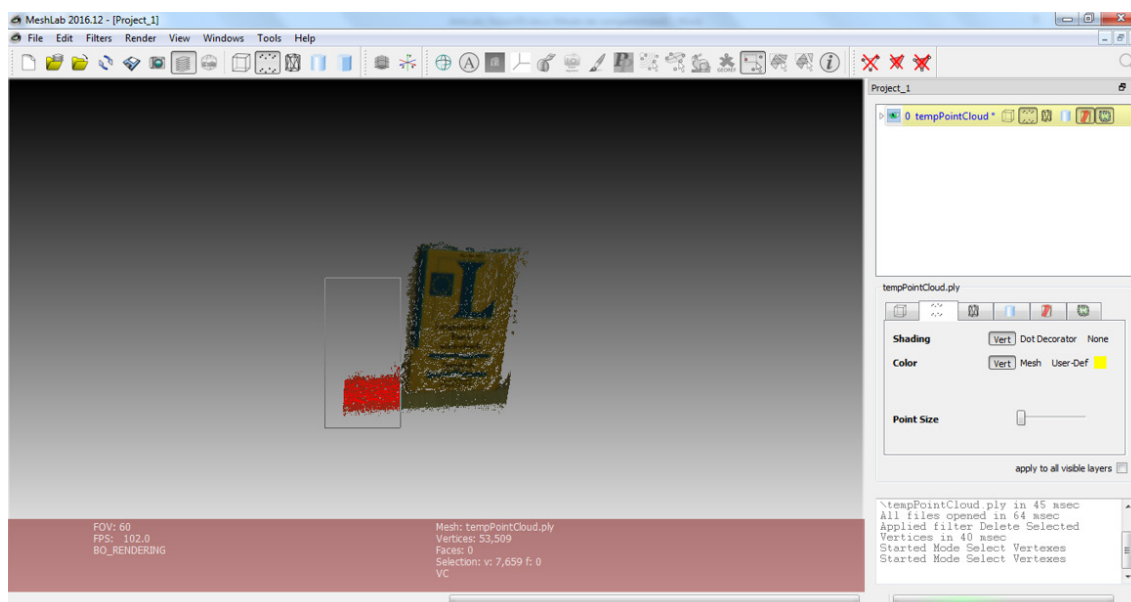


Figure 6. Processing the point cloud using MeshLab / Procesado de nube de puntos utilizando MeshLab

allows the gathering of the point cloud in color of the scene through a calibration process (Hamzah & Ibrahim, 2016; Krig, 2014; Szeliski, 2011; Wöhler, 2013). From this point cloud, the obtaining of the length ( $L$ ) between two points ( $P1, P2$ ) in 3D belonging to the object is done via the Euclidean distance (1). It is important to notice that the selection of these parameters is mandatory in the graphical interface to perform the calculation displayed in (1):

$$L = \sqrt{(P1 - P2) \cdot (P1 - P2)} \quad (1)$$

In order to obtain the area of the surface, we used the Delaunay technique (De Loera, Rambau, & Santos, 2010) to gather the set of triangles defining the surface of the object. By using the algorithm described below, we obtained an estimation of the area of the defined surface by the object's point cloud.

*Entrada:* **CT**, variable de conjunto de triángulos;  
*Salida:* **Area**, variable para Área;

**Area** = 0;

Para cada triángulo **T** dentro de **CT**

**P1, P2, P3** = **Vertices(T)**;

**PC** = **ProductoCruz((P2 - P1), (P3 - P1))**;

**Area** = **Area** + 0.5\*(**Norma(PC)**);

Fin Para

**Imprimir(Area)**;

It is possible to process the point cloud *before* obtaining the length and area through a button that has the interface in the second module. This button opens the MeshLab open source software (Cignoni et al., 2008) to work more easily with the point clouds and to be able to eliminate points that do not belong to the object, sub-map points, and other tasks (see **FIGURE 6**).

Once the description of the design and implementation processes of the hardware-software tool for the stereo vision system is done, we proceed to perform some tests to assess its accuracy with the estimation of lengths and areas for certain objects; testing its benefits and limitations at the same time. This tool, together with the image database for the calibration and reconstruction processes can be obtained from Ferrin (2017).

### III. Results

In order to assess the described methodology and tools, we performed a calibration process by using a lattice of

lab<sup>TM</sup> para estéreo calibración; la interfaz permite, además de obtener las matrices intrínsecas y extrínsecas, visualizar los errores promedios de reproyección por cada par de imágenes, y la ubicación espacial de las cámaras en relación con cada uno de los puntos de calibración en el espacio.

El segundo módulo (**FIGURA 5**) consta de un conjunto de herramientas secuenciales que permiten realizar las medidas finales de longitud y área superficial de objetos. Para esto es necesario cargar el par de imágenes del objeto a estudiar y las matrices de calibración (obtenidas en el primer módulo). Luego se selecciona la técnica de correspondencia que se va a utilizar –por ahora se utiliza *Block Matching* (Hamzah, Hamid, & Salim, 2010) y *Random Walk* (Lee, Lee, Lim, & Suh, 2015)–. Es importante mencionar en este punto que el aplicativo permite incorporar nuevas funciones para correspondencia de imágenes, simplemente ubicando el *script* correspondiente dentro de la carpeta *AlgoritmosCorrespondencia* que se encuentra dentro de la carpeta Biblioteca. El código fuente del proyecto se puede ver en Ferrin (2017).

El *script* debe tener como parámetros de entrada dos imágenes en escala de grises (tamaño  $M \times N$ , tipo *uint8*) y como parámetro de salida una imagen del mismo tamaño de las imágenes de entrada, pero de tipo *single*. Una vez aplicada la técnica de correspondencia de imágenes y valiéndose de los parámetros de calibración, el sistema, a través de un proceso de triangulación (Hamzah & Ibrahim, 2016; Krig, 2014; Szeliski, 2011; Wöhler, 2013), permite obtener la nube de puntos a color de la escena. A partir de esta nube de puntos, la obtención de la longitud  $L$ , entre dos puntos  $P1$  y  $P2$ , en 3D, pertenecientes al objeto se obtiene mediante la distancia euclidiana (Ec. 1), una vez se han seleccionado los mismos con ayuda de la interfaz gráfica:

$$L = \sqrt{(P1 - P2) \cdot (P1 - P2)} \quad (1)$$

Para obtener el área de la superficie se recurre a la técnica de Delaunay (De Loera, Rambau, & Santos, 2010) para obtener el conjunto de triángulos que definen la superficie del objeto. Con ayuda del siguiente algoritmo se obtiene una estimación del área de la superficie definida por la nube de puntos del objeto.

*Entrada:* **CT**, variable de conjunto de triángulos;  
*Salida:* **Area**, variable para Área;

**Area** = 0;

Para cada triángulo **T** dentro de **CT**

**P1, P2, P3** = **Vertices(T)**;

**PC** = **ProductoCruz((P2 - P1), (P3 - P1))**;

**Area** = **Area** + 0.5\*(**Norma(PC)**);

Fin Para

**Imprimir(Area)**;

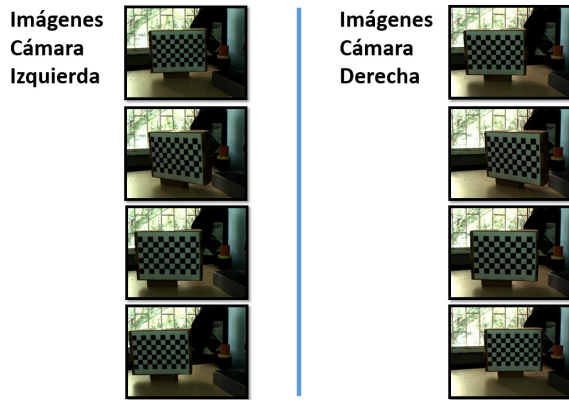


Figure 7. A subset of four pairs of images for calibration / Un subconjunto de cuatro pares de imágenes para calibración

Es posible procesar la nube de puntos antes de obtener la longitud y el área a través de un botón que posee la interfaz del segundo módulo, el cual abre el software de código abierto MeshLab (Cignoni et al., 2008) para trabajar de manera más fácil con la nube de puntos, y poder así eliminar puntos que no pertenecen al objeto, sub-muestrear puntos, etc. (ver FIGURA 6).

Una vez descrito el proceso de diseño e implementación de la herramienta hardware-software del sistema de visión estéreo, se procede, en la siguiente sección, a realizar algunas pruebas para evaluar su exactitud al momento de estimar longitudes y áreas de objetos, y de esta forma resaltar sus bondades y limitaciones. Esta herramienta, al igual que la base de datos de imágenes para calibración y reconstrucción, puede ser obtenida de Ferrin (2017).

### III. Resultados

Para evaluar la metodología y la herramienta descritas se procede a realizar un proceso de calibración, utilizando un damero de 11x6 con tamaño de cuadrícula de 24 mm. Se obtuvieron 16 vistas variando la orientación del damedero.

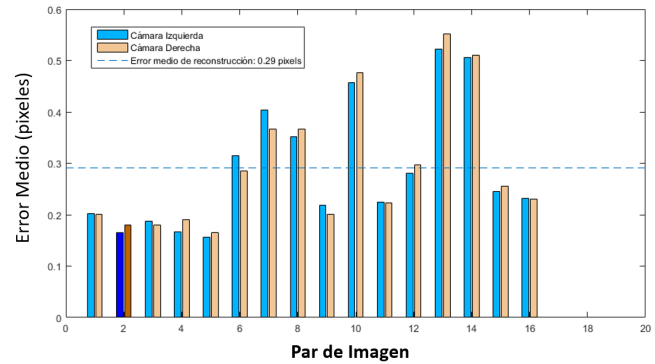


Figure 8. Reprojection average error for each pair of images / Error promedio de reproyección por cada par de imagen

11x6 with grid size of 24 mm. We obtained 16 views by modifying the orientation of the lattice inside the region where the objects to be reconstructed were located. In FIGURE 7, it is possible to observe four pairs of images of the set of 16 available.

In TABLE 1, we summarize the different parameters obtained after the calibration process.

From the previous result, we reprojected the obtained points for the calibration process and a general reprojection average error of 0.29 pixels is obtained, as the reader might see in FIGURE 8.

As per the information in FIGURE 9, it is possible to infer that the lowest reprojection error is obtained in the fifth pair of images (0.17 pixels) and the higher one was achieved in the third pair of images (0.55 pixels).

The extrinsic parameters allow to visualize the location and orientation of the cameras in relation to the location and orientation of the lattice on each of the 16 views (see FIGURE 10). The baseline obtained through the norm of the translation vector between the cameras is 48.2 mm, this

Table 1. Parameters of stereo calibration / Parámetros para la calibración estéreo

Parameter / Parámetro	Left camera / Cámara izquierda	Right camera / Cámara derecha
Intrinsic matrix / Matriz intrínseca	$\begin{bmatrix} 1177.4 & 0 & 374.7 \\ 0 & 1188.4 & 234.3 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1177.1 & 0 & 391.7 \\ 0 & 1185.9 & 230.5 \\ 0 & 0 & 1 \end{bmatrix}$
Radial coefficients / Coeficientes radiales	$\begin{bmatrix} 0.0978 \\ -4.0535 \\ 34.4441 \end{bmatrix}$	$\begin{bmatrix} -0.1313 \\ 2.5777 \\ -18.2521 \end{bmatrix}$
Tangential coefficients / Coeficientes tangenciales	$\begin{bmatrix} -0.0034 \\ 0.0041 \end{bmatrix}$	$\begin{bmatrix} -0.0030 \\ 0.0108 \end{bmatrix}$
Rotation matrix / Matriz de rotación	$\begin{bmatrix} 0.9996 & -0.0176 & 0.0233 \\ 0.0176 & 0.9998 & -0.0020 \\ -0.0233 & 0.0024 & 0.997 \end{bmatrix}$	
Translation vector / Vector de traslación (mm)		$\begin{bmatrix} -48.2361 \\ 0.3119 \\ 0.4282 \end{bmatrix}$

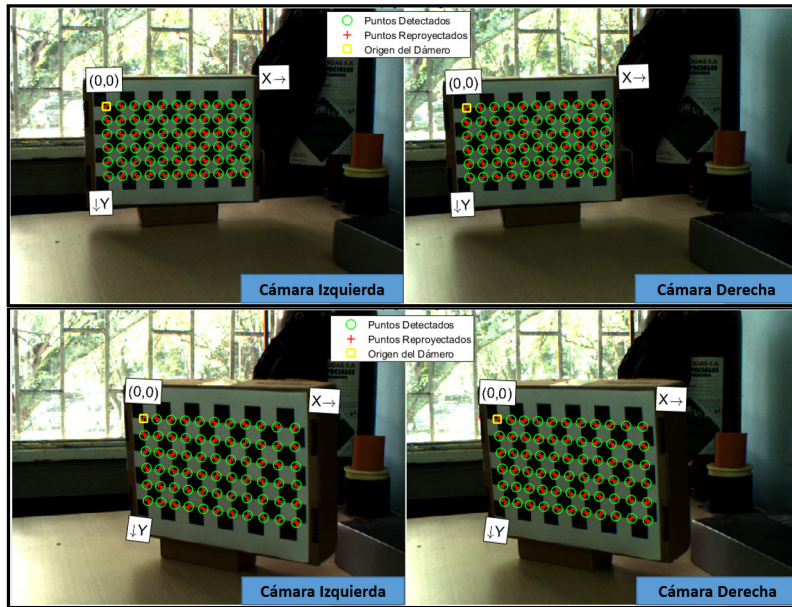


Figure 9. Reprojected points for the fifth pair (up) and thirteenth pair (down) of images / Puntos reprojectados para los pares de imagen quinto (arriba) y décimo tercero (abajo)

means that it has been possible to recover the baseline with a relative error of 0.4%. The reader might remember that the system baseline is 48 mm (see **FIGURE 1**).

Once the calibration of the stereo system is performed, we proceeded to estimate the lengths and areas on three objects with cylindrical form (O1 and O2) and one with squared form (O3) (see **FIGURE 11**).

#### Effect of the image corresponding technique in the estimation of the length and area

Even when the images are rectified, the occlusions and homogeneities do not allow to have a perfect dense correspondence for the images. For this reason, the application allows the easy incorporation of techniques in the state of

ro dentro de la región donde se pondrán los objetos a reconstruir. En la **FIGURA 7** se observan cuatro pares de imágenes del conjunto de 16 disponibles.

En la **TABLA 1** se resumen los diferentes parámetros obtenidos después del proceso de calibración.

A partir del resultado anterior, se reprojectan los puntos obtenidos para el proceso de calibración y se obtiene un error promedio general de reproyección de 0.29 píxeles, tal como se observa en la **FIGURA 8**.

Como se puede observar en la **Figura 9**, el error de reproyección más bajo se obtiene en el quinto par de imágenes (0.17 píxeles) y el más alto en el décimo tercer par de imagen (0.55 píxeles).

Los parámetros extrínsecos permiten visualizar la ubicación y orientación de las cámaras con relación a la ubicación y orientación del damero en cada una de las 16 vistas (ver **FIGURA 10**). La línea base obtenida a partir de la norma del vector de traslación entre las cámaras es 48.2 mm, ello significa que se ha podido recuperar la línea base con error relativo del 0.4%. Recuerdese que la línea base del sistema es de 48 mm (ver **FIGURA 1**).

Una vez calibrado el sistema de visión estéreo, se procede a estimar longitudes y áreas sobre tres objetos con forma cilíndrica, O1 y O2, y cuadrilátera, O3 (ver **FIGURA. 11**).

#### Efecto de la técnica de correspondencia de imágenes sobre la estimación de longitud y área

Aún cuando las imágenes son rectificadas, las occlusiones y homogeneidades no permiten tener una correspondencia densa de imágenes perfecta. Por tal motivo, el aplicativo permite la fácil incorporación de técnicas del estado del arte en correspondencia densa de imágenes en el pipeline

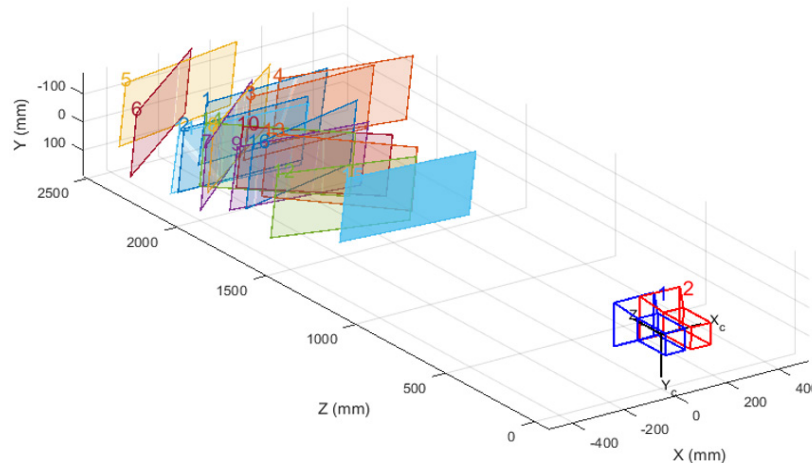


Figure 10. Location and orientation of the cameras in the space / Ubicación y orientación de las cámaras en el espacio



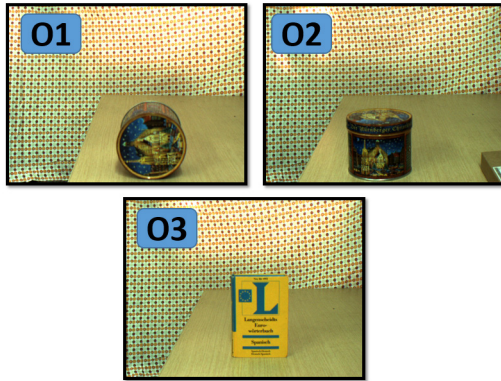


Figure 11. Three objects for the tests of the estimation of length and area / Tres objetos para pruebas de estimación de longitud y área

de ejecución. Para ello basta con crear la función, tal como se explicó en la metodología, para poder ver el efecto sobre la reconstrucción 3D del objeto.

En la FIGURA 12 se observa la mejora en la reconstrucción 3D de los objetos cuando se aplica la técnica de *Random Walk*, en lugar de la de *Block Matching*. Las Tablas 2, 3 y 4 muestran las estimaciones de medidas de los objetos O1, O2 y O3 para ambas técnicas, a una distancia de 35 cm del sistema de visión estéreo.

De los resultados anteriores se deduce que la técnica *Random Walk* presenta, en general, menor nivel de error absoluto que la de *Block Matching*, al momento de obtener medidas de longitud y área. Por los resultados expuestos por Lee et al., (2015) se sabe que la técnica *Random Walk* es más exacta que *Block Matching*, aun cuando esta última es mucho más rápida.

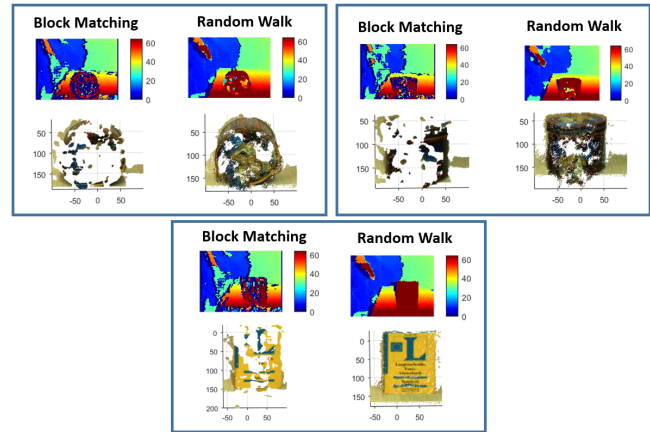


Figure 12. Disparity map and 3D reconstruction of the O1, O2, and O3 objects using block matching and random walk / Mapa de disparidad y reconstrucción 3D de los objetos O1, O2 y O3 utilizando Block Matching y Random Walk

the art related to dense correspondence of images in the execution pipeline. To do so, the creation of the function (explained in the methodology) is required in order to see the effect in the 3D reconstruction of the object.

In FIGURE 12, we present the improvements in the 3D reconstruction of the objects when the random walk technique is applied instead of the block matching one. Tables 2, 3, and 4 show the estimations of measurements for the O1, O2, and O3 objects for both techniques at a distance of 35 cm of the stereo vision system.

Table 2. Estimation of length and area measurements for O1 / Estimación de medidas de longitud y área para O1

Technique / Técnica	Height / Altura (mm)	Area / Área (mm <sup>2</sup> )
Reference / Referencia	115.0	10.387.4
Block Matching	123.7	26.847.1
Random Walk	110.5	16.416.7

Table 3. Estimation of length and area measurements for O2 / Estimación de medidas de longitud y área para O2

Technique / Técnica	Height / Altura (mm)	Area / Área (mm <sup>2</sup> )
Reference / Referencia	95.0	10.925.0
Block Matching	105.2	18.494.3
Random Walk	92.1	15.562.1

Table 4. Estimation of length and area measurements for O3 / Estimación de medidas de longitud y área para O3

Technique / Técnica	Height / Altura (mm)	Base / Base (mm)	Area / Área (mm <sup>2</sup> )
Reference / Referencia	150.0	100.0	15.000.0
Block Matching	163.1	119.7	21.154.2
Random Walk	157.3	109.3	19.938.5

From the previous results, the reader might infer that the random walk technique presents—in general—a lower level of the absolute error compared to the block matching one when measurements of length and area are involved. By considering the exposed results by Lee et al., (2015), it is known that the random walk technique is more accurate than the block matching one, even though the latter one is faster.

#### *Variation of the error in the estimation of length and area in function of the distance to the stereo vision sound*

We continued with the evaluation of the variations in the absolute error for the estimation of the height, base, and area of the O3 object in function of the distances to the stereo vision system. We worked with the random walk technique and with working distances varying from 35 cm to 80 cm, with 5 cm steps. **TABLE 5** summarizes the obtained results.

The absolute mean error in the height estimation is 6.3 mm, with a standard deviation of 1.2 mm. For the base, we obtained an average value of 8.9 mm and a standard deviation of 3.1 mm. This http://mac-torrent-download.net/none of the cases, the obtained values were greater than 1 cm of absolute value (see **FIGURE 13**). On the other hand, we obtained an average absolute value of 5.042 mm<sup>2</sup> for the area; that means, more than 50 times 1 squared centimeter (see **FIGURE 14**).

In conclusion, for the dimensions of the studied object (O3), it is possible to see that the absolute error of the estimated lengths (height and base) is not greater than 1 cm

#### *Variación del error en la estimación de longitud y área en función de la distancia al sistema de visión estéreo*

Se procede ahora a evaluar las variaciones del error absoluto en la estimación de la altura, la base y el área del objeto O3 en función de la distancias al sistemas de visión estéreo. Se fija la técnica *Random Walk* y se trabajan distancias desde 35 cm hasta 80 cm, con incrementos de 5 cm. La **TABLA 5** resume los resultados obtenidos.

El error promedio absoluto en la estimación de la altura es de 6.3 mm, con una desviación estándar de 1.2 mm. Para la base se obtiene un promedio 8.9 mm y una desviación estándar de 3.1 mm. Es decir que en ningún caso se supera 1 cm de error absoluto (ver **FIGURA 13**). Por otra parte, para el área se obtiene un error absoluto promedio de 5.042 mm<sup>2</sup>, es decir más de 50 veces 1 cm<sup>2</sup> (ver **FIGURA. 14**).

En conclusión, para las dimensiones del objeto estudiado (O3), se puede observar que el error absoluto de las longitudes estimadas (altura y base) no supera 1 cm (en el caso pesimista se obtendría un error relativo de 10% para la longitud más pequeña, esto es la de la base), mientras que el área presenta un error absoluto comparable al área de O3 (error relativo 34 %).

#### IV. Conclusiones

Se desarrolló un sistema hardware-software de reconstrucción 3D mediante visión estéreo para la estimación semi-automática de la longitud y el área de objetos. La interfaz gráfica permitió obtener muy fácilmente las medidas de longitud y área de tres objetos.

Se pudo comprobar que una mejor técnica de correspondencia de imágenes permite obtener una mejor nube de puntos de los objetos y, en consecuencia, obtener mejores estimaciones de longitud y área. Por otro lado, las estimaciones de longitud se obtuvieron con menores errores relativos

Table 5. Results of area, base, and height for different values of distance to the stereo vision system / Resultados de altura, base y área para diferentes valores de distancia al sistema de visión estéreo

Technique / Técnica	Height / Altura (mm)	Base / Base (mm)	Area / Área (mm <sup>2</sup> )
Reference / Referencia	150.0	100.0	15000.0
80	157.0	115.1	21171.1
75	157.7	111.3	20591.1
70	160.9	104.8	20315.4
65	158.5	113.4	20591.5
60	156.4	110.4	20075.3
55	154.2	107.1	19653.2
50	155.3	107.7	19502.3
45	156.7	106.0	18742.2
40	155.2	106.6	19847.2
35	157.3	109.3	19938.5

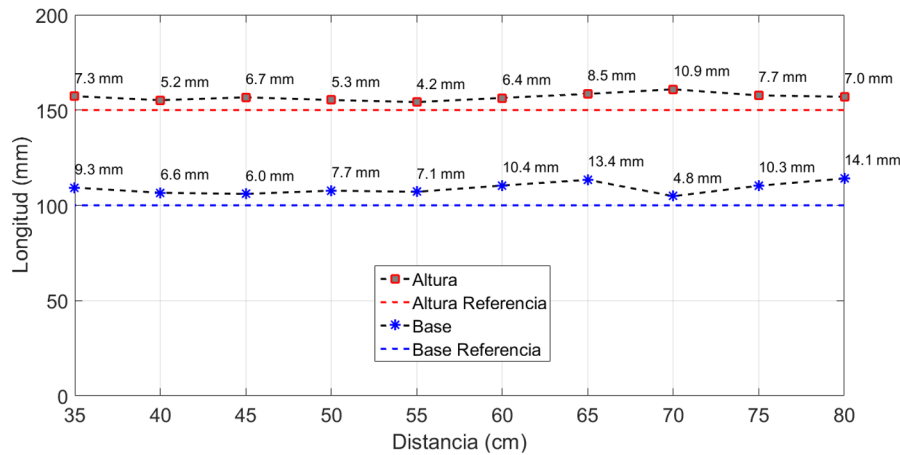


Figure 13. Absolute errors for each estimation of height and base in function of the distance to the stereo vision system / Errores absolutos para cada estimación de altura y base en función de la distancia al sistema de visión estereó

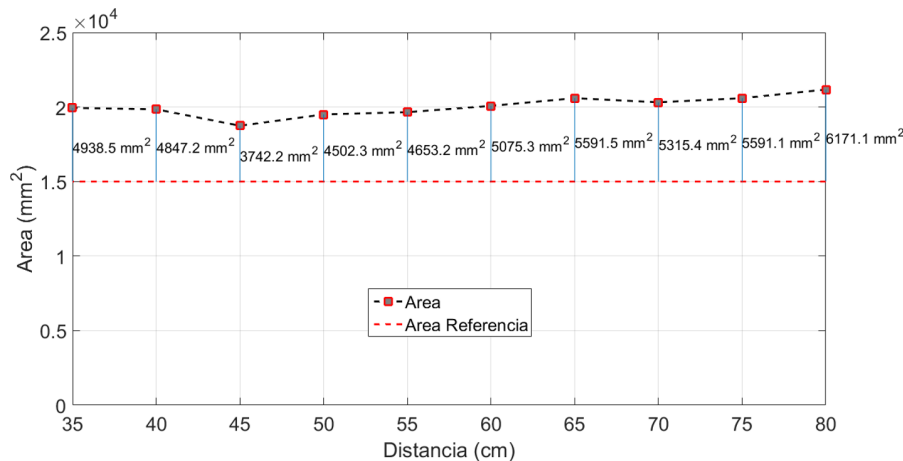


Figure 14. Absolute errors for each estimation of the area in function of the distance to the stereo vision system / Errores absolutos para cada estimación de área en función de la distancia al sistema de visión estereó

que las de área. Lo anterior sugiere que es necesario buscar mejores técnicas de correspondencia de imágenes para obtener nubes de puntos de mejor calidad y así poder reducir los errores relativos en la estimación de estas importantes variables para el control de calidad de piezas en la industria de manufactura.

La herramienta implementada les permitirá a los investigadores y estudiantes investigar e implementar sistemas de reconstrucción 3D mediante visión estereó para estimación de longitud y área de objetos, con sus propias cámaras y técnicas de correspondencia de imágenes.

## V. Agradecimientos

A la Decanatura Asociada de Investigaciones de la Institución Universitaria Antonio José Camacho por el apoyo administrativo y económico al proyecto PI00110 “Generación y análisis de imágenes estereoscópicas a partir de imágenes bidimensionales” y al grupo de investigación Percepción y Sistemas Inteligentes [PSI] de la Universidad del Valle por su asesoría Científica. *ST*

(in the worst case, we would obtain a relative error of 10% for the smallest length, i.e., the base), whilst the area presented an absolute error comparable to the area of O3 (relative error of 34%).

## IV. Conclusions

We developed a hardware-software system for the 3D reconstruction through stereo vision used to perform the semiautomatic estimation of the length and area of various object. The graphical interface allowed the gathering in a simple way of these measurements for three objects.

We were able to compare that a better technique for the image corresponding process allows the gathering of a better point cloud and, consequently, an achievement of better estimations of length and area. Alternatively, the length estimations were obtained with lower relative error than the area ones. This suggests that it is necessary to seek new techniques of image corresponding to obtain point clouds with better quality,

trying to reduce the relative errors in the estimation of these important variables for the quality control of pieces in the manufacturing industry.

The implemented tool will allow researchers and students to research and implement 3D reconstruction systems through stereo vision to estimate the length and area of objects with their own cameras and image corresponding techniques.

## V. Acknowledgement

The authors of this paper would like to express their gratitude to the associated deanery of research in the *Institución Universitaria Antonio José Camacho* for its administrative support and funding to the PI00110 project named “generation and analysis of stereoscopic images through bi-dimensional images” and to the Perception and Intelligent Systems [PSI] research group of the *Universidad del Valle* for its scientific advice. *ST*

## References / Referencias

- Bodkin, B. H. (2012). *Real-Time Mobile Stereo Vision*. University of Tennessee. Retrieved from [http://trace.tennessee.edu/utk\\_gradthes/1313](http://trace.tennessee.edu/utk_gradthes/1313)
- Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., & Ranzuglia, G. (2008). MeshLab: an Open-Source Mesh Processing Tool. *Sixth Eurographics Italian Chapter Conference*, (pp. 129-136). doi:10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136
- De Loera, J. A., Rambau, J., & Santos, F. (2010). *Triangulations* (Vol. 25). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-12971-1
- Demant, C., Streicher-Abel, B., & Garnica, C. (2013). *Industrial image processing*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-33905-9
- Ferrin, C. (2017). *UNIAJC: Proyecto visión estéreo*. Retrieved from [https://github.com/cdfbdex/UNIAJC\\_VISION-ESTEREO](https://github.com/cdfbdex/UNIAJC_VISION-ESTEREO)
- Foix, S. & Aleny, G. (2011). Lock-in time-of-flight (ToF) cameras: A survey. *IEEE Sensors Journal*, 11(3). doi: 10.1109/JSEN.2010.2101060
- Hagen, H., Disch, A., Ehret, J., Klein, R., Sascha, K., Zeckzer, D., & Michael, M. (n.d.). *Visual inspection methods for quality control in automotive engineering*. Visualization, 2004 IEEE. doi:10.1109/VISUAL.2004.111
- Hamzah, R. A., Hamid, A. M. A., & Salim, S. I. M. (2010). The solution of stereo correspondence problem using block matching algorithm in stereo vision mobile robot. *2010 Second International Conference on Computer Research and Development*, (pp. 733-737). doi:10.1109/ICCRD.2010.167
- Hamzah, R. A. & Ibrahim, H. (2016). Literature survey on stereo vision disparity map algorithms. *Journal of Sensors*, 2016. Art. 8742920. doi:10.1155/2016/8742920
- Hansard, M., Lee, S., Choi, O., & Horaud, R. (2013). *Time of flight cameras: Principles, methods, and applications*. London, UK: Springer. <https://doi.org/10.1007/978-1-4471-4658-2>
- Kim, M. Y., Ayaz, S. M., Park, J., & Roh, Y. (2014). Adaptive 3D sensing system based on variable magnification using stereo vision and structured light. *Optics and Lasers in Engineering*, 55, 113-127. doi:10.1016/j.optlaseng.2013.10.021
- Krig, S. (2014). *Computer vision metrics: Survey, taxonomy, and analysis*. Berkeley, CA: Apress. <https://doi.org/10.1007/978-1-4302-5930-5>
- Lee, S., Lee, J. H., Lim, J., & Suh, I. H. (2015). Robust stereo matching using adaptive random walk with restart algorithm. *Image and Vision Computing*, 37, 1-11. doi:10.1016/j.imavis.2015.01.003
- Szeliski, R. (2011). *Computer vision: Algorithms and applications*. London, UK: Springer. doi:10.1007/978-1-84882-935-0
- Wöhler, C. (2013). *3D Computer vision*. London, UK: Springer. doi:10.1007/978-1-4471-4150-1
- Zhang, S., Wang, C., & Chan, S. C. (2013). A new high resolution depth map estimation system using stereo vision and kinect depth sensing. *Journal of Signal Processing Systems*, 79(1), 19-31. doi:10.1007/s11265-013-0821-8

## CURRICULUM VITAE

**Bryan García** Electronic Engineer (2016) and student of the Master in Engineering with emphasis in Automation of the *Universidad del Valle* (Cali, Colombia). His interest areas are: artificial vision, thermography, machine learning, and signals/images processing / Ingeniero Electrónico (2016) y estudiante de la Maestría en Ingeniería con énfasis en Automática de la de la Universidad del Valle (Cali, Colombia). Sus áreas de interés son: visión artificial, termografía, aprendizaje de máquina y procesamiento de señales y de imágenes.

**Carlos Diego Ferrín Bolaños** Physics Engineer (*Universidad del Cauca*, Popayan – Colombia); Master in Electronics Engineering (*Universidad del Valle*, 2015) and student of the Doctorate in Engineering with emphasis in Electrics and Electronics of the *Universidad del Valle*. He was beneficiary of the “young researchers” program of Colciencias (2011) and his interest areas are: artificial vision, machine learning, and processing of signals, images, and point clouds / Ingeniero Físico (Universidad del Cauca, Popayán-Colombia, 2010); Magister en Ingeniería Electrónica (Universidad del Valle, 2015) y estudiante del Doctorado en Ingeniería, con énfasis en Eléctrica y Electrónica, de la Universidad del Valle. Fue Joven Investigador (Colciencias, 2011). Sus áreas de interés son la visión artificial, aprendizaje de máquina y el procesamiento de señales, imágenes y nubes de puntos.

**Jorge Humberto Erazo Aux** Electronic Engineer (2014), Master in Engineering with emphasis in Electronics (2010) and student of the Doctorate in Engineering with emphasis in Electrics and Electronics of the *Universidad del Valle*. Full time and auxiliary professor affiliated to the engineering faculty of the *Institución Universitaria Antonio José Camacho* (Cali-Colombia). Thermography professional level I and II of the Infrared Training Center – ITC (2007 and 2011). His areas of interest are: thermography, artificial vision, digital signal processing, and pattern recognition / Ingeniero Electrónico (2004), Magister en Ingeniería con énfasis en Electrónica (2010) y estudiante de Doctorado en Ingeniería con énfasis en Eléctrica y Electrónica de la Universidad del Valle (Cali - Colombia). Profesor auxiliar de tiempo completo adscrito a la Facultad de Ingeniería de la Institución Universitaria Antonio José Camacho (Cali-Colombia). Termógrafo nivel I y II del *Infrared Training Center* - ITC (2007 y 2011). Sus áreas de interés son: termografía, visión artificial, procesamiento digital de señales y reconocimiento de patrones.