MULTI-SPECTRAL REMOTE SENSING IMAGE REGISTRATION VIA SPATIAL RELATIONSHIP ANALYSIS ON SIFT KEYPOINTS

Mahmudul Hasan[1], *Xiuping Jia*[1], *Antonio Robles-Kelly*[2,3,1], *Jun Zhou,*[2,3,1] , *Mark R. Pickering*[1]

[1]School of Engineering and Information Technology, University College, The University of New South Wales, Australian Defence Force Academy, Canberra, ACT, Australia.
[2]National Information and Communications Technology Australia, Locked Bag 8001, Canberra, ACT, Australia
[3]College of Engineering and Computer Science, The Australian National University, Canberra, Australia

1. INTRODUCTION

Multi-sensor image registration is a challenging task in remote sensing. Considering the fact that multi-sensor devices capture the images at wide range of frequencies and at different time, multi-spectral image registration is necessary for data fusion of the images. Several conventional methods for image registration suffer from poor performance due to their sensitivity to scale and intensity variation. The scale invariant feature transform (SIFT) [1] is widely used for image registration and object recognition to address these problems. However, directly applying SIFT to remote sensing image registration often results in a very large number of feature points but small number of matching with a high false alarm rate. This is due to the fact that spatial information is not considered during the SIFT-based matching process. This paper proposes a method to improve SIFT-based matching by taking advantage of neighborhood information. The proposed method generates more correct matching points as the relative structure in different remote sensing images are almost static.

2. PROBLEMS FOR SIFT FEATURE MATCHING

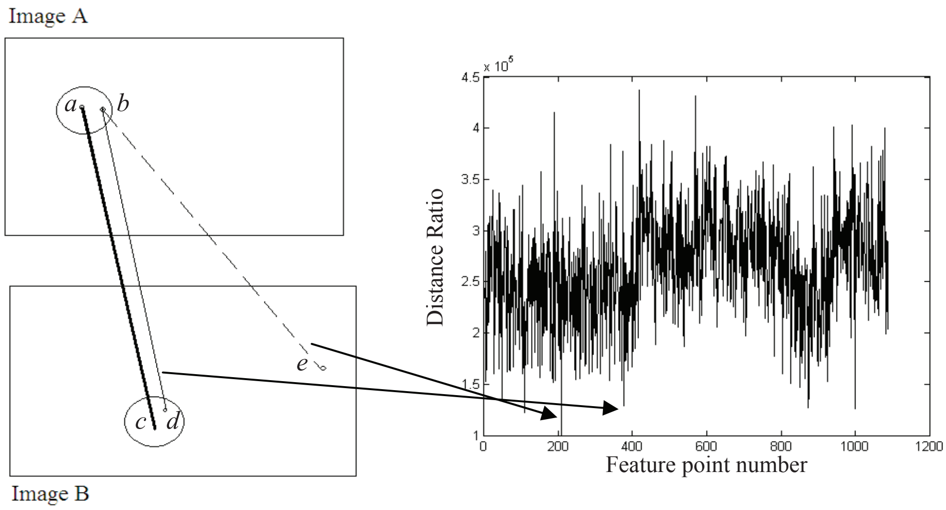
The standard way of applying SIFT to image registration is as follows. The first step is to detect feature points in the scale space [2] using a Difference of Gaussian (DoG) filter. Then a 128-element feature descriptor is generated for each feature point using statistics of the gradient directions which are scale and rotation invariant. These descriptors are used to find matching points by calculating the ratio of the Euclidian distance between every feature point in the images to be registered.

Lowe [1] did not consider the neighborhood relationship of feature points in the spatial space of the images as the method was targeted at object retrieval where the locations, poses, and spatial relations of the objects to be retrieved can be quite different in two images. On the contrary, in the case of remote sensing image registration, we assume in most cases that the spatial relationship of the objects within an image does not experience a significant change within another image subject to rotation and scaling. For example, Qiaoliang [3], Yi [4], and Vural [5] suggested modifications to SIFT for better matching accuracy by imposing scale and orientation restriction. What makes the spatial relationship more important is that similar feature descriptors may be found in many locations, such as from buildings with similar shapes, which is common in remote sensing images. Thus, imposing a location restriction in feature point matching is the underlying principle of the technique we propose in this paper.

To illustrate the issue, we show an example in Figure 1. SIFT matching is applied to images A and B. The bold line shows a pair *(a, c)* of matched featured points in the two images. The dotted line shows the best match *e* of another feature point *b* in image A, while the correct match should be point *d*. In the proposed technique, *e* is not selected as a matched feature point for *b* because the spatial distance between points *c* and *e* is too large.

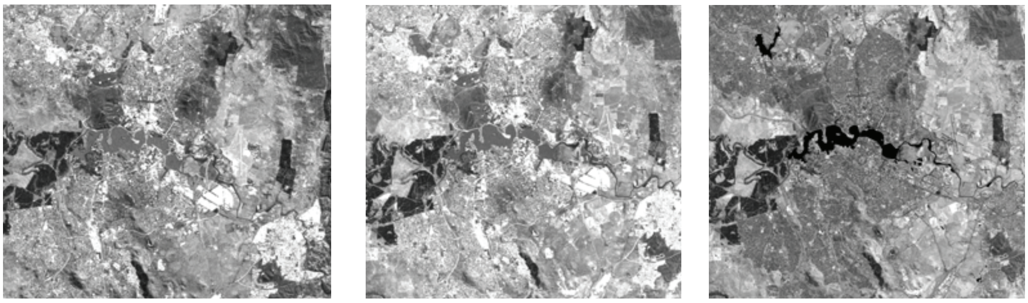
3. MATCHING FEATURE POINT USING STRUCTURAL INFORMATION

As we mentioned above, in multi-sensor remote sensing images, the spatial relationship between objects remains approximately the same. Thus, if we can find the matching position of a feature point, we can predict the matching position of the neighboring feature points. Note that each feature point is associated with a scale and an orientation via SIFT, so from a pair of matched feature points the scale difference for surrounding points can be predicted. In Figure 1, feature points *a* and *c* are matched while the counterpart for neighboring feature point *b* cannot be decided because the SIFT descriptors for points *d* and *e* are almost equally different from the SIFT descriptor of point *b*. This problem is made worse by the fact that a lot of similar descriptors can be found in typical remote sensing images. The idea to solve this problem is as follows. Still considering the example in Figure 1, assume points *a* and *c* are already matched with high confidence that the match is correct. We can predict that the feature points around *a* (shown in the circular window) can be found around *c*. So, for point *b* we only search the neighborhood of point *c* for a matching descriptor, which results in a correct match at point *d*. This process is iterated to recover more matching feature points and hence a more accurate registration. To find the final matching points, we used the RANSAC algorithm [6].



(a) (b)

Figure 1: (a) Feature point matching by SIFT; a to c, b to e. (b) Euclidian distance ratio of all the feature point on image B for the feature point b on image A.



(a) (b) (c)

Figure 2: (a) Band 1 of LANDSAT image taken in 2001 around Canberra with image size 703×654. (b) Band 1 and (c) Band 5 of LANDSAT image taken in 2000 around Canberra with image size 849×773.

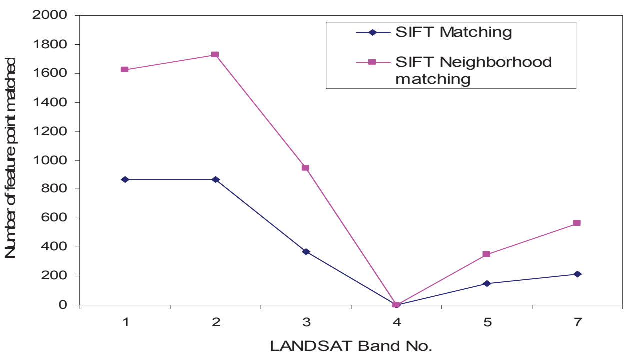


Figure 3: Number of feature points correctly matched in image registration between band 1 of LANDSAT image taken in 2001 and band 1 to 5 and 7 of LANDSAT image taken in 2000.

To evaluate the performance of the proposed technique we conducted an experiment using band 1 of a LANDSAT image taken in 2001 and bands 1, 2, 3, 4, 5 and 7 of a LANDSAT image taken in 2000 as shown in Figure 2. Both SIFT and our method (SIFT Neighborhood matching) were applied to find matching feature points. The year 2001 band 1 image produced 9901 feature points and the year 2000 data produced 7383, 8240, 8087, 8337, 7870, 7186 feature points for bands 1 to 6 respectively. The RANSAC method was applied to remove the false matches. Figure 3 shows the numbers of feature points correctly matched using the two algorithms. Both the method was unable to register band 1 of LANSAT image taken in 2001 with band 4 of LANDSAT image taken in 2000 because band 1 is captured at visible blue and band 4 is captured at near infra-red channel that caused a big amount of variation in pixel intensities between the images. SIFT produced less than 11% matching feature points, while the proposed neighborhood matching algorithm improved the performance to almost twice this amount in most cases for a neighborhood window with a radius of 100 pixels.

4. REFERENCES

[1] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision,* vol. 60, pp. 91-110, 2004.

[2] T. Lindeberg, *Scale-space theory in computer vision*: Kluwer Academic Publishers, 1994.

[3] L. Qiaoliang, W. Guoyou, L. Jianguo, and C. Shaobo, "Robust Scale-Invariant Feature Matching for Remote Sensing Image Registration," *Geoscience and Remote Sensing Letters, IEEE,* vol. 6, pp. 287-291, 2009.

[4] Z. Yi, C. Zhiguo, and X. Yang, "Multi-spectral remote image registration based on SIFT," *Electronics Letters,* vol. 44, pp. 107-108, 2008.

[5] M. F. Vural, Y. Yardimci, and A. Temizel, "Registration of Multispectral Satellite Images With Orientation-Restricted SIFT," in *IEEE International Geoscience & Remote Sensing Symposium (IGARSS)*, 2009.

[6] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM,* vol. 24, pp. 381-395, 1981.