Registration of Remote Sensing Images with Steerable Pyramid Transform and Robust SIFT Features

Xiangzeng Liu1*, Zheng Tian1, 2, Haixia Xu2, Chengcai Leng1, Feiyue He1
1 Department of Applied Mathematics, Northwestern Polytechnical University, Xi’an 710129, China;
Email: dongfangyingdiyi@163.com
2 School of Computer Science, Northwestern Polytechnical University, Xi’an 710129, China;
3 State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing Applications Chinese Academy of
Sciences, Beijing, 100101, China

Abstract—This paper proposes an automatic registration technique based on steerable pyramid transform and robust shift invariant feature transform (SIFT) features, which can deal with the large variations of scale, rotation and illumination of the images. First, the steerable pyramid transform is used to the two input images and the sub-band images along certain orientations are obtained at each layer. Then, a robust SIFT algorithm is developed to calculate the initial transformation parameters for registration of these obtained images. Finally, the final refined parameters determined by using gradual optimization are adopted to achieve the registration result. The effectiveness of the proposed method is demonstrated by the experimental results.

Index Terms—Image registration, Steerable pyramids transform, robust SIFT, Remote sensing image

I. INTRODUCTION

Image registration is a process of determining the point-to-point correspondence between two images of the same scene, which are acquired by different sensors or by the same sensor at different times or with different parameters [1]. Various orthogonal wavelets were tested to register affine transformed images. Several registration algorithms that combined wavelet-based pyramid with other similarity measures were proposed in references [2,3]. However, orthogonal wavelet transforms are lack of translation and rotation invariance. Freeman et al. [4] proposed a steerable pyramid transform, which is multi-scale, multi-orientation, and self-inverting image decomposition, and it has the advantage that the sub-bands are both translation and rotation invariant [5]. These characteristics make it useful for registering remote sensing images, and therefore the registration with steerable pyramid transform will be more stable and robust under large rotations and image noise [6].

The crucial point of feature-based methods is to adopt discriminative and robust feature descriptors that are invariant to the assumed differences between the two input images, and therefore the extraction of invariant features is very important to registration results. Lowe [7,8] presented the SIFT method to extract distinctive invariant features from images. These features are invariant to image scale and rotation, and provide robust matching across a substantial range of affine distortion, addition of noise, and changes in illumination. Mikolajczyk et al. [9] compared the performance of several descriptors for local interest regions, and concluded that the performance of the SIFT-based descriptors was best. Therefore, the SIFT method has been successfully applied in remote sensing image registration [10,11]. However, the SIFT algorithm degrades under the large scale variances, rotations and changes in viewpoint.

To achieve robust registration of remote sensing images, we propose an automatic registration method based on steerable pyramid transform and robust SIFT features. The main contribution of this paper can be divided into two aspects. First, the images along certain orientations are used for initial matching at each layer, which are obtained by applying the steerable pyramid transform to the two input images. Second, a robust SIFT algorithm is developed, in which the improved main orientation of each keypoint is adopted and a new measure is used to determine the correspondence between keypoints. The new measure not only considers the distance between keypoints descriptors but also considers the scales, orientations and the cross correlation of the neighborhood of keypoints. It enhances the robustness of the proposed algorithm to large changes in viewpoint.

The rest of this paper is organized as follows. Section II gives an overview of the steerable pyramid transform and SIFT algorithm. In Section III, we describe our algorithm for remote sensing images registration. Experimental results and conclusions are given in Section IV and V respectively.

II. STEERABLE PYRAMID TRANSFORM AND SIFT ALGORITHM

A. Steerable Pyramid Transform

The steerable pyramid transform introduced by Freeman et al. [4] is a linear multi-scale, multi-orientation image decomposition that provides a useful front-end for image-processing and computer vision applications. The “steerable filter” refers to a class of filters, in which a filter of arbitrary orientation can be synthesized as a linear combination of a set of “basis
filters”. For any function \( f(x, y) \), \( f^\theta(x, y) \) is \( f(x, y) \) rotated through an angle \( \theta \) about the origin. We call \( f(x, y) \) is steerable if it satisfies the following equation:

\[
 f^\theta(x, y) = \sum_{j=1}^{M} k_j(\theta) f^\theta_j(x, y),
\]

where \( k_j(\theta) \) are the interpolation functions \( j=1, \ldots, M \).

More details about steerable pyramid can be found in reference [4].

In our algorithm, four orientation band-pass components and two-level decomposition are adopted in the steerable pyramid algorithm. The interpolation functions \( k_j(\theta) \) can be expressed as follows based on Theorem 1 in [4]:

\[
k_j(\theta) = \frac{1}{4} [2 \cos(\theta - \theta_j) + 2 \cos(3(\theta - \theta_j))],
\]

where \( \theta_j = j \pi / 4 \), \( j=1, 2, 3, 4 \). According to (1), an image of arbitrary orientation can be synthesized as a linear combination of the four orientation band-pass components at each layer. Therefore, in a fixed orientation, the structural information of the reference and sensed images can be used for image registration.

B. SIFT Algorithm

SIFT algorithm was proposed in [8] as a method to extract and describe key-points, which are robust to scale, rotation and change in illumination. There are five steps to implement the SIFT algorithm:

1) Scale-space extrema detection: The first stage searches over scale space using a Difference of Gaussian (DoG) function to identify potential interest points that are invariant to scale and orientation.

2) Keypoint localization: The location and scale of each candidate point are determined and the keypoints are selected based on measures of stability.

3) Orientation assignment: One or more orientations are assigned to each keypoint location based on local image gradient directions.

4) Keypoint descriptor: A feature descriptor is created by computing the gradient magnitude and orientation at each image sample point in a region around the keypoint location. These samples are then accumulated into orientation histograms summarizing the contents over location. These samples are then accumulated into each image sample point in a region around the keypoint by computing the gradient magnitude and orientation at each layer. Therefore, in a fixed orientation, the structural information of the reference and sensed images can be used for image registration.

5) The correspondence of feature points can be determined by taking the ratio of distance for the descriptor vector from the closest neighbor to the distance of the second closest.

For more details about the SIFT algorithm, readers can refer to [8]. In our algorithm, the step 3) and 5) of the SIFT algorithm are modified, which will be used to extract features from the obtained images at each layer of the decomposition with the steerable pyramid.

III. ROBUST SIFT MATCHING ALGORITHM IN STEERABLE PYRAMID

In this section, first, a robust SIFT algorithm (RSIFT) is developed to extract and match image features. Further, the robust SIFT matching algorithm in steerable pyramid (S-RSIFT) is proposed. To achieve robust registration of remote sensing images, we adopt more robust feature descriptor and similarity measure than the original SIFT algorithm. The robust SIFT matching algorithm (RSIFT) can be obtained by modifying the SIFT algorithm:

In step 3), we propose to use the Prewitt operators on each keypoint. The computation of the gradient and the orientation computation are:

\[
P_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix},
\]

\[
P_y = \begin{bmatrix} -1 & 1 \end{bmatrix},
\]

\[
dx(x, y) = P_x \ast L(x, y),
\]

\[
dy(x, y) = P_y \ast L(x, y),
\]

\[
m(x, y) = \sqrt{(dx(x, y))^2 + (dy(x, y))^2},
\]

\[
\theta(x, y) = \tan^{-1}(dy(x, y)/dx(x, y)),
\]

\[
\theta'(x, y) = \begin{cases} \theta(x, y), & \theta(x, y) \in [0, 180] \\ 360 - \theta(x, y), & \theta(x, y) \in (180, 360) \end{cases},
\]

where \( \ast \) is convolution operation, the modified descriptor not only retains more information but also has the effect of smoothness.

In step 5), a new similarity measure is proposed. Firstly, we can get the rough correspondence between keypoints by using the step 5) of the SIFT algorithm. Assume the two rough correspondence point sets are:

\[
B = \{b_1, b_2, \ldots, b_n\}
\]

and

\[
C = \{c_1, c_2, \ldots, c_n\},
\]

where each keypoint \( b_i = (x_i, y_i) \) and its corresponding point \( c_j = (x_j, y_j), \theta_i, s_i, \theta_j, s_j \) denotes the scale and orientation of \( b_i \) and \( c_j \), \( s_{ij} = s_i / s_j \) and \( \theta_{ij} = \theta_i - \theta_j \) expresses the scale ratio and the orientation difference between them.

Then, the scale ratio histogram and the orientation difference histogram are formed, and the highest peaks in the two histograms correspond to the estimated value of scale ratio \( s \) and orientation difference \( \theta \), respectively. Assume that \( W_1 \) and \( W_2 \) are two correlation windows of size \((2w+1) \times (2w+1)\) centered on \( b_i \) and \( c_j \), which can be represented as \( A_i \) and \( B_j \):

\[
A_i = I(x_i + su \cos \theta + sv \sin \theta, y_i + sv \cos \theta - su \sin \theta)
\]

\[
B_j = J(x_j + u, y_j + v)
\]

where \( u, v \in [-w, w] \), \( A_i \) is the element of \( W_1 \) centered on \( b_i \) and \( B_j \) is the element of \( W_2 \) centered on \( c_j \).

Finally, the new similarity measure between keypoints \( I_i \) and \( J_j \) can be represented as follows:

\[
D(i, j) = \frac{C_u + 1}{2} \left( \frac{1}{1 + \Delta_u(i, j)} \right) \left( (1 + \Delta_u(i, j)) \cdot r(i, j) \right)^{-1}
\]

where \( \Delta_u(i, j) = d(i, j) / d(1, 1) \) and \( r(i, j) = \cos \theta - \cos \theta' \).

215
where \( C_{i,j} \) is the normalized cross correlation (NCC), which is used in [12]. \( \Delta_i(i,j) = |s - s_{ij} | \) is the error of scale ratio and \( \Delta_i(i,j) = |\theta - \theta_{ij} | \) is the error of orientation difference between keypoint \( I_i \) and \( J_j \), \( r(i,j) \) is the Euclidean distance between the descriptor vector for \( I_i \) and the descriptor vector for \( J_j \). The best candidate match for each keypoint is found by maximizing \( D(i,j) \) between the keypoint \( I_i \) and \( J_j \).

Further, the implementation details of the robust SIFT matching algorithm in steerable pyramid (S-RSIFT) can be described as follows:

1) Keypoints are extracted from the two input images by DoG operator, and their orientations are determined by step 3) in RSIFT algorithm. Then, the orientation histogram is formed from the gradient orientations, and the main orientation of the two input images are determined by the peak in the histogram, respectively.
2) Apply the steerable pyramid transform to the two input images to level two with four orientation (0, \( \pi/4 \), \( \pi/2 \), \( 3\pi/4 \)) band-pass components respectively. At each layer, synthesize the four different orientation band-pass sub-images to a fixed orientation image, whose orientation is the main orientation of the input image obtained from step 1), respectively.
3) We assume the transform between the two images is affine transform. Three sets of initial transformation parameters can be derived using RSIFT algorithm to the synthetic images at each level and the non-oriented band-pass sub-images (NOBSI) of level 1, which are filtered on the two input images respectively.
4) The root mean-square error (RMSE) is used for evaluating the matching result as follows:

\[
RMSE = \sqrt{\frac{\sum_{m=1}^{m}(ax_i + by_i + c - X_j)^2 + (dx_i + ey_i + f - Y_j)^2}{m}} \tag{8}
\]

where \( m \) means the total number of matching points. The refined transformation parameters can be chosen by minimizing the RMSE. The finally refined transformation parameters are used to achieve the registration of the reference and sensed images.

The proposed algorithm is a coarse-to-fine procedure, and not only the stability and multi-orientation of the steerable pyramid filters are used, but also the robust of the RSIFT algorithm is employed. The effectiveness of the proposed algorithm will be demonstrated by the experimental results in next section.

IV. EXPERIMENTS AND RESULTS

In this section, we apply our algorithm to two sets of images. The first set is the satellite image of Beijing Bird’s Nest and its affine transformed image (Fig. 1). This set is used to show the implementation and the accuracy of our algorithm. The second set is composed of two Land-sat TM images, which are used to show the registration results of images from different sensors.

The experiments consist of two parts. In the first part, we show how the proposed algorithm is applied to the satellite image of Beijing Bird’s Nest and its affine transformed image in Fig. 1. In this experiment, the real transformation parameters (RTP) are shown in Table I. The main orientation of (a) and (b) is 0.15 and 0.25 radian, which is calculated from the first step of the proposed algorithm, respectively. (d) shows the matching results of the non-oriented band-pass sub-images (NOBSI) of level 1, (e) and (f) shows the matching results of the synthetic images at level 1 and level 2, respectively. From (d), (e) and (f), we can see different match points can be extracted from different layers, which make full use of the structural information of the two input images and enhance the robustness of the proposed algorithm to large rotations. The optimization process of transformation parameters are shown in Table I. TPL1 and TPL2 denotes the transformation parameters of level 1 and level 2, respectively. RFTP denotes the refined transformation parameters based on TPL1 and TPL2, which are chosen by minimizing the RMSE, where the matching points used to calculate the RMSE are from (e). TPLN1 denotes transformation parameters of NOBSI of level 1. FRFTP denotes the refined transformation parameters based on TPN1 and TPL2, where the matching points used to calculate the RMSE are from (d). From Table I, we can see the final transformation parameters (FRFTP) is the optimal and our algorithm yields sub-pixel accuracy.

In the second part, we apply the proposed algorithm to two Land-sat TM images. The registration results of two Land-sat TM images (Land-sat TM band 0 and band 8) are shown in Fig. 2. The transformation parameters are \( T = [0.966, 0.259, 249.99, -0.259, 0.966, 7.06] \). To compare the registration accuracy, we consider the root mean-square error (RMSE) for each set.
mean square error of intensity (RMSEI) and the correlation (corr) between the overlapping areas of registered image pairs, which are used in [13].

The registration accuracy in terms of the RMSEI and corr for Fig. 2 are shown in Table II. The results show that the proposed algorithm outperforms the other two algorithms.

V. CONCLUSION

In this paper, we present a remote sensing image registration method with RSIFT in steerable pyramid. The method takes advantage of the orientations information with steerable transform and the robust of the RSIFT algorithm. The experimental results show that the proposed algorithm returns better performance as compared to SIFT+SVD methods and SIFT methods. However, the accuracy of the proposed algorithm is still affected by the large viewpoint difference between images in some degree. Therefore, it is necessary to construct a more robust descriptor for image registration.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (No.60375003), the Aeronautics and Astronautics Basal Science Foundation of China (No.03153059) and the Science and Technology Innovation Foundation of Northwestern Polytechnical University (No.2007KJ01033).

REFERENCES


