ROBUST ESTIMATION OF A MULTI-CAMERA SYSTEM MOTION PARAMETERS USING INTER-CAMERA MOUNTING PARAMETERS

Mehdi Mazaheri[a], Ayman Habib[b]
[a] Department of Geomatics Engineering, University of Calgary,
2500 University Drive NW, Calgary, Alberta, T2N 1N4, CANADA- mmazaher@ucalgary.ca
[b] Lyles School of Civil Engineering, Purdue University,
550 Stadium Mall DriveWest Lafayette, Indiana, 47907-2051, USA– ahabib@purdue.edu

ABSTRACT

Accurate Indoor 3D models could be a valuable asset for many applications such as building reconstruction, rescue operations, and Building Information Management (BIM). Through image processing, 3D models can be generated from high quality images captured by off-the-shelf digital cameras. To acquire redundant data and produce real scale models, a multi-camera system can be used. This paper presents a methodology for robust estimation of motion parameters of a multi-camera system. First, the interior orientation parameters and mounting parameters among the cameras are estimated through a single-step procedure using images covering a calibration test field. Then, synchronized images are taken at a given time interval while the system is moving through building corridors. In the next step, features are extracted and matched between all possible intra-epoch and inter-epoch image pairs. False matches across Epipolar lines are filtered out in two steps: a) using the known mounting parameters for the intra-epoch matches and b) while estimating inter-epoch Relative Orientation (RO) parameters. In the proposed methodology, the system rotation (R) and translation vector (r) between successive epochs are estimated through rotation and translation compatibility constraints that involve the mounting parameters among the individual cameras on board the system. Knowing the system motion parameters (R and r), outliers along and across Epipolar lines can be detected. Preliminary results have shown that the system motion parameters are reliably estimated using the proposed methodology.

KEYWORDS: multi-camera system, motion estimation, rotation averaging

INTRODUCTION

An accurate 3D model of indoor urban environments can be a valuable asset to cultural heritage documentation, generation of virtual environments, city planning, urban design, and fire & police planning (Chow, 2014). A 3D model will help us to find an appropriate path to or out of a building in case of emergencies, especially when vision is limited due to smoke or blockage in a part of building. Accurate 3D models facilitate metrology – e.g., area and volume computation, which are needed for reconstruction and maintenance purposes. Augmented reality, video games, robot navigation, and disaster management are some of the other applications in which indoor 3D models can play an important role.

One of the major methods to generate accurate and textured 3D models is imagery. Compared to laser scanning, imaging sensors are more attainable and cost significantly lower. These days, a few hundred dollar digital camera is capable of taking high quality images at different conditions, which is suitable for accurate 3D modelling. However, extensive post processing is required to extract 3D information out of 2D images. By using multiple cameras, larger area of the scene is captured and redundant data is recorded, which improves efficiency of 3D reconstruction algorithms. In addition, true scale recovery of scene geometry is directly possible by taking the rigid inter-camera mounting parameters into account. The mounting parameters are the translation and rotation between a reference camera and the remaining ones, which are estimated together with the interior orientation parameters of the individual cameras through a multi-camera system calibration (Detchev et al., 2014). The system calibration should also address the synchronization of the different cameras especially when dealing with a dynamic scene or a moving multi-camera system.

Estimation of system trajectory is the most important and challenging part of 3D reconstruction. The system trajectory is described by the Exterior Orientation Parameters (EOPs) of the reference camera throughout the data acquisition campaign. The EOPs of the remaining cameras are derived from those associated with the reference camera using the mounting parameters. The system trajectory can be optimized through a bundle adjustment that keeps the inter-camera relative orientation fixed.

Various methods to estimate the trajectory of single/multi- camera systems have been studied for Simultaneous Localization And Mapping (SLAM), Visual Odometry, and Structure from Motion (SfM). A simple method to reconstruct the trajectory is 2D to 2D correspondence. In this method, a new image in a given sequence is oriented relative to the previous image by correspondent image points and rescaled using common points within the existing model. Other approaches to estimate the system trajectory, which is based on 2D/3D correspondences, are reviewed in Scaramuzza and Fraundorfer (2011). While the trajectory is being reconstructed, a local bundle adjustment can be used to optimize the EOPs and 3D points, which is called a sliding-window-based bundle adjustment (Fraundorfer et al., 2010).

One way to robustly estimate the system motion is to incorporate all the possible relative orientations – which are up to a scale – to find the absolute orientation. Govindu (2001, 2004, 2006) published series of papers for robust motion estimation from a set of pair-wise relative orientation parameters. For a sequence of *N* images, at most *N* (*N*-1)/2 relative rotation matrices R_i^j between image pairs (*i*, *j*) can be computed. The absolute rotation of every image *i* relative to the world coordinate system (R_i^w) can be robustly estimated while considering the compatibility constraint $R_j^w R_i^j = R_i^w$. This problem is also called multiple rotation averaging, which is reviewed in Hartley et al. (2013)

In this research, a methodology is developed to robustly estimate the rotation and true scale translation of the reference camera, using the rigid inter-camera mounting parameters and the pair-wise relative orientation of the cameras between successive epochs. Knowing the reference camera motion, the EOPs of all the cameras at any epoch can be computed. The rest of paper describes the proposed methodology and the experimental result using a real dataset.

PROPOSED METHODOLOGY

Figure 1 shows reference and slave cameras at two different epochs. As mentioned earlier, the system motion is defined by the rotation and translation of the reference camera c_r from time t_1 to t_2, which are denoted by $R_{c_r(t_2)}^{c_r(t_1)}$ and $r_{c_r(t_2)}^{c_r(t_1)}$, respectively. The rigid rotation and translation between the reference camera (c_r) and a slave camera (e.g., c_i) is defined by $R_{c_i}^{c_r}$ and $r_{c_i}^{c_r}$ (mounting parameters). Assuming the *n* mounted cameras have overlapping field of view, and the displacement between two successive epochs (t_1 and t_2) is short, n^2 sets of relative orientation parameters between the cameras can be estimated. The possible pair-wise relative orientation parameters are depicted in Figure 1.

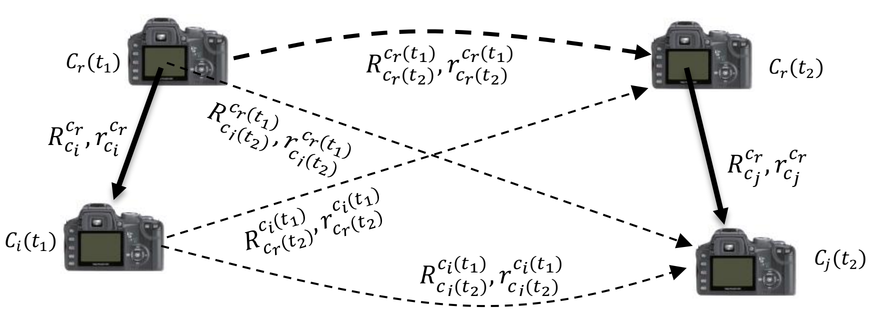


Figure 1: Reference and slave cameras at two different epochs t_1 and t_2, and the possible sets of pair-wise relative orientation parameters

To find the relative orientation between the cameras at different epochs, the Essential matrix is estimated by a RANSAC and 5-point algorithm (Li and Hartley, 2006), and decomposed to rotation and translation components. At this stage, outlier feature matches across the Epipolar lines are filtered out. Using the mounting parameters, all the pair-wise relative rotation matrices between the cameras at different epochs can be used to derive the relative rotation matrix of the reference camera between epochs t_1 and t_2 – i.e., $R_{c_r(t_2)}^{c_r(t_1)}$. Therefore, n^2 independent estimates of the rotation matrix $R_{c_r(t_2)}^{c_r(t_1)}$ are obtained. The estimates are then averaged to estimate the system rotation robustly. For the true scale translation of the reference camera ($r_{c_r(t_2)}^{c_r(t_1)}$), a system of linear equations is formed using all the estimated translations between the camera pairs at different epochs. Knowing the system motion parameters, outlier matches along the Epipolar lines can be identified – which are not detectable by stereo Epipolar geometry. In the next two sections, the concept of rotation and translation estimation using the derived sets of relative orientation parameters are explained.

Rotation Estimation

In order to robustly estimate the rotation of the reference camera from t_1 to t_2, $R_{c_r(t_2)}^{c_r(t_1)}$, compatibility constraints between the estimated pair-wise rotation matrices are established. Significant overlap among the images between successive epochs, allow for the estimation of the relative orientation parameters between a reference/slave camera at time t_1 and a slave/reference camera at time t_2. Therefore, four types of rotation constraints between the cameras are possible – i.e., $[c_r(t_1), c_r(t_2)]$, $[c_r(t_1), c_j(t_2)]$, $[c_i(t_1), c_r(t_2)]$, $[c_i(t_1), c_j(t_2)]$, where *i*, *j* are slave camera indices. Each compatibility constraint yields to an estimate of the rotation matrix between $c_r(t_1)$ and $c_r(t_2)$ – i.e., $R_{c_r(t_2)}^{c_r(t_1)}$.

In general, n^2 estimates for the system rotation between two epochs ($R_{c_r(t_2)}^{c_r(t_1)}$) can be derived. The problem of averaging estimates of a rotation matrix is called single rotation averaging (Hartley et al., 2013). In this research, angle-axis representation of the rotation matrix is used for averaging. Each estimate of $R_{c_r(t_2)}^{c_r(t_1)}$ is decomposed to an angle of rotation θ around a unit axis v. Care must be taken that $2\pi - \theta$ and $-v$ also defines the same rotation, so the decomposition is not unique. In this work, the magnitude of θ is always small between successive epochs, so the decomposition with positive rotation angle is selected. However, some of the rotation estimates could be outlier due to weak estimation of relative orientation in poor textured areas or in case of short baseline. A rotation estimate is considered as outlier when the angle of rotation θ falls outside a given confidence interval. After removing the outliers, the inlier sets of angle-axis are averaged to robustly evaluate the rotation matrix $\bar{R}_{c_r(t_2)}^{c_r(t_1)}$.

Translation Estimation

After computing the system rotation ($\bar{R}_{c_r(t_2)}^{c_r(t_1)}$), the true scale translation of the reference camera between two successive epochs is estimated, which is denoted by $\bar{r}_{c_r(t_2)}^{c_r(t_1)}$. Similar to the rotation estimation, four translation compatibility constraints can be established. Each compatibility constraint is a vector summation that involves the unknown translation vector $\bar{r}_{c_r(t_2)}^{c_r(t_1)}$, and the estimated translation of the camera pairs between t_1 and t_2. Each translation compatibility constraint is a set of three equations in four unknowns – the three elements of translation $\bar{r}_{c_r(t_2)}^{c_r(t_1)}$ and a scale factor λ for the translation of the camera pairs between two epochs. Therefore, for n^2 possible compatibility constraints, a system of linear equations is formed with $3n^2$ equations in n^2+3 unknowns, which directly solves for the three elements of $\bar{r}_{c_r(t_2)}^{c_r(t_1)}$ and the n^2 scaling factors.

EXPRIMENT AND RESULTS

The data acquisition system consists of three cameras mounted on a cart and a laptop equipped with a software that controls the cameras and records synchronized images (Figure 2-a). The cameras are tightly fixed while being slightly tilted inward to increase the images overlap. The interior orientation and mounting parameters of the cameras are estimated through a one-step bundle adjustment procedure (Detchev et al., 2014), using images of the test field shown in Figure 2-b, which is specifically designed for multi-camera system calibration.

(a)

(b)

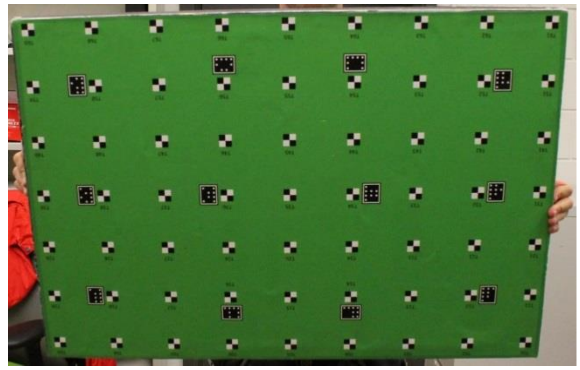
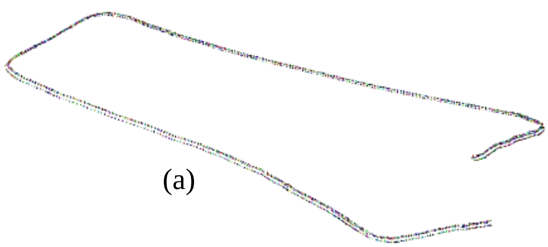
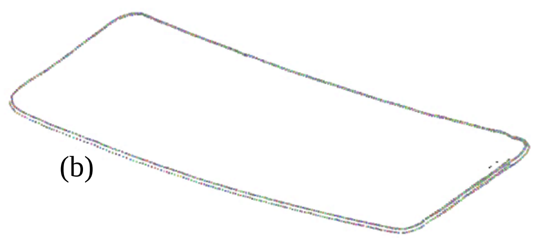


Figure 2: a) The multi-camera data acquisition system and b) calibration test field

The multi-camera system was pushed through a corridor loop and 1656 images (3 cameras*552 epochs) were captured every 1.5 seconds. In this experiment, nine sets of relative orientation parameters could be estimated between each camera at one epoch and the three cameras at the next epoch. As mentioned earlier, the nine sets of relative orientation parameters are related to the reference camera (top camera) and averaged to robustly estimate the rotation of the reference camera between successive epochs ($\bar{R}_{c_r(t_2)}^{c_r(t_1)}$). The translation of the reference camera between successive epochs ($\bar{r}_{c_r(t_2)}^{c_r(t_1)}$) are then estimated by solving a system of linear equations. Knowing the robust motion parameters of the reference cameras between the successive epochs ($\bar{R}_{c_r(t_2)}^{c_r(t_1)}$, $\bar{r}_{c_r(t_2)}^{c_r(t_1)}$), the EOPs of all the images can be computed. Figure 3-a shows the estimated trajectory of the system using the proposed methodology. In this figure, the beginning of the trajectory loop is quite far from the end, which is due to accumulation of errors during the sequential estimation of the trajectory. This error accumulation is inevitable for such a long sequence, but could be minimized through a global bundle adjustment using all the images. Figure 3-b shows the optimized trajectory after global bundle adjustment. The proposed strategy can successfully reconstruct the approximate trajectory, to be optimized by global bundle adjustment.



(a)



(b)

Figure 3: a) The estimated trajectory by the proposed method and b) optimized trajectory after bundle adjustment

CONCLUSIONS AND FUTURE WORK

In this paper, the concept of trajectory estimation for a multi-camera system using pair-wise relative orientation parameters was described. The constant inter-camera mounting parameters allows for relating the estimated pair-wise relative orientation parameters to the reference camera, and robustly evaluate the system motion parameters. Knowing the system motion parameters will further allow for the removal of outliers and provide reliable approximations for a global bundle adjustment to estimate accurate trajectory.

In poor textured areas or very small movement of the system (e.g., by rotating the multi-camera system around itself), the relative orientation parameters are weakly estimated. In this case, detection of bad estimates requires a more sophisticated algorithm, which will be the focus of our future work. In addition, other rotation averaging methods will be investigated and tested.

REFERENCES

Chow, J., 2014. *Multi-Sensor Integration for Indoor 3D Reconstruction*. PhD. Thesis, University of Calgary.

Detchev, I., Mazaheri, M., Rondeel, S., and Habib, A., 2014. Calibration of multi-camera photogrammetric systems. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XL-1*: 101–108.

Fraundorfer, F., Scaramuzza, D., and Pollefeys, M., 2010. A constricted bundle adjustment parameterization for relative scale estimation in visual odometry. *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pp. 1899–1904.

Govindu, V., 2006. Robustness in Motion Averaging. P. J. Narayanan, S. Nayar, & H.-Y. Shum (Eds.), *Computer Vision – ACCV 2006*. Springer Berlin Heidelberg, Vol. 3852, pp. 457–466.

Govindu, V. M., 2001. Combining two-view constraints for motion estimation. *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, Vol. 2, pp. 218–225.

Govindu, V. M., 2004. Lie-algebraic averaging for globally consistent motion estimation. *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, Vol. 1, pp. 684–691.

Hartley, R., Trumpf, J., Dai, Y., and Li, H., 2013. Rotation Averaging. *International Journal of Computer Vision*, *103*(3): 267–305.

Li, H., and Hartley, R., 2006. Five-Point Motion Estimation Made Easy. *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, Vol. 1, pp. 630–633.

Scaramuzza, D., and Fraundorfer, F., 2011. Visual odometry Part I: The First 30 Years and Fundamentals. *Robotics & Automation Magazine, IEEE*, *18*(4): 80–92.