Journal of Geomatics Vol. 13, No. 1, April 2019

An approach for generation of multi temporal co-registered optical remote sensing images from Resourcesat-2/2A sensors

Indranil Misra*, Vivek Sharma, S. Manthira Moorthi and Debajyoti Dhar
Space Applications Centre, ISRO, Ahmedabad 380015, India
*E-Mail: indranil@sac.isro.gov.in

(Received: Jan 02, 2019; in final form: Jun 17, 2019)

Abstract: Multi temporal remote sensing data availability is crucial for remote sensing community to study the planet earth. The multi temporal data can be used for scientific studies only after sub-pixel level image registration of all multi date acquisitions. Multi temporal co-registered remote sensing data is essential for many real time applications such as crop forecasting, forestry, inland water mapping, change detection and time series analysis. Multi temporal co-registered data is the first pre-processing step for generation of Analysis Ready Data (ARD) cube that allows immediate analysis with minimal user efforts. Image Registration at sub-pixel level requires optimal geometric transformation of all datasets such that data stack is geometrically aligned over each other. The major step in image registration is feature detection to generate a collection of tie points with outliers removed, match the feature points and finally estimate the transformation parameters. In this paper, we present an approach for multi temporal image registration that employs Scale Invariant Feature Transform (SIFT) technique along with a segmented affine based transformation model for different image segments to correct the data at geographic coordinate space to achieve sub-pixel level geometric accuracy. The key purpose of this work is to generate co-registered data stack for image analysis. Resourcesat-2/2A (RS-2/RS-2A) LISS-3 data from Indian Remote Sensing Satellite (IRS) is used for multi temporal co-registration task. The RS-2/2A LISS-3 data have spatial resolution of 24 meters and combination of both RS-2 and RS-2A LISS-3 gives better temporal repetivity to cover the same region in less number of days. The technique developed is tested with LISS-3 data of same region acquired during time interval from 2012 to 2018.

Keywords: LISS-3, SIFT, ARD, Affine Transform, Image Registration, Multi Temporal

1. Introduction

Remote Sensing data acquisitions at different timeline encourages researchers to study the feature changes that can be used for diverse space applications such as agricultural monitoring, disaster management and has the capability to solve complex problems related to earth surface studies (Haiganga and Zhou, 2008). The multi temporal data analysis requires robust preprocessing of data stack to generate sub-pixel level image registration accuracy. This open new door for development of next generation methods of image registration for multi spectral remote sensing data. The critical steps in image registration is to detect stable tie points by removing outliers and estimation of transformation model that establishes a mapping between the images. The image registration task is the most important requirement of remote sensing world where large amount of multi temporal satellite images need to be co-registered for every kind of scientific studies. The data stack generated should be evaluated for both relative and absolute geometric accuracies. The relative geometric accuracy helps us to know the image-to-image registration performance of the stack whereas absolute geometric accuracy of the stack can be evaluated with reference images or ground control points (GCP) to measure the location of an object in data with respect to its true location on the earth.

1.1 Feature detection and description using SIFT

The image registration process starts with efficient feature detection mechanism. Detecting the appropriate feature point is the crux of image registration problem (Aksakal, 2013; Zitova and Flusser, 2003). In literature, many kind of interest point detectors are reported to detect stable feature points (Schmid et al., 2000). But Scale Invariant Feature Transform (SIFT) is one of the robust feature detection technique that gives better performance and it is invariant to scale, rotation, illumination conditions and image noise (David, 1999). The advantage of SIFT over other feature detector is that it is not only detect the feature points but it also provides the description of the feature point in comparison to the other feature detectors like Harris Corner. SIFT is also capable to detect feature points at sub-pixel level which is needed for accurate transformation parameter estimation.

The SIFT algorithm starts with construction of scale space extrema (David, 1999) using Gaussian Kernel (G). The eq. 1 shows the relation between Smooth Image (L) and Input Image (I) at point (x, y)

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

Where, $G(x, y, \sigma) = \frac{1}{2\pi\sigma} e^{-(x^2+y^2)/2\sigma^2}$ (2)

It has been found that Laplacian of Gaussian (LOG) is able to provide stable features and give excellent notion of scales but it is computationally costly. So the Difference of Gaussian (DOG) pyramid is created which is considered to be the close approximation of Laplacian of Gaussian (LOG) (David, 1999).

The next task is to scan each DOG image to locate the minimum and maximum around all the neighboring points including the scale (Zheng et al., 2008). Figure 1 shows the DOG pyramid generated from scale space and extrema located from the DOGs.

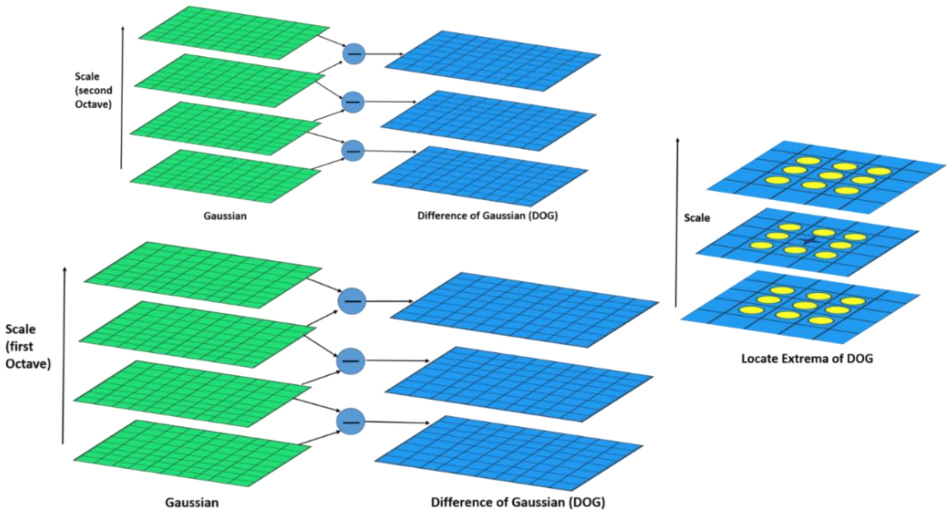

Figure 1: DOG pyramid and extrema localization

The potential stable feature point at sub-pixel level localization and initial outliers can be removed by using Taylor Series approximation

$$D(x) = D + \frac{\partial D^T}{\partial x}x + \frac{1}{2}x^T\frac{\partial^2 D^T}{\partial x^2}x \quad (3)$$

and differentiate set to zero

$$\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2}\frac{\partial D}{\partial x} \quad (4)$$

to get location in terms of (x, y, σ).

The outliers that are still left can be removed by doing contrast limit filtering and edge response elimination. The key points orientation is computed using gradient of each blurred image. The orientation assignment achieves rotation invariance. It mainly computes central derivatives, gradient magnitudes (m) and direction (θ) of smooth image (L) at the scale of keypoint (x,y). The weighted direction histogram in a neighborhood of a keypoint in form of bin is created and finally select the peak as direction of the keypoint.

$$m(x,y) = \sqrt{\left(L(x+1,y) - L(x-1,y)\right)^2 + \left(L(x,y+1) - L(x,y-1)\right)^2} \quad (5)$$

$$\theta(x,y) = \tan^{-1}(\ (L(x,y+1) - L(x,y-1))/\ (L(x+1,y) - L(x-1,y))) \quad (6)$$

The descriptor is built by sampling the point around keypoints. Rotate the gradients and coordinates by previously computed orientation. Then separate the region into sub-region. The histogram is created for each sub-region with specified bins. The descriptor is stored as element vector for each keypoints. The feature points are detected using SIFT independently for both input image and reference image. The matching finds the nearest neighbor i.e. a keypoint with minimum Euclidean Distance (Rabin et al., 2008).

2. Segmented affine transformation estimation model

The input remote sensing image is radiometrically corrected and each pixel is tagged with latitude/longitude (lat/lon) information. In this approach, the matched featured points are divided into multiple segments and affine transformation model is computed for each segment. The affine transform (Zhou, 2016) for a segment is shown as:

$$aff(k) = \{a_{ok}, a_{1k}, a_{2k}, a_{3k}, a_{4k}, a_{5k}\} \quad (7)$$

Where k = segment number and a_{ik} = affine transform parameters

The tagged input (lat, lon) of image point (x, y) transformed to corrected geo-point (lat_{cor}, lon_{cor}) depending upon the segment number k of the point (x, y).

$$\begin{bmatrix} lat_{cor} \\ lon_{cor} \end{bmatrix} = \begin{bmatrix} a_{ok} & a_{1k} \\ a_{2k} & a_{3k} \end{bmatrix}\begin{bmatrix} lat \\ lon \end{bmatrix} + \begin{bmatrix} a_{4k} \\ a_{5k} \end{bmatrix} \quad (8)$$

Figure 2 shows an optical remote sensing data divided into four segments in pixel wise direction. Each segment contains matched feature points which is used for affine model estimation for that segment. Segment size is determined on the basis of input image size and pixel size. The four different affine models estimated for the segmented image regions and geographic grid is modified using its corresponding affine transformation model shown in eq. 8 to generate a single final modified geometric grid that maintains the continuity of transitions in the image segments. The output grid generated with segmented affine model approach reduces internal distortion in the final geo-corrected data and provide sub-pixel level registration accuracy with respect to the reference image.

3. Processing workflow developed

The approach described in above sections need to be realized in a data processing workflow for generation of the multi temporal co-registered products. The input to the workflow is multi spectral radiometrically corrected remote sensing data with tagged geographic coordinate information. Using the latitude and longitude coordinates, corresponding reference tile need to be fetched from reference database as shown in figure 3.

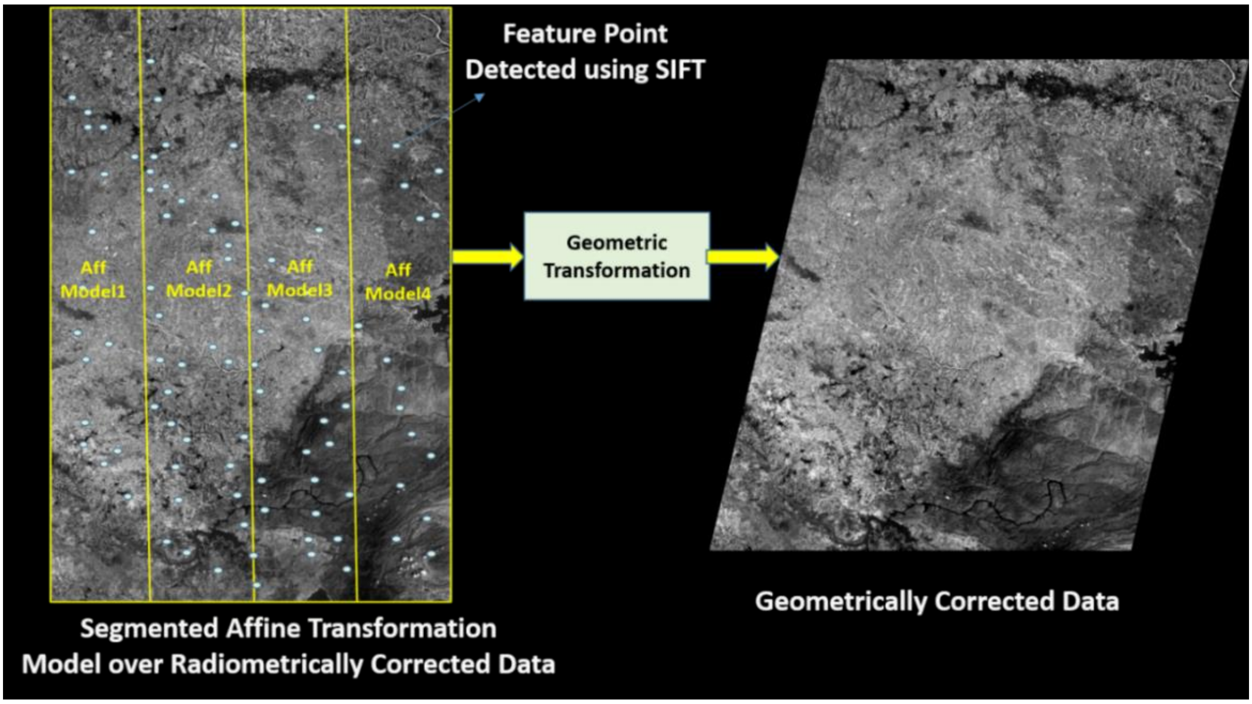

Figure 2: Segmented affine transformation model over remote sensing data

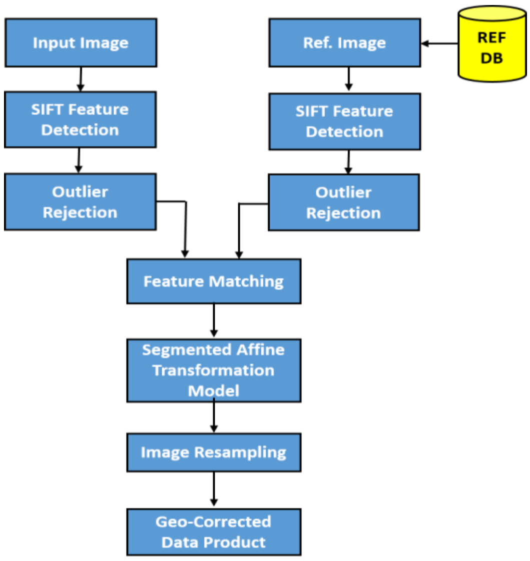

Figure 3: Data processing workflow

The reference tiles are stored as standard geo-tiff ortho-rectified images in UTM map projection. The absolute location accuracy of reference images is less than 12 meters and it remain same for both RS-2/RS-2A. SIFT features are detected in both the images and matched to get the set of key points required for estimation of affine transformation. The segmented affine transformation models are generated as per the approach in section-III. The input image is geometrically transformed using image resampling technique at final stage of data processing to generate geometrically corrected data product. Since resampling kernel is used for input to output transformation. Figure 3 shows the processing workflow developed for multi temporal image registration. The execution time of the algorithm is less than a minute for a scene of 6000 lines*6000pixels image size

4. Data used and results achieved

The processing workflow developed is tested with recent acquisitions of LISS-3 data on board Resourcesat-2A. Resourcesat-2A satellite was launched by ISRO in Dec. 2016. The LISS-3 sensor is primarily used for earth resource monitoring that includes agriculture, forest and other natural resource observation and planning. The details of sensor, data used for testing and its geo-meta information is shown in table 1. Figure 4 shows the false color composite (FCC) image of the study area acquisition by Resurcesat-2A LISS-3 (L3). The multi date acquisitions of the same study area (Figure 4) is considered for image registration exercise.

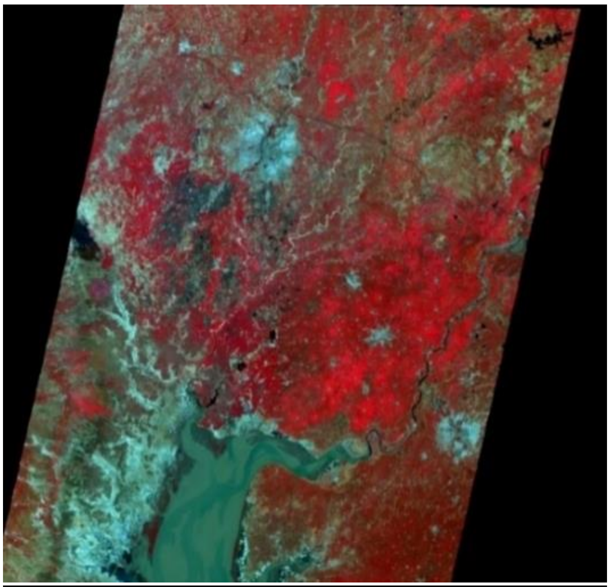

Figure 4: RS-2A LISS-3 FCC image (Path/Row: 93/56)

Table 1: LISS-3 sensor and data details

Spatial Resolution (Meters)	24.0
Swath (Kilometers)	141
Spectral Bands (Microns)	0.52-0.59 0.62-0.68 0.77-0.86 1.55-1.70
Quantisation (Bits)	10
Date Rate (MBPS)	105
Revisit Time (in Days)	24
Path/Row Scene (Referencing Scheme) Used for testing	93/56
Region (Cities) Covered	Ahmedabad, Gandhinagar, India
Multi Date Cloud Free Data Acquisitions	1. 25th Dec 2016 2. 18th Jan 2017 3. 31st March 2017 4. 18th May 2017
Geographic Corner Coordinates (lat/lon) of Scene	**UL:** 23d24'16.04"N, 71d46'20.12"E **UR:** 23d25'54.53"N, 73d37'49.12"E **LR:** 21d53'27.37"N, 73d38'43.88"E **LL:** 21d51'56.03"N, 71d48'28.95"E

The RS-2A LISS-3 multi date datasets of the scene information tabulated in table 1 are processed using the workflow developed. The performance of registration is examined both visually and quantitatively for detailed analysis. The visual inspection is carried out by swiping the registered product over reference tile. All the dates are seen at same region with reference tile and the geometric fidelity shows all the images are registered over reference image. Figure 5 shows the image swipes of all the four dates listed in table 1 with reference at sub-pixel level zoom view.

The image swipes help us to visualize the registration performance at particular regions in the image, which is not enough to quote the final registration accuracy. The quantitative parameters are needed to conclude the overall image registration performance. The evaluation of registered product can be done by computing root mean square error (RMSE). The RMSE is calculated by mapping the relation of stable matched feature points obtained from registered data and reference data [9]. The RMSE computed is an indicator to show that relative geometric accuracy between the images. The SIFT detector is used here to fetch the feature points. The matched correspondence is tabulated in table 2 and it shows the RMSE is within a pixel for all the datasets with respect to reference.

Table 2: Registration Performance Table

Date of Acquisition	No. of Control Points	RMSE (in pixels)
25th Dec 2016	300	0.44
18th Jan 2017	454	0.38
31st Mar 2017	345	0.42
18th May 2017	286	0.46

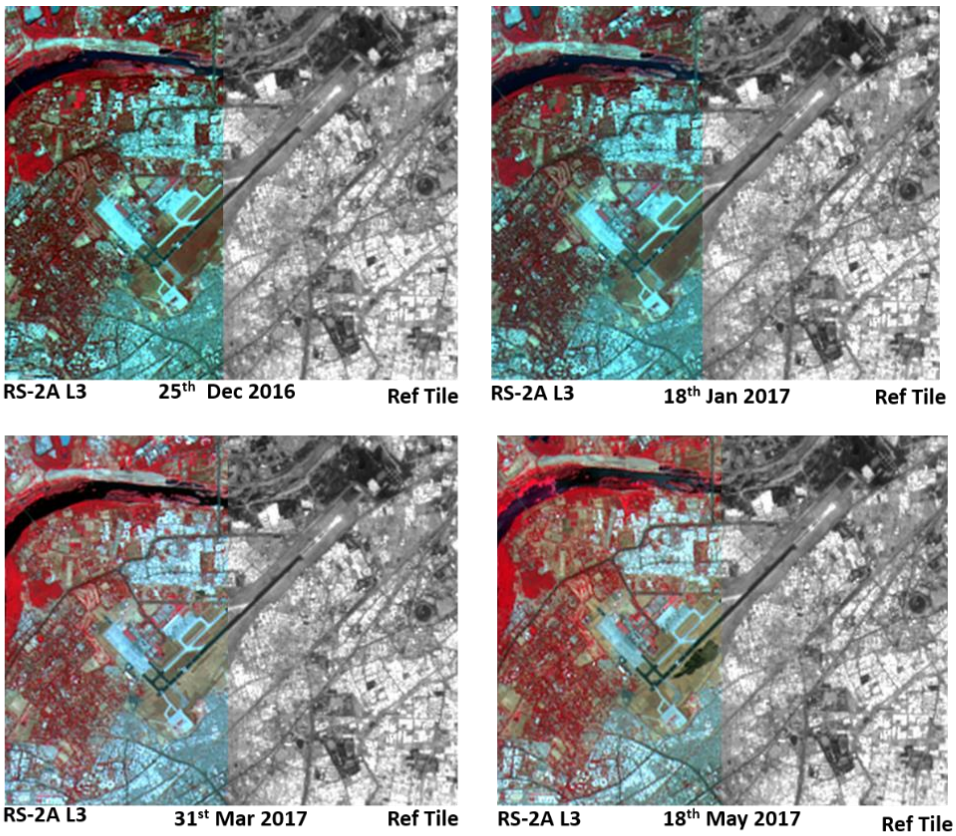

Figure 5: Multi date image swipes (Overlay) against Ref Tile

In addition, multi temporal products in combination of Resourcesat 2 and 2A from the year 2012 to 2018 generated using the processing workflow covering mixed terrain also includes partially cloudy data. The number of multi temporal products considered for analysis is 24 across the span of 6 years acquisition of Resourcesat mission. It is observed that all the data products are well registered within a 0.5 pixel across the images with Band-3 of 23-May-12 as reference shown in figure 6. Also absolute location accuracy of all multi temporal data products is within a pixel. The performance of the registration is checked at multiple patches dominated by urban area, hilly terrain, agriculture land and vegetation cover as shown in figure 7. The RMSE in pixel is computed for the patches with respect to the reference and shown in table 3. All 24 LISS-3 multi temporal registered data products stack meets the specifications of less than 0.25 pixels band-to-band registration (BBR) between the multi spectral channels and also internal distortion is at sub-pixel level which is needed for any kind of change detection studies and analysis. The approach developed is tested in multi temporal coastal scenes dominated by ocean and it has been found that relative registration error is at sub-pixel level registration accuracy

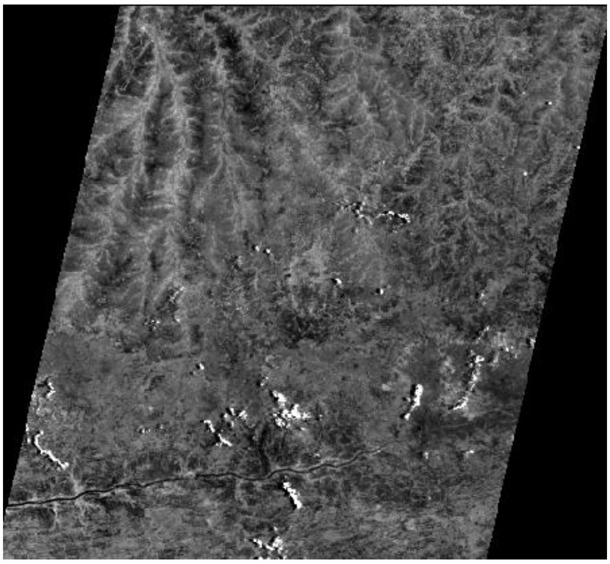

Figure 6: RS2 LISS-3 BAND-3 23 May 2012 (Path/Row: 93/56)

Table 3: Registration Accuracy at different patches

Regions	**No. of Control Points**	**RMSE (in pixels)**
Urban	60	0.47
Hilly	23	0.53
Agriculture	18	0.42
Vegetation	34	0.45

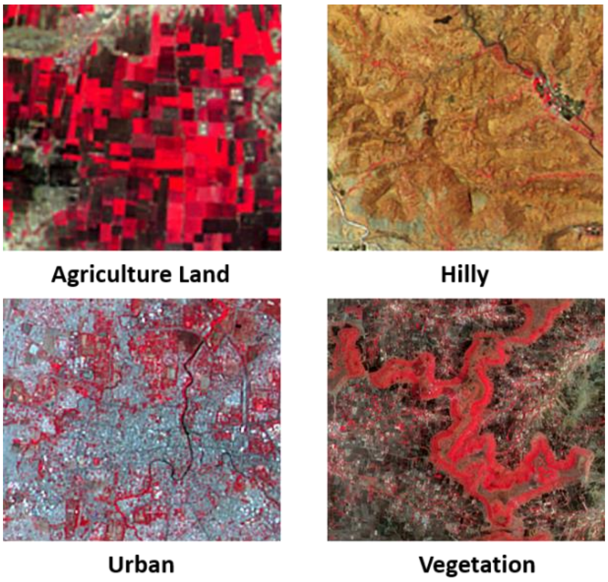

Figure 7: Multiple patches extracted from image to check registration accuracy

5. Conclusion and future work

The image registration for multi temporal data sets is achieved at sub-pixel level accuracy, which is evaluated both visually and quantitatively. SIFT with segmented affine transformation model is the approach developed for co-registration of multi date remote sensing data acquisitions. The co-registered data becomes analysis ready data product which can be used for any kind of earth monitoring studies. The workflow need to be tested with many other multi date acquisitions covering different terrains. The next task is to compare SIFT based approach developed with other feature detection techniques. At present the processing workflow handles same sensor multi temporal data. The future work would include multi-sensor data registration from different satellite platforms and different sensors.

Acknowledgements

The authors acknowledge Shri D K Das, Director, Space Applications Centre (SAC), ISRO for his encouragement. The authors also thank other team members of optical data processing for their support during the course of work.

References

Aksakal, S.K. (2013). Geometric accuracy investigations of SEVIRI High Resolution Visible (HRV) Level 1.5 Imagery, Remote Sensing, (5), 2475-2491; doi: 10.3390/rs5052475.

David, G. L. (1999). Object recognition from local scale-invariant features, International Conference on Computer Vision, 1150-1157.

Haiganga, M.G and Q. Zhou (2008). A review of multi-temporal remote sensing data change detection algorithms, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XXXVII. Part B7. Beijing, pp. 757-762.

Rabin, J., J. Delon and Y. Gousseau (2008). A statistical approach to the matching of local features, IEEE ICPR (40, 41).

Schmid, C., R. Mohr and C. Bauckhage (2000). Evaluation of interest point detectors, International Journal of Computer Vision, 37(2), 151–172.

Zheng, Z., C. Zhiguo and X. Yang (2008). Multi-spectral remote image registration based on SIFT, Electronics Letters, 44, 107-108.

Zhou, J. (2016). A novel image registration algorithm using SIFT feature descriptors, International Conference on Smart City and Systems Engineering, November 25-26, 2016.

Zitova, B. and J. Flusser (2003). Image registration methods: A survey, Journal of image and Vision, Elsevier, 21, 977-1000.