



Desarrollo de un prototipo de escaneo móvil, empleando iluminación estructurada que permita la reconstrucción 3D usando una geometría bifocal

Development of a mobile scanning prototype, using structured illumination that allows 3D reconstruction using bifocal geometry



Ontuña Narváez Joel Sebastián¹, Campuzano Nieto Gonzalo Bayardo¹

1. Universidad de Cuenca, Cuenca

joelontu@gmail.com

DOI:



Resumen

El presente artículo pretende evidenciar cómo es el desarrollo de un prototipo para escanear cuyo aporte es la iluminación estructurada para reconstruir en 3D por medio de geometría bifocal. Detallar el procedimiento para escaneo móvil permite presentar un insumo para replicar este prototipo. Además, se detallan el algoritmo de segmentación de línea láser dentro del procesamiento de datos. Por tanto, su aplicabilidad permite mayor agilidad en la reconstrucción 3D, ahorrando procesos en la localización de puntos medulares.

Palabras clave: *fotometría esférica, imágenes, colinealidad, coplanaridad*

Abstract

This article aims to demonstrate the development of a prototype for scanning whose contribution is the structured illumination to reconstruct in 3D by means of bifocal geometry. Detailing the procedure for mobile scanning allows to present an input to replicate this prototype. In addition, the laser line segmentation algorithm within the data processing is detailed. Therefore, its applicability allows greater agility in 3D reconstruction, saving processes in the location of medial points.

Key words: *fotometría esférica, imágenes, colinealidad, coplanaridad*

Introducción

El escáner láser es usado en diversos campos e investigaciones académicas, por ejemplo, en el diseño de productos, industria automotriz y ciencia médica. También se utiliza para registrar edificios, especialmente en lugares a los que las personas no pueden acceder debido a los riesgos de seguridad (Ebrahim, 2014).

En el año 2012, se realizó un estudio sobre el uso de un sistema láser usando la visión estéreo binocular (dos cámaras paralelamente alineadas), que trata de explicar cómo se puede obtener información 3D a partir de escenas 2D, donde se encuentra proyectada una línea láser (Lv & Zhang, 2012).

Una visión estéreo binocular es un caso estándar; el caso más general implica que las cámaras no se encuentren perfectamente alineadas y que deban pasar por un proceso de re-proyección para que las imágenes obtenidas de ambas cámaras sean coplanares y se encuentren perfectamente alineadas. El proceso señalado anteriormente es conocido como rectificación de imágenes. Al aplicar esta rectificación a un par de imágenes estéreo se puede perder algo de información de la escena y lo que se requiere buscar es la manera de minimizar la pérdida de información. La línea láser es un factor que puede agilizar el proceso de búsqueda de píxeles tanto de la imagen izquierda como de la imagen derecha que pertenecen a un mismo punto en la escena.

Metodología

1) a) Posición de las cámaras:

En *Visión por Computador*, de Domingo Mery, se dice que: “En el análisis bifocal se

tiene un sistema de visión con dos cámaras, o bien una sola cámara que toma dos imágenes del objeto de estudio en dos tiempos distintos, suponiendo que en ese tiempo la cámara o el objeto se han movido” (Mery, 2004).

En el gráfico que se visualiza a continuación, se puede notar que se hace uso del caso más general de la geometría bifocal. En la práctica, las cámaras no estarán perfectamente alineadas ni colineales (Kaehler & Bradski, 2016).

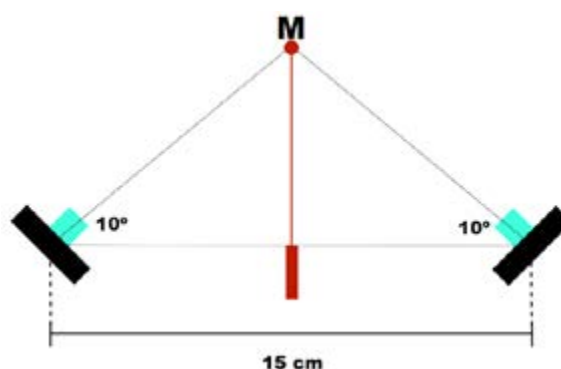


Figura 1 Geometría bifocal general

Más adelante, se verá más a detalle el proceso que se debe seguir para obtener un par de imágenes perfectamente alineadas y colineales; este proceso es conocido como rectificación de imágenes.

b) Calibración estéreo de Cámaras:

La calibración es un proceso que consiste en hallar los parámetros intrínsecos y extrínsecos de una cámara para poder obtener una relación entre las unidades del mundo real y las unidades de la cámara, es decir, milímetros y píxeles respectivamente (Ontuña, Marco Teórico, 2019).

Existen diversos métodos de calibración de cámaras, pero el más usado es el método propuesto por Zhang, que toma varias fotos de un patrón de tablero de ajedrez. Aquí se utilizó un patrón de ajedrez de 7X1 que se puede ver en el trabajo de titulación de Joel Ontuña (Ontuña, Metodología, 2019). De aquí en adelante se lo nombrará el patrón de calibración. Se toma fotos del patrón de calibración en distintas posiciones con ambas cámaras de tal forma que se pueda ver tanto en la cámara izquierda como en la cámara derecha.

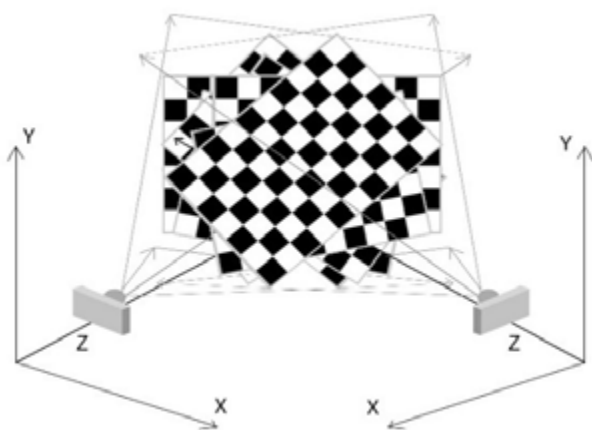


Figura 2 Calibración de cámaras usando un tablero de ajedrez (Krutikova, Sisojevs, & Kovalovs, 2017)

Después de este proceso, se obtienen los parámetros intrínsecos de las cámaras (distancia focal, centro óptico y punto principal), que sirven para poder hallar las matrices de proyección P_1 y P_2 de la cámara izquierda y derecha, respectivamente.

$$P = K[R|T] = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (1)$$

La matriz K representa a la matriz de los parámetros intrínsecos (Salmerón & Martínez, 2016), mientras que la matriz R y T son las matrices de rotación y traslación, respectivamente.

La cámara izquierda es tomada como referencia en el sistema de coordenadas en el mundo, entonces, su matriz de rotación equivaldrá a la matriz identidad de 3x3 y el vector de traslación T será equivalente a $[0 \ 0 \ 0]^T$. Las matrices de rotación y de traslación de la cámara derecha equivaldrán al movimiento dado con respecto a la cámara izquierda.

Adicionalmente, se obtienen los parámetros de distorsión $\{k_1, k_2, k_3, \dots, x_c, y_c\}$, que ayudan a quitar la distorsión tipo ojo de pez que persiste en las cámaras (Zisserman, 2003).

A partir de las matrices de proyección, se puede calcular la matriz fundamental que relaciona la geometría de ambas vistas, es la matriz que representa la geometría epipolar (Zisserman, 2003). Dicha geometría relaciona un par de vistas y la relación que existe entre dos puntos de distintos planos de imagen que son resultados de la proyección de un punto 3D, que se denomina restricción epipolar. (Gonzales, Cancelas, Álvarez, & Enguita, 2016)

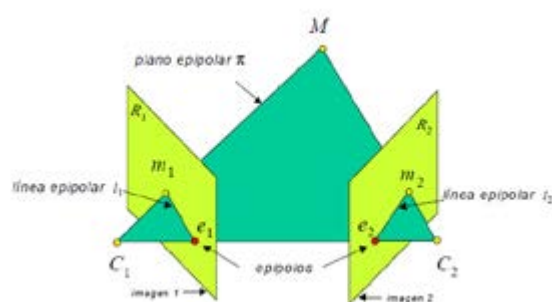


Figura 3 Representación de la geometría epipolar (Mery, 2004)

2) Rectificación de imágenes: Lo que se busca con los algoritmos de rectificación es aplicar una transformación a un par de imágenes, de tal manera que, las líneas epipolares queden paralelas al eje horizontal y además sean colineales. Al tener esto, se simplifica operaciones de búsqueda de puntos correspondientes. (Gonzales, Cancelas, Álvarez, & Enguita, 2016)

Lo que se busca hacer en este caso general es transformar las imágenes, de tal manera que estas sean coplanares (Klette, 2014). Para poder tener las imágenes coplanares se hace uso del algoritmo de Bouguet que ayuda a minimizar los cambios que se producen al re-proyectar las imágenes de ambas cámaras. (Kaehler & Bradski, 2016)

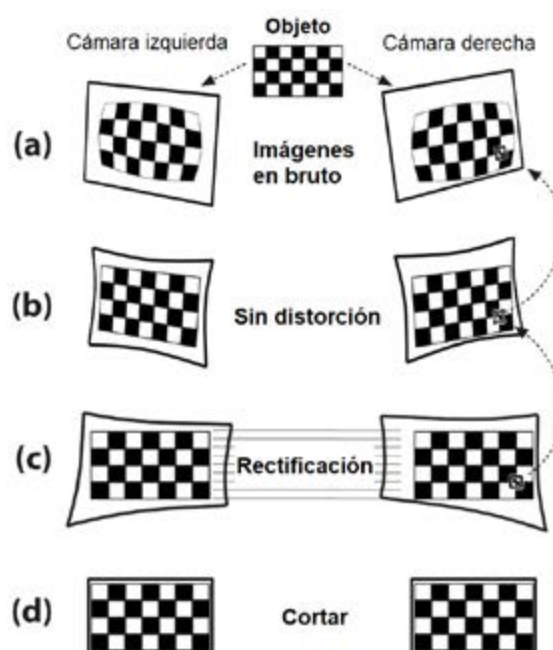


Figura 4 Proceso de rectificación de imágenes (Ontuña, Marco Teórico, 2019)

Matriz de rectificación:

$$R_{rect} = \begin{bmatrix} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \end{bmatrix} \quad (2)$$

donde:

$$\begin{aligned} \vec{e}_1 &= \frac{\vec{T}}{\|\vec{T}\|} \\ \vec{e}_2 &= \frac{1}{\sqrt{T_x^2 + T_y^2}} (-T_y, T_x, 0)^T \\ \vec{e}_3 &= \vec{e}_1 \times \vec{e}_2 \end{aligned} \quad (3)$$

A continuación, se describe gráficamente el proceso de rectificación:

3) Procesamiento de imágenes:

Como se hace uso de la iluminación estructurada entonces lo que se busca obtener de la imagen es únicamente la línea láser, por lo tanto, se realiza una búsqueda de los pixeles más luminosos de la imagen.

Para tomar las fotos de la escena, se necesita que haya una iluminación que vaya de los 0 lx a los 3 lx.

Algoritmo de segmentación de línea láser:

Objetivo: Hallar los pixeles que corresponden a la línea láser.

Solución:

1. Aplicar un desenfoque Gaussiano a la imagen en gris.
2. Binarizar la imagen aplicando un umbral de intensidad de 70.
3. Aplicar operaciones de morfología para obtener la línea láser por completo.



Figura 5 Segmentación de la línea láser (Ontuña, Pruebas y Resultados, 2019)

En cada fila va a existir más de un píxel que pertenezca a la línea láser y lo que se requiere es solo un píxel de cada fila, entonces, la solución es hallar el píxel central de cada fila (Ontuña, Metodología, 2019).

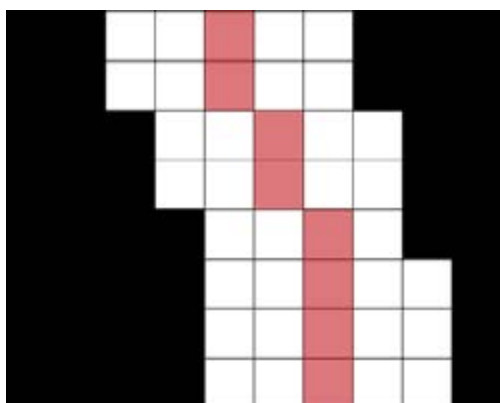


Figura 6 Los pixeles rosados representan al píxel central de una fila (Ontuña, Metodología, 2019)

4) Correspondencia de puntos: Sabiendo que m_1 y m_2 son pixeles pertenecientes a la luz láser, el píxel m_1 con coordenadas $[x_1, y_1]$ de la imagen izquierda será correspondiente con el píxel m_2 con coordenadas $[x_2, y_2]$ de la imagen derecha, si y solo si, $y_1 = y_2$. Esto pasa debido a que las imágenes son colineales gracias a la rectificación aplicada.

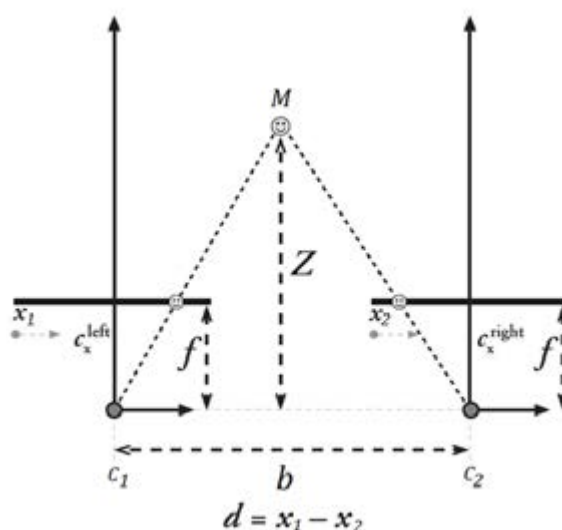


Figura 7 Correspondencia de puntos en imágenes rectificadas (Kaehler & Bradski, 2016)



Figura 8 Puntos correspondientes en ambas imágenes (Ontuña, Pruebas y Resultados, 2019)

5) Reconstrucción 3D: Al tener los puntos correspondientes de las imágenes se procede a realizar la reconstrucción 3D con estos puntos y la matriz de proyección, usando el método de triangulación que se simplifica en una sola ecuación.

$$Z = \frac{fb}{x_1 - x_2} \quad (4)$$

Donde Z es la profundidad, f es la distancia focal, y b es la distancia entre ambas cámaras. La ecuación nace a partir de la semejanza de triángulos (Kaehler & Bradski, 2016).

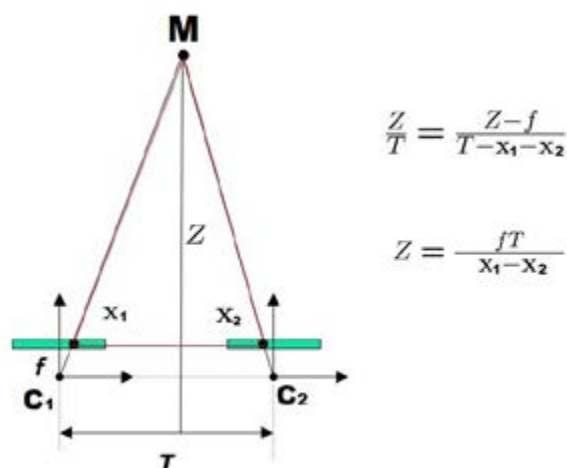


Figura 9 Método de la Triangulación

Resultados

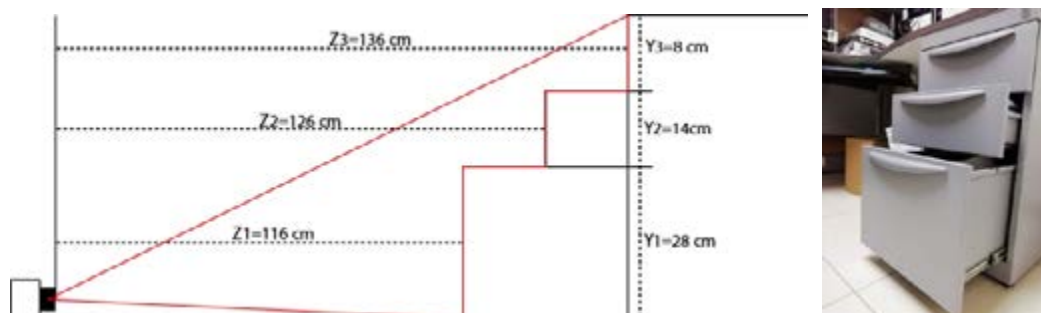


Figura 10 Ubicación del sistema frente al objeto (Ontuña, Pruebas y Resultados, 2019)

Se hizo una prueba a diferentes distancias para poder determinar el error presente en el sistema. Al observar estos resultados se puede notar que

mientras haya más separación entre el sistema y el objeto, el porcentaje de error se eleva más.

Tabla 1 Datos reales contra datos obtenidos de la profundidad (Ontuña, Pruebas y Resultados, 2019)

| | medida real | número de datos obtenidos | media de datos obtenidos (Profundidad) | desviación estándar | porcentaje de error |
|--------------------------------------|-------------|---------------------------|--|---------------------|---------------------|
| Z_1 | 116 cm | 129 | 111.792 cm | 0.709 | 3.627% |
| Z_2 | 126 cm | 57 | 118.642 cm | 0.958 | 5.839% |
| Z_3 | 136 cm | 33 | 126.567 cm | 0.751 | 6.935% |
| Media del Porcentaje de error | | | | | 5.467% |

Con respecto a la altura se realizó algo similar:

Tabla 2 Datos reales contra datos obtenidos de la altura (Ontuña, Pruebas y Resultados, 2019)

| | medida real | dato obtenido (Altura) | porcentaje de error |
|--------------------------------------|-------------|------------------------|---------------------|
| Y_1 | 28 cm | 28.385 cm | 1.375% |
| Y_2 | 14 cm | 13.237 cm | 5.444% |
| Y_3 | 8 cm | 7.694 cm | 3.828 % |
| media del porcentaje de error | | | 3.549 % |

Escaneo con porcentaje de error de 3.730%:



Figura 11 Resultados de escaneo

Conclusiones

Gracias a la implementación de la iluminación estructurada se pudo agilizar la reconstrucción 3D, puesto que se ahorran procesos que son necesarios para hallar puntos correspondientes, pero no fue solamente gracias a eso, sino también al uso de algoritmos de rectificación.

El sistema puede ser implementado en un robot móvil o en un dron con el objeto de poder escanear por completo un interior.

Uno de los problemas que siempre existirán en sistemas con iluminación estructurada es la oclusión, pues la luz láser no se verá en zonas tapadas por otros objetos y no puede cubrir por completo los puntos ciegos.

Referencias

- [1] Ebrahim, M. (2014). 3D LASER SCANNERS: HISTORY, APPLICATIONS, AND FUTURE. *ResearchGate*. doi:10.13140/2.1.3331.3284
- [2] Gonzales, R. C., Cancelas, J. A., Álvarez, I., & Enguita, J. M. (2016). VISIÓN 3D: ESTEREOSCOPIA. En E. Alegre, G. Pajares, & A. d. Escalera, *Conceptos y Métodos en Visión por Computador* (págs. 247-264). Grupo de Visión del CEA.
- [3] Kaehler, A., & Bradski, G. (2016). *Learning OpenCV 3 computer vision in C++ with the OpenCV library*. O'Reilly Media, Inc.
- [4] Klette, R. (2014). *Concise Computer Vision*. London: Springer.
- [5] Krutikova, O., Sisojevs, A., & Kovalovs, M. (2017). Creation of a Depth Map from Stereo Images of Faces for 3D Model Reconstruction. *Science Direct*, 452-459. doi:10.1016/j.procs.2017.01.159
- [6] Lv, Z., & Zhang, Z. (2012). Build 3D Scanner System based on Binocular Stereo Vision. *JOURNAL OF COMPUTERS*, 399-404.
- [7] Mery, D. (2004). *Visión por Computador*. Santiago de Chile: Universidad Católica de Chile.
- [8] Ontuña, J. (2019). Marco Teórico. En J. Ontuña, *Desarrollo de un prototipo de escaneo móvil montado en un Quadcopter empleando iluminación estructurada que permita la reconstrucción 3d de interiores usando una geometría bifocal* (págs. 8-49). Quito: Universidad Central del Ecuador.
- [9] Ontuña, J. (2019). Metodología. En J. Ontuña, *Desarrollo de un prototipo de escaneo móvil montado en un Quadcopter empleando iluminación estructurada que permita la reconstrucción 3d de interiores usando una geometría bifocal* (págs. 50-73). Quito: Universidad Central del Ecuador.
- [10] Ontuña, J. (2019). Pruebas y Resultados. En J. Ontuña, *Desarrollo de un prototipo de escaneo móvil montado en un Quadcopter empleando iluminación estructurada que permita la reconstrucción 3d de interiores usando una geometría bifocal* (págs. 74-93). Quito: Universidad Central del Ecuador.
- [11] Salmerón, A. J., & Martínez, E. I. (2016). VISIÓN 3D: MODELO DE CÁMARA. En E. Alegre, G. Pajares, & A. d. Escalera, *Conceptos y métodos en visión por computador* (págs. 229 - 246). España: Grupo de Visión del CEA.
- [12] Zisserman, R. H. (2003). *Multiple View Geometry in Computer Vision*. Cambridge: Cambridge University Press.

Recibido: 20 de febrero de 2022

Aprobado: 5 de mayo de 2022

