

Three-Dimensional Reconstruction of Multiplatform Stereo Data With Variance Component Estimation

Yanfeng Gu, *Member, IEEE*, Zhimin Cao, and Ye Zhang

Abstract—In this paper, we address a problem of 3-D reconstruction with generalized stereo data from multiple platforms of remote sensing. Nowadays, rational function model (RFM)-based 3-D reconstruction with stereo images obtained from a single platform of remote sensing like a satellite or an airborne platform has been widely investigated, but there are little attentions to be paid to the problem of 3-D reconstruction with stereo images from multiple platforms in the existing literature. In order to make full use of the generalized stereo images from different platforms with different rigorous sensor models for 3-D reconstruction, we need to form the least squares estimation model of the corresponding RFM-based forward-intersection task after collecting observations from different platforms. However, resolutions of the stereo images from different platforms are greatly different so that the observations in the corresponding least squares problem are mathematically seriously unbalanced. To solve this problem for achieving precise reconstruction, we first model how the spatial resolution of the observation images of different platforms changes pixel by pixel and then embed the variance-component-estimation technique into the RFM-based 3-D reconstruction procedure to adaptively adjust weights for different observations. Experiments are conducted on simulated and real data sets. Experimental results show that the proposed algorithm can efficiently fulfill the 3-D reconstruction task for multiplatform stereo images with noticeable improvement over the classical RFM-based 3-D reconstruction method in terms of precision.

Index Terms—Generalized stereo images, multiplatform, rational function model (RFM), variance component estimation (VCE), 3-D reconstruction.

I. INTRODUCTION

THREE-dimensional reconstruction of land covers is always one of the most important applications in the field of remote sensing. The mission of 3-D reconstruction is to restore the 3-D information of a given scene or object. To fulfill this mission, remote sensing data, particularly optical images, with as small pixel size as possible and as large scene size as possible are always required. Since the launch of the IKONOS satellite in 1999, the resolution of satellite imagery has achieved less than 1 m. For the GeoEye-2 satellite which will be launched in 2013, the resolution will even be 0.25 m. Moreover, the resolution of airborne optical images can be less than 0.05 m.

Manuscript received April 7, 2013; revised July 22, 2013; accepted August 26, 2013. Date of publication October 2, 2013; date of current version February 27, 2014. This work was supported by in part by the Natural Science Foundation of China under Grant 61371180 and in part by the Fundamental Research Funds for the Central Universities under Grant HIT.NSRIF.2010095.

The authors are with the School of Electrical and Information Engineering, Harbin Institute of Technology, Harbin 150001, China.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2013.2280381

With these high-resolution satellite images or the very high resolution airborne images, a lot of methods have been proposed for the 3-D reconstruction of the ground scene up to now. By using stereoscopic images acquired from satellites with rigorous sensor model, the digital elevation model can be generated by the classical collinearity equation and refined by using ground control points (GCPs) [1]–[9]. After the launch of the IKONOS-2 satellite, more and more satellite imagery vendors deliver rational polynomial coefficients (RPCs) of the corresponding rational function model (RFM) as the geometric parameters of the imagery to the user instead of the complex sensor model. For its sensor-independent character, the RFM in the form of the quotient of two polynomials is used widely for photogrammetric processing.

From then on, for 3-D reconstruction, more and more attentions have been paid to the RFM-based methods. Generally speaking, the development of RFM-based 3-D reconstruction methods can be divided into three different stages. The first stage of RFM-based methods began from Yang's investigation [10], in which the author presented the performance of the second-order and the three-order RFM-based 3-D reconstruction with a stereo-image pair from the SPOT5 satellite. After that, the classical forward-intersection RFM-based (C-RFM for short) 3-D reconstruction method was proposed by Di *et al.* [11]. In their paper, the IKONOS stereo-image pair was used to deal with shoreline mapping and with change detection by using the forward-intersection method. In [12], Tao and Hu discussed the RFM-based 3-D reconstruction methods in detail and also tested their method with satellite images and aerial images, respectively.

The second stage of using RFM-based 3-D reconstruction methods is to enhance the performance of the C-RFM method with the stereo-image pair from a single platform by using some known GCPs [13]–[20]. There are two ways to enhance the performance of the C-RFM method. The first way is to remove calibration errors in the RFM by exploiting accurate GCPs; the corresponding research works include [13]–[18]. The most typical research works are [16] and [17], in which Fraser and Hanley achieved the submeter level of 3-D reconstruction accuracy both in horizontal coordinates and vertical coordinates by using bias-corrected RPCs. The second way is to refine the results of the C-RFM method in object space by using known GCPs. As a typical research, Niu *et al.* [19] exploited different transformation models in both object and image spaces to improve the accuracy of results of the C-RFM method by using known GCPs.

Nowadays, on the one hand, we cannot obtain the stereo-image pair from a single platform which is required by the

existing 3-D reconstruction models in case of some special applications. On the other hand, more and more imagery sources, which describe a similar ground scene or object, are available from satellite or airborne platforms with different angles and resolutions. Therefore, it is necessary to integrate data acquired by different platforms and to integrate data obtained at different times, different spaces, different viewing angles, and different sensors. Difference from times, spaces, viewing angles, and sensors can be finally considered as the ways to obtain images with different spatial resolutions of the same or approximately same ground scene or object. Generally speaking, those images can be potentially used as generalized stereo-image pair (GSP). Here, we call this as the third stage of the RFM-based 3-D reconstruction method. This stage is involved in RFM-based multipatform 3-D reconstruction. Under these circumstances, the stereo-image pair is always composed of images with different resolutions and with different noise levels. To address this problem, there are several methods proposed in the literature. Li *et al.* [21], [22] analyzed the geopositioning accuracy of different ways to integrating IKONOS and QuickBird images for shoreline mapping. However, the reconstruction method that they used is only the C-RFM method with system error reduced by using affine transform in image space. The real solution for multipatform 3-D reconstruction is not investigated intrinsically. Hu *et al.* [20] investigated the exploitation of GSP from overlapped IKONOS and QuickBird image pairs for 3-D feature extraction. To do this, the authors assigned different weights (1 for the IKONOS image and $(1/0.6)^2$ for the QuickBird image) to the observations from different images according to the mean pixel size of the corresponding satellite images. It can be said that they had revealed the essence for multipatform 3-D reconstruction to a certain extent. However, it is unreasonable to assign a same weight to each pixel of the satellite image according to the mean resolution, particularly for the image acquired with a large pitch angle or large field of view (FOV). Moreover, the corresponding weight matrix is not adjusted with consideration of the undulated variance components during the iterative procedure of 3-D reconstruction. Furthermore, the correlation among observations from multipatform is neglected.

Generally, the key mathematical model of C-RFM 3-D reconstruction is an ordinary least squares (OLS) problem or a weighted least squares (WLS) problem. Unfortunately, for the multipatform case, the classical conditions for OLS [37] are no longer satisfied. The corresponding mathematical model is changed to a generalized least squares (GLS) one but neither OLS nor WLS. In this paper, we try to give a comprehensive analysis for this problem. Essentially, it is a problem caused by combining heterogeneous data for estimating unknown parameters. Therefore, the crucial procedure is to determine the proper or optimal components of the corresponding covariance matrix. Obviously, variance component estimation (VCE) technique is a suitable tool for dealing with this problem. For this tool, there are a lot of efforts made for improving VCE in the past decades [23]–[26]. As to the applications of VCE technique, Koch and Kusche [27] got the appropriate weight for the determination of the gravity field of the Earth by combining different types of satellite data. In [28], the

authors constructed a new orthogonal complement matrix of the coefficient matrix in the Gauss–Markov model to improve the performance of the traditional VCE method, and the improved VCE was used for the GPS stochastic model successfully. By integrating interferometric synthetic aperture radar (InSAR) and GPS measurements for 3-D surface displacement mapping [29], the VCE method was used to weight the InSAR and GPS measurements in deriving 3-D surface displacement. The corresponding experimental results with both simulated and real data sets illustrated that the proposed method in [29] can significantly enhance the accuracies of the estimated 3-D displacements.

The objective of this paper is to make an attempt to give a deep analysis to the problem of 3-D reconstruction for GSP from multipatform in the mathematic view and then to propose a general-type solution for this problem. For doing these, we choose RFM (which is a famous solution for the general sensor model) as basis for the so-called general-type solution to derive the key factor for dealing with this problem. We found that the key factor is to introduce information about the detailed pixel size of observations to embed the VCE technique in the procedure of 3-D reconstruction. Therefore, there are two different situations with different available information for the detailed pixel size analysis discussed. Finally, a general-type solution which is called generalized RFM method (G-RFM for short) is proposed.

The rest of this paper is organized as follows. The RFM sensor model and the corresponding 3-D reconstruction methods for comparison with the proposed G-RFM method were introduced in Section II. In Section III, first, the mathematic model of the proposed G-RFM method for GSP from multipatform was given, and then, detailed pixel size analysis which is the key factor for the realization of G-RFM was given. A 3-D reconstruction algorithm was presented to solve the proposed G-RFM model at the end of Section III. Experiments were conducted on simulated and real data sets, and the corresponding results were illustrated in Section IV. At last, some remarkable conclusions of this paper were drawn in Section V.

II. RFM-BASED 3-D RECONSTRUCTION

A. RFM

The RFM is a kind of general sensor model. It can be used to replace the rigorous sensor model to construct the geometric relationship between the image space and the corresponding object space. The RFM (forward or inverse) takes a form of the quotient of two polynomials as given in

$$f = \frac{F_1(X, Y, Z)}{F_2(X, Y, Z)} \quad (1)$$

where f can be either the normalized row or column of the specified image pixel or one of the normalized horizontal coordinates of the specified point in object space. Accordingly, X and Y represent the corresponding normalized horizontal geodetic coordinates or the corresponding normalized row or column value. Z represents the normalized height of the specified point in object space. F_1 and F_2 are two polynomials which

have the same form with different parameters as described in [11]–[20]

$$F(X, Y, Z) = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^p c^{r(i,j,k)} X^i Y^j Z^k$$

$$i + j + k \leq \max(m, n, p) \quad (2)$$

where $m, n, p \in \{1, 2, 3\}$; $c^{r(i,j,k)}$ are RPCs (where $r(i, j, k) = 0, 1, 2, \dots, 19$). For satellite and airborne optical sensors, (m, n, p) , the order of polynomials, is generally set to (3,3,3).

By giving a certain number (e.g., the number is 39) of GCPs and the corresponding image points, the RPCs can be calculated by using the iterative solution described in [13]. This solution is also described in Appendix I in detail with RPC00B format [38].

B. Classic RFM-Based 3-D Reconstruction

For a given scene or object, a stereo-image pair collected by a single platform can be used for the 3-D reconstruction of the scene or object by using the RFM. Let (r_l, c_l) , (r_r, c_r) be the (normalized) pixel coordinates of the (normalized) object point (P, L, H) in the stereo images. Then, by the first-order Taylor expansion, the following equation set in matrix form can be obtained:

$$\mathbf{y}_c = \mathbf{A}_c \mathbf{x} + \mathbf{b}_c \quad (3)$$

where

$$\mathbf{y}_c = \begin{bmatrix} r_l - r_{l0} \\ c_l - c_{l0} \\ r_r - r_{r0} \\ c_r - c_{r0} \end{bmatrix} \quad \mathbf{A}_c = \begin{bmatrix} \frac{\partial r_l}{\partial P} & \frac{\partial r_l}{\partial L} & \frac{\partial r_l}{\partial H} \\ \frac{\partial c_l}{\partial P} & \frac{\partial c_l}{\partial L} & \frac{\partial c_l}{\partial H} \\ \frac{\partial r_r}{\partial P} & \frac{\partial r_r}{\partial L} & \frac{\partial r_r}{\partial H} \\ \frac{\partial c_r}{\partial P} & \frac{\partial c_r}{\partial L} & \frac{\partial c_r}{\partial H} \end{bmatrix} \bigg|_{(P_0, L_0, H_0)}$$

$$\mathbf{x} = \begin{bmatrix} \Delta P \\ \Delta L \\ \Delta H \end{bmatrix} \quad \mathbf{b}_c = \begin{bmatrix} e_{lr} \\ e_{lc} \\ e_{rr} \\ e_{rc} \end{bmatrix}$$

and the subscript c represents “C-RFM”; $(r_{l0}, c_{l0}, r_{r0}, c_{r0})$ and (P_0, L_0, H_0) are the initial values of the corresponding coordinates in image space and object space, and $(e_{lr}, e_{lc}, e_{rr}, e_{rc})$ are the residual parts.

Therefore, the (weighted) least squares solution of (3) can be obtained as given in [12]

$$\bar{\mathbf{x}} = [\Delta P \ \Delta L \ \Delta H]^T = (\mathbf{A}_c^T \mathbf{W}_c \mathbf{A}_c)^{-1} \mathbf{A}_c^T \mathbf{W}_c \mathbf{b}_c \quad (4)$$

where $[\Delta P \ \Delta L \ \Delta H]^T$ are the corrections between the real object coordinates (P, L, H) and the initial coordinates (P_0, L_0, H_0) , and \mathbf{W}_c is a weight matrix which is an identity matrix here for the resolutions of the stereo-image pair which are the same.

The reconstruction process described above is under the condition that the stereo-image pair is composed of two images with the same resolution and same normalized parameters. If the normalized parameters of the stereo-image pair are not the same, the normalized parameters should be added into (3) and (4) just as described in [12].

C. Weighted RFM-Based 3-D Reconstruction (W-RFM) With GSPs

After GSP is acquired with RFMs available, the 3-D reconstruction task can be addressed. Specifically, the 3-D coordinates of a point in object space can be calculated from the corresponding pixel coordinates in GSP. Let (r_{m1}, c_{m1}) , (r_{m2}, c_{m2}) be the pixel coordinates about the unknown object point (P, L, H) in GSP. For calculating the unknown coordinates (P, L, H) , the way employed by the C-RFM method described in Section II-B can be used. Therefore, the following equation in matrix form can be obtained:

$$\mathbf{y}_m = \mathbf{A}_m \mathbf{x} + \mathbf{b}_m \quad (5)$$

where

$$\mathbf{y}_m = \begin{bmatrix} r_{m1} - r_{m1}^0 \\ c_{m1} - c_{m1}^0 \\ r_{m2} - r_{m2}^0 \\ c_{m2} - c_{m2}^0 \end{bmatrix} \quad \mathbf{A}_m = \begin{bmatrix} \frac{\partial r_{m1}}{\partial P} & \frac{\partial r_{m1}}{\partial L} & \frac{\partial r_{m1}}{\partial H} \\ \frac{\partial c_{m1}}{\partial P} & \frac{\partial c_{m1}}{\partial L} & \frac{\partial c_{m1}}{\partial H} \\ \frac{\partial r_{m2}}{\partial P} & \frac{\partial r_{m2}}{\partial L} & \frac{\partial r_{m2}}{\partial H} \\ \frac{\partial c_{m2}}{\partial P} & \frac{\partial c_{m2}}{\partial L} & \frac{\partial c_{m2}}{\partial H} \end{bmatrix} \bigg|_{(P_0, L_0, H_0)}$$

$$\mathbf{x} = \begin{bmatrix} \Delta P \\ \Delta L \\ \Delta H \end{bmatrix} \quad \mathbf{b}_m = \begin{bmatrix} e_{r_m1} \\ e_{c_m1} \\ e_{r_m2} \\ e_{c_m2} \end{bmatrix}$$

and the subscript m represents “multiplatform”; (r_{m1}^0, c_{m1}^0) , (r_{m2}^0, c_{m2}^0) are the corresponding initial values of pixel coordinates in GSP calculated by (P_0, L_0, H_0) ; $(e_{r_m1}, e_{c_m1}, e_{r_m2}, e_{c_m2})$ which are the residual parts.

Therefore, the corresponding (weighted) least squares solution will be

$$\bar{\mathbf{x}} = [\Delta P \ \Delta L \ \Delta H]^T = (\mathbf{A}_m^T \mathbf{W}_m \mathbf{A}_m)^{-1} \mathbf{A}_m^T \mathbf{W}_m \mathbf{b}_m. \quad (6)$$

It can be seen that (6) is exactly the same as (4) in terms of expression. However, the deep truths of these two equations have notable difference. The corresponding observations of (4) are collected from a single platform with even the same resolution. However, for (6), the difference of the resolutions of the corresponding observations cannot be ignored even if they are at the same resolution level. Therefore, the weight matrix \mathbf{W}_m should not to be set to the identity matrix here. Obviously, the weights for different observations need to be set according to the corresponding resolutions.

For assigning weights for different observations from different platforms according to their resolutions, the weight matrix used in [20] for integrating IKONOS (1-m average resolution) and QuickBird (0.6-m average resolution) satellite images to realize the task of 3-D reconstruction [with the IKONOS image as observation from the first platform and the QuickBird image as observation from the second platform corresponding to (5)] is

$$\mathbf{W}_m = \text{diag} \left[1 \ 1 \ \left(\frac{1}{0.6} \right)^2 \ \left(\frac{1}{0.6} \right)^2 \right]. \quad (7)$$

It can be seen that the abstracted form of the weight matrix used in [20] is

$$\mathbf{W}_m = \text{diag} \left[\left(\frac{1}{\text{mean}(\text{Res}_{m1}^r)} \right)^2 \left(\frac{1}{\text{mean}(\text{Res}_{m1}^c)} \right)^2 \right. \\ \left. \times \left(\frac{1}{\text{mean}(\text{Res}_{m2}^r)} \right)^2 \left(\frac{1}{\text{mean}(\text{Res}_{m2}^c)} \right)^2 \right] \quad (8)$$

where $\text{mean}(\text{Res}_{mi}^r)$ ($i = 1, 2$) denotes the mean row size of the i th platform, and $\text{mean}(\text{Res}_{mi}^c)$ ($i = 1, 2$) denotes the mean column size of the i th platform.

However, based on the Gauss–Markov theorem, the best linear unbiased estimation of the general linear model described in (5) is

$$\bar{\mathbf{x}} = [\Delta P \ \Delta L \ \Delta H]^T = (\mathbf{A}_m^T \boldsymbol{\Sigma}_{\mathbf{y}_m}^{-1} \mathbf{A}_m)^{-1} \mathbf{A}_m^T \boldsymbol{\Sigma}_{\mathbf{y}_m}^{-1} \mathbf{b}_m \quad (9)$$

it means that the weight matrix will be

$$\mathbf{W}_m = \text{diag} \left[\left(\frac{1}{\text{var}(\text{Res}_{m1}^r)} \right)^2 \left(\frac{1}{\text{var}(\text{Res}_{m1}^c)} \right)^2 \right. \\ \left. \times \left(\frac{1}{\text{var}(\text{Res}_{m2}^r)} \right)^2 \left(\frac{1}{\text{var}(\text{Res}_{m2}^c)} \right)^2 \right] \quad (10)$$

where $\text{var}(\text{Res}_{mi}^r)$ ($i = 1, 2$) denote the variance of the row size of the i th platform and $\text{var}(\text{Res}_{mi}^c)$ ($i = 1, 2$) denote the variance of the column size of the i th platform.

Under the assumption that observations in (5) are independent which is universally accepted in the literature and according to the theory described in [30], the optimal weights for WLS or GLS will be

$$\mathbf{W}_m = \text{diag} \left[\frac{\text{mean}(\text{Res}_{m1}^r) \text{mean}(\text{Res}_{m1}^c)}{(\text{var}(\text{Res}_{m1}^r))^2 (\text{var}(\text{Res}_{m1}^c))^2} \right. \\ \left. \times \frac{\text{mean}(\text{Res}_{m2}^r) \text{mean}(\text{Res}_{m2}^c)}{(\text{var}(\text{Res}_{m2}^r))^2 (\text{var}(\text{Res}_{m2}^c))^2} \right] \quad (11)$$

and considering that the standard deviation of interest point extraction, which is either manual or automatic, should be at subpixel level [31]–[33], the optimal weighted matrix described in (11) will be rearranged as

$$\mathbf{W}_m = \text{diag} \left[(\text{var}(\text{Res}_{m1}^r))^{-1} \eta \cdot (\text{var}(\text{Res}_{m1}^c))^{-1} \right. \\ \left. \times \kappa \cdot (\text{var}(\text{Res}_{m2}^r))^{-1} \lambda \cdot (\text{var}(\text{Res}_{m2}^c))^{-1} \right] \quad (12)$$

where η , κ , and λ are three unknown parameters which can be set manually by experience or can be set by solving the following optimal condition in (13) using training data:

$$[\eta^* \kappa^* \lambda^*] = \arg \min_{\eta \in R_\eta, \kappa \in R_\kappa, \lambda \in R_\lambda} \|\mathbf{x} - \bar{\mathbf{x}}\|_2 \quad (13)$$

where R_η , R_κ , R_λ are ranges of η , κ , and λ which can be easily selected under the consideration of the relationship of

resolution among the tested pixels and $\|\mathbf{x} - \bar{\mathbf{x}}\|_2$ denotes the Euclidean norm of the difference between the ground truth and the estimated coordinates of the training data.

Nevertheless, can the weighted matrix taking form as (10), (11), or (12) really address the problem caused by observations from different platforms? From the following analysis, we can see that there should be a much better choice existing.

III. GENERALIZED 3-D RECONSTRUCTION FOR MULTIPLATFORM

All the aforementioned analyses are based on the following assumptions: 1) The resolutions of the stereo-image pair used for 3-D reconstruction are the same or at a same level, and 2) the pixel sizes of an image are all equal. However, these assumptions cannot be satisfied if the used stereo-image pair is acquired from different platforms/sensors or one or more images of the used stereo-image pair are acquired with a large viewing angle. Under these circumstances, the difference of pixel size cannot be neglected. Moreover, the relationship between pixel size and its variance is a negative exponent function [34], [35]. Obviously, the corresponding difference of variances can also not be ignored. Consequently, the formed least squares problem becomes a seriously unbalanced one. As a result, it is unreasonable to set the weight matrix to the identity matrix or as the form of (10) or (11). The corresponding weight matrix must be estimated by considering the difference of pixel size and the corresponding difference of variances of images used for 3-D reconstruction.

A. Mathematical Model of G-RFM 3-D Reconstruction Method

For addressing the problem about 3-D reconstruction with GSP from multiplatforms in depth, we need to deeply analyze this problem from its mathematical model—general least squares estimation which can be expressed as

$$\mathbf{Y} = \mathbf{A}\boldsymbol{\gamma} + \mathbf{E} \quad (14)$$

where $\mathbf{Y} = [y_1 \ y_2 \ \dots \ y_n]^T$, $\mathbf{Y} \in \mathbf{R}^{n \times 1}$ is an $n \times 1$ vector of the measurements, $\mathbf{A} \in \mathbf{R}^{n \times m}$ is an $n \times m$ ($n > m$) design matrix with full column rank, $\boldsymbol{\gamma} = [\gamma_1 \ \gamma_2 \ \dots \ \gamma_m]^T$, $\boldsymbol{\gamma} \in \mathbf{R}^{m \times 1}$ is an $m \times 1$ vector of unknown parameters, and $\mathbf{E} = [e_1 \ e_2 \ \dots \ e_n]^T$, $\mathbf{E} \in \mathbf{R}^{n \times 1}$ is an $n \times 1$ vector about errors of the measurements.

If all of the measurements are acquired from a same source, the mathematical expectation of \mathbf{E} would be $E\{\mathbf{E}\} = \mathbf{O} = [0 \ 0 \ \dots \ 0]^T \in \mathbf{R}^{n \times 1}$, and the covariance matrix of \mathbf{E} would be $\mathbf{Q}_E = \mathbf{Q}_Y = \sigma^2 \mathbf{I}_n$ without considering the correlation between these measurements. This case is called equal precision (EQ) here. On the contrary, if the measurements are acquired from multiplatforms (or multisensors), it is called the unequal precision (UEQ) case which is the right one concerned here. Furthermore, in practice, even if we can assume that observations from different sources are uncorrelated, the correlation among observations from a same source cannot be ignored.

The corresponding mathematic model of the UEQ case can be viewed as a multi-EQ case

$$\begin{cases} \mathbf{Y}_1 = \mathbf{A}_1\gamma + \mathbf{E}_1 \\ \mathbf{Y}_2 = \mathbf{A}_2\gamma + \mathbf{E}_2 \\ \dots \\ \mathbf{Y}_k = \mathbf{A}_k\gamma + \mathbf{E}_k \end{cases} \quad (15)$$

where $\mathbf{E}_i \sim (0, \Sigma_{\mathbf{E}_i})$ with $\Sigma_{\mathbf{E}_i}$ being a positive definite covariance matrix and n_i being the number of measurements from the i th platform/sensor, k is the number of platforms/sensors, and \mathbf{E}_i is uncorrelated with \mathbf{E}_j . Then, (15) can be rewritten in a form that is the same as that of the standard linear model

$$\ddot{\mathbf{Y}} = \ddot{\mathbf{A}}\gamma + \ddot{\mathbf{E}} \quad (16)$$

where $\ddot{\mathbf{Y}} = [\mathbf{Y}_1 \ \mathbf{Y}_2 \ \dots \ \mathbf{Y}_k]$, $\mathbf{Y}_i = [y_{i1} \ y_{i2} \ \dots \ y_{in_i}]$, the mathematical expectation of $\ddot{\mathbf{E}}$ is zero, and the covariance matrix of $\ddot{\mathbf{E}}$ is

$$\mathbf{Q}_{\ddot{\mathbf{E}}} = \mathbf{Q}_{\ddot{\mathbf{Y}}} = \begin{pmatrix} \Sigma_{\mathbf{E}_1} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Sigma_{\mathbf{E}_2} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \dots & \Sigma_{\mathbf{E}_{k-1}} & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \Sigma_{\mathbf{E}_k} \end{pmatrix}. \quad (17)$$

For both the EQ case and the UEQ case, the GLS estimation, which is the best linear unbiased, efficient, and consistent estimation, can be expressed as

$$\begin{aligned} \tilde{\gamma} &= (\mathbf{A}^T \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{Y} \\ \text{or } \tilde{\gamma} &= (\ddot{\mathbf{A}}^T \mathbf{Q}_{\ddot{\mathbf{E}}}^{-1} \ddot{\mathbf{A}})^{-1} \ddot{\mathbf{A}}^T \mathbf{Q}_{\ddot{\mathbf{E}}}^{-1} \ddot{\mathbf{Y}}. \end{aligned} \quad (18)$$

For the problem of multiplatform 3-D reconstruction, the design matrix $\ddot{\mathbf{A}}$ and the measurements $\ddot{\mathbf{Y}}$ are known. Therefore, we turn to finding a way to estimate the elements of the corresponding covariance matrix.

Let (19) be defined as shown at the bottom of the page, where k is the number of platforms, n_k is the number of observations from a single platform, $\sigma_{i-uu}^2 (1 \leq u \leq n_k; 1 \leq i \leq k)$ denotes the variance component of the u th observation from the i th platform, $\sigma_{i-jvs}^2 (1 \leq i \leq n; 1 \leq j \leq n_k; 1 \leq s \leq n_k)$ denotes the covariance component between the j th and the s th observations from the i th platform, $N = \sum_{i=1}^k n_i$ is the total number

of observations, and Q_r^i and Q_{jvr}^i with $i = 1, 2, \dots, nj$, $r = 1, 2, \dots, n_i$ are known factors or the initialization values of factors for the corresponding variance or covariance components. In matrix form, there is

$$\mathbf{Q} = \sum_{i=1}^N \tilde{\mathbf{Q}}_i \theta_i \quad (20)$$

where $M = M_1 + M_2 = \sum_{i=1}^k n_i + \sum_{i=1}^k (n_i \cdot (n_i - 1)/2) = \sum_{i=1}^k (n_i \cdot (n_i + 1)/2)$ is the total number of Q_r^i and Q_{jvr}^i in (19) (M_1 is the number of Q_r^i , and M_2 is the number of Q_{jvr}^i) and θ_i is the corresponding variance or covariance components. If θ_i is a variance component, the form of $\tilde{\mathbf{Q}}_i$ will be

$$\tilde{\mathbf{Q}}_i = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \dots & \vdots \\ 0 & \dots & Q_r^i & \dots & 0 \\ \vdots & \dots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix} \quad (21)$$

and if θ_i is a covariance component, the form of $\tilde{\mathbf{Q}}_i$ will be

$$\tilde{\mathbf{Q}}_i = \begin{bmatrix} 0 & \dots & Q_{jvr}^i & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ Q_{jvr}^i & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix}. \quad (22)$$

It can be seen that the original problem is turned to obtain variance/covariance components θ_i in (20). Therefore, the VCE technique [23]–[26] can be exploited here to calculate the corresponding variance/covariance components. For using the VCE technique, $\mathbf{P} = \mathbf{Q}^{-1} |_{\theta=[\theta_1 \ \theta_2 \ \dots \ \theta_{nn}]=[1 \ 1 \ \dots \ 1]} = \sum_{i=1}^{nn} \tilde{\mathbf{P}}_i$ with $\tilde{\mathbf{P}}_i$ having the same form of (21) or (22) as the corresponding variance or covariance component.

To successfully use the G-RFM method, it only needs to initialize the weighting factors in (19). Obviously, it can be acquired just like (10) to reflect the effect caused by unbalanced pixel sizes with consideration of correlation among observations. Therefore, it is necessary to calculate the accurate pixel size of each point of GSP under different situations.

$$\mathbf{Q}_{\ddot{\mathbf{E}}} = \mathbf{Q}_{\ddot{\mathbf{Y}}} = \mathbf{Q} = \begin{bmatrix} Q_1^1 \sigma_{1-11}^2 & \dots & Q_{1vn_1}^1 \sigma_{1-1vn_1}^2 & & \\ \vdots & \ddots & \vdots & & \\ Q_{1vn_1}^1 \sigma_{1-1vn_1}^2 & \dots & Q_{n_1}^1 \sigma_{1-n_1n_1}^2 & & \\ & & & \ddots & \\ & & & & Q_1^k \sigma_{k-11}^2 & \dots & Q_{1vn_k}^k \sigma_{k-1vn_k}^2 \\ & & & & \vdots & \ddots & \vdots \\ & & & & Q_{1vn_k}^k \sigma_{k-1vn_k}^2 & \dots & Q_{n_k}^k \sigma_{k-n_kn_k}^2 \end{bmatrix} \quad (19)$$

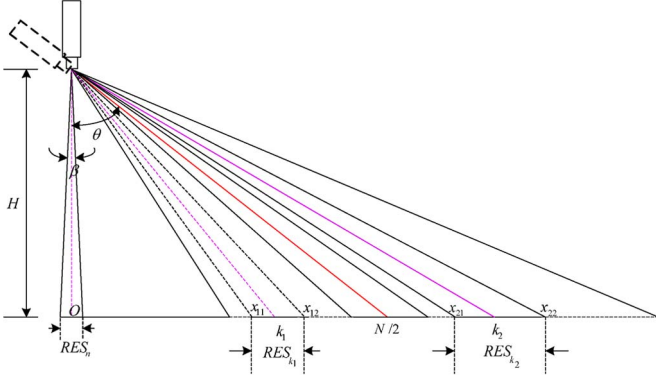


Fig. 1. Pixel size analysis with sufficient information.

B. Pixel Size Analysis

In practice, there are two different situations to be considered. For the first situation, sufficient information about the imaging condition for the used stereo-image pair can be obtained, i.e., sensor orientation parameters or the general sensor imaging model. For the second situation, only the mean pixel sizes of the nadir view and the corresponding viewing angle (if not known, it can be easily estimated by information existing in the image, such as shadow) of the used stereo-image pair can be obtained. Under these two situations, different methods for accuracy pixel size analysis will be conducted.

For the first situation, the diagram of imaging geometry in remote sensing is shown in Fig. 1. In Fig. 1, N denotes the total number of a line (column) of the image, H is the distance from the sensor to the ground, θ is the viewing angle of the sensor; β is the angle of the instantaneous FOV (IFOV for short), RES_n is the pixel size of the nadir view, and RES_{k_1} and RES_{k_2} are the corresponding pixel sizes of k_1 th ($1 \leq k_1 \in \mathbb{Z}^+ < N/2$) and k_2 th ($N/2 \leq k_2 \in \mathbb{Z}^+ \leq N$) pixels.

From Fig. 1, we can find that the distances from nadir point O to x_{11} and to x_{12} are

$$S_{11} = H \cdot \tan \left(\theta - \left(\frac{N}{2} - k_1 \right) \cdot \beta - \frac{\beta}{2} \right) \quad (23)$$

$$S_{12} = H \cdot \tan \left(\theta - \left(\frac{N}{2} - k_1 \right) \cdot \beta + \frac{\beta}{2} \right) \quad (24)$$

where $\beta = 2 \arctan(RES_n/2H)$.

Therefore, the size of the k_1 th pixel is

$$\begin{aligned} RES_{k_1} &= S_{12} - S_{11} \\ &= H \cdot \tan \left(\theta - \left(\frac{N}{2} - k_1 \right) \cdot \beta + \frac{\beta}{2} \right) \\ &\quad - H \cdot \tan \left(\theta - \left(\frac{N}{2} - k_1 \right) \cdot \beta - \frac{\beta}{2} \right). \end{aligned} \quad (25)$$

Similarly, the size of the k_2 th pixel is

$$\begin{aligned} RES_{k_2} &= H \cdot \tan \left(\theta - \left(\frac{N}{2} - k_2 \right) \cdot \beta + \frac{\beta}{2} \right) \\ &\quad - H \cdot \tan \left(\theta - \left(\frac{N}{2} - k_2 \right) \cdot \beta - \frac{\beta}{2} \right). \end{aligned} \quad (26)$$

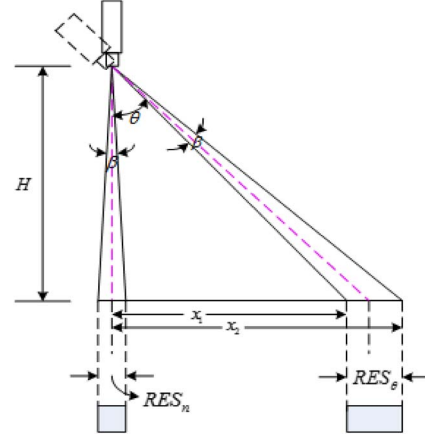


Fig. 2. Pixel size analysis without sufficient information.

The analysis above is under the condition that $\theta \geq 0^\circ$. If $\theta < 0^\circ$, the pixel sizes of the k_1 th and k_2 th pixels are

$$\begin{aligned} RES_{k_1} &= H \cdot \tan \left(\left| \theta - \left(\frac{N}{2} - k_1 \right) \cdot \beta - \frac{\beta}{2} \right| \right) \\ &\quad - H \cdot \tan \left(\left| \theta - \left(\frac{N}{2} - k_1 \right) \cdot \beta + \frac{\beta}{2} \right| \right) \end{aligned} \quad (27)$$

$$\begin{aligned} RES_{k_2} &= H \cdot \tan \left(\left| \theta - \left(\frac{N}{2} - k_2 \right) \cdot \beta - \frac{\beta}{2} \right| \right) \\ &\quad - H \cdot \tan \left(\left| \theta - \left(\frac{N}{2} - k_2 \right) \cdot \beta + \frac{\beta}{2} \right| \right). \end{aligned} \quad (28)$$

For the second situation that there is no sufficient information to do accuracy pixel size analysis, the coarse relationship between pixel sizes of the image obtained from a pitch angle and from nadir which is illustrated in Fig. 2 will be used.

In Fig. 2, H is the distance from the sensor to the ground, θ is the viewing angle of the sensor, β is the angle of the IFOV, RES_n is pixel size of the nadir view, and RES_θ is the pixel size of the image with viewing angle θ . Among these parameters, only RES_n and θ are known. Moreover, the objective of the analysis is to obtain the relationship between RES_θ and RES_n . From Fig. 2, we can get that

$$RES_n = 2H \tan \left(\frac{\beta}{2} \right) \quad (29)$$

$$\begin{aligned} RES_\theta &= x_2 - x_1 \\ &= H \cdot \tan \left(\theta + \frac{\beta}{2} \right) - H \cdot \tan \left(\theta - \frac{\beta}{2} \right). \end{aligned} \quad (30)$$

Therefore

$$\begin{aligned} \frac{RES_\theta}{RES_n} &= \frac{H \cdot \tan \left(\theta + \frac{\beta}{2} \right) - H \cdot \tan \left(\theta - \frac{\beta}{2} \right)}{2H \cdot \tan \left(\frac{\beta}{2} \right)} \\ &= \frac{\frac{\tan(\theta) + \tan(\beta/2)}{1 - \tan(\theta) \tan(\beta/2)} - \frac{\tan(\theta) - \tan(\beta/2)}{1 + \tan(\theta) \tan(\beta/2)}}{2 \tan \left(\frac{\beta}{2} \right)} \end{aligned}$$

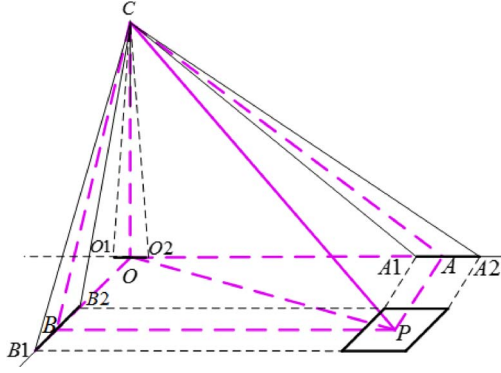


Fig. 3. Resolution analysis under the most general case.

$$\begin{aligned}
 &= \frac{(1+\tan^2(\theta)) \tan(\frac{\beta}{2})}{1-\tan^2(\theta) \tan^2(\frac{\beta}{2})} \\
 &= \frac{\tan(\frac{\beta}{2})}{\tan(\frac{\beta}{2})} \\
 &= \frac{1+\tan^2(\theta)}{1-\tan^2(\theta) \tan^2(\frac{\beta}{2})} \\
 &= \frac{1}{\cos^2(\theta)} \cdot \frac{1}{1-\tan^2(\theta) \tan^2(\frac{\beta}{2})} \approx \frac{1}{\cos^2(\theta)} \cdot (31)
 \end{aligned}$$

Based on the analysis above, let us turn to the most general case that the used image is acquired by a sensor with elevation angle θ and azimuth angle ϕ as shown in Fig. 3.

When $\theta = \angle OCP$ and $\phi = \angle BPO$, the height of the sensor above nadir point O is the distance from point C to point O. From Fig. 3, the relative viewing angle $\theta_{along} = \angle OAC$ in the direction of along track and the corresponding one $\theta_{across} = \angle OBC$ in the direction of across track can be calculated as follows:

$$\theta_{along} = \angle OAC = \arctan(\tan(\theta) \cdot \cos(\phi)) \quad (32)$$

$$\theta_{across} = \angle OBC = \arctan(\tan(\theta) \cdot \sin(\phi)). \quad (33)$$

Then, the parameter θ in (22)–(28) for pixel size analysis for the along track and across track will be substituted by θ_{along} and θ_{across} , respectively. Obviously, in this case, there is certainly correlation between the pixel sizes in along track and in across track which cannot be neglected.

Based on the analysis above, detailed (column) pixel size analysis results of the QuickBird image were illustrated in Appendix III.

C. Realization of G-RFM Method

Until now, all theory details for the feasibility of the proposed G-RFM method were deduced in former sections. Moreover, the detailed realization of the G-RFM method can be given. Here, it is assumed that the stereo-image pair is composed of two images obtained from two platforms, and the objective is to reconstruct object point (P, L, H) corresponding to pixels (r_{m1}, c_{m1}) and (r_{m2}, c_{m2}) in the two images. The work flow of G-RFM is shown in Fig. 4.

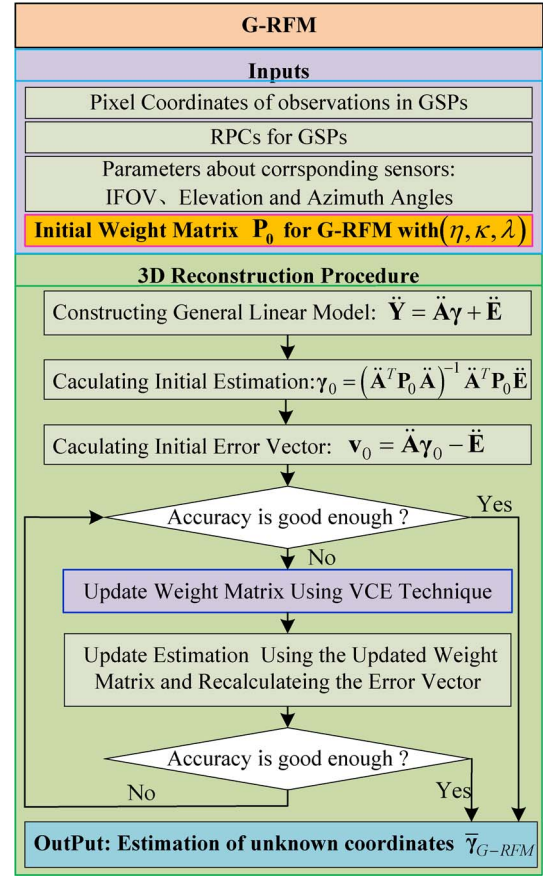


Fig. 4. Work flow of the proposed G-RFM method.

For more details, there are two places which have to be explained.

1) *Initial Weight Matrix P_0* : In Fig. 4, we can see that one of the important inputs is P_0 which will take the form as described in

$$P_0 = \begin{bmatrix} 1 & \eta \cdot \rho_1 & 0 & 0 \\ \eta \cdot \rho_1 & \eta \cdot \rho_1 & 0 & 0 \\ 0 & 0 & \kappa \cdot \rho_2 & \frac{\kappa+\lambda}{2} \cdot \left(\frac{\rho_2}{\rho_3}\right) \\ 0 & 0 & \frac{\kappa+\lambda}{2} \cdot \left(\frac{\rho_2}{\rho_3}\right) & \lambda \cdot \rho_3 \end{bmatrix} \quad (34)$$

where $\rho_i (i = 1, 2, 3)$ is the ratio of the $(i+1)$ th observed pixel size of the test point calculated by using (25)–(33) with consideration of the azimuth angle. Furthermore, in practice, the relationship of these parameters is set to $1 \geq \rho_1 \geq \rho_2 \geq \rho_3$ by arranging observations in descending order according to pixel size.

2) *Update Weight Matrix Using VCE Technique*: This is the kernel part of G-RFM. The VCE technique is used to accurately estimate the variance/covariance components of the weight matrix. As mentioned earlier, there are several types of VCE technique. The one used in this paper is the general Helmert form described in [25]. For using it iteratively, we must first initialize a matrix formed as in

$$R_0 = \ddot{A} (\ddot{A}^T Q_0^{-1} \ddot{A})^{-1} \ddot{A}^T Q_0^{-1} - I = \ddot{A} (\ddot{A}^T P_0 \ddot{A})^{-1} \ddot{A}^T P_0 - I \quad (35)$$

and then, the iterative procedure can be carried out as described in Appendix II.

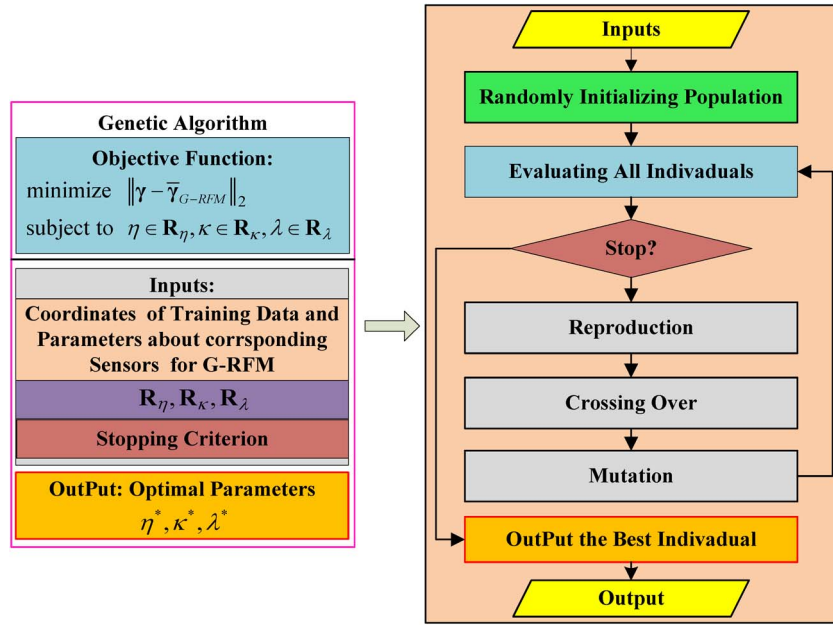


Fig. 5. GA for finding optimal parameters η^* , κ^* , λ^* .

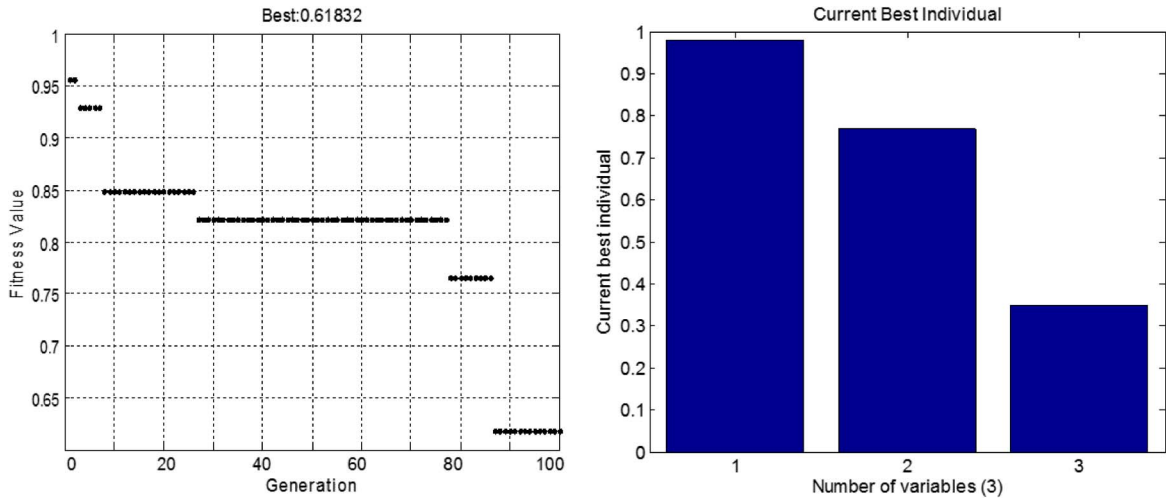


Fig. 6. Illustration of GA optimization results for finding optimal parameters η^* , κ^* , λ^* .

For achieving the best performance, the best configuration of parameters η , κ , λ in (34) should be determined. Obviously, this is a complex optimization problem. Based on statistical results, we find that the solution space of this problem is nonsmooth with multiple local optimums. Genetic algorithms (GAs) are the most successful global optimization tool based on the well-known theory of natural selection and evolutionary processes [39], [40]. GAs are a kind of stochastic but directed search method in a parallel iterative manner. Moreover, GAs do not require any prior knowledge or special properties (e.g., smoothness, convexity, unimodality, or existence of derivatives) of the objective function of the optimization problem. Therefore, GAs have been successfully applied to solve the complex optimization problem in many science and engineering branches [42]–[45]. Consequently, the task for finding the best configuration of parameters η , κ , λ is dealt with GAs illustrated in Fig. 5.

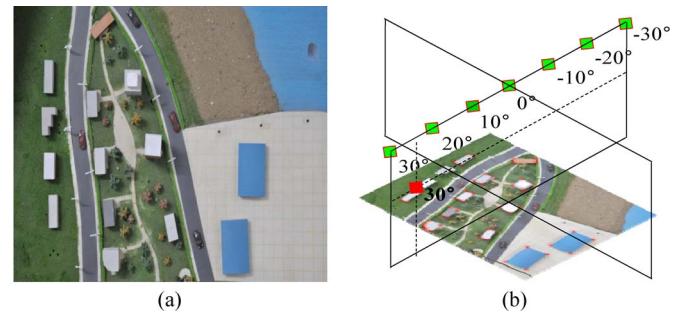


Fig. 7. Test data obtained from semisimulation platform. (a) Nadir view of the test area. (b) Illustration of camera positions.

For stereo-image-pair-based photogrammetry methods, vertical accuracy is the most concerned one, and horizontal accuracy is always better than vertical accuracy. Therefore, only the height components of training points were used in GAs. That is

TABLE I
RMSE COMPARISON OF EXPERIMENT 2 FOR THE CORRESPONDING REAL SCENE (UNIT: METERS)

	C-RFM			W-RFM			OW-RFM			G-RFM-2		
	P	L	H	P	L	H	P	L	H	P	L	H
-30°	0.140	0.273	0.182	0.132	0.269	0.200	0.143	0.275	0.205	0.083	0.092	0.098
-20°	0.181	0.287	0.205	0.179	0.288	0.203	0.185	0.276	0.212	0.203	0.281	0.227
-10°	0.495	0.465	0.766	0.320	0.398	0.579	0.317	0.365	0.596	0.322	0.398	0.406
0°	0.183	0.331	0.319	0.192	0.351	0.336	0.197	0.359	0.314	0.197	0.259	0.240
10°	0.261	0.352	0.382	0.254	0.353	0.404	0.242	0.328	0.397	0.205	0.130	0.276
20°	0.287	0.439	0.447	0.287	0.411	0.438	0.281	0.405	0.442	0.228	0.347	0.388

TABLE II
COMPARISON OF TIME CONSUMPTION OF
EXPERIMENT 2 (UNIT: SECONDS)

	-30°	-20°	-10°	0°	10°	20°	average
C-RFM	0.812	0.810	0.845	0.812	0.814	0.820	0.818
W-RFM	0.815	0.815	0.830	0.814	0.813	0.817	0.817
OW-RFM	0.815	0.816	0.834	0.812	0.814	0.815	0.818
G-RFM-2	1.423	1.467	1.839	1.432	1.503	1.437	1.517

to say, the objective function for GA is taken as the form in (13) with tiny adjustment as shown in the following:

$$[\eta^* \kappa^* \lambda^*] = \arg \min_{\eta \in R_\eta, \kappa \in R_\kappa, \lambda \in R_\lambda} |H - \bar{H}|. \quad (36)$$

On the whole, for the realization of G-RFM in practice, there are two separate parts included: The first is using GAs to obtain optimal parameters $\eta^*, \kappa^*, \lambda^*$ for G-RFM offline. Second, with the optimal parameters $\eta^*, \kappa^*, \lambda^*$, G-RFM with the work flow described in Fig. 4 can be carried out for any test point in the corresponding GSP.

In Fig. 6, there is an illustration of the optimization results about Experiment 4 under the situation that sufficient imaging information is available, and in the next section, for avoiding procrastination, only the experimental results with optimized parameters are given.

D. Computational Complexity Analysis

The number of floating point operations (flops) which is commonly used in the literature is chosen to measure the computational complexity. For the proposed G-RFM method taking in an iterative way, only the computational complexity for a single iteration is analyzed. In addition, for multiplication and inversion of the matrix being the heaviest computational burden in G-RFM, only these two kinds of operations are taken into consideration. The following computational complexity analysis is all based on the analysis described in Sections III-A and C.

For a particular iteration, there are mainly three parts (as shown in Appendix II): 1) constructing matrices \mathbf{T}_s and \mathbf{W}_s ; 2) updating variance and covariance components $\hat{\theta}_s$; and 3) calculating the unknown parameters $\hat{\mathbf{X}}_s$ and the corresponding residual vector \mathbf{V}_s . As to the first part, it needs $M(N^2 + 2N) + (M(M-1)/2)(2N^2 + 4N)$ and $M_1 + 2M_2$ flops to construct matrices \mathbf{T}_s and \mathbf{W}_s , respectively. For the second part, it needs $O(M^3)$ flops to calculate $\hat{\theta}_s$ using GLS



Fig. 8. Stereo images of ISPRS Avenches data set. (a) image_5873. (b) image_5874.

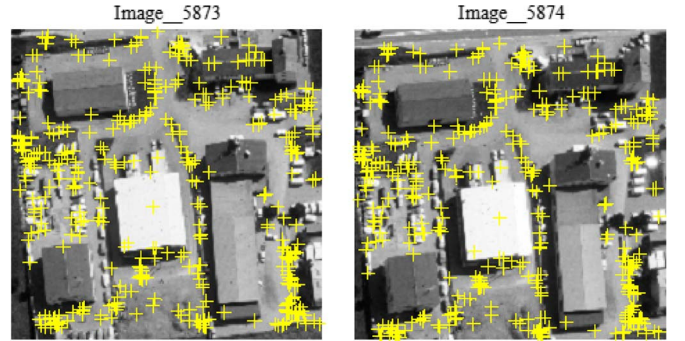


Fig. 9. Selected 450 pairs of GCPs.

estimation. At last, for the third part, there are also about $O(N^3)$ flops to estimate the unknown parameters and to get the corresponding residuals, so overall, the total computational complexity of a single iteration of the proposed G-RFM is

$$\begin{aligned} T &= M(N^2 + 2N) + \frac{M(M-1)}{2}(2N^2 + 4N) + M_1 + 2M_2 \\ &\quad + O(M^3) + O(N^3) \\ &= O(M^2) + O(M^3) + O(M) + O(M^3) + O(N^3) \\ &= O(M^3). \end{aligned} \quad (37)$$

In contrast, the computational complexity of traditional methods without VCE technique embedded is only $O(N^3)$ to calculate the unknown parameters and the residuals.

Obviously, the proposed G-RFM method requires more computational complexity than the traditional methods if there are too many observations used. Nevertheless, in practice, there are commonly two images used for deriving the 3-D information in object space. It means that there are only four observations

TABLE III
RMSE COMPARISON OF EXPERIMENT 1 (UNIT: METERS)

	1v2			1v4			1v8		
	P	L	H	P	L	H	P	L	H
C-RFM	0.3417	0.4889	0.7686	0.4238	0.6029	0.9768	0.5655	0.8163	1.1972
W-RFM	0.3056	0.4930	0.7574	0.4352	0.6057	0.9688	0.4358	0.7321	1.0832
OW-RFM	0.2610	0.5205	0.7469	0.3125	0.5476	0.8843	0.3462	0.6632	1.0221
G-RFM-1	0.1047	0.1198	0.1612	0.1005	0.1335	0.1903	0.1051	0.1369	0.1635

(two observations from a single image) for estimating the 3-D coordinates of the unknown point in object space. Furthermore, observations from different images can be safely assumed uncorrelated. Therefore, the number of unknown covariance components will reduce again, and for consideration of the high development of the corresponding hardware technique, the computational burden of the proposed G-RFM method cannot affect its application in practice.

IV. EXPERIMENTAL RESULTS

To validate the proposed G-RFM method under two different situations about pixel size analysis described in Section III, four experiments were conducted on simulated and real data sets. All the RFMs (third order) used in these experiments were solved by the method described in [13] with accuracy within $1e-5$ m \sim $1e-9$ m; the detail of this method is illustrated in Appendix I, and all the used GCPs with even distribution were selected by the method proposed in [36].

There are four different comparison methods used.

1) *C-RFM Method* [12]: The reconstruction formula is $\bar{\mathbf{x}} = [\Delta P \ \Delta L \ \Delta H]^T = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{b}$ with weight matrix $\mathbf{W} = \mathbf{I}$.

2) *Weighted RFM-Based Reconstruction Method (W-RFM)*: The key reconstruction formula was identical to that of C-RFM with the weight matrix taking the form of (10).

3) *Optimal Weight RFM-Based Reconstruction Method 2 (OW-RFM)*: The key reconstruction formula was identical to that of C-RFM with the weight matrix taking the form of (11).

4) *G-RFM Method—G-RFM*: It is the proposed method. For the first situation without sufficient information available, the corresponding name of the proposed method is called “G-RFM-1”; for the second situation with sufficient information known, the corresponding name of the proposed method is called “G-RFM-2.”

A. Experiment 1: Multiplatform GSP With Different Viewing Angles (Simulated Data)

The test data for the same scene from different platforms were simulated data obtained by a Nikon D300 camera with 4288×2848 pixels from a semisimulation platform (a 1 : 500 sand-table platform with measure accuracy of 3-D coordinate at millimeter level and half-degree accuracy of camera viewing angles) in our laboratory as shown in Fig. 7. In Fig. 7, the small blocks with green or red color are different positions of the camera. Therefore, images corresponding to a red block and any image corresponding to a green block can be used to form a stereo-image pair. The nadir pixel size of image captured

TABLE IV
COMPARISON OF TIME CONSUMPTION OF
EXPERIMENT 1 (UNIT: SECONDS)

	1v2	1v4	1v8	average
C-RFM	7.54	7.37	7.46	7.457
W-RFM	7.62	7.60	7.48	7.567
OW-RFM	7.62	7.62	7.50	7.580
G-RFM-1	12.67	13.45	13.56	13.227

from position of green block at 0° [$\alpha = 0^\circ$ $\beta = 0.0061^\circ$ in (25) and (26)] in Fig. 5 of this semi-simulation platform was 0.0143 cm (*0.072 m for the corresponding real scene*). Fifty-one pairs of GCPs were carefully selected for solving RPCs and 3-D reconstruction.

The corresponding experimental results were illustrated in Tables I and II.

B. Experiment 2: Multiplatform Stereo-Image Pair With Different Scales (Semisimulated Data)

The test data used in this experiment were the International Society for Photogrammetry and Remote Sensing (ISPRS) Avenches aerial data set downloaded from <http://www.isprs.org/data/avenches/default.aspx> with a 0.075-m resolution as shown in Fig. 8. The corresponding parameters of the rigorous sensor were included in the data set.

In this experiment, 450 pairs of GCPs were selected by the method used in [36] which is an integration of scale-invariant feature transform (SIFT) and maximally stable extremal regions (MSER) for coarse registration followed by the random sample consensus (RANSAC) technique and manual edition for removing the misregistered points, and the results were shown in Fig. 9. All of the 3-D coordinates in object space of these GCPs were calculated by the provided sensor parameters, and they were then used for solving RPCs. The image_5873 was down-sampled by factors 2, 4, and 8, and then, the original image_5873 and the corresponding down-sampled images formed three different stereo-image pairs with the original image_5874.

The results of this experiment were evaluated by root of mean square error (RMSE) as illustrated in Table III, and the corresponding time consumption was shown in Table IV.

C. Experiment 3: Real Multiangle Aerial Stereo Pairs (Real Data)

In this experiment, the test data were five multiangle aerial images with 0.25-m resolution selected from Zone4 of the Amiens Summer data set acquired from http://isprs.ign.fr/home_en.htm. The approximate camera positions of these

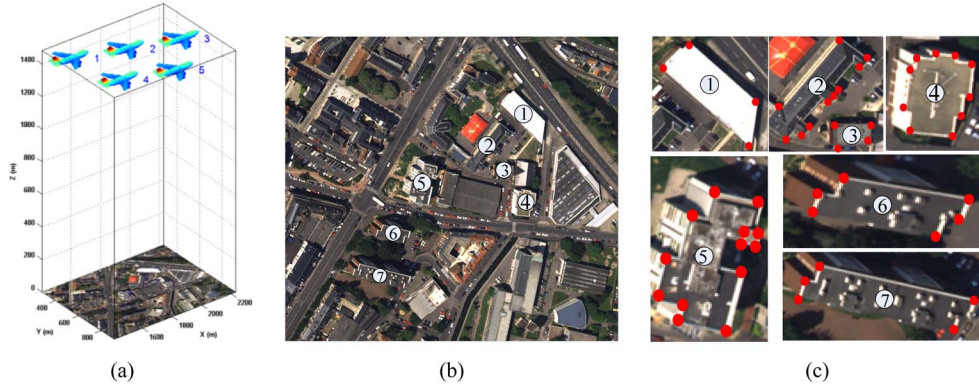


Fig. 10. Data used in this experiment. (a) Camera positions of the used multiview aerial images. (b) One of the test images of Amiens city. (c) Corresponding GCPs.

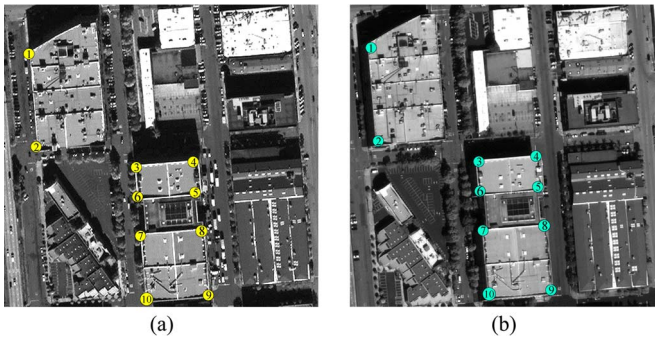


Fig. 11. Test area with GCP identified. (a) QuickBird satellite image. (b) WorldView2 satellite image.

images were illustrated in Fig. 10(a). One of these images was shown in Fig. 10(b), and there were a total of 52 manually selected GCPs in each image as shown in Fig. 10(c). The 3-D coordinates were calculated by the calibration parameter provided (see Fig. 11).

In this experiment, the total five images with different viewing angles composed ten GSPs. The corresponding experimental results were illustrated in Tables V and VI. In these two tables, the expression “*ivj*” denoted that the stereo-image pair is composed of images acquired in locations *i* and *j*.

D. Experiment 4: Real Satellite Stereo Pairs From Different Satellite (Real Data)

In this experiment, the test data set was the one used for the 2012 Data Fusion Contest. From this data set, two panchromatic satellite images about San Francisco acquired by QuickBird at 11-11-2007 and by WorldView2 at 09-10-2011 were selected to verify the performance of the G-RFM method. In Table VII, the key features of these two images were listed.

The test area used in this experiment was illustrated in Fig. 11.

The corresponding height values of the GCPs were obtained from the LiDAR data of the used data set with the vertical accuracy being 0.12 m at 95% confidence level for all land cover types. The accurate uniform horizontal coordinates of these GCPs were solved by employing the RPCs of these two satellite images. The corresponding results were illustrated in Tables VIII and IX.

E. Discussion

In this paper, the theory of 3-D reconstruction with GSP from different platforms/sensors was deeply analyzed, and four different experiments were conducted to evaluate the performance of the proposed G-RFM method. As discussed in the analysis in the former sections, there are two different situations as to the availability of imaging information: 1) There is sufficient imaging information available, and 2) only the mean pixel size and viewing angle are available.

The first experiment was used to evaluate the performance of the G-RFM method under the first situation by using simulated data with different viewing angles. Theoretically, by doing detailed pixel analysis and VCE estimation, the effect caused by the viewing angle can be eliminated. For this experiment, the corresponding experimental results of this experiment were illustrated in Table I. It can be seen that, as a whole, the proposed G-RFM method can really address the problem of 3-D reconstruction caused by the change of the viewing angle more successfully than classical methods. However, for GSP with the image from -20° used, the best method is not G-RFM but C-RFM, but it is the only exception.

The second experiment was used to test the performance of the G-RFM method under the second situation by using simulated data with different scales. As illustrated in Table III, the results of the G-RFM method were much better than that of other methods with strong stability. Certainly, the GSPs used in this experiment were designed to have obviously unbalanced resolutions which is very much in accordance with the problem addressed in this paper. Therefore, the results were a matter of cause.

For evaluating the validity of the proposed G-RFM to real data, Experiments 3 and 4 were conducted with GSPs composed by real aerial and satellite images under both situations, respectively. From Tables V and VIII, it can be seen that, under both situations, the performance of G-RFM was all better than that of the classical methods. Particularly for the first situation, G-RFM can dramatically outperform the classical methods. A little exception in Experiment 3 appeared with case “3v5” where the best method is C-RFM.

For the unexpectation that occurred in these experiments, obviously, G-RFM was trapped into local optimum in these cases, and these unexpectations all occurred in Experiments 1 and 3.

TABLE V
RMSE COMPARISON OF EXPERIMENT 3 (UNIT: METERS)

	C-RFM			W-RFM			OW-RFM			G-RFM-1			G-RFM-2		
	P	L	H	P	L	H	P	L	H	P	L	H	P	L	H
1v2	0.038	0.066	0.263	0.031	0.057	0.260	0.029	0.057	0.257	0.042	0.050	0.202	0.040	0.052	0.153
1v3	0.108	0.062	0.319	0.105	0.059	0.310	0.115	0.056	0.302	0.064	0.055	0.225	0.054	0.042	0.166
1v4	0.082	0.100	0.240	0.077	0.092	0.234	0.075	0.093	0.212	0.047	0.046	0.196	0.033	0.039	0.125
1v5	0.065	0.051	0.207	0.058	0.078	0.209	0.041	0.072	0.198	0.045	0.057	0.183	0.049	0.056	0.168
2v3	0.106	0.096	0.242	0.097	0.083	0.232	0.091	0.084	0.230	0.052	0.059	0.218	0.049	0.050	0.161
2v4	0.079	0.091	0.235	0.079	0.088	0.236	0.073	0.080	0.239	0.041	0.042	0.211	0.028	0.036	0.176
2v5	0.056	0.047	0.222	0.050	0.045	0.214	0.045	0.045	0.209	0.037	0.046	0.200	0.044	0.032	0.156
3v4	0.052	0.075	0.190	0.051	0.072	0.192	0.049	0.076	0.201	0.044	0.049	0.190	0.049	0.041	0.154
3v5	0.100	0.065	0.121	0.098	0.059	0.137	0.095	0.050	0.136	0.060	0.053	0.135	0.059	0.057	0.131
4v5	0.048	0.046	0.194	0.045	0.041	0.188	0.043	0.046	0.192	0.043	0.039	0.184	0.042	0.032	0.166

TABLE VI
COMPARISON OF TIME CONSUMPTION OF EXPERIMENT 3 (UNIT: SECONDS)

	1v2	1v3	1v4	1v5	2v3	2v4	2v5	3v4	3v5	4v5	average
C-RFM	0.832	0.835	0.832	0.836	0.834	0.842	0.831	0.835	0.832	0.841	0.835
W-RFM	0.833	0.834	0.835	0.840	0.834	0.843	0.834	0.836	0.832	0.837	0.835
OW-RFM	0.833	0.835	0.834	0.836	0.835	0.840	0.836	0.835	0.833	0.836	0.835
G-RFM-1	1.432	1.436	1.433	1.440	1.439	1.442	1.441	1.432	1.445	1.442	1.438
G-RFM-2	1.437	1.442	1.439	1.457	1.455	1.490	1.501	1.446	1.472	1.463	1.460

TABLE VII
KEY FEATURES OF THE USED SATELLITE IMAGES

	QuickBird	WorldView2
Height of Satellite	450 km	770 km
Elevation Angle of Satellite	61.2°	67.9°
Azimuth Angle of Satellite	10.1°	257.5°
Mean Off-Nadir View Angle	25.9°	19.6°
Mean Row Size	0.779 m	0.533 m
Mean Column Size	0.697 m	0.497 m

In Experiment 1, the used images were acquired by Nikon D300 with very high resolution, and the corresponding error of manually selected correspondence points may be beyond knowledge in common sense (in remote sensing). In Experiment 3, the used aerial images were acquired from almost the same height, and for images used in the unexpectation case, viewing angles from the sensor to the corresponding test points were almost symmetrical. It means that the unbalance effect discussed in this paper is very weak. Therefore, the possible reasons may be that the corresponding configurations of parameters R_η , R_κ , R_λ were unsuitable or that the observations (coordinates of the test points) used in these cases were just suitable for C-RFM. The real reason and the intrinsic relationship between imaging geometry (sensor height, viewing angle, convergence angle, etc.) and performance of G-RFM, particularly the selection of parameters R_η , R_κ , R_λ , will be studied in future.

As to computational complexity, in addition to the theoretical analysis in Section III-D, the corresponding time consumption of these experiments had been given to reflect computational complexity indirectly. It can be seen that the good reconstruction performance of the proposed G-RFM method is dependent on the sacrificed computational complexity.

V. CONCLUSION

In this paper, the G-RFM method for the problem of 3-D reconstruction with GSP from multiplatforms/multisensors is proposed. By strict mathematical analysis, the intrinsic mathematical model for this problem is deduced, i.e., GLS estimation problem. Furthermore, the way to accurately address this estimation problem is given based on strict analysis, and it is the main contribution of this paper to embed VCE technique into the classical RFM-based 3-D reconstruction procedure with detailed pixel size analysis. Experimental results illustrated that the proposed G-RFM method is an efficient method for integrating GSP from multiplatform with different resolutions. By using the proposed G-RFM method, a lot of redundant remote sensing images disused could be used for 3-D reconstruction, and the existing 3-D reconstruction results can be efficiently improved.

APPENDIX I

A. RFM in RPC00B Format [38]

RPC00Bs express the normalized column and row values in an image, (r_n, c_n) , as a ratio of polynomials of the normalized geodetic latitude, longitude, and height, (P, L, H) .

Normalized values are used instead of actual values in order to minimize numerical errors in the calculation. The scales and offset of each parameter are selected so that all normalized values fall in the range $[-1, 1]$. The normalization used is as follows:

$$P = \frac{(\text{Latitude} - \text{LAT_OFF})}{\text{LAT_SCLAE}}$$

$$L = \frac{(\text{Longitude} - \text{LONG_OFF})}{\text{LONG_SCLAE}}$$

TABLE VIII
RMSE COMPARISON OF EXPERIMENT 4 (UNIT: METERS)

	C-RFM	W-RFM	OW-RFM	G-RFM-1	G-RFM-2
P	0.5337	0.5315	0.5317	0.4803	0.4916
L	0.3174	0.3981	0.2282	0.4832	0.6433
H	1.0859	1.0874	1.0526	1.0242	0.6183

TABLE IX
COMPARISON OF TIME CONSUMPTION OF
EXPERIMENT 4 (UNIT: SECONDS)

	C-RFM	W-RFM	OW-RFM	G-RFM-1	G-RFM-2
Time	0.0150	0.0150	0.0160	0.0302	0.0310

TABLE X
RPC TERMS

i	$p_i(P, L, H)$	i	$p_i(P, L, H)$
1	1	11	PLH
2	L	12	L ³
3	P	13	LP ²
4	H	14	LH ²
5	LP	15	L ² P
6	LH	16	P ³
7	PH	17	PH ²
8	L ²	18	L ² H
9	P ²	19	P ² H
10	H ²	20	H ³

$$H = \frac{(Height - HEIGHT_OFF)}{HEIGHT_SCLAE}$$

$$r_n = \frac{(Row - LINE_OFF)}{LINE_SCLAE}$$

$$c_n = \frac{(Column - SAMP_OFF)}{SAMP_SCLAE}.$$

Each polynomial is up to the third order in (P, L, H) , having as many as 20 terms. The rational functions are

$$r_n = \frac{\sum_{i=1}^{20} LINE_NUM_COEF_i \cdot p_i(P, L, H)}{\sum_{i=1}^{20} LINE_DEN_COEF_i \cdot p_i(P, L, H)}$$

$$c_n = \frac{\sum_{i=1}^{20} SAMP_NUM_COEF_i \cdot p_i(P, L, H)}{\sum_{i=1}^{20} SAMP_DEN_COEF_i \cdot p_i(P, L, H)}.$$

$LINE_NUM_COEF$, $LINE_DEN_COEF$, $SAMP_NUM_COEF$, and $SAMP_DEN_COEF$ are 20-term vectors of coefficients that are given in the RPC00B file. $p(P, L, H)$ is a 20-term vector with the following terms.

2. Method for Solving RPCs With GCPs [13]

For the rational function model

$$row_{RFM} = r_n$$

$$= \frac{Num(P, L, H)}{Den(P, L, H)}$$

$$= \frac{\sum_{i=1}^{20} LINE_NUM_COEF_i \cdot p_i(P, L, H)}{\sum_{i=1}^{20} LINE_DEN_COEF_i \cdot p_i(P, L, H)}$$

$$= \frac{\sum_{i=1}^{20} a_i \cdot p_i(P, L, H)}{\sum_{i=1}^{20} b_i \cdot p_i(P, L, H)}.$$

The difference between the values of row_{RFM} and the real value of row (row) is shown at the bottom of the page.

Let

$$B = [1LPH \cdots L^2HP^2HH^3] \cdot [1b_2 \cdots b_{20}]^T$$

$$\mathbf{J} = [a_1a_2 \cdots a_{20}b_2 \cdots b_{20}]^T.$$

$$e_{row} = row_{RFM} - row$$

$$= \frac{[1LPH \cdots L^2HP^2HH^3] \cdot [a_1a_2a_3 \cdots a_{18}a_{19}a_{20}]^T}{[1LPH \cdots L^2HP^2HH^3] \cdot [b_1b_2b_3 \cdots b_{18}b_{19}b_{20}]^T} - row$$

$$= \frac{[1LPH \cdots L^2HP^2HH^3] \cdot [a_1a_2a_3 \cdots a_{18}a_{19}a_{20}]^T - row \cdot [1LPH \cdots L^2HP^2HH^3] \cdot [b_1b_2b_3 \cdots b_{18}b_{19}b_{20}]^T}{[1LPH \cdots L^2HP^2HH^3] \cdot [b_1b_2b_3 \cdots b_{18}b_{19}b_{20}]^T}.$$

$$= \frac{[1LPH \cdots L^2HP^2HH^3 - row \cdot L - row \cdot P \cdots - row \cdot H^3] \cdot [a_1a_2 \cdots a_{20}b_2 \cdots b_{20}]^T}{[1XYZ \cdots X^3Y^3Z^3] \cdot [b_1b_2b_3 \cdots b_{18}b_{19}b_{20}]^T}$$

$$- \frac{row}{[1LPH \cdots L^2HP^2HH^3] \cdot [b_1b_2b_3 \cdots b_{18}b_{19}b_{20}]^T} (b_1 = 1)$$

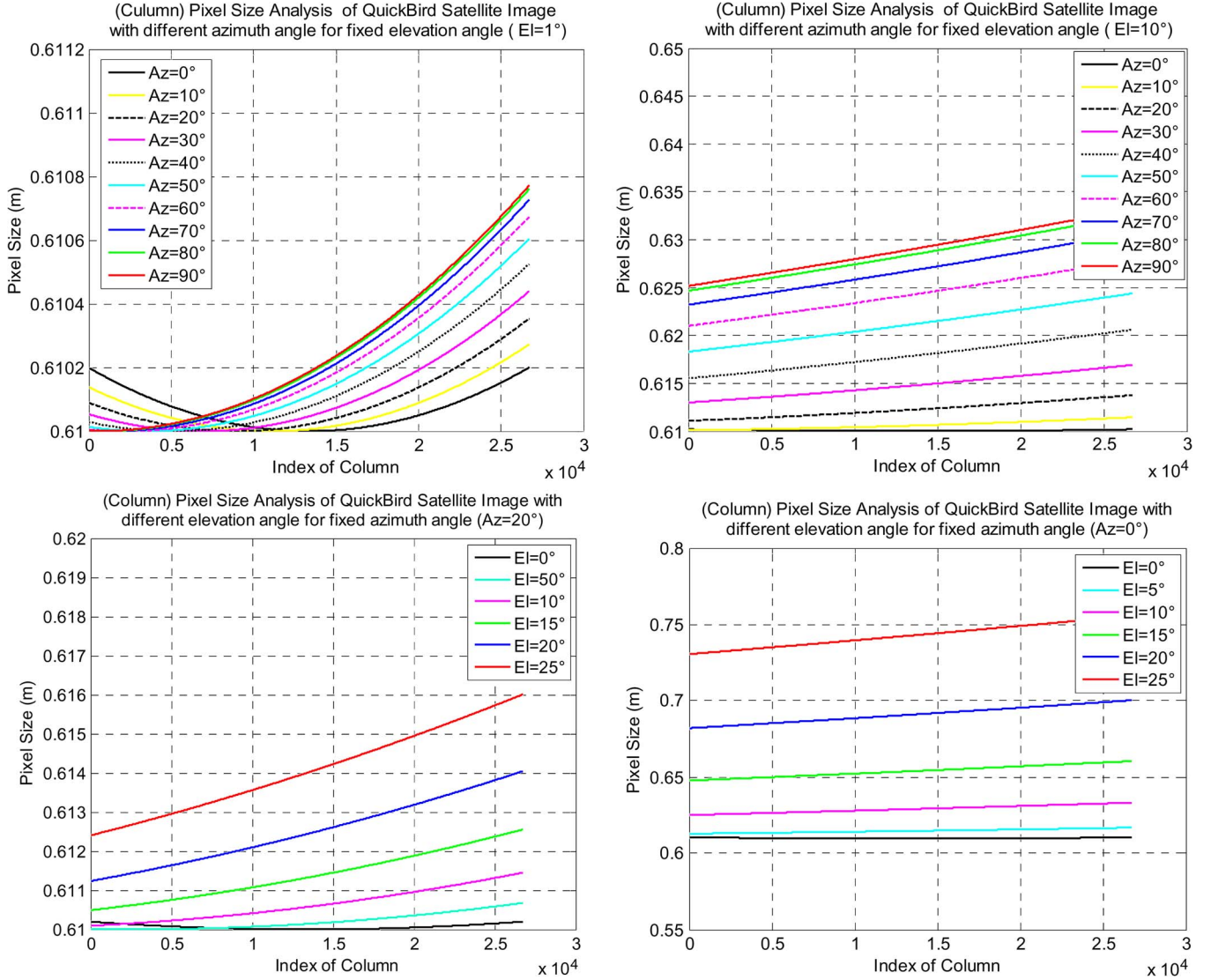


Fig. 12. Detailed (column) pixel size analysis results of QuickBird image.

There will be an equation

$$e_{row} = \left[\frac{1}{B} \frac{L}{B} \frac{P}{B} \frac{H}{B} \cdots \frac{H^3}{B} - \frac{row \cdot L}{B} - \frac{row \cdot P}{B} - \frac{row \cdot H}{B} \cdots - \frac{row \cdot H^3}{B} \right] \cdot \mathbf{J} - \frac{row}{B}.$$

Given $n(n \geq 39)$ GCPs with coordinates in object space known, there will be

$$\begin{bmatrix} e_{row1} \\ e_{row2} \\ \vdots \\ e_{rown} \end{bmatrix} = \begin{bmatrix} \frac{1}{B_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{B_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{B_n} \end{bmatrix} \cdot \begin{bmatrix} 1 & L_1 & \cdots & H_1^3 & -row_1 \cdot L_1 & \cdots & -row_1 H_1^3 \\ 1 & L_2 & \cdots & H_2^3 & -row_2 \cdot L_2 & \cdots & -row_2 H_2^3 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & L_n & \cdots & H_n^3 & -row_n \cdot L_n & \cdots & -row_n H_n^3 \end{bmatrix}$$

$$\cdot \mathbf{J} - \begin{bmatrix} \frac{1}{B_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{B_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{B_n} \end{bmatrix} \cdot \begin{bmatrix} row_1 \\ row_2 \\ \vdots \\ row_n \end{bmatrix}$$

or, in matrix format

$$\mathbf{E}_{row} = \mathbf{W}_{row} \mathbf{M} \mathbf{J} - \mathbf{W}_{row} \mathbf{R}.$$

Therefore, the solution of RPCs (i.e., \mathbf{J}) can be got by using the least squares estimation method as

$$\mathbf{J} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{R}.$$

The corresponding RPCs for the column model can also be obtained by the method described above.

$$\mathbf{T}_s = \begin{bmatrix} \text{tr}(\mathbf{R}_{s-1}^T \tilde{\mathbf{P}}_1 \mathbf{R}_{s-1} \tilde{\mathbf{Q}}_1) & \text{tr}(\mathbf{R}_{s-1}^T \tilde{\mathbf{P}}_1 \mathbf{R}_{s-1} \tilde{\mathbf{Q}}_2) & \cdots & \text{tr}(\mathbf{R}_{s-1}^T \tilde{\mathbf{P}}_1 \mathbf{R}_{s-1} \tilde{\mathbf{Q}}_s) \\ \text{tr}(\mathbf{R}_{s-1}^T \tilde{\mathbf{P}}_2 \mathbf{R}_{s-1} \tilde{\mathbf{Q}}_1) & \text{tr}(\mathbf{R}_{s-1}^T \tilde{\mathbf{P}}_2 \mathbf{R}_{s-1} \tilde{\mathbf{Q}}_2) & \cdots & \text{tr}(\mathbf{R}_{s-1}^T \tilde{\mathbf{P}}_2 \mathbf{R}_{s-1} \tilde{\mathbf{Q}}_s) \\ \vdots & \vdots & \ddots & \vdots \\ \text{tr}(\mathbf{R}_{s-1}^T \tilde{\mathbf{P}}_M \mathbf{R}_{s-1} \tilde{\mathbf{Q}}_1) & \text{tr}(\mathbf{R}_{s-1}^T \tilde{\mathbf{P}}_M \mathbf{R}_{s-1} \tilde{\mathbf{Q}}_2) & \cdots & \text{tr}(\mathbf{R}_{s-1}^T \tilde{\mathbf{P}}_M \mathbf{R}_{s-1} \tilde{\mathbf{Q}}_s) \end{bmatrix}$$

$$\mathbf{W}_s = [\mathbf{V}_{s-1}^T \tilde{\mathbf{P}}_1 \mathbf{V} \mathbf{V}_{s-1}^T \tilde{\mathbf{P}}_2 \mathbf{V} \cdots \mathbf{V}_{s-1}^T \tilde{\mathbf{P}}_M \mathbf{V}]^T$$

$$\hat{\boldsymbol{\theta}}_s = (\mathbf{T}_s)^+ \mathbf{W}_s$$

$$\hat{\mathbf{X}}_s = (\ddot{\mathbf{A}}^T \mathbf{Q}_s^{-1} \ddot{\mathbf{A}})^{-1} \mathbf{Q}_s^{-1} \ddot{\mathbf{Y}}$$

$$\mathbf{R}_s = \ddot{\mathbf{A}} (\ddot{\mathbf{A}}^T \mathbf{Q}_s^{-1} \ddot{\mathbf{A}})^{-1} \ddot{\mathbf{A}}^T \mathbf{Q}_s^{-1} - \mathbf{I}$$

$$\mathbf{V}_s = \mathbf{R}_s \ddot{\mathbf{Y}}$$

APPENDIX II

Iterative process of VCE technique to acquire variance components in (17) of mathematical model described in (13).

From the first step $s = 1$, do the process shown equation at the top of the page until meeting the stopping criterion.

The used iterative process above is a general Helmert method described in [25].

APPENDIX III

The detailed (column) pixel size analysis results of the QuickBird satellite image whose nadir ground sampling distance (GSD) is 0.61 m were illustrated in the following figures.

ACKNOWLEDGMENT

The authors would like to thank ISPRS, DigitalGlobe, Asrum Services, and U.S. Geological Survey for providing the imagery used in this study and the IEEE Geoscience and Remote Sensing Society Data Fusion Technical Committee for organizing the 2012 Data Fusion Contest.

REFERENCES

- [1] D. J. Gagan, "Practical aspects of topographic mapping from SPOT imagery," *Photogram. Rec.*, vol. 12, no. 69, pp. 349–355, Apr. 1987.
- [2] V. Kratky, "Rigorous photogrammetric processing of SPOT images at CCM Canada," *ISPRS J. Photogram. Remote Sens.*, vol. 44, no. 2, pp. 53–71, Oct. 1989.
- [3] L.-C. Chen and L.-H. Lee, "Least squares prediction using on-board data in bundle adjustment for SPOT imagery," in *Proc. IEEE IGARSS*, Jul. 1989, pp. 450–453.
- [4] E. M. Yasser and K. Novak, "Precision rectification of SPOT imagery using the direct linear transformation model," *Photogram. Eng. Remote Sens.*, vol. 62, no. 1, pp. 67–72, Jan. 1996.
- [5] T. Ono, S. Hattori, H. Hasegawa, and S. Akamatsu, "Digital mapping using high resolution satellite imagery based on 2D affine projection model," *Int. Archives Photogram. Remote Sens.*, vol. XXXIII, pt. B3, pp. 672–677, 2000.
- [6] T. Toutin, "DEM generation from new VIR sensors: IKONOS, ASTER and Landsat-7," in *Proc. IEEE IGARSS*, Aug. 2001, vol. 3, pp. 973–975.
- [7] M. Schmidt, R. Goossens, G. Menz, A. Altmaier, and D. Devriendt, "The use of CORONA satellite images for generating a high resolution digital elevation model," in *Proc. IEEE IGARSS*, Aug. 2001, vol. 7, pp. 3123–3125.
- [8] T. Tadono, M. Matsuoka, N. Futamura, J. Takaku, H. Suzuki, T. Iijima, M. Shimada, T. Igarashi, and R. Shibasaki, "High resolution DEM generation from ALOS PRISM data-algorithm development and evaluation," in *Proc. IEEE IGARSS*, Jun. 2002, vol. 1, pp. 405–407.
- [9] J. Takaku, N. Futamura, T. Iijima, T. Tadono, M. Shimada, and R. Shibasaki, "High resolution DEM generation from ALOS PRISM data," in *Proc. IEEE IGARSS*, Sep. 2004, vol. 7, pp. 4548–4551.
- [10] X. Yang, "Accuracy of rational function approximation in photogrammetry," in *Proc. Process. ASPRS Annu. Conv.*, May 22–26, 2000, [CD-ROM].
- [11] K. Di, R. Ma, and R. Li, "Deriving 3-D shorelines from high resolution IKONOS satellite images with rational functions," in *Proc. Process. ASPRS Annu. Conv.*, Apr. 24–27, 2001, [CD-ROM].
- [12] C. V. Tao and Y. Hu, "3D reconstruction methods based on the rational function model," *Photogram. Eng. Remote Sens.*, vol. 68, no. 7, pp. 705–714, Jul. 2002.
- [13] C. V. Tao and Y. Hu, "A comprehensive study of the rational function model for photogrammetric processing," *Photogram. Eng. Remote Sens.*, vol. 67, no. 12, pp. 1347–1357, Dec. 2001.
- [14] Y. Hu and V. Tao, "Updating solutions of the rational function model using additional control information," *Photogram. Eng. Remote Sens.*, vol. 68, no. 7, pp. 715–724, Jul. 2002.
- [15] G. Dial and J. Grodecki, "Block adjustment with rational polynomial camera models," in *Proc. ACSM-ASPRS Annu. Conv.*, Washington D.C., USA, Apr. 22–26, 2002, [CD-ROM].
- [16] C. S. Fraser and H. B. Hanley, "Bias compensation in rational functions for Ikonos satellite imagery," *Photogram. Eng. Remote Sens.*, vol. 69, no. 1, pp. 53–57, Jan. 2003.
- [17] C. S. Fraser and H. B. Hanley, "Bias compensation RPCs for sensor orientation of high-resolution satellite imagery," *Photogram. Eng. Remote Sens.*, vol. 71, no. 8, pp. 909–915, Aug. 2005.
- [18] V. Nagasubramanian, P. V. Radhadevi, R. Ramachandran, and R. Krishnan, "3D reconstruction with rational function model," *J. Indian Soc. Remote Sens.*, vol. 36, no. 1, pp. 27–35, Mar. 2008.
- [19] X. Niu, J. Wang, K. Di, J.-D. Lee, and R. Li, "Geometric modeling and photogrammetric processing of high resolution satellite imagery," in *Proc. Process. XXth Congr. ISPRS*, Jul. 12–23, 2004, [CD-ROM].
- [20] Y. Hu, V. Tao, and A. Croitoru, "Understanding the rational function model: Methods and applications," in *Proc. 20th ISPRS Congr.*, Istanbul, Turkey, 2004, pp. 1–6, DVD.
- [21] R. Li, F. Zhou, X. Niu, and K. Di, "Integration of Ikonos and QuickBird Imagery for geopositioning accuracy analysis," *Photogram. Eng. Remote Sens.*, vol. 73, no. 9, pp. 1067–1074, Sep. 2007.
- [22] R. Li, S. Deshpande, X. T. Niu, F. Zhou, K. Di, and B. Wu, "Geometric integration of aerial and high-resolution satellite imagery and application in shoreline mapping," *J. Marine Geodesy*, vol. 31, no. 3, pp. 143–159, 2008.
- [23] C. Rao, "Estimation of variance and covariance components—MINQUE theory," *J. Multivariate Anal.*, vol. 1, no. 3, pp. 257–275, Sep. 1971.
- [24] K.-R. Koch, *Parameter Estimation and Hypothesis Testing in Linear Models*, 2nd Ed. New York, NY, USA: Springer-Verlag, 1999, ch. 3, pp. 225–237.

- [25] X. Cui, *Generalized Surveying Adjustment*. Wuhan, China: Wuhan Univ. Press, 2001, ch. 3, pp. 74–106.
- [26] P. J. G. Teunissen and A. R. Amiri-Simkooei, “Least-squares variance component estimation,” *J. GeoDesy*, vol. 82, pp. 65–82, 2008.
- [27] K.-R. Koch and J. Kusche, “Regularization of geopotential determination from satellite data by variance components,” *J. GeoDesy*, vol. 76, no. 5, pp. 259–268, May 2002.
- [28] B. Li, Y. Shen, and L. Lou, “Efficient estimation of variance and covariance components: A case study for GPS stochastic model evaluation,” *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 203–210, Jan. 2011.
- [29] J. Hu, Z. W. Li, Q. Sun *et al.*, “Three dimensional surface displacements from InSAR and GPS measurements with variance component estimation,” *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 754–758, Jul. 2012.
- [30] K. Nolsøe, J. N. Nielsen, and H. Madsen, *Optimal Weights in Prediction Error and Weighted Least Squares Methods*. Lyngby, Denmark: Tech. Univ. Denmark, 2000, pp. 6–7.
- [31] V. Rodehorst and A. Koschan, “Comparison and evaluation of feature point detectors,” in *Proc. 5th Int. Symp. Turkish-German Joint Geodetic Days Geodesy Geoinformation Service Daily Life*, Berlin, Germany, 2006, pp. 1–8.
- [32] A. Heyden and K. Rohr, “Evaluation of corner extraction schemes using invariance methods,” in *Proc. 13th IEEE Int. Conf. Pattern Recog.*, Aug. 1996, vol. 1, pp. 895–899.
- [33] F. Mokhtarian and F. Mohanna, “Performance evaluation of corner detectors using consistency and accuracy measures,” *Comput. Vis. Image Understanding*, vol. 102, no. 1, pp. 81–94, Apr. 2006.
- [34] R. Zhang, A. W. Warrick, and D. E. Myers, “Variance as a function of sample support size,” *Math. Geology*, vol. 22, no. 1, pp. 107–121, Jan. 1990.
- [35] J. B. Collins and C. E. Woodcock, “Geostatistical estimation of resolution-dependent variance in remotely sensed images,” *Photogram. Eng. Remote Sens.*, vol. 65, no. 1, pp. 41–50, Jan. 1999.
- [36] Y. Gu, Z. Cao, and Y. Zhang, “Multi information based ground control points selection method,” in *Proc. IEEE IGARSS*, Jul. 2011, pp. 4344–4347.
- [37] C.-M. Kuan, *Classical Least Squares Theory*. Taipei, Taiwan: National Taiwan Univ., 2004, ch. 3, pp. 51–52. [Online]. Available: http://homepage.ntu.edu.tw/~ckuan/pdf/et01/et_Ch3.pdf
- [38] *DigitalGlobe Imagery Support Data (ISD) Documentation*, DigitalGlobe Co. Company, Longmont, CO, USA, 2011.
- [39] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI, USA: Univ. Michigan Press, 1975.
- [40] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA, USA: Addison-Wesley, 1989.
- [41] F. Valdez and P. Melin, “Comparative study of particle swarm optimization and genetic algorithms for complex mathematical functions,” *J. Autom., Mobile Robot. Intell. Syst.*, vol. 2, no. 1, pp. 43–51, 2008.
- [42] C. E. Zoumas, A. G. Bakirtzis, J. B. Theocharis, and V. Petridis, “A genetic algorithm solution approach to the hydrothermal coordination problem,” *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1356–1364, Aug. 2004.
- [43] G. Leng, T. M. McGinnity, and G. Prasad, “Design for self-organizing fuzzy neural networks based on genetic algorithms,” *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 6, pp. 755–766, Dec. 2006.
- [44] T. Celik, “Change detection in satellite images using a genetic algorithm approach,” *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 386–390, Apr. 2010.
- [45] Y. Zanjireh, A. H. Rezaie, and H. Amindavar, “Multi component signal decomposition based on chirplet pursuit and genetic algorithms,” *Appl. Acoust.*, vol. 74, no. 12, pp. 1333–1342, Dec. 2013.



Yanfeng Gu (M'09) was born in Jiamusi, China, in 1977. He received the B.E. degree from the Dalian University of Technology, Dalian, China, in 1999 and the M.E. and Ph.D. degrees from the Harbin Institute of Technology, Harbin, China, in 2001 and 2005, respectively.

He is currently a Full Professor with the School of Electrical and Information Engineering, Harbin Institute of Technology, Harbin. Since 2011, he has been a Visiting Scholar in the University of California, Berkeley, CA, USA. His research interests include advanced signal processing, machine learning, sparse representation, and their applications to image processing, particularly remote sensing and medical imaging. He has published more than 60 peer-reviewed papers and four book chapters, and he is the inventor or coinventor of seven patents.



Zhimin Cao received the B.S. degree in electronic information engineering from Daqing Petroleum Institute, Daqing, China, in 2006 and the M.S. degree in information and communication engineering from the Harbin Institute of Technology, Harbin, China, in 2008, where he is currently working toward the Ph.D. degree in the School of Electronic and Information Engineering.

He is also a Lecturer with the School of Electronic Science, Northeast Petroleum University, Daqing, China.



Ye Zhang received the B.S. degree in communication engineering and the M.S. and Ph.D. degrees in communication and electronic system from the Harbin Institute of Technology (HIT), Harbin, China, in 1982, 1985, and 1996, respectively.

In 1985, he joined HIT as a Teacher. Between 1998 and 1999, he was a Visiting Scholar at the University of Texas at San Antonio, San Antonio, TX, USA. His research interests are remote sensing hyperspectral image analysis and processing, image/video compression and transmission, and multi-source information collaboration processing and applications. Currently, he is a Professor and Doctoral Supervisor in information and communication engineering. He is the Director of the Institute of Image and Information Technology in the School of Electronic and Information Engineering, HIT.