International Journal of Computer Applications (0975 – 8887)
Volume 80 – No 14, October 2013

Remote Sensing Image Matching using Sift And Affine Transformation

Elsa Kuriakose
Dept. of Electronics and Communication.
Sree Narayana Gurukulam College of Engineering

Anjaly Viswan
Assistant Professor
Sree Narayana Gurukulam College of Engineering
Ernakulam , India

ABSTRACT
This paper presents remote sensing image matching using sift algorithm and affine transformation. The novelty in our approach is to find the features in the reference image and then match the input image with that of reference image using Affine Transformation. Both synthetic and real data have been considered in this work for the evaluation of the proposed methodology. After registering the image, the outliers are removed. A speeded up affine invariant detector is proposed in this paper for local feature extraction. The experimental results show that the proposed algorithm decreases the redundancy of key points and speeds up the implementation. It is able to account for differences in spectral content, rotation, scale, translation, different viewpoint, and change in illumination. The proposed technique improves the computational efficiency and decrease the storage requirement.

General Terms
Image Registration, Scaling coefficients, Algorithms, Image Transformation.

Keywords
Image segmentation, scale invariant feature transform, affine transformation, remote sensing .

1. INTRODUCTION
Remote sensing applications is one of the fields where further research on AIR is required. However under the scope of remote sensing applications, one of the major problems is related to the radiometric content (due to multisensory or multispectral pair of image). Moreover scale is frequently known, as most satellite images are provided with sufficiently accurate scale information. Similar comments may be applicable regarding differences in the view angle, as satellite images are roughly acquired at an altitude of 600 km. The main concept regarding remote sensing image matching is to obtain an accurate set of tie points and then apply the transformation function which is most suitable to the pair of images to be registered. The extraction and matching algorithm together identify common features in multi sensor situations from which tie points can be derived. From image segmentation methods, set of objects can be extracted present on an image. Image segmentation comprises a wide variety of methods, either for monochrome or color images. Most image segmentation methods can be classified according to their

nature : Histogram thresholding, feature space clustering, region based approaches, edge detection approaches, fuzzy approaches, neural networks, physics based approaches and any combination of these. SIFT algorithm proposed by David Lowe was the most representative image feature extraction method. It detected feature points using DOG (Difference-of – Gaussian) operator in scale space. An affine transformation does not necessarily preserve angles or lengths, though it does preserve ratios of distances between points lying on a straight line.

(a) (b)

Fig 1 (a) Reference image and (b)Input image

2. PROPOSED ALGORITHM
The important steps used in this work are given below.

1) Read the reference image and input image.

2) Extract the principal components from images.

3) Apply segmentation

4) SIFT approach

- 4.1 - create the scale space
- 4.2 - keypoint localization
- 4.3 - orientation and assignment
- 4.4 – keypoint descriptor

5) Obtain the keypoints from reference image and input image.

6) Apply Affine Transformation.

7) Image matching

8) Comparison using transformation function.

2.1 Input Image And Reference Image
Figure 1 shows example from the image sets used to register the detectors. In this dataset, there are 4 different changes in imaging condition: viewpoint changes, scale changes, JPEG compression and angular rotation.

Consider (X_{REF}, Y_{REF}) as the cordinates of a point from the reference image and (X_{NEW}, Y_{NEW}) as (pixel, line) of the corresponding point in the new image to be registerd.The relation between (X_{REF}, Y_{REF}) and (X_{NEW}, Y_{NEW}) may be written as

$$X_{REF} = f\,(X_{NEW}, Y_{NEW}) \quad (1)$$
$$Y_{REF} = f\,(X_{NEW}, Y_{NEW}) \quad (2)$$

where f and g are the functions that describe the relation between the coordinates of two images.

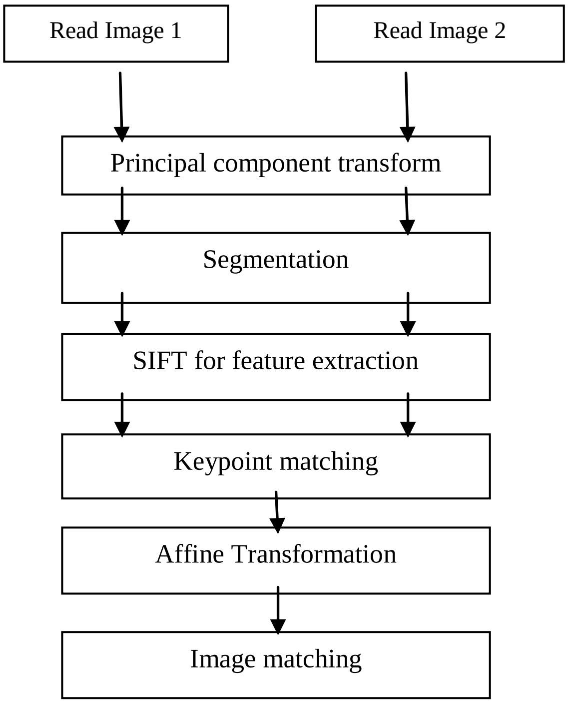


Fig 2. Main steps of the proposed methodology for image registration.

2.2) Data Reduction Through PCA

Pricipal components are still a basic tool for image description used in numerous applications. PCA allows for reduction of number of bands of an image through a linear combination of them. Additionally it allows for considering one single band, which explains the majority of image variability.This step should account for minimizing the loss of important information for the later stages of the methodology.

Let A be an image of size p x q x r pixels, where r is the number of spectral bands. This image stack I is rearranged as a new image A_{vector} of size pq x r pixels, where each column of A_{vector} corresponds to the stacked pixels of the corresponding band of A.

Let us consider v = (v1 , v2 , v3 ,….., vr)T , which is a r – element vector formed by the values of a particular pixel position across the r spectral bands, i.e each line of Avector . Let pv = M{v} as the expected value of v and Qv = M{(v – pv)(v – pv)T} as the respective covariance matrix , let m be a matrix whose rows are formed from the eigenvectors of Qv in such an order that the first row of Z is the eigenvector corresponding to the largest eigenvalue and so on.

The principal component transform

$$C_v = Z\,(v - p_v) \qquad (3)$$

PCA is a widely known method of data reduction.

2.3 Segmentation By Region Growing Technique

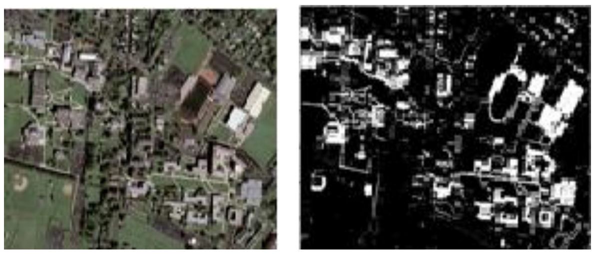
Image segmentation is the process of subdividing an image into its constituent part or objects in an image. It divides the image into nonintersecting regions such that each region is homogeneous and the union of two adjacent regions is not homogeneous. Thus it is possible to analyse the objects to get extra information. Segmentation simplifies the image since it significantly reduces the number of pixel values.

Consider an image which have set of pixels R. Region growing technique divides R into subregions (R_1 , R_2 , R_3 , .R_n). The union of all these regions (R_1 , R_2 , R_3 , ….R_n) UR_i should give the original image R. The region R_i should be connected.

$R_i \cap R_j = \emptyset$ for $i \neq j$ for all $i = 1,2,\ldots.n$
$P(R_i)$ = TRUE for $i = 1,2,\ldots n$
P (R_i U R_j) = FALSE for any adjacent region R_i and R_j.

P (R_i) is a logical predicate defined over the points in set R_i and Ø is a null set.

A large number of segmentation methods can be found in the literature ,but there is no single method which can be considered good for all images, nor are all methods equally good for a particular type of image. The existing methods include gray-level thresholding, iterative pixel classification, surface based segmentation, edge detection and method based on fuzzy set theory. Automatic image segmenataion is the most reliable method since the threshold value for segmentation is automatically selected from the histogram. It is applied to images with a bimodal histogram. It also provides a good result for unimodal or multimodal histograms where a precise delineation of the objects present on the scene is not a requirement. The procedure utilizes only the zeroth- and the first –order cumulative moments of the gray-level histogram. After bilevel thresholding, the image R pixels are assigned as 0 or 1. Fuzzy *c*-means (FCM) and hard *c*-mean(HCM) account for the presence of different classes of objects on an image. However, these fuzzy-based methods require a priori information, in particular the need for a previous definition of the number of clusters.



(a) (b)

Fig 3. (a) original image and (b) segmented image.

2.4 SIFT

The SIFT (Scale Invariant Feature Transform) is an approach for extracting distinctive invariant features from images. It is widely used in image matching. SIFT descriptor detects the extreme points through the whole scale space. Only 20-50 percent of detected keypoints are useful in image matching between two images. Four major stages in SIFT are (1)scale space local extrema detection. (2)accurate keypoints localization.(3)orientation assignment.(4)keypoints descriptor.

First step, the features locations are determined as the local extrema of Difference of Gaussians (DOG pyramid).
Let I (x, y) be an image and S (x, y, σ) the scale space of I, which is defined as

$$S(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \qquad (4)$$

Where * is the convolution operation in x and y and G (x, y, σ) is a variable-scale Gaussian defined as

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2 + y^2)/2\sigma^2} \quad (5)$$

The scale-space extrema detection starts with the detection of local maxima and minima of D(x,y,σ) where

$$D(x,y,\sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$$

$$= S(x, y, k\sigma) - S(x, y, \sigma) \quad (6)$$

The detection is performed by searching over all scales and image locations in order to identify potential interest points that are invariant to scale and orientation.

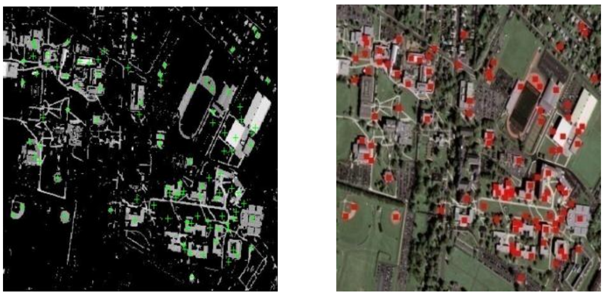
Second step is to accurately localize the keypoints. This is performed by rejecting those keypoints, which have low contrast or are poorly localized along an edge. Unstable extremas are detected by considering a threshold over the extremum of the Taylor expansion of *D(x,y,σ)* .

Once the SIFT- feature location is determined , next step is to assign orientation to each feature based on local image gradient. The gradient's absolute value and direction are given by

$$m(x,y) = \sqrt{(S(x+1,y) - S(x-1,y))^2 + (S(x,y+1) - S(x,y-1))^2} \quad (7)$$

$$\theta(x,y) = \tan^{-1}\left(\frac{S(x,y+1) - S(x,y-1)}{S(x+1,y) - S(x-1,y)}\right) \quad (8)$$

The final stage of the SIFT approach is to build descriptor for each keypoints. The gradient magnitudes and orientations within each feature are computed by and weighted by appropriate Gaussian window, and the coordinate of each pixel and its gradient orientation are rotated relative to the keypoints orientation.

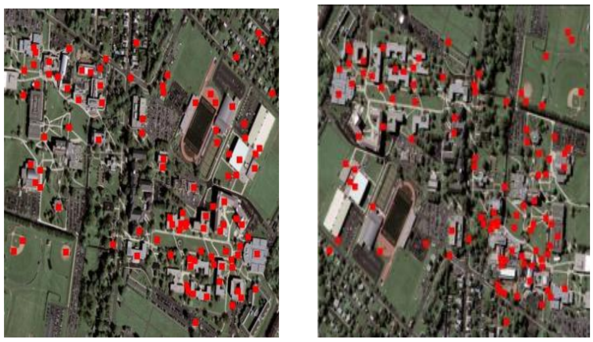


(a) (b)

Fig 4. (a) Segmented image and (b) SIFT features

2.5 Keypoint Matching

The best candidate match for each keypoint is found by identifying its nearest neighbor in the database of keypoints from training images. The nearest neighbor is defined as the keypoint with minimum Euclidean distance for the invariant descriptor. For each keypoint, if the ratio of closest to second-closed neighbor (d_{ratio}) is lower than a certain ratio threshold,the nearest neighbor is a correct match. When this ratio threshold is decreased , the number of matches will be reduced and the matching will be more stable and reliable. Since several distortion effects may be present in an acquired image, it is desirable to have a reference image with as little distortions as possible.



(a) (b)

Fig 5. (a) SIFT features of reference image and (b) SIFT features of input image

2.6. Affine Transformation

In geometry , an affine transformation is a function between affine spaces which preserves the affine structure. In affine coordinates , which include Cartesian coordinates in Euclidean spaces, each output coordinate of an affine map is a linear function of all output coordinates. Any affine transformation is equivalent to a linear transformation followed by a translation. It maps pixel intensity values located at position (x1, y1) in an input image into new pixel position (x0 ,y0) in the reference image. In the finite-dimensional case , if the linear map is represented as a multiplication by a matrix A and the translation as the addition of a vector b, an affine map f acting on a vector x can be represented as

$$\vec{y} = f(\vec{x}) = A\vec{x} + \vec{b} \quad (9)$$

The general affine transformation is given by

$$\begin{vmatrix} x_0 \\ y_0 \end{vmatrix} = A \times \begin{vmatrix} x_1 \\ y_1 \end{vmatrix} + B \quad (10)$$

By defining only the B matrix, this transformation can carry out pure translation:

$$A = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \; B = \begin{vmatrix} b_1 \\ b_0 \end{vmatrix} \quad (11)$$

Pure rotation uses the A matrix and is defined as :

$$A = \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix}, \quad B = \begin{vmatrix} 0 \\ 0 \end{vmatrix} \quad (12)$$

Similarly, pure scaling is :

$$A = \begin{vmatrix} a_{11} & 0 \\ 0 & a_{22} \end{vmatrix}, \quad B = \begin{vmatrix} 0 \\ 0 \end{vmatrix} \quad (13)$$

The affine transformation preserves the collinearity relation between the points, that is point which lie on same line continue to be collinear after the transformation. An affine transformation does not necessarily preserve angles or lengths, though it does preserve ratios of distances between points lying on the straight line. Also, sets of parallel lines remain parallel after an affine transformation.

2.7 Final Set of Tie Points

The final set of tie points is composed by the initial matching candidates. The performance of the proposed methodology for AIR was evaluated through measures which allow for an objective and automatic evaluation of the image registration process quality. These measures includes a) number of redundant points (Nred), regarding the transformation function used on the image registration , it is better to use as much CPs (control points) as possible. b) RMSall , Given a residual (rxi , ryi) for a certain CP, the RMS is computed over the norm (defined as the distance from the origin to its location) of the N residuals

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(rx_i^2 + ry_i^2)} \qquad (14)$$

In order to allow for general comparison , the RMS should be computed over the normalized (to the pixel size) residuals , leading to RMS_{all}.

3. CONCLUSION

In this paper, integration of SIFT algorithm with matching algorithm based on affine transformation is implemented. The feature points improves the robust of registration algorithm greatly. Based on the similarity measure defined, the similarity between every point in the sensed image and a point in the reference image can be easily obtained , which represents the matching degree of geographic points. Based on this similarity measure and integrating with noticeably improvement on feature match process, a new registration approach is constructed. The proposed method is applicable for Satellite and research applications and medical applications.

4. RESULT

In automatic image matching, it is necessary that all stages are automatic, which includes image segmentation phase. However, fully automatic image segmentation is still a present subject of research, in particular for remote sensing images. Nevertheless, in order to provide some sensitivity analysis of the proposed image registration methodology on the segmentation method, the k-means clustering technique was also considered as an alternative to the thresholding method. Despite the fact that it involves some parameters "number of clusters" and "squared euclidean" etc. The application of different segmentation methods than thresholding method is also a valid alternative .

The proposed methodology for image matching was compared with SIFT approach. They were applied to the first principal component of the images . Although SIFT has achieved results comparable to our proposed method using dratio = 0.4, it will results in RMSall = 35.00. But our method clearly outperformed SIFT, in particular for the medium spatial resolution pairs. TABLE I shows the comparison results of two images considered.

The application of PCA allowed for a proper reduction of each image dimension since the first principal component explained in all situations more than 85% of the total variance, without compromising the registration accuracy. The use of remaining principal components does not provide meaningful information for later segmentation and may therefore be discarded in the later stages of the proposed methodology. Other methods of data reduction could have been used for this purpose, such as independent component analysis or the projection pursuit, among the linear projections, or other non linear projection methods such as curvilinear component analysis. However, PCA is still the mostly used method for reducing the number of spectral bands and has provided good results under the scope of the proposed methodology.

(a)

(b)

Fig 6. (a) satellite image and (b) city image.

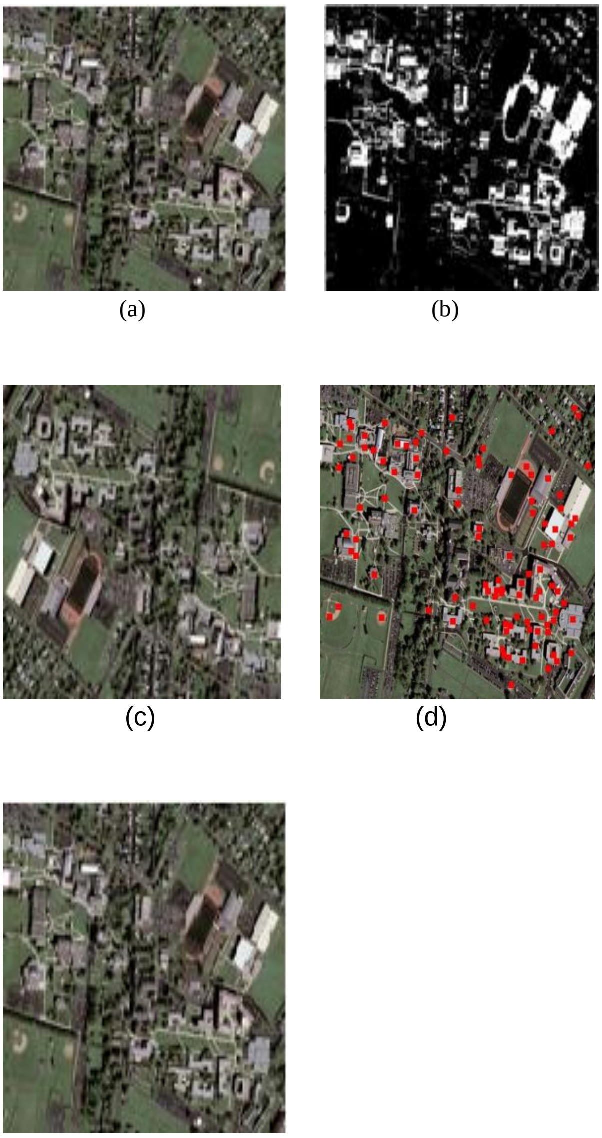
TABLE I

Measures N_{red} and RMS_{all} of images in Fig 4. For given d_{ratio}

PAIR	AIR METHOD	D_{ratio}	N_{red}	RMS_{all}
a	PROPOSED SIFT	0.6 0.3	469 205	.295 .134
b	PROPOSED SIFT	0.9 0.6	15 11	0.467 1.09

5. ANALYSIS

In the proposed method , a fully automatic image registration is done by combining PCA, image segmentation, SIFT and affine transformation. It allows for the registration of a pair of images with different pixel size, translation and rotation effects, and to some extent with different spectral content. Furthermore, it has shown robustness against an automatic choice of the involved parameters, which is a highly desirable characteristic of this class of method.

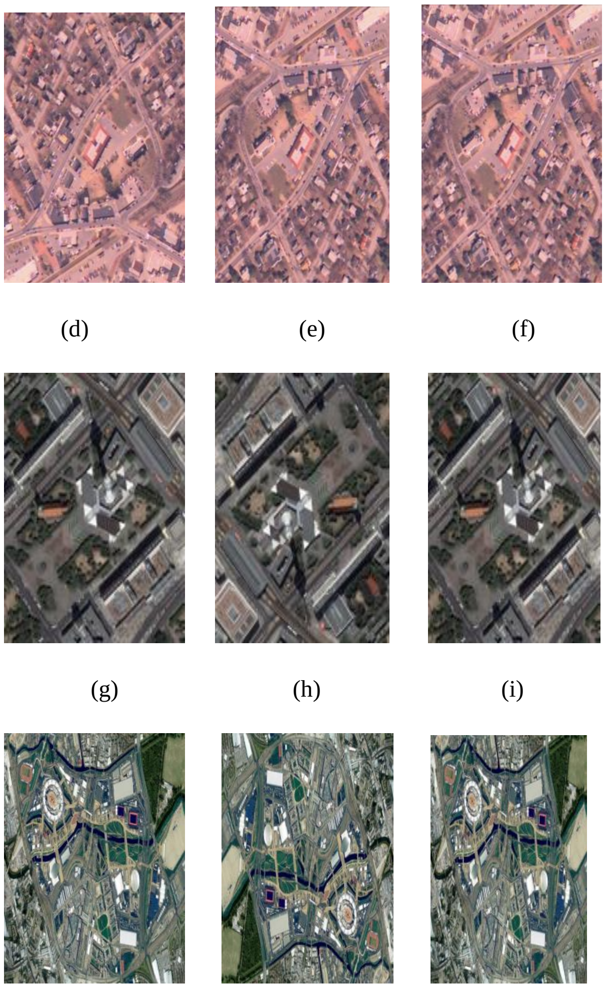


(a) (b) (c) (d) (e)

Fig 7. (a) Reference image. (b) after segmentation (c) Input image. (d) feature detection in the input image (e) input image after registration.

6. REGISTERED IMAGES: A FEW EXAMPLES

(a) (b) (c)



(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

Fig.6 : A few more examples of reference image, input image and registered image
fig (a) reference images, (b) input image, (c)registered image
fig (d) reference images, (e)input image, (f)registered image
fig (g) reference images , (h) input image, (i)registered image
fig (j) reference images , (k) input image, (l)registered image

7. REFERENCES

[1] Hernani Gonclaves, Lui Corte-Real , Jose A. Gonclaves "Automatic image registration through image segmentation and SIFT". IEEE Transaction on *Geoscience and Remote Sensing*., vol. 49, no. 7, pp. july 2011.

[2] J.S. Beis and D.G. Lowe, "Shape Indexing using approximate nearest neighbor search in high-dimensional spaces," in *Proc. Conf. Comput. Vis. Pattern Recog*., Washington,DC,1197, pp.1000-10006

[3] L.G. Brown, Ä survey of image registration techniques,"*Comput. Surv*., vol. 24, no. 4, pp.325-376, Dec. 1992.

[4] H. D. Cheng, X. H. Jiang, Y. Sun, and J. Wang, "Color image segmentation: Advances and prospectus," *Pattern Recog*., vol. 34, no. 12, pp.2259-2281, Dec. 2001

[5] L. Cheng, J. Gong, X. Yang, C. Fan, and P. Han, "Robust affine invariant feature extraction for image matching," *IEEE Geosci. Remote Sens*. Lett., vol. 5, no. 2,pp. 246-250, Apr. 2008

[6] L. M. G. Fonseca and B. S. Manjunath, "Registration techniques for multisensor remotely sensed imagery," Photogram. Eng. Remote Sens., vol.62, no. 9,pp.1049-1056, Sep. 1996.

[7] H. Gonclaves, J. A . Gonclaves, 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, Apr. 2009.

[8] D. G. Lowe, "Object recognition from local scale-invariant features,"in Proc. Int. Conf. Computer Vision, Corfu, Greece, 2008, pp. 1150-1157.

[9] A. Goshtashby, G. C. Stockman, and C. V. Page ."A region based approach to digital image registration with sub-pixel accuracy,"*IEEE Trans. Geosci. Remote Sens*., vol. GRS -24, no. 3, pp. 390-399, May 1986.

[10] L. Journaux , I. Foucherot, and P.Gouton, "Reduction of the number of spectral bands in Landsat images: A comparison of linear and non linear methods, " *Opt. Eng*., vol. 45, no. 6, p. 067002, June 2006.

[11] J. Ma, J. C.–W. Chan, and F. Canters, "Fully automatic subpixel image registration of multiangle CHRIS /Proba data," *IEEE Trans. Geosci. Remote Sens*., vol. 48, no. 7,pp. 2829-2839, Jul. 2010.

[12] J. Inglada and A. Giros, " On the possibility of automatic multisensor image registration,"*IEEE Trans. Geosci . Remote Sens.*, vol 42, no. 10, pp. 2104-2120, Oct. 2004.

[13] N. Otsu," A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. SMCB-9, no. 1, pp.62- 66, Jan . 1979.

[14] J. Ning, L.Zhang, D.Zhang, and C.Wu, "Interactive image segmentation by maximal similarity based region merging, "Pattern Recognit., vol. 43, no. 2, pp. 445-456, Feb. 2010.