Vol. 8(27), Jan. 2018, PP. 3705-3711

Contents list available at IJMEC

International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)

Journal Homepage: www.aeuso.org

PISSN: 2411-6173 - EISSN: 2305-0543

Minimization of Reprojection Error in Scene Metric Reconstruction from Multiple Images

Abdulkadir Iyyaka Audu[1*] and Abdul Hamid Sadka[2]

[1] *Computer Engineering Department, Faculty of Engineering, University of Maiduguri*

[2] *Department Electronic and Computer, Engineering, Brunel University, London, United Kingdom*

Phone Number: *+234-0709252528*

****Corresponding Author's E-mail:*** lm324fairchild@gmail.com

Abstract

One problem which has remained deeply rooted in multiview geometry is the reconstruction of a scene starting with multiple two-dimensional representation of different parts of the same scene. This is usually formulated as a non-linear problem which has to be solved using an iterative optimization process starting from a sub-optimal solution obtained by using linear methods. In this paper, the main focus is to start from first principle and determine the optimization technique which can minimize repropjection error during metric reconstruction of a three-dimensional scene from multiple image samples.

Keywords: *Metric reconstruction, sparse bundle adjustment, multiple images, centre of projection, reprojection error.*

1. Introduction

The objects of interest in the world are usually a set of points. The intensity acquisition of these points is achieved through the non-singular transformation capability of cameras which are almost usually based on signal processing principles, measurement, and algorithms [1]. Although important research and developmental effort have been targeted at camera fabrication based on state-of-the art technology, there still exist practical limits on the intrinsic parameters such as focal length. [2]. For example, the dependence of two important parameters depth of field (DOF) and field of view (FOV) on camera focal length f, working distance u, magnification μ, circle of confusion c, f-number of camera lens f_n, and sensor physical size are expressed in (1) and (2) respectively [2].

$$DOF = \frac{2f^2 f_n c\,(\mu+1)}{f^2\mu^2 - f_n^2 c^2} \tag{1}$$

$$FOV = 2\arctan\left(\frac{d\,(u-f)}{2uf}\right) \tag{2}$$

It is explicitly obvious from (1) and (2) that localization and matching errors [1] are deeply engraved in camera measurements. It turns out that the data acquired with such cameras are difficult to work

Article History:

Received Date: Apr. 23, 2017

Accepted Date: Oct. 04, 2017

Available Online: Jan. 04, 2018

with especially in scene reconstruction and other certain critical application areas, and often does not give accurate result. It is important that these errors be adequately attended to and considerably reduced in order to improve accuracy, reliability, and computational performance issues in image processing applications such as three-dimensional (3D) reconstruction.

This work is intended to investigate the visual reconstruction of a scene from multiple images acquired using a camera with a single center of projection (SCP) as opposed to multiple centers of projection (MCP) which is depicted in Figures 1 and 2 respectively. Scene visual reconstruction by way of bundle adjustment attempts to recover a model of a 3D scene from multiple images [3]. As part of this, it usually also recovers the poses (positions and orientations) of the cameras that took the images, and information about their internal parameters [4, 5, 6, 7, 8, 9, 10]. Bundle adjustment can cope with any model that predicts the values of some known measurements or descriptors on the basis of some continuous parametric representation of the world, which is to be estimated from the measurements. It can incorporate several types of image observations along with their associated precisions and it can provide statistical information regarding the quality of the solution. Bundle adjustment tools can accommodate different sensor types, e.g. frame cameras and linear array scanners. Bundle adjustment methods can easily handle missing data, i.e. points which are not visible in a given image due to occlusion or insufficient object coverage. They can also handle redundant data, i.e. points which are visible in any number of images [11, 12, 13, 14, 15, 16, 17, 18].

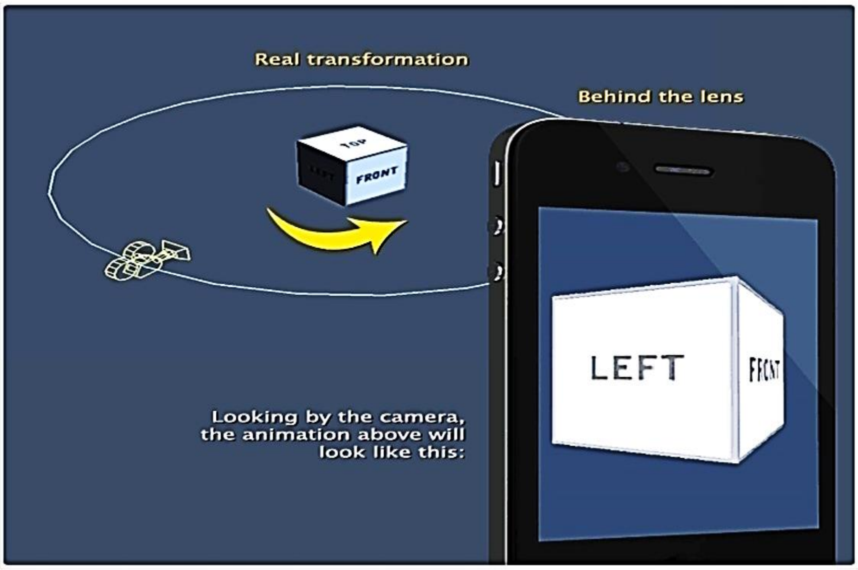

Figure 1: Single center of projection.

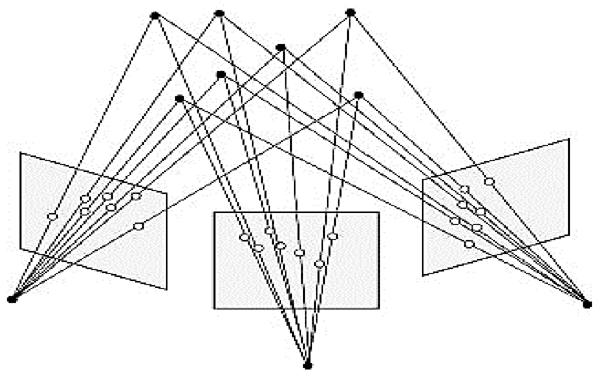

Figure 2: Illustration of multi-view with cameras with different center of projection.

Since different parts of the scene are observed from the same point of different directions with a single camera, the determination of point correspondences becomes a major challenge. Also the camera parameters (intrinsic and extrinsic) and the 3D reconstructed points contribute to the computational complexity of the process [19]. The number of the camera parameters is directly decided by the number of images but the number of the 3D reconstructed points is largely decided by the image resolution and the scene texture. This problem is addressed by using optimization algorithm in the form of bundle adjustment algorithm (SBA) to maximize the likelihood of reconstruction. One interesting aspect of this study is that apart from intensity variation, there is a strong texture variation of the scene.

SBA is an active research area in computer vision. It provides a concise and straightforward introduction as well as a detailed coverage on the unique research problem of re-projection error minimization in a wide range of stereo imagery. It adequately addresses the problems of high computational and memory costs associated with least squares optimization techniques used in the early years of bundle adjustment. This is in view of the fact that the Jacobian of some parameters exhibit sparsity which can be tailored to improve computational speed. This concept has been successfully applied in parameter refinement problems and has allowed inconsistencies between stereo pairs to be removed. It has also enhanced the realistic descriptions and modeling of other sophisticated imaging system application [20, 11, 21].

There are two trends which have been widely adopted in scene reconstruction algorithms based on SBA. One situation is when the scene is fixed and the camera motion is around it. This is of course simple to implement with hand-held cameras, otherwise a specialized rig is required. When a hand–held camera is used the recovery of homographies that map images to each other and consequently allow the images to be transformed and combined becomes a huge challenge. Alternatively, the scene object can be placed on a turn-table and images are acquired at regular angular interval with a fixed camera. This is only realizable if the scene is a portable object or cast object.

Understandably, [22], the stereo image data obtained by High-Resolution Imaging Science Experiment (HiRISE) and used in the characterization of the surface, subsurface, and atmosphere of Mars have been optimized using SBA, in view of the scientific relevance of the project. The project is aimed at obtaining high-precision topographic information which is at the core of most space exploration programmes.

SBA is also known to have been used in image mosaicking [23]. This approach facilitates an accurate 3D reconstruction from multiple images of the same. This procedure features result which is statistically optimal based on the enforcement of hard geometric constraints. Frame decimation discussed in [24] provides for an automatic method which decides on the frame rate for any image sequence to be used in a structure from motion problem. This idea makes use of a coarse to fine, optical flow-based video mosaicking algorithm.

Expansion procedure is adopted in [25] to output a qualitative dense set of patches on the surface of an object. Feature points obtained by matching multiple images are used to generate initial patches. These are continuously expanded until dense patches are obtained. The optimality of the patches is determined by local photometric consistency and global visibility constraints.

In simultaneous localization and mapping (SLAM) problem, [26], classical bundle adjustment (BA) technique has been used to precisely describe and estimate the position of a mobile robot on a constructed map of the same environment.

Cost function is very critical in the formulation of SBA. The comprehensive theoretical background of SBA given in [27] defines a robust cost function as the square sum of all the dimensions of an error vector function. This is contrary to the conventional method of approximating the cost function locally with a quadratic Taylor expansion. The discussion also provides an entirely different proposition in which it is investigated that error-buildup can be described as a function of accuracy that is obtainable through the use of bundle adjustment and helps to improve the reliability of camera tracking. The discussion also focuses on how bundle adjustment facilitates real-time application.

It is observed that the definition of variables in bundle adjustment is critical when large error propagation is needed in order to correct global error at loop closure. This will involve constraining the expected solution. An adaptive bundle adjustment is proposed in [28] which works in a metric-space defined by a connected Riemannian manifold. This is aimed at addressing the problem of single coordinate frame which is responsible for the high computational cost of bundle adjustment.

Since line features from image correspondences is an integral part of scene modeling, augmented reality, and visual servoing [29], the 'two points' and the 'two plane' overparameterization which can cause gauge freedoms and/or internal constraints is addressed with plucker coordinates so that the feature lines can be represented in 3D.

In [18], solution instability due to the linear dependencies between parameters when perspective algorithms (collinearity based) are used has been identified to be common in imaging situations where long focal length lenses and narrow FOV play a prominent role. This problem is addressed by the use of a scaled orthographic projection model based on linear algebraic formulations. Using quaternions, The mathematical model developed uses quaternions which translate to partial derivatives as well as the inner constraint equations for a scaled orthographic bundle adjustment.

In the method adopted by [30], fast direct Cholesky decomposition techniques are employed to solve SBA problem with sparse linear subproblem.

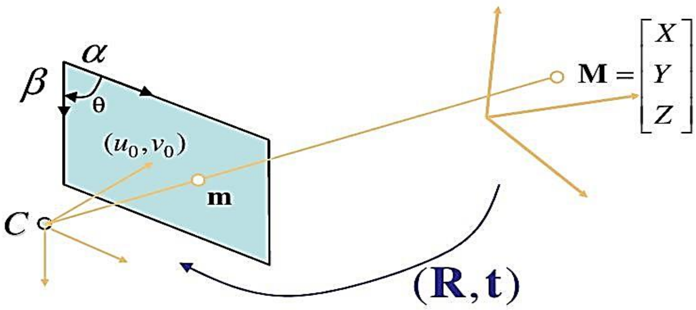

Figure 3: Perspective Camera Projection.

The technique involves the use of a compressed representation of large sparse matrices for efficiently handling the block data structures of SBA to take advantage of this representation. The performance evaluation is proved to surpass what is known to be obtainable from the method used in [1].

2. Bundle Adjustment Problem Formulation

Bundle adjustment is a technique used to compute the maximum likelihood estimate of structure and motion from image feature correspondences. It exploits the sparse primary structure of the problem, where connections exist just between points and cameras. This is a non-linear system problem which have to be solved using an iterative optimization process starting from a sub-optimal solution obtained by using linear methods.

A camera can be modeled in several different ways. Affine and orthographic projections are sometimes useful for distant cameras, and more exotic models such as push-broom and rational polynomial cameras are needed for certain applications [11]. Other camera models can be derived from it. But in addition to pose (position and orientation), and simple internal parameters such as focal length and principal point, real cameras also require various types of additional parameters to model internal aberrations such as radial distortion. However, perspective projection as shown in Figure 3, [31], is the most prominent.

Perspective projection is the linear mapping between the extended coordinates of any world point *M* and its corresponding image point *m* such that collinearity property exist between *M, m*, and *C* (centre of projection). This can be expressed as

$$\Lambda m = QM = P \tag{3}$$

where *M* is 4 × 1 vector and *m* is 3 × 1 vector. Λ is an arbitrary scale factor. Q is a 3 × 4vector referred to as the projection matrix. Therefore, in reality, an object or structure consists of several *M* points and image of such an object or structure will consist of a corresponding number of *m* points. An important characteristic which can be exhibited by a perspective camera is for the first three left columns of *M* to be non-singular. It therefore means *Q* can be further decomposed such that

$$\Lambda m = K \left[R | \quad t \right] M \tag{4}$$

K is 3 × 3 upper triangular matrix. It is called camera or calibration matrix of the camera. It comprises of the optical properties of the camera namely: Focal length, principal point and aspect ratio. *R* is an orthogonal 3 × 3 matrix and *t* a 3 × 1 vector. R and t are collectively referred to as the camera's extrinsic orientation and correspond, respectively, to the rotation and translation that make up the rigid transformation from the world to the camera coordinate frame [1]. The coordinate system *C* attached to the camera is related to the world coordinate system through a rotation *R* followed by a translation t.

For a multi-view setting, consider having M_j scene points captured by several cameras described by Q^i. Assuming the projection of M_j point due to Q^i camera is m^i_j. Starting from multi-view image samples, multi-view 3D reconstruction involves the determination of M_j and Q^i such that (3) is satisfied as expressed in (5).

$$\Lambda m_j^i = Q^i M_j \tag{5}$$

A significant challenge in SBA is that (5) is not exactly satisfied. m^i_j has inherent noise superimposed during the measurement process. Therefore, for every image point m_j^i, a predictive model $\Lambda m_j^i = m(Q^i, M_j)$, [11], is required

such that

$$\Delta \hat{m}_j^i \equiv m_j^i - m\left(Q^i, M_j\right) \tag{6}$$

$$f\left(Q, M\right) = \sum_{i=1}^{c} \sum_{j=1}^{d} V_{ij} || m_j^i - m\left(Q^i, M_j\right) ||^2 \tag{7}$$

3. Discussion

Equation (6) is referred to as re-projection error. Hence the problem of scene reconstruction and camera parameter estimation boils down to the minimization of the re-projection error between the image locations of observed and predicted image points, which is expressed as the sum of squares of a large number of nonlinear, real-valued functions. The objective function for the minimization problem defined in the context of bundle adjustment is expressed in (7), where *c* is the number of scene points and *d* is the number of cameras. V_{ij} is a visibility weight which equals 1 if a scene point *j* can be seen in camera *i*, otherwise it equals 0. If the unexpected variation in pixel coordinates is modeled as Gaussian noise with zero mean, then (7) becomes a statistical nonlinear model. Using the condition of linear independence of the columns of Q^i, (7) can be expressed as a set of linear equations as in (8)

$$\left(Q^i\right)^T Q^i \hat{f} = \left(Q^i\right)^T m_j^i \tag{8}$$

Conclusions

The parameters of (8) can be estimated using least-squares algorithm like Levenberg-Marquardt (LMA) [32]. However, for a large set of object points and camera parameters which, constitute the unknown contributing to the minimized re-projection error, the system represented by (8) becomes overdetermined. The computational cost will then have cubic complexity [28]. Therefore, a specialized LMA known as SBA is required in order to seek a minimal solution.

References

[1] M. I. Lourakis, A. A. Argyros, Sba: A software package for generic sparse bundle adjustment, ACM Transactions on Mathematical Software (TOMS) 36 (1) (2009) 2.

[2] A. V. Bhavsar, A. Rajagopalan, Towards unrestrained depth inference with coherent occlusion filling, International journal of computer vision 97 (2) (2012) 167–190.

[3] N. BA rlin, P. Grussenmeyer, Bundle adjustment with and without damping, The Photogrammetric Record 28 (144) (2013) 396–415.

[4] G. Wang, J. Wu, Z. Ji, Single view based pose estimation from circle or parallel lines, Pattern Recognition Letters 29 (7) (2008) 977–985.

[5] A. Ansar, K. Daniilidis, Linear pose estimation from points or lines, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (5) (2003) 578–589.

[6] F. Duan, F. Wu, Z. Hu, Pose determination and plane measurement using a trapezium, Pattern Recognition Letters 29 (3) (2008) 223–231.

[7] K. Rahbar, H. R. Pourreza, Inside looking out camera pose estimation for virtual studio, Graphical Models 70 (4) (2008) 57–75.

[8] L. Quan, Z. Lan, Linear n-point camera pose determination, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (8) (1999) 774–780.

[9] H. F. Ugurdag, S. GA ren, F. Canbay, Gravitational pose estimation, Computers and Electrical Engineering 36 (6) (2010) 1165–1180.

[10] S. Zhang, Y. Ding, K. Hao, D. Zhang, An efficient two-step solution for vision-based pose determination of a parallel manipulator, Robotics and Computer Integrated Manufacturing 28 (2) (2012) 182.

[11] B. Triggs, P. F. McLauchlan, R. I. Hartley, A. W. Fitzgibbon, Bundle adjustment - a modern synthesis, Vision algorithms: theory and practice, Springer, 2000, pp. 298–372.

[12] M. yi Shen, Z. yu Xiang, J. lin Liu, Vision based terrain reconstruction for planet rover using a special binocular bundle adjustment, Journal of Zhejiang University SCIENCE A 9 (10) (2008) 1341–1350.

[13] L. Reyes, E. Bayro-Corrochano, Simultaneous and sequential reconstruction of visual primitives with bundle adjustment, Journal of Mathematical Imaging and Vision 25 (1) (2006) 63–78.

[14] M. Lhuillier, Incremental fusion of structure-from-motion and gps using constrained bundle adjustments, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (12) (2012) 2489–2495.

[15] R. I. Hartley, Euclidean reconstruction from uncalibrated views, Applications of invariance in computer vision, Springer, 1994, pp. 235–256.

[16] D. D. Lichti, C. Kim, S. Jamtsho, An integrated bundle adjustment approach to range camera geometric self-calibration, ISPRS Journal of Photogrammetry and Remote Sensing 65 (4) (2010) 360–368.

[17] B. Fodor, C. Kazo, Z. Janko, L. Hajder, Normal map recovery using bundle adjustment, IET Computer Vision 8 (1) (2014) 66.

[18] K. F. Blonquist, R. T. Pack, A bundle adjustment approach with inner constraints for the scaled orthographic projection, ISPRS Journal of Photogrammetry and Remote Sensing 66 (6) (2011) 919–926.

[19] X. Liu, W. Gao, Z.-Y. Hu, Hybrid parallel bundle adjustment for 3d scene reconstruction with massive points, Journal of Computer Science and Technology 27 (6) (2012) 1269–1280.

[20] M. Hu, G. Penney, M. Figl, P. Edwards, F. Bello, R. Casula, D. Rueckert, D. Hawkes, Reconstruction of a 3d surface from video that is robust to missing data and outliers: Application to minimally invasive surgery using stereo and mono endoscopes, Medical image analysis 16 (3) (2012) 597–611.

[21] X. Liu, W. Gao, Z.-Y. Hu, Hybrid parallel bundle adjustment for 3d scene reconstruction with massive points, Journal of Computer Science and Technology 27 (6) (2012) 1269–1280.

[22] R. Li, J. Hwangbo, Y. Chen, K. Di, Rigorous photogrammetric processing of hirise stereo imagery for mars topographic mapping, Geoscience and Remote Sensing, IEEE Transactions on 49 (7) (2011) 2558–2572.

[23] P. F. McLauchlan, A. Jaenicke, Image mosaicing using sequential bundle adjustment, Image and Vision Computing 20 (9) (2002) 751–759.

[24] D. NistA˜©r, Frame decimation for structure and motion, 3D Structure from Images-SMILE 2000, Springer, 2001, pp. 17–34.

[25] Y. Furukawa, J. Ponce, Accurate, dense, and robust multiview stereopsis, Pattern Analysis and Machine Intelligence, IEEE Transactions on 32 (8) (2010) 1362–1376.

[26] K. Konolige, K. Konolige, M. Agrawal, Frameslam: From bundle adjustment to real-time visual mapping, IEEE Transactions on Robotics 22 (5) (2008) 1066–1077.

[27] C. Engels, H. StewA˜©nius, D. NistA˜©r, Bundle adjustment rules, Photogrammetric computer vision 2.

[28] D. Sibley, C. Mei, I. Reid, P. Newman, Adaptive relative bundle adjustment., in: Robotics: science and systems, 2009.

[29] A. Bartoli, P. Sturm, Structure-from-motion using lines: Representation, triangulation, and bundle adjustment, Computer Vision and Image Understanding 100 (3) (2005) 416–441.

[30] W. G. Kurt Konilige, Sparse sparse bundle adjustment.

[31] G. Medioni, S. B. Kang, Emerging topics in computer vision, Prentice Hall PTR, 2004.

[32] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, Journal of the Society for Industrial and Applied Mathematics 11 (2) (1963) 431–441.