

SURF Based Matching for SAR Image Registration

Ujwal Kumar Durgam

Department of Electronics and
Communication Engineering,
National Institute of Technology,
Rourkela, India.
Email: ujwalkumar06@gmail.com

Sourabh Paul

Department of Electronics and
Communication Engineering,
National Institute of Technology,
Rourkela, India.
Email: sourabhpaul26@gmail.com

Umesh C. Pati

Department of Electronics and
Communication Engineering,
National Institute of Technology,
Rourkela, India.
Email: ucpati@nitrrkl.ac.in

Abstract—Sub-pixel accuracy in registration of synthetic aperture radar (SAR) images is still a challenging task in remote sensing applications. Speeded Up Robust Feature (SURF) is one of the most popularly used method for feature detection and description of SAR images. But using SURF alone in registration cannot give accurate matching in corresponding features, as it contains many wrong correspondences called outliers. RANSAC (Random Sample Consensus) an outlier removal technique is used to remove those outliers. Even then some outliers still exist which degrade the registration quality. In this paper, a novel algorithm is proposed to remove those remaining outliers by limiting the RMSE to less than 0.5 in registration process. Firstly, SURF based feature matching is performed between image pairs to get the corresponding features, then RANSAC is used to remove most of the outliers obtained from SURF feature matching. Then, the proposed method is applied to still refine the matched features obtained after RANSAC.

Keywords—*Speeded Up Robust feature; Affine transformation; Random sample consensus.*

I. INTRODUCTION

Synthetic Aperture Radar (SAR) image registration has a wide range of remote sensing applications like change detection, image fusion, urban growth detection, damage identification, border monitoring, traffic studies etc. All these applications require an accurate alignment of images to sub-pixel accuracy. SAR images obtained in a particular region through different sensors at different times are used for registration. The process of registration is classified into two types: feature based methods and intensity based methods [1-2], [11]. As the SAR images contain speckles, intensity based registration is computationally complex and likely to produce errors. Due to the distinctive features in SAR image, feature based methods are always preferred over intensity based methods.

There are many image registration algorithms proposed, of which most of them are based on SIFT (Scale-invariant feature transform). SIFT extracts distinctive features in images that are invariant to image scale and rotation. These features can be used for reliable matching of images, robust to noise, change in illumination, and 3D camera viewpoint. The descriptors accurately extract scale and rotation invariant characteristics around key-points. When SAR-based methods are used for

registration, they usually suffer from unavoidable speckles. Schwind et.al [3] proposed SIFT-Octave (SIFT-OCT) which reduces the effect of speckles by skipping the first octave of the scale space pyramid. Shunhu Wang [4] proposed Bilateral filter SIFT (BF-SIFT), where an Anisotropic scale space is constructed using bilateral filter, compared to the linear scale space built by Gaussian kernel, the anisotropic scale space built by bilateral filter preserves more details at coarser scales. So, more features which are located precisely can be detected. Bangsong Wang [5] presented Uniform SIFT which extracts enough robust and reliable uniformly distributed features by the strategies of optimal feature selection based on voronoi diagram and feature scale space proportional extractions. Flora Dellinger [6] proposed SIFT algorithm by using a new gradient definition yielding an orientation and a magnitude that are robust to speckle noise. Ruihua Liu, Yanguang Wang [7] implemented an R-SURF method to register SAR images.

In this paper, SURF (Speeded up robust feature) based registration is performed for better performance and accuracy. The use of integral images reduces time complexity. Initially, corresponding features are matched from image pairs using standard SURF. Then the outliers are removed by RANSAC method. Finally, proposed algorithm is introduced to refine the matched features. These final matched points are used to calculate the affine transformation parameters. Then sensed image is transformed in accordance to the transformation parameters.

This paper is organized as follows: Section II provides the details of SURF feature detector and descriptor. Section III represents the proposed outlier removal algorithm to obtain the final set of corresponding points between the reference and sensed images. Section IV presents simulation results of the proposed method. Finally, Section V offers conclusion.

II. SPEEDED UP ROBUST FEATURE

SURF [8] is a scale and rotation invariant interest point detector and descriptor of images inspired from SIFT. SURF takes the advantage of integral images to compute convolution, and moreover only 64 dimensions of descriptor results in less computational and matching time. Steps for SURF feature matching are given as follows:

A. Interest points detection

Each point $X = (x, y)$ of an image I , is passed through Laplacian of Gaussian filter of various sizes with corresponding standard deviation σ in different directions. The responses of the filter are denoted by $L_{xx}(X, \sigma)$, $L_{yy}(X, \sigma)$ and $L_{xy}(X, \sigma)$ in x , y , and xy direction respectively and are normalized by the size of the filter. This process takes the advantage of integral images and the computational time is thus independent of the size of the filter. The hessian matrix at a point X at a scale σ is defined as

$$H = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix} \quad (1)$$

In practical, as proposed by H. Bay in [8], the Gaussian is discretized by approximation, and the corresponding convolution of Laplacian of the approximated Gaussian with image is denoted by D_{xx} , D_{xy} and D_{yy} respectively. The determinant of the hessian matrix is approximated to

$$\text{Det}(H(X, \sigma)) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (2)$$

B. Suppression of non-maximal points

Interest points are detected by suppressing the non-maximal points. A filter of size 9×9 is considered as initial scale layer, with scale $s=1.2$ (Laplacian of Gaussian with $\sigma=1.2$), the image is repeatedly convolved with increasing filter size to 15×15 , 21×21 , 27×27 ..., and their corresponding filter responses are obtained composing an octave. The process is repeated for larger scales, the step between the consecutive filter sizes should also scale accordingly. Hence, for each new octave, the difference in the consecutive filter size in an octave is doubled (going from 6 in 1st octave, 12 in 2nd, 24 in 3rd and so on). Hence, in 2nd octave, 15×15 , 27×27 , 39×39 filters are considered. These octaves built a scale space of $\text{Det}(H)$ image for localizing the interest points. In each octave, the interest point is localized by applying non-maximal suppression in $3 \times 3 \times 3$ neighborhood; a point 'X' is selected as an interest point if its $\text{Det}(H(X, \sigma))$ is greater of its neighboring 8 pixels in its scale and also greater of its 3×3 neighbors in its neighboring scales. i.e., for a point in scale 15, its $\text{Det}(H(X, \sigma))$ has to be greatest of all its 8 neighbors in scale 15 as well as greatest of the 3×3 pixels in scale 9 and 21 for the 1st octave. The similar procedure is followed to obtain interest points in higher octaves resulting in feature points of an image in different octaves.

C. Feature direction

A circular region of radius $6*s$ (where s is the scale of the image in which the feature is detected) around the feature point is constructed. On each point lying within this circular region, Haar wavelet responses of size $4*s$ in X and Y directions are obtained. Working on integral images, this takes less computational time. The dominant orientation is estimated by summing the horizontal and vertical wavelet responses within a rotating wedge, covering an angle of 60 degrees in the wavelet response space. The resulting maximum is then chosen to describe the orientation of the interest point descriptor.

D. Keypoint Descriptor

For constructing the descriptor, the first step consists of constructing a square region of size $20s \times 20s$ centered at the feature point orientated in the obtained feature direction. This region is further divided into 4×4 square sub regions, in each square sub region, the sum of the Haar wavelet responses of size $2*s$ are obtained in X and Y direction, denoted by Σdx and Σdy respectively. Hence, each sub region has a four dimensional descriptor vector v , where $v = (\Sigma dx, \Sigma dy, |\Sigma dx|, |\Sigma dy|)$. This results in a vector of length $4*16$ i.e., 64 dimensional descriptor. The feature points are obtained over test and reference images by using the above described SURF, obtaining 64 dimensional descriptor for each feature point on both images.

E. Feature Matching

The feature points detected on the reference image have to be matched with the feature points on the test image. The process of matching is made possible by the descriptor available for each feature point, the feature points with minimum descriptor distance are considered to be matched. Let $dr(i)$ and $dt(j)$ denote the descriptors or feature vectors of the i^{th} and j^{th} key point in reference and test images respectively. For a feature i with descriptor $dr(i)$, the best match $\hat{j}(i)$ is given as

$$\hat{j}(i) = \arg \min_j \|dr(i) - dt(j)\| \quad (3)$$

$$\|dr(i) - dt(\hat{j}(i))\| < T \|dr(i) - dt(j)\| \quad \forall j \neq \hat{j}(i) \quad (4)$$

As suggested by Lowe [10], $\hat{j}(i)$ from Eq. 3 is a valid match point if it satisfies Eq. 4. Where T is defined as the ratio of the distance between the 1^{st} neighbor to the 2^{nd} neighbor. The threshold value of T suggested by Lowe is 0.8

III. PROPOSED OUTLIER REMOVAL ALGORITHM

Feature matching operation between the reference and sensed images is performed by utilizing the values of T . Lower value of T provides higher correct rate but reduces the number of matched points whereas higher values of T , increases the number of matched features with the expense of correct rate. Matched features obtained by using T values contain many outliers. So, RANSAC is used to remove most of them but still some outliers exist which degrade the registration quality. In this paper, we have proposed a method which removes the remaining outliers one by one in an iterative process.

The matched features obtained using RANSAC are refined by eliminating a matched pair in each iteration and calculating the RMSE. If an outlier exists in a set of matched features, then the residue of the outlier exists in RMSE calculation, then RMSE will be more. In other words, if the contribution of outlier exists, then the RMSE will be more. Hence, every time leaving a pair of points, RMSE is calculated for rest of the points. The point on whose removal the RMSE is

least is the outlier of that iteration. So, that pair is completely discarded and the remaining set of matched features is used in the next iteration. This process continues until RMSE value is less than our limited set point. Expecting sub-pixel accuracy of registration the set point is set to 0.5. The set of matched features for which our set point is reached is the final set of corresponding points used in affine transformation to calculate the transformation parameters.

A. Transformation Model

(x', y') And (x, y) are the co-ordinates of reference and sensed image respectively. The transformation parameters are obtained from affine transformation matrix given as

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_x \\ a_{21} & a_{22} & a_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (5)$$

Where $(a_{11}, a_{12}, a_{21}, a_{22})$ together represent scale, rotation, shearing, and (a_x, a_y) are the translation parameters. These transformation parameters are used for transforming the sensed image to get aligned with the reference image.

IV. SIMULATION AND ANALYSIS

The simulation of the proposed algorithm is done with the help of Matlab 2014a. The first set of images shown in Fig.1 consists of Fig. 1(a) which is a segment of 800×800 pixels obtained from JERS-1 sensor with 30 meter resolution viewing the wetland regions of circumpolar North America acquired in the summer of 1998, has been taken as the reference image. And Fig. 1(b) which is a segment of 800×800 pixels of the same area taken in winter of 1997/1998 with a rotation of 20 degrees has been selected as the sensed images. Fig. 1(c) and (d) are SURF features shown on reference and sensed images respectively. Fig. 1(e) shows the checkerboard representation of the registered images.

In this experiment, the matched features obtained after refining with RANSAC provides RMSE of 0.8502 for $T=0.7$. On using our proposed method 0.4994 of RMSE is obtained, thus registering the images with sub pixel accuracy.

The next set of images shown in Fig. 2 consists of Fig. 2(a) which is a section of 800×800 pixels obtained from JERS-1 of 30 meter resolution covering the area of Amazon river Basin, from the Atlantic to the Pacific, acquired during the generally low flood time of the Amazon River in September - December 1995, has been taken as the reference image and Fig. 2(b) which is a section of 800×800 of the same area in 1996 with a rotation of 20 degree and translation of -325 and 120 pixels in X and Y direction respectively, has been taken as the sensed image. Fig. 2(c) and (d) are SURF features shown on reference and sensed images respectively. Fig. 2(e) shows the checkerboard representation of registered images.

In this experiment, the matched features obtained after refining with RANSAC provides RMSE value of 0.7084 for $T=0.6$. lower values of T is considered here as the number of matched points in this case are more compared to the previous

experiment, thereby considering higher correct rate points. On using our proposed method RMSE of 0.4826 is obtained, thus registering the images with sub pixel accuracy.

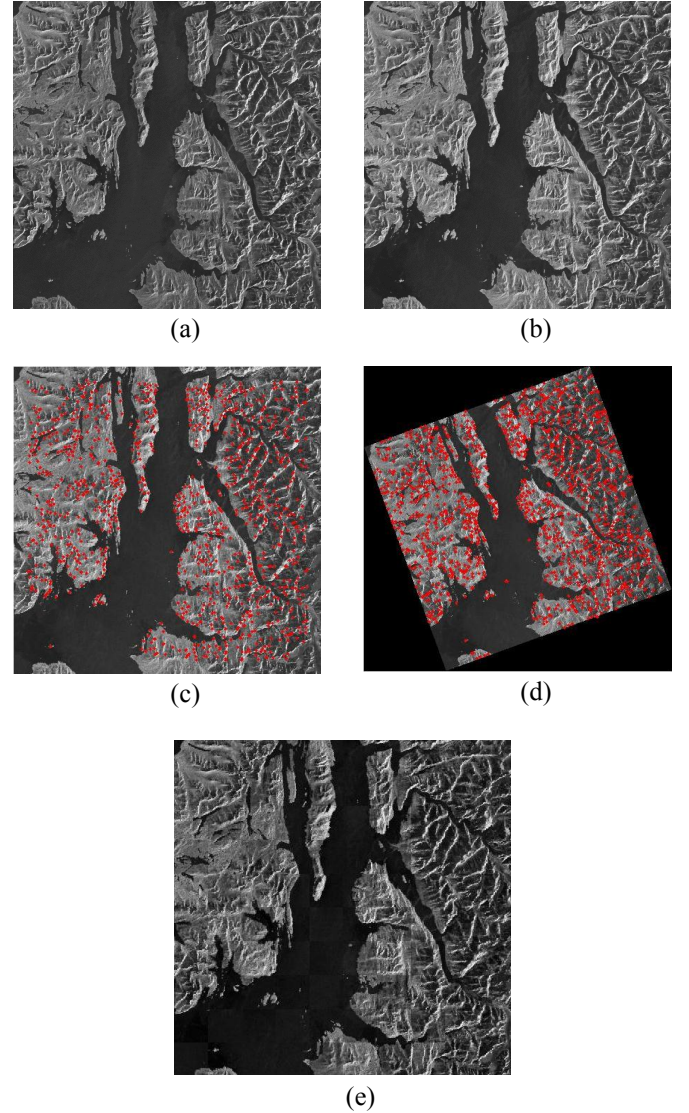


Fig. 1. JERS-1 Images of wetland regions of circumpolar North America in summer and winter. (a) Reference image, (b) Sensed image, (c) Feature points on reference image, (d) Feature points on sensed image, (e) Checkerboard mosaicked image of (a) and (b).

The third pair of images shown in Fig. 3 are very high resolution (2 meter) TerraSAR-X images viewing a location in Barcelona captured on 1st April 2014. Fig. 3(a) which is a section of 500×500 , has been taken as the reference image and Fig. 3(b) which is obtained on simulated rotation of 20 degrees and with a translation of -175 and 115 in X and Y direction respectively of the reference, has been taken as the sensed image. Fig. 3(c) and (d) are SURF features shown on reference and sensed images respectively. Fig. 3(e) shows the checkerboard representation of registered images.

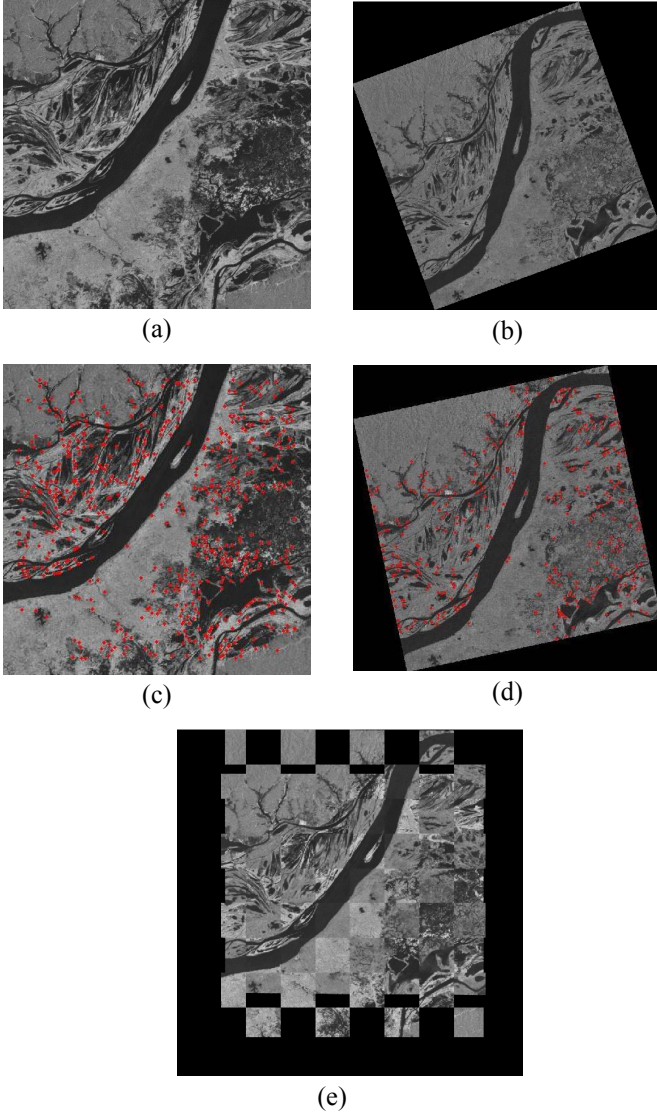


Fig. 2. JERS-1 Images of a location in Amazon river basin during low flood and high flood. (a) Reference image, (b) Sensed image, (c) Feature points on reference image, (d) Feature points on sensed image, (e) Checkerboard mosaicked image of (a) and (b).

In this experiment, the matched features obtained after refining with RANSAC provides RMSE value of 0.8318 for $T=0.7$. On using our proposed method 0.4984 of RMSE is obtained, thus registering the images with sub-pixel accuracy.

In feature-based registration method, the quality of image registration is measured using RMSE factor. RMSE is given by the Eqs. (6) - (8)

$$dx(i) = x'_i - a_{11}x_i - a_{12}y_i - a_x \quad (6)$$

$$dy(i) = y'_i - a_{21}x_i - a_{22}y_i - a_y \quad (7)$$

$$RMSE = \sqrt{\frac{1}{k} \sum_{i=1}^k dx^2 + dy^2} \quad (8)$$

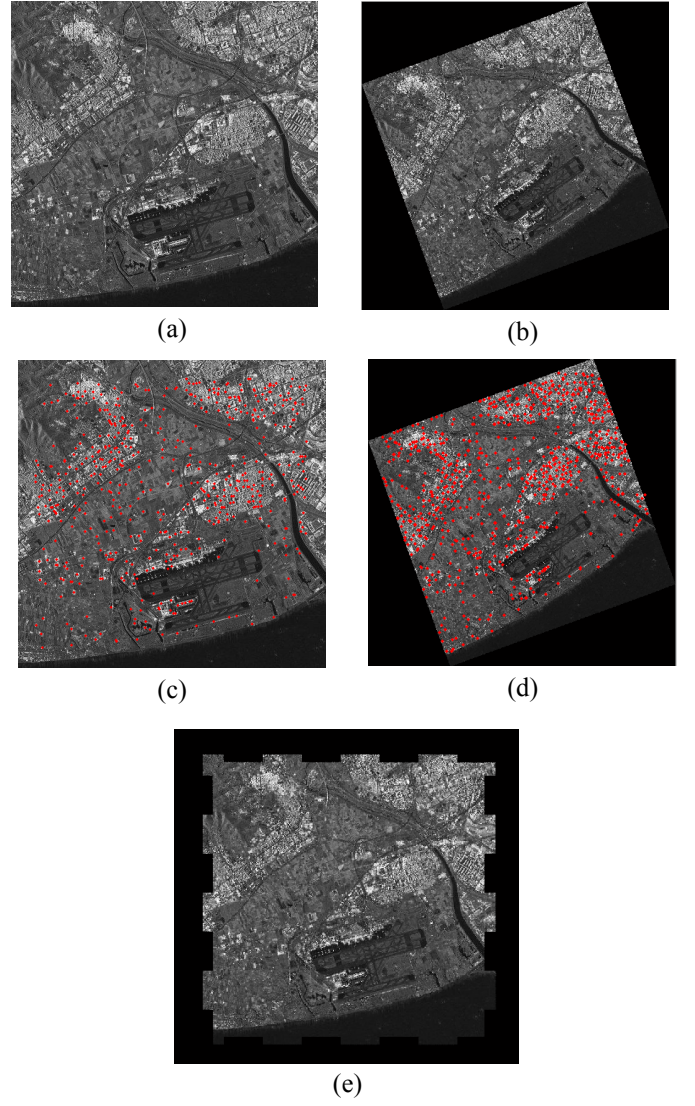


Fig. 3. Terra-SAR Images of a location in Barcelona. (a) Reference image, (b) Sensed image, (c) Feature points on reference image, (d) Feature points on sensed image, (e) Checkerboard mosaicked image of (a) and (b).

Here, k is the total number of matched key-points, dy and dx are the parameters defined by the x , y difference value between the matched points in the reference image and transformed points of the sensed image using affine transformation parameters. From Eq. 8 RMSE gives the root of the mean square deviation of the sensed points from the reference points. Another important parameter RMS_{LOO} (Root Mean Square Leave One Out) is also popularly used metric for quality measurement in registration. The details on this parameter are given in [9].

TABLE I shows the comparison of registration quality between the proposed method and the R-SURF method proposed by Ruihua Liu, Yanguang Wang in [7] which registers images using SURF and RANSAC. From TABLE I, it can be observed that proposed method achieve sub-pixel accuracy compared to other one.

V. CONCLUSION

In this paper, SURF is used for feature detection as well as feature matching to register the SAR images. But the matched features obtained from SURF contain outliers. RANSAC is used to eliminate those outliers. Finally, the proposed method is introduced to eliminate the outliers which still degrade the registration even after RANSAC outlier removal. Experimental

results show that the proposed method provides very high sub-pixel accuracy by limiting the value of RMSE less than 0.5. Images of different types and resolutions with transformation differences are used to analyze the performance of proposed algorithm. It provides better registration accuracy for two or more images having different illumination and transformation differences.

TABLE I. REGISTRATION PARAMETERS COMPARISON BETWEEN THE PROPOSED AND R-SURF METHOD

Image Pair	Threshold T	RMSE _{all} for R-SURF Method	RMSE _{LOO} for R-SURF Method	RMSE _{all} for Proposed Method	RMSE _{LOO} for Proposed Method
1	0.8	0.8782	0.8804	0.4977	0.5005
	0.7	0.8502	0.8526	0.4994	0.5024
	0.6	0.8340	0.8363	0.4973	0.5001
2	0.7	0.74438	0.8154	0.4910	0.4967
	0.6	0.7084	0.7285	0.4826	0.5011
3	0.8	0.8264	0.8296	0.4954	0.4980
	0.7	0.8318	0.8355	0.4984	0.5025

ACKNOWLEDGMENT

The authors would like to thank The Department of Science and Technology, Government of India for the financial support under FIST program (Grant No. SR/FST/ETI-020/2010) to setup Virtual and Intelligent Instrumentation Laboratory in the Department of Electronics and Communication Engineering, National Institute of Technology, Rourkela in which the research work has been carried out.

REFERENCES

- [1] B. Zitova, and J. Flusser, "Image registration methods: a survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.
- [2] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, pp. 325–376, Dec. 1992.
- [3] P. Schwind, S. Suri, P. Reinartz, and A. Siebert, "Applicability of the SIFT operator for geometrical SAR image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 31, no. 8, pp. 1959–1980, Mar. 2010.
- [4] S. Wang, H. You, and K. Fu, "BFSIFT: A Novel Method to Find Feature Matches for SAR Image Registration," *IEEE Trans. Lett. Geosci. Remote Sens.*, vol. 9, no. 4, pp. 649–653, Jul. 2012.
- [5] B. Wang, J. Zhang, L. Lu, G. Huang, and Z. Zhao, "A Uniform SIFT-Like Algorithm for SAR Image Registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 12, no. 7, pp. 1426–1430, Jul. 2015.
- [6] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SARSIFT: A SIFT-Like Algorithm for SAR Image," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 453–466, Jan. 2015.
- [7] R. Liu, Y. Wang, "SAR Image Matching base on Speeded Up Robust Feature," *IEEE Int. Conf. on Intelligent Systems*, vol. 4, pp. 518–522, 2009.
- [8] H. Bay, T. Tuytelaars, and L. V. Gool, "Surf: Speeded up robust features," *Computer vision*, pp. 404–417, 2006.
- [9] H. Goncalves, J. Goncalves, and L. Corte-Real, "Measures for an objective evaluation of the geometric correction process quality," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 292–296, 2009.
- [10] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [11] S. Paul, D. Ujwal, and U. C. Pati, "Landsat enhanced thematic mapper plus image registration using sift," *IEEE Int. Conf. on ICIIP*, Shimla, 2015, accepted.