



Polibits

ISSN: 1870-9044

polibits@nlp.cic.ipn.mx

Instituto Politécnico Nacional

México

Rodríguez Zalapa, Omar; Hernández Zavala, Antonio; Huerta Ruelas, Jorge Adalberto  
Sistema de medición de distancia mediante imágenes para determinar la posición de una  
esfera utilizando el sensor Kinect XBOX

Polibits, vol. 49, 2014, pp. 59-67

Instituto Politécnico Nacional

Distrito Federal, México

Disponible en: <http://www.redalyc.org/articulo.oa?id=402640463008>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

# Sistema de medición de distancia mediante imágenes para determinar la posición de una esfera utilizando el sensor Kinect XBOX

Omar Rodríguez Zalapa, Antonio Hernández Zavala,  
Jorge Adalberto Huerta Ruelas

**Resumen**—En este documento se presenta un método para medir la distancia del centroide de un objeto segmentado en una imagen de color con respecto a un punto de referencia fijo. El algoritmo se probó mediante una secuencia de imágenes de color, analizando más de 100 posiciones verticales diferentes de una esfera alojada en el interior de una columna cilíndrica transparente de acrílico con diámetro y longitud constante. El algoritmo propuesto integra técnicas de corrección por balance de blancos y de calibración de la cámara con sus parámetros intrínsecos, además, se prueba un nuevo método de segmentación en color utilizado para calcular distancias del mundo real a partir de imágenes en color RGB. Los resultados obtenidos reflejan una alta confiabilidad ya que el 100% de las mediciones realizadas tuvo un error menor a 1.64% con un nivel de precisión más alto que el instrumento utilizado de referencia, en un rango de distancia de 0 a 1340 mm.

**Palabras clave**—Fotogrametría, medición de distancias en imágenes, sensor visual de distancias, metrología visual de simple vista, segmentación, seguimiento de objetos.

## Distance Measurement System using Images to Determine the Position of a Sphere using the XBOX Kinect Sensor

**Abstract**—This paper presents a method to measure the distance from the centroid of a segmented object in a color image with respect to a fixed reference point into the image. The algorithm was tested using a color image sequence by analyzing over 100 different vertical positions of a ball housed inside a transparent acrylic cylindrical column of constant diameter and length. The proposed algorithm integrates techniques of correction by white balance and calibration of the camera with its intrinsic parameters; in addition, a new color segmentation method is tested to calculate real-world distances into color images RGB. The results show high reliability because 100% of measurements

had a relative error in percentage less than 1.64%, with a higher level of precision than the reference instrument used in a distance range from 0 to 1340 mm.

**Index Terms**—Photogrammetry, measuring distances into images, visual sensor of distances, simple view metrology, segmentation, tracking objects.

### I. INTRODUCCIÓN

El desarrollo de cámaras digitales con alta tecnología y día a día con mejores capacidades, ha permitido que en diferentes aplicaciones se utilicen como una alternativa a la visión humana, en diferentes tareas como, en la supervisión de personas mediante sistemas “inteligentes” que pueden detectar, contar, identificar y seguir la trayectoria de las personas [1], [2]; en sistemas de inspección visual automatizados de productos para el control de su calidad en las empresas de manufactura [3], entre muchas otras aplicaciones que crecen día a día.

En particular, el empleo de cámaras como sensores de medición de distancia ha tenido diferentes aplicaciones. Por ejemplo, su utilización para permitir la navegación autónoma de robots terrestres [4]; en arquitectura como un instrumento de medición en interiores para obtener las dimensiones de paredes y pisos, así como para la ubicación correcta de muebles con el propósito de diseño de interiores, en exteriores para medir el tamaño y la posición de ventanas y puertas [5], y muchas más. Una imagen o secuencia de imágenes trae consigo una cantidad muy grande de información geométrica acerca de la escena representada, se han desarrollado diferentes técnicas para la construcción de escenarios 3D a partir de imágenes en 2D [6].

Se han desarrollado técnicas y métodos de descomposición de imágenes para su representación en el espacio del mundo real. Thomas Bucher [7], describe un método para mapear una imagen a coordenadas del mundo real y obtener así, una aproximación de la altura de objetos, longitudes y cambios de posición; basándose en un pequeño grupo de parámetros de fácil estimación a partir de características de los objetos o marcas en la escena, esto sin la necesidad de requerir alguno de los parámetros intrínsecos de la cámara.

Manuscrito recibido el 11 de marzo de 2014; aceptado para la publicación el 30 de mayo del 2014.

Los autores están en el Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada C.I.C.A.T.A., Unidad Querétaro, México (correos: omar.rodriguez.2013@ieee.org, anhernandez@ipn.mx, jhuertar@ipn.mx)

Lázaro et al. [8], presentan la caracterización de la variación de intensidad de niveles de grises y su análisis mediante FFT (*Fast Fourier Transform*: transformada rápida de Fourier), en imágenes tomadas para medir la distancia entre un diodo emisor de infrarrojo y el centro de una cámara. El método propuesto se aplicó para hacer una estimación de distancias en el rango de 420 a 800 cm, logrando una exactitud sobre el 3%.

Al área de investigación que corresponde el obtener mediciones de distancias a partir de imágenes se le llama *Metrología Visual* y puede ser clasificada en dos tipos: *Metrología de simple vista* y *metrología de múltiples vistas* [9].

Entre las ventajas de la metrología visual sobre otras técnicas es que sólo se requiere una vista del objeto (capturada en una imagen) para hacer una medición, por lo cual se considera un método no invasivo, fácil de utilizar, con mayor cantidad de información, posibilidad de determinar muchas distancias en el sistema en base a una secuencia de imágenes y registro histórico para análisis posterior. Aunque, en ciertas aplicaciones, el resultado de la medición se requiere lo más pronto posible con respecto al momento en el que ésta se realizó, y el procesamiento digital de la imagen siempre consumirá un tiempo que se debe tomar en cuenta en el sistema de medición.

El problema de medir las dimensiones de objetos de manera directa con instrumentos como por ejemplo, cinta métrica, flexometro, regla, calibrador (Vernier), micrómetro, etc., es que el objeto tiene que estar disponible físicamente para colocar el instrumento de medición sobre él.

Existen otros métodos para medir distancias sin contacto llamados activos, que requieren la activación de un emisor para generar ya sea un ultrasonido, un rayo de luz infrarroja o un láser; éste emisor, se debe direccionar hacia un punto específico para medir mediante un receptor, las características de retorno de la señal emitida y así, poder determinar la distancia existente entre el instrumento y un punto sobre un objeto remoto. En la actualidad, éste tipo de instrumentos de medición de distancias digitales pueden traer incorporadas las funciones de registro histórico de las mediciones en chips de memoria interna o externa al instrumento, así como algunos tienen la posibilidad de comunicación con una computadora mediante puertos de comunicación como RS232, USB o Ethernet.

En metrología visual de múltiples vistas, se requieren dos o más cámaras dispuestas espacialmente en una forma particular, y se requiere conocer las propiedades específicas de cada cámara individual para obtener mediciones exactas. El hecho de que se necesite tomar y analizar más de dos imágenes de la misma escena también hace que el sistema sea más complejo de programar y de usar. En este trabajo se determinó utilizar *metrología visual de simple vista* ya que implica mayor facilidad y menor costo económico y computacional.

En la sección II, se presenta la descripción del problema y la arquitectura de los elementos que intervienen para medir la distancia que hay del centroide de una esfera dentro de una columna cilíndrica de acrílico a la base de la misma, mediante la utilización de imágenes de color. En la sección III, se explica la propuesta de solución. En la sección IV, se muestran los resultados obtenidos y finalmente en la sección V, se presenta algunas conclusiones.

## II. DESCRIPCIÓN DEL PROBLEMA

Se requiere diseñar e implementar un sistema de medición para determinar la distancia que hay entre una esfera, que se ubicará en posiciones fijas y estables, a lo largo de una columna de 1395 mm de longitud, y un punto de referencia fijo, llamado base de la columna. El desplazamiento incremental vertical controlado de la esfera se llevó a cabo mediante un hilo amarrado a ella que salía por la parte superior de la columna y la mantenía suspendida en la posición deseada. Una vez validada la precisión y exactitud del sistema de medición, se integrará a un sistema automático de control de posición de una esfera levitada neumáticamente.

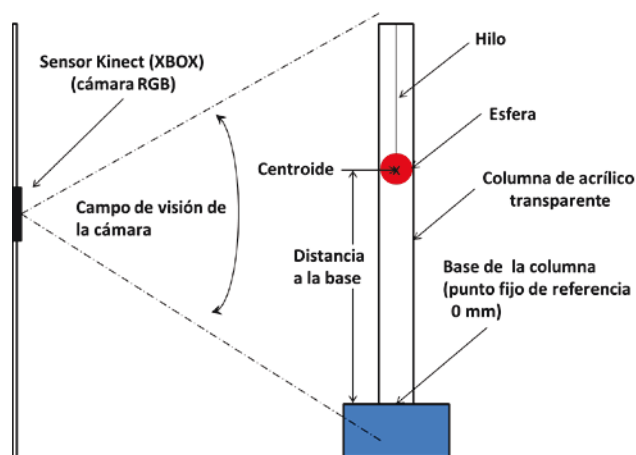


Fig. 1. Arquitectura del sistema de medición.

En la fig. 1, se muestra la localización de cada uno de los componentes del sistema visual de medición de distancias. La captura de imágenes de color se llevó a cabo mediante el sensor Kinect XBOX 360 por dos razones principales: a) Bajo costo (en comparación con una cámara digital industrial con similares características); b) Configuración ágil y simple para que pueda ser utilizado con una computadora mediante un cable USB (se requiere una fuente externa de alimentación para la cámara). Además, se cuenta con un soporte robusto mediante el acceso abierto a diferentes tipos de drivers, librerías y desarrollo de aplicaciones listas para ser utilizadas y con acceso libre al código fuente, todo esto para diferentes plataformas como Mac, UNIX, y Windows.

Las características destacadas del sensor XBOX 360 de Microsoft para éste proyecto se muestran en la tabla I. Las

características completas de éste dispositivo se han publicado en diferentes documentos [12, 13].

En la tabla I, se indica que el campo de visión del sensor es de  $57^\circ$  en horizontal y  $43^\circ$  en vertical, considerando que el sensor se dispone en una posición horizontal como la que se muestra en la fig. 2. Para el sistema de medición se determinó colocar el sensor en dirección vertical como se muestra en la fig. 1, obteniendo de esta manera un campo de visión sobre el eje vertical de  $57^\circ$ . Se ajustó la posición definitiva del sensor sobre un tubo vertical de metal de tal manera, que su campo de visión cubriera la totalidad de la columna de acrílico por la que se desplaza la esfera.

TABLA I.  
CARACTERÍSTICAS PRINCIPALES DEL SENSOR KINECT XBOX 360

Característica	Valor
Campo de visión angular	$57^\circ$ horizontal, $43^\circ$ vertical
Rango de inclinación física	$\pm 27^\circ$
Máximo valor de flujo de datos	Aprox. 30 cuadros por segundo
Resolución de imágenes de color	$640 \times 480$ pixeles (VGA)
Tipo de conexión del dispositivo	USB (+ fuente externa de energía)

La columna cilíndrica de acrílico transparente se encuentra separada a una distancia de 1425 mm del centro de la cámara, y las dimensiones de la columna cilíndrica son 1395 mm de altura con un diámetro externo de 76.2 mm y un grosor de paredes de 3 mm; de manera externa la mitad de la columna cilíndrica se cubrió con una película de vinil negro, dejando el lado descubierto en dirección del eje focal de la cámara.



Fig. 2. Posición horizontal del sensor Kinect XBOX 360® [14].

La base tiene dimensiones de 299x248x295 mm y es del mismo material que la columna. La esfera tiene un diámetro de 60 mm y es de poliestireno expandido (unicel) pintada de color rojo, está alojada dentro de la columna de acrílico y puede deslizarse libremente en dirección vertical. En la parte inferior interna de la base de la columna se alojara como trabajo posterior una tarjeta de control de velocidad para un motor de c.d. (corriente directa) con aspas, que inyectara aire para elevar la esfera dentro de la columna. El aire saldrá por la parte superior, y a su vez se colocara una trampa para que no se salga la esfera.

Para probar y validar la exactitud y precisión del sistema de

medición de distancias mediante imágenes de color, se consideraron condiciones estables y controladas para la ubicación de la esfera en posiciones fijas específicas a lo largo de la longitud de la columna, esto se logró mediante la suspensión de la esfera con un hilo amarrado a ella, cuya longitud se ajustaba y éste salía por la parte superior de la misma; como se puede observar en la fig. 1.

### III. PROPUESTA DE SOLUCIÓN

De manera previa a la captura de imágenes y su procesamiento, se ajustó la orientación de la columna y del sensor a un eje vertical de  $90^\circ$  y la base se niveló con respecto a un eje horizontal de  $0^\circ$ .

Para poder realizar una medición de distancia en una secuencia de imágenes capturadas por el sensor, se debe considerar el modelo de cámara oscura (*pinhole*) [15] del plano de la cámara que se encuentra representado en la fig. 3. Aquí se muestra que un punto  $X$  en el espacio real 3D, es representado en el plano de la imagen como  $x$ . Las coordenadas Euclidianas ( $X, Y, Z$ ) definen la ubicación del punto  $X$  en el espacio real, así como  $(x, y)$  definen la posición del mismo punto en el plano de la imagen en la cámara denotado por  $x$ .

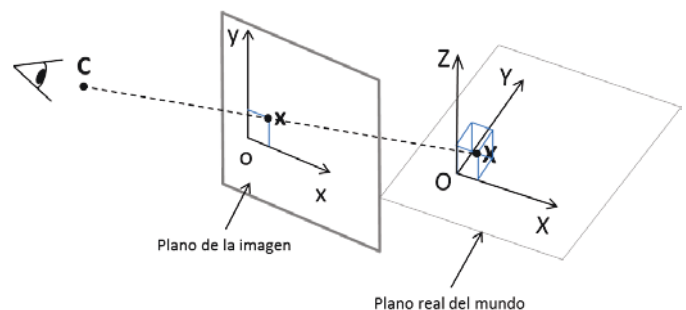


Fig. 3. Modelo "pinhole" (o cámara oscura) que representa el plano 2D de la imagen en la cámara con respecto al espacio real 3D del mundo. Un punto  $X$  en el espacio real es generado como  $x$  en el plano de la cámara. El punto  $C$  representa el centro de la cámara [15].

El proceso de calibración para obtener los parámetros intrínsecos y extrínsecos de la cámara de color incorporada en el sensor, se llevó a cabo siguiendo la metodología del ToolBox para MatLab desarrollado por Jean-Yves Bouguet [16]; obteniendo los siguientes resultados con incertidumbres de los parámetros intrínsecos (modelo de la cámara).

Longitud Focal:

$$fc = [523.95439 \ 520.91487] \pm [2.45522 \ 2.02515]$$

Punto principal:

$$cc = [319.50000 \ 239.50000] \pm [0.00000 \ 0.00000]$$

Sesgo:

$$\alpha_{p\_c} = [0.00000] \pm [0.00000] \Rightarrow$$

$$\text{ángulo de pixel ejes} = 90.00000 \pm 0.00000 \text{ grados}$$

*Distorsión:*

$$k_c = [0.13660 \quad -0.29048 \quad 0.00217 \quad -0.00161 \quad 0.00000] \pm [0.02046 \quad 0.06266 \quad 0.00137 \quad 0.00108 \quad 0.00000]$$

$$\text{Error de pixel: } err = [0.83738 \quad 0.83172]$$

Los valores mostrados de los parámetros intrínsecos de la cámara son los que directamente arroja el software de calibración con formato propio del autor. Y estos valores son los que se tomaron en cuenta para llevar a cabo las correcciones necesarias en las imágenes capturadas. Además, debemos considerar otras operaciones de ajuste y calibración inicial tomando en cuenta alguna imagen *muestra* del escenario real de medición, como la que se indica en la fig. 5 (a), que muestra el objeto de estudio inmerso en el escenario real con cierta inclinación inducida que el programa de computadora desarrollado corregirá. En la fig. 4. se muestra el diagrama de flujo del algoritmo completo del proceso de ajuste inicial para realizar la medición de distancias en imágenes de color.

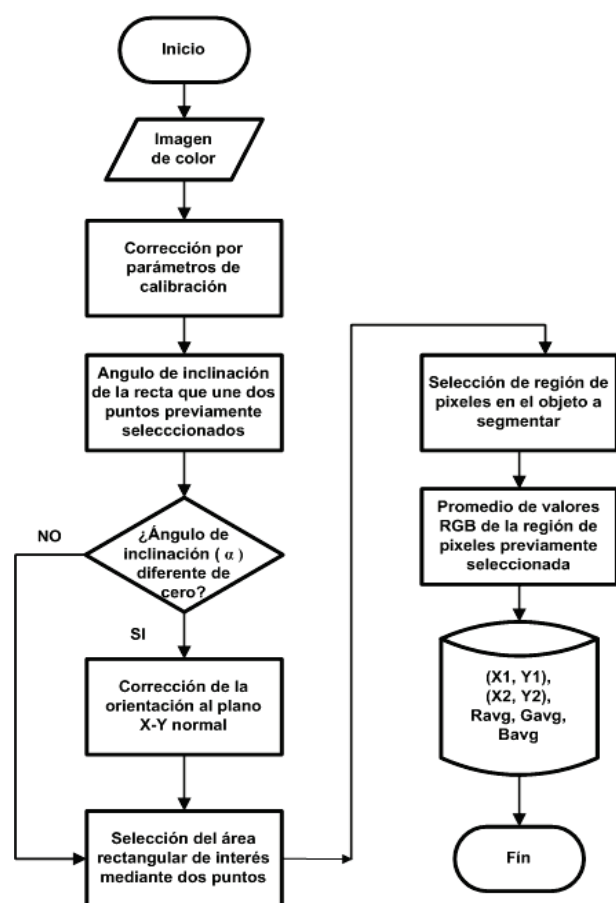


Fig. 4. Diagrama de flujo del proceso de ajuste inicial para realizar medición de distancias en imágenes de color.

Se consideró incluir como parte del ajuste inicial una operación para determinar la posible desviación del plano horizontal con respecto a 0°, que pudieran tener objetos característicos de una imagen muestra del escenario real de

medición. Para llevar a cabo esto, el usuario selecciona dos puntos sobre un borde de línea recta de algún objeto que se supone está en el plano horizontal (0°), y se calcula el ángulo de inclinación ( $\alpha$ ) de la línea recta que une a estos puntos como se indica en el diagrama de flujo mostrado en la fig. 4. Si éste ángulo es diferente de cero se realiza una corrección de la imagen mediante una rotación plana, en magnitud y sentido indicado por el ángulo  $\alpha$ .

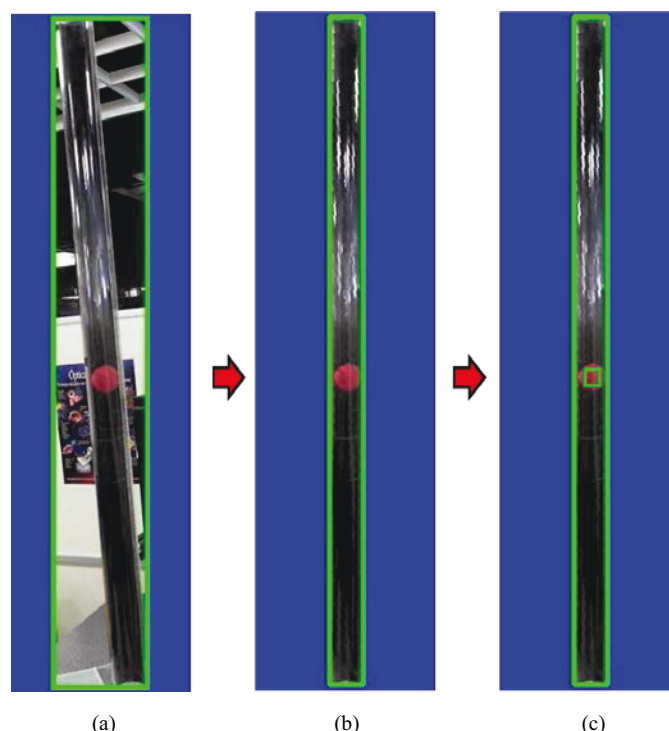


Fig. 5. Se muestran las operaciones de ajuste inicial: (a) y (b) Corrección de la orientación al plano X-Y normal y selección del área rectangular de interés y (c) selección de región de píxeles en el objeto a segmentar.

Posteriormente el usuario selecciona manualmente el área rectangular de interés (área que encierra la trayectoria de desplazamiento del objeto: ver la fig. 5 (b)). De la misma manera, se selecciona una región rectangular de píxeles en el objeto a segmentar, para obtener el promedio de los valores RGB (ver la fig. 5 (c)). El resultado del proceso completo del ajuste inicial se muestra en la fig. 5, por cuestión de presentación, las tres imágenes con resolución de 640x480 píxeles se han girado 90° y recortado en el mismo factor de ancho y alto quedando con un valor de 4.5 cm de ancho por 17 cm de alto.

La técnica de segmentación propuesta extraerá el objeto de interés en una secuencia de operaciones sobre cada imagen de entrada. Considerando que hay variación continua en la uniformidad de la intensidad luminosa (debido a variación de la luz natural del medio ambiente, variación de la luz artificial de lámparas, apertura y cierre de puerta, así como, diferentes niveles de luz reflejada en superficies lisas reflejantes como vidrios, piso, mesas, etc. y en la propia columna cilíndrica de

acrílico) que afecta el color de las diferentes regiones sobre la imagen, provocando intersecciones entre regiones adyacentes; el algoritmo de segmentación desarrollado resolverá esto, para extraer con una buena fidelidad el objeto de interés.

Una vez realizado el proceso de ajuste inicial, el proceso para determinar la distancia del centro de la esfera a la base de la columna, es el que se indica en el diagrama de flujo mostrado en la fig. 6. En él se indica que una vez que una imagen es capturada (posterior al ajuste inicial), se realizan las siguientes tres operaciones de corrección:

- Corrección por parámetros de calibración intrínsecos de la cámara de color.
- Corrección por balance de blancos. Consiste en realizar un balance automático de tonos blancos sobre la imagen, dado que la intensidad luminosa y el tono de diferentes fuentes de luz afecta el color de los objetos [17] y que el cerebro humano y la retina son capaces de percibir y determinar el color de un objeto bajo diferentes condiciones de iluminación, llamándole a esta habilidad *constancia de color* [18, 19].
- Corrección de la orientación de la imagen al plano X-Y normal. Se aplica sólo si en las operaciones de ajuste inicial, se determinó un ángulo  $\alpha$  de desviación diferente de  $0^\circ$ .

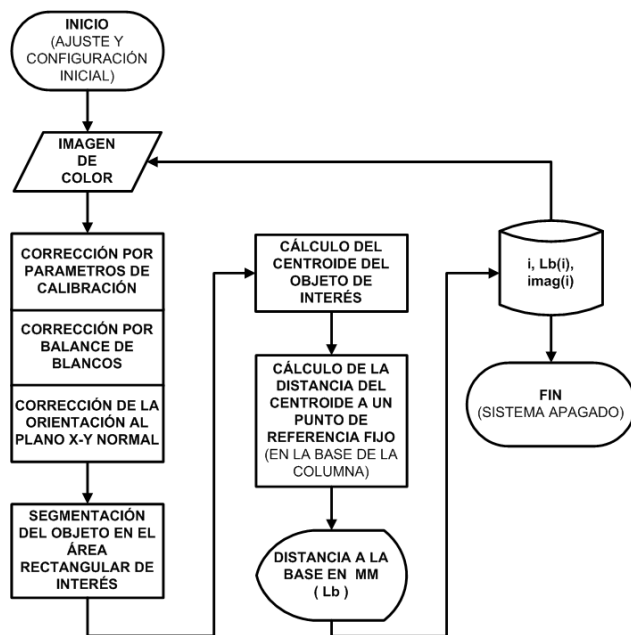


Fig. 6. Diagrama de flujo del proceso de medición de distancia en imágenes de color.

El sensor XBOX Kinect no trae incorporado el autobalance de blancos para las imágenes RGB capturadas mediante su cámara de color, por lo que se elaboró un programa para que en cada una de imágenes capturadas se realice el balance automático de blancos y así obtener los colores más aproximados a colores que correspondan al mundo real,

mismo que es llamado desde el programa principal de procesamiento de imágenes, que lleva a cabo el cálculo de la distancia del centroide de la esfera a la base de la columna. Los sub-programas (funciones) y el programa principal fueron desarrollados en MatLab versión 7.10.0.499 (R2010a).

La selección y extracción automática del área de interés en la imagen de color, para llevar a cabo el proceso de segmentación de la esfera, se realiza tomando en cuenta las coordenadas de los dos puntos seleccionados en el ajuste inicial. El algoritmo de segmentación se lleva a cabo en los siguientes pasos (que forman parte del algoritmo mostrado en la fig. 6):

- Calcular en el área de interés, la distancia euclidiana entre los valores  $RGB$  de cada pixel ' $z$ ' en la posición  $(x, y)$ , y los valores promedio  $RGB$  de los pixeles ' $a$ ' obtenidos en el ajuste inicial.

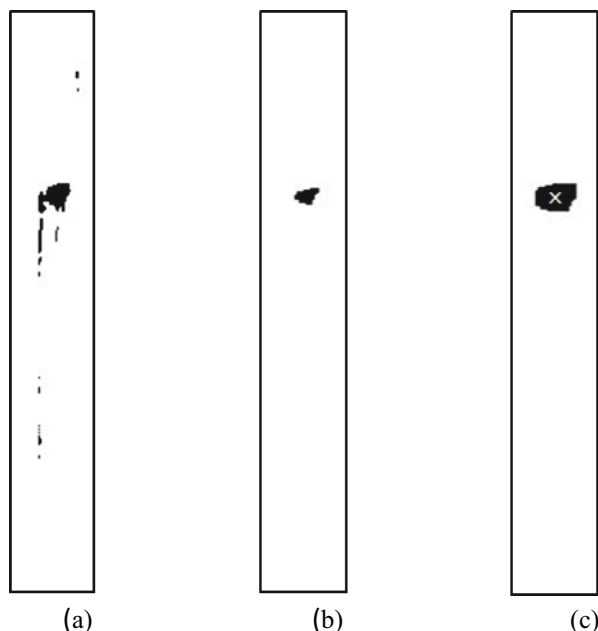
$$D(Z_{RGB}, a_{RGB_{avg}}) = \sqrt{(z_R - R_{avg})^2 + (z_G - G_{avg})^2 + (z_B - B_{avg})^2} \quad (1)$$

donde  $Z_{RGB}$  son los valores  $RGB$  de cada pixel ' $z$ ' en la posición  $(x, y)$ ,  $a_{RGB_{avg}}$  son los valores promedio  $RGB$  de los pixeles ' $a$ ' obtenidos en el ajuste inicial. Así también  $z_R$ ,  $z_G$ , y  $z_B$  son los valores correspondientes de los canales Rojo, Verde y Azul del pixel ' $z$ ' en la posición  $(x, y)$  en la región de interés y  $R_{avg}$ ,  $G_{avg}$ ,  $B_{avg}$  son los valores promedio de los canales Rojo, Verde y Azul de la subregión del objeto a segmentar seleccionada en el ajuste inicial.

- Si el valor de intensidad rojo ( $R$ ) es mayor que los correspondientes valores verde ( $G$ ) y azul ( $B$ ) en cada posición  $x, y$  del pixel en el área de interés, y se cumple además que la distancia de éstos valores  $RGB$  a los valores promedio  $RGB_{avg}$  (calculada en el paso anterior) sea menor que un valor de umbral preestablecido; se verifica entonces que además la distancia del canal rojo al canal verde y azul en cada pixel en dicha posición sea mayor a un valor de umbral preestablecido. Si se cumplen estas tres condiciones, los valores  $RGB$  del pixel en la posición  $(x, y)$  se deja sin cambio y en caso contrario se les asigna un valor de cero; el resultado se convierte a una imagen binaria, esto se ilustra en la fig. 7 (a).
- Se aplican funciones predefinidas de MatLab para extraer las características de detección de elementos conectados y obtener así, el elemento conectado con un área mayor a un valor preestablecido para posteriormente, aplicar técnicas morfológicas de erosión y dilatación para eliminar artefactos de ruido en los objetos segmentados. El resultado logrado en este paso se ilustra en la fig. 7 (b) y 7 (c) respectivamente.
- Para el objeto segmentado en los pasos 2–3 descritos anteriormente, se determina la posición  $(x, y)$  del



centroide de la esfera. En la fig. 7 (c), se puede observar una cruz sobre el objeto segmentado (esfera) dentro de la columna, indicando con una cruz blanca la posición del centroide.



**Fig. 7.** Resultado de la secuencia del algoritmo de segmentación y localización del centroide: a) 1ª. Fase, discriminación de píxeles y binarización de imagen, b) 2ª. Fase, Selección de mayor área de elementos conectados y erosión, c) 3ª. Fase, Dilatación y localización del centroide

Finalmente se determina la distancia en milímetros que hay del centroide del objeto segmentado a un punto de referencia fijo en la base de la columna. Esto se lleva a cabo, estableciendo una relación de proporcionalidad entre la distancia medida en píxeles del centroide al extremo de la base de la columna y la longitud total de la columna en píxeles y en mm; conociendo estos valores a partir de las operaciones de ajuste inicial del sistema de medición. Se imprimen y se guardan como resultados la distancia del centroide de la esfera a la base de la columna, el número de imagen y su nombre.

Como se indica en la fig. 6, el ciclo se mantiene en forma continua y empieza de la misma manera cuando se ha capturado una nueva imagen. El proceso de captura y análisis se detiene cuando ya no hay imágenes capturadas o el sistema se apaga.

#### IV. RESULTADOS OBTENIDOS

Se capturaron 134 posiciones fijas incrementales de la esfera con intervalos de 10 mm a partir de la base de la columna (posición 1: 0 mm) y hasta una altura de 1340 mm (posición 134). Cada posición representó un experimento individual de medición en el que se tomaron 5 imágenes de color RGB, para su procesamiento mediante el algoritmo

propuesto y poder determinar así, la distancia del centroide de la esfera a la base de la columna.

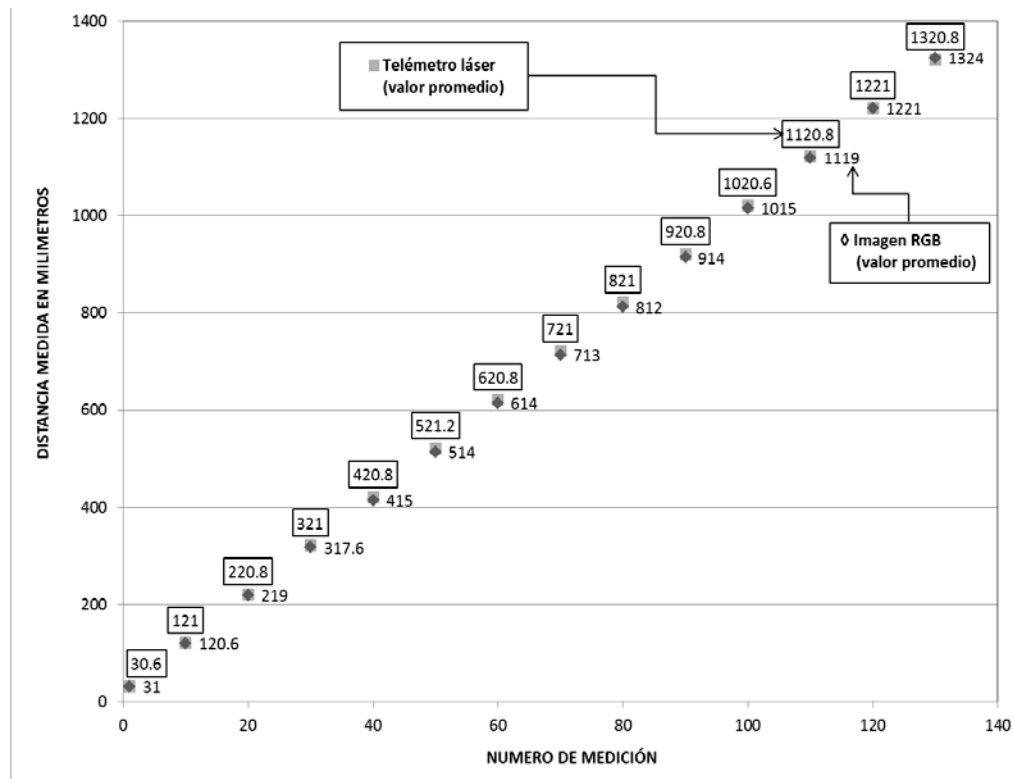
Para ubicar la esfera en la posición fija deseada y como referencia de comparación de las mediciones realizadas, se utilizó un Telémetro Láser Bosch DLE40 [20], que se niveló y ajustó para que quedara de manera fija y estable en la parte inferior de la columna dentro de su base, manteniendo las mismas condiciones de operación del instrumento entre cada lectura de distancia. Una vez que la esfera se posicionaba en forma manual, en la distancia deseada (ajustando la longitud del hilo que permitió suspender la esfera y mantenerla de manera estable en la altura deseada y con ayuda de la lectura del telémetro), el sensor Kinect recibía la orden de capturar 5 imágenes y se procesó cada una de manera independiente para obtener un valor de distancia (mediante el algoritmo mostrado en la fig. 6); así también, por cada bloque de cinco imágenes se registraron también las cinco mediciones correspondientes con el telémetro láser para cada una de las posiciones de la esfera, este proceso se repitió de manera secuencial hasta completar  $134 \times 5$  mediciones, correspondientes al número total de posiciones de la esfera.



**Fig. 8.** La vista frontal del Telémetro Láser Bosch DLE40 [20].

Al finalizar el proceso, se almacenó en un archivo de texto el resultado de las mediciones, en éste archivo se registró el número de la medición, los cinco valores de distancia correspondientes a cada una de las imágenes capturadas y su promedio correspondiente a cada posición de la esfera; para su análisis posterior. En la fig. 8, se muestra una foto de la vista frontal del instrumento de medición de referencia y en la tabla II, se indican sus características de medición así como algunas especificaciones técnicas. Para una referencia completa consultar [20].

Como se menciona en el manual de operación [20], el aparato de medición ha sido proyectado para medir distancias en forma manual de longitudes, alturas, separaciones, y para calcular superficies y volúmenes. El aparato de medición es adecuado para trazar medidas en la construcción tanto en interiores como en exteriores. En el proyecto se utilizó como referencia para verificar el grado de precisión y exactitud de las mediciones de distancia en imágenes RGB.



**Fig. 10.** Gráfica comparativa de las mediciones de distancia realizadas con imágenes RGB, con respecto a las mediciones realizadas con el Telémetro Láser Bosch DLE40; mostrando 14 posiciones representativas de un total de 134.

En la fig. 9, se muestran los resultados obtenidos al realizar mediciones independientes una de otra de 134 posiciones diferentes de la esfera dentro de la columna; la gráfica muestra 14 posiciones representativas de la trayectoria vertical completa.

TABLA II.

CARACTERÍSTICAS DE MEDICIÓN Y ESPECIFICACIONES TÉCNICAS DEL TELÉMETRO LÁSER BOSCH DLE40.

Especificación	Valor
Clase del láser	2
Tipo de láser	635 nm, < 1mW
Rango de medición (en interiores)	0.05 – 40 m
Precisión de medición (típica)	± 1.5 mm
Pilas de consumo	4 x LR03 (AAA) (incluidas)
Peso aprox. con batería	0.18 kg
Dimensiones	58 x 100 x 32 mm

En el eje X se muestra el número de la medición realizada individualmente, en cada posición y en el eje Y se muestra el valor de la distancia medida de dicha posición en milímetros (mm) utilizando en su representación el valor promedio del resultado de las 5 imágenes RGB y de la misma manera el valor promedio de las 5 lecturas del Telémetro Láser respectivamente.

Como se puede observar las mediciones de distancia hechas mediante las imágenes RGB se aproximan a las lecturas

tomadas en el instrumento. Como se observa en la gráfica, los valores promedio correspondiente al telémetro láser se muestran encerrados en un recuadro para diferenciarlos de los valores correspondientes al resultado de distancia promedio medida mediante las imágenes RGB.

Se calculó la magnitud del error relativo porcentual para cada una de las 134 posiciones diferentes mediante la ecuación (2),

$$\% E_{rel.} = \frac{E_{abs.}}{V_{real}} \times 100 = \frac{|V_{real} - V_{med.}|}{V_{real}} \times 100 \quad (2)$$

donde  $\% E_{rel.}$  es igual al porcentaje de error relativo,  $E_{abs.}$  es el error absoluto que se obtiene al calcular la diferencia absoluta entre el valor real de referencia ( $V_{real}$ ) y el valor medido ( $V_{med.}$ ), en nuestro caso, el valor real de referencia es la lectura de distancia del telémetro láser y el valor medido es la distancia resultante en la imagen RGB.

En la gráfica representada en la fig. 11, se puede observar la distribución poblacional del error relativo porcentual de las 134 mediciones realizadas independientemente una de otra. Se puede notar como se tiene una desviación máxima de 1.64%. El 100% de las mediciones realizadas tuvieron un porcentaje de error relativo que se mantuvo en un rango del 0 a 1.64%; en un rango de distancia de 0 a 1340 mm sobre el eje Z, paralelo al plano de imagen en la cámara (ver la fig. 3).

Las magnitudes del error absoluto y del error relativo porcentual es variable de una medición a otra ya que hay factores no controlados que afectan las mediciones: a) Nivel



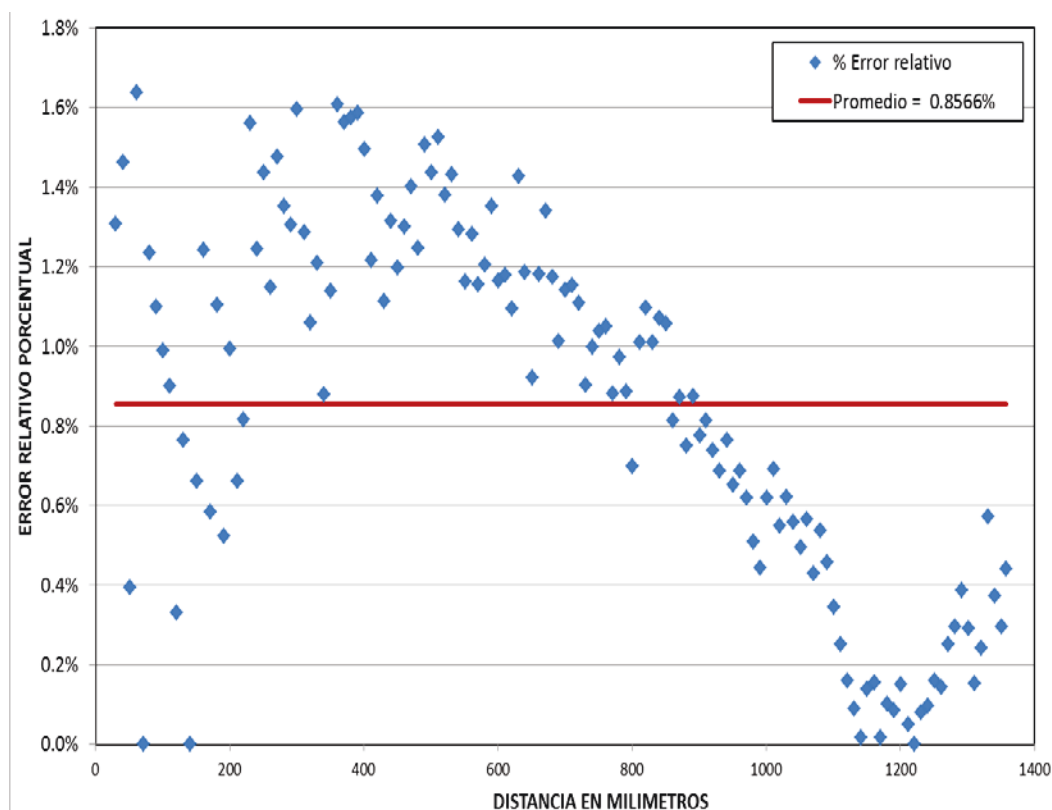


Fig. 11. Grafica que muestra la distribución poblacional del error relativo porcentual para un total de 134 mediciones realizadas.

de iluminación de escenario, b) reflexión y refracción no homogénea de la luz en la columna de acrílico transparente ya que el objeto de interés se encuentra dentro de ésta, y c) exactitud y precisión del instrumento de referencia (Telémetro Láser) al realizar cada medición.

En la tabla III, se muestran los valores máximo y mínimo del total de las desviaciones típica estándar (calculadas para cada 5 mediciones correspondientes a cada posición de la esfera), utilizando ambos métodos, el telémetro láser y el sistema de medición mediante imágenes RGB.

TABLA III.

VALOR MÁXIMO Y MÍNIMO DE LA DESVIACIÓN TÍPICA ESTÁNDAR, OBTENIDAS DE 134 POSICIONES DE LA ESFERA Y REALIZANDO 5 MEDICIONES POR POSICIÓN, UTILIZANDO AMBOS MÉTODOS: TELÉMETRO LÁSER Y SISTEMA MEDIANTE IMÁGENES RGB, INDICANDO TAMBIÉN LA OCURRENCIA DE ESTOS VALORES. LA  $\sigma$  REFIERE AL VALOR DE DESVIACIÓN TÍPICA ESTÁNDAR.

Método	Máximo $\sigma$	Ocurrencia (repetición del valor)	Mínimo $\sigma$	Ocurrencia (repetición del valor)
Imágenes RGB	1.96	2	0	109
Telémetro láser	0.49	21	0	74

Como se observa en la tabla III, el máximo valor de desviación típica estándar es mayor en el método de imágenes RGB que en el método de telémetro láser pero solamente se obtuvo este valor en dos grupos de mediciones

correspondientes a dos posiciones; mientras que utilizando el método del telémetro la ocurrencia de la máxima desviación típica estándar fue de 21. Y el mínimo valor de desviación típica estándar en ambos métodos fue de cero, sin embargo, fue más alta en el método de imágenes RGB en un 47.3% con respecto al método del telémetro láser. Indicando con esto que, el método de imágenes RGB tiene una precisión más alta que el método de telémetro láser para medir distancias bajo las mismas condiciones del experimento realizado.

## V. COCLUSIONES

El sistema de medición propuesto para determinar la posición de la esfera, dio como resultado que el 100% de las mediciones realizadas tuvieron un error relativo porcentual en un margen de 0 a 1.64%, con un buen nivel de precisión. Debido a esto, se han generado buenas expectativas para su incorporación en sistemas de monitoreo y control automático de procesos.

Para la integración de éste sistema de medición a un sistema automático de control de posición de una esfera levitada neumáticamente, se deben tomar en cuenta los siguientes factores en su diseño: a) La esfera se encontrará en constante movimiento y su posición no se mantendrá fija, (aun cuando la velocidad del motor se mantenga constante); b) la estabilidad en una posición deseada dependerá del grado de turbulencia del aire a su alrededor y c) el movimiento será

errático y oscilante (girá alrededor de su centro de gravedad en diferentes direcciones y se moverá con desplazamientos en los tres ejes  $X$ ,  $Y$ ,  $Z$ ), para una velocidad definida del motor de c.d. Por lo que, como trabajo futuro es importante realizar el análisis del desempeño de éste sistema de medición de distancia en el entorno dinámico real en el que se integrará.

#### REFERENCIAS

- [1] H. M. Dee, S. A. Velastin, "How close are we to solving the problem of automated visual surveillance?," *Machine Vision and Applications*, vol. 19, no. 5–6, 2008, pp. 329–343.
- [2] P. Vera, D. Zenteno, J. Salas, "Counting Pedestrians in Bidirectional Scenarios Using Zenithal Depth Images," *Proceedings of 5th Mexican Conference, MCPR*, Querétaro, Mexico, junio 26–29, 2013, pp. 84–93.
- [3] R. T. Chin, C. A. Harlow, "Automated visual inspection: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, 1982, pp. 557–573.
- [4] E. Royer et al., "Monocular vision for mobile robot localization and autonomous navigation," *International Journal of Computer Vision*, vol. 74, no. 3, 2007, pp. 237–260.
- [5] A. Criminisi, I. Reid, A. Zisserman, "A plane measuring device," *Image and Vision Computing*, vol. 17, no. 8, 1999, pp. 625–634.
- [6] Y. Wan et al., "A Study in 3D-Reconstruction Using Kinect Sensor," *8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, IEEE, 2012, pp. 1–7.
- [7] T. Bucher, "Measurement of distance and height in images based on easy attainable calibration parameters," *Proceedings of the IEEE in Intelligent Vehicles Symposium, IV*, 2000, pp. 314–319.
- [8] J. L. Lázaro et al., "Sensor for distance estimation using FFT of images," *Sensors*, 2009, vol. 9, no. 12, p. 10434–10446.
- [9] A. Criminisi, *Accurate visual metrology from single and multiple uncalibrated images*. Springer, 2001.
- [10] P. Wright, "Single View Metrology: Measuring the Dimensions of Objects From a Single Image." University of York, Computer Science Department, 2006.
- [11] R. J. Moffat, "Describing the uncertainties in experimental results," *Experimental thermal and fluid science*, vol. 1, no. 1, 1988, pp. 3–17.
- [12] MSDN, Microsoft, <http://msdn.microsoft.com/en-us/library/jj131033>, consultado el 30/08/2013.
- [13] Kinect, <http://wiki.ros.org/kinect>, consultado el 23/01/2012.
- [14] Wikipedia, <http://commons.wikimedia.org/wiki/File:Xbox-360-Kinect-Standalone.png>, recuperado el 15/06/2013.
- [15] A. Criminisi, I. Reid, A. Zisserman, "A plane measuring device," *Image and Vision Computing*, vol. 17, no. 8, 1999, pp. 625–634.
- [16] J.-Y. Bouguet, "Camera calibration toolbox for Matlab", [http://www.vision.caltech.edu/bouguetj/calib\\_doc/2004](http://www.vision.caltech.edu/bouguetj/calib_doc/2004).
- [17] R. Ramanath et al., "Color image processing pipeline," *Signal Processing Magazine*, IEEE, vol. 22, no. 1, 2005, pp. 34–43.
- [18] S. Zeki, Semir. *A Vision of the Brain*. Blackwell scientific publications, 1993.
- [19] M. Ebner. *Color constancy*. Wiley, 2007.
- [20] Bosch DLE40. <http://www.boschmexico.com.mx/telemetro-laser-dle-40-professional>, recuperado el 01/04/2013.
- [21] R. Doskocil et al., "Measurement of distance by single visual camera at robot sensor systems," *MECHATRONIKA, 15th International Symposium*, IEEE, 2012.