Coarse-to-Fine Registration of Remote Sensing Optical Images using SIFT and SPSA Optimization

Sourabh Paul[1], Umesh C. Pati[2]

Department of Electronics and Communication, NIT Rourkela, Rourkela-769008, Odisha, India.

{[1] sourabhpaul26@gmail.com ; [2] ucpati@nitrkl.ac.in }

Abstract. Sub-pixel accuracy is the vital requirement of remote sensing optical image registration. For this purpose, a coarse-to-fine registration algorithm is proposed to register the remote sensing optical images. The coarse registration operation is performed by the scale-invariant feature transform (SIFT) approach with an outlier removal method. The outliers are removed by the Random sample consensus (RANSAC) algorithm. The fine registration process is performed by maximizing the mutual information between the input images using the first order simultaneous perturbation stochastic approximation (SPSA) along with the second order SPSA. To verify the effectiveness of the proposed method, experiments are performed by using three sets of optical image pairs.

Keywords: Image registration; Scale invariant feature transform (SIFT); Simultaneous perturbation stochastic approximation (SPSA).

1 Introduction

Remote sensing image registration is the process of geometrical alignment of remotely sensed images captured by same or different sensors. Image registration methods can be classified as intensity-based methods [1] and feature-based methods [2]. In intensity-based method, the pixels intensities of the image pair are used to measure the similarity between the images. The similarity measurement matrices used for intensity-based methods are maximum likelihood, least square matching, cross-correlation, mutual information etc. In intensity-based method, an optimization technique is utilized with a similarity matrix to speed up the registration operation. In feature-based methods, the robust features such as corner, edges, contours etc. are extracted and the matching operation is performed.

Scale invariant feature transform (SIFT) [3] is well known approach used for the feature-based image registration. SIFT is a very effective approach to extract the distinctive invariant features and it can be used to perform the matching operation between the images. Regardless of its robustness, some problems still arise such as the uneven distribution of the matched features and the existence of outliers in the matched pairs. So, different modified SIFT algorithms have been proposed in the past to improve the performance of SIFT approach. Goncalves et al. [4] proposed an automatic image registration algorithm (IS-SIFT) using image segmentation and SIFT. In [4], the bivariate histogram is utilized to remove the outliers obtained from the SIFT feature matching operation. Sedaghat et al. [5] presented a uniform robust scale invariant feature transform (UR-SIFT) algorithm to uniformly distribute the extracted features. In [5], uniform SIFT features are matched through cross matching technique to increase the number of matched pairs. Gong et al. [6] developed a coarse-to-fine registration scheme using SIFT and mutual information. In [6], coarse registration was performed by the standard SIFT with reliable outlier method and the fine registration was implemented by maximizing the mutual information using a modified Marquardt-Levenberg optimization technique. Zhang et al. [7] proposed a coarse-to-fine registration algorithm to register the large size very high resolution images using SIFT, Oriented FAST and Rotated BRIEF (ORB) feature matching. In [8, 9], improved SIFT based matching is performed to registered the remote sensing images.

The mutual information is a popular similarity matrix used for the intensity-based image registration methods. Cole-Rodes et al. [10] presented a first order Simultaneous Perturbation Stochastic Approximation (SPSA) optimization with the maximization of mutual information criterion to register the remote sensing images. But, the first order SPSA converges slowly if the result is very close to the optimum solution. In order to increase the computational speed of the optimization technique, Cole-Rodes et al. [11] utilized a second order SPSA algorithm. Suri et al. [12] presented a coarse-to-fine registration algorithm to register the IKONOS and TERRA SAR images through image segmentation and maximization of mutual information criterion using the first order SPSA optimization technique.

Although a variety of feature-based and intensity-based registration algorithms have been proposed to register the remotely sensed optical images, sub-pixel accuracy is still a vital challenge in remote sensing image registration. In this paper, we have proposed a coarse-to-fine registration method to obtain the sub-pixel accuracy between the registered images. Motivated by methods [6] and [7], we have performed the coarse registration by using the SIFT feature matching. The matches obtained by SIFT matching contain many outliers. Random sample consensus (RANSAC) [13] is implemented to refine the matches. The fine registration scheme is conducted by maximizing the mutual information with the first and second order SPSA optimization approach. Inspired by the accuracy and the convergence rate of the algorithms [10] and [11], the first order and second order SPSA are utilized for the fine registration process.

Rest of the paper is organized as follows: Section 2 presents the initial coarse registration method followed by Section 3, which provides the fine registration algorithm. The simulation results are discussed in Section 4. Finally, Section 5 offers a conclusion.

2 Coarse Registration

The coarse registration process is performed by standard SIFT matching with RANSAC-based outlier removal process. The SIFT algorithm is implemented for feature extraction and matching of the remote sensing optical images. The algorithm consists of five major steps:

2.1 Scale-space extrema detection

The extrema points are detected in each scale of every octave of the Difference of Gaussian (DOG) images by comparing every pixel to its eight surrounding neighbors of current scale, nine neighbors in the upper and lower scales. A particular key point is selected if it is larger or smaller than all of its 26 neighbors.

2.2 Key-points localization

The location and scale of every feature point is estimated by 3D quadratic function. For eliminating the poorly localized features principal curvature analysis is performed. The features with the contrast value less than 0.03 (proposed by Lowe [3]) are discarded.

2.3 Orientation Assignment

The local gradients directions are estimated to compute the dominant orientation of a feature point. Therefore, the feature points are invariant to any rotation of the image.

2.4 Key-points Descriptor Formation

The gradient magnitudes and orientations in a 16x16 location grid around the key-point location are computed to form the 128 elements descriptor.

2.5 Feature Matching

It is performed by calculating the minimum Euclidean distance between the descriptors of input images. But, directly using this minimum distance criteria produces many false matches. So, Lowe [3] proposed d_{ratio} factor to improve the correct rate. The ratio of the Euclidean distance of first nearest neighbor to the second nearest neighbor is defined as d_{ratio}. By choosing a high d_{ratio} value, the number of matched pairs can be increased, but the correct rate decreases [6]. A very small value of d_{ratio} can improve the correct rate, but in this case, the number of matches decrease. So, a proper selection of d_{ratio} is an important factor for remote sensing image registration. In our proposed method, the value of d_{ratio} is set to 0.6. But, a number of outliers still exists in the matched pairs set. So, RANSAC [13] algorithm is implemented to remove the remaining outliers.

3 Fine Registration

The fine registration is performed by maximizing the mutual information using the first order SPSA along with the second order SPSA optimization techniques. In [11], Cole-Rodes et al. proposed a switching scheme between first order and second order SPSA to register the remote sensing images in a multiresolution framework. The second order SPSA provides a fast convergence, whereas the first order is more robust to get a solution which is close to the optimum one from a further away point. The algorithm is very effective when the input images have low distortion between them. But, it fails to register the images if a large distortion occurs between the images. Our proposed coarse registration method coarsely align the input images. Using our coarse registration process, a very low distortion can be obtained between coarsely registered images. So, we have utilized SPSA in the fine registration process to improve the accuracy in registration.

3.1 Mutual Information

Mutual information is the measurement of relative entropy between two functions. Let A and B are the two input images for registration. Then, mutual information between A and B is defined as

$$MI = H(A) + H(B) - H(A,B) \tag{1}$$

where $H(A,B)$ is the joint entropy and $H(A)$ and $H(B)$ represent the marginal entropy of A and B respectively.

3.2 Multiresolution Approach

To speed up the fine registration operation the SPSA optimization is used in a multi-resolution framework. The reference and sensed images are decomposed through the Simoncelli steerable filter. The low pass bands of the filter are iteratively registered from the coarsest level to the finest level by maximizing the mutual information using the SPSA optimization. In our proposed method, we have used four levels of decomposition and the image size gets halved in each level starting from finest to coarsest level. The values of transformation parameters, which provides a maximum mutual information between the reference and sensed image, is used as an initial transformation for the next finer level.

3.3 First-order SPSA

Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm was introduced by Spall [14] to find a root of the multivariate gradient equation using gradient approximation. It is popularly used to solve the optimization problem in remote sensing image registration [11]. In the first order SPSA, the update rule for the transformation parameters is given as

$$\theta_{n+1} = \theta_n + a_n g_n \tag{2}$$

where the gradient vector $g_n = [(g_n)_1 (g_n)_2 ......(g_n)_p]^T$ for the p-dimensional parameter space is estimated as

$$(g_n)_i = \frac{L(\theta_n + c_n\Delta_n) - L(\theta_n - c_n\Delta_n)}{2c_n(\Delta_n)_i} \quad \text{for } i=1,2\ldots p \tag{3}$$

where L is the objective function that has to be optimized.

According to Bernoulli's distribution, every element $(\Delta_n)_i$ of the vector (Δ_n) takes a value of +1 or -1. a_n and c_n are the positive sequences of the form

$$a_n = \frac{a}{(A+n+1)^\alpha} \quad \text{and} \quad c_n = \frac{c}{(n+1)^\gamma} \quad \text{where } 0<\gamma<\alpha<1 \tag{4}$$

where A, a and c are the constants of the optimization. As suggested in [10], these parameters values are A=100, c=0.5, a=12, α=0.602 and γ=0.101.

3.4 Second-order SPSA

The first order SPSA optimization converges slowly when it is very closer to the optimum solution. In [15], Spall presented a second order SPSA optimization technique to speed up the convergence. The update rule for the second order SPSA optimization algorithm is given as

$$\theta_{n+1} = \theta_n + a_{2n}\bar{H}_n^{-1}g_n\,, \quad \bar{H}_n = f_n(\hat{H}_n) \tag{5}$$

$$\hat{H}_n = \frac{n}{n+1}\hat{H}_{,n-1} + \frac{1}{n+1}H_n \tag{6}$$

where H_n is the per iteration estimate of the Hessian of L, and a_{2n} is a constant. Eq. (6) shows a recursive calculation of the per-iteration Hessian estimate $\hat{H}$. In Eq. (5), f_n is the transformation function for which $\hat{H}$ becomes invertible. Further details can be found in [15]. In our proposed method, the value of a_{2n} is set to 0.5.

3.5 Transformation Model

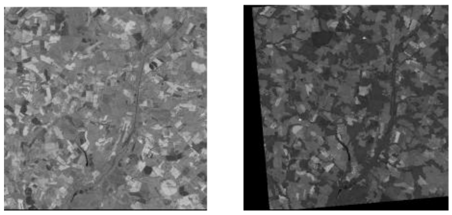
Let $f_r\,(u', v')$ and $f_s\,(u, v)$ are the reference and sensed images respectively. To register the images, a geometric transformation is estimated so that image $f_s\,(u, v)$ gets aligned perfectly with $f_r\,(u', v')$. So, the transformation gives the spatial relationship between the image pair. Here, affine transformation model is used as a geometric transformation function. The affine model is given as

$$\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \alpha_x \\ a_{21} & a_{22} & \alpha_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{7}$$

where ($a_{11}, a_{12}, a_{21}\ a_{22}$) combinedly represent the rotation, scale and shear differences and (α_x, α_y) are the translation parameters.

4 Simulation and Analysis

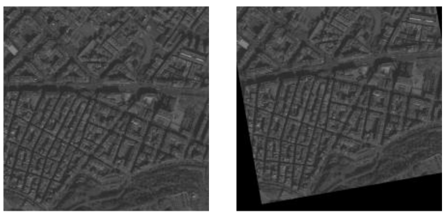
Three sets of optical images are selected for the experimental result analysis. The images are taken from USGS Earth Explorer (http://earthexplorer.usgs.gov). We have compared our proposed method with the other two remote sensing optical image registration algorithms named as IS-SIFT [4] and UR-SIFT [5]. All the experiments are performed on an Intel core i7-4770, 3.40 GHz CPU and 4 GB of physical memory computer using MATLAB 2014a.



(a) (b)

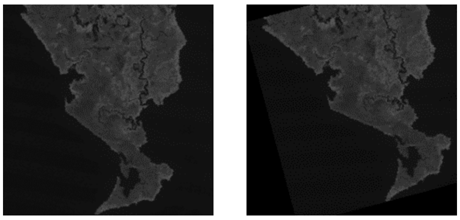
Fig. 1. Images of Baltimore, USA. (a) Reference image, (b) Sensed image

The first pair of images are obtained from LANDSAT Enhanced Thematic Mapper Plus (ETM+) sensor (resolution 30 m) on July 10, 2001 at the region of Baltimore. A section of 500x500 pixels in band 4 (0.76-0.90 μm) is taken as reference image and a segment of same size in the band 5 (1.55-1.75 μm) with 5 degree simulated rotation and translation of 10 pixels in X direction and 20 pixels in Y direction, is selected as sensed image. Fig 1(a) and Fig. 1(b) show the reference image and the sensed image respectively. The second pair of images is taken from the OrbView-3 on June 13, 2004 (resolution 1 m.) which cover the region of Barcelona in. A section of size 800x800 pixels taken is considered as the reference image and a segment with the same size with 10 degree simulated rotation and translation of -20 pixels in X direction and -35 pixels in Y direction, is selected as sensed image. Fig 2(a) and Fig. 2(b) show the reference image and the sensed image respectively. The third pair of images is taken from the Earth Observer-1 Advanced Land Imager (ALI) on May 4, 2013 over an area of Chesapeake Bay. A section of size 256x256 pixels taken is taken as the reference image and a segment with the same size with 15 degree simulated rotation and translation of 20 pixels in X direction and -10 pixels in Y direction, is selected as sensed image. Fig 3(a) and Fig. 3(b) show the reference image and the sensed image respectively.



(a) (b)

Fig. 2. Images at the region of Barcelona in Spain. (a) Reference image, (b) Sensed image.



(a) (b)

Fig. 3. Images at the region of Chesapeake Bay. (a) Reference image, (b) Sensed image.

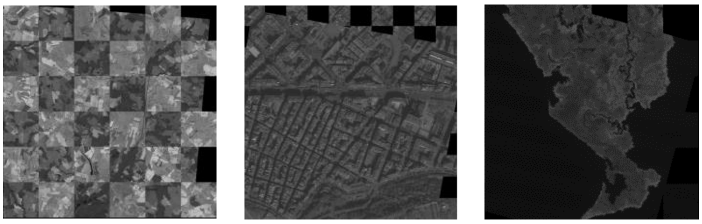
Table 1. Registration parameters (a_{11}, a_{12}, a_{21}, a_{22}, α_x, α_y) comparison between different Methods

Pair	Method	a_{11}	a_{12}	a_{21}	a_{22}	α_x	α_y
1	Ground Truth	0.9962	-0.0872	0.0872	0.9962	10.00	20.00
	IS-SIFT	0.9943	-0.0915	0.0917	0.9945	8.91	20.75
	UR-SIFT	0.9958	-0.0860	0.0862	0.9956	10.84	20.67
	Proposed Method	0.9959	-0.0878	0.0879	0.9960	9.43	20.29
2	Ground Truth	0.9848	-0.1736	0.1736	0.9848	-20.00	-35.00
	IS-SIFT	0.9844	-0.1740	0.1739	0.9841	-20.17	-34.43
	UR-SIFT	0.9846	-0.1738	0.1734	0.9850	-20.07	-34.91
	Proposed Method	0.9847	-0.1737	0.1735	0.9849	-20.03	-35.04
3	Ground Truth	0.9659	-0.2588	0.2588	0.9659	20.00	-10.00
	IS-SIFT	0.9604	-0.2508	0.2511	0.9608	20.78	-10.87
	UR-SIFT	0.9626	-0.2524	0.2531	0.9629	20.56	-10.67
	Proposed Method	0.9637	-0.2542	0.2548	0.9641	20.42	-10.45

Table 2. RMSE, MI and the associated computational time comparison between different Methods

Pair	Method	RMSE	MI	Time (Sec)
1	Ground Truth	-	1.0587	-
	IS-SIFT	0.9230	1.0381	36
	UR-SIFT	0.6701	1.0403	65
	Proposed Method	0.3205	1.0432	32
2	Ground Truth	-	1.2000	-
	IS-SIFT	0.4324	1.1877	112
	UR-SIFT	0.0315	1.1984	121
	Proposed Method	0.0285	1.1987	105
3	Ground Truth	-	0.8200	-
	IS-SIFT	0.8190	0.7681	22
	UR-SIFT	0.6133	0.7941	32
	Proposed Method	0.4246	0.8072	17

Table 1 shows the transformation parameters comparison between IS-SIFT [4], UR-SIFT [5] and the proposed method for different image pairs. Table 2 presents mutual information (MI), root mean square error (RMSE) and associated computational time comparison between different methods. It can be clearly observed that the proposed method provides less RMSE value compared to the other methods and the registration parameters are very close to the ground truth values in case of our proposed algorithm. The results obtained by using the UR-SIFT and IS-SIFT algorithms are acceptable but, still accuracy is less than the proposed method. IS-SIFT method is less accurate than the other two algorithms because the number of matches are less. The d_{ratio} criterion of IS-SIFT eliminates a number of correct correspondences. Although the number of matches obtained in coarse registration scheme of our proposed method is less than the UR-SIFT, still accuracy is comparatively better in our method. The transformation obtained from the coarse registration is further refined by the fine registration method. Therefore, the proposed method provides high accuracy in registration. Moreover, the proposed method takes less computational time compared to IS-SIFT and UR-SIFT methods. In IS-SIFT, significant computational time is required for the segmentation the input images. The uniform distribution of features and cross-matching algorithm of UR-SIFT need more computational time than the proposed method. The registered images obtained by using the proposed method are shown in Fig. 4. From the visual representation, it is clear that the edges and the regions of the registered images are perfectly aligned.



(a) (b) (c)

Fig. 4. Checkerboard mosaiced image of (a) pair 1, (b) pair2 and (c) pair 3

5 Conclusion

In this paper, we have proposed a coarse-to-fine registration method to register remote sensing optical images. The coarse registration is performed by using the standard SIFT approach with RANSAC-based outlier removal technique. The matched features obtained by using SIFT approach are refined through RANSAC algorithm. So, the result of coarse registration is very close to the optimum solution. The transformation parameter values obtained in the coarse registration scheme are used as an initial solution for the SPSA optimizer in the fine registration scheme. The fine registration is performed by the maximization of mutual information using the SPSA optimization in a multi-resolution framework. The simulation result shows that the proposed method provides comparatively higher accuracy in registration of remote sensing optical images compared to the other existing algorithms.

References

[1] Zitova, B., Flusser, J.: Image Registration Methods: A Survey. Image Vis. Comput. 21, 977–1000 (2003)

[2] Brown, L. G.: A survey of image registration techniques. ACM Comput. Surv. 24(4), 325–376 (1992)

[3] Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International journal of computer vision 60(2), 91-110 (2004)

[4] Goncalves, H., Corte-Real, L., Goncalves J. A.: Automatic image registration through image segmentation and SIFT. IEEE Trans. Geosci. Remote Sens. 49(7), 2589-260 (2011)

[5] Sedaghat, A., Mokhtarzade, M., Ebadi, H.: Uniform robust scale-invariant feature matching for optical remote sensing images. IEEE Trans. Geosci. Remote Sens. 49(11), 4516-4527 (2011)

[6] Gong, M., Zhao, S., Jiao, L., Tian, D, Wang: A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. IEEE Trans. Geosci. Remote Sens. 52(7), 4328-4338 (2014)

[7] Zhang, Y., Zhou, P., Ren, Y., Zou, Z.: GPU-accelerated large-size VHR images registration via coarse-to-fine matching. Comput. Geosci. 66, 54-65 (2014)

[8] Wu Y., Ma, W., Gong, M., Su, L., Su, Jiao, L : A novel point matching algorithm based on fast sample consensus for image registration. . IEEE Trans. Geosci. Remote Sens. Lett. 12(1), 43-47 (2015)

[9] Sedaghat, A., Ebadi, H: Remote sensing image matching based on adaptive binning SIFT descriptor. IEEE Trans. Geosci. Remote Sens. 53(10), 5283-5293 (2015)

[10] Cole-Rhodes, A. A., Johnson, K. L., LeMoigne, J., Zavorin, I.:Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. IEEE Trans. Image Proces. 12(12), 1495-1511 (2003).

[11] Cole-Rhodes, A. A., Johnson, K. L., LeMoigne, J.: Image registration using a 2nd order stochastic optimization of mutual information. Proc. IGARS. 6, 4038-4040 (2003)

[12] Suri, S., Reinartz, P.: Mutual-information-based registration of TerraSAR-X and Ikonos imagery in urban areas. IEEE Trans. Geosci. Remote Sens. 48(2), 939–949 (2010)

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

[14] Spall, J. C.: Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. IEEE Trans. Automat. Contr. 37, 332-341 (1992)

[15] Spall, J. C.: Accelarated second-order stochastic optimization using only function measurements. Proc. DAC. 1417-1424 (1997)