Progress In Electromagnetics Research M, Vol. 34, 55–62, 2014

A Quantitative Evaluation Method of Ground Control Points for Remote Sensing Image Registration

Wenting Ma*, Jian Yang, Xia Ning, and Wei Gao

Abstract—Ground control point (GCP) extraction is an essential step in automatic registration of remote sensing images. However the lack of quantitative and objective methods for analyzing the GCP quality becomes the bottleneck that prevents the broad development of automatic image registration. Although several measurements for evaluating the number, accuracy and distribution of GCPs have been proposed in recent years, some of them are redundant and the evaluation of dispersion is not effective enough. In this paper, a method for an objective and quantitative evaluation of GCP quality is proposed. The proposed method consists of three parts: measurement calculation, cost function calculation and final validation. In the first part, two new measurements are proposed to evaluate the number, dispersion and isotropy of GCPs, and the root mean square of GCP residuals using leave-one-out method (RMS_{loo}) is used to evaluate the accuracy. In the second part, seed cost functions are utilized to transform the measurements into a limited value range as well as to be desired on the ascending direction. Subsequently, all the seed cost functions are combined by a total cost function to provide an integrated evaluation. In the third part, the GCP scenario is validated by the accepted threshold depending on the value of the total cost function. To evaluate the performance of the proposed method, experiments using four typical emulated scenarios of GCP distribution and two sets of real GCPs in SAR images are considered. The results demonstrate that the proposed GCP evaluation method performs more effectively than the existing methods, especially in the evaluation of dispersion quality.

1. INTRODUCTION

Remote sensing images are required to be registered before many applications are implemented [1], such as change detection [2], data fusion [3], target recognition [4], target detection [5], and so on. Ground control point (GCP) extraction is an essential step in the registration process [6]. However, in the past several decades, only the root mean square (RMS) of GCP residuals and the number of GCPs are broadly used to evaluate the GCP quality [7]. The evaluation of GCP distribution quality is usually carried out manually which inevitably leads to low efficiency and sometimes unfair judgement, or even skipped which usually leads to unrecognition of ill-distributed GCPs and undesired registration results. Therefore, diversified and adequate measurements for an objective and quantitative evaluation of GCP quality are greatly needed. On the other hand, the lack of criterions to combine the measurements for providing a final validation makes the evaluation process away from automatic applications. Whereas, the automatic validation has great significance for automatic registration of remote sensing images as well as dealing with a large number of datasets.

In recent years, several researches about evaluation of GCP quality have been reported. Buiten and van Putten [8] did a lot of researches on analysis and testing of GCP residuals, but it was far away from an integral and adequate evaluation method for the lack of diversified quality issues. Gonçalves et al. [7] proposed a comprehensive method containing measurements for evaluating the number, accuracy and distribution of GCPs, as well as the distribution of GCP residuals. Moreover, he also used a cost

Received 29 September 2013, Accepted 28 November 2013, Scheduled 1 December 2013
* Corresponding author: Wenting Ma (mwt09@mails.tsinghua.edu.cn).
The authors are with the Department of Electronic Engineering, Tsinghua University, Rohm Building 10-101, Beijing, China.

function combining these measurements to provide a final validation. The evaluation method proposed by Gonçalves (Gonçalves method for short) is considered to be the most comprehensive and effective method for evaluating GCP quality so far, and has been broadly applied in the registration of remote sensing images [1, 9–11]. However, there is redundancy between some of the measurements in the Gonçalves method, such as RMS_{all} (the root mean square of residuals), RMS_{loo} (the root mean square of residuals using leave-one-out method) and B_{pp} (the bad point proportion), P_{quad} (the measurement of how equally the residuals distribute across the quadrants) and S_{kew} (the measurement of how much the residuals present a pronounced scatter along an axis). In addition, some are not valid enough, such as S_{cat} (the measurement of GCP dispersion), which is nearly around 1 in most scenarios and fails to evaluate the GCP dispersion [1, 9]. Therefore, the method containing effective and simplified measurements for evaluating GCP quality is greatly in need.

In this paper, an evaluation method consisting of RMS_{loo} and two new measurements for evaluating the number, dispersion and isotropy of GCPs is proposed. The method can evaluate the quality of GCPs at almost the same aspects as the Gonçalves method but using fewer measurements. Subsequently, seed cost functions are utilized to transform the measurements into a limited value range as well as to be desired on the ascending direction, and then combined by a total cost function to provide an integrated evaluation. According to the value of the total cost function, the GCP scenario is validated by the accepted threshold. In the paper, the accepted thresholds of all the measurements and cost functions are suggested. The assumption in this paper is that the available information is reduced to the minimum without any metadata, which is the same as in [7]. Experiments using emulated and real sets of GCPs demonstrate the effectiveness of the proposed method in evaluating the GCP quality, especially dispersion quality, and that the validations by the proposed method are more reasonable than the existing methods.

2. THE PROPOSED METHOD

Considering (x_i, y_i) as (pixel, line) and (x_i', y_i') as the geographical coordinates of the i-th GCP, the predicted geographical coordinates of (x_i, y_i), denoted by (x_i'', y_i''), and the residuals (rx_i, ry_i), can be written as Eq. (1a) and Eq. (1b) respectively.

$$\begin{cases} x_i'' = f(x_i, y_i) \\ y_i'' = g(x_i, y_i) \end{cases} \tag{1a}$$

$$\begin{cases} rx_i = x_i'' - x_i' \\ ry_i = y_i'' - y_i' \end{cases} \tag{1b}$$

where $f(x, y)$ and $g(x, y)$ are transformation functions, which can be estimated by the N pairs of GCPs using the least square (LS) algorithm (assuming that there are N pairs of GCPs).

The procedure of the proposed evaluation method is depicted in Fig. 1. In the procedure, there are three parts: measurement calculation, cost function calculation and final validation.

2.1. N_{class}

The number and dispersion of GCPs are important aspects related to the estimation of transformation functions. The number of GCPs should be larger than a minimum value related to the order of transformation functions, and should be as large as possible. Meanwhile, GCPs need to be far enough from each other in order to avoid the badly-conditioned matrix in the LS algorithm which may lead to undesired transformation functions as well as undesired registration results.

N_{class} is proposed to evaluate the two aspects of GCPs, and combines the roles of N_{red} and S_{cat} in [7]. Assuming that the minimum accepted distance between GCPs is $d_{\min}$, each GCP is considered to play the same role with all the GCPs around it within $d_{\min}$, and the contributions of them are considered to be equivalent. Depending on the minimum accepted distance, GCPs are clustered into N_{class} categories.

In this process, many kinds of clustering methods can be used, such as some simple ones (the nearest and farthest distance clustering methods, the threshold segmentation clustering method and so

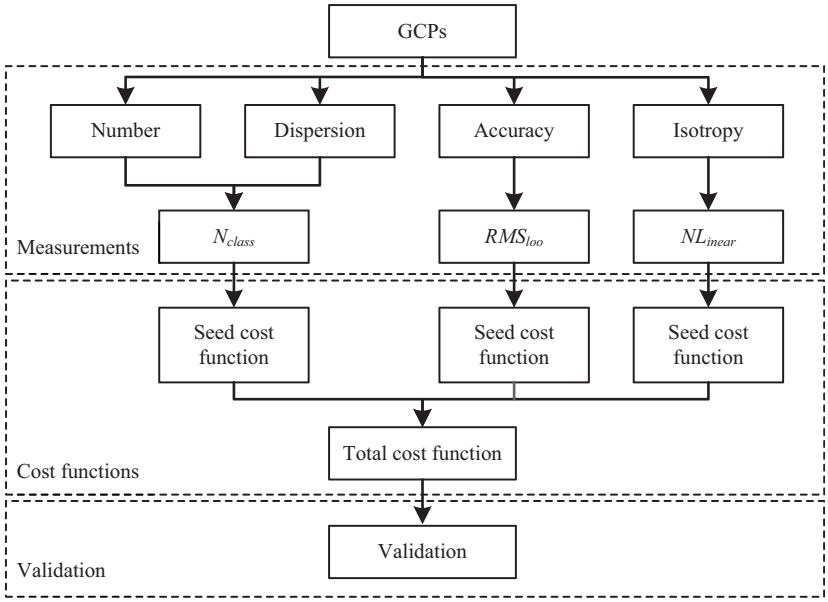

Figure 1. Procedure of the proposed evaluation method.

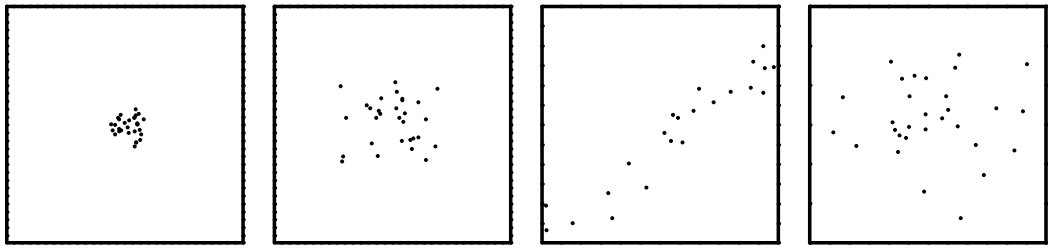

Figure 2. Representative emulated scenarios of GCP distribution in images of 300×300 pixel size.

Table 1. N_{class} of GCP scenarios in Fig. 2 using the farthest distance clustering method.

$d_{\min}$	N_{class}			
	(a)	(b)	(c)	(d)
5	22	27	29	30
10	9	23	26	28
20	3	16	20	23
40	1	9	10	15

on) which have low complexity and low computational cost but cannot obtain perfect clustering results, and some complex methods (the K-means clustering method [12], the fuzzy clustering method [13], the graph theory-based clustering methods [14] and so on) which can obtain better clustering results but have higher computational cost. When calculating N_{class}, we concern about the number of clustering centers, instead of the accuracy of clustering results. Therefore the simple clustering methods are enough for N_{class} calculation.

For a given $d_{\min}$, GCP scenarios with large number and dispersion can get a large N_{class}. And for a given GCP scenario, different values of $d_{\min}$ will lead to different values of N_{class}. To illustrate the influence of $d_{\min}$ on N_{class}, four typical scenarios of 30-GCP distribution (in Fig. 2) are emulated. Obtained values of N_{class} using the farthest distance clustering method are presented in Table 1, when the values of $d_{\min}$ are 5, 10, 20, 40. In practical, $d_{\min}$ is set depending on the requirement.

2.2. RMS_{loo}

The accuracy of GCPs is greatly related to the accuracy of transformation functions and registration results. RMS_{loo} [7] is proposed by Gonçalves to evaluate this aspect of GCPs, which means the root mean square of GCP residuals using leave-one-out (LOO) method and is defined as Eq. (2). In the equation, the residuals of each GCP pair are calculated using the transformation functions estimated by the other $N-1$ GCP pairs, instead of all the GCP pairs as in RMS_{all} [7]. Compared with RMS_{all}, RMS_{loo} has a special advantage that it can recognize the situation with a large value when the residuals of most GCP pairs are very small but the residuals of some are large. Therefore, a small RMS_{loo} means that all the residuals of GCPs are small.

$$RMS_{loo} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(rx_i^2 + ry_i^2\right)} \tag{2}$$

2.3. NL_{inear}

In some situations, such as in Fig. 2(c), GCPs are isotropous (presenting a pronounced distribution around an axis), which also usually leads to an ill-conditioned matrix during the transformation function calculation. The NL_{inear} is proposed to measure the isotropy and is carried out using the correlation coefficient of GCPs. When the number of GCPs is larger than 20, the Pearson correlation coefficient is used, otherwise the Spearman correlation coefficient is used. NL_{inear} is defined as 1 subtracting the absolute value of the Pearson correlation coefficient or the Spearman correlation coefficient to measure the isotropy of GCPs, which is different from S_{kew} in the Gonçalves method measuring the isotropy of the residuals.

2.4. Cost Functions

In order to combine the three measurements and provide an integrated evaluation result, a total cost function is proposed. Before the combination, as shown in Fig. 1, seed cost functions are needed for the following reasons. Firstly, the measurements need to be transformed into a limited value range and to be desired on the ascending direction. Secondly, the relative weighting of each measurement to the others in the total cost function should be able to be adjusted depending on the requirement of applications.

In most conditions, N_{class} is required to be larger than a threshold, otherwise it will lead to a rapidly grading performance of the subsequent applications or even an invalid result. Moreover, though N_{class} is considered to be the larger the better, too much larger than the threshold do not bring any further contribution. RMS_{loo} is required to be smaller than a threshold, and the smaller the better. Taking above demands into consideration, the shapes of the seed cost functions of N_{class} and RMS_{loo} are required to be like step functions, but with gentler and trailing edges. And Eqs. (3a) and (3b) presenting S-shape curves (as in Fig. 3) are good templates of the seed cost functions. Considering that the value of NL_{inear} is limited on the interval [0, 1] and is desired on the ascending direction, itself can be the seed cost function.

$$C_{Nclass} = \frac{2}{pi}\mathrm{atan}\left(\frac{N_{class}}{N_0}\right)^{\alpha n} \tag{3a}$$

$$C_{RMSloo} = -\frac{2}{pi}\mathrm{atan}\left(\frac{RMS_{loo}}{RMS_0}\right)^{\alpha r} + 1 \tag{3b}$$

where exponents αn and αr are positive parameters related to the steepness of the function curves, and the parameters N_0 and RMS_0 are positive parameters related to the steepness and displacement of the function curves.

To provide an integrated evaluation, the seed cost functions are combined using multiplication as in Eq. (4). The parameters αn, N_0, αr, RMS_0 can adjust the relative weightings of N_{class} and RMS_{loo} to NL_{inear}, which are set depending on the requirement of applications. Based on the simulations described

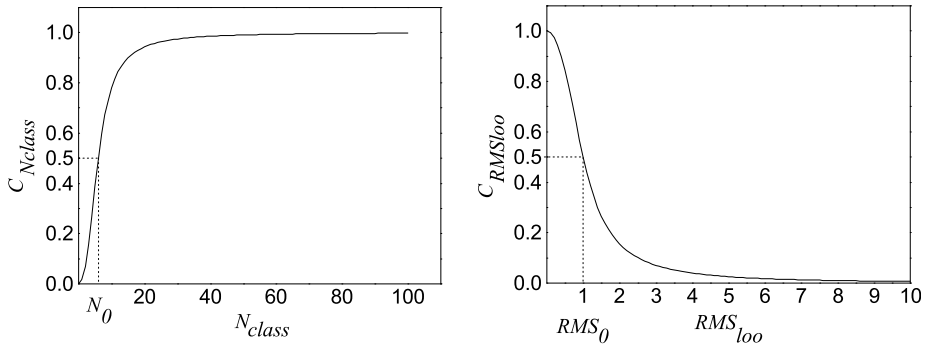

Figure 3. Seed cost functions. (a) Seed cost function of N_{class} when $\alpha n = 2$, $N_0 = 6$. (b) Seed cost function of RMS_{loo} when $\alpha r = 2$, $RMS_0 = 1$.

in Section 3.2, the parameters are empirically suggested: $\alpha n = 2$, $N_0 = 6$, $\alpha r = 2$, $RMS_0 = 1$.

$$\begin{aligned} C_{ost} &= C_{Nclass} * C_{RMSloo} * NL_{inear} \\ &= \left[\frac{2}{pi}\text{atan}\left(\frac{N_{class}}{N_0}\right)^{\alpha n}\right]\left[-\frac{2}{pi}\text{atan}\left(\frac{RMS_{loo}}{RMS_0}\right)^{\alpha r}+1\right]NL_{inear} \end{aligned} \tag{4}$$

3. APPLICATION OF THE PROPOSED METHOD

In this section, four typical scenarios (in Fig. 2) of possible GCP distribution are emulated and two sets of real GCPs (in Fig. 4) in SAR images are selected to illustrate the application of the proposed evaluation method. In the following experiments, the farthest distance clustering method is used in N_{class} calculation, and the values of αn, N_0, αr, RMS_0 are the same as suggested in Section 2.4.

3.1. Accepted Thresholds

Considering the second-order polynomials as transformation functions which are common in the registration of remote sensing images, N_{class} is required larger than 6. RMS_{loo} should be less than 1 in the case of subpixel accuracy. With respect to NL_{inear}, one may consider around 0.6 as the minimum accepted threshold present on the GCP isotropy. Therefore, when parameters of the cost functions are set as suggested, C_{Nclass} and C_{RMSloo} are required to be larger than 0.5, and correspondingly the accepted threshold of the total cost function is 0.15. Regarding other different transformation functions and parameters, the accepted thresholds can be easily obtained. Care should be taken that, different values of the three measurements may result in the same value of C_{ost}, so the measurements also need to be analyzed individually.

3.2. Application of the Proposed Method to Emulated GCP Scenarios

Four typical scenarios of 30-GCP distribution which are common in practical are emulated, and all of them are assumed in images of 300×300 pixel size. The location of GCPs (row and column) in Fig. 2 are all randomly generated, according to a uniform distribution on the interval [120, 180] for Fig. 2(a) and on the interval [75, 225] for Fig. 2(b) and around a line with 20 standard deviation for Fig. 2(c) and on the interval [0, 300] for Fig. 2(d). Figs. 2(a) and 2(c) are undesirable GCP distributions, because GCPs in Fig. 2(a) are too close to each other and GCPs in Fig. 2(c) are nearly around a preference axis. Figs. 2(b) and 2(d) are both desirable GCP distributions, and Fig. 2(d) is better owing to the larger coverage area of GCPs. Moreover, the GCP residuals in the four scenarios are randomly generated.

Regarding the scenarios shown in Fig. 2, obtained values of the measurements in the proposed method are shown in Table 2. Though there are 30 GCPs in all of the four scenarios, the values of N_{class} are greatly different and agree with the distributions shown in Fig. 2. The measurement NL_{inear}

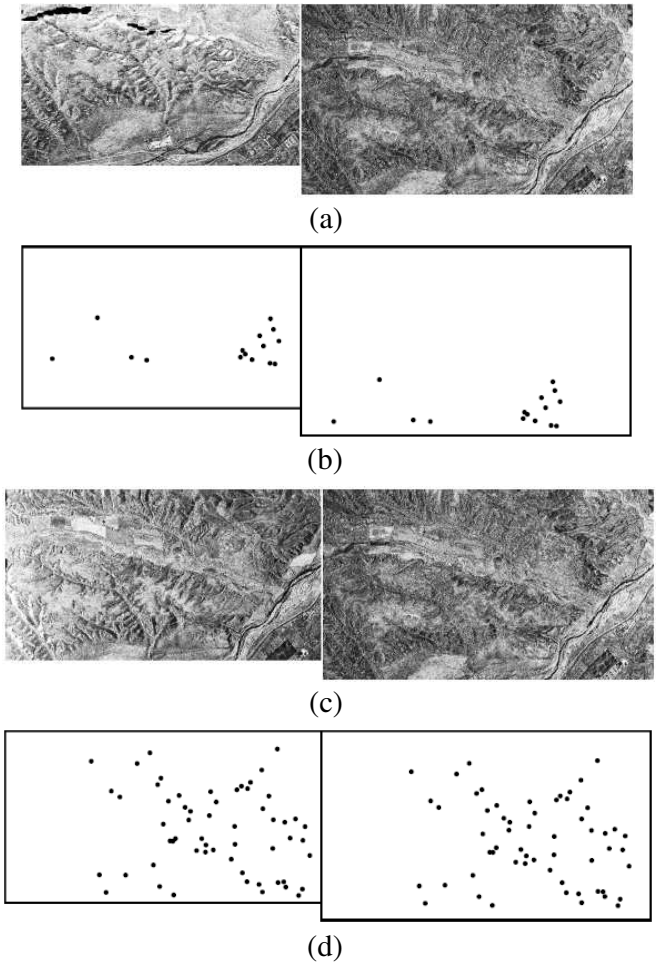

Figure 4. Two sets of real GCPs in SAR images. (a) and (c) are two pairs of SAR images for GCP extraction. (b) and (d) present the GCPs extracted from (a) and (c) respectively.

Table 2. Values of the proposed measurements evaluating the emulated GCP scenarios shown in Fig. 2 and real GCP scenarios in Fig. 4, where $d_{\min} = 20$, $\alpha n = 2$, $N_0 = 6$, $\alpha r = 2$, $RMS_0 = 1$ and validation with value 1 denotes acceptance.

Measurements	Emulated GCPs in Fig. 2				Real GCPs in Fig. 4	
	(a)	(b)	(c)	(d)	(a)	(b)
N_{class}	3	16	20	23	12	51
C_{Nclss}	0.16	0.91	0.94	0.96	0.84	0.99
RMS_{loo}	0.79	0.76	0.83	2.22	2.77	2.79
C_{RMSloo}	0.65	0.67	0.62	0.13	0.31	0.29
NL_{inear}	0.95	0.99	0.05	1.00	0.90	0.81
C_{ost}	0.10	0.60	0.03	0.12	0.2318	0.2424
validation	0	1	0	0	1	1

Table 3. Values of the Gonçalves measurements evaluating emulated GCP scenarios shown in Fig. 2 and real GCP scenarios in Fig. 4, where validation with value 1 denotes acceptance.

Measurements	Emulated GCPs in Fig. 2				Real GCPs in Fig. 4	
	(a)	(b)	(c)	(d)	(a)	(b)
N_{red}	18	18	18	18	3	49
RMS_{all}	0.70	0.68	0.75	1.98	2.54	2.75
RMS_{loo}	0.79	0.76	0.83	2.22	2.77	2.79
P_{quad}	0.10	0	0.17	0	0.03	0
B_{pp}	0.13	0.03	0.20	0.20	0.56	0.56
S_{kew}	0.06	0.01	0.95	0	0.10	0.22
S_{cat}	1	1	1	1	0.99	1
ϕ	0.41	0.37	0.55	0.75	1.004	0.99
validation	1	1	1	0	0	0

has been proved to be able to successfully detect a preference axis as in Fig. 2(c). The results of the illustration indicate that the failing of N_{class} in Fig. 2(a), NL_{inear} in Fig. 2(c) and RMS_{loo} in Fig. 2(d) all have great influence on the values of the total cost function.

The performance of the proposed method is also contrasted with the most comprehensive one — the Gonçalves method. Obtained values of the Gonçalves measurements evaluating the GCP scenarios in Fig. 2 are shown in Table 3. The thresholds of N_{red}, RMS_{all}, RMS_{loo}, P_{quad}, B_{pp}, S_{kew}, S_{cat} and ϕ are 12, 1, 1, 0.95, 0, 0.4, 0.95 and 0.591 respectively for second-order polynomials as transformation functions, and the values of all the measurements are desired on the descending direction except N_{red} [7]. The results of the illustration indicate that the values of S_{cat}, are all around 1 on the unaccepted interval which fails to reflect the GCP dispersion. The contrast between Table 2 and Table 3 indicates that the proposed measurement for evaluating GCP dispersion — N_{class} performs much better than S_{cat}. About half of the validations of the scenarios in Fig. 2 by the proposed method and the Gonçalves method are unanimous. Fig. 2(a) is judged accepted and Fig. 2(d) unaccepted by the Gonçalves method, while the validations of the proposed method are opposite. From Fig. 2, it can be obviously known that the validations of the proposed method are more reasonable. The experiment demonstrates that the proposed method performs more effective and reasonable than the Gonçalves method, especially in the evaluation of the GCP dispersion.

3.3. Application of the Proposed Method to Real GCP Scenarios

In this experiment, the proposed method and the Gonçalves method are applied to two sets of real GCPs extracted using BFSIFT [12] in SAR images. The left and right SAR images in Fig. 4(a) are 454×793 and 535×941 pixel size respectively, and the left and right SAR images in Fig. 4(c) are 484×900 and 535×941 pixel size respectively. The GCPs extracted from Figs. 4(a) and 4(c) are presented in Fig. 4(b) and Fig. 4(d) respectively, and obtained values of the proposed measurements and the Gonçalves measurements are presented in Table 2 and Table 3. The two sets of GCPs are both considered to have accepted dispersion quality and validated accepted by the proposed method. However, the Gonçalves method considers that they both have bad dispersion quality and validates them unaccepted. From Fig. 4 it can be obviously known that the validations of the proposed method agree with the fact better. The experiment demonstrates that when applied to real GCPs, the proposed method can evaluate GCP quality effectively, and even better than the Gonçalves method.

4. CONCLUSION

In order to provide an objective, quantitative and effective evaluation of GCP quality, we proposed an evaluation method in this paper, which consisted of the root mean square of residuals using leave-one-out method (RMS_{loo}) and two new measurements evaluating the number, dispersion and isotropy of GCPs. In the method, seed cost functions were utilized to transform the measurements into a limited value range as well as to be desired on the ascending direction. All the seed cost functions were combined by a total cost function, which could be used as a single quantitative and objective analysis of GCP quality. The GCP scenario was validated by the accepted threshold depending on the value of the total cost function. In addition, the accepted thresholds of all the measurements and cost functions were suggested. Compared with the Gonçalves method, the measurements in the proposed method could evaluate almost the same aspects of GCP quality but using fewer measurements. Experiments using emulated and real sets of GCPs demonstrated that the proposed method performed more effectively than the Gonçalves method, especially in the evaluation of dispersion quality.

REFERENCES

1. Fan, B., C. Huo, C. Pan, and Q. Kong, "Registration of optical and SAR satellite images by exploring the spatial relationship of the improved SIFT," *IEEE Geosci. Remote Sens. Lett.*, Vol. 10, No. 4, 657–661, 2013.
2. Moser, G. and S. B. Serpico, "Unsupervised change detection from multichannel SAR data by Markovian data fusion," *IEEE Trans. on Geosci. Remote Sens.*, Vol. 47, No. 7, 2114–2128, 2009.
3. Alparone, L., S. Baronti, A. Garzelli, and F. Nencini, "Landsat ETM+ and SAR image fusion based on generalized intensity modulation," *IEEE Trans. on Geosci. Remote Sens.*, Vol. 42, No. 12, 2832–2839, 2004.
4. Du, L., H. Liu, Z. Bao, and M. Xing, "Radar HRRP target recognition based on higher order spectra," *IEEE Trans. on Signal Process.*, Vol. 53, No. 7, 2359–2368, 2005.
5. Ma, W. and J. Yang, "Target detection algorithm for polarimetric SAR images using GOPCE," *Proc. IEEE CIE Int. Conf. Radar*, 507–509, 2011.
6. Chureesampant, K. and J. Susaki, "Automatic GCP extraction of fully polarimetric SAR images," *IEEE Trans. on Geosci. Remote Sens.*, Vol. 52, No. 1, 137–148, 2014.
7. Gonçalves, H., J. A. Gonçalves, and L. Corte-Real, "Measures for an objective evaluation of the geometric correction process quality," *IEEE Geosci. Remote Sens. Lett.*, Vol. 6, No. 2, 292–296, 2009.
8. Buiten, H. J. and B. van Putten, "Quality assessment of remote sensing image registration — Analysis and testing of control point residuals," *ISPRS J. Photogramm. Remote Sens.*, Vol. 52, No. 2, 57–73, 1997.
9. Gonçalves, H., L. Corte-Real, and J. A. Gonçalves, "Automatic image registration through image segmentation and SIFT," *IEEE Trans. on Geosci. Remote Sens.*, Vol. 49, No. 7, 2589–2600, 2011.
10. Gonçalves, H., J. A. Gonçalves, and L. Corte-Real, "Automatic image registration based on correlation and Hough transform," *Proc. SPIE Image Signal Process. Remote Sens. XIV*, 71090J, 2008, doi: 10.1117/12.800351.
11. Wang, S., H. You, and K. Fu, "BFSIFT: A novel method to find feature matches for SAR image registration," *IEEE Geosci. Remote Sens. Lett.*, Vol. 9, No. 4, 649–653, 2012.
12. Bradley, P. and U. Fayyad, "Refining initial points for K-means clustering," *Proc. 15th Int. Conf. Machine Learning*, 91–99, 1998.
13. Höppner, F., F. Klawonn, and R. Kruse, *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis, and Image Recognition*, Wiley, New York, 1999.
14. Hartuv, E. and R. Shamir, "A clustering algorithm based on graph connectivity," *Inf. Process. Lett.*, Vol. 76, No. 4–6 175–6181, 2000.