

## IMPLEMENTACIÓN DE FOTOGRAMETRÍA DIGITAL APLICADA A IMÁGENES AÉREAS

Frete, Hector; [hectorfretesac@gmail.com](mailto:hectorfretesac@gmail.com); Gómez, Marcos;  
[gomezredondomarcos@gmail.com](mailto:gomezredondomarcos@gmail.com)

Universidad Nacional de Asunción.

Facultad de Ingeniería.

Laboratorio de Sistemas de Potencia y Control.

La fotogrametría digital aplicada a imágenes aéreas ha ganado gran interés en los últimos años debido a la diversidad de aplicaciones que ofrece, además de los abaratamientos de costos en la fabricación de cámaras fotográficas de alta resolución y vehículos aéreos no tripulados. En este contexto, el propósito de este trabajo es la implementación de un algoritmo de reconstrucción automática de escenarios tridimensionales a partir de un par de imágenes aéreas que cumpla dos criterios principales. En primer lugar, que el algoritmo sea de bajo costo computacional. En segundo lugar, que permita realizar mediciones de segmentos delimitados por puntos del escenario fácilmente identificables. El algoritmo recibe como entrada un par de imágenes aéreas y tiene como salida un modelo tridimensional del escenario en forma de nube de puntos. Luego es puesto a prueba con pares de fotografías capturadas por un vehículo aéreo no tripulado, utilizando para la evaluación de los resultados comparaciones de longitudes de segmentos medidas primeramente en el escenario fotografiado y posteriormente en la nube de puntos obtenida. En pruebas realizadas con fotografías capturadas a hasta 20m de altura se obtienen errores en el orden de centímetros (hasta 4 cm) con tiempos de procesamiento notablemente cortos.

Palabras claves: Fotogrametría digital, Reconstrucción 3-D, Nubes de puntos

## Introducción

En las últimas décadas los vehículos aéreos no tripulados o drones han ganado bastante espacio en diversas aplicaciones debido a su disponibilidad comercial y precios relativamente bajos. Entre las áreas de aplicación más importantes se pueden citar la agricultura de precisión, vigilancia, transporte de medicinas en lugares remotos o de difícil acceso, entre otras Haidari et al. (2016); Tariq, Gillani, Qureshi y Haneef (2017); Zhang y Kovacs (2012). Una de las principales aplicaciones es la toma de fotografías aéreas, dado que los drones permiten un barrido rápido, preciso y económico de una superficie en comparación con métodos terrestres, satelitales y aéreos tripulados como refleja el estudio de Tariq, Osama y Gillani (2016). Estas imágenes pueden tener usos de gran interés y alto impacto, mediante la aplicación de la fotogrametría digital como afirma Yang, Lin y Liu (2016).

Resulta de particular interés la obtención de modelos tridimensionales debido a la gran cantidad de información que puede obtenerse de estos modelos y su diversidad de aplicaciones, como cálculo de curvas de nivel, medición de áreas y volúmenes, entre otros. La creciente capacidad de procesamiento de los computadores personales, la constante reducción de sus costos de producción y las ventajas propias del procesamiento digital han despertado un gran interés por lograr la reconstrucción automática de escenarios.

Por lo anteriormente mencionado, el presente trabajo se enfoca en la implementación de un algoritmo para la reconstrucción de escenarios mediante la obtención de modelos tridimensionales en forma de nubes de puntos a partir de imágenes aéreas capturadas con una cámara previamente calibrada. Para ello se utilizan técnicas de visión artificial en un proceso de reconstrucción altamente automatizado. Dicho proceso debe contemplar tres etapas fundamentales. La primera etapa consiste en la detección, descripción y pareo de puntos de interés, como el propuesto por Leutenegger, Chli y Siegwart (2011). La segunda, en el cálculo de la geometría del escenario, como en el algoritmo de los cinco puntos de Nistér (2004) para la obtención de la matriz esencial, que posibilita el cálculo de la posición relativa entre cámaras. Por último, la triangulación de los puntos, etapa en la cual son calculadas las coordenadas tridimensionales de los puntos identificados en el escenario. Para esta última etapa Hartley y Zisserman (2004) detallan en su libro los distintos métodos disponibles. Además de estas etapas fundamentales, se propone una etapa opcional de densificación de la nube de puntos, en la cual se agrega detalle a la nube obtenida, facilitando las mediciones.

## Objetivos

### Objetivo General

- Implementar un algoritmo de reconstrucción tridimensional a partir de imágenes aéreas, de bajo costo computacional, que permita realizar mediciones del escenario reconstruido.

### Objetivos Específicos

- Implementar algoritmos de acondicionamiento y procesamiento de imágenes.
- Realizar el pareo de puntos homólogos entre distintas imágenes de los escenarios fotografiados.
- Obtener modelos tridimensionales, en forma de nubes de puntos, de los escenarios fotografiados.
- Evaluar los resultados obtenidos.

### Materiales y Métodos

El algoritmo recibe como entrada un par de imágenes aéreas y tiene como salida una nube de puntos que representa la distribución del escenario fotografiado en el espacio. Ambas imágenes de entrada deben ser capturadas con la misma cámara. Si desea utilizarse la etapa opcional de densificación, las imágenes deben ser capturadas de modo que las posiciones de cámara se encuentren a la misma altura respecto a un nivel de referencia previamente determinado y entre ellas exista una rotación máxima de  $6^\circ$  para su correcto funcionamiento.

En primer lugar, se calibra la cámara mediante el método propuesto por Zhang (2000). Esta calibración, además de los parámetros internos de la cámara necesarios para la reconstrucción, provee coeficientes de distorsión que permiten modelar la distorsión ocasionada por la lente de la cámara. Mediante dichos coeficientes se remueve la distorsión de las imágenes antes de iniciar el proceso de reconstrucción.

Con la distorsión eliminada, se procede a realizar la detección y descripción de puntos de interés en las imágenes, esto es, puntos fácilmente detectables en las imágenes debido a las variaciones de color en sus entornos. Esto se realiza por medio del método BRISK, propuesto por Leutenegger et al. (2011). Una vez obtenidos los descriptores de los puntos de interés, para cada punto obtenido en la primera imagen se busca el punto de la segunda imagen que represente al mismo punto en el espacio tridimensional. Esto se logra comparando los descriptores obtenidos para la primera imagen con los obtenidos para la

segunda imagen y aplicando posteriormente el criterio de Lowe (2004). A los pares de puntos obtenidos de esta manera se denominan puntos homólogos.

Con los puntos homólogos es posible calcular la matriz esencial, matriz que encapsula la información de las posiciones de la cámara al capturar las imágenes y la restricción epipolar. Esta restricción puede plantearse matemáticamente de la siguiente forma:

$$\tilde{x}E\tilde{x}' = 0$$

**E:** matriz esencial.

$\tilde{x}$ : coordenadas normalizadas del punto en la primera imagen.

$\tilde{x}'$ : coordenadas normalizadas del punto en la segunda imagen.

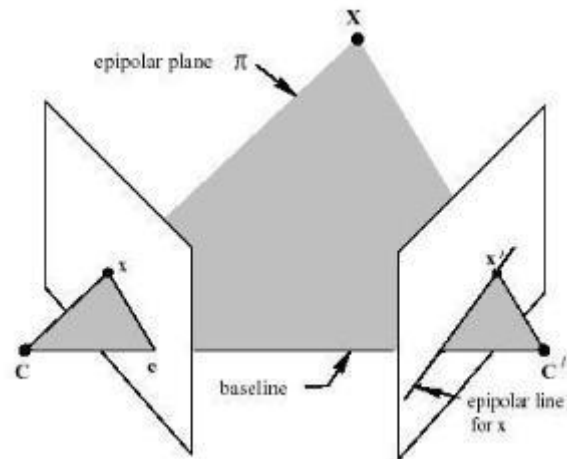


Figura 1: Restricción epipolar. Fuente: Hartley y Zisserman (2004)

Geoméricamente, la restricción epipolar implica que, para cualquier punto de la primera imagen, su correspondiente punto homólogo se encuentra sobre una recta de la segunda imagen denominada recta epipolar (ver Figura 1).

El cálculo de esta matriz se realiza mediante el algoritmo de los cinco puntos, descrito en Nistér (2004). Además, se eliminan los pares de puntos que no satisfacen la restricción epipolar para la matriz esencial calculada. Posteriormente, se establece la primera posición cámara como el origen de coordenadas del sistema y se calcula a partir de la matriz esencial la rotación entre las posiciones de cámara y un vector unitario que representa la traslación entre dichas posiciones. Esto puede expresarse matemáticamente de la forma:

$$P = K[I|0] \quad P = K[R|t], \text{ donde}$$

**P:** Matriz de cámara en la posición 1.

**P':** Matriz de cámara en la posición 2.

**K:** Parámetros internos de la cámara.

**I:** Matriz identidad.

**R:** Matriz de Rotación.

**t:** Vector de traslación.

Debido a que solo es posible extraer la dirección de la traslación, la reconstrucción sólo puede obtenerse a un factor de escala, por lo que es necesario conocer al menos una dimensión del escenario para obtener las medidas reales en el modelo tridimensional obtenido.

En la siguiente etapa se realiza la triangulación, obteniendo así las coordenadas tridimensionales de los puntos que se ajustaron al modelo. Para ello se utilizan las matrices de proyección calculadas en la etapa anterior según las siguientes ecuaciones:

$$x = PX \qquad x' = P'X$$

Siendo  $X$  las coordenadas tridimensionales del punto.

La etapa de densificación consiste en agregar puntos al modelo utilizando como partida los puntos que se adaptan a la matriz esencial. Para ello se realiza la búsqueda de nuevos pares de puntos en los vecindarios de los puntos anteriormente obtenidos de manera similar a la propuesta por Lhuillier y Quan (2012).

Las coordenadas tridimensionales así obtenidas se exportan a un archivo con extensión PLY propuesto por Turk (1994).

Para la implementación del algoritmo se utilizó el lenguaje de programación C++ y la librería OpenCV 3.2 en el entorno de desarrollo Visual Studio 2017.

Para la evaluación del algoritmo se tomaron fotografías aéreas de un escenario con el dron Autel X-Star Premium (Figura 2), realizando mediciones en dicho escenario para contrastarlas con



*Figura 2: Drone Autel X-Star Premium utilizado para las pruebas experimentales. Fuente: Autel Robotics*

las mediciones realizadas en los modelos tridimensionales obtenidos. Fueron tomados dos grupos de fotografías, uno a 15 metros de altura y otro a 20 metros de altura. Estas alturas fueron elegidas considerando las limitaciones de la resolución de la cámara (12 megapíxeles). El criterio de evaluación utilizado fue el de conservación de la norma euclidiana, esto es, las longitudes de distintos segmentos en el escenario.

El procesamiento de las imágenes obtenidas fue realizado utilizando un núcleo del procesador Intel(R) Core™ i7-6700HQ. Las mediciones en los modelos tridimensionales fueron realizadas mediante el programa CloudCompare.

## Resultados y discusión

En la Figura 3 se observa un par de imágenes capturado a 15m de altura, en tanto que en la Figura 4 se tiene un par de imágenes a 20m de altura. Estas imágenes fueron capturadas



Figura 3: Par de imágenes a 15m de altura. Fuente: Elaboración propia.

atendiendo las restricciones requeridas por la etapa de densificación, mediante la cual es posible obtener modelos tridimensionales con aproximadamente 38000 puntos, suficientes para realizar las mediciones en las nubes de puntos obtenidas.



Figura 4: Par de imágenes a 20m de altura. Fuente: Elaboración propia.



Figura 5: Vista superior de la nube obtenida con las imágenes a 15m de altura. Fuente: Elaboración propia.



Figura 6: Vista frontal de la nube obtenida con imágenes a 15m de altura. Fuente: Elaboración propia.

En la Figura 5 se observa la vista superior de la nube de puntos obtenida con el par de imágenes a 15m de altura, en tanto que en la Figura 6 se observa una vista frontal de la misma nube.

La Figura 7 presenta una vista lateral del modelo tridimensional obtenido para el mismo par de imágenes, mientras que en la Figura 8 se observan las superficies generadas a partir de la nube de puntos utilizando el algoritmo de Poisson que provee el programa CloudCompare.

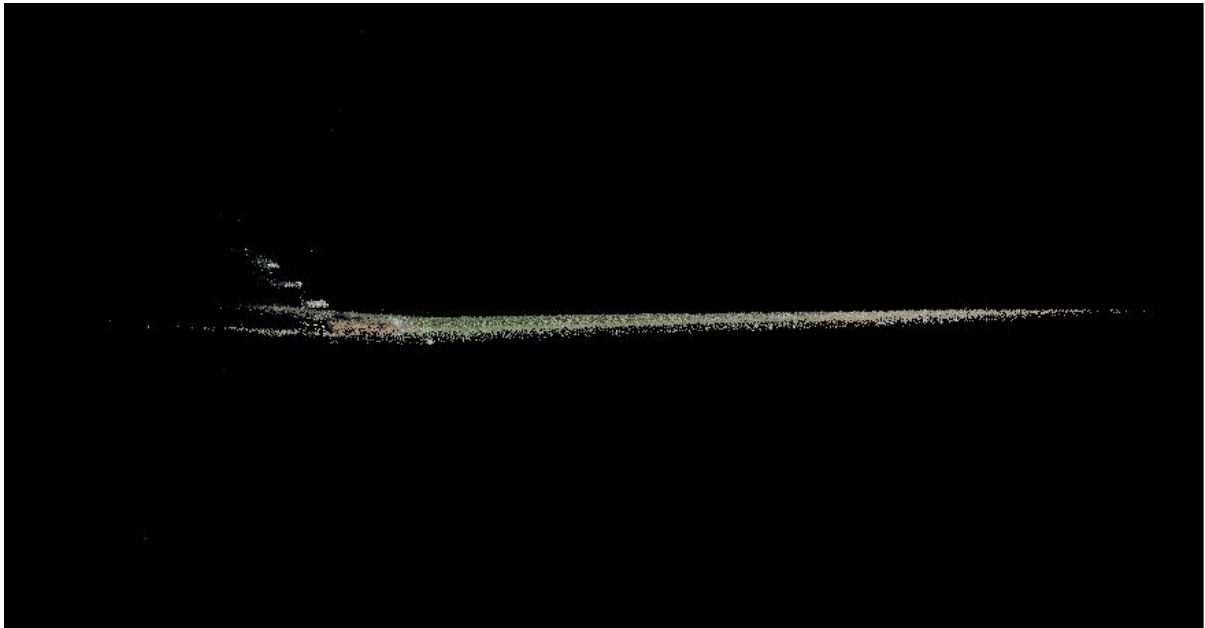


Figura 7: Vista lateral de la nube obtenida con imágenes a 15m de altura. Fuente: Elaboración propia

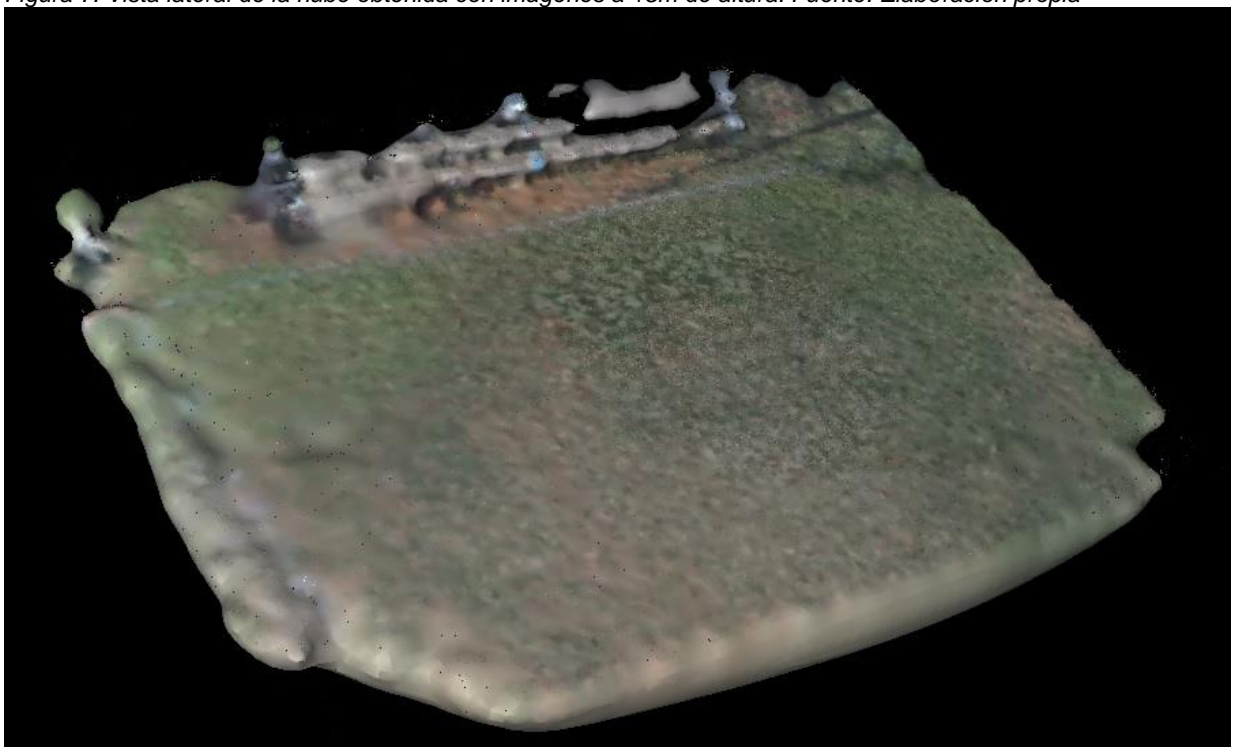


Figura 8: Vista isométrica de la superficie generada por el algoritmo de Poisson del programa CloudCompare. Fuente: Elaboración propia.



En la Figura 9 se aprecian las mediciones realizadas en el escenario y la nomenclatura utilizada para diferenciar los segmentos. Los segmentos con la denominación “v” consisten exclusivamente en alturas entre los niveles de las gradas de la gradería presente en el

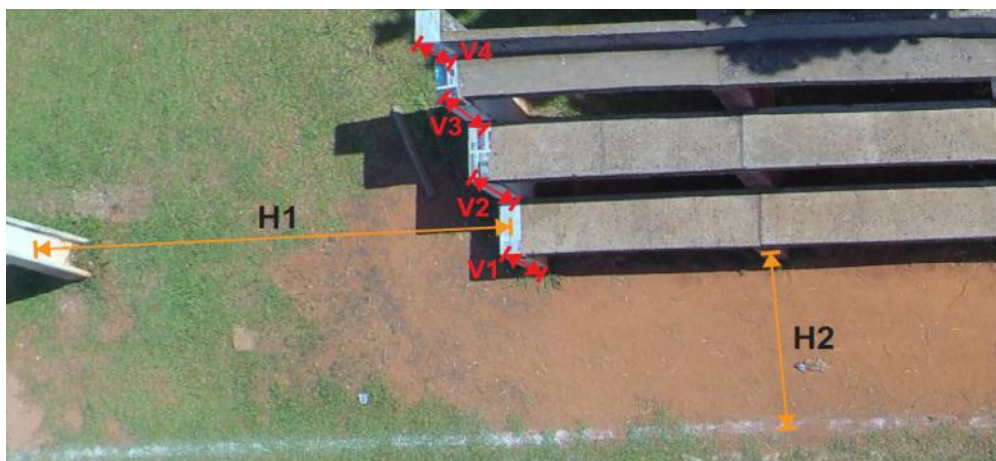


Figura 9: Segmentos utilizados para la evaluación de los resultados. Fuente: Elaboración propia.

escenario, en tanto que los segmentos con denominación “h” no poseen variaciones de altura en su longitud.

Los resultados de estas mediciones para 15m de altura se recogen en la Tabla 1.

Segmento	Valor medido [cm]	Valor real [cm]	Error [cm]	Error porcentual [%]
H1	487,56	487	0,56	0,11
H2	184,86	180	4,86	2,7
V1	65,16	66	0,84	1,27
V2	57,59	56	1,59	2,84
V3	53,64	56	-2,36	4,21
V4	29,41	30	-0,59	1,97

Tabla 1: Resultados obtenidos con las imágenes a 15m de altura. Fuente: Elaboración propia.

En la Tabla 2 se recogen los resultados para las imágenes a 20m de altura.

Segmento	Valor medido [cm]	Valor real [cm]	Error [cm]	Error porcentual [%]
H1	482,59	487	-4,41	0,90
H2	184,2	180	4,2	2,33
V1	62,31	66	-3,69	5,59
V2	57,19	56	1,19	2,12
V3	53,3	56	-2,7	4,82
V4	30,2	30	-0,2	0,66

Tabla 2: Resultados obtenidos con las imágenes a 20m de altura. Fuente: Elaboración propia.

Se observan mayores errores en los segmentos H1 y H2, esto se debe principalmente a la dificultad para identificar los puntos que determinan dichos segmentos en las nubes de puntos. Los errores en las demás mediciones pueden atribuirse al método de triangulación utilizado, debido a que éste es muy sensible al ruido introducido por el sensor de la cámara. Aún así, la exactitud de los resultados se encuentra en un margen aceptable considerando el corto tiempo de procesamiento necesario para la obtención de los resultados. Este tiempo depende principalmente de la densificación, debido a que esta etapa posee orden  $O(n) = n \log(n)$ , donde  $n$  es el número de puntos deseados en la salida. Para los pares de imágenes de prueba y el procesador utilizado se obtuvo un tiempo promedio de 16 segundos sin la etapa de densificación (alrededor de 6000 puntos) y un tiempo promedio de 1 minuto y 16 segundos con la etapa de densificación (alrededor de 30000 puntos).

## Conclusiones

Se logró implementar un algoritmo de reconstrucción automática de escenarios a partir de un par de imágenes aéreas, cumpliendo los requerimientos de bajo costo computacional y la posibilidad de realizar mediciones en los modelos tridimensionales obtenidos. Los tiempos de procesamiento se encuentran muy por debajo de los requeridos por otros algoritmos, con una exactitud aceptable.

La resolución de la cámara resulta determinante para la calidad de los modelos tridimensionales obtenidos, debido a que limita el nivel de detalle que puede obtenerse en la reconstrucción. Así también es importante la calidad del sensor de la cámara, debido a que puede introducir ruido que repercute en la exactitud de la nube obtenida.

Cabe destacar que el algoritmo constituye una base para el desarrollo de métodos más complejos y actualmente se está modificando el algoritmo para la utilización tres o más

imágenes en el proceso de reconstrucción y la incorporación de puntos de control en los escenarios a reconstruir, buscando mejorar la exactitud y precisión de los modelos tridimensionales obtenidos.

## **Bibliografía**

Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools. Obtenido de <https://github.com/opencv/opencv>

CloudCompare (versión 2.9) [GPL software]. (2018). Obtenido de <http://www.cloudcompare.org/>

Haidari, L. A., Brown, S. T., Ferguson, M., Bancroft, E., Spiker, M., Wilcox, A., Ambikapathi, R., Sampath, V., Connor, D. L., y Lee, B. Y. (2016). The economic and operational value of using drones to transport vaccines. *Vaccine*, 34(34).

Hartley, R. I. y Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*. Cambridge University Press, segunda edición.

Imagen del drone X-Star Premium. (s.f.). Recuperado de <https://www.autelrobotics.com/>.

Leutenegger, S., Chli, M., y Siegwart, R. Y. (2011). Brisk: Binary robust invariant scalable keypoints. En: 2011 International Conference on Computer Vision, Nov.

Lhuillier, M. y Quan, L. (2000). Robust Dense Matching Using Local and Global Geometric Constraints. En: 15th International Conference on Pattern Recognition (ICPR '00), Septiembre. Vol. 1.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60.

Nistér, D. (2004). An efficient solution to the five-point relative pose problem. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(6).

Tariq, A., Osama, S. M., y Gillani, A. (2016). Development of a low cost and light weight uav for photogrammetry and precision land mapping using aerial imagery. En: 2016 International Conference on Frontiers of Information Technology (FIT), Dec.

Tariq, A., Gillani, S. M. O. A., Qureshi, H. K., y Haneef, I. (2017). Heritage preservation using aerial imagery from light weight low cost unmanned aerial vehicle (uav). En: 2017 International Conference on Communication Technologies (ComTech), April.

Turk, G. (1994). The PLY Polygon File Format. Recuperado de <https://web.archive.org/web/20161204152348/http://www.dcs.ed.ac.uk/teaching/cs4/www/graphics/Web/ply.html>

Yang, Y., Lin, Z., y Liu, F. (2016). Stable imaging and accuracy issues of low-altitude unmanned aerial vehicle photogrammetry systems. *Remote Sensing*, 8(4).

Zhang, C. y Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agriculture*, 13(6).

Zhang, Z. (2000) A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(11):1330–1334.