Multistrip Bundle Block Adjustment of ZY-3 Satellite Imagery by Rigorous Sensor Model Without Ground Control Point

Yongjun Zhang, Maoteng Zheng, Xiaodong Xiong, and Jinxin Xiong

Abstract—Extensive applications of Zi-Yuan 3 (ZY-3) satellite imagery of China have commenced since its on-orbit test was finished. Most of the data are processed scene by scene with a few ground control points (GCPs) for each scene; this conventional method is mature and widely used all over the world. However, very little work has focused on its application to super large area blocks. This letter aims to study mapping applications without GCPs for a super large area, which is defined as a range of interprovincial or even nationwide areas. The automatic matching and bundle block adjustment (BBA) software developed by our research team are applied to deal with two blocks of ZY-3 three-line-camera imagery which covers most of the provinces in eastern China. Our comparison analysis of different data processing methods and the geolocation accuracies of the overlapped areas between adjacent strips are presented in this letter, as well as the possibility of nationwide BBA. The preliminary test results show that multistrip BBA without GCPs can achieve accuracies of about 13–15 m in both planimetry and height, which means that nationwide BBA is considered practical and feasible.

Index Terms—Ground control point (GCP), multistrip data, nationwide bundle block adjustment (BBA), three-line camera (TLC), ZY-3.

I. INTRODUCTION

A large number of high-resolution optical satellites have been launched in recent years all over the world. Zi-Yuan 3 (ZY-3) is the first civilian mapping satellite capable of stereo observation in China, utilizing an onboard three-line camera (TLC) which consists of three independent cameras looking forward, nadir, and backward, respectively. The forward and backward cameras are tilted at ±22° with respect to the nadir camera, thereby forming an ultimate 0.8 base-to-height ratio. The ground sampling distance (GSD) of the forward and backward cameras is about 3.2 m, and the nadir camera is about 2.1 m. The direct georeferencing accuracy is about 25 m after calibration by a multistrip on-orbit calibration method [1]. Using the single-scene orientation method, at least four ground control points (GCPs) are needed for orientation of each scene to achieve an acceptable accuracy. In addition, the orientation accuracies of these scenes are usually different because of the accuracies of the GCPs and their distributions, and errors therefore occur in the mosaicked image by adjacent scenes. To deal with a very large area block, a large number of GCPs are needed, which would require a great deal of manpower and financial resources. It is also difficult to acquire field control points in some inaccessible areas because of the harsh environment. The number of GCPs is decreased by using bundle block adjustment (BBA) with these small scenes. Yet, the BBA with full strips cannot only decrease the number of GCPs (even no GCP is needed) but also the number of unknowns. This method is quite suitable for orientation in super large area such as interprovincial area or even the whole China. In this letter, two large blocks consisting of ZY-3 multistrips imagery are investigated. High-precision orbit observation data are integrated with a rigorous orientation model to conduct BBA without GCPs. The proposed method does not require manpower for GCP field surveying and GCP identification (ID) on images. The method is also fully automatic since no GCPs are needed, which can significantly improve the processing efficiency.

II. RELATED WORKS

A spaceborne three-line sensor was first applied in the German Modular Optoelectronic Multispectral Stereo Scanner (MOMS-02). The orientation image model proposed by Hoffman in 1982 was adopted in the BBA of the MOMS-02 imagery, the camera focal length, and the charge-coupled device (CCD) translation, and rotation parameters also were introduced in this model. The test block was an area of 340 km × 37 km with a few evenly distributed GCPs. Accuracies of about 1.0 GSD in both planimetry and height were reported [2]. The accuracies were further improved to 12 m in planimetry and 7 m in height after orbit mechanic conditions were introduced in the BBA. Accuracies of 8 m in planimetry and 10 m in height were achieved in MOMS-2P, which was carried on the Russia Space Station. The area of the test block was about 50 km × 105 km [3].

The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) designed by Japan is carried on Terra, which was developed by the National Aeronautics and Space Administration and launched in 1998. ASTER is capable of stereo observation with a base-to-height ratio of 0.6, and the GSD is about 15 m. Dowman and Michials introduced Kepler orbit constraints into a strict physical model with some GCPs and eventually achieved an accuracy of about 1.0 pixel [4],...
accuracies were about 10–20 m in planimetry and 5–10 m in direct georeferencing with HRS imagery, and the positioning accuracies were about 10–20 m in planimetry and 5–10 m in height [7]. Korns et al. introduced a self-calibration strategy into BBA by using a three-order polynomial with a few GCPs to compensate for errors in both image space and object space and achieved an overall accuracy of about 2 m in both planimetry and height [8]. Toutin achieved an accuracy of 0.5 GSD with a proper number of GCPs, and the accuracy of the produced digital surface model was about 5.2 m [9]. Bouillon et al. achieved accuracies of 8.4 m in planimetry and 4.5 m in height without GCPs after a long-term observation period and calibration with sufficient test fields located all over the world [10]. Massera et al. included 180,607 images and about 3000 GCPs in a big block in an attempt to conduct a global BBA using SPOT5 HRS imagery; the whole process required about 3.5 h, but the overall accuracy was not clearly reported [11].

The Panchromatic Remote-sensing Instrument for Stereo Mapping (PRISM) sensor on the Advanced Land Observing Satellite (ALOS), developed by Japan and launched in 2006, is also a three-line push broom sensor and is capable of stereo observation. It consists of three independent cameras looking at forward (tilted at 23.8°), nadir, and backward (tilted at −23.8°), respectively, and the GSD is about 2.5 m [12]. Kocaman and Gruen adopted a piecewise polynomial model to represent the exterior orientation parameters (EOPs) while a total of 30 additional parameters were introduced to compensate for the interior orientation parameters (IOPs) errors of PRISM. After correction with a few GCPs, accuracies of 0.5 GSD in planimetry and 0.3 GSD in height were obtained [13]. Weser et al. applied a third-order spline to represent the orbit and attitude observations, and six translation parameters were introduced to compensate for the orbit and attitude errors. After BBA with evenly distributed GCPs, an accuracy at subpixel level was obtained [14]. Similar to Kocaman and Gruen, Rottensteiner et al. performed BBA based on a strip concept. The unknown parameters of the scenes belonging to the same strip were combined to decrease the redundant parameters, and the required number of GCPs was reduced accordingly [15]. This research team had finally completed the entire Australian continent mosaic [23].

Lutes adopted the rational function model (RFM) to deal with CARTOSAT-1 imagery, and five GCPs were used to perform BBA based on an RFM affine transformation model. The positioning accuracies were reported to be about 2–4 m in planimetry and 5–10 m in height [16]. Baltisavias et al. used the same method and achieved an accuracy of about 0.5 GSD in both planimetry and height [17]. Radhadevi adopted a rigorous model and introduced polynomials to compensate for EOP and IOP errors and reported 3–4 m accuracy while using only one GCP [18]. Michalis and Dowman introduced the coplanarity condition into BBA with orbit constraints, which finally enabled achieving accuracy at subpixel level [19].

Super high-resolution satellites, such as GeoEye-1/2, WorldView-1/2, QuickBird, and IKONOS, have been launched in recent years. These satellites often have higher resolution than the above motioned satellites and are also capable of stereo observation. Toutin et al. adopted a hybrid model that combined RFM and a rigorous physical model to process a stereo image pair of WorldView-1/2 and achieved positioning accuracies of about 2.6 m in planimetry and 2.1 m in height [20]. Eisenbeiss et al. used a few GCPs to conduct BBA based on an RFM affine transformation model with QuickBird and IKONOS imagery and achieved accuracies of 0.5 m in planimetry and 0.5 m in height in bold terrain and 1.0 m in height in urban and mountainous areas [21].

All methods dealing with high-resolution satellite imagery mentioned earlier were tested in small areas except for the global adjustment with SPOT5 imagery, but this global adjustment applied a large number of GCPs distributed almost all over the world [11]; for the entire Australia continent mosaic through long-strip ALOS PRISM orientation, this technique has reduced the ground control requirements, but a few GCPs are still needed [15], [23]. In this letter, a rigorous physical model was used to deal with the multistrip ZY-3 three-line imagery. A large block covering several provinces in China was established to perform BBA without any GCP. The positioning accuracy and nationwide BBA ability will be discussed in the following section.

III. MATHEMATICAL MODEL

In order to generate photogrammetric products, establishing the mathematical relationship between the image space and object space is needed. Most of the sensor models dealing with satellite imagery can be classified into two types: rigorous physical models and general models. A general model actually uses a rational function to describe the relationship between image space and object space, such as RFM. This model is very fast and convenient and can efficiently hide the exact orbit information; thus, most of the commercial satellites release their 1A/1B products with a rational polynomial coefficient (RPC) instead of orbit and attitude data. However, RPC is only suitable for small areas like 100 km × 100 km. In a rigorous physical model, EOPs represent the position and attitude of the sensor; and the collinearity equation is the basic mathematical model. It is a relatively strict sensor model [22]. This model is not restricted by the size of the block area, and additional parameters can be easily introduced into it. It is often applied in on-orbit calibration, bundle adjustment of large areas, and sensor geolocation accuracy analysis. In this letter, a rigorous physical model is adopted to reconstruct the geometric relationship between image space and object space. BBA of a large block area of ZY-3 TLC imagery is performed and assessed. The feasibility of nationwide BBA is also discussed.

A. Exterior Orientation Model

For line array imagery, the mathematical relationship between the image point coordinates in an image space and the corresponding object point coordinates in an object space can be expressed by the following equation:

\[ P_{\text{img}} + \delta x = \lambda \cdot R_{cs}^T \cdot \left[ R_{s}^T (t) \cdot R_{\text{out}}^T \cdot [P_{\text{obj}} - S(t)] + \Delta L_{cs} \right] \]  
\[ P_{\text{img}} = (x, y, -f)^T \]  
\[ \delta x = (x_c, y_c, \Delta f)^T \]
where $P_{\text{img}}$ is the image point coordinates; $x$ denotes the image point coordinates along track; $y$ denotes the image point coordinates across track; $f$ is the focal length; $\delta_x$ denotes the correction parameters of the image point coordinates; $x_c$ denotes the correction parameters of the image point coordinates along track; $y_c$ denotes the correction parameters of the collinearity equation for the image point coordinates across track; $\Delta f$ denotes the correction parameters of the focal length; $\lambda$ is the scale factor; $R_{cs}$ is the rotation matrix from the camera coordinates to the satellite coordinate system; $R_{ow}(t)$ is the rotation matrix from the satellite coordinate system to the orbit coordinate system, which consists of three attitude angles of Roll, Pitch, Yaw; $R_{ow}$ is the rotation matrix from the orbit coordinate system to the geocentric coordinate system (here, referred to as WGS84); $P_{\text{obj}}$ denotes the object coordinates in the WGS84 coordinate system; $S(t)$ denotes the satellite coordinates in the WGS84 coordinate system; $t$ is the imaging time; and $\Delta L_{cs}$ is the vector of the translation parameters from the camera projection center to the satellite center.

A trajectory model is applied to represent the changing of the satellite’s position and attitude while scanning. A systematic error compensation model is adopted in this letter and is expressed as follows:

$$
S_X(t) = S_X^0(t) + \Delta S_X
$$
$$
S_Y(t) = S_Y^0(t) + \Delta S_Y
$$
$$
S_Z(t) = S_Z^0(t) + \Delta S_Z
$$
$$
\text{Roll}(t) = \text{Roll}^0(t) + \Delta \text{Roll}
$$
$$
\text{Pitch}(t) = \text{Pitch}^0(t) + \Delta \text{Pitch}
$$
$$
\text{Yaw}(t) = \text{Yaw}^0(t) + \Delta \text{Yaw}
$$

(4)

where $S_X(t)$, $S_Y(t)$, $S_Z(t)$, Roll$(t)$, Pitch$(t)$, and Yaw$(t)$ are the sensor position and attitude parameters; $S_X^0(t)$, $S_Y^0(t)$, $S_Z^0(t)$, Roll$^0(t)$, Pitch$^0(t)$, and Yaw$^0(t)$ are the initial values of the sensor position and attitude derived by the observations of the sensor position and attitude; $\Delta S_X$, $\Delta S_Y$, $\Delta S_Z$, $\Delta \text{Roll}$, $\Delta \text{Pitch}$, and $\Delta \text{Yaw}$ are the compensation parameters of the observations of the sensor position and attitude; and $t$ is the corresponding imaging time.

The initial values of the sensor position and attitude were preprocessed before the experiment. The postprocessed sensor position data were provided by other Global Positioning System (GPS) research teams, where the position data were processed by a differential positioning approach using the original carrier phase observations of onboard GPS receivers and base stations on the ground. The accuracy of these postprocessed GPS data was claimed to be about 0.5 m and can be assumed as the true value of the sensor position [1]. The attitude data are transmitted along with the imagery from space; thus, only three attitude compensation parameters ($\Delta \text{Roll}$, $\Delta \text{Pitch}$, and $\Delta \text{Yaw}$) were set as unknowns in each strip during BBA.

### B. Interior Orientation Model

According to [1], correction parameters $x_c$ (along track) and $y_c$ (across track) can be expressed by the following equations:

$$
x_c = \begin{cases} 
0 < c < N_1 \\
N_1 < c < N_1 + N_2 \\
N_1 + N_2 < c < N \\
N + c & \text{if } y_0 + y_1 c > 0
\end{cases}
$$

$$
y_c = \begin{cases} 
0 < c < N_1 \\
N_1 < c < N_1 + N_2 \\
N_1 + N_2 < c < N \\
N + c & \text{if } y_0 + y_1 c > 0
\end{cases}
$$

(5)

where $x_0$, $x_1$, and $x_2$ are the translation errors along track of the three sub-CCD arrays in the nadir camera; $y_0$, $y_1$, $y_2$, $y_3$, $y_4$, and $y_5$ are the translation and scaling errors across track of the three sub-CCD arrays in the nadir camera; $c$ is the CCD column ID in the whole CCD array; $N_1$ and $N_2$ are the actual pixel numbers captured by the first and second CCD arrays, respectively, in the whole mosaic image as shown in Fig. 1; $N$ is the pixel number of the whole image captured by the nadir camera; and $c$ is the pixel column ID in the whole image. More information can be found in [1].

The calibrated parameters in previous work [1] are directly used in this letter, and only the EOPs are introduced as unknowns to perform BBA without any GCP.

### IV. Experiments and Analysis

#### A. Test Block 1

Test block 1 consists of eight strips of ZY-3 TLC imagery, and the length of each strip is about 3000 km. The ground coverage is shown in Fig. 2. A total number of 139 field points were identified as check points, mainly distributed in the middle part of the block. The accuracy of these points is at the centimeter level. Automatic image matching software and BBA software named DPGrid.Sat, which was developed by our research team at Wuhan University, were used to extract tie points and conduct BBA, respectively. A total number of 2.26 million tie points were extracted in this test block.

#### B. Test Block 2

Test block 2 consists of five strips of ZY-3 TLC imagery, and the ground coverage is shown in Fig. 2. A total number of 50 field points were identified to check the accuracy of BBA without GCPs. The check points were distributed in the south and north of this block. A total number of 60800 tie points were extracted with DPGrid.Sat. The tie point extraction method can be found in [24].

#### C. Auxiliary Data

The orbit observation data were postprocessed using the original carrier phase observations of onboard GPS receivers and base stations on the ground. Attitude data were measured by a star tracker and transmitted with the imagery from the satellite.

#### D. Weight Strategy

The weights of the observations of position and attitude were not fixed during adjustment. The initial weights of the image points were set at 1.0. In this letter, there are no other observations, such as GPS, attitude, or additional parameters are used.
In case such observations are available, they would be weighted by the square of the ratio of its accuracy against the accuracy of the image point observations. After iteration during adjustment, the weight of each image point was recalculated according to the new image point observation weights. The weights of the compensation parameters of the sensor attitude in (4) were set at a large value to ensure that the estimated compensation parameters of sensor attitude would not result in unreasonably large values that would certainly compromise the level of accuracy.

E. Experiment

The results of direct forward intersection and BBA without GCPs are compared in this section.

1) Tests in Block 1: As can be seen in Table I, the direct intersection accuracies were about 25 m in planimetry and 9.0 m in height, which are in accordance with [1]. The root-mean-square error (rmse) of the image point residuals, which represents the internal accuracy, was improved from 1.0 ~ 2.0 pixels to better than 1.0 pixel after BBA without GCPs. The planar intersection accuracies were about 25 m in planimetry and 9.0 m in height, which are in accordance with [1]. The root-mean-square error (rmse) of the image point residuals, which represents the internal accuracy, was improved from 1.0 ~ 2.0 pixels to better than 1.0 pixel. The height accuracy was improved from about 20.6 m to about 13 m while to better than 1.0 pixel after BBA without GCPs. The planar intersection accuracies were about 25 m in planimetry and 9.0 m in height, which are in accordance with [1]. The root-mean-square error (rmse) of the image point residuals, which represents the internal accuracy, was improved from 1.0 ~ 2.0 pixels to better than 1.0 pixel.

2) Tests in Block 2: After BBA, the internal accuracy improved from 2.0 – 3.0 pixels to better than 1.0 pixel. The external accuracy also improved by 10 – 20 m in plane. The height of the rmse of the check point residuals increased slightly, which is in agreement with the results in test block 1. These check points were distributed in the south and north parts of the block area, respectively. The accuracy levels of north and south areas, shown in Table II, are very close to each other and indicate that the accuracies in the whole block are stable.

3) Discrepancies Between Adjacent Strips in the Overlapped Area: The tie points which had multistrip overlap were detached into subtie points. If tie point P had image points in three different strips, for example, then the tie point was split into three subtie points with each of them having image points at only one strip. These subtie points were projected to the object space, and three object points $P_1(X_1, Y_1, H_1)$, $P_2(X_2, Y_2, H_2)$, and $P_3(X_3, Y_3, H_3)$ were obtained independently. Then, the discrepancies $\Delta X, \Delta Y, \Delta H$ were calculated by the following equations:

\[
\begin{align*}
\Delta X &= \sqrt{\frac{(X_1 - X_2)^2 + (X_1 - X_3)^2 + (X_2 - X_3)^2}{3}} \\
\Delta Y &= \sqrt{\frac{(Y_1 - Y_2)^2 + (Y_1 - Y_3)^2 + (Y_2 - Y_3)^2}{3}} \\
\Delta H &= \sqrt{\frac{(H_1 - H_2)^2 + (H_1 - H_3)^2 + (H_2 - H_3)^2}{3}}.
\end{align*}
\]

As demonstrated in Table III, BBA without GCPs had smaller discrepancies than direct intersection, particularly in the height direction.

4) Ability of Data Processing for Super Large Areas: The aforementioned experiments were performed in Windows XP 32-b operating system, with RAM of about 1 GB and hard drive memory of about 500 GB. As the number of strips became larger, more memory was required. It was found that at least 200 strips were needed to form a super large area that could fully cover the whole territory of China. So, the massive memory issue caused by such large areas was taken into consideration. The predicted memory requirements of nationwide BBA experiments applying the aforementioned model are listed in Table IV.

TABLE I

<table>
<thead>
<tr>
<th>Strategies</th>
<th>RMSE of Image point (pixel)</th>
<th>RMSE of Check point residuals (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct intersection</td>
<td>1.683</td>
<td>1.262</td>
</tr>
<tr>
<td>BBA without GCP</td>
<td>0.338</td>
<td>0.709</td>
</tr>
</tbody>
</table>

TABLE II

<table>
<thead>
<tr>
<th>Strategies</th>
<th>RMSE of image point (pixel)</th>
<th>Check Point Area</th>
<th>RMSE of check point residuals (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct intersection</td>
<td>1.784</td>
<td>North</td>
<td>23.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>South</td>
<td>34.9</td>
</tr>
<tr>
<td>BBA without GCP</td>
<td>0.622</td>
<td>North</td>
<td>10.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>South</td>
<td>11.0</td>
</tr>
</tbody>
</table>

TABLE III

<table>
<thead>
<tr>
<th>Strategy</th>
<th>$\Delta X$ (m)</th>
<th>$\Delta Y$ (m)</th>
<th>$\Delta H$ (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Intersection</td>
<td>3.20</td>
<td>4.05</td>
<td>12.09</td>
</tr>
<tr>
<td>BBA without GCP</td>
<td>3.30</td>
<td>3.07</td>
<td>6.53</td>
</tr>
</tbody>
</table>
We directly formed a normal matrix without forming any error equations. As can be seen in Table IV, only 360 MB of memory is needed, which means that the nationwide BBA experiment is fully feasible and practical on a common computer.

### V. CONCLUSION

A total of 13 strips of ZY-3 TLC imagery have been used to perform BBA with the rigorous physical model in this letter. The sensor position data were already postprocessed using the original carrier phase observations of the onboard GPS receivers and base stations on the ground; thus, the translation parameters were treated as known in BBA. Direct intersection BBA without GCPs was tested, and the snap accuracy and potential ability of super large area BBA were assessed. Based on the test results shown in this letter, the following conclusions can be drawn.

The accuracies of direct intersection of ZY-3 TLC imagery are about 20–30 m in planimetry and 13–15 m in height, which is in accordance with [1].

Accuracies of about 13–15 m in both planimetry and height can be achieved by BBA without GCPs. The height accuracy has no improvement after BBA; an accurate weight determination method or a few GCPs might help to improve height accuracy, yet this still need more tests.

According to our analysis of the massive data process required, BBA of super large areas is feasible and practical.

### ACKNOWLEDGMENT

The authors would like to thank the China Centre for Resources Satellite Data and Application for providing all test data for this project.

### REFERENCES


