3D MODEL RECONSTRUCTION FROM A SINGLE IMAGE WITH IMU DATA

*Tzu-Fei Chen [a], Huan Chang [b] and Fuan Tsai [*c]*

[a,b] *Graduate student, Department of Civil Engineering, National Central University,*
300,Jhongda Rd., Jhongli, Taoyuan 32001, Taiwan; Tel: +886-3-4227151#57614.
[a] *E-mail:vivy7385@hotmail.com*
[b] *E-mail:1984chang@gmail.com*

[c]*Associate Professor, Center for Space and Remote Sensing Research, National Central University,*
300,Jhongda Rd., Jhongli, Taoyuan 32001, Taiwan; Tel: +886-3-4227151#57619
E-mail: ftsai@csrsr.ncu.edu.tw

KEY WORDS: Inertial measurement unit (IMU), Single image reconstruction, Depth map, accuracy assessment

ABSTRACT: This research developed a method for reconstructing 3D scenes from a single photograph with depth map and inertial measurement unit (IMU). The proposed method uses IMU to provide the relative attitudes of Exterior Orientation Parameters (EOP) and assumes the camera position is the origin. Depth map provides the relative distance from feature points to the camera center. Therefore, collinearity condition equations can be constructed to determine 3D positions of feature points from a single image. In this study, there are four steps for 3D scene reconstruction from a single image: (1) camera and boresight calibration; (2) depth map generation; (3) feature points calculation; (4) 3D scene model reconstruction. The proposed method begins with angles calibration between camera boresight and IMU system. Broesight angles can be predicted using the result of image back projection and simultaneously recorded orientation data from the IMU. Using the relative distance provided from the depth map, coordinates of the extracted feature points are determined from collinearity condition equations. Finally, the feature points are used to reconstruct 3D models. Textures are then applied to the constructed models for visualization. In order to evaluate the quality of reconstructed model, this research simulates the effects caused by different errors to estimate the accuracy of the models.

1. INTRODUCTION

Photogrammetry has many advantages as a technique for the acquisition of three-dimensional models for virtual reality. But the traditional photogrammetric process to extract 3D geometry from multiple images is often considered too labour-intensive. On the other hand, single view reconstruction (SVR) techniques have a potential to extract 3D geometry from an image should provide a more economic and efficient way for building model reconstruction. Many researchers have dealt with the problem of 3D reconstruction from a single image. The majority of the researchers have managed to solve this problem using a calibrated camera (Delage et al., 2006) or a partially calibrated camera (a camera that is set to a known height (Lee et al., 2009). However, the internal and external parameters of the camera play the vital role in the reconstruction. In this paper gain the exterior orientation parameters with extra equipment IMU data. According to El-Sheimy (1996), the transformations between camera and IMU have to be taken into account before catch the image.

It's important for photogrammetry about ground coordinate, image coordinate and exterior orientation. The difference set will get to result for different purpose. This paper's purpose is getting to exterior orientation with extra equipment IMU data and reconstructs the model from depth map of single image. According to Wolf & Dewitt (2000) and He Weixin (2002), the collinearity condition is the condition using known exterior orientation parameters that any object point, and its photo image all lie along a straight line in three- dimensional space. In order to solve the coordinate of the object, in field of the computer vision, depth map from the single image could provide us an extra condition we need. According to Ashutos et al. (2008) proposed a method to reconstruct 3-D depth from a single image by applying supervised learning to predict the value of the depth map as a function of the image. However, in this research would like to provide a novel method to reconstruct the scene from a single photograph with depth map and inertial measurement unit (IMU).

2. The proposed method

The proposed method comprises four major parts for 3D scene reconstruction from a single image. The first part is to calibrate the IOP of camera and to establish transformation between sensor frames. The second one is to generate the depth map from a single image in this part. The third is to calculate feature points. The final part is to reconstruct 3D scene model. Fig. 1 illustrates the work flow of the methodology model.

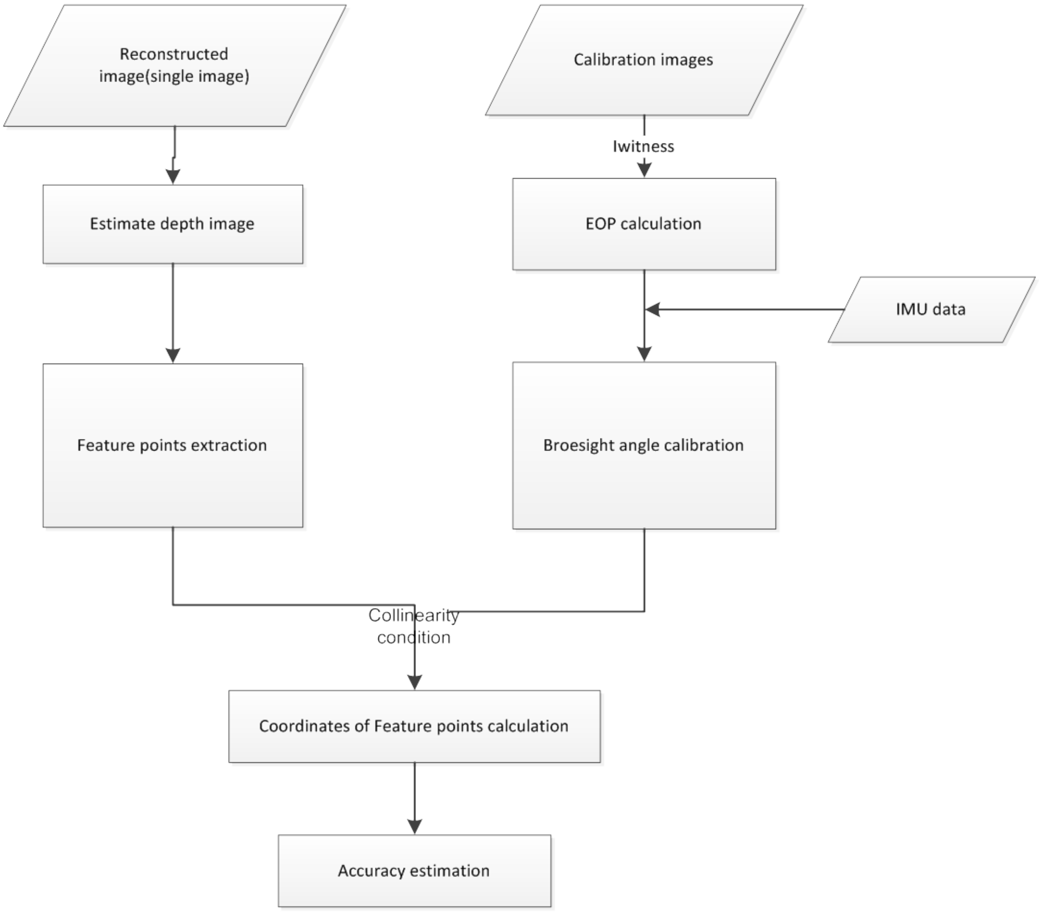

Fig.1. Work flow of methodology model

2.1 camera and boresight calibration

The interior orientation is required for the reconstruction of the bundle of rays. As the result, the first step is camera calibration. The focal length (**f**) of the camera in this case is determined by calibration before catching the photograph, and this study ignored the lens distortion.

After camera calibrating, an extra equipment, IMU, to attain three elements of exterior orientation parameters (**ω, φ, κ**) of a photograph. Boresight angle calibration between different sensor frames has to be taken into account. Broesight angles can be predicted using the result of image back projection and simultaneously recorded orientation data from the IMU. In addition to transformations between sensors, rotations between different sensor frames are also considered. The INS b-frame (gyro frame) cannot be aligned with the c-frame. According to El-Sheimy (1996), the constant rotation $\mathbf{R}_{\mathbf{c}}^{\mathbf{b}}$ between the two frames is again obtained by calibration. In this case, $\mathbf{R}_{\mathbf{c}}^{\mathbf{m}}(\mathrm{t})$ can be written as:

$$\mathbf{R}_{\mathbf{c}}^{\mathbf{m}}(\mathbf{t}) = \mathbf{R}_{\mathbf{b}}^{\mathbf{m}}(\mathbf{t}) \cdot \mathbf{R}_{\mathbf{c}}^{\mathbf{b}} \qquad \textbf{[1]}$$

Where $\mathbf{R}_{\mathbf{c}}^{\mathbf{b}}$ is the rotation between the camera c-frame and the INS b-frame as determined from a calibration process.

2.2 depth map generation

Depth map provides the relative distance from feature points to the camera center. To obtain a reliable depth map use a single view of the input image taking into account. It requires a rough depth classified by Ashutos et al. (2008). An approximated depth map can be generated by applying supervised learning to predict the value of the depth. Depth map from the single image could provide us an extra condition we need. The distance from object to camera center are defined as depth **(d)**, and the formulas such as [3] shows.

$$d = \sqrt{(X - X^c)^2 + (Y - Y^c)^2 + (Z - Z^c)^2} \quad [2]$$

2.3 3D model reconstruction

After feature extracting, the coordinates of the extracted feature points are determined from collinearity condition equations and with depth information from depth map. The collinearity condition is using known exterior orientation parameters provide from IMU data that any object point, and its photo image all lie along a straight line in three- dimensional space. And the formulas such as [1] shows.

$$x = -f\frac{m_{11}(X - Xc) + m_{12}(Y - Yc) + m_{13}(Z - Zc)}{m_{31}(X - Xc) + m_{32}(Y - Yc) + m_{33}(Z - Zc)}$$

$$y = -f\frac{m_{21}(X - Xc) + m_{22}(Y - Yc) + m_{23}(Z - Zc)}{m_{31}(X - Xc) + m_{32}(Y - Yc) + m_{33}(Z - Zc)} \text{............[1]}$$

Where

$m_{11} = \cos\varphi \cos\kappa$

$m_{12} = \sin\omega \sin\varphi \cos\kappa + \cos\omega \sin\kappa$

$m_{13} = \cos\omega \sin\varphi \cos\kappa + \sin\omega \sin\kappa$

$m_{21} = -\cos\varphi \sin\kappa$

$m_{22} = -\sin\omega \sin\varphi \sin\kappa + \cos\omega \cos\kappa$

$m_{23} = \cos\omega \sin\varphi \sin\kappa + \sin\omega \cos\kappa$

$m_{31} = \sin\varphi$

$m_{32} = -\sin\omega \cos\varphi$

$m_{33} = \cos\omega \cos\varphi$

2.4 Accuracy of the model estimation

In order to evaluate the quality of reconstructed model from single image, this research compares the geometrical performances between the model form single image and simulated real model. The experimental results are to evaluate the quality of reconstructed model from single image.

3. Experiment result

The test data image size is 537 rows and 1300 columns, and its image pixel size is 4μm*4μm; focal length is

57mm. Depth from camera center is 50m roughly. Because of test area is simulated from created 3D model, The source of check points came from there to estimate the accuracy of the experiments .The experiments compare the geometrical performances between the model form single image and real model. The experimental results are to evaluate the quality of reconstructed model from single image.

3.1 compare accuracy in using real depth and using depth map

In this study, we use the three equations to solve the three unknown object coordinates (X, Y, Z). The serious problem is that there has no degree of freedom in this algorithm. As the result, it is important that the depth map should be gain from high accuracy, or it may have the expected heavy error in precision. Table 1 show the accuracy in using real depth from real model and rough depth from depth map.

Table 1: compare the accuracy in using real depth and using depth map

RMSE(m)	Reference	Proposed method
X	0.065	0.141
Y	0.047	0.193
Z	0.014	2.025

3.2 The accuracy in different depth resolution of depth map

According to above result, the depth map plays the vital part in the accuracy. Because the depth map uses several gray levels to express the distance from feature point to camera center, different level scale will result in accuracy changing. In this section, we discuss the effect in different depth resolution in 18cm/per gray value, 36 cm/per gray value, 56 cm/per gray value, 74 cm/per gray value, 92 cm/per gray value. Table 2 shows the experimental results.

Table 2: compare the accuracy in different depth resolution of depth map

Resolution in depth map (cm/per gray value)	18	36	56	74	92
RMSE (m)	0.086	0.095	0.232	0.262	0.438

4. CONCULSION AND FUTURE WORKS

This study developed algorithms combining the field in computer vision and close range photogrammertry and required only a single image with IMU and depth map to reconstruct 3D building models. Experimental results demonstrated that the accuracy in the reconstruction of 3D building models can be determined by the reliable accuracy in depth map. Future improvement of the developed system will focus on the capability of reconstructing in increasing the degree of freedom.

5. ACKONWLEDGMENTS

This project was supported in part by National Science Council of Taiwan under grant Nos.NSC-98-2221-E-008-097-MY2.

6. References

Criminisi, A., Reid, I. and Zisserman, A., 2000. Single view metrology. International Journal of Computer Vision 40 (2), pp. 123–148.El-Sheimy, N., 1996, The development of VISAT a mobile survey system for GIS applications, Ph.D. Thesis, The University of Calgary, Canada.Saxena, A., Sun, M. and Ng, A.Y., 2008. Make3D: Learning 3D Scene Structure from a Single Still Image. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(5), pp. 824-840.Van Den Heuvel, F. A., 1998. 3D reconstruction from a single image using geometric constraints, ISPRS Journal of Photogrammetry & Remote Sensing, 53, pp. 354–368.Wolf, P. R. and Dewitt, B. A., 2000. Elements of Photogrammetry with Applications in GIS. MC Graw Hill, New York.Wolf, P. R. and Ghilani, C. D., 2006. Adjustment Computation Spatial Data analysis. John Wiley & Sons, New York.