AUTOMATED MULTI-SOURCE REMOTE SENSING IMAGE REGISTRATION BASED ON PHASE CONGRUENCY

Yuanxin YE[a, *], Lian XIONG[a], Jie SHAN[ab]

[a]School of Remote Sensing and Information Engineering, Wuhan University,
NO 129, Luoyu Road, Wuhan 430079, China
[b]School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA

KEY WORDS: Remote Sensing, Image Registration, Phase Congruency, Image Pyramid, Registration Accuracy

ABSTRACT:

Automated registration of multi-source remote sensing data, e.g., optical, SAR, LiDAR data, could be difficult due to their heterogeneity. This paper proposes a method based on the concept of phase congruency, a measure of feature significance robust to change in illumination and contrast. We first calculate the phase congruencies respectively for the input image and the reference image to reduce local illumination and contrast difference caused by the heterogeneity or radiometric variation. Minimum moment of phase congruency is used to select feature points in the input image, followed by a normalized cross correlation to determine their correspondences in the reference image. Prior georeferencing information and image pyramid guide the above matching process, which is then refined by means of least squares matching (LSM) method. The proposed method uses random sample consensus (RANSAC) to remove the large correspondence errors caused by e.g., shadow and shielding. Finally, image registration is achieved by determining a projective transformation between the two images based on the final set of correspondences. The proposed method is evaluated with multi-source remote sensing images, including optical images, SAR images and LiDAR data. The registration quality is evaluated by using manually collected check points. The results demonstrate that the proposed method is robust and can achieve a registration accuracy which can be comparable to that produced by manual registration.

1. INTRODUCTION

Image registration is a fundamental problem in image processing, which is to align two or more images of the same scene often taken at different times, from different viewpoints, or by different sensors. It is also a basic step for image mosaic, object identification, image fusion, change detection etc. As heterogeneity or radiometric variation always exists among multi-source remote sensing images, especially the images acquired in different spectral bands or by different sensors (e.g., optical, SAR, LIDAR), automated registration of multi-source remote sensing images could be a difficult task.

Methods of automated image registration can be generally classified into two categories: area-based and feature-based (Zitova and Flusser 2003). The area-based method use the similarity between pixel intensities to determine the alignment between two images, while in the feature-based method, the features (such as edges, ridges, and corners) are extracted firstly, and then the images are aligned by using the similarity of these features.

For multi-source remote sensing images, a number of registration methods have been proposed by researchers. (Zhang *et al.* 2007) proposed a fast matching method for remote sensing images, in which feature points were extracted by Harris operator and a wavelet-pyramid structure was used to speed up the matching process. (Li *et al.* 2006) employed Scale Invariant Feature Transform (SIFT)(Lowe 2004) to the registration of remote sensing images and demonstrated that it was suitable for images with geometric distortion and scale variation. (Yi *et al.* 2008) presented an enhanced SIFT named SR-SIFT (Scale Restriction Scale Invariant Feature Transform),

which improved the correct matching ratio by modifying the dominant orientation of feature points and introducing the scale restriction in the matching process. However, SIFT may be fragile to the registration of images with significant radiometric variation (e.g. optical and SAR)(Kelman *et al.* 2007). (Cole-Rhodes *et al.* 2003) proposed a multi-resolution registration method by optimization of mutual information using a stochastic gradient, and stated it was reliable for multi-spectral remote sensing images. (Suri and Reinartz 2010) also employed mutual information as the similarity measure for the registration of the SAR image and optical images, and their experiments showed that mutual information was robust to the registration of multi-source remote images with radiometric variation. However, since mutual information-based methods are very time-consuming, it may be a restriction in practice. (Wong and Clausi 2007) utilized phase congruency model with local illumination and contrast invariance to the registration of remote sensing images, they first extracted the feature points by the phase congruency minimum moment, and then the matching process was accomplished by using phase congruency moment-based patches as local feature descriptors. Nevertheless, the repeatability of the feature points extracted between multi-source remote sensing images was not taken into count. Low repeatability causes mismatches. To tackle this problem, we presented an automatic registration method based on phase congruency for multi-source remote sensing images. We first calculate the phase congruencies respectively for the input image and the reference image. Minimum moment of phase congruency is applied to select feature points in the input image, followed by a normalized cross correlation (NCC) to determine their correspondences in the reference image. Prior georeferencing information and image pyramid guide the above

---

* Corresponding author. Yuanxin YE, yeyuanxin@whu.edu.cn

matching process, followed by a refinement by means of the least square matching (LSM) method. Large correspondence errors are removed by RANSAC. Finally, the input image is registered to the reference through a projective transform.

2. METHODOLOGY

2.1 Phase Congruency

Phase congruency is a model of feature detection based on frequency domain. It postulates that features in an image are perceived at points where the Fourier components are maximal in phase. (Morrone and Owens 1987) showed that this model successfully explained a number of psychophysical effects in human feature perception, and it is invariant to illumination and contrast conditions.

Given a one-dimensional signal $I(x)$, and its Fourier expansion is

$$I(x) = \sum_n A_n \cos(n\omega x + \varphi_{n_0}) = \sum_n A_n \cos(\varphi_n(x)) \tag{1}$$

where A_n is the amplitude of n-th Fourier component, ω is a constant (usually 2π), φ_{n_0} is initial phase of n-th Fourier component, $\varphi_n(x)$ is local phase of the Fourier component at location x.

Phase congruency function is defined as:

$$PC(x) = \max_{\overline{\varphi}(x) \in [0, 2\pi]} \frac{\sum_n A_n(x) \cos(\varphi_n(x) - \overline{\varphi}(x))}{\sum_n A_n(x)} \tag{2}$$

where the value of $\overline{\varphi}(x)$ maximizes this equation is the amplitude weighted mean local phase angle of all the Fourier components. As it is rather awkward using this equation for computation, (Kovesi 1999) proposed a scheme to calculate the phase congruency by using logarithmic Gabor wavelet, which not only considered noise compensation and frequency spread, but also extended phase congruency into two dimensions. Thus, the phase congruency at each location in the image is defined as:

$$PC(x,y) = \frac{\sum_o \sum_n W_o(x,y) \lfloor A_{no}(x,y) \Delta\Phi_{no}(x,y) - T_o \rfloor}{\sum_o \sum_n A_{no}(x,y) + \varepsilon} \tag{3}$$

$$\Delta\Phi_{no}(x,y) = \cos(\phi_{no}(x,y) - \overline{\varphi}(x,y)) - \left| \sin(\phi_{no}(x,y) - \overline{\phi}(x,y)) \right| \tag{4}$$

where (x, y) is the coordinate of the location, $W_o(x,y)$ represents the weighting factor based on frequency spread, $A_{no}(x,y)$ and $\phi_{no}(x,y)$ are the amplitude and phase at (x, y) at wavelet scale n respectively, $\overline{\phi}(x,y)$ is the weighted mean local phase, $\lfloor \ \rfloor$ denotes that the enclosed quantity is equal to itself when its value is positive, and zero otherwise, T is the noise threshold, and ε is a small constant to avoid division by zero. For detailed information about phase

congruency, readers are referred to the literature (Kovesi 1999).

The phase congruency maps of TM band 3 (visible) and band 5 (infrared) are shown in Fig.1. We can find that though there is an obvious local luminance and contrast difference caused by the radiometric variation between the two images, their phase congruency maps are similar visually to some extent. Therefore, we tried to obtain the correspondences by using cross-correlation technique in the phase congruency map, which will be described in next section.
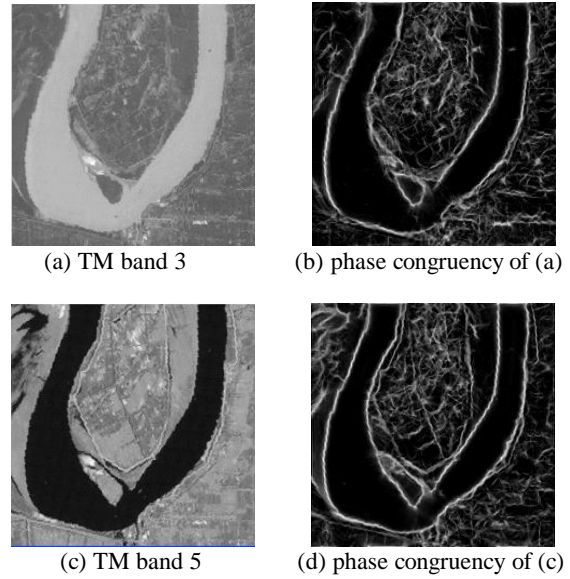

| (a) TM band 3 | (b) phase congruency of (a) |
| (c) TM band 5 | (d) phase congruency of (c) |

Figure 1: TM band3, band5 and their phase congruency maps

2.2 Feature Point Extraction and Matching

The minimum moment of phase congruency indicates the feature point (corner) significance (Kovesi 2003) and can be calculated as the expression in (5)

$$m = \frac{1}{2}\left(\sum_\theta \left[(p(\theta)\sin(\theta))^2 + (p(\theta)\cos(\theta))^2 \right] - \sqrt{4\left(\sum_\theta (P(\theta)\sin(\theta))(P(\theta)\cos(\theta))\right)^2 + \left(\sum_\theta \left[(P(\theta)\cos(\theta))^2 - (P(\theta)\sin(\theta))^2 \right]\right)^2}\right) \tag{5}$$

where $P(\theta)$ is the phase congruency at orientation θ, high values of m indicate high feature point strength. Usually, the points are selected as feature points when their m values is greater than a given threshold value.

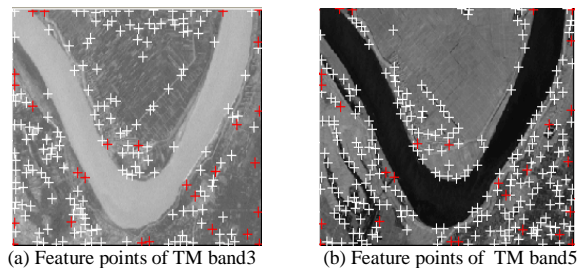| (a) Feature points of TM band3 | (b) Feature points of TM band5 |

Figure 1: Extracted feature points by using the minimum moment of phase congruency

Due to the radiometric variation existing among multi-source remote sensing images，it is very different to extract the feature points with high repeatability. It is shown in Fig. 2 that 193 and 266 feature points are extracted respectively from the TM band 3 and band 5 images of the same area by using the minimum moment of phase congruency. However there are only 27 correspondences with red marking among these feature points, the repeatability is only 13.9 percent.

Since the repeatability of feature points extracted between multi-source remote sensing images is low, our schemes of feature point extraction and matching is as follows. Feature points are initially extracted only in input image. In the reference image we determine a search window，in which every pixel is treated as candidate for matching. After that the point pair having the maximum similarity is regarded as the corresponding point pair. NCC is used as the similarity measure for matching, and it is calculated as the expression in (6):

$$C(x,y) = \frac{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}(A(x,y)-\overline{A})(B(x,y)-\overline{B})}{\sqrt{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}(A(x,y)-\overline{A})^2 \sum_{c=0}^{M-1}\sum_{r=0}^{N-1}(\overline{B}(x,y)-\overline{B})^2}} \quad (6)$$

where $A(x,y)$ and $B(x,y)$ are the phase congruency value of the template window in an image A and in an image B respectively. $\overline{A}$ and $\overline{B}$ represent the mean phase congruency value of the template window, respectively, and M and N are number of rows and columns in the template window.

We select a template window from the input image (TM band 5). Then the NCC is respectively calculated in the original images and in the phase congruency maps for x-direction shift (-20 to 20 pixels) within a search area in the reference image (TM band 3). Fig. 3 shows the above process and the corresponding NCC curve.

(a) Reference image   (b) Input image

(c) NCC in the original images   (d) NCC in phase congruency maps
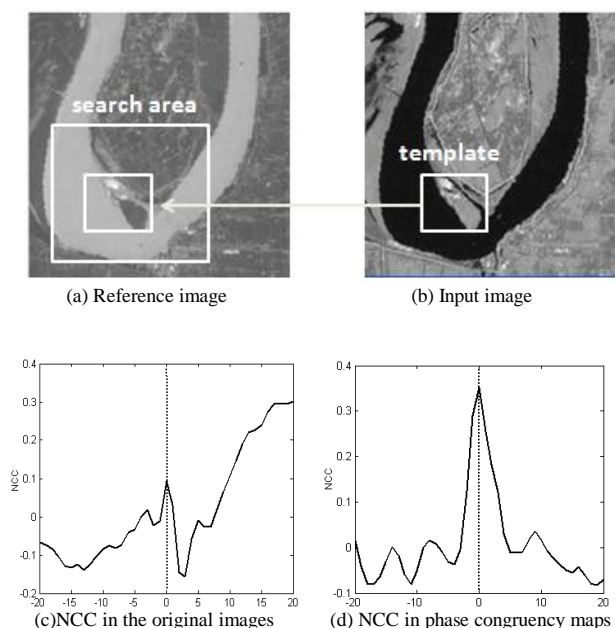
Figure 3: The matching process and the NCC curve

From Fig 3, we can find that the correct match can be obtained when the NCC applied in the phase congruency maps achieves maximum. In contrast, it fails to use the NCC in the original images. Therefore, NCC applied in the phase congruency maps is more robust for the remote sensing images with radiometric variation.

It is computationally expensive that all the pixels in the reference image are regarded as the candidates for matching. With the help of priori georeferencing information, a relatively small search area can be determined in the reference image. Image pyramid is used to guide the matching process, we first find the coarse correspondences in the low resolution, and then refine them until they are matched in the original image. The Gaussian pyramid is employed in our method due to its universality and simplicity，namely，the original image is filtered by the Gaussian function and down-sampled. Once the correspondences have been found, LSM method (Gruen 1985) is applied to achieve subpixel accuracy.

Due to the fixed size of the template window to calculate NCC, it is vulnerable to the images with different resolutions. However, the resolution of remote image is usually available, if the images are at different spatial resolutions, the image of the higher resolution need to be down-sampled to that of the lower resolution.

2.3 Geometric Transform

After the correspondences have been found, it is necessary to determine a geometric transform between the reference image and the input image. Common geometric transform in remote sensing image registration include similarity, affine, projective, polynomial and piecewise linear, etc. The selection of geometric transform depends on the types of relative geometric distortions exhibited in the input image. Since the input remote sensing images in our experiment have been geometrically rectified by using trajectory and pose data, a projective transform (eq 7) can be used to handle most common global geometric distortions. A more complex geometric transform would be necessary for handing more complex distortions caused by spatially variant terrain relief. RANSAC (Fischler and Bolles 1981) is used to remove the large correspondence errors.

$$x = \frac{a_0 + a_1 x' + a_2 y'}{1 + a_3 x' + a_4 y'}$$
$$y = \frac{a_5 + a_6 x' + a_7 y'}{1 + a_3 x' + a_4 y'} \quad (7)$$

where (x,y) is the coordinate in the reference image, and (x',y') is the coordinate in the input image.

2.4 Registration

The proposed registration method consists of the following steps:

(1) The phase congruencies are calculated respectively for the input image and the reference image to reduce local contrast and illumination differences caused by radiometric variation between the images.

(2) The feature points are extracted in the input image by the phase congruency minimum moment. In order to ensure the even distribution of the feature points, we divide the input image into $n \times n$ grids region. In every grid region, the minimum moment value is ordered from large to small, and the strongest k points are selected as the feature points, where k is number of feature points desired.

(3) In the reference image we determine a search area by priori georeferencing information. NCC is used as the similarity measure for matching combined with the search strategy based on the Gaussian image pyramid. The point pairs of which NCC are maximal and greater than a certain threshold t are regarded as the correspondences. Once the correspondences have been found, LSM is applied to achieve subpixel accuracy.

(4) We select the projective transform as the final geometric transformation model, and apply RANSAC to remove the correspondences with large errors.

(5) The input image is rectified by the projective transformation and then resampled by bilinear interpolation algorithm.

3. EXPEREMENT

Four sets of remote sensing images are used to evaluate the proposed method. The test set 1 includes multi-temporal Landsat TM band 3. In the test set 2, the images of different spectra are TM band 5 （infrared） and ETM+ band 3 （visible）. The test set 3 consists of aerial images and LIDAR intensity images. Test set 4 images are respectively visible image and SAR image. There is a radiometric variation among these images, especially the last two test sets. The details of the test sets are listed in Table 1.

For all the test sets, the input images were divided into 5×5 gird region. We extracted 10 feature points in every gird region. The NCC threshold was set to 0.6, and the size of template window was 25×25.

Test	Reference image				Input image			
	sensor	resolution	size(pixel)	date	sensor	resolution	size(pixel)	date
test 1	Landsat 4-5 TM,band:3	30m	745×689	2005-9-11	Landsat 4-5 TM,band:3	30m	694×644	2005-5-6
test 2	Landsat 7 ETM+,band:5	30m	752×751	2004-9-17	Landsat 4-5 TM,band:3	30m	745×691	2005-5-6
test 3	Aerial	0.6m	505×350	2009-7-22	LIDAR	1m	502×345	2009-7-22
test 4	Landsat 7 ETM+,band:3	30m	754×747	2008-9-15	ALOS PALSAR,band:L	10m	660×698	2008-8-6

Table 1. The parameters of the test sets

A set of 25 control points (CPs) was selected manually for each test case. Among them, 15 pairs were used for determining the geometrical transformation, while the reminding 10 pairs were regarded as check points. The root mean squared error (RMSE) was used for accuracy assessment. For evaluation purpose, SR-SIFT algorithm (Yi *et al.* 2008) and Automatic Registration of Remote-Sensing Images (ARRSI) algorithm (Wong and Clausi 2007) were also applied. The registration statistics are shown in Table 2.

From Table 2, we could find that the proposed method was successful for the registration of all the test sets. As the first two test sets are respectively multi-temporal and multi-spectrum images from the same types of sensors, their radiometric variation is relatively small. Their registration accuracies both achieved within 1 pixel. The correspondences and the registration results of the first two test sets are shown in Fig.4 and Fig.5, respectively. For the last two test sets, Due to significant radiometric variation existing between the images, their registration accuracies respectively achieved 1.63 pixels and 1.42 pixels, which was slightly lower than the registration accuracy of the first two test sets. Fig.6 and Fig.7 respectively shows their registration results. These results demonstrate the effectiveness of the proposed method in registering multi-source remote sensing images.

Methods	Test 1		Test 2		Test 3		Test 4	
	Number of correspondences	RMSE (pixel)	Number of correspondences	RMSE (pixel)	Number of correspondences	RMSE (pixel)	Number of correspondences	RMSE (pixel)
manually	15	0.69	15	0.73	15	1.52	15	1.23
SR-SIFT	31	1.03	11	1.45	Fail	Fail	Fail	Fail
ARRSI	42	1.33	25	1.60	12	2.57	21	4.56
the proposed	61	0.56	33	0.79	17	1.63	29	1.42

Table 2. The comparison of registration accuracy of different methods

The proposed method outperformed the other registration methods for all test sets, and the registration accuracy was comparable to that produced by manual registration. SR-SIFT improved the SIFT by modifying the dominant orientation of feature points and increasing the scale restriction in the matching process. It can register the first two test sets. However, it obtained the less correspondences and the lower registration accuracy than the proposed methods due to the effect of radiometric variation. For the last two tests, SR-SIFT even failed to register them. The reason is that the feature descriptor of SR-SIFT is based on the local gradient information, which is vulnerable for registering the images with significant radiometric variation. ARRSI method performs the registration in the phase congruency map as the proposed method does. It extracts feature points both in the reference image and the input image. As a result, the low repeatability of feature points (described in section 2.2) increased the possibility of mismatch and had poorer performance than our method. According to the above analysis, the proposed method is robust for the registration of multi-source remote sensing images.
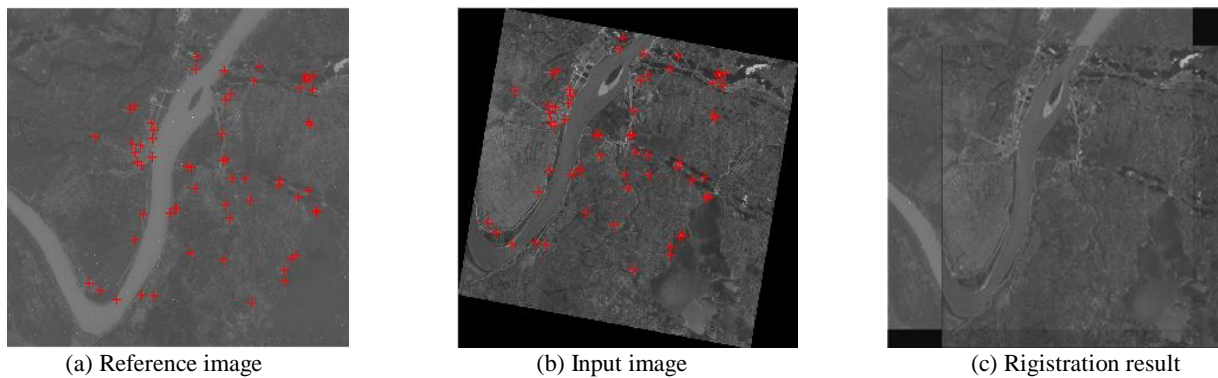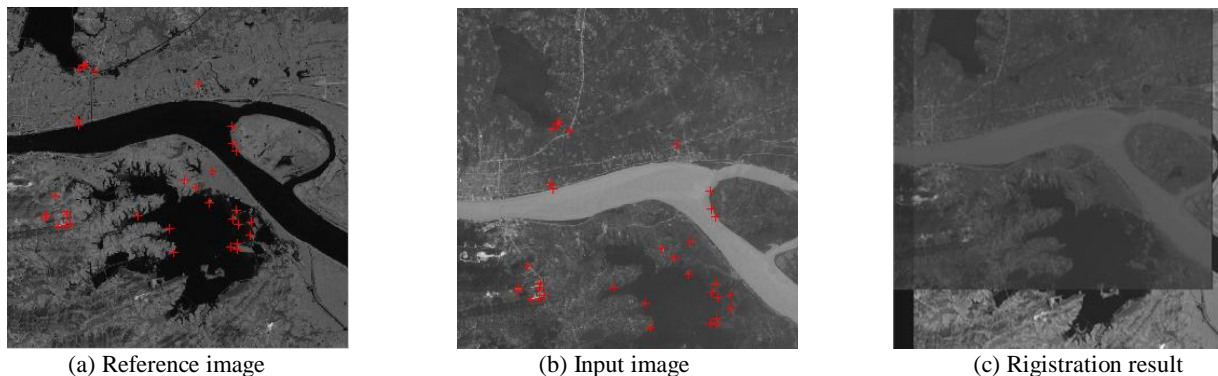
| (a) Reference image | (b) Input image | (c) Rigistration result |

Figure 4: Image registration from test set 1

| (a) Reference image | (b) Input image | (c) Rigistration result |

Figure 5: Image registration from test set 2

| (a) Reference image | (b) Input image | (c) Rigistration result |

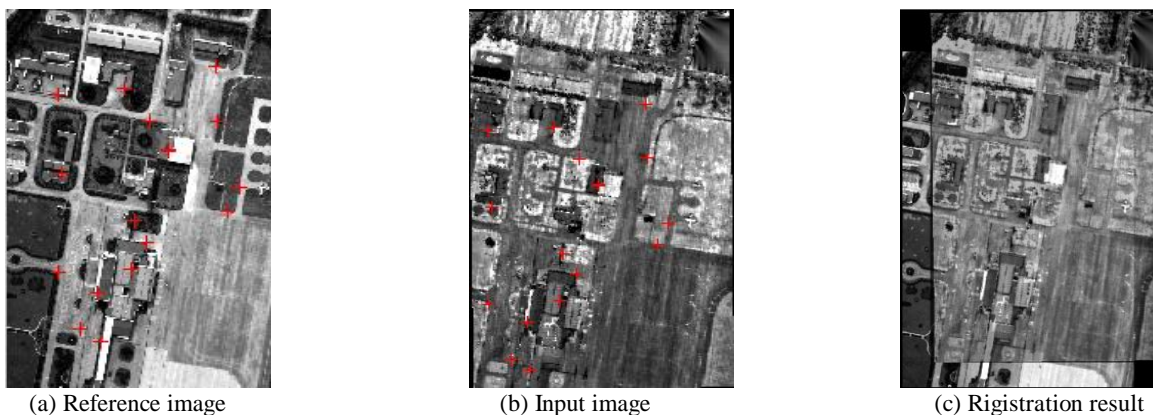Figure 6: Image registration from test set 3

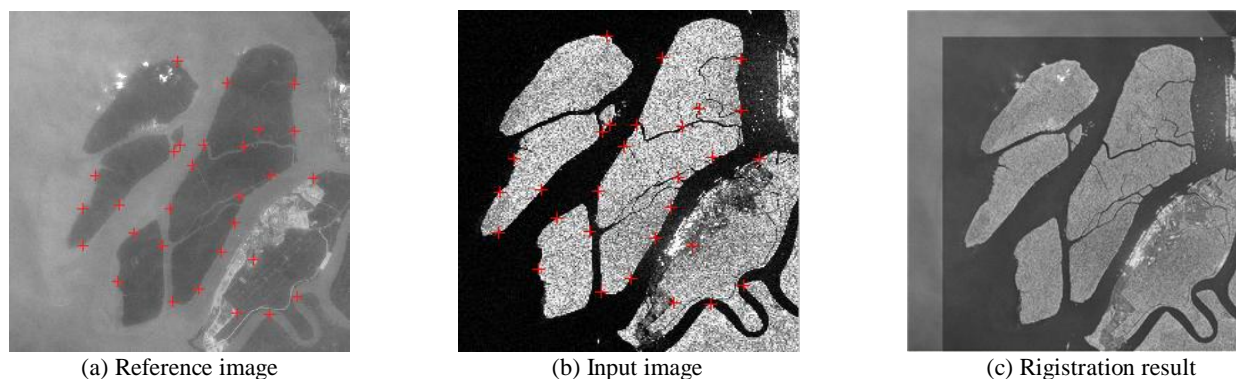| (a) Reference image | (b) Input image | (c) Rigistration result |

Figure 7: Image registration from test set 4

4. CONCLUSION

This paper proposed an automated registration method for multi-source remote images based on phase congruency. The proposed method has utilized phase congruency to reduce the effect of local luminance and contrast difference caused by the radiometric variation between the two images. A variety of techniques, including NCC, image pyramid structure, LSM and RANSAC, are used in the matching process, which is automatic without manual intervention. Experimental results demonstrate that the proposed method is robust for the registration of multi-source remote sensing images. As NCC is calculated in a square template window and a project transform is used as the final geometric transform, the proposed method may be vulnerable to the registration of the images with large geometric deformation and scale variation. The future work will include introducing image scale-space theory and moment invariants technique in the matching process，and use the sensor model or the more complex geometric transform (e.g. rational polynomial function) to hand more complex distortions caused by spatially variant terrain relief.

ACKNOWLEDGEMENT

REFERENCE

Cole-Rhodes, A. A., Johnson, K. L., LeMoigne, J., *et al*, 2003. Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. IEEE Transactions on Image Processing 12(12), pp. 1495-1511.

Zhang, D., Yu, L. and Cai, Z., 2007. Matching method based on Harris operator and wavelet pyramid. Journal of Zhejiang University(Science Edition) 34(4), pp. 465-468.

Fischler, M. A. and Bolles, R. C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24(6), pp. 381-395.

Gruen, A., 1985. Adaptive least squares correlation: a powerful image matching technique. South African Journal of Photogrammetry, Remote Sensing and Cartography 14(3), pp. 175-187.

Kelman, A., Sofka, M. and Stewart, C. V., 2007. Keypoint descriptors for matching across multiple image modalities and non-linear intensity variations. IEEE Conference on Computer Vision and Pattern Recognition, pp. 3257-3263.

Kovesi, P., 1999. Image features from phase congruency. Videre: Journal of Computer Vision Research 1(3), pp. 1-26.

Kovesi, P., 2003. Phase congruency detects corners and edges. in Proceedings of the International Conference on Digital Image Computing, Australia, pp. 309-318.

Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), pp. 91-110.

Morrone, M. C. and Owens, R. A., 1987. Feature Detection from Local Energy. Pattern Recognition Letters 6(5), pp. 303-313.

Suri, S. and Reinartz, P., 2010. Mutual-Information-Based Registration of TerraSAR-X and Ikonos Imagery in Urban Areas. IEEE Transactions on Geoscience and Remote Sensing 48(2), pp. 939-949.

Wong, A. and Clausi, D. A., 2007. ARRSI: Automatic registration of remote-sensing images. IEEE Transactions on Geoscience and Remote Sensing 45(5), pp. 1483-1493.

Li, X., Zheng, L. and Hu, Z., 2006. SIFT Based Automatic Registration of Remotely-sensed Imagery. Journal of Remote Sensing 10(6), pp. 829-835.

Yi, Z., Cao, Z. and Yang, X., 2008. Multi-spectral remote image registration based on SIFT. Electronics Letters 44(2), pp. 107-108.

Zitova, B. and Flusser, J., 2003. Image registration methods: a survey. Image and Vision Computing 21(11), pp. 977-1000.