

EVALUACIÓN DE TÉCNICAS DE ESTIMACIÓN ÓPTICA DE MEDIDAS DE PROFUNDIDAD

ASSESSMENT OF ESTIMATION TECHNIQUES OPTICS DEPTH MEASURES

Msc. (c) J.E. Vivas, Ing. S.C. Imbachi, y Msc. J.F. Florez Marulanda*

***Universidad del Cauca.** Departamento de Electronica, Instrumentation y Control, I+D
Grupo de Automática industrial.
Calle 5 # 4 – 70, Popayán, Cauca, Colombia.
Tel.: (+57) (2) 8209800
eduardovivas, simbachi, jflorez@unicauca.edu.co

Resumen: Se presenta un estudio comparativo de tres métodos de medición de profundidad por procesamiento óptico y digital de imágenes, analizando diferentes distancias a 0.67, 1.5, 2, 2.5, y 4 m de dos pares de sistemas de medición conformados por un sistema de visión estéreo, un escáner laser y un sistema de captación de movimiento integrado (Kinect), con el objetivo de conocer el adecuado rango de operación de cada sistema. Se observó que el de mayor precisión en la medida es el escáner laser con un error porcentual general del 1.38%.

Palabras clave: Medición, Profundidad, Procesamiento, Comparativo, Óptico.

Abstract: It discloses a comparative study of three methods of measuring depth by optical and digital image processing, analyzing different distances (0.67, 1.5, 2, 2.5, and 4 m) on two pairs of measurement systems comprised of a system of stereo vision, a laser scanner and a motion capture system integrated (Kinect). From which it was observed that the highest measurement accuracy is the laser scanner with an overall percentage error of 1.38%.

Keywords: Measurement, Depth, Processing, Comparative, Optical.

1. INTRODUCCIÓN

Los sistemas ópticos adquieren información del entorno de múltiples formas (Chen y Brown, 2000), de forma clásica por características visuales del ambiente (color, intensidad, contraste, iluminación, etc.) (Chen y Brown, 2000, Torrontegi, 2010, Mattocchia, 2011); por extracción de características en el rango no visible como el infrarrojo (Muñoz, Imbachi y Florez, 2014, Lili, 2010), al igual que combinación de los dos anteriores métodos (Zhang, 2012). Una medida de gran interés de adquisición por medios ópticos es la profundidad (Florez, 2009); ya que ayuda a la

orientación de sistemas robóticos (Fu, 1989) comparación de nivel, entre otras (Revelo, Usuma y Florez, 2012). Se han utilizado distintos métodos para este objetivo, desde la visión estéreo (Khoshelham y Elberink, 2012) y configuraciones adecuadas de este mismo (Montalvo, 2009), utilización de elementos especializados en esta medición como el radar laser (Muñoz, Imbachi y Florez, 2014); hasta adecuación de sistemas diseñados para otra funcionalidad dándose el caso del Kinect (Khoshelham y Elberink, 2012, Han, Shao y Xu, 2013).

En este artículo se expondrá la estimación de medidas por sistemas ópticos (Chen y Brown, 2000), comparando dos pares de ellos el sistema de visión estéreo (Mandelbaum y McDowel, 1998) con el escáner láser (Lohmann, Koch, y Schaeffe, 2000) y el Kinect (Matyunin y Vatolin, 2011) con el escáner láser. En el presente trabajo se da a conocer tres tipos de métodos ópticos de medición de la profundidad, al igual que la comparación de estos; con lo cual se facilita la elección y adecuación de cada uno de ellos para casos específicos que se presenten en problemas reales.

En la sección de metodología y materiales, se presenta el hardware y software usados en la investigación y su implementación. En resultados y discusión se plantea toda la información obtenida en las distintas pruebas, y manejo de ella, donde se evidencia que el de mayor precisión es el escáner láser. Finalizando con un análisis complementario de los resultados en la sección de conclusiones (Caballero A. et al., 2013).

2. METODOLOGIA Y MATERIALES

Para el desarrollo del estudio se utilizaron: un radar láser Allen Bradley referencia Safe Zone Scanner Laser multizone (LIDAR) (Allen Bradley 2005), dos cámaras Life Cam HD- 5000 (Revelo, Usuma y Florez, 2012), un Kinect (Khoshelham y Elberink, 2012), un cable serie RS232, una fuente de alimentación de 12 V / 3 Amp, un convertor USB – Serie y un computador con procesador Intel core i5 de 2.3 GHz y memoria RAM de 4GB. Dividiendo su implementación en un componente hardware (Fig. 1) que realiza la unión y alineación de los sensores para obtener la misma información (Allen Bradley 2005), esto en dos montajes: visión estéreo - escáner láser (Fig. 1.a) y Kinect – escáner láser (Fig. 1.b).

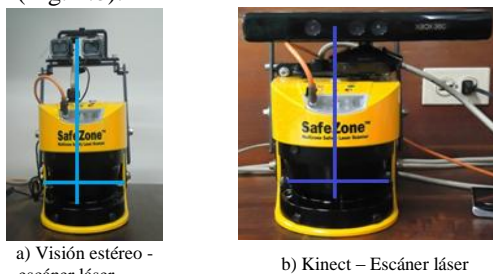


Fig. 1. Componente hardware para captación de datos.

El componente software de la implementación se desarrolló en Open Cv (Creus, 2011) y en lenguaje C++ con interfaz Qt (Burrus, 2011, Campbell, Sukthankar, and Nourbakhsh 2004), con

un diseño modular para la captación y procesamiento de la información proveniente de los sensores.

Procesamiento visión estéreo

La señal entregada por el sistema de visión estéreo, reconstrucción 3D del entorno (Fig. 2), entra al Sub Módulo A1 donde se desarrolla el procedimiento para obtener la matriz de distancias a un objeto presente en una escena, la señal que entrega este módulo, Matriz de Distancias, pasa al Sub Módulo B1 donde se determina el perfil de distancias a usar, correspondiente al perfil observado por el escáner láser, representado en la señal Perfil de Distancias y la señal $\delta_{\text{visión}}$, que representa el vector con la desviación estándar de las medidas de distancias obtenidas del perfil correspondiente



Fig. 2. Procesamiento modular del sistema de visión estéreo.

Procesamiento Láser

La señal en 2D (Fig. 3), entregada por el escáner láser, entra al Sub Módulo A2 donde se lleva a cabo un procedimiento para obtener el vector de profundidades consignado en la señal Perfil de Profundidades; esta señal entra al Sub Modulo B2 que realiza la conversión de profundidades a distancias, representado en la señal Perfil de Distancias Escáner Láser y calcula la desviación estándar de las medidas de distancia, determinada por la señal $\delta_{\text{láser}}$.



Fig. 3. Procesamiento modular del sistema láser.

Procesamiento Kinect

Los algoritmos empleados para la obtención de datos de profundidad se basaron en el software de código abierto RGBDemo, que tiene como objetivo proporcionar una herramienta para trabajar con el Kinect (Ballester y Pheatt, 2013) y desarrollar programas de visión artificial independientes, sin los inconvenientes de integrar bibliotecas de uso privativo (Burrus, 2011).

RGBDemo utiliza las librerías Openni y libfreenect; tiene archivos de calibración generales para el Kinect, pero permite realizar una calibración personalizada para cada Kinect con

libfreenect; y realiza una calibración del Kinect, basado en el método de visión estéreo desarrollado con las librerías de OpenCv y obtiene los datos de profundidad de los objetos presentes en una escena, mediante los algoritmos calibrar-Kinect-ir.cpp y viewer.cpp.

En viewer.cpp se encuentra las opciones Show IR, Dual IR/RGB, en el menú Capture, y Grab all frames, en el menú File, que permiten la captura y almacenamiento de imágenes tomadas con el Kinect. Dual IR/RGB permite activar la captura de imágenes con el sensor IR y la cámara RGB simultáneamente (**¡Error! No se encuentra el origen de la referencia.** 4). La opción Grab all frames guarda las imágenes en un directorio accesible para el algoritmo de calibración.

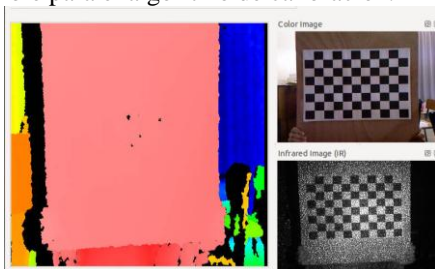


Fig. 4. Captura realizada con el Kinect, opción Dual IR/RGB. Interfaz RGBDemo

El grupo de imágenes duales capturadas se usa por el algoritmo calibrar-Kinect-ir.cpp para generar los archivos de calibración calibration_rgb.yaml y calibration_depth.yaml, que contienen los parámetros intrínsecos y extrínsecos de las cámaras RGB e IR respectivamente; estos son utilizados para calcular la distancia a la que se encuentran los objetos del Kinect (Burrus, 2011).

La señal detección de objetos (Fig. 5), entregada por el software del Kinect, entra al Sub Modulo A3 que realiza un procedimiento para obtener la señal Matriz de Distancias, que almacena los datos de las distancias a un objeto en una escena, en una matriz de 640x480 datos; esta señal entra en el Sub Modulo B3 que determina el perfil de distancias que corresponde con el perfil del Kinect, denotado con la señal Perfil de Distancias del láser y entrega la señal δ_{Kinect} que representa la desviación estándar de las distancias medidas.



Fig. 5. Procesamiento modular del sistema Kinect.

3. RESULTADOS Y DISCUSIÓN

Las medidas captadas por los pares de sensores fueron contrastadas y organizadas de acuerdo al rango de medición de cada uno de ellos.

Visión estéreo – escáner láser

Se empleó un escenario de pruebas (Fig. 6), constituido por una mesa de 70 cm sobre la que se colocó dos cubos de cartón uno de ellos con una de sus caras perpendiculares hacia el eje de proyección y el otro con una leve inclinación, para tomar los datos con el sistema de visión estéreo – escáner láser. Del mismo se realizó una captura por la cámara RGB de la cámara de visión estéreo (Fig. 6.a); y la correlación del escáner láser se indica con una línea (Fig. 6.b). Se usó el mismo escenario de pruebas para las distancias de 0.68, 1.5 y 2 m.



a) Captura con cámara RGB.



b) Correlación con el escáner láser

Fig. 6. Escenario de pruebas para visión estereo – escáner láser a 0.68, 1.5 y 2 m

A. 1. Medidas de prueba

Los resultados obtenidos al realizar la medida a 0.68, 1.5 y 2 m, se ordenaron (Tabla 1) conforme a las distancias promedio tomadas con la visión estéreo (Z Visión), el escáner láser (Z Láser) y el error absoluto asociado a cada uno de ellos (EA visión y EA Láser).

Tabla 1 Resultado promedio de un dato visión estéreo – escáner láser a 0.68, 1.5 y 2 m

Medida (m)	Z Visión (m)	EA Visión (m)	Z Láser (m)	EA Láser (m)
0.68	0,37	0,31	0,7	0,02
1.5	0,71	0,79	1,54	0,04

2	0,38	1,62	2,01	0,01
---	------	------	------	------

De acuerdo a la Tabla 1 se observa que la medición de profundidad: a 68 cm por visión estéreo es de 31 cm, mientras que por medio del escáner láser esta presenta un error de 2 cm. A 1.5 m por visión estéreo presenta un error de 79 cm, mientras que el escáner láser tiene un error en la medida de 4 cm. A 2 m la visión estéreo tiene un error en la medición de 162 cm y el escáner láser un error de 1 cm.

Kinect – escáner láser

Para el Kinect – escáner láser se cambió de escenario de pruebas para la medición a 0.67 m (Fig. 7.a), este está conformado por una mesa de 70 cm con 4 cubos de cartón todos con una de sus caras perpendicular al eje de proyección de los sensores, la prueba a 2.5 m se hace en un escenario distinto (Fig. 8.a) constituido por diferentes asientos, una mesa de 70 cm y dos cubos de cartón, y a 4.0 m se usó el escenario anterior a una mayor distancia (Fig. 9.a); en todos ellos se analiza la imagen en escala de colores que se obtiene del análisis con el sensor infrarrojo del Kinect (Fig. 7.b, 8.b, 9.b), donde se observan las zonas oscuras que dan origen a datos erróneos en el Kinect junto al perfil de línea de la escena que correlaciona con el láser.



a) Captura con cámara RGB

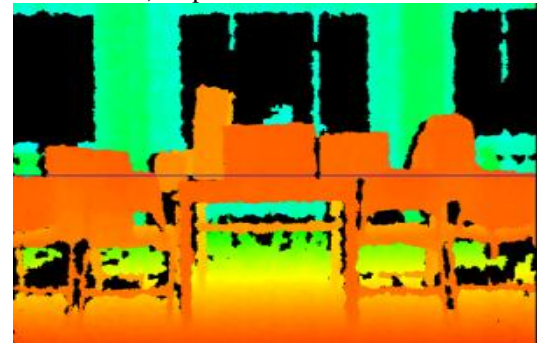


b) Imagen en escala de colores con línea de correlación láser

Fig. 7. Escenario de pruebas para el Kinect – Escáner láser a 0.67 m



a) Captura con cámara RGB.

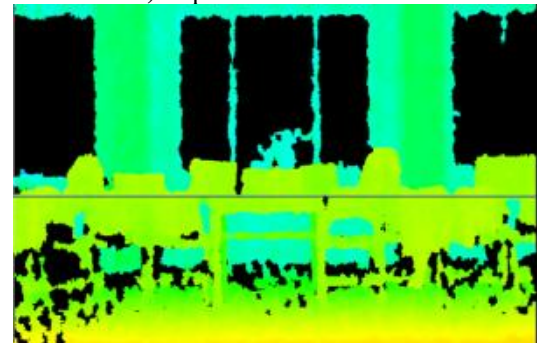


b) Imagen en escala de colores con línea de correlación láser

Fig. 1. Escenario de pruebas de Kinect – Escáner láser a 2.5 m



a) Captura con cámara RGB.



b) Imagen en escala de colores con línea de correlación láser

Fig. 9. Escenario de pruebas de Kinect – Escáner

láser a 4 m.

B. 1. Medidas de prueba

Los datos de las pruebas se ordenaron (Tabla 2) conforme a las distancias promedio tomadas con el Kinect (Z Kinect), el escáner láser (Z Láser) y el error absoluto asociado a cada uno de ellos (EA Kinect y EA Láser).

Tabla 2. Resultado promedio de un dato a 0.67 m

Medida (m)	Z Kinect (m)	EA Kinect (m)	Z Láser (m)	EA Láser (m)
0,67	0,66	0,01	0,68	0,01
2,5	2,54	0,03	2,52	0,01
4	4,12	0,11	4,02	0,01

De la Tabla 2, en la prueba a 0.67 m se observa que la medición tanto del Kinect como del escáner láser tienen un error de 1 cm; al realizar la medición a 2.5 m el Kinect da un error en la medida de 3 cm y el escáner láser un error absoluto de 1 cm; alcanzando la mayor diferencia entre las medidas a 4 m ya que el Kinect tiene un error de 11 cm y el láser un error de 1 cm.

Los resultados obtenidos en las distintas medidas, se analizaron y ordenaron (Tabla 3) presentando el error porcentual de cada medición.

Tabla 3. Resultados globales de las mediciones

Contraste de error	Medida (m)			Promedio (%)
	0,68	1,50	2,00	
Visión estéreo (%)	45,59	52,67	84,00	60,75
Escáner Laser (%)	2,94	2,67	0,5	2,04
Contraste de error	Medida (m)			Promedio (%)
	0,67	2,51	4,01	
Kinect (%)	1,49	1,2	2,74	1,81
Escáner Laser (%)	1,49	0,4	0,25	0,72

Se observa de la Tabla 3 que los elementos de mediciones ordenados conforme a su precisión son el Escáner láser, Kinect, Visión estéreo. Este resultado es debido a que el escáner láser es un elemento especializado para la medición de la profundidad y no basa su procesamiento por análisis digital de imágenes sino por tratamiento óptico de las medidas de su láser interno y

posterior respuesta digital; por otro lado el Kinect es la unión de procesamiento óptico y digital de imágenes, ya que analiza el entorno por imágenes RGB y el despliegue de una matriz de puntos infrarrojos que los percibe por un sensor monocromático CMOS (Moreno et al., 2013), sin embargo posee un rango de visión limitado; por otra parte el sistema de visión estéreo sin compensación analiza el entorno por métodos de comparación de iluminación, contraste y reconstrucción de relieve y forma, lo que únicamente converge en un procesamiento digital de imágenes dependiente de las características visuales que tenga el ambiente a su alrededor.

4. CONCLUSIONES

Se realizó medidas de profundidad a 0.68, 1.5 y 2 m con un sistema de visión estéreo incorporado a un escáner láser, posterior se hicieron medidas a 0.67, 2.5 y 4 m con un Kinect incorporado a un escáner láser, finalmente se comparó la información de los dos grupos de profundidades conforme a los dos sistemas de medición. En el primer grupo, el sistema de visión estéreo es el que presenta errores altos ya que utiliza estimaciones visuales para la reconstrucción 3D y subsiguiente análisis de profundidad por procesamiento digital. En el segundo grupo de profundidades, no hay mucha diferencia entre los errores del Kinect y el escáner láser, pero las variaciones de incertidumbre del Kinect se deben a la adquisición de datos por imágenes bidimensionales y matriz de proyección de puntos infrarrojos para la profundidad, los cuales presentan puntos muertos por el procesamiento en escala de colores de los objetos.

Por otro lado, el escáner láser en los dos grupos de medidas tiene gran precisión, ya que analiza su entorno no por imágenes de proyección en los ejes X, Y, Z (método de visión estéreo) o imágenes bidimensionales en el plano XY y matrices de proyección en Z (Kinect), sino reconstrucción matricial del entorno en el plano XZ.

AGRADECIMIENTO

Se agradece a la Universidad del Cauca por el apoyo prestado en la realización de este trabajo.

REFERENCIAS

Ballester, J., & Pheatt, C. (2013). Using the Xbox Kinect sensor for positional data acquisition. *American Journal of Physics*, 81(1), 71. <http://doi.org/10.1119/1.4748853>

- Campbell, J., Sukthankar, R., & Nourbakhsh, I. (2004). Visual Odometry Using Commodity Optical Flow. *Intelligent Robots and Systems (IROS)*, 1008–1009.
- Lohmann, P., Koch, A., & Schaeffer, M. (2000). Approaches to the filtering of laser scanner data. *International Archives of Photogrammetry and Remote Sensing*, 33(B3/1; PART 3), 540–547.
- Mandelbaum, R., McDowell, L., Bogoni, L., Reich, B., & Hansen, M. (1998). Real-time stereo processing, obstacle detection, and terrain estimation from vehicle-mounted stereo cameras. *Proceedings of the Fourth IEEE Workshop on Applications of Computer Vision. WACV'98*, (1), 288–289.
<http://doi.org/10.1109/ACV.1998.732909>
- Matyunin, S., Vatolin, D., Berdnikov, Y., & Smirnov, M. (2011). Temporal filtering for depth maps generated by Kinect depth camera. *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 3DTV-CON 2011 - Proceedings*.
<http://doi.org/10.1109/3DTV.2011.5877202>
- Caballero Amaury, Velasco Gabriel, Pardo García A. (2013). DIFFERENTIATIONS OF OBJECTS IN DIFFUSE DATABASES. *Revista colombiana de tecnologías de Avanzada*. 2 (22). Pág. 131 – 137.
- Revelo-Luna, D. A., Usama, F. D., & Flórez-Marulanda, J. F. (2012). Reconstrucción 3D de escenas mediante un sistema de visión estéreo basado en extracción de características y desarrollado en openCV. *Ingeniería Y Universidad*, 16(2), 485–500.
- Creus Solé, A. (2011). "Generalidades" (Octava Edición ed.). Impreso en Colombia: Alfaomega.
- Burrus, N. (2011). RGBDemo. (Manctl) Recuperado el Agosto de 2012, de Nicolas Burrus Home Page:
<http://labs.manctl.com/rgbdemo/index.php/Main/HomePage>
- Ballester, J., & Pheatt, C. (2013). Using the Xbox Kinect sensor for positional data acquisition. *American Journal of Physics*, 81(1), 71.
<http://doi.org/10.1119/1.4748853>
- Campbell, J., Sukthankar, R., & Nourbakhsh, I. (2004). Visual Odometry Using Commodity Optical Flow. *Intelligent Robots and Systems (IROS)*, 1008–1009.
- Lohmann, P., Koch, A., & Schaeffer, M. (2000). Approaches to the filtering of laser scanner data. *International Archives of Photogrammetry and Remote Sensing*, 33(B3/1; PART 3), 540–547.
- Mandelbaum, R., McDowell, L., Bogoni, L., Reich, B., & Hansen, M. (1998). Real-time stereo processing, obstacle detection, and terrain estimation from vehicle-mounted stereo cameras. *Proceedings of the Fourth IEEE Workshop on Applications of Computer Vision. WACV'98*, (1), 288–289.
<http://doi.org/10.1109/ACV.1998.732909>
- Matyunin, S., Vatolin, D., Berdnikov, Y., & Smirnov, M. (2011). Temporal filtering for depth maps generated by Kinect depth camera. *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 3DTV-CON 2011 - Proceedings*.
<http://doi.org/10.1109/3DTV.2011.5877202>
- Revelo-Luna, D. A., Usama, F. D., & Flórez-Marulanda, J. F. (2012). Reconstrucción 3D de escenas mediante un sistema de visión estéreo basado en extracción de características y desarrollado en openCV. *Ingeniería Y Universidad*, 16(2), 485–500.
- Moreno Rubio J., Jiménez López A, Barrera Lombana N. (2013). El amplificador de potencia de carga sintonizada. *Revista colombiana de tecnologías de Avanzada*. 2(22). Pág. 9 – 13.