Camera Rotation and Translation Recovery using Optical Flow Based Method for 3D Reconstruction from Un-Calibrated Images

Article Info

Received: 23/4/2017
Accepted: 28/5/ 2017
Published online: 1/6/2017

[1,2] Intan Syaherra Ramli, [2]Rahmita Wirza O.K Rahmat, [2] Ng Seng Beng, [2]Puteri Suhaiza Sulaiman and [2] Hizmawati Madzin

[1]Computer Science Department, Faculty of Computer Science and Mathematics, Universiti Teknologi MARA (UiTM)

[2] Multimedia Department, Faculty of Computer Science & Information Technology, Universiti Putra Malaysia (UPM)

isramli1986@gmail.com

ISSN 2231-8844

ABSTRACT

Camera rotation and translation recovery from un-calibrated images is a very challenging task as it could affect the accuracy of 3D point reconstruction process. Existing method based on Structure from Motion (SfM) using epipolar geometry and fundamental matrix is a popular method to recover camera intrinsic and extrinsic parameter. However, this method requires complex process and high in computational time. Other than that, GIS/GPS information is not stable and could affect the accuracy and reliability. SfM method using optical flow can be used to recover camera rotation and translation. However, existing methods only focus to recover camera rotation and translation on particular angle and the assumption is the object position always perpendicular to the center of camera. Even though optical flow method is less complex but it still encounters the problem with accuracy. Therefore, this paper discusses a preliminary experiment to recover camera rotation and translation using optical flow based method. In the experiment, various camera position and optical flow pattern were observed. Result shows that world camera position can be recovered from optical flow vector. This experiment is a prerequisite step in order to propose an accurate and efficient method in 3D reconstruction from un-calibrated images.

Key words: **3D reconstruction, Camera Rotation & Translation, Optical flow**

1. Introduction

3D reconstruction has been discussed widely in computer graphic research as the 3D content is the main element in virtual and augmented reality application. Moreover, virtual and augmented reality applications have tremendous impact in various fields such as entertainment, tourism, architectural and education. 3D reconstruction involve the process to capture the real object either using active or passive method. Active method requires special hardware such as structured light and laser that involve interference of the real object (Izadi et al., 2011). On the other hand, passive method only uses a normal camera sensor and image understanding in 3D reconstruction process (Kang et al., 2013; Zakaria et al.,2010; Karami et al., 2014). Moreover, 3D reconstruction using software such as 3D studio Max, Blender or Maya requires training and skills.

Passive method using image based sequence is very popular method in 3D reconstruction. This input can be acquired using single camera or multi–cameras. Image acquisitions using multi-cameras is high in cost and tedious since the process require camera calibration in fix position. Therefore, image acquisition using single moving camera has the demand as it low in cost. People can use their normal consumer camera or Smartphone to capture the image sequence. Moreover, this camera can be easily mounted on moving vehicles to capture the large scale of images (Haala et al., 2013; Nex et al.,2013) . However, 3D reconstruction using single moving camera has more challenges due to various image rotation and translation (Nguyen et al., 2013; Pollefeys et al., 2004). This process related to camera rotation and translation recovery process which is one of the important steps in 3D reconstruction using feature based method.

Therefore, this paper presents a preliminary experiment to recover camera rotation and translation using optical flow based method. This experiment is a prerequisite step in order to propose an accurate and efficient optical flow method to recover the camera rotation and translation from un-calibrated images. Next section in this paper will discuss about related work, camera rotation and translation using optical flow method, preliminary experiment result, discussion and future work.

2. Literature Review

Based on literature, there are three main methods for 3D reconstruction from un-calibrated images either using feature, volumetric or hybrid methods. Feature based method involves feature extraction from selected point of interest (Toldo et al. 2015 ; Pollefeys et al. 2004; El et al. 2014; Zhao et al. 2010). Meanwhile, volumetric based methods can be either image from silhouette or space carving (Zhao et al. 2010; Liang and Wong 2010). This approach has limitation on object concavities and high in computation time as it involves voxel representation (Nguyen et al. 2013 and Zhao et al. 2010). On the other hand, hybrid method involves combination of both feature based and volumetric methods (Nguyen et al. 2013).

Therefore, this research focuses on feature based method which is widely used and able to detect feature from complex object. This method involves few steps such as feature extraction and matching, camera rotation and translation recovery, 3D point reconstruction and merging process. Camera rotation and translation recovery is very important in order to relate the image and world coordinates. This information is required to reconstruct the accurate 3D point. In general, Structure from Motion (SfM) based on epipolar geometry and fundamental matrix, optical flow and GIS/GPS are the existing methods that have been used to recover the camera rotation and translation for un-calibrated images.

- **SfM based on Epipolar geometry & Fundamental Matrix**

Existing method based on Structure from Motion (SfM) using epipolar geometry and fundamental matrix is a popular method to recover camera intrinsic and extrinsic parameter between two views. This method only relies on the cameras internal parameters and relative pose. Fundamental matrix is the algebraic representation of epipolar geometry that derives from the mapping between a point and its epipolar line. However, this method has taking longer time to refine camera parameter for each additional view (Pollefeys et al., 2004). Moreover, it requires complex process and high in computational time (Kang et al.2013 ;Pollefeys et al.,2004).

Therefore, there are research that use essential matrix that specialized the fundamental matrix to the case of normalized image (Nguyen et al., 2013; Ji et al. 2004). This matrix has fewer degree of freedom and additional properties compared to fundamental matrix. For example, some research has restricted that the rotation must be about the same axis and equal angles (Ji et al. 2004). In other research, the focal length information was extracted from EXIF tag (Zhao et al. 2010). However, EXIF tag is not always reliable in all images (Pollefeys et al.2004).

- **SfM based on Optical Flow**

Another method to recover camera rotation and translation is SfM method using optical flow. This method provides the relation between corresponding points using gradient based method. It detected the movement of a region by calculating where it moved in the image space. The assumptions were the pixel intensities of an object do not change and the neighboring pixels had similar motion. This method is fast and reliable to recover camera rotation and translation (Kanatani., 1988; Ohta , 201;Ming-Xin et al. 2012).

However, existing methods only focus to recover camera rotation and translation on particular angle and the assumption is the object position always perpendicular to the center of camera (Diskin et al. 2013; Kanatani 1988; Ohta 2015;Ming-Xin et al. 2012). Moreover, there are research that only focused on tracking the global estimation of translation and rotation of camera in z-axis rotation (Ming-Xin et al. 2012). This situation is not always reliable for single moving camera. In addition, even though optical flow method is less complex to recover camera rotation and translation but it still encounters the problem with accuracy (Ohta 2015).

- **GIS/GPS**

Besides, there are research that have utilized Geographical Information System (GIS) or Global Positioning System (GPS) to recover camera position (Tanathong and Lee, 2014; Kume et al., 2010; Guo and Gao, 2013). This method is fast because the information about camera position can be retrieved directly from the sensor. However, this method suffers from unstable connection from satellite that could affect the accuracy (Tanathong and Lee, 2014). Previous research mentioned that actual error of GPS positioning grows to the 10 meter level (Kume et al. 2010). Moreover, direct GPS usage for camera pose affects the robustness in feature match. Therefore, prior information in SfM method is required (Guo and Gao,2013).

Based on literature, SfM method using optical flow able to recover camera rotation and translation between two corresponding point in fast and less complexity. However, this approach still has issue related with accuracy. For example, false match that considered as outliers need to be rectified. This is to ensure that the camera rotation and translation on image plane is accurate.

3. Camera rotation and translation using optical flow method

This section discuss about the methodology for camera rotation and translation recovery in 3D reconstruction from un-calibrated images. The process is a part of the whole process in 3D reconstruction from un-calibrated images as shown in Fig. 1.

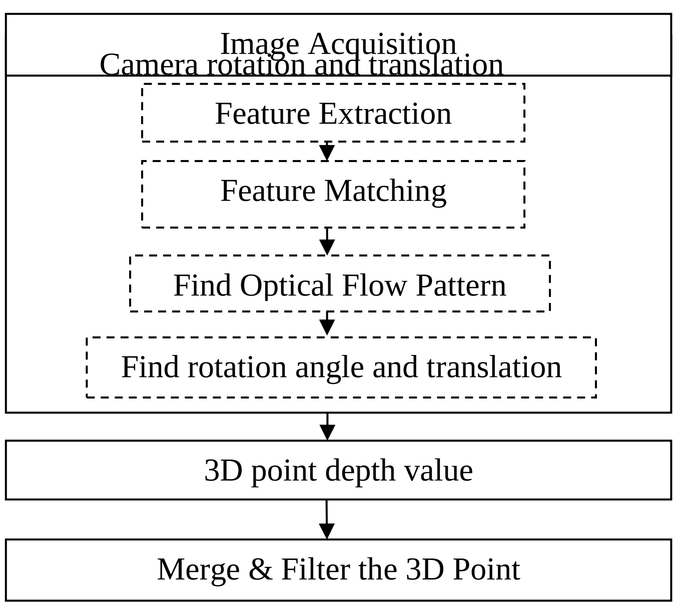
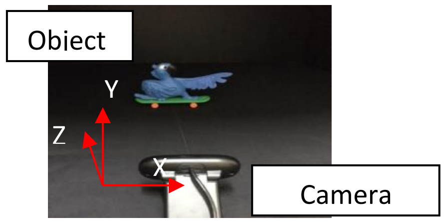
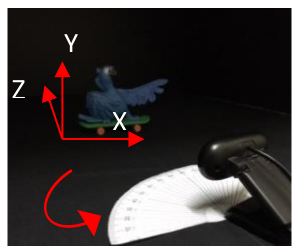


Fig.1. Methodology

3.1 Image Acquisition

This step involves the process to acquire image from various camera position. Experiment was set up using controlled environment to reduce the variables with black background. Image sequences were capture using normal user camera from various camera rotation and translation as shown in Table 1. This image sequences were later will be used to observe the optical flow pattern.

Table 1: Image Acquisition

Image Acquisition	
Different Translation (x , y and z)	
Different Rotation (x , y and z)	

3.2 Camera rotation and translation recovery

In this step, features were extracted from two image sequence using Harris Corner Detection. This process used to select a set of point of interest from the images. After that, the selected features between two images were matched using Pyramidal Lucas Kanade optical flow method. This method consider a pixel $I(x,y)$ in the first frame that moved by distance (dx,dy) in the next frame after dt time. Since this research dealt with image only, therefore time was not needed. Optical flow vectors are used to identify the movement of the camera in the world space. This information later will be used to find the transformation ratio between image plane and world space. This relation is very useful in order to reconstruct the 3D point equation in the next step.

3.3 Preliminary experiment result

This section discuss about the finding on preliminary experiment for camera translation and rotation recovery. Fig. 2 shows features that have been extracted and matched from two images. Meanwhile, Table 2 and 3 show the changes on optical flow pattern when the camera translate and rotate at y-axis camera rotation.

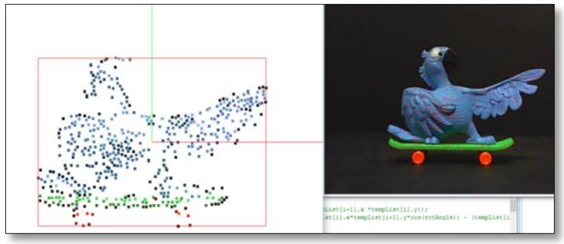
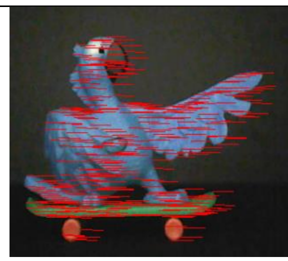
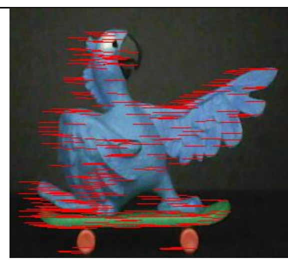
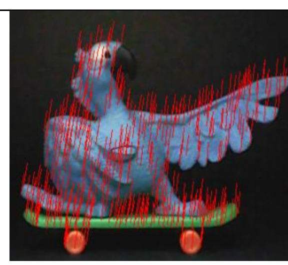
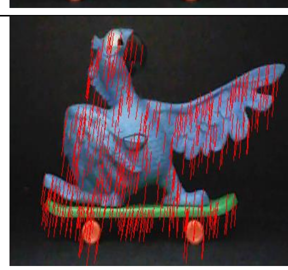


Fig.2. Feature extraction and Matching

Table 2: Camera Translation Recovery

Camera Translation	Optical Flow Pattern	Average Translation
Left (x-axis)		Tx : +6.6321 Ty: -0.0572
Right (x-axis)		Tx: -6.7257 Ty: +0.0546
Upward (y-axis)		Tx: 0.6689 Ty: 4.8915
Downward (y-axis)		Tx: -0.6562 Ty: -4.8637

Forward (z-axis)		Tx: 0.7203 Ty: 0.6527
Backward (z-axis)		Tx: -0.6535 Ty: -0.5903

Table 3: Camera Rotation Recovery

Camera Rotation	Optical Flow Pattern	Average Translation
Left (y-axis)		Tx: +14.3748 Ty: -0.17320
Right (y-axis)		Tx:-10.3344 Ty: +0.3650

Based on the preliminary experiment, camera translation and rotation can be recovered from optical flow vector. This can be obtained from the average translation of x (Tx) and y (Ty) as in Eq. 1.1 and 1.2. Fig.3 shows the algorithm that has been developed to identify the world camera direction from two images.

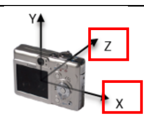
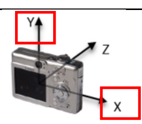
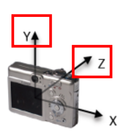
$$\mathrm{Tx} = u_2 - u_1 \qquad (1.1)$$
$$\mathrm{Ty} = v_2 - v_1 \qquad (1.2)$$

```
if ( (Tx = = +) && (Ty= = -)) {
        if (|Ty|<=0.1) {
          Camera translate to the left on x-axis
        else
           Camera rotate to the left on y-axis}
else if (( (Tx = = -) && (Ty= = +)) {
        if (|Ty|<=0.1) {
            Camera translate to the right on x-axis
        else
          Camera rotate to the right on y-axis}
else if (( (Tx = = +) && (Ty= = +)) {
        if (|Tx/Ty| = =1) {
            Camera translate forward on z-axis
        else
          Camera translate upward on y-axis}
else if (( (Tx = = -) && (Ty= = -)) {
        if (|Tx/Ty| = =1) {
            Camera translate backward on z-axis
        else
          Camera translate downward on y-axis}}
```

Fig. 3. Algorithm to identify the camera direction from two images

Once the camera movements have been identified, the next step involves finding the value of rotation angle in image plane based on specific formula as in Table 4.

Table 4: Camera Rotation Angle

Camera Rotation		Formula
y-Axis		$z_2 = z_1\cos\theta - x_1\sin\theta$ $x_2 = z_1\sin\theta + x_1\cos\theta$ $y_2 = y_1$
z-Axis		$y_2 = x_1\sin\theta + x_1\cos\theta$ $x_2 = x_1\cos\theta - y_1\sin\theta$ $z_2 = z_1$
x-Axis		$y_2 = y_1\cos\theta - z_1\sin\theta$ $z_2 = y_1\sin\theta + z_1\cos\theta$ $x_2 = x_1$

4. Conclusion & future work

This paper focuses on preliminary experiment on camera rotation and translation recovery based optical flow feature matching. This is one of the important step to relate the image transformation between image plane and world space. The information is valuable to find 3D point in the next step of 3D reconstruction from un-calibrated images. This task become more challenging since the images rotation and translation derive from various camera position. Existing methods that used to recover camera rotation and translation still encounter the problem with higher processing time, complexity and accuracy. Moreover, current methods based on optical flow only focus on certain position of camera rotation angle and translation. This situation is not always reliable for single moving camera. Therefore, this research aim to propose a new method to recover camera rotation and translation based on optical flow for 3D reconstruction from un-calibrated images.

Result from the preliminary experiment shows that camera translation in all axes and y-axis rotation can be recovered from optical flow vector. Further experiment will be conducted for camera rotation in x and z axis. From this result, the translation value and rotation angle will be recovered using specific formula. After that, the 3D point equation will be reconstructed based on relation between image plane and world space camera coordinates. Finally, all sets of points will be merged together to represent the real 3D object. In future work, the 3D point accuracy and performance will be validated based on benchmark data reconstructed from calibrated camera setting. The proposed method also will be compared with the previous methods.

5. Acknowledgement

Financial support from Fundamental Research Grant Scheme (FRGS) by the ministry of Higher Education (MOHE) is gratefully acknowledged.

References

Diskin, Y., and Asari,V. K. (2013). Dense 3D point-cloud model using optical flow for a monocular reconstruction system. 2013 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), 1–6.

El,S. and Saaidi, A. (2014) Multi-view Passive 3D Reconstruction : Comparison and Evaluation of Three Techniques and A New Method for 3D Object Reconstruction, 194–201.

Guo, F.-S. ,and Gao. (2013). WBatch Reconstruction from UAV Images with Prior Information. Acta Automatica Sinica, 39(6),834–845.

Haala, N., Cramer, M., and Rothermel, M. (2013). Quality Of 3d Point Clouds From Highly Overlapping Uav Imagery. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL(September), 4–6.

Izadi, S., Davison, A., Fitzgibbon, A., Kim, D., Hilliges, O., Molyneaux, D., and Freeman, D. (2011). Kinect Fusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera. In Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST ’11, 559.

Ji, Q., and Dai, S. (2004). Self-calibration of a rotating camera with a translational offset. IEEE Transactions on Robotics and Automation, 20(1), 1–14.

Kanatani, K. I. (1988). Transformation of Optical Flow by Camera Rotation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 10(2), 131–143.

Kang, L., Wu, L., Chen, X., and Yang, Y.-H. (2013). Practical structure and motion recovery from two uncalibrated images using Constrained Adaptive Differential Evolution. Pattern Recognition, 46(5), 1466–1484.

Karami, M., Afrouzian, R., Kasaei, S., and Seyedarabi, H. (2014). Multiview 3D Reconstruction based on Vanishing Points and Homography, 367–370.

Kume, H., Taketomi, T., Sato, T., and Yokoya, N. (2010). Extrinsic Camera Parameter Estimation Using Video Images and GPS Considering GPS Positioning Accuracy. 20th International Conference on Pattern Recognition.

Liang, C. and Wong, K. Y. (2010). 3D reconstruction using silhouettes from unordered viewpoints. Image and Vision Computing, 28(4), 579–589.

Ming-Xin, J., Zhi-Jing S., and Hong-Yu, W. (2012). Real-Time Object Tracking Algorithm With Cameras Mounted On Moving Platforms, International Journal of Image and Graphics.

Nex, F., and Remondino, F. (2013). UAV for 3D mapping applications: a review. Applied Geomatics, 6(1), 1–15.

Nguyen, H., Wunsche, M., Delmas, B. C., Lutteroth, P. C., and Van Der Mark, W. (2013). High resolution 3D content creation using unconstrained and uncalibrated cameras. 2013 6th International Conference on Human System Interactions, HSI 2013, 637–644.

Ohta,N.(2015).3-D Reconstruction from optical flow based on sem iparametricmodel.http://www.researchgate.net/publication/228831969.

Pollefeys, L. M., Van Gool, M., Vergauwen, F., Verbiest, K., Cornelis, J. Tops, & R. Koch.(2004). Visual modeling with a hand-held camera. International Journal of Computer Vision, 59(3), 207–232.

Snavely, N., Seitz, S. M., and Szeliski, R. (2008). Modeling the world from Internet photo collections. International Journal of Computer Vision, 80(2), 189–210.

Tanathong, S., & Lee, I. (2014). Translation-based KLT tracker under severe camera rotation using GPS/INS data. IEEE Geoscience and Remote Sensing Letters, 11(1), 64–68.

Toldo, R., Gherardi, R., Farenzena, M., and Fusiello, A.(2015). Hierarchical structure-and-motion recovery from uncalibrated images. Computer Vision and Image Understanding, 140, 127–143.

Zakaria, A. M., and Said, M. A.(2010). 3D Reconstruction of a scene from multiple un-calibrated images using close range photogrammetry. 2010 International Symposium on Information Technology, 1, 1–5.

Zhao, L., Liu, S., Li, J., and Xu, H.(2010). An approach for 3D reconstruction from uncalibrated images. In International Conference on Challenges in Environmental Science and Computer Engineering, CESCE 2010, Vol. 1, pp. 390–393.