

APPLICATION AND PERFORMANCE ANALYSIS OF A NEW BUNDLE ADJUSTMENT MODEL

Y. Sun ^{a,b}, X. Liu ^a, R. Chen ^b, J. Wan ^b, Q. Wang ^b, H. Wang ^c, Y. Li ^d, L. Yan ^{b,*}

^a Chinese Academy of Surveying & Mapping, Beijing, 100039, China - syb51@163.com

^b Beijing Key Lab of Spatial Information Integration and Its Application, Peking University, Beijing, 100871, China - lyan@pku.edu.cn

^c Institute of surveying and mapping PLA information engineering university, Zhengzhou, 450052, China – zzch0908@163.com

^d China Top RS Technology Co. Ltd, 16 North Taiping Road, Beijing 100039, China – toprs-lyc@casm.ac.cn

KEY WORDS: Bundle Adjustment, Parallax Angle, Convergence, Aerial Photogrammetry, Nonlinear Optimization

ABSTRACT:

As the basis for photogrammetry, Bundle Adjustment (BA) can restore the pose of cameras accurately, reconstruct the 3D models of environment, and serve as the criterion of digital production. For the classical nonlinear optimization of BA model based on the Euclidean coordinate, it suffers the problem of being seriously dependent on the initial values, making it unable to converge fast or converge to a global minimum. This paper first introduces a new BA model based on parallax angle feature parametrization, and then analyses the applications and performance of the model used in photogrammetry field. To discuss the impact and the performance of the model (especially in aerial photogrammetry), experiments using two aerial datasets under different initial values were conducted. The experiment results are better than some well-known software packages of BA, and the simulation results illustrate the stability of the new model than the normal BA under the Euclidean coordinate. In all, the new BA model shows promising applications in faster and more efficient aerial photogrammetry with good convergence and fast convergence speed.

1. INTRODUCTION

As entering the digital age, photogrammetry has been widely used in nearly all walks of life. The bundle adjustment (BA) model serves as one of the most effective methods for the modern high precision measurement positioning. In photogrammetry, the BA model is one of the most rigorous aerial triangulation methods, which can optimize the posture of the cameras and the three-dimensional coordinates of encrypted dots. This model is also important in generating the digital products and has been applied in many fields (Chaturvedi, 2000; Mikhail et al., 2001; Sauerbier, 2004).

The classic method of the BA model expresses three-dimensional points under the Euclidean coordinate system, and takes the collinear equation as the optimized adjustment equation (Ackermann, 1984). However, a large number of experiments show that this representation method is only effective for close feature points. If the feature points in the environment are too far in distance to access large parallax from the image, great errors and uncertainties of these feature points may emerge in the depth direction, which will cause algorithm divergence. Actually, this situation is ubiquitous in the real world, because the three axes of the Euclidean coordinate system share the same data "dimension" (Yuan, 2009). When they have incremental values in the same order of magnitude, the relative increment value of Z-axis is far less than the lateral relative increment, which can cause the relative error of Z-axis parameters too tiny compared to the lateral ones. In other words, although the three-dimensional relative error matrix is full rank and irrelevant, it still has weak irrelevance due to the tiny value of Z-axis. If an effective operation is to be undertaken, the Z-axis parameters need to be amplified to adapt to the lateral parameters, which can cause the amplification of the error of the Z-axis at the same time.

Since BA is a high dimensional nonlinear optimization problem, in which the multiple iterations need to be solved through Taylor series expansion of linearization, an appropriate initial value is necessary. The more accurate, the initial value is, the greater, the chance of convergence to global optimum will be, and the faster, the convergence speed will be. On the contrary, an unreasonable initial value may lead to optimization problems such as falling into local optimum and divergence, which entails more iterative times.

To solve the problem of slow convergence speed of the BA algorithm based on the Euclidean coordinate system, a large number of ground control points or more accurate orientation information are always used, which are not directly from cameras (Yuan et al., 2009). A differential GPS dynamic positioning technique has once been successfully used to measure the instantaneous spatial location of the camera station, and been used for aerial triangulation which led to a reformation in aerial photogrammetry. However, the technique still has not gotten rid of the bondage of ground control points (Moré, 1978). In addition, many scholars have begun to try different numerical optimization algorithms such as the Gauss-Newton (GN) algorithm, the Levenberg-Marquart (LM) algorithm and the conjugate gradient algorithm (Hartley and Zisserman, 2003; Wu et al., 2011). Some modern optimization methods have also been used in photogrammetry, with the development of artificial intelligence, to solve the problem. However, without changing the collinear equation in Euclidean space, the problem of slow convergence speed could still not be solved essentially.

As for the hard numerical matching between the radius vector of the elevation and the lateral parameters caused by Euclidean coordinates, a new BA model based on parallax angle feature parametrization, ParallaxBA, has recently been proposed (Zhao

* Corresponding author

et al., 2015). This model can work without the consideration of the dimensional irrelevance between the elevation and the lateral parameters, and reduce the relative error of the radius vector. In many cases, normal BA based Euclidean XYZ feature Parameterization will lead to a result containing an ill-conditioned equation and an objective function with very small gradients under certain conditions, while the ParallaxBA can address these bottlenecks (Zhao et al., 2015). In this paper, the applications (especially in aerial photogrammetry) and the performance of the new BA model are discussed, and the convergence performance of the model is particularly analysed, since the divergence and the slow convergence speed is a problem studied for long in the aerial photogrammetry. Experiments and simulations are conducted as well to apply the new BA model and verify its performance.

2. METHODOLOGY

In fact, ParallaxBA was proposed first in the field of computer vision instead of aerial photogrammetry (Zhao et al., 2015). This new BA model is merely a free-net BA model that cannot achieve geo-location for the triangulated tie points, so it fails in supporting some applications for location and mapping in photogrammetry (Yan et al., 2017). To apply it in the aerial photogrammetry, Ground Control Points (GCPs) need to be added into the ParallaxBA model, where tie points and GCPs are represented in the form of angles and XYZ, respectively (Yan et al., 2017). After GCPs being added into the BA functions, the ParallaxBA model can then be used to acquire the spatial three-dimensional points and be applied in aerial triangulation. This section simply introduces the method in the new BA model, including the creation of the observation function and the solution of the functions.

2.1 Observation Function

To conduct BA, the observation functions need to be constructed where different coordinates can be used to represent the feature points. Aiming at accurately expressing the different types of feature points which are close far and almost no depth information in the three-dimensional space, the parallax angle is used to express the different types of feature points such as the feature points that are observed for only one time. The basic idea of the hybrid feature parameterization in the new BA model is to express the location of spatial 3D feature points using the azimuth angle, the elevation angle and the parallax angle through combining two camera centres as anchor, whereas the GCPs are still expressed by XYZ. As the representation of the GCPs is normal like in the classic BA, it will not be repeatedly introduced here. The following is the process of the representation in the form of angles.

Suppose that the feature point is observed only once, and the camera centre which first observes the feature point is defined as the main anchor denoted as t_m therefore, the feature could be expressed as:

$$F_j = [\psi_j \quad \theta_j]^T \quad (1)$$

where F_j = feature point

ψ_j, θ_j = azimuth angle, altitude angle

$[\psi_j, \theta_j]$ = vector direction the main anchor t_m to the feature point F_j in the P_0 global coordinate (shown as Figure 1)

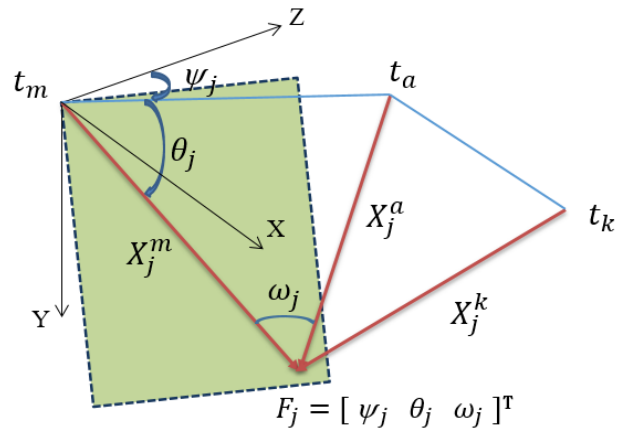


Figure 1. The parameterization of 3D feature points based on the parallax angle.

For the parameterization of the first observed feature point, there is no depth information. If the feature point F_j is observed for twice or more than twice, two camera centres of them are selected as the anchors, wherein one is the main anchor denoted as t_m , and the other is secondary anchor denoted as t_a . Therefore, the feature point F_j in 3D can be expressed by azimuth, elevation angle and parallax angle ω_j :

$$F_j = [\psi_j \quad \theta_j \quad \omega_j]^T \quad (2)$$

where ω_j = parallax angle

Compared with the feature point observed for once, the additional parallax angle ω_j represents the angle from vector x_j^m to vector x_j^a . x_j^m and x_j^a are the vector from the main anchor t_m to feature point F_j , and that from the secondary anchor t_a to feature point F_j , respectively.

Supposing that the main anchor and secondary anchor of the feature point in 3D are t_m and t_a , respectively, the coordinate of the image point of the feature point F_j under the camera P_i can be expressed as:

$$\begin{bmatrix} u_j^i \\ v_j^i \end{bmatrix} = \begin{bmatrix} \frac{x_j^i}{t_j^i} \\ \frac{y_j^i}{t_j^i} \end{bmatrix} \quad (3)$$

where u, v = image coordinate

In this formula,

$$\begin{bmatrix} x_j^i \\ y_j^i \\ t_j^i \end{bmatrix} = \begin{cases} KR_m \mu_j^m, & \text{if } i = m \\ KR_i \tilde{x}_j^m, & \text{if } i \neq m \end{cases} \quad (4)$$

where x, y, t = homogeneous image point coordinate
 K = intrinsic parameter matrix

$$R_i = r(\alpha_i, \beta_i, \gamma_i) \quad (5)$$

where R_i = rotation matrix of P_i , which represents the function of the Euler angle $[\alpha_i \beta_i \gamma_i]^T$

The vector u_j^m is the unit vector from the main anchor t_m to the feature point F_j , which can be calculated by the following formula:

$$x_j^m = v(\psi_j, \theta_j) = \begin{bmatrix} \sin \psi_j \cos \theta_j \\ \sin \theta_j \\ \cos \psi_j \cos \theta_j \end{bmatrix} \quad (6)$$

$\tilde{x}_j^i (i \neq m)$ represents the standardized vector from the camera P_i to the feature point F_j , which can be calculated by the following formula:

$$\tilde{x}_j^i = \sin(\omega_j + \varphi_j) \frac{\|t_a - t_m\| \mu_j^m - \sin \omega_j (t_i - t_m)}{\|t_a - t_m\|} \quad (7)$$

where φ_j = angle from vector $t_a - t_m$ to the vector u_j^m

Therefore, φ_j can be calculated by the inner product of $t_a - t_m$ and u_j^m :

$$\varphi_j = \arccos(\mu_j^m \cdot \frac{t_a - t_m}{\|t_a - t_m\|}) \quad (8)$$

In conclusion, the eq. (4) to (8) are the observation equations of the least-square optimization of the new BA model.

2.2 Solution of the Least-square Optimization

The essence of the BA is just the least-square optimization, which can be in the function below:

$$\|\mathcal{E}\|_{Q_z^{-1}}^2 = (z - f(X))^T Q_z^{-1} (z - f(X)) \quad (9)$$

where $f(X)$ = observation function (formula (4))

$X = [\alpha, \beta, \gamma, X_s, Y_s, Z_s, \phi, \theta, \omega]$, which contains all variables that need to be optimized

$z = [u, v]$, which is the observation variable to be optimized

Q_z^{-1} = weight of the feature point

When given the initial values of all the variables X_0 , the solution of the optimization can be finally transformed to the solution of the following formula (Zhao et al., 2015):

$$J^T Q_z^{-1} J \Delta = J^T Q_z^{-1} (z - f(X_0)) \quad (10)$$

Where J = first derivative of the observation values with respect to all the unknown variables

To solve the nonlinear equation (Formula (10)), the Gauss-Newton method and the Levenberg-Marquart method are widely used in photogrammetry (Hartley and Zisserman, 2003). In this paper, these two solutions are also applied to verify the convergence performance of the new BA model. Since the applications and the performance of the BA model are focused and payed more attention here, the details and some other parts of the method are not introduced, and more can be referred from Zhao and Yan's work (Yan et al., 2017; Zhao et al., 2015).

3. EXPERIMENT AND RESULT

To discuss the applications (especially in aerial photogrammetry) and analyse the performance of the new BA model, two datasets are processed in the experiment. The first dataset is a group of aerial images -- the Toronto dataset -- released by the International Society for Photogrammetry and Remote Sensing (ISPRS) and the other one is a UAV dataset obtained at a flying altitude of 138.632 meters, covering an area of 26245.5 square meters.

3.1 The Result of Toronto Dataset

The Toronto dataset contains 13 images captured by the UCD cameras, 139648 three-dimensional feature points and 297097 projected points. To better analyse the performance of the ParallaxBA model based on the parallax angle with GN and LM, the performance of G2O (Grisetti et al., 2011) and sSBA (Konolige and Garage, 2010), which are known as two efficient BA software packages, are compared. Since sSBA only offers the LM module, so only the LM process of it is compared. Note that the LM algorithm is realized in different methods in these three software packages, and the one in ParallaxBA is the same as the sSBA. In addition, the performance of the same software packages differs when running in Windows and Linux, therefore the result in two systems are compared. Since the G2O and sSBA show very low efficiency in Windows, their result in that system are not presented.

The MSE, the iterations, the number of functions and the run time are shown in Table 1, which are produced in the processing of the data with different software packages given the same initial values. Figure 2 shows the MSE curve of each iteration. The BA in G2O with GN cannot converge because of the singularity of the normal equation, while the MSE reaches 153.13 after 200 iterations in G2O with LM. In sSBA and ParallaxBA, the MSEs all reach 0.048656, but it needs 64 iterations in the former when only 8 and 20 iterations in the other. The ParallaxBA packages are more efficient as well, because their run time are far less than the others. Figure 3 shows the result of ParallaxBA -- the reconstructed 3D features (blue points) and the camera centres (triangular cones).

		G2O GN	G2O LM	sSBA	ParallaxBA GN	ParallaxBA LM
Initial MSE		2991803.22	2991803.22	2991803.22	2991803.27	2991803.27
Convergence MSE		N/A	153.13	0.048656	0.048656	0.048656
iterations		N/A	200	64	8	20
Total time (second)	Win	N/A	N/A	N/A	2.64	5.90
	Linux	N/A	66.9	17.6	3.36	8.4

Table 1. The convergence of the Toronto dataset in G2O, sSBA and ParallaxBA

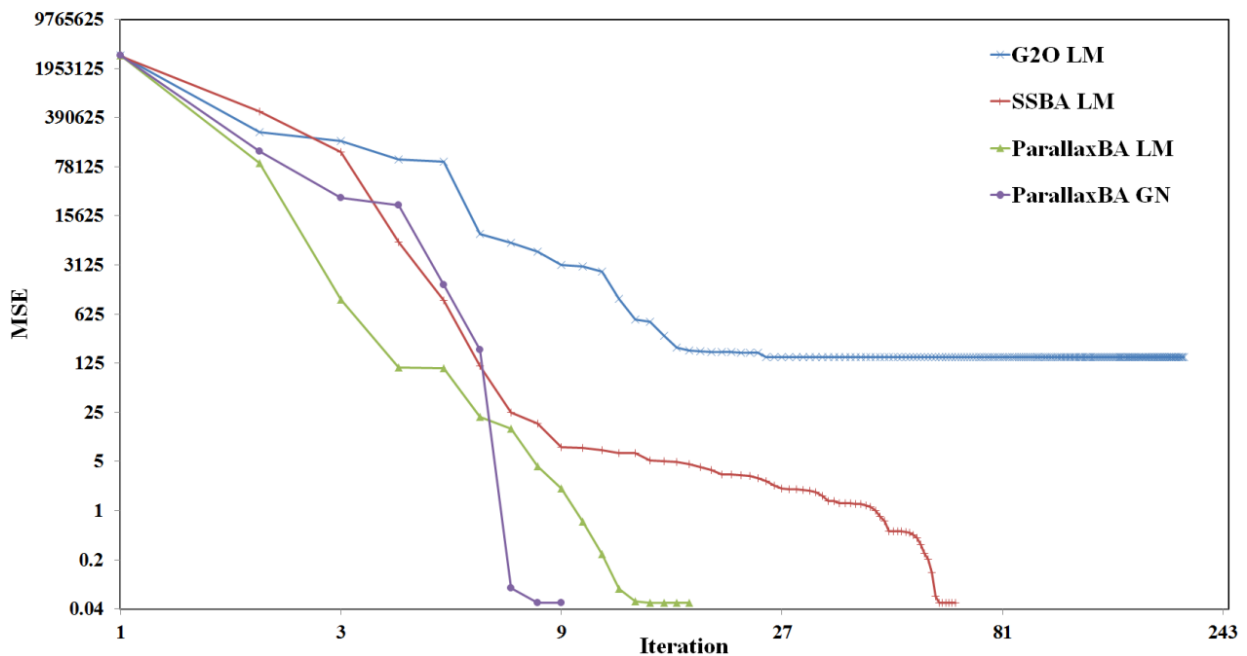


Figure 2. The MSE curve of each iteration



Figure 3. The reconstructed 3D features (blue points) and the camera centres (triangular cones)

3.2 The Result of UAV Dataset

The UAV dataset contains 40 images captured by the DMC-GH4 cameras, whose focal length is 3926.65 pixels. The image resolution is 4608 x 2592 and the ground resolution is 0.0241902

m/pixel. In the processing, 14891 tie points and 71907 projections are acquired by feature extraction and matching methods. After bundle adjustment for these parameters, the final RMSE of the reprojection image points converges to 0.561714 pixels from the initial 9.57711 pixels within only 0.426 seconds.

The ortho-rectification and mosaic of the UAV images based on the BA results -- the optimized camera parameters and the feature points -- is shown in Figure 4. Most structures in the stitched image show good shapes in accordance with the real conditions, such as the straight road edge and the building edges, though few distortions exist, especially around the boundary of the stitched image.



Figure 4. The ortho-rectification and mosaic result of the UAV dataset

All the above results have validated the effectiveness and the fast processing speed of the ParallaxBA model. For the good performance of the new BA model, it shows promising future in applications in aerial photogrammetry.

4. DISCUSSION

The divergence and the convergence speed problem has been studied for long in aerial triangulation, thus the convergence performance of the ParallaxBA model is particularly discussed in this section. The convergence of BA greatly depends on the initial value (Mikhail et al., 2001). To analyse the convergence of the ParallaxBA model in a more comprehensive way, simulations are used here to conduct BA under different initial values. The normal BA (BA based on the Euclidean coordinate system) is also conducted for comparison.

The simulated data is obtained under aerial photography condition, which contains 4 flight strips and 90 images (7680 x 13824 in size, without any geometric distortion) in total. The simulated movement trajectory is shown in Figure 5. The camera angle elements and line elements can be calculated through relative orientation, which are assumed as the optimal values. The Gauss noise under different levels is then added to the optimal values to simulate different initial values.

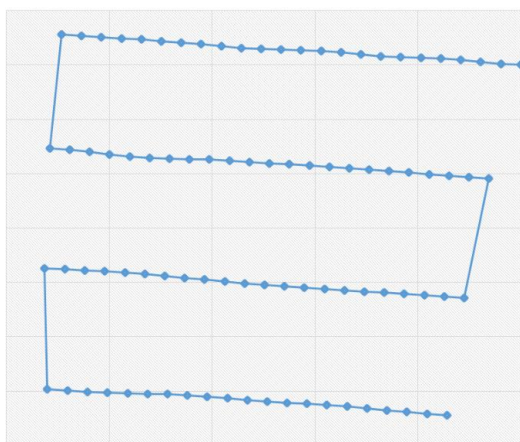


Figure 5. The simulated Movement trajectory

4.1 Convergence under Different Initial Values of Camera Angle Elements

When given the same initial MSE, add the Gauss noise of 0.03, 0.05, 0.08, 0.10, 0.13 and 0.15 degrees to the initial values of camera angle elements and convergence performance is shown in Table 2 and Figure 6. Obviously, the two kinds of BA converge to the same value when the error level (noise) is 0.03 degrees. However, as the error level keeps on increasing, the ParallaxBA model converges to a smaller MES than the normal BA when reaching the largest iteration.

The error level	BA model	Initial MSE	Convergence MSE	Iteration
0.03	Normal BA	116325.924	0.083715487	41
	ParallaxBA	116323.691	0.08371552	7
0.05	Normal BA	288298.4	20.3781	300
	ParallaxBA	288298.9	0.083716	7
0.08	Normal BA	885426.4	144.9745	300
	ParallaxBA	885427.9	0.083716	7
0.10	Normal BA	1585482	323.7692	300
	ParallaxBA	1585484	0.083716	7
0.13	Normal BA	3319862	238.5598	300
	ParallaxBA	3319862	0.083716	7
0.15	Normal BA	5966617	206.4639	300
	ParallaxBA	5966583	0.083716	11

Table 2. The convergence of the ParallaxBA model and the normal BA model given different initial values of camera angle elements

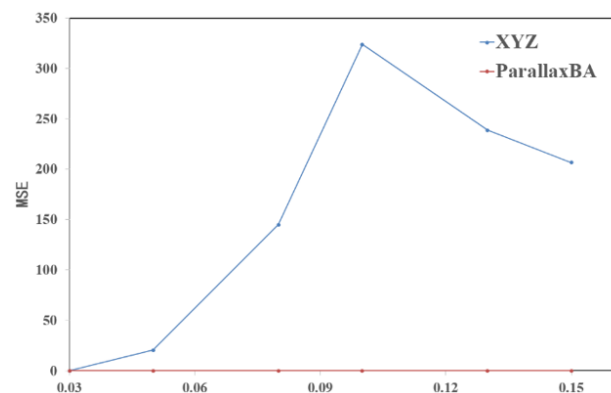


Figure 6. The converging MSE of the ParallaxBA model and the normal BA model (XYZ) given different initial values of camera angle elements

4.2 Convergence under Different Initial Values of Camera Line Elements

Assume that the distance between the cameras is unit one, then add the Gauss noise of 0.1, 0.2, 0.3, 0.4 and 0.15 to the initial values of camera line elements. The convergence performance is shown in Table 3 and Figure 7 when given the same initial MSE. Obviously, the two kinds of BA converge to the same value when the error level (noise) is 0.3. However, as the error level keeps on increasing, the ParallaxBA model converges to a smaller MES than the normal BA when reaching the largest iteration.

The error level	BA model	Initial MSE	Convergence MSE	Iteration
0.1	Normal BA	49018.209	0.08371549	29
	ParallaxBA	49018.426	0.0837155	7
0.2	Normal BA	113900.8	0.083715	31
	ParallaxBA	113900.9	0.083716	7
0.3	Normal BA	225860.4	0.083715	44
	ParallaxBA	225859.6	0.083716	9
0.4	Normal BA	391300	15.66996	300
	ParallaxBA	391298.9	0.083716	9
0.5	Normal BA	613232.2	42.72522	300
	ParallaxBA	613229.9	0.083716	10

Table 3. The convergence of the ParallaxBA model and the normal BA model given different initial values of camera line elements

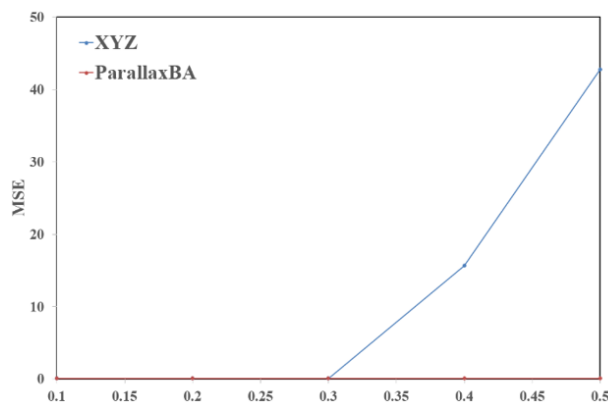


Figure 7. The converging MSE of the ParallaxBA model and the normal BA model (XYZ) given different initial values of camera line elements

5. CONSLUSION

Traditional least-square bundle adjustment methods of aerial triangulation are established based on the Euclidean coordinate system. A large number of experiments and studies have shown that there exist strong correlations between the unknowns in traditional models, which can lead to poor convergence and slow convergence speed. However, the new BA model introduced in this paper show promising results in photogrammetry especially in the aerial triangulation. The experiment results have illustrated the good performance of the BA model, whose MSE reaches 0.048656 after just 8 (using GN) and 20 (using LM) iterations. The simulation results under different initial values have also proved the stable convergence performance of the introduced BA model, in which its MSEs after iterations keep the same small, whereas the normal BA ones become larger as the initial error increases. Therefore, the performance of the new BA model is better than the normal one, and the new BA model can be applied more in the aerial photogrammetry and be further studied.

ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China (Grant Nos. 41571432).

REFERENCES

- Ackermann, F., 1984. Utilization of navigation data for aerial triangulation. International Archives of Photogrammetry and Remote Sensing, ISPRS Commission III 25, 1-9.
- Chaturvedi, A., 2000. Accuracy assessment of DTM data: a cost effective approach for a large scale digital mapping project. International Archives of Photogrammetry and Remote Sensing 33, 105-111.
- Grisetti, G., Kümmerle, R., Strasdat, H., Konolige, K., 2011. g2o: A general Framework for (Hyper) Graph Optimization, IEEE international conference on robotics and automation (ICRA), pp. 3607-3613.
- Hartley, R., Zisserman, A., 2003. Multiple view geometry in computer vision. Cambridge university press.
- Konolige, K., Garage, W., 2010. Sparse Sparse Bundle Adjustment, BMVC. Citeseer, pp. 102.101-102.111.
- Mikhail, E.M., Bethel, J.S., McGlone, J.C., 2001. Introduction to modern photogrammetry. New York.
- Moré, J.J., 1978. The Levenberg-Marquardt algorithm: implementation and theory, Numerical analysis. Springer, pp. 105-116.
- Sauerbier, M., 2004. Accuracy of automated aerotriangulation and DTM generation for low textured imagery. ETH, Eidgenössische Technische Hochschule Zürich, Institut für Geodäsie und Photogrammetrie.
- Wu, C., Agarwal, S., Curless, B., Seitz, S.M., 2011. Multicore bundle adjustment, Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, pp. 3057-3064.
- Yan, L., Chen, R., Sun, H., Sun, Y., Liu, L., Wang, Q., 2017. A novel bundle adjustment method with additional ground control point constraint. Remote Sensing Letters 8, 68-77.
- Yuan, X., 2009. Quality assessment for GPS-supported bundle block adjustment based on aerial digital frame imagery. The Photogrammetric Record 24, 139-156.
- Yuan, X., Fu, J., Sun, H., Toth, C., 2009. The application of GPS precise point positioning technology in aerial triangulation. ISPRS Journal of Photogrammetry and Remote Sensing 64, 541-550.
- Zhao, L., Huang, S., Sun, Y., Yan, L., Dissanayake, G., 2015. ParallaxBA: bundle adjustment using parallax angle feature parametrization. The International Journal of Robotics Research 34, 493-516.