

A SAR IMAGE REGISTRATION METHOD BASED ON SIFT ALGORITHM

W. Lu ^{a,b}, X. Yue ^{b,c}, Y. Zhao ^{b,c}, C. Han ^{b,c,*}

^a College of Resources and Environment, University of Chinese Academy of Sciences,
Beijing, 100149, China

^b Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences,
Beijing, 100094, China - (luwd, yuexj, zhaoyh, hancm)@radi.ac.cn

^c Department of Airborne remote sensing, Sanya Institute of Remote Sensing,
Sanya, 572029, China

Commission III, WG III/2

KEY WORDS: image registration, scale-invariant feature transform (SIFT), Wallis filtering, adaptive smoothing, image filtering

ABSTRACT:

In order to improve the stability and rapidity of synthetic aperture radar (SAR) images matching, an effective method was presented. Firstly, the adaptive smoothing filtering was employed for image denoising in image processing based on Wallis filtering to avoid the follow-up noise is amplified. Secondly, feature points were extracted by a simplified SIFT algorithm. Finally, the exact matching of the images was achieved with these points. Compared with the existing methods, it not only maintains the richness of features, but also reduces the noise of the image. The simulation results show that the proposed algorithm can achieve better matching effect.

1. INTRODUCTION

Synthetic aperture radar (SAR) is widely used in the military field with its all-weather, all-time, strong penetration and other advantages. Side view imaging can enlarge the measuring range. SAR image registration is a key step for image processing. Because of the differences in image mechanism, images acquired by distinct sensors perhaps appear different intensities and texture patterns. Especially for SAR images, noise and deformable objects of images also make image registration become more difficult. Image preprocessing is significant before image analysis. Meanwhile proposing robust image registration method has paid more and more attention. Generally, image registration methods can be classified into two categories, namely, intensity-based methods and feature-based methods. Compared with intensity-based methods, feature-based methods depending on matching distinctive features have better effectiveness and precision.

Among feature-based methods, an algorithm named scale-invariant feature transform (SIFT) gets more and more attention in the remote sensing fields because of their distinguished performance (Schwind et al., 2000; Lowe, 2004; Wang et al., 2012). Due to its efficiency, this method is widely used in the computer vision field to match images or to recognize and localize objects. Its invariances to scale changes, rotations, and translations and its robustness to illumination changes and affine distortion make it suitable for different kinds of applications such as object retrieval, image indexing, stitching, registration, or video tracking. Therefore, the SIFT algorithm and its variants are a great option for remote sensing images. Such algorithms have been applied mostly to satellite optical images since they have similar characteristics to images acquired with ordinary cameras. Several registration methods use keypoints that SIFT extracts as control points to determine deformation models. Li et al. (2009) took into account the specificity of remote sensing images and introduced a new

matching criterion with scale and orientation restrictions. A multilevel SIFT matching approach is proposed by Huo et al. (2012) to register high resolution images, with the help of Random Sample Consensus (RANSAC). Sedaghat et al. (2011) adapted the algorithm to obtain key points that are uniformly distributed in space and filter the mismatches by applying a projective model. The SIFT algorithm also has some assets for the retrieval of remote sensing images or for classification applications. Zhang et al. (2011) proposed a SAR image registration method based on the divided regions; Xia (2014) combined the corner match with prior information. Object detection is another application field of the SIFT algorithm. Single buildings are detected on very high resolution optical images by Sirmacek and Unsalan (2009) by using SIFT key points, multiple subgraph matching, and graph cut methods. Tao et al. (2011) performed airport detection by considering both clustered SIFT key points and region segmentation. Although the SIFT algorithm has proved its efficiency in the application of optical remote sensing, the situation for SAR images is different. The SAR image is corrupted by strong multiplicative noise, which is known as speckle, and data processing is difficult. Especially, the SIFT algorithm does not work well on this type of image. Therefore, we are supposed to take into account the potential application of SAR images in the local context, it is important to adapt to this new method. Several improvements have been proposed to apply the algorithm to more cases. Some suggest making the images prefilter or denoise to reduce the influence of speckle noise. Others remove some invariances, or modify some steps of the algorithm, to improve the performances. Spatial relationships between key points are considered by Lv et al. (2011) and Fan et al. (2013) to suppress false matches. Performances of these newly proposed algorithms are still relatively insufficient, and the number of correct matches is not sufficient to consider other applications than image registration. However, many of them do not consider the statistical specificities of speckle noise.

* Corresponding author

In this paper, we propose a new algorithm improved based on the SIFT algorithm before. Firstly, the adaptive smoothing filter is conducted on SAR image pairs, which is an adaptive smoothing algorithm based on image gradient. According to the mutation characteristics of gray value of image pixels, we adaptively change the filter weights and sharpen the image edge during the process of smoothing, which better eliminate the contradiction between filtering noise smoothing and sharpening the edge, with great performance on filtering and easy real-time processing. Secondly, the images were filtered through Wallis filter, which utilizes the Gauss smoothing operator in the calculation of mean and variance of local gray value. Moreover, it enhances the useful information of image, suppresses noise, increases the image signal-to-noise ratio, extracts the extremely fuzzy texture information and effectively improves the precision. Finally, the SIFT algorithm is conducted to extract the feature points that are used to match the images accurately. The specific process is described as follows: 1) the scale-invariant feature points in the image are extracted; 2) the detected feature points are described, and the feature descriptors are formed; 3) the original image and the SIFT descriptor of the target image are matched. Comparing with the existing methods, our one not only keeps the richness of image features, but effectively reduces image noise with great accuracy of feature point detection and image matching. Through the experiment of three groups of SAR images, we adjust the parameters to achieve the best results with many trial-and-error tests. The experimental results show that the proposed algorithm can achieve good matching results.

2. PROPOSED METHOD

2.1 Wallis filtering

Wallis filtering is a kind of local image transformation, the aim of Wallis filter is to map gray scale mean and variance of image into given gray scale mean and variance. Through this filtering, the contrast of the image becomes more uniform. Because this filter uses a smooth operator in computing local gray scale mean and variance, it can enhance much useful information of image and weaken noise to some extent, which increases the SNR (signal to noise ratio) of image. The general formula of Wallis filter is:

$$\begin{cases} f(x, y) = g(x, y)r_1 + r_0 \\ r_1 = \frac{cs_f}{(cs_g + (1-c)s_f)} \\ r_0 = bm_f + (1-b-r_1)m_g \end{cases} \quad (1)$$

Where r_1 is a multiplicative parameter and r_0 is an additive parameter. If $r_1 > 1$, transformation (1) becomes a high pass filter; else if $r_1 < 1$, transformation (1) is a low pass filter. m_g expresses gray scale mean of image in a certain area of a pixel and can be regarded as the DC component of image signal. s_g can be called the AC component of image signal, which means gray scale variance of image in a certain area of a pixel. $g(x, y)$ and $f(x, y)$ are original image and the filtered image. m_f is a target value of gray scale mean, and it must be a median of dynamic image range. s_f is a target value of gray scale variance, and it determines the contrast of image. c is a contrast expansion constant of image, $c \in [0, 1]$, it should increase with the window increasing in the process. b is an image intensity coefficient. When $b \rightarrow 1$, the mean of image is forced to m_f , and when $b \rightarrow 0$, the mean of image is compelled to m_g . Therefore, in order to maintain the gray mean of the original image, b should be used the smaller value. Multiplicative coefficient r_1 decides the

performance of this filter. And its relationship with others parameter can be rewritten in the following:

$$\frac{1}{r_1} = \frac{1}{c^2} + \frac{s_g}{s_f} \quad (2)$$

To the typical Wallis filter, $c=1$, $b=1$, hence:

$$f(x, y) = [g(x, y) - m_g] \cdot \frac{s_f}{s_g} + m_f \quad (3)$$

2.2 Adaptive smoothing filtering

In order to avoid the use of Wallis filter to enhance the image, and increase the noise information on the image, this paper uses adaptive smoothing filter to denoise the image before Wallis filtering. Saint-Marc et al.(1991) proposed adaptive smoothing filtering method, which is used to smooth the image while still retaining the image discontinuities. The basic idea of this method is using a locally weighted iterative convolution template and the original image signal, the process has the properties of anisotropic diffusion, the weighted coefficient of each pixel in each iteration is changed, it is the gradient function of the pixel, reflect the continuity of image pixel gray value. Adaptive smoothing filter is an iterative process, the maximum number of iterations is k , the specific steps of the algorithm are as follows:

- 1) Initial iteration times $n = 0$, set the maximum number of iterations k and the value of the parameter h ;
- 2) Compute the gradient of the n iteration: $G_x^{(n)}(x, y)$ and

$$G_y^{(n)}(x, y):$$

$$\begin{cases} G_x^{(n)}(x, y) = \frac{1}{2}(I^{(n)}(x+1, y) - I^{(n)}(x-1, y)) \\ G_y^{(n)}(x, y) = \frac{1}{2}(I^{(n)}(x, y+1) - I^{(n)}(x, y-1)) \end{cases} \quad (4)$$

- 3) Compute the window weight coefficients of the n iteration: $\omega^{(n)}(x, y)$

$$\omega^{(n)}(x, y) = \exp\left(-\frac{[G_x^{(n)}(x, y)]^2 + [G_y^{(n)}(x, y)]^2}{2h^2}\right) \quad (5)$$

- 4) Compute the convolution:

$$I^{(n+1)}(x, y) = \frac{\sum_{i=-1}^1 \sum_{j=-1}^1 I^{(n)}(x+i, y+j) \omega^{(n)}(x+i, y+j)}{\sum_{i=-1}^1 \sum_{j=-1}^1 \omega^{(n)}(x+i, y+j)} \quad (6)$$

- 5) Whether $n = k$, if $n = k$, Stop the iteration; otherwise $n = n + 1$, go to step 2).

The weight coefficient W is obtained by the image gradient and the parameter h , and the parameter H value determines the degree of the edge detail information in the image smoothing process. h larger, the edge of the details will be smooth out; h smaller, the edge and the noise are both retained, no smoothing effect. Therefore, the h value plays a decisive role in the final image smoothing effect. Parameter h usually takes 1.5 to 2 times the standard deviation.

2.3 SIFT algorithm

Through a Gaussian differential function, the SIFT algorithm generates a scale space, and selects local extreme point as the

candidate feature point, then removes the unstable low contrast and edge response points, and precisely positions feature points. Image data is sampled down, and the image pyramid is constructed. Gaussian convolution kernel can achieve the scale transformation. Two-dimensional image can be formed as image scale space with the convolution to different kernel Gaussian filters. As follows:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (7)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (8)$$

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (9)$$

Wherein, $I(x, y)$ is an input image, G is the Gaussian kernel function, σ is a variable kernel. Extreme point of Gaussian Laplace can provide stable characteristic for the image. Gaussian Laplace function can be approximated by Gaussian differential function. And the extreme points can be obtained by Gaussian differential function. Gaussian function fits to the three-dimensional quadratic function to achieve the positioning accuracy of the feature points. Gaussian differential function can be expanded by Taylor formula on candidate extreme points. Detecting extreme point of differenced Gaussian has a strong edge response. There is a proportional relationship between the principal curvatures of the edge points and its eigenvalue of the Hessian matrix. Calculate the eigenvalue of the Hessian matrix to determine whether it is the stable point.

Lowe proposed the method to construct feature points' descriptors using the gradient distribution of the feature points' neighbor pixels. Gradient magnitude and direction of pixel (x, y) are expressed as follows: $m(x, y)$ and $\theta(x, y)$ represent the magnitude and orientation of the gradient, respectively. L is the scale of each key point.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (10)$$

$$\theta(x, y) = \arctan \frac{L(x+1, y) - L(x-1, y)}{L(x, y+1) - L(x, y-1)} \quad (11)$$

The gradient distribution is counted by the center pixel of the feature points. Histogram about the gradient direction is divided into 36 columns, and the peak value indicates as the main direction of the feature points. Rotate the axis of the image to the main direction of the feature points. Use Gaussian weighted principal to calculate the accumulated value of the pixel in the window in each direction of the gradient. Neighbor window is divided into 16 sub-blocks of 4×4 , and statistic the accumulated value of gradient histogram of 8 directions in each sub-block, then generate a seed point. The final descriptor is a vector holding 128 elements.

Euclidean distance between feature descriptors is the similarity measure used to match two features. The two closest features in the input image are determined for every reference image feature. If the Euclidean distance of the second-closest match is smaller than the threshold (0.6) times the distance of the closest match, the input feature is accepted as a match. Since comparing the distance of all features with each other would be too expensive, the approximate nearest neighbor best-bin-first (BBF) algorithm is used.

3. EXPERIMENTS AND RESULTS

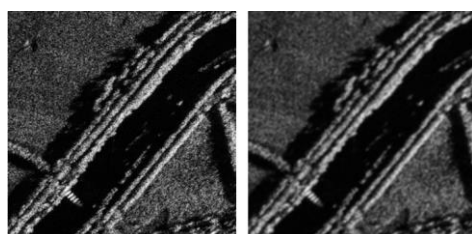
This paper proposes an image registration algorithm based on SIFT, this algorithm uses adaptive filtering and Wallis filtering

to process the image before extracting the feature. Adaptive filter is designed to eliminate noise, Wallis filtering is to increase the image texture information. Specific steps are as follows:

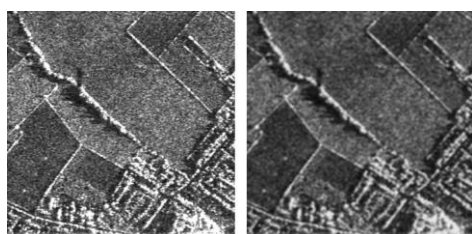
- 1) The image is denoised by adaptive smoothing filtering, detailed procedures are shown in section 2.2.
- 2) The image is divided into a fixed size, nonoverlapping rectangular window area, the scale of the window should correspond to the scale of the texture pattern to be enhanced, and the m_g and s_g of the window are calculated.
- 3) According to the parameters that input, the multiplicative coefficient r_1 and the additive coefficient r_0 of each region Wallis filter are calculated. the parameters m_f and s_f values are set to 127 and the value from 40 to 70. The parameter s_f should be reduced with the decrease of the regional scale, so as to avoid the saturation of the gray value of a large number of pixels (beyond [0,255]). The range of image contrast expansion constant value c is [0,1]. In this paper, the value [0.75,1] can be used to prevent the low gray value of image that is enhanced. Image brightness coefficient b ranges [0,1], this paper uses the value ranges [0.5,1] that this paper used can maintain the original image of the local gray mean.
- 4) According to the formula (1), r_1 and r_0 of each region are used to calculate the gray value of the Wallis transform of the central point of the region, and we can remember them as matrix R , Mean value and standard deviation of gray value after the transformation of all regions are calculated and recorded as m_R and s_R .
- 5) The original image $g(x, y)$ is subjected to Wallis filtering transform by pixel. Since each window does not overlap, the coefficients of any pixel on image are obtained by bilinear interpolation and the new gray value of all pixels is calculated according to equation (1).
- 6) According to equation (3), the image that was filtered with Wallis will be subjected to a global classical Wallis filter transform again. Where m_g and s_g are replaced by m_R and s_R . The process can further enhance the degree of image sharpening and increase the image contrast.
- 7) Extract the SIFT feature points and match the images that were filtered, output the images.

Wallis filtering is applied to the SIFT matching process. The experiment uses three pairs of images for feature extraction, which contains a representative of the surface features information, so that the experiment has a good persuasion. The image size used in this experiment is 200 size 200×200 pixels. Firstly, the improved Wallis filter is used to enhance the image. Wallis filter parameters are set to: Rectangular area size 39×39 pixels, m_f and s_f were 60 and 127 respectively, b and c were 0.6 and 0.75 respectively. Then, the enhanced image is matched by SIFT, and the results are compared with the original SIFT feature extraction. In order to verify the effect of the adaptive smoothing filter in the Wallis filtering process, two SAR images were selected to perform the adaptive smooth filtering denoising experiment. During the experiment, the k was 4, and the h was about 1.75 times the standard deviation. Figure 1 is the original image and its corresponding adaptive smooth filtered image. We objectively evaluate the filtering effect using PSNR (peak

signal noise ratio), NMSE (normalized mean square error) and SSIM (structural similarity), the statistical results shown in Table 1. As it shown that the smaller the ENMS, the better the filtering effect, the bigger the RPSN, the better the filtering effect, and the closer the SSIM to 1, the better the image structure is. The experimental results show that the adaptive filtering algorithm not only can filter out the noise, but also protect the edge of the image, whether it is from the visual effect of the filtered image or the image quantitative evaluation index after denoising, is conducive to the Wallis filter processing.



(a) Data 1



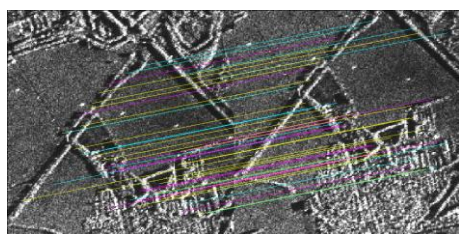
(b) Data 2

Figure 1. Original image with adaptive smoothing filtered image

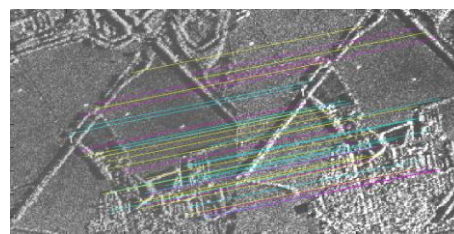
image	E_{NMS}	R_{PSN}	S_{SIM}
Data1	0.00024	38.78	0.96
Data2	0.00132	32.23	0.94

Table 1. image filtering results evaluation statistics

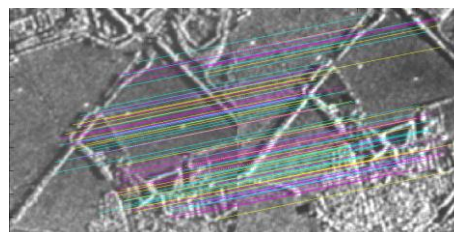
Image enhancement processing and feature extraction results are shown in Figure 2. Figure 2 (a) shows the matching results for the original image. The original image that the gray scale is uniform, but part of the edge area contrast is not obvious. The images that used the Wallis filtering and Wallis filtering combined with adaptive filtering to enhance are shown in (b) and (c) respectively.



(a) Diagram of SIFT matching points connection of original image



(b) Diagram of SIFT matching points connection after Wallis filtering



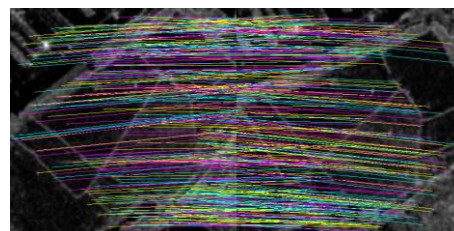
(c) Diagram of SIFT matching points connection after filtering twice

Figure 2. The matching result of SIFT algorithm to the enhanced image

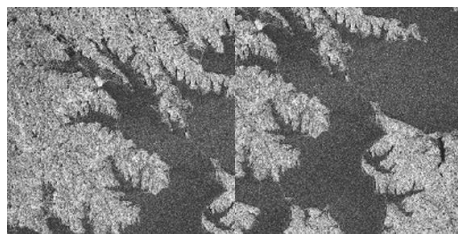
It can be seen from the experimental results that the original image is enhanced and the region of the original image with a small change in gray scale is significantly enhanced by Wallis filtering. The difference between the two algorithms is not clearly differentiated by visualization, but the results of feature extraction with SIFT are quite different. The number of matching points is 50, 45, and 102. Some points on the edge line have not been extracted, because the contrast on the original image is weak. However, the number of feature points matched on the image after Wallis filtering enhancement is the least, because the image has not been denoised, the contrast and noise are magnified at the same time. It is better to extract the feature points on the image after two filtering enhancement, corner and edge points are well extracted. The other two sets of original images and matching results are shown in Figures 3.



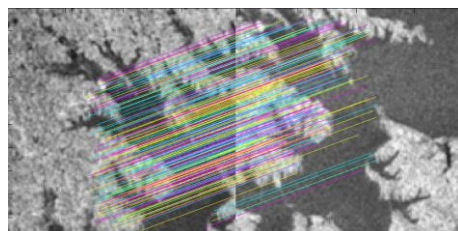
(a) Original image, data 2



(b) Diagram of SIFT matching points connection after filtering twice



(c) Original image, data 3



(d) Diagram of SIFT matching points connection after filtering twice

Figure 3. Original image and the matching result of SIFT algorithm to the enhanced image

Finally, the RANSAC (Random Sample Consensus) algorithm is used to filter the point pairs of SIFT matching and remove the wrong matching points so that the correct matching rate is obtained. The comparison of SIFT algorithm and the proposed algorithm is shown in Table 2. The results show that the SIFT algorithm can effectively extract the stable matching point, but its point is less and the correct rate is not very high. While using Wallis filtering combined with adaptive smoothing filtering before the SIFT match, the number of point pairs and the correct rate has been significantly improved.

No.	matching points	correct points	correct rate (%)
	SIFT/Proposed method	SIFT/Proposed method	SIFT/Proposed method
1	50/110	45/108	88/98
2	228/329	203/319	89/97
3	83/305	71/278	86/91

Table 1. Result of traditional SIFT algorithm and proposed algorithm

4. CONCLUSION

Under the condition of serious speckle noise and similar texture structure in SAR images, the method based on SIFT features can still be matched to a large number of stable points. This paper selects 3 groups of different categories of SAR image data and uses SIFT algorithm combined with adaptive smoothing filtering and Wallis filtering to match the images, and then the RANSAC algorithm is used to eliminate the error of the matching points, in order to improve the accuracy of matching points. The experimental result shows that the SAR image SIFT algorithm can match accurately to the stable characteristics and has certain robustness. Together with Wallis filtering and adaptive filtering algorithm, it can effectively cut down the image noise and enhance the image texture, and the effective number of matching points and correct rate are increased, so this method can be applied to the subsequent processing of SAR image data.

ACKNOWLEDGEMENTS

The author would like to thank the anonymous reviewers for their valuable comments and suggestions. This work was supported by the key research and development plan of Hainan province under Grant NO. SY16ZY02132.

REFERENCES

- Schwind, P., Suri, S., Reinartz, P., and Siebert, A., 2010. Applicability of the SIFT operator to geometric SAR image registration. *Int. J. Remote Sens.*, vol. 31, no. 8, pp. 1959–1980.
- Lowe, D. G., 2004. Distinctive image features from scale-invariant key points. *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110.
- Wang, S., You, H., and Fu, K., 2012. BFSIFT: A novel method to find feature matches for SAR image registration. *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 649–653.
- Li, Q., Wang, G., Liu, J., and Chen, S., 2009. Robust scale-invariant feature matching for remote sensing image registration. *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 287–291.
- Huo, C., Pan, C., Huo, L., and Zhou, Z., 2012. Multilevel SIFT matching for large-size VHR image registration. *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 171–175.
- Sedaghat, A., Mokhtarzade, M., and Ebadi, H., 2011. Uniform robust scale invariant feature matching for optical remote sensing images. *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4516–4527.
- Zhang, B., Tian, Z., Yan, W., 2011. An SAR Image Registration Algorithm Based on Segmentation-derived Regions. *Chinese journal of engineering mathematics*. 28(1), pp. 7–14.
- Xia, G., 2014. Research on image sequence alignment method for radar of WAS SAR-GMTI large angle of view. Xi'an Electronic and Science University.
- Sirmacek, B. and Unsalan, C., 2009. Urban-area and building detection using SIFT keypoints and graph theory. *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1156–1167.
- Tao, C., Tan, Y., Cai, H., and Tian, J., 2011. Airport detection from large IKONOS images using clustered SIFT keypoints and region information. *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 1, pp. 128–132.
- Lv, W., Yu, W., Wang, J., and Wang, K., 2011. SAR image matching based on SIFT keypoints and multi-subregions information. in *Int. Asia-Pac. Conf. Synthetic Aperture Radar*, pp. 1–4.
- Fan, B., Huo, C., Pan, C., and Kong, Q., 2013. 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, pp. 657–661.
- Saint-Marc, P., Chen, J.S., and Medioni, G., 1991. Adaptive smoothing: a general tool for early vision. *IEEE Trans. Pattern Analysis and Machine Intelligence*. 13(6): pp. 514–529

Revised June 2017