Progress In Electromagnetics Research M, Vol. 37, 73–82, 2014

Modified RANSAC for SIFT-Based InSAR Image Registration

Yang Wang[1, *], Haifeng Huang[1], Zhen Dong[1], and Manqing Wu[1, 2]

Abstract—In this paper, we propose a modified version of the Random Sample Consensus (RANSAC) method for Interferometric Synthetic Aperture Radar (InSAR) image registration based on the Scale-Invariant Feature Transform (SIFT). Because of speckle, the "maximization of inliers" criterion in the original RANSAC cannot obtain the optimal results. Since in InSAR image registration, the registration accuracy is in inverse proportion to number of residues. Therefore, we modify the old criterion with a new one — Minimization of residues — to obtain the optimal results. We tested our method on a variety of real data from different sensors, and the experimental results demonstrated the validity and robustness of the proposed method.

1. INTRODUCTION

Interferometric synthetic aperture radar (InSAR) has been widely used for digital elevation model (DEM) generation [1–6]. This technique is based on the generation of an interferogram by image registration of two single-look complex (SLC) images of the same scene taken from two-orbit or two-time measurements. Although plenty of research has been done on the topic of image registration such as medical images, and many excellent methods have been put forward, many of these methods cannot be directly applied to InSAR image registration because of its special characteristics such as speckle, layover and shadow. In the last decades, a lot of methods [7–11] have been proposed for InSAR image registration among which the cross-correlation method is the most widely used one. This process usually consists of two steps: coarse registration and fine registration. The procedures of the two steps are the same except for an oversampling step in the latter to reach subpixel level. Image registration is realized by maximizing the real or complex cross-correlation of two images. However, owing to a high computation load, pixel-by-pixel registration over the entire image is time consuming. In real cases, registration of some evenly distributed control points based on windows is used instead. After the control points identification, matching windows, and searching windows centered on these points are set up in the master and slave images, respectively. Relative offsets of the control points are obtained when the highest correlation coefficient is achieved by cross-correlation of the two windows. After that, interpolation is used to get the relative offset of other pixels. As a matter of fact, cross-correlation method is a point-based method which uses cross-correlation function as its registration measure function. However, control points in cross-correlation method do not have theoretical meaning (such as corners, extreme points) or robust enough to be detected in different images. Furthermore, cross-correlation function is not robust under some conditions such as rotation.

In order to solve aforementioned problems, a famous method called Scale-Invariant Feature Transform (SIFT) was proposed in 1999 [12] and optimized in 2004 [13]. SIFT operator, unlike cross-correlation, extracts stable keypoints which are of high repeatability and will not change too much under different imaging conditions, at the extrema of the Laplacian of the image in the image scale space representation. Then, it relies on the scale space of the image to construct feature descriptors,

Received 22 April 2014, Accepted 18 June 2014, Scheduled 25 June 2014

* Corresponding author: Yang Wang (wycbx8384@163.com).

[1] School of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China. [2] China Electronics Technology Group Corporation (CETC), Beijing 100000, China.

which are accredited to be very distinctive and invariant to scale change, rotation, and even affine transformation. All these good characteristics make SIFT a very powerful tool for many applications such as object recognition, robot localization and mapping, and panorama stitching. When applied to image registration, Random Sample Consensus (RANSAC) [14] is often used with SIFT to remove outliers (mismatched pairs of points). Although a lot of image registration results by SIFT and modified versions with RANSAC are reported [15–20], few works have been done on InSAR image registration and little attention has been paid to RANSAC. Furthermore, the "maximization of inliers" criterion of the original RANSAC is not optimal for InSAR image registration for the number of residues (NOR) [21] of the interferogram obtained through this criterion is not the fewest. In order to improve the registration accuracy, we modify RANSAC with a new criterion and propose the modified RANSAC for InSAR image registration.

This paper is organized as follows. Section 2 makes elaborate descriptions about every step included in our proposed method. Section 3 presents experimental results from different image pairs. Conclusions are drawn in Section 4.

2. PROPOSED METHOD

A framework of the proposed method is shown in Figure 1. First, features are extracted from the master and slave images by SIFT. Second, an initial set of matching points is obtained by nearest neighborhood distance ratio (NNDR). Third, the final set of matching points is obtained by modified RANSAC. Finally, after getting the transformation by the final set of matching points and the registered slave image by this transformation, the interferogram is obtained by the multiplication of the master and the conjugated registered slave images, and the interferometric phase is extracted by the inverse tangent from the interferogram. Details are described as follows.

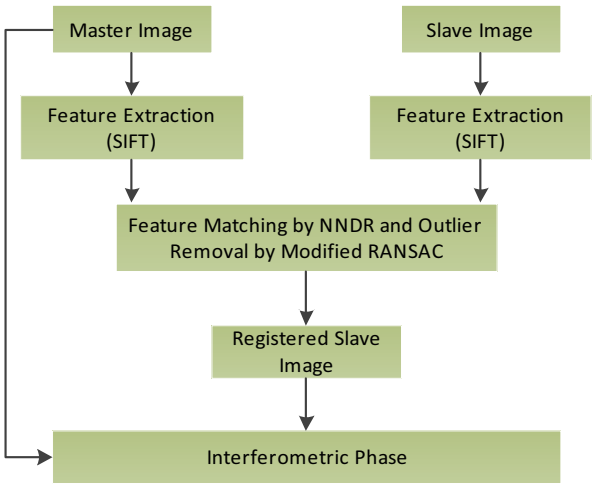


Figure 1. Framework of the proposed method.

2.1. SIFT

Generally speaking, the entire processing chain of SIFT can be summarized into five main steps. The first is the keypoints detection in the difference of Gaussian images, which are produced by subtracting adjacent Gaussian images. The second is the keypoints subpixel localization which is very suitable for InSAR image registration, and unstable keypoints elimination. The third is the keypoints main orientation assignment, which makes these points rotation-invariant. Then, descriptors are constructed based on weighted gradients in the local image regions. Finally, keypoints matching are realized through these descriptors by NNDR, which compares the ratio of the Euclidean distance of the closest match to the second closest one with a given threshold. Lowe suggests that this threshold be set to 0.8 [13]. According to our experience, the best results could be achieved when the threshold is set to 0.82. We

will use this value in the rest of the paper. To reduce the processing time and decrease the influence of speckle, we skip the first scale-space octave as proposed by [15].

2.2. Transformation Model

The transformation between master and slave images should be carefully chosen, because improper transformation may lead to bad results. For InSAR image registration, master and slave images are acquired from the same satellite and sensor with the same imaging parameters, and in most cases affine transformation can accurately express the relation between the two images; therefore, affine transformation is taken as the transformation model in this paper.

2.3. Modified RANSAC

RANSAC is an algorithm for robust fitting of models in the presence of many data outliers. To get a reasonable result, different iterations are needed according to the dimensions of the data. The main steps can be concluded as follows:

1) Select M pairs of points from the initially matched points set at random. M is different due to different transformation models. For affine transformation, M is no less than 3.
2) Estimate the transformation parameter $\vec{x}$ from the selected M pairs of points.
3) Find how many pairs of points fit the model with parameter vector $\vec{x}$ within a given tolerance. Call this inliers.
4) Repeat 1) to 3) n times and take the biggest set of inliers as the final matching points set.
   In step 4), a threshold always exists for number of iterations (call it N). When we iterate 1) to 3) more than N times, the number of inliers will converge on a fixed value (call it I). With I we can calculate the transformation model, register the image, and obtain the interferometric phase. After that, residues are calculated for evaluation. The better the registration, the fewer the residues. However, residues obtained via I are not the fewest because of speckles. Unlike noise, speckles are the results of coherent processing on SAR data and cannot be correctly removed by RANSAC. Therefore, there are always speckles in I which makes the transformation model not optimal or even erroneous. In order to tackle this problem, we make modifications as follows.
5) After 1st iteration from 1) to 3), inliers are obtain, call this M_1. We can calculate the transformation model, register the image, obtain the interferometric phase, and calculate the number of residues, call this R_1. After 2nd iteration, we obtain M_2, R_2, and:
$$\begin{cases} R_{optimal} = R_2 & M_{optimal} = M_2 \quad \text{if} \quad R_1 > R_2 \\ R_{optimal} = R_1 & M_{optimal} = M_1 \quad \text{if} \quad R_1 < R_2 \end{cases} \tag{1}$$
6) Repeat 1) to 3). After each iteration, we obtain M_n, R_n (n is the number of iteration), and compare R_n with $R_{optimal}$:
$$\begin{cases} R_{optimal} = R_{optimal} & M_{optimal} = M_{optimal} \quad \text{if} \quad R_n > R_{optimal} \\ R_{optimal} = R_n & M_{optimal} = M_n \quad \text{if} \quad R_n < R_{optimal} \end{cases} \tag{2}$$
   If $R_{optimal}$ does not change after some iterations (such as 30), $M_{optimal}$ is regarded as the optimal inliers and the corresponding interferometric phase is the final image registration result.

3. EXPERIMENTS

In this section, we first demonstrate the validity of the proposed method using data of Etna Mountain, and then test its robustness with real data from different sensors. The performance of the proposed method is verified in terms of qualitative and quantitative evaluations with real data. The results obtained by cross-correlation (fine registration) are also presented. The main purpose of doing so is to see whether the proposed method can reach subpixel level. The control points of cross-correlation method are evenly chosen in the entire image with an interval of 20 pixels and the vectors of relative offset which are larger than two times the median value of neighboring ones are removed.

3.1. Validity

The SIR-C/X-SAR mission, supported by NASA, DLR, and ASI, took its two flights in April 9 to April 20 in 1994 and September 30 to October 11 in the same year respectively. The real data, X band, are about Etna Mountain of Italy. The master and slave images with the size of 1024×1024 are shown in Figures 2(a) and (b), respectively.

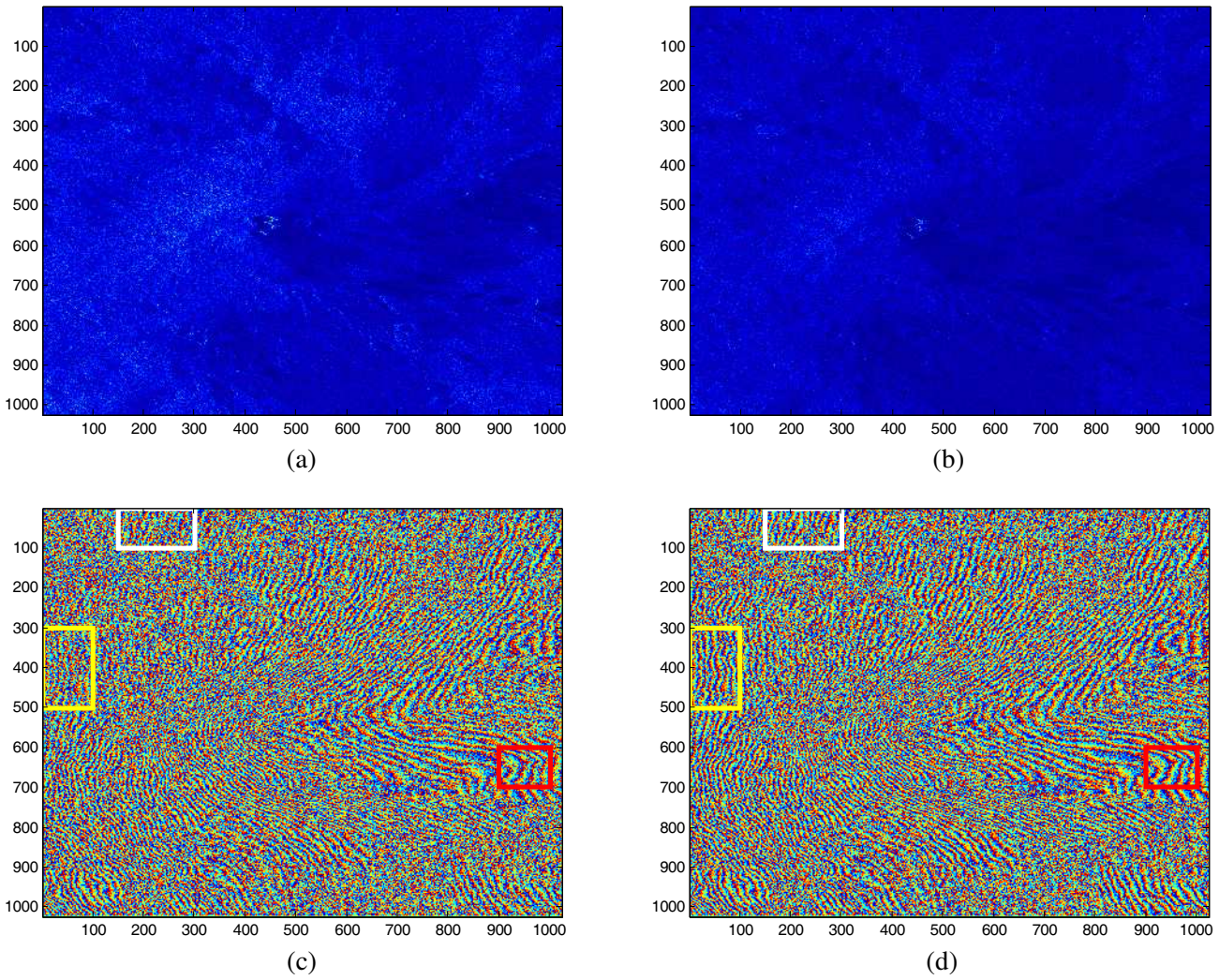


Figure 2. Master (a) and slave (b) images of Etna Mountain. (c) Interferometric phase obtained by SIFT and original RANSAC. (d) Interferometric phase obtained by SIFT and modified RANSAC. Numbers of residues of regions labeled out by yellow lines in (c) and (d) are 5320 and 3169 respectively. Numbers of residues of regions labeled out by white lines in (c) and (d) are 3992 and 2667 respectively. Numbers of residues of regions labeled out by red lines in (c) and (d) are 1837 and 970 respectively.

Interferometric phases obtained by original and modified RANSAC are shown in Figures 2(c) and (d) respectively. We can see that the number of residues decreases to some extent by modified RANSAC and this reduction is global. Table 1 shows the performance comparison between original and modified RANSAC. After modification to original RANSAC, the NOR decreases from 223603 to 207837 and the MCC increases from 0.5790 to 0.6086. It is clearly indicated that our modification is effective and better than the original RANSAC.

Table 1. Image registration performance comparison on Etna Mountain.

Method	Original RANSAC	Modified RANSAC
NOR	231566	207837
MCC	0.5622	0.6086

We calculate MCC using Guarnieri's method [22], and window size is 9×9.

3.2. Robustness

Details about the image pairs acquired from different sensors, topography, and frequency bands are listed in Table 2. The master images of this data set are shown in Figure 3.

Table 2. Descriptions of each SAR image pair.

Data	Sensor	Topography	Band	Acquired Date
1	TerraSAR-X	Mountain	X	2008.3.10 & 2008.3.21
2	Alos Palsar	Mountain	L	2007.7.1 & 2007.8.16
3	Alos Palsar	Forestry	L	2007.3.13 & 2007.4.28
4	RadarSat-2	Urban area	C	2008.5.4 & 2008.5.28
5	Sir-C	Mixture	L	1994.10.9 & 1994.10.10

Images acquired earlier are used as master images.

Interferometric phases obtained after image registration are presented in Figure 4. From visual effect, results obtained by the proposed method are better than those obtained by original RANSAC. For data 1, the interference fringes in the lower part of Figure 4(b) is clearer than the same region

Table 3. Performance comparison of methods.

Image Pair	Method	NOR	MCC
Data 1	CC	427102	0.6365
	Original	469624	0.5975
	Proposed	408163	0.6687
Data 2	CC	1400084	0.3758
	Original	1622266	0.3217
	Proposed	1419601	0.3720
Data 3	CC	386541	0.3057
	Original	425387	0.2458
	Proposed	398650	0.2889
Data 4	CC	611674	0.4647
	Original	632730	0.4602
	Proposed	573941	0.5518
Data 5	CC	1016066	0.7663
	Original	1217355	0.7094
	Proposed	896909	0.7950

We calculate MCC using Guarnieri's method [22], and window size is 9×9.
"CC" refers to cross-correlation method (fine registration).

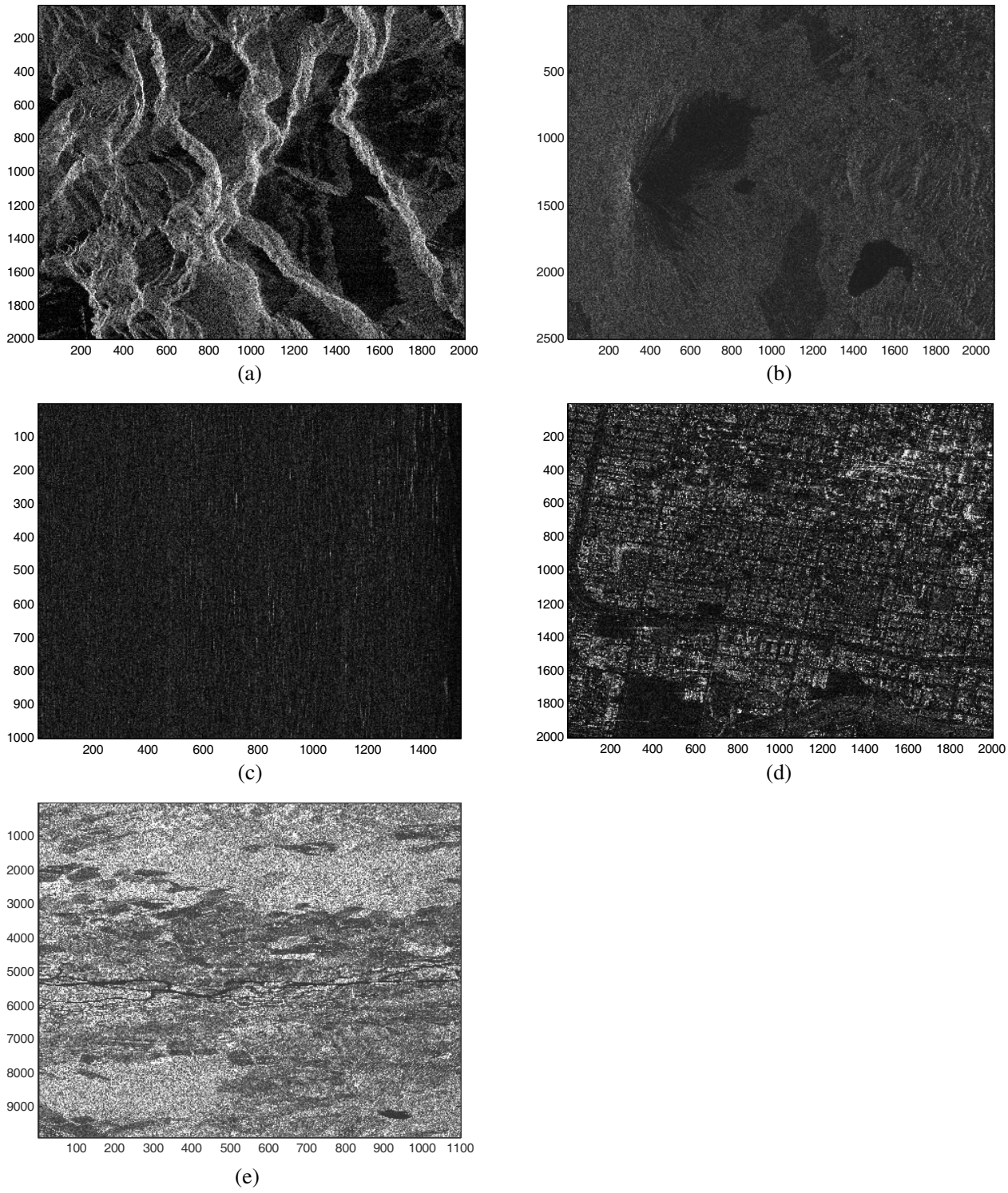
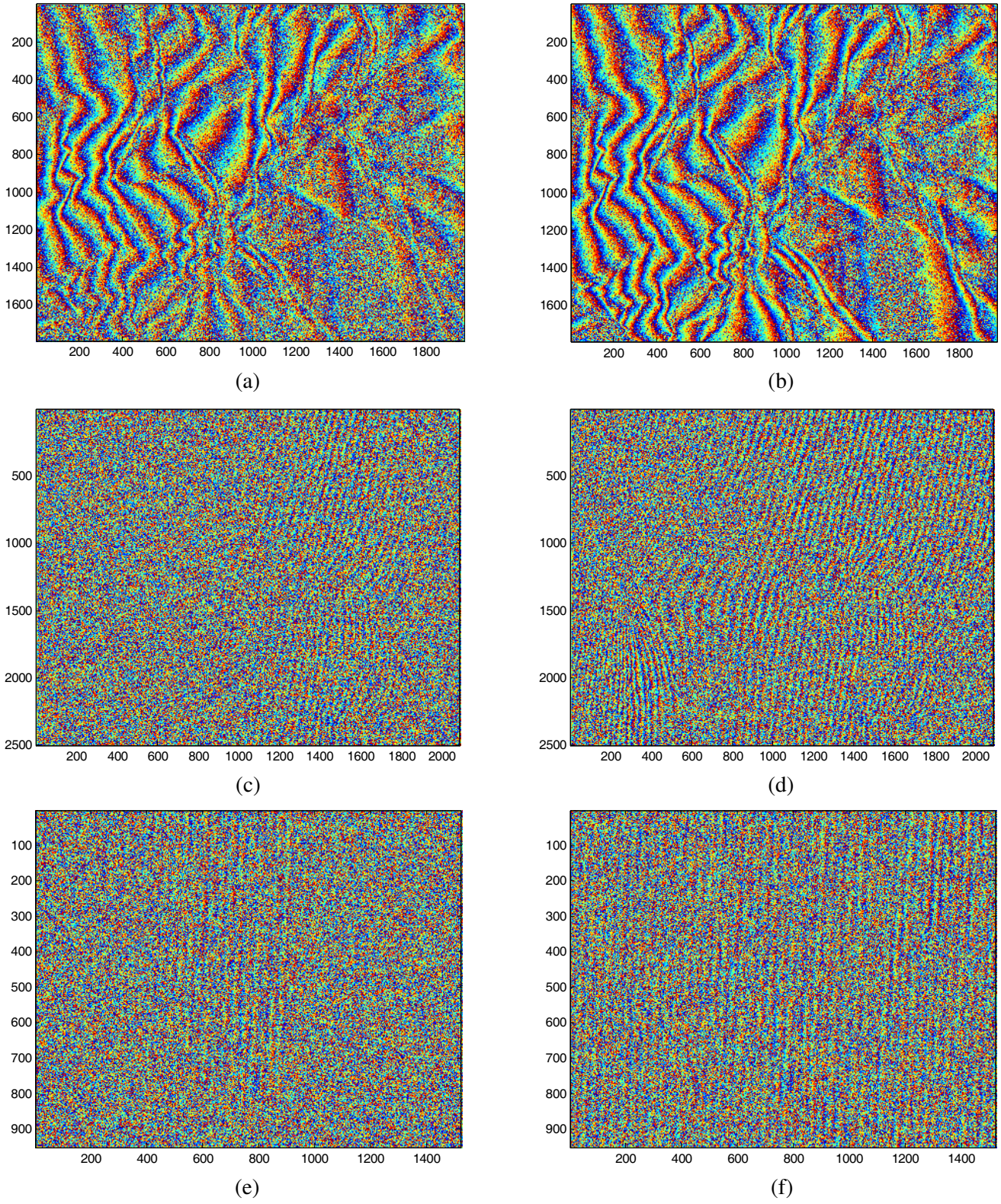


Figure 3. Master images of data sets described in Table 2. (a) Master image of data 1, 2000×2000 pixels, is about Grand Canyon. (b) Master image of data 2, 2500×2088 pixels, is about Fuji Mountain. (c) Master image of data 3, 1000×1536 pixels, is about Amazon rain forest. (d) Master image of data 4, 2000×2000 pixels, is about urban area of Canada. (e) Master image of data 5, 9900×1100 pixels, is about Selenga Delta.

in Figure 4(a). We expect the MCC of data 2 to be low due to the long time decorrelation. More interference fringes are obtained around the crater in Figure 4(d) when compared with Figure 4(c). In addition to time decorrelation, volume decorrelation is also serious for data 3. More interference



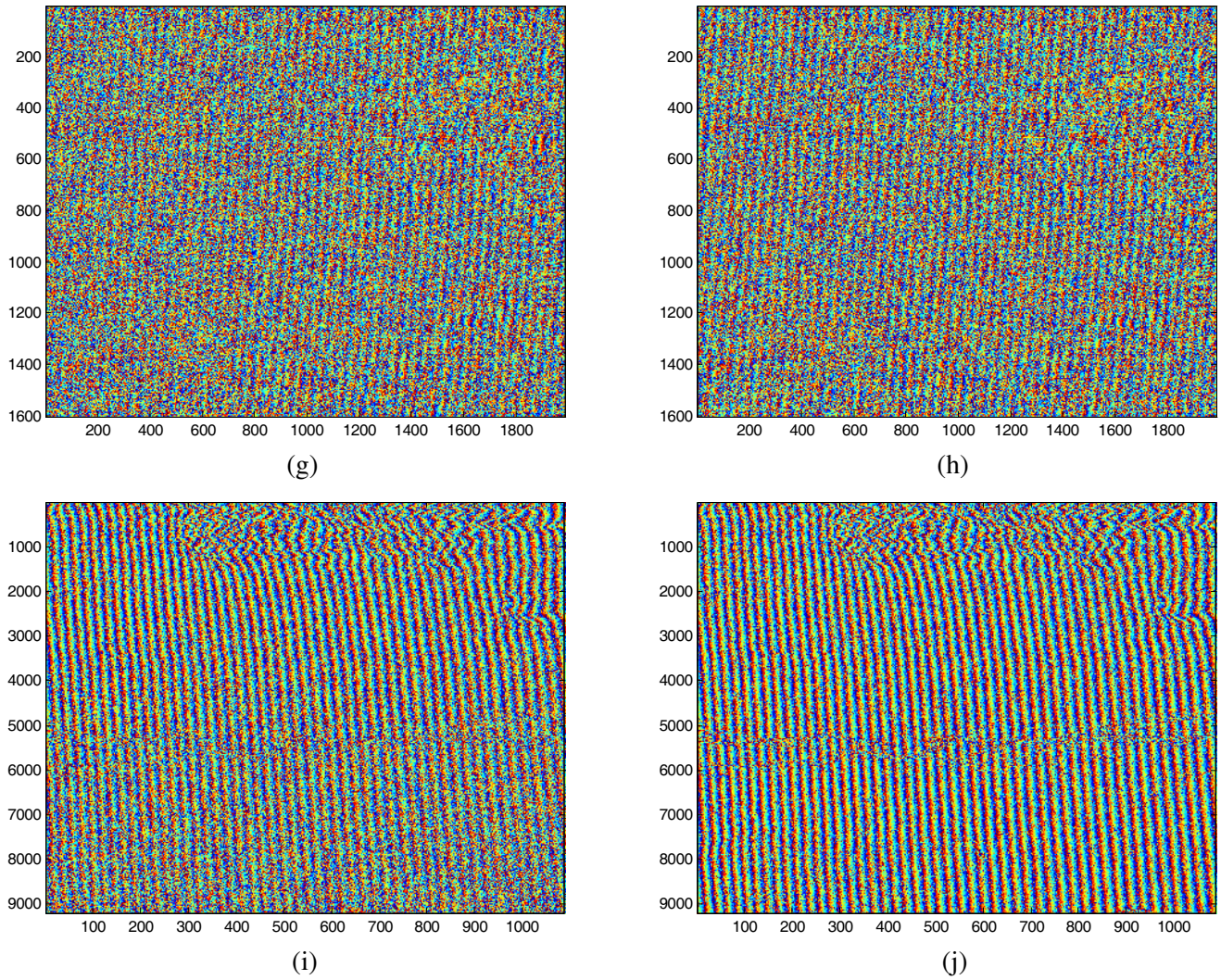


Figure 4. Interferometric phases obtained after registration by original RANSAC and proposed method. (a), (c), (e), (g), and (i) are interferometric phases obtained by original RANSAC on data 1 to 5 accordingly. (b), (d), (f), (h), and (j) are interferometric phases obtained by proposed method on data 1 to 5 accordingly. The non-overlap areas have been cut.

fringes are obtained in the left and right part of Figure 4(f) when compared with Figure 4(e). Although not obvious because of weak scattering of the roads, improvements still can be seen in the left part of Figure 4(h) when compared with the same region in Figure 4(g). Due to low time decorrelation, the interference fringes are very clear for data 5. The lower part of Figure 4(i) is noisy and the corresponding area shown in Figure 4(j) is very clear. Quantitative evaluations made in terms of NOR and MCC are listed in Table 3. From this table, we can see that the proposed method, with higher mean cross-correlation and fewer residues, is better than the original RANSAC.

Comparisons with the original RANSAC on data 1 to data 5 have shown that NOR has decreased by 13.1%, 12.5%, 6.2%, 9.4%, and 26.3%, respectively, and MCC has increased by 12%, 15.6%, 17.5%, 20%, and 12.1%, respectively. We also introduce results from the cross-correlation method. From comparisons between the proposed method and the cross-correlation method, some conclusions can be drawn. The proposed method outperforms the cross-correlation method on data 1, 4, and 5. The reasons are that the image pairs have low time decorrelation or varied topography such as mountains and buildings; therefore, the proposed method can extract more and stable keypoints from the images. On the other hand, the cross-correlation method uses evenly distributed control points for registration.

Although correlation coefficients are maximized on these points, correlation coefficients of other points are indirectly obtained by interpolation, therefore the MCC of the entire images cannot be maximized. On data 2 and data 3, however, due to serious time decorrelation and volume decorrelation, not enough keypoints can be extracted but results obtained by the proposed method are still comparable with the ones obtained by cross-correlation in terms of NOR or MCC. From the relations between registration accuracy and NOR or MCC, we can say that the proposed method has reached the subpixel level and can be used for InSAR image registration.

4. CONCLUSIONS

In this paper, an efficient method has been developed for InSAR image registration. First, we introduce a modified RANSAC to obtain more accurate image registration. This technique uses "minimization of residues" as the new criterion, which is more suitable and problem-specific for InSAR image registration. Then, an experiment on Mount Etna verifies that this modification can obtain results with fewer residues and higher MCC. Finally, the proposed method and cross-correlation method are tested on a variety of real data. In short, proposed method is better for high correlation data (stable scatterers, such as building and mountain, or low time decorrelation data) than cross-correlation method.

The proposed method still has some limitations. The construction of SIFT descriptor mainly depends on the main orientation of keypoints and this main orientation depends on the gradients of neighboring pixels. The gradients are sensitive to noise. Therefore, in high correlation data (noise level is low), SIFT descriptor is reliable and superior to cross-correlation function. In low correlation data, however SIFT descriptor is seriously affected by high level noise. Besides, iteration in RANSAC is time consuming. Future work will include developing a more robust SIFT descriptor to noise and a more advanced RANSAC to improve the efficiency.

ACKNOWLEDGMENT

The authors would like to thank PIER's Editor and the anonymous reviewers for their constructive feedback on an earlier version of this paper. This work is supported by the National Natural Science Foundation of China (grant: 61002031).

REFERENCES

1. Bamler, R. and P. Hartl, "Synthetic aperture radar interferometry," *Inverse Problems*, Vol. 14, R1–R54, 1998.
2. Hanssen, R. F., *Radar Interferometry: Data Interpretation and Error Analysis*, Kluwer Academic, 2001.
3. Li, C. and D. Y. Zhu, "A residue-pairing algorithm for InSAR phase unwrapping," *Progress In Electromagnetics Research*, Vol. 95, 341–354, 2009.
4. Wu, B. I., M. C. Yeung, Y. Hara, and J. A. Kong, "InSAR height inversion by using 3-D phase projection with multiple baselines," *Progress In Electromagnetics Research*, Vol. 91, 173–193, 2009.
5. Li, S., H. Xu, and L. Zhang, "An advanced DSS-SAR InSAR terrain height estimation approach based on baseline decoupling," *Progress In Electromagnetics Research*, Vol. 119, 207–224, 2011.
6. Liu, Q., S. Xing, X. Wang, J. Dong, D. Dai, and Y. Li, "The "slope" effect of coherent transponder in InSAR DEM," *Progress In Electromagnetics Research*, Vol. 127, 351–370, 2012.
7. Stone, H. S., M. T. Orchard, E. C. Change, and S. A. Martucci, "A fast direct Fourier based algorithm for subpixel registration of images," *IEEE Trans. Geosci. Remote Sens.*, Vol. 39, No. 10, 2235–2243, 2001.
8. Lin, Q., J. F. Vesecky, and H. A. Zebker, "New approaches in interferometric SAR data processing," *IEEE Trans. Geosci Remote Sens.*, Vol. 30, No. 3, 560–567, 1992.
9. Gabriel, K. and R. M. Goldstein, "Crossed orbit interferometry: Theory and experimental results from SIR-B," *Int. J. Remote Sens.*, Vol. 9, No. 5, 857–872, Sep. 1988.

10. Natsuaki, R. and A. Hirose, "SPEC method — A fine coregistration method for SAR interferometry," *IEEE Trans. Geosci Remote Sens.*, Vol. 49, No. 1, 28–37, 2011.
11. Li, D. and Y. H. Zhang, "A fast offset estimation approach for InSAR image subpixel registration," *IEEE Geosci Remote Sens Letters.*, Vol. 9, No. 2, 267–271, 2011.
12. Lowe, D. G., "Object recognition from local scale-invariant features," *IEEE International Conference on Computer Vision*, Vol. 2, 1150–1157, Kerkyra, Greece, Sep. 20–27, 1999.
13. Lowe, D. G., "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, Vol. 60, No. 2, 91–110, 2004.
14. Fischler, M. A. and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, Vol. 24, No. 6, 381–395, 1981.
15. Schwind, P., S. Suri, P. Reinartz, and A. Siebert, "Applicability of the SIFT operator for geometrical SAR image registration," *Int. J. Remote Sens.*, Vol. 31, No. 8, 1959–1980, 2010.
16. Morel, J. M. and G. S. Yu, "ASIFT: A new framework for fully affine invariant image comparison," *SIAM J. Imag. Sci.*, Vol. 2, No. 2, 438–469, 2009.
17. Li, Q. L., G. Y. Wang, J. G. Liu, and S. B. Chen, "Robust scale-invariant feature matching for remote sensing image registration," *IEEE Geosci. Remote Sens. Lett.*, Vol. 6, No. 2, 287–291, 2009.
18. Suri, S., P. Schwind, J. Uhl, and P. Reinartz, "Modifications in the SIFT operator for effective SAR image matching," *Int. J. Image Data Fus.*, Vol. 1, No. 3, 243–256, 2010.
19. Wang, S., H. You, and K. Fu, "BFSIFT: A novel method to find feature matches for SAR image registration," *IEEE Trans. Geosci. Remote Sens.*, Vol. 9, No. 4, 649–653, 2012.
20. Goncalves, H., L. Corte-Real, and J. Goncalves, "Automatic image registration through image segmentation and SIFT," *IEEE Trans. Geosci Remote Sens.*, Vol. 49, No. 7, 2589–2600, 2011.
21. Goldstein, R. M., H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Science*, Vol. 23, No. 4, 713–720, 1988.
22. Guarnier, M. I. and C. Prati, "A quick and dirty coherence estimator for data browsing," *IEEE Trans. Geosci. Remote Sensing.*, Vol. 35, 660–669, May 1997.