Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 171 (2020) 2390–2399

Procedia Computer Science

www.elsevier.com/locate/procedia

Third International Conference on Computing and Network Communications (CoCoNet'19)

Remote Sensing Image Registration Methodology: Review and Discussion

Ms. Priyanka S. Tondewad[1], Ms. Manisha P. Dale[2]

[1] *Electronics & Telecommunication Department, AISSMS IOIT, Sangamvadi Pune - 411001, India*
[2] *Electronics & Telecommunication Department, AISSMS IOIT, Sangamvadi Pune - 411001, India*

Abstract

Image registration is the process of overlaying two or more multi sensor or multi temporal or multi resolution images of the same scene. Image registration is very crucial preprocessing step in any remote sensing image processing application. This paper presents a brief review of evolution of different image registration methodologies with challenges involved in it. Image registration basically aligns the two images (reference and sensed image) geometrically. In the first part of this paper, classical methods will be discussed which are mainly based on area based (Intensity based) approach and feature based approach. Ample amount of work has already been done in both methods in manual and automatic manner. In the second part all recent approaches will be discussed and compared with results, on the basis of registration accuracy, number of correct correspondence. In recent studies Convolutional Neural Network with deep learning algorithms are used in combination with traditional methods to boost the overall performance.

Keywords : Convolutional Neural Network (CNN), Remote Sensing (RS),Automatic Image Registration(AIR).

1. Introduction

RS image registration have very wide area of applications such as change detection for disaster management and prevention, landscape planning, image mosaicking, weather forecasting, agriculture monitoring and most importantly security and surveillance related military operations. Remote Sensing image registration is a method which finds control points from both input images to match with the same point on ground. In registration process one input image is called reference image and other as moving or sensed image. Basically registration method is to superimpose the pixels of reference image to the sensed image by aligning the image into common co-ordinate system. Misalignment between the two images may be due to different viewpoints, different sensor positions, object movement or

1877-0509 © 2020 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the scientific committee of the Third International Conference on Computing and Network Communications (CoCoNet'19).
10.1016/j.procs.2020.04.259

deformation [24]. Effective analysis of the observation area can be done, when the images are aligned accurately therefore image registration plays very vital role in input image pre-processing.

About two decades back the image basic registration methods started evolving. Primarily the RS image registration can be done by area- based (Intensity based), feature-based or hybrid methods. Area based methods are more concerned about use of image intensities, whereas feature based methods concentrates on feature extraction and further processing. While in case of feature based approach the control points are estimated for a perfect match between a reference and sensed image. The whole focus is on the spatial relations or various descriptors of features. Now a days Convolutional Neural Network algorithms are used for image registration purpose. First part of this paper will give brief idea about various versions of both area and feature based registration methods and second part will have all the recent trends.

2. Input Image Set

Before going to the actual registration process one should know about what are the different types of images that can be used in registration as per the application domain:

a. Multi Temporal Images: Two or more images of the same or different sensors captured on different instant of time.
Outcome of using these type of images is to analyze change detection.

b. Multi Sensor Images: Two or more images of different sensors can give better exposure to the analysis of region of interest. As different sensors gives different information as the various bands it can work on like infrared, near infrared, microwave and other bands. Each band carries various details in accordance with the our application we can register the Multispectral (MS) image with Panchromatic (PAN) or MS with Synthetic Aperture Radar (SAR) or MS with Hyperspectral (HS) to enhance the details of particular band.

c. Different Viewpoint: Two or more images captured by same type of sensors but from different viewpoints so as to make a 3D realization of some particular scene.

With these image data sets other important factors one need to take care of is noise present in the image, clouds coverage, accuracy of the sensors used, angle of precision according to the application.

3. Image Registration Methodology

Two decades back traditional image geo-referencing began with manual image registration process, where human need to locate the ground control points (GCP) on both sensed & reference image by visual observation like intersection of roads. After finding GCPs next steps of feature detection and matching done by human operators. Performing such operations on remotely sensed (RS) images is very time consuming and tedious job as RS images are of large size.[1]

Problems faced in manual image registration:

a. Finding and locating accurate control points.
b. More GCPs required to improve registration accuracy.
c. Expert human operators required.
d. Difficult to match timing constraints.

These problems lead to develop semi or fully automated registration methods [2]. An automatic image registration algorithms can be broadly classified as area based and feature based automatic image registration (AIR) methods.

3.1. Area Based Methods

Area based methods are also called as intensity based methods as it works directly on intensity values irrespective of

structural analysis. Area based methods use the specific sized blocks (rectangle or circular) of both images and compare them based on statistical values. These statistical values are used to measure similarity between two images. After some more development in area based methods the different similarity measures being used like correlation coefficient (CC), Sequential similarity detection [1] [27]. These methods are easy to implement. Few methods are implemented using Fourier Transform (FT & Fast FT)[1][28]. Fourier based methods are more robust against the frequency dependant noise and non-uniform time varying illuminations. The center point of block was considered as control point (or tie point) for comparison. This eventually led to less accurate results as intensity values are affected by many factors like time of acquiring the image, sensors used, illumination, different angle, sometimes shadow of the objects also affects the pixel value. Also, processing the image block wise is not suitable when the image is having lot of deformations. There after Mutual Information (MI) method came into to existence, this method is more suitable for multi sensor data sets. In [30] a combined strategy of MI and correlation is used as similarity measure for a multiresolution image set. Author used wavelet based multiresolution pyramid. Results shown that MI based methods are more suitable for subpixel registration [30].
Therefor few algorithms were failed to operate on different types of input images, misaligned images or different gray level images[1][2].
It's been observed that as the deformation level increases the computational time required.

3.2. Feature Based Methods

Feature based methods concentrates more on features than the intensity. There are ample amount of different sensors developed till now and every sensor has different set of characteristics therefore in spite of having numerous algorithms to work on multi sensor data sets there are few challenges in each stage. This section will brief about basic steps involved in feature based registration with challenges:

3.2.1 Feature Extraction

Feature extraction is identification of different region like closed boundary contours, corners, forest, road, river, urban area and other salient features using various segmentation and edge detection methods. These features are represented by the point called control points (CPs). Earlier, combined corner and edge detector based on local autocorrelation is used to find the features, this method is invented by Chris Harris and Mike Stephens in 1988[3]. Famously known by the name Harris corner detector, is mostly used for 3D image interpretation. Modified Harris corner detectors are used in RS image registration to reduce the computation time as it ignores the low gradient value pixels [31]. There after Correlation Coefficients are used for feature extraction

The Scale Invariant Feature Transform (SIFT) can extract local features with scale and rotation invariant in scale space these advantages of SIFT makes it the most popular preprocessing stage in many algorithms [4] [5] [15] [24] [26]. Only SIFT is not sufficient for RS images as compared to other image applications. SIFT gives uneven distribution of features [6] and concentrates on local features [23]. Therefore various variants of SIFT or modified SIFT algorithms are preferred for RS images. In Ref. [6], to beat the disadvantage of SIFT Ultra Robust SIFT (UR-SIFT) is developed. Similarly in Ref. [4] Uniform Robust SIFT is implemented to improve on uniform feature finding. Accurate and uniform feature detection is possible with this method. The Non-subsampled Contourlet Transform (NCST) is used to separate the high and low frequency components and then SIFT is used for better performance [5]. As SIFT is good at finding local features the missed out global features makes registration more discussed about finding both local & global features. In Ref. [23] Watanabe Fan et. al., implemented Uniform Nonlinear Diffusion based Harris feature extractor which proved to be robust to speckle noise in SAR images and nonlinear intensity variations in multimodal images. The other set of methods are comprises of Locally Linear Transform (LLT) [33], performs well specifically in the nonrigid and affine cases. Following are the most common challenges in feature detection

a. Features should be robust to noise and coveTPSrage of scene.
b. More number of features are required for better results.
c. Detected features should be distinctive.

3.2.2 Feature Matching

In this step correspondence between the extracted features from both input images are matched using various feature matching or similarity measure algorithms like Euclidian distance measure [4], gradient based algorithm [2], cross correlation based method, line-point invariant via dual matching [4] and many more methods are available. One of the biggest challenge at this stage is to increase the correct feature matching pairs. SIFT in combination with optimized Local Self Similarity measure Mutual Information (LSS_MI) is tested in [26].

3.2.3 Transformation

This is an important step to map the information from one image input to other to get accurately aligned images. The transformation technique should be able to select the method of transformation which will remove only the spatial distortions between images due to differences in acquiring the image [1]. General categories of transformation are rigid, affine, projective, perspective and global polynomial [1][8].

Rigid transformations is combination of rotation, translation and scaling operations. These are more suitable when the rigid body formation is important.

Affine transformations works in Cartesian coordinate systems, it is comprised of rotation, scaling and translation operations. Affine transformations are less prone to the distorted images. This transformation is linear transformation, while mapping the overall geometry of the points remains same. These are the globally used transforms in registration process.

The projective and perspective transformations are more suitable when user have the knowledge of distance between sensor and object. These transformations are used to remove the distortions occurred due to different projections of sensors at varying distances.

Generally polynomial transformations are useful to resolve low frequency distortions because of their unpredictable behavior when the degree of polynomial increases.

Bajcsy et al[9][29] introduced a different approach to the RS image registration is with some complex or local distortions is without using any mathematical mapping functions, where the estimation is of geometric deformation which is nothing but search for "best" suitable parameters. This approach is called elastic registration. In this method image is considered as rubber sheet, where external forces for stretching and internal forces for smoothing are applied to make them aligned to each other with minimum stretching and bending. Limitation of this method is that it is less efficient for localized deformation mapping.

3.2.3 Resampling and Outlier Removal

The mapping function formed in previous step is used to transform the sensed image to registered image. The transformation can be considered in two manner as forward or backward. When each pixel from sensed image is directly transformed using mapping functions then it's called as forward method. This method is difficult to implement therefore backward method is more preferred. It takes inverse of the estimated mapping function at particular pixel in registered image coordinates with respect to the same pixel of sensed and reference image. Image interpolation is performed in order to avoid holes or overlaps in output image.

The most frequently used interpolants are like bicubic, nearest neighbor, bilinear functions, quadratic splines, cubic B-spline, Gaussian, truncated sinc and variants of spline functions. Inglada, Jordi Muron, Vincent,Pichard, amien, Feuvrier, Thomas [10] worked on finding the effect of interpolants and similarity function on the artifacts introduced at subpixel level. It has been observed that despite of blurring effect of B-spline interpolators, it performs better for disparity map estimation. Also these artifacts in subpixels can be reduced by preprocessing, which smoothens the secondary interpolated image [6]. In Ref. [2], Thin Plate Spline(TPS) interpolation is tested on different sets of images, shown that TPS is beneficial when subpixel accuracy is a concern. Also TPS reduces the local deformations when provided CPs are uniformly distributed [23].

4. Comparison of Various Methodologies

In Ref. [11], simple edge detection based CP selection is done. Matching is done with invariants based similarity measure along with this tine-plate spline interpolation is used. This algorithm is tested on SPOT and SAR image data set resultant registration accuracy observed is less than 0.3 pixels at individual CPs and RMSE of less than 0.2 pixels. Cloud removal was one of the challenge that time, Ik-Hyun Lee and T. Mahmood [12] worked on cloudy images for robust registration using two step segmentation, Ultra Scale SIFT (US-SIFT) for CP selection and k-means clustering. Results shown the segmentation accuracy increased by 93.26% and registration accuracy by 24.83%. On the similar lines Morgan McGuire and Harold Stone [13] tried fractional and binary masks at low frequency wavelet bands to find the correct match points. Unmasked algorithms use to miss the CPs in those occluded regions. It's been noted that fractional masks are more efficient than binary masks one or both input images are occluded.

In Ref. [14], [15] Goncalves, Luts Corte-Real [3] proposed AIR algorithm which combines image segmentation and well known SIFT approach to obtain CPs, complemented by a robust procedure of outlier removal.

M. Gong, S. Zhao, L. Jiao, D. Tian and S. Wang discussed a novel Coarse-to-Fine Scheme for Automatic Image Registration Based on SIFT and MI in ref. [7]. The preregistration (Coarse registration) process is implemented by robust SIFT approach to remove outliers and fine tuning (fine registration) is achieved by the maximization of MI using modified Marquardt-Levenberg search strategy in a multiresolution framework. Experimental work on optical-optical combination shows that the accuracy achieved is one hundredth of a pixel.

Regarding the registration of multiview images with the difference in the acquisition angle and the terrain elevation, the thin – plate spline model is more suitable than affine transformation model, to handle the effects introduced by different acquisition angle and terrain elevation [16].

5. Recently Developed Methodology

Convolutional Neural Networks (CNN) and Deep Neural Networks (DNN) are been used in various image related applications as these network architectures are good at finding low level and abstract feature [17]. In the recent studies CNN and DNN are tested with some trained data sets like VGG16, ImageNet having 1000 – 1500 images for feature extraction and/or image classification. In [18] Yang, Zhuoqian Dan, Tingting Yang, Yang constructed a feature descriptor using the output of certain layers of VGG-16 network. Simple convolution, pooling and fully connected layers, based on visualization convolution filters and some trail-and-error experimentations gives single layer output as a feature, on the similar lines several networks to be selected to build feature descriptors. These feature points are generated at the center of specific sized image patches. Data set used in this literature is UAV and satellite images from Google Earth. Feature matching precision test result for satellite image dataset is 71.7 % using traditional SIFT method while using NN method its 95.65% and RMSE 12.63 pixels which is almost half of the other methods, therefor results are acceptable. The illustration of the result can be observed by the figures 1 to 6 f or both data sets. Ye, Famao, Su, Yanfei, Xiao, Hui, Zhao, Xuqing Min, Weidong [19] built a customized data set to fine-tune the VGG-16 model and ImageNet to obtain CNN features and combined conventional SIFT features. There after combined features are fed to PSO-SIFT algorithm to register the images. This method gives comparable results with the conventional algorithm as, RMSE is 0.7321 whereas tradition SURF algorithm RMSE is 0.80. Despite of getting acceptable results NN are still challenging in case RS image analysis because of following difficulties [16], [17], [18]

- Difficult to handle high quality RS images
- Less availability of training data
- Deciding the depth of NN
- Timing complexity

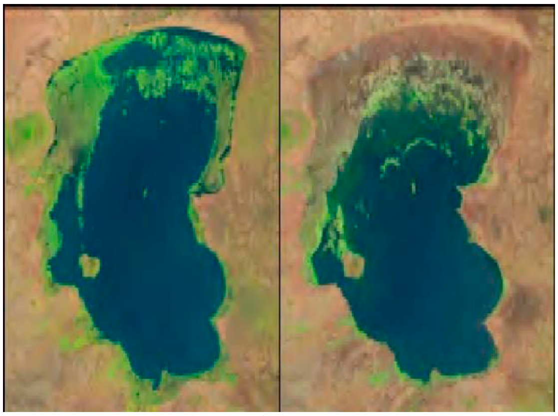

Fig. 1

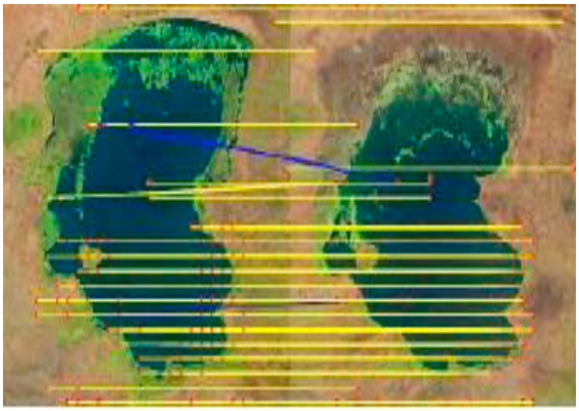

Fig. 2

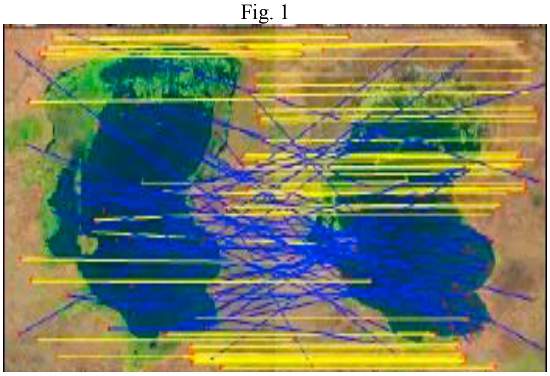

Fig. 3

Fig. 4

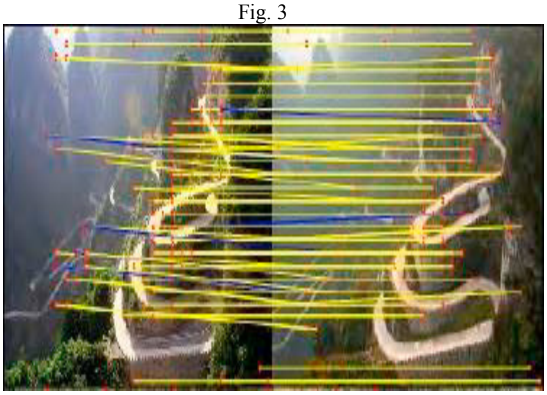

Fig. 5

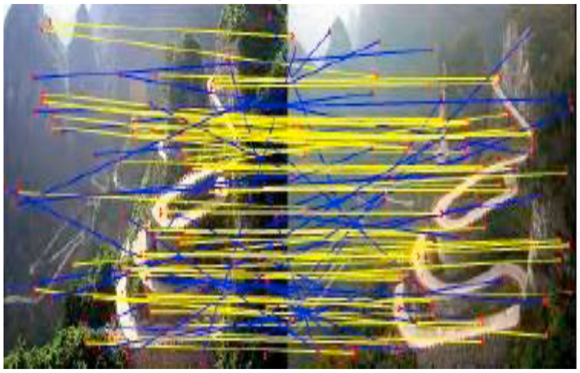

Fig. 6

Fig. 1. Multi-temporal Google Earth dataset [17] Fig. 2. Feature matching using deep learning algorithm [17] Fig. 3. Feature matching using SIFT algorithm [17] Fig. 4. Multi-temporal UAV image dataset [17] Fig. 5. Feature matching using deep learning algorithm [17] Fig. 6. Feature matching using deep learning algorithm [17].

For many important applications the fusion of optical and Synthetic aperture radar (SAR) images are used. As SAR images are captured by radar these images are not affected by clouds, weather conditions which makes it useful for many applications. The only difficulty with SAR images is the presence of speckle noise. In recent 3-4 years phase congruency method has been introduced to work on SAR and optical image registration. Phase congruency (PC) is nothing but finding the key points of high degree of order in Fourier domain. PC is robust to speckle noise, image illumination and contrast. In Ref. [20], basic PC and SAR SIFT is used which increased the number of correct CPs. After that, PC and weighted SIFT is used in [21] and PC feature detection along with spatial constraint matching is implemented in [22], to acquire both spatial and spectral features which resulted into satisfactory number of correct CPs, improved RMSE as compared to other state-of-art methods. Spatial constraints helps to remove outliers very effectively but with increased timing complexity. In [23], Jianwei Fan et. al. have designed UND-harris feature extraction method and used it along with PC structural descriptors which greatly reduced the effect of spackle noise on feature extraction. This method is more suitable for images having more structural details.

6. Algorithm Evaluation

Root Mean Square Error (RMSE) is the most popularly known and used accuracy measure. Different variants of RMSE like RMSE all, RMSE live one, Ratio of correct CPs to total CPs identified is also called feature matching ration gives the precision and accuracy of algorithm.

7. Conclusion and Discussion

In last two decades a lot of work has been done in RS image registration manual or automatic but still there are some problems to get correct correspondence due to some natural reasons. From the various literatures it has been seen that for feature detection SIFT is most frequently used method but SIFT cannot be applied directly to the remote sensing images therefore numerous forms of SIFT in combination of information measures are used to get good and enough number of CPs. These SIFT based methods suffer from the problems of insufficient feature points and high outlier ratio under severe appearance change. Therefore it's important to concentrate on both local and global features. Next challenging task is to find correct feature match and transformation estimation with trade-off of large computational time. Another critical point is how to differentiate between image deformation and actual change in the scene. Neural Networks in combination with traditional methods gives satisfactory results with the benefit of less computation time. Although CNN based methods are not efficient at this stage but these method can give quick results for larger datasets as compared to traditional approach. It is required to work on different CNN like AlexNet, GoogleNet with various different data sets as well as with different combinations of traditional methods to CNN to get more improved results. Following is summary in table 1.

Table 1. Summary of results referred

Author	Data set	Methodology	Results	Conclusion
D.Fedorov,L. M.G. Fonseca,C.Kenney, B.S. Anjunath [18]	Various combinations of RS images	Area based registration	Satisfactory results for different combinations.	Small number of control points reduces the registration speed exponentially. Wavelet methods is efficient
Bentoutou, Youcef Taleb, Nasreddine Kpalma, Kidiyo Ronsin, Joseph [2]	SPOT and SAR	Invariant relations and the elastic thin-plate spline (TPS) interpolation	RMSE < 0.3 pixels, achieved at each individual CP RMSE < 0.2 pixels, obtained for the ten CPs	TPS interpolation is fast method. Yields accurate global and local registrations.
Hong, Gang Zhang, Yun [19]	IKONOS (PAN), Quickbird (MS)	Combination of feature-based and area-based matching DWT & redundant wavelet transform ("à trous")	For IKONOS pair RSME = 3.21, For IKONOS & QuickBird Pair RSME = 0.95	Tested & compared, 1st , 2nd and 3rd order polynomial transformations & triangle based methods in both cases, the RMSE of the triangle-based result is the smallest. Large number of control points must be selected manually.
Peter Bunting, Richard Lucas, Fr´ed´eric Labrosse [20]	(LiDAR,HyMap,AIRSAP)	Bunting's arear based algorithm and cross correlation method	Bunting algorithm, results in most reliable results across a range of image modalities, scales and transformations.	To improve the existing algorithm to identify CPs when large errors are inputted into the network before propagation

Author	Data set	Methodology	Results	Conclusion
Fouad, M. M. Dansereau, R. M. Whitehead, A. D. [21]	LANDSAT and IKONOS	Intensity-based registration approach Using Huber M-estimator	Geometric Registration Accuracy increases by 88.5% with average increase of 33.1% in computation time	Huber M-estimator is not as vulnerable as LSE (Least square estimator) to outliers.
Gonçalves, Hernâni Corte-Real, Luís Goncalves, José A. [13]	Landsat 5 & Hyperion images	Image segmentation & SIFTand Bivariate histogram for feature matching	LANDSAT pair RMSall = 0.54, dratio = 0.3 Hyperion pair RMSall = 0.35, dratio = 0.8	The cc information can discriminate local features that have similar local appearance under noisy environment.
Xie, Jun Li, Baohua Han, Wei Bao, Jinhe Gu, Feng Guo, Fei [22]	SPOT PAN image pair	A tiny facet primitive triangulation algorithm based on SIFT key points	Obtained registration accuracy is 87%	The accuracy rate of the paper was significantly improved than the conventional method
Fouad, M. M. Dansereau, R. M. Whitehead, A. D. [4]	AVIRIS data	Non subsampled Contourlet Transform (NSCT) and SIFT	Correct CP matching rate = 92.55%	Method is more effective for infrared and visible image registration. Improved correct CP matching rate compared to SIFT.
Yuhao Zhou, Qiuze Yu,Sunni Hua,Wen Yin, Yuanxiang Li [35]	IKONOS Pair	SIFT AND THIN-PLATE SPLINE (TPS)	RSME = 0.7566 Feature matching improved by 33% MI increased by 17%	Experimental results on two images of different sensors show the effectiveness of the proposed approach.
Li, Yingying Liu, QingjieJing, Linhai Liu, Shuo Miao, Fengxian [25]	SPOT IMAGE Pair	SIFT and optimization of local self-similarity mutual information (LSS MI)	Inlier percentage for Rigid its 90.70%, Non-rigid 99.75%, affine 99.75%.	Need to test the effectiveness of our LLT for 3-D feature matching by extending the feature extraction method to 3-D images.
Chen, Shuhan Li, Xiaorun Zhao, Liaoying [26]	Hypersion & LANDSAT	Cramér–Rao lower bound (CRLB) and leave-one-out cross-validation (LOOCV) approach	RMSE = 0.5921, Inlying CPs Percentage 91.31%	Registration accuracy does not need any ground truth to be determined & characterizes individual pairs of registered images, taking into account their inherent structure.
Xianmin Wang, Qizhi Xu [5]	QuickBird, WorldView-2	Canny detector and line-Point Invariant	Correct Matching Rate 84.1%	False keypoint matches are greatly reduced and the correct match rate is significantly enhanced
Wahed, M., Gh.S. El-tawel El-karim, A.Gad [15]	Landsat TM pair	Combination of Steerable Pyramid Transform, SIFT, Affine transform	RMSE 0.11, PSNR = 33.47	Ability to increase the number of matched points
Maoguo Gong Shengmeng Zhao Licheng Jiao Dayong Tian Shuang Wang [24]	SPOT IMAGE Pair	SIFT & MI	RMSE = 0.031, MI = 3.467	Improves the accuracy of registration and achieves good computational efficiency
Cui, Song Zhong, Yanfei [28]	LiDAR and hyperspectral image	multi-scale phase consistency (MS-PC)	RMSE = 2.98	Requires less tie points.

Author	Data set	Methodology	Results	Conclusion
Fan, Jianwei Wu, Yan Wang, Fan Zhang, Peng Li, Ming [27]	SAR pair	SIFT and local search refinement (LSR)	RMSE = 0.63	More robust and efficient. But may not work when images have a significant difference in intensity
Ye, Famao Su, Yanfei Xiao, Hui Zhao, Xuqing Min, Weidong [17]	MS and SAR	VGG16 model to extract features and low level SIFT & high level CNN fusion	RMSE = 0.7321	Both the aligning accuracy and the number of correct correspondences is acceptable.
Yang, Zhuoqian Dan, Tingting Yang, Yang [16]	-pretrained VGG network. -UAV dataset -Google earth dataset	VGG-16 network, expectation maximization (EM) , Gaussian mixture model (GMM)	RMSE = 12.63 pixel	Improved accuracy in feature pre matching test.
S. Cui and Y. Zhong [28]	LiDAR & Visible band images	Multi scale phase congruency	RMSE = 2.23	Method is robust to radiation differences
S. Jiang, B. Wang, X. Zhu, M. Xiang, X. Fu, and X. Sun [21]	SAR & Optical images	- PC SIFT - PC gradient SIFT	Correct CP ratio=83% Correct CP ratio = 77.8%	Reduced Euclidian distance & increased correct CPs ratio

References

[1] X. Dai and S. Khorram, “A Feature-Based Image Registration Algorithm Using Improved Chain-Code Representation Combined with Invariant Moments,” vol. 37, no. 5, pp. 2351–2362, 1999.

[2] Y. Bentoutou, N. Taleb, K. Kpalma, and J. Ronsin, “An automatic image registration for applications in remote sensing,” *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 9, pp. 2127–2137, 2005.

[3] C. Harris and M. Stephens, “A Combined Corner And Edge Detector,”AVC, pp. 147–152, 1988.

[4] S. Paul, U. C. Pati, and S. Member, “Using Modified Uniform Robust SIFT,” vol. 13, no. 9, pp. 1300–1304, 2016.

[5] H. Qingqing, Y. Jian, W. Chengyi, C. Jingbo, and M. Yu, “Sensing Image Using NSCT and SIFT”, IGARSS, vol. 2, no. 1, pp. 2360–2363, 2012.

[6] “Multi-Sensor Optical Remote Sensing Image Registration Based On Line-Point Invariant", Xianmin Wang , Qizhi Xu Beijing Key Laboratory of Digital Media School of Computer Science and Engineering , Beihang University , Beijing 100191 , China, pp. 2364–2367, 2016.

[7] L. Gottesfeld, “A Survey of Image Registration”, ACM Comput,ng Surveys, no. 4, 1992.

[8] “Introduction To Remote Sensing Image Registration" Jacqueline Le Moigne Software Engineering Division , NASA Goddard Space Flight Center , Greenbelt , MD 20771 , USA,” pp. 2565–2568, 2017.

[9] B. Zitová and J. Flusser, “Image registration methods: A survey,” *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, 2003.

[10] J. Inglada, V. Muron, D. Pichard, and T. Feuvrier, “Analysis of artifacts in sub-pixel remote sensing image registration,” *Rev. Fr. Photogramm. Teledetect.*, vol. 45, no. 184, pp. 29–34, 2006.

[11] X. Otazu, M. González-Audícana, O. Fors, and J. Núñez, “Introduction of Sensor Spectral Response into Image Fusion Methods. Application to Wavelet-Based Methods,” *IEEE Trans. Geosci. Remote Sens.*, 2005.

[12] I. H. Lee and M. T. Mahmood, “Robust registration of cloudy satellite images using two-step segmentation,” *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1121–1125, 2015.

[13] M. McGuire and H. S. Stone, “Techniques for multiresolution image registration in the presence of occlusions,” *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1476–1479, 2000.

[14] H. Gonçalves, L. Corte-Real, and J. A. Goncalves, “Automatic image registration through image segmentation and SIFT,” *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 7, pp. 2589–2600, 2011.

[15] H. Wen and X. You Sheng, “An improved SIFT operator-based image registration using cross-correlation information,” *Proc. - 4th Int. Congr. Image Signal Process. CISP 2011*, vol. 2, pp. 869–873, 2011.

[16] M. Wahed, G. S. E. -, and A. G. El-karim, “Automatic Image Registration Technique of Remote Sensing Images,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 4, no. 2, pp. 177–187, 2013.

[17] L. Zhang, L. Zhang, and B. O. Du, “Deep Learning for Remote Sensing Data,” IEEE Geoscience and remote sensing magazine, june, 2016.

[18] Z. Yang, T. Dan, and Y. Yang, “Multi-temporal remote sensing image registration using deep convolutional features,” *IEEE Access*, vol. 6, pp. 38544–38555, 2018.

[19] F. Ye, Y. Su, H. Xiao, X. Zhao, and W. Min, "Remote Sensing Image Registration Using Convolutional Neural Network Features," IEEE Geoscience And Remote Sensing Letters, Vol. 15, No. 2, February 2018
[20] S. Cui and Y. Zhong, "Multi-modal remote sensing image registration based on multi-scale phase congruency," *2018 10th IAPR Work. Pattern Recognit. Remote Sensing, PRRS 2018*, pp. 1–5, 2018.
[21] S. Jiang, B. Wang, X. Zhu, M. Xiang, X. Fu, and X. Sun, "Registration Of SAR And Optical Images By Weighted SIFT Based On Phase Congruency," no. 2, pp. 8885–8888, 2018.
[22] W. Ma, Y. U. E. Wu, Q. Su, and Y. Zhong, "Remote Sensing Image Registration Based on Phase Congruency Feature Detection and Spatial Constraint Matching," *IEEE Access*, vol. 6, pp. 77554–77567, 2018.
[23] J. Fan, Y. Wu, M. Li, W. Liang, and Y. Cao, "SAR and Optical Image Registration Using Nonlinear Diffusion and Phase Congruency Structural Descriptor," IEEE Transactions On Geoscience And Remote Sensing pp. 1–12, 2018.
[24] Maoguo Gong, Shengmeng Zhao, Licheng Jiao, Dayong Tian, and Shuang Wang, "A Novel Coarse-to-Fine Scheme for Automatic Image Registration Based on SIFT and Mutual Information," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4328–4338, 2014.
[25] Y. Li, Q. Liu, L. Jing, S. Liu, and F. Miao, "A genetic-optimized multi-angle normalized cross correlation SIFT for automatic remote sensing registration," *Int. Geosci. Remote Sens. Symp.*, vol. 2016-Novem, pp. 2586–2589, 2016.
[26] S. Chen, X. Li, and L. Zhao, "Multi-source remote sensing image registration based on sift and optimization of local self-similarity mutual information," *Int. Geosci. Remote Sens. Symp.*, vol. 2016-Novem, pp. 2548–2551, 2016.
[27] J. Fan, Y. Wu, F. Wang, P. Zhang, and M. Li, "Representation of Image Patch Feature for SAR Image Registration," *Ieee Trans. Geosci. Remote Sens.*, pp. 1–13, 2016.
[28] S. Cui and Y. Zhong, "Multi-modal remote sensing image registration based on multi-scale phase congruency," *2018 10th IAPR Work. Pattern Recognit. Remote Sensing, PRRS 2018*, pp. 1–5, 2018.
[29] W. Ma, J. Zhang, Y. Wu, L. Jiao, H. Zhu, and W. Zhao, "A Novel Two-Step Registration Method for Remote Sensing Images Based on Deep and Local Features," *IEEE Trans. Geosci. Remote Sens.*, vol. PP, pp. 1–10, 2019.
[30] L. M. G. Fonseca and B. S. Manjunath, "Registration techniques for multisensor remotely sensed imagery," Photogramm. Eng. Remote Sensing, vol. 562, pp. 1049–1056, Sept. 1996.
[31] R. Bajcsy, S. Kovacic, "Multiresolution elastic matching", Computer Vision, Graphics and Image Processing 46 (1989) 1–21.
[32] A. A. Cole-Rhodes, K. L. Johnson, J. Le Moigne, and I. Zavorin, "Mul- tiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient," *IEEE Trans. Image Process., vol.* 12, pp. 1495–1511, Dec. 2003.
[33] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," Int. J. Comput. Vis., vol. 37, no. 2, pp. 151–172, Jun. 2000.
[34] Jiayi Ma, Huabing Zhou, Ji Zhao, Yuan Gao, Junjun Jiang, and Jinwen Tian, "Robust Feature Matching for Remote Sensing Image Registration via Locally Linear Transforming", *IEEE Transactions on Geoscience and Remote Sensing 2015,* vol. 53, Issue 12 pp. 6469-6481.
[35] Yuhao Zhou, Qiuze Yu,Sunni Hua,Wen Yin, Yuanxiang Li, "An Automatic Global-To-Local Image Registration Based On SIFT And Thin-Plate Spline (TPS)", IGARSS, IEEE 2013.