Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography
Vol. 37, No. 2, 55-62, 2019
https://doi.org/10.7848/ksgpc.2019.37.2.55

ISSN 1598-4850(Print)
ISSN 2288-260X(Online)
Original article

Conjugate Point Extraction for High-Resolution Stereo Satellite Images Orientation

Oh, Jae Hong [1] · Lee, Chang No [2]

Abstract

The stereo geometry establishment based on the precise sensor modeling is prerequisite for accurate stereo data processing. Ground control points are generally required for the accurate sensor modeling though it is not possible over the area where the accessibility is limited or reference data is not available. For the areas, the relative orientation should be carried out to improve the geometric consistency between the stereo data though it does not improve the absolute positional accuracy. The relative orientation requires conjugate points that are well distributed over the entire image region. Therefore the automatic conjugate point extraction is required because the manual operation is labor-intensive. In this study, we applied the method consisting of the key point extraction, the search space minimization based on the epipolar line, and the rigorous outlier detection based on the RPCs (Rational Polynomial Coefficients) bias compensation modeling. We tested different parameters of window sizes for Kompsat-2 across track stereo data and analyzed the RPCs precision after the bias compensation for the cases whether the epipolar line information is used or not. The experimental results showed that matching outliers were inevitable for the different matching parameterization but they were successfully detected and removed with the rigorous method for sub-pixel level of stereo RPCs precision.

Keywords : Satellite Image, RPCs, Stereo, Relative Orientation, Bias-compensation, Quasi-GCPs

1. Introduction

It is necessary to generate epipolar resampled images to extract precise topographic information from high-resolution stereo satellite data. Conventionally, GCPs (Ground Control Points) are used to improve the sensor modeling accuracy as Fraser and Ravanbakhsh (2009) showed that one GCP can be helpful to achieve one-pixel level of positional accuracy for GeoEye-1 data through RPCs (Rational Polynomial Coefficients) bias compensation approach. But the results can be different depending on the satellite data.

Precision of the sensor modeling between the stereo data should be secured even for target sites where the access is limited or any other reference data is not available. Without any ground control, it is not feasible to improve absolute positional accuracy of the data but it is possible to improve the relative accuracy. In other words, the consistency between the RPCs of the stereo data can be improved even without any GCP. This procedure is called the relative orientation. In the approach, the conjugate points are required to generate quasi-GCPs by the space intersection, followed by the reprojection of the quasi-GCPs back to the stereo data to analyze the errors in the image space. The relative orientation is the bias-compensation of the RPCs by minimizing those errors (de

Received 2019. 01. 14, Revised 2019. 03. 08, Accepted 2019. 04. 24
1) Member, Dept. of Civil Engineering, Korea Maritime and Ocean University (E-mail: jhoh@kmou.ac.kr)
2) Corresponding Author, Member, Dept. of Civil Engineering, Seoul National University of Science and Technology (E-mail: changno@seoultech.ac.kr)

Franchis *et al.*, 2004; Ghuffar, 2016; Gong and Fritsch, 2017; Oh and Lee, 2018).

The main ingredient of the relative orientation is the conjugate point set that are well distributed over the entire image region without outliers. The automatic conjugate point extraction is required because the manual operation is labor-intensive. In this study, we applied the method consisting of the key point extraction, the search space minimization based on the epipolar line, and the rigorous outlier detection based on the RPCs bias compensation modeling. Key point extraction is carried out using Harris corner detector (Harris and Stephens, 1988) which performs well with simple implementation. The similarity computation is based on the normalized cross correlation with Fourier transform (Oh *et al.*, 2011).

The focus of the study is minimization of the search space and the outlier detection. The base search space is defined by the iterative ground-space projection using RPCs with the ground height range. But the projection based on RPCs has errors depending on the satellite with that the search space should be expanded. Another important information in the stereo data is the epipolar geometry. We generated epipolar line corresponding to a key point and we used the closeness to the line for another threshold to filter out mismatches. It is possible that some match points still survive the search space minimization. Therefore we applied the rigorous data snooping method (Baarda, 1968) to detect and remove the points. We tested different parameters for the conjugate point extraction for Kompsat-2 across track stereo data.

This paper is structured as follows: the proposed method will be explained in Section 2; experimental results are presented in Section 3; the summary and conclusion follow in Section 4..

2. Methodology

The proposed method includes the key point extraction, the search space minimization based on the epipolar line, and the rigorous outlier detection based on the RPCs bias compensation modeling.

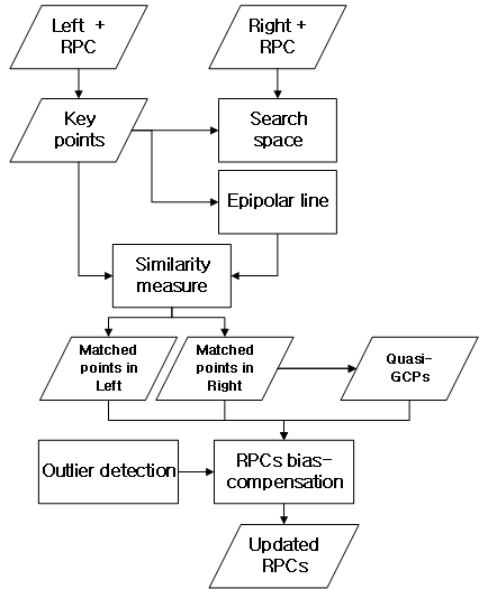


Fig. 1. Flow chart of the study

2.1 Key points extraction

Harris Corner Detector is a corner detection operator to extract key corner points (Harris and Stephens, 1988) by improving Moravec's corner detector. Given input image *I*, the derivatives are computed for the structure tensor as shown in Eq. (1) that the Harris response *R* is computed from as Eq. (2). Then, the local maxima is sought as corners within a 3 by 3 window. Finally, a percent threshold of the local maxima is used to extract the corner location.

$$M=\begin{bmatrix}\sum_W I_x^2 & \sum_W I_x I_y \\ \sum_W I_x I_y & \sum_W I_y^2\end{bmatrix} \tag{1}$$

$$R=\det M-k(traceM)^2=\lambda_1\lambda_2-k(\lambda_1+\lambda_2)^2 \tag{2}$$

M is the structure tensor, I_x, I_y are the derivatives of the input image *I*, *W* is the window size over the area x, y, *R* is the corner response measure, *k* is an empirically determined constant, λ_1, λ_2 are eigenvalues of *M*.

2.2 Search space minimization

As depicted in Fig. 2, a point *p* is extracted in the left image

and the conjugate point q would be located somewhere in the right image. Searching q in the entire image should be not only time-consuming, but it increases the possibility of mismatches. Therefore it is important to minimize the search space. In the study, we utilized two geometric constraints; the ground elevation range and the closeness to the epipolar line. The epipolar line is generated using the projection of a left image point into the ground space with certain elevation range, followed by the projection of the ground points into the right image space (Oh *et al.*, 2011).

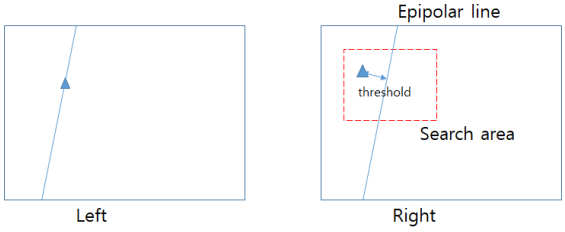


Fig. 2. Search space minimization

First, we have the information on the ground elevation range which can define the search space in the right image. In other words, p in the left image is projected to the ground space of certain height range, and then projected into the right image space for the search space. The positional error of the projection should be considered.

To this end, we need the sensor modeling of high-resolution satellite data as presented in Eq. (3) that enables the corresponding image coordinates computation given ground point coordinates.

$$l = RPC_l(\phi,\lambda,h) = \frac{a^T u}{b^T u} L_S + L_0, \qquad (3)$$
$$s = RPC_s(\phi,\lambda,h) = \frac{c^T u}{d^T u} S_S + S_0$$

with

$$u = [1 \;\; V \;\; U \;\; W \;\; VU \;\; VW \;\; UW \;\; V^2 \;\; U^2 \;\; W^2 \;\; UVW \;\; V^3 \;\; VU^2 \;\; VW^2 \;\; V^2U \;\; U^3 \;\; UW^2 \;\; V^2W \;\; U^2W \;\; W^3]^T$$

$$U = \frac{\phi-\phi_0}{\phi_S}, V = \frac{\lambda-\lambda_0}{\lambda_S}, W = \frac{h-h_0}{h_S}, Y = \frac{l-L_0}{L_S}, X = \frac{s-S_0}{S_S}$$

$$a = [a_1 \; a_2 \ldots a_{20}]^T, b = [b_1 \; b_2 \ldots b_{20}]^T, c = [c_1 \; c_2 \ldots c_{20}]^T, d = [d_1 \; d_2 \ldots d_{20}]^T$$

where X, Y normalized image coordinates, U, V, W normalized ground coordinates, ϕ, λ, h latitude, longitude, ellipsoidal height, l, s image coordinates, $\phi_0, \lambda_0, h_0, S_0, L_0$ offsets for latitude, longitude, ellipsoidal height, image sample, line), and $\phi_S, \lambda_S, h_S, S_S, L_S$ scales.

The computation of the ground coordinates from image coordinates can be carried out using Eq. (4). In the equation, the normalized horizontal ground coordinates U, V are iteratively computed.

$$\begin{bmatrix} V \\ U \end{bmatrix} = \begin{bmatrix} V^0 \\ U^0 \end{bmatrix} + \begin{bmatrix} dV \\ dU \end{bmatrix} \qquad (4)$$

$$\begin{bmatrix} dV \\ dU \end{bmatrix} = \begin{bmatrix} \frac{\partial Y}{\partial V} & \frac{\partial Y}{\partial U} \\ \frac{\partial X}{\partial V} & \frac{\partial X}{\partial U} \end{bmatrix}^{-1} \begin{bmatrix} Y - Y^0 \\ X - X^0 \end{bmatrix}$$

$$\frac{\partial Y}{\partial V} = \frac{\partial Y}{\partial u^T}\frac{\partial u}{\partial V}, \; \frac{\partial Y}{\partial U} = \frac{\partial Y}{\partial u^T}\frac{\partial u}{\partial U},$$
$$\frac{\partial X}{\partial V} = \frac{\partial X}{\partial u^T}\frac{\partial u}{\partial V}, \; \frac{\partial X}{\partial U} = \frac{\partial X}{\partial u^T}\frac{\partial u}{\partial U}$$

$$\frac{\partial Y}{\partial u^T} = \frac{(b^T u)a^T - (a^T u)b^T}{(b^T u)^2},$$
$$\frac{\partial X}{\partial u^T} = \frac{(d^T u)c^T - (c^T u)d^T}{(d^T u)^2}$$

$$\frac{\partial u}{\partial U} = [0 \; 0 \; 1 \; 0 \; V \; 0 \; W \; 0 \; 2U \; 0 \; VW \; 0 \; 2VU \; 0 \; V^2 \; 3U^2 \; W^2 \; 0 \; 2UW \; 0]^T,$$
$$\frac{\partial u}{\partial V} = [0 \; 1 \; 0 \; 0 \; U \; W \; 0 \; 2V \; 0 \; 0 \; UW \; 3V^2 \; U^2 \; W^2 \; 2VU \; 0 \; 0 \; 2VW \; 0 \; 0]^T$$

where, U^0, V^0 are the initial normalized ground coordinates.

2.3 Normalized cross correlation

Since the geometric and spectral difference between the stereo satellite images are rather small and the epipolar geometry is well established, there is no need for SIFT (Scale Invariant Feature Transform) or other sophisticated and computationally expensive feature matching method. The well known normalized cross-correlation can be used for the stereo image matching using Eq. (5). Normalized cross correlation is an area-based matching. The image intensity within $w \times w$ window area is used to measure the similarity. $w \times w$ area around a point p in the left image is selected and compared with $w \times w$ areas in the right image for matched point q.

$$NCC=\frac{\sum_{i=1}^{w}\sum_{j=1}^{w}\left(L_{ij}-\overline{L}\right)\left(R_{ij}-\overline{R}\right)}{\sqrt{\left[\sum_{i=1}^{w}\sum_{j=1}^{w}\left(L_{ij}-\overline{L}\right)^2\right]\left[\sum_{i=1}^{w}\sum_{j=1}^{w}\left(R_{ij}-\overline{R}\right)^2\right]}} \quad (5)$$

L is the transformed image patch in the left epipolar image, and R is the right image patch within the search region, both are in the size of $w\times w$; $\overline{L}, \overline{R}$ average of all intensity value in the patches.

A problem is the high computational cost such that FFT (Fast Fourier Transformation) can be utilized (Oh *et al.*, 2011). $\overline{R}$ can be computed using FFT over the search area in the right image. Decomposition of the numerator of Eq. (5) gives Eq. (6) where each part can be computed using FFT.

$$\sum_{i=1}^{w}\sum_{j=1}^{w}\left(L_{ij}-\overline{L}\right)\left(R_{ij}-\overline{R}\right)=\sum_{i=1}^{w}\sum_{j=1}^{w}R_{ij}\left(L_{ij}-\overline{L}\right)-\overline{R}\sum_{i=1}^{w}\sum_{j=1}^{w}\left(L_{ij}-\overline{L}\right) \quad (6)$$

The decomposition of the denominator yields Eq. (7). $\left[\sum_{i=1}^{w}\sum_{j=1}^{w}\left(L_{ij}-\overline{L}\right)^2\right]$ is computed once for each p, the other terms can be computed using FFT.

$$\begin{aligned}&\left[\sum_{i=1}^{w}\sum_{j=1}^{w}\left(L_{ij}-\overline{L}\right)^2\right]\left[\sum_{i=1}^{w}\sum_{j=1}^{w}\left(R_{ij}-\overline{R}\right)^2\right]=\\&\left[\sum_{i=1}^{w}\sum_{j=1}^{w}\left(L_{ij}-\overline{L}\right)^2\right]\left[\sum_{i=1}^{w}\sum_{j=1}^{w}R_{ij}^2-2\overline{R}\sum_{i=1}^{w}\sum_{j=1}^{w}R_{ij}-\sum_{i=1}^{w}\sum_{j=1}^{w}\overline{R}^2\right]\end{aligned} \quad (7)$$

2.4 RPCs bias-compensation

RPCs from the image providers include errors such that the coefficients should be improved by compensating the biases. Eq. (8) shows the RPCs bias-compensation based on the affine modelings.

$$\begin{aligned}l' &= l+A_0+A_1l+A_2s\\ s' &= s+B_0+B_1l+B_2s\end{aligned} \quad (8)$$

where A, B are coefficients for shift, drift and systematic errors, l, s are computed image coordinates (reprojected coordinates) and l', s' are bias compensated image coordinates (the matched point coordinates).

2.5 Outlier detection

The image matching should include some mismatches. Therefore in this study we utilized the outlier detection based on Baarda's data snooping (Baarda, 1968). RANSAC is a technique to estimate parameters of a model through iterations from a large set of observations that contain outliers, but this method rigorously examines each observation in a small set of observations. The observation equation can be formed as Eq. (9) and the quadratic form of residuals after solving the equation can be computed as Eq. (10). Eq. (8) is used to set up Eq. (9) where unknown vectors ξ includes the coefficients A, B.

$$y=M\xi+e,\quad \sim e(0,\sigma_0^2P^{-1}) \quad (9)$$

where, y is observation vector, M is the design matrix, ξ is the unknown vectors (the coefficients in case of Eq. (8)), e is the error vector with standard deviation of unit vector σ_0^2 and weight matrix P.

$$\Omega=\tilde{e}^{T}P\tilde{e} \quad (10)$$

Inclusion of an outlier in the observation equation extends Eq. (9) to Eq. (11) and the quadratic form of residuals becomes Eq. (12).

$$y=M\xi+\eta_j\xi_0^{(j)}+e,\quad \sim e(0,\sigma_0^2P^{-1}) \quad (11)$$

where, η_j is the outlier value, $\xi_0^{(j)}=\left[0\ 0\ \cdots\ \underbrace{1}_{j-th}\ \cdots\ 0\ 0\right]^T$

$$R_j=\left(\xi_0^{(j)}\right)^2\left(\eta_j^TPQ_{\tilde{e}}P\eta_j\right) \quad (12)$$

The test statistics for the outlier detection is carried out as Eq. (13).

$$\begin{aligned}&T=\frac{R_j(n-m-1)}{\Omega-R_j}\ \text{and}\\&\begin{cases}T_j\le F_\alpha(1,n-m-1):Accept\ H_0(no\ \ outlier)\\ T_j>F_\alpha(1,n-m-1):Reject\ H_0(no\ \ outlier)\end{cases}\end{aligned} \quad (13)$$

3. Experiment

3.1 Test data

We tested two Kompsat-2 images (each covers 15x15 km) over Daegu in South Korea with one meter level spatial resolution. The test area consists of urban, agriculture, and mountain features. The images were acquired in Dec.

31, 2009 and Feb. 07, 2010, respectively that they show radiometric dissimilarity due to different acquisition dates. The convergence angle between the two images is 19.75 degrees which is relatively smaller than the optimal range 30~50 degrees.

Table 1. Specification of tested data

Product level	Level 1R	
Acquisition date	2009-12-31/ 2010-02-07	
ID	#1	#2
Image size [pixels]	Sample 15,000 Line 15,500	Sample 15,000 Line 15,500
Incidence/ Azimuth	14.15°/256.17°	5.8°/91.7°
GSD (col/row)	0.77/0.83m	0.77/0.82m

3.2 Key point extraction

Harris corner point extraction was applied to image #1 using $k = 0.06$ after the Gaussian smoothing. We used a loosened local maxima threshold such as 10% to extract key point locations. Therefore, total 173,200 dense key points were extracted well distributed over the entire image region. Much dense points are extracted over the urban area. Smaller thresholds should produce much less and more key points in the rural and urban area, respectively leading to the biased point distribution. Therefore we adopted the approach of a loosened threshold with sampling the result for better key point distribution on the image. For the study, we sampled by selecting every 100th point for 1,732 points as shown in Fig. 3 that shows the well-distribution of the points.

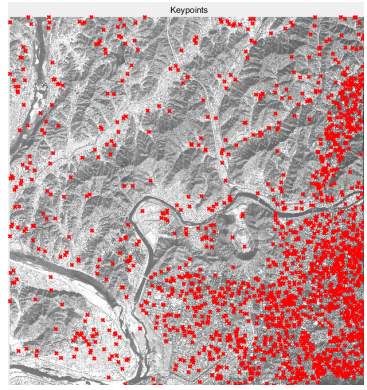


Fig. 3. Tested Kompsat-2 stereo data and the extracted keypoints (sampled points every 100th)

3.3 Conjugate point matching

The ground height range was computed as 0~332 meters from the RPCs that are used for the iterative image-ground projections. We established the search space corresponding to the height range with margins of RPCs positional errors. Kompsat-2 data have a reported horizontal accuracy of 80 meters at 90% significance (Seo *et al.*, 2008) such that we established the search space 300 pixels wider than the boundary determined by the ground height range. The closeness to the epipolar line was similarly set to 300 pixels.

We applied different window sizes for the key point matching such as 15, 21, 25, 31, 35, and 41 pixels. In addition, we tested two cases whether the epipolar line information is used or not. The number of matches is plotted in Fig. 4 for both cases. Fig. 4 shows that the number of matched points decreases from 1643 to 1360 for the case of not using epipolar information and the number of matched points ranges 1325~1365 points for the case of using epipolar information. From the difference we can assume the discrepancy occurs because some points are filtered outside the epipolar threshold. Note that most points showed the distance to the epipolar line was about 106 pixels.

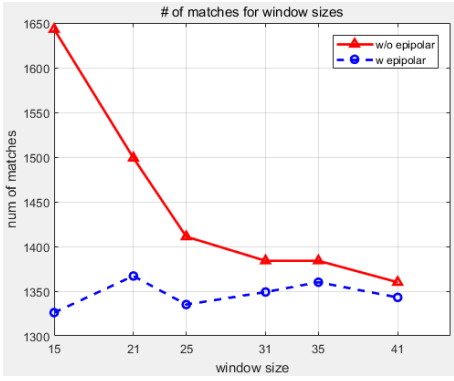
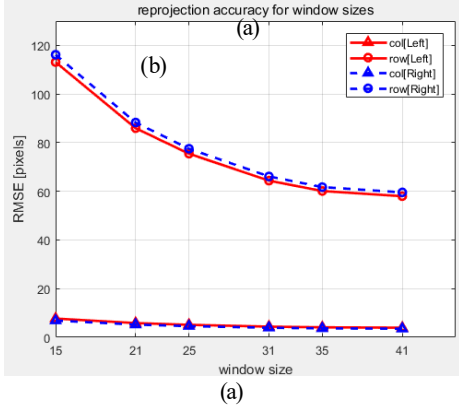
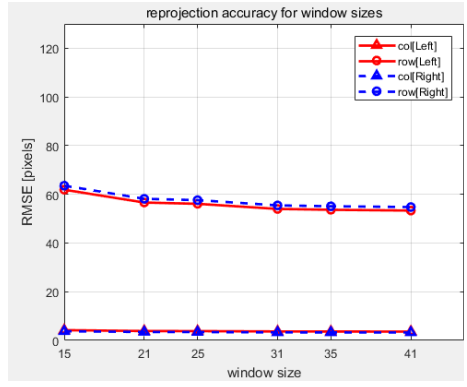


Fig. 4. The number of matches as the window size increases whether or not using epipolar information

In Fig. 5, the reprojection accuracy is shown that is $l'-l, s'-s$ of Eq. (8) where l, s are the reprojected coordinates and l', s' are the matched point coordinates. In both cases of epipolar information, significant errors are observed along the row coordinates both in #1 and #2 images while smaller errors are shown along the column coordinates. But using epipolar information showed better reprojection accuracy. Note that the errors show the inconsistency between the stereo data set. In Fig. 6 the error plot at the conjugate points for image #1 is presented for the two cases. We can observe that large error points in Fig. 6(a) have been removed in Fig. 6(b).

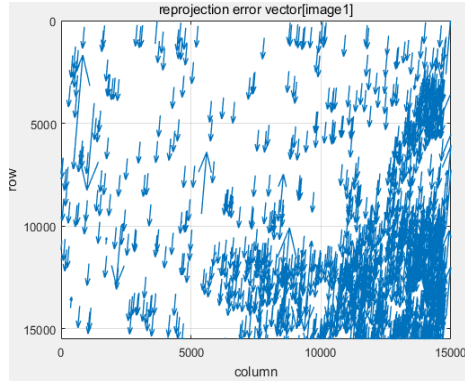


(a)

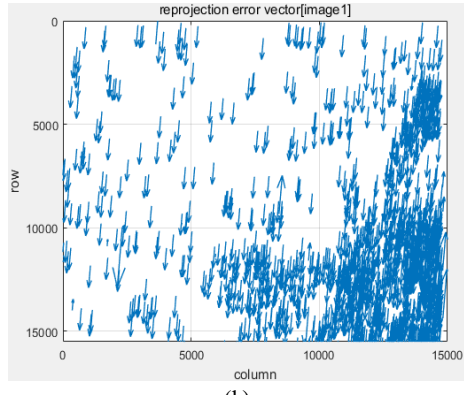


(b)

Fig. 5. The reprojection accuracy matches as the window size increases. (a) without epipolar information, (b) with epipolar information



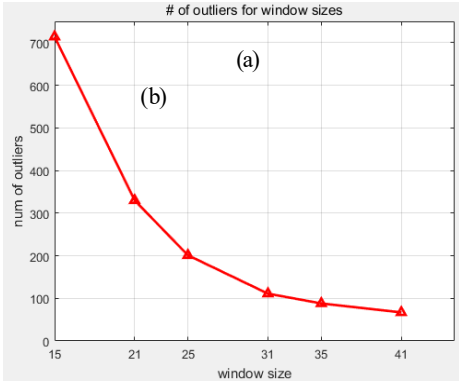
(a)



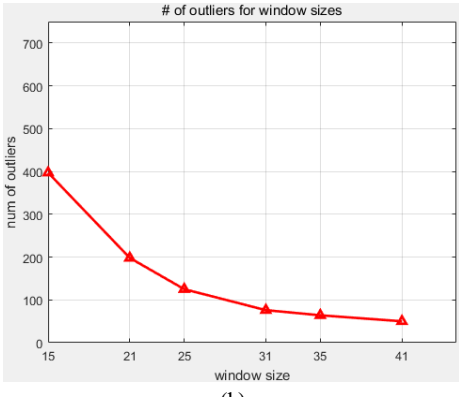
(b)

Fig. 6. The reprojection error plot as the window size increases for image 1 (a) without epipolar information, (b) with epipolar information

Next, the reprojection errors were modeled using the bias compensation model Eq. (8). For outlier detection, we set the confidence level as 99.99%. The number of detected outliers are plotted in Fig. 7 and the reprojection accuracy after removing the outliers are shown in Fig. 8. The number of outliers significantly decreases as the window size increases. The case of not using epipolar information showed that more outliers were detected in the process. It is because points outside the epipolar line threshold were detected in this step as matching outliers. After removing the outlier points, the reprojection errors dropped to less than one pixel along both row and line directions. The result shows that the rigorous outlier detection was helpful even if the epipolar information was not used.



(a)



(b)

Fig. 7. Outlier detection as the window size increases. (a) without epipolar information, (b) with epipolar information

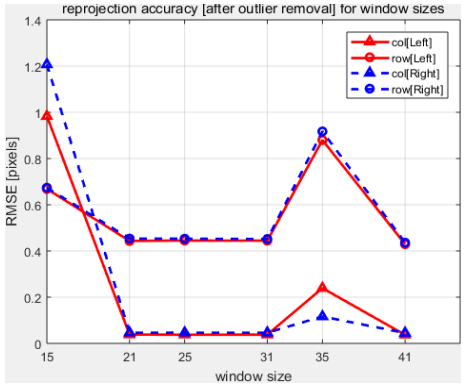


Fig. 8. The reprojection accuracy after removing outliers as the window size increases (same results for both epipolar cases)

The point distribution after removing outliers for matching window size 41 pixels are shown in Fig. 9 that shows survived and removed points with colors.

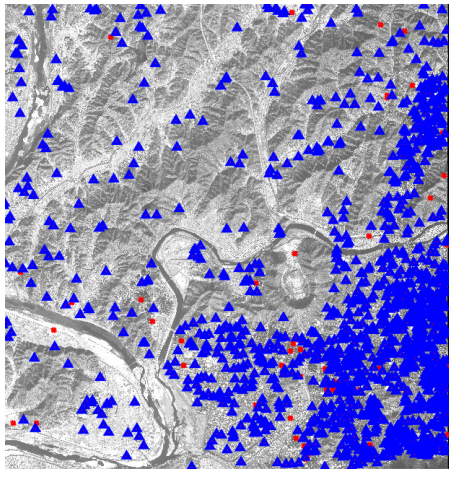
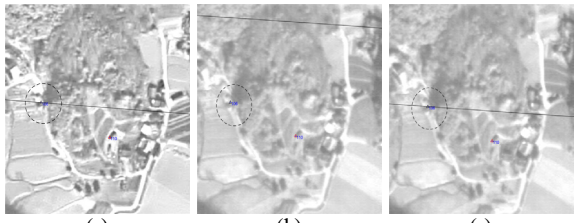


Fig. 9. Extracted conjugate points (blue) and removed outliers (red) case of window size 41 pixels

Based on the relative orientation, the reprojection errors were modeled and compensated for the compensated RPCs. Fig. 10 shows the epipolar line discrepancy before and after the bias compensation. Fig. 10(b) shows that the conjugate point is far from the epipolar line in the case of the original RPCs. This shows the geometric inconsistency between the stereo data. However, after the bias compensation, the

conjugate point is well located on the epipolar line as can be shown in Fig. 10(c). In other words, the consistency between the stereo data has been improved.



(a) (b) (c)

Fig. 10. Epipolar line before and after RPCs bias-compensation (a) a point with epipolar line in #1 image, (b) corresponding epipolar line in #2 image (before RPCs compensation), (c) corresponding epipolar line in #2 image (after RPCs compensation)

4. Conclusion

This study investigated the automatic conjugate point extraction method including the key point extraction, the search space minimization based on the epipolar line, and the rigorous outlier detection based on the RPCs bias compensation modeling. The different matching window sizes were tested and the result showed that the number of outliers significantly decreases as the window size increases and the reprojection errors dropped to less than one pixel along both row and line directions after removing the outlier points. The epipolar information was helpful to filter some mismatches in the matching process even though the rigorous outlier detection could detect most mismatches. The outliers in the image matching was inevitable for the different matching parameterization but they were successfully detected and removed with the rigorous method for sub-pixel level of stereo RPCs precision.

Acknowledgements

This research was supported by a grant(19SIUE-B148326-02) from CAS 500-1/2 Image Acquisition and Utilization Technology Development Program funded by Ministry of Land, Infrastructure and Transport of Korean government

References

Baarda, W. (1968), *A Testing Procedure for Use in Geodetic Networks*, Netherlands Geodetic Commission, Publications on Geodesy, New Series, Delft.

de Franchis, C., Meinhardt-Llopis, E., Michel, J., Morel, J.M., and Facciolo, G. (2014), An automatic and modular stereo pipeline for pushbroom images, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information* Sciences, 5-7 September 2014, Zurich, Switzerland, pp. 49-56.

Fraser, C.S. and Ravanbakhsh, M. (2009), Georeferencing accuracy of GeoEye-1 imagery, *Photogrammetric Engineering & Remote Sensing*, Vol. 75, No. 6, pp. 634-638.

Ghuffar, S. (2016), Satellite stereo based digital surface model generation using semi global matching in object and image space, *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 12-19 July, Prague, Czech Republic, pp. 63-68.

Gong, K. and Fritsch, D. (2017), Relative orientation and modified piecewise epipolar resampling for high resolution satellite images, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 6–9 June 2017, Hannover, Germany, pp. 579-586.

Harris, C. and Stephens, M. (1988), A combined corner and edge detector, *Proceedings of the Alvey Vision Conference*, 31 August – 2 September 1988, Manchester, United Kingdom, pp. 147–151.

Oh, J.H., Toth, C., and Grejner-Brzezinska, D.A. (2011), Automatic georeferencing of aerial images using stereo high-resolution satellite images, *Photogrammetric Engineering & Remote Sensing*, Vol. 77, No. 11, pp. 1157-1168.

Oh, J.H. and Lee, C.N. (2018), Relative orientation of stereo images for epipolar image resampling without any ground control point, *The 39th Asian Conference on Remote Sensing*, 15-19 October, Kuala Lumpur, Malaysia.

Seo, D.C., Yang, J.Y., Lee, D.H., Song, J.H., and Lim, H.S. (2008), Kompsat-2 direct sensor modeling and geometric calibration/validation, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3-11 July, Beijing, China, Vol. XXXVII, Part B1.