International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.2,No.12, 2012
Published online: Dec 25, 2012
DOI: 10.7321/jscse.v2.n12.2

Registration of Remote Sensing Image with Polar Sin Transform-based Descriptor

[1,2*] Meng YI, [2]Bao-long GUO
[1,] *Department of Electrical and Computer Engineering and Center for Automation Research, UMIACS, University of Maryland, College park, MD, 20742, USA, yimeng@umiacs.umd.edu*
[*2,] *School of Electromechanical Engineering, Xidian University, Xi'an 710071, China,blguo1199@163.com*

Abstract. **In this paper a fast registration of remote sensing image with polar sin transform based descriptor is proposed. The main characteristics of the proposed method include: (1) Several new methods are presented to realize fast and accurate interest points extraction under various different scenes, include image scale space construction based on integral image, interest points controlling based on uniform distribution, (2) Through comparing the euclidean distance of the PST (polar sin transform) descriptors defined on the corner neighborhoods, the corresponding matches are established. Experiment results illustrate that the proposed algorithm carries out real-time image registration and is robust to large image translation, scaling and rotation.**

Keywords: **image registration; integral image; uniform distribution; polar sin transform;**

* *Corresponding address:*
Meng YI ,
Department of Electrical and Computer Engineering and Center for Automation Research,
UMIACS, University of Maryland, College park,
MD, 20742, USA
Email: yimeng@umiacs.umd.edu

1. Introduction

Nowadays, image registration has been widely applied in many computer vision domains, including robot localization, remote sensing, medical image analysis, pattern recognition, enhanced visualization and ground moving target detection and tracking [1-3]. As showing good invariant under a certain mathematical transformation and grayscale distortion, Image registration algorithm based on the Local invariant feature descriptor has been draw widely attention in recent years [4-6]. Feature point matching, a critical and challenging process has been widely studied. According to them, there are essentially two parts to any feature-based image registration algorithm: invariant feature detection and invariant Feature matching.

Various methods for detecting control points in an image have been developed. This involves detecting salient features in the two images to be registered. Tinne Tuytelaars et al. [7] has surveyed and compared various point detectors, finding the Harris detector [8] to be the most repeatable. However, the number of control points detected by the Harris algorithm is large, but few accurate correspondences are sufficient to register overlapping images by the projective transformation, so it is necessary to reduce the number of initial interest points. Mikolajczyk [9] proposed the multi-scale Harris with multiple scale selection, however this method is time-consuming. In this paper, we simulate the gaussian convolution using the integral image and rapid construct the image of multi-scale space. Then we detect the corner points by applying scene-adaptation technique. Scene-adaptation controls the number of feature points accurately while ensure the quality of the feature points, thereby reducing the complexity of the feature description and matching.

Corner matching is another important step in corner-based image registration. The most popular methods are distribution-based descriptors and filter-based descriptors (SIFT [10], GLOH [11] and PCA-SIFT [12], Steerable filters [13]), which extract distinctive feature that are invariant to noise and viewpoint. In recent years, these methods are successfully applied in remote sensing sequences registration owing to its good characteristics of being invariant to image scaling and rotation. However, Firstly, these methods are time-

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.2,No.12, 2012
Published online: Dec 25, 2012
DOI: 10.7321/jscse.v2.n12.2

consuming, which means that it will takes longer time than others matching step; Secondly, due to the noise and occlusion, some matching points will be more accurate than others [15]. In this study, polar sin transform is adopted [14]. Compared with other methods, this algorithm can distinguish more accurate matching points from the less accurate ones and then register the images as accurately as possible.

The flowchart of the proposed algorithm framework is shown in Fig.1. This paper is structured as follows. In section 2, the proposed registration algorithm is describes in detail. Section 3 presented and evaluated the experimental results of the performance of the algorithm. Section 4 concludes.

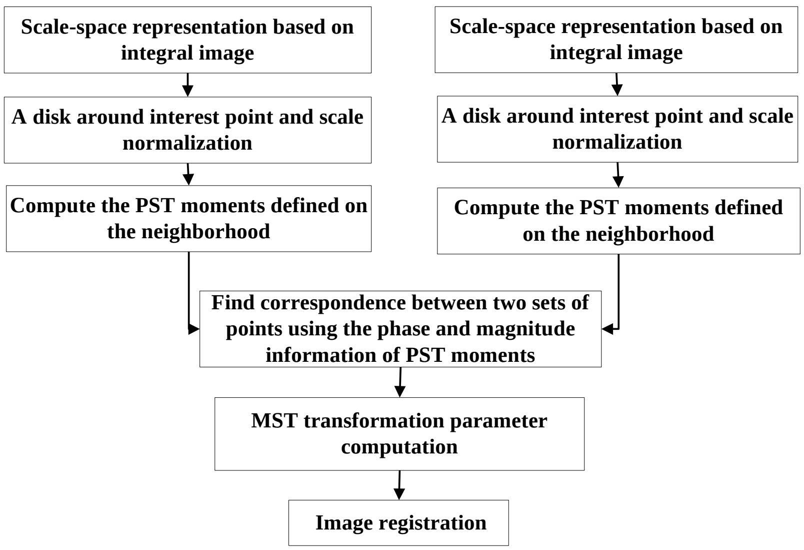


Figure 1. Flowchart of the proposed image registration algorithm

2. Aproach

Knowing that two overlapping images are related by the affine transformation, a 2-D model can well trade off the accuracy and computational complexity for image registration.

Assuming (x, y) represents a point in the base image and (X, Y) represents the same point in the image overlapping the base image, the projective transformation between corresponding points in the images can be written as:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = s\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \tag{1}$$

Where s represent the scale factor, θ is the rotation angle, $(\Delta x, \Delta y)$ represent the translation of the horizontal and vertical direction. When the rotation angle is less than $5°$, then $\cos\theta \approx 1$ 、 $\sin\theta \approx \theta$. Therefore, the problem of the above projective transformation model is transformed into following equations.

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.2,No.12, 2012
Published online: Dec 25, 2012
DOI: 10.7321/jscse.v2.n12.2

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} 1 & -\theta \\ \theta & 1 \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix} \quad (2)$$

Having the coordinates of 4 corresponding points in the images, the unknown parameters of the transformation can be determined by substituting the corresponding points into (2) and solving the obtained system of 4 linear equations.

2.1 Feature Points Extraction

2.1.1 Scale-Space Representation based on Integral Image

Feature detection algorithm use a number of invariant features in multi-scale, therefore, the scale-space needs to be offered before feature detection. In the most general case, convolution with Gaussian kernels and their derivatives form a class of low-level operators for scale-space representation. However, the time to compute the gaussian convolution will increase dramatically.

According to theorem of the central limit, continue convoluting with the function itself will ultimately generate a gaussian function. Therefore, rather than convolve the image with a gaussian function, this article constructs the image of multi-scale space combining with the rectangular convolution kernel function and down-sampling, then we speed up the calculation using the integral map.

A two-dimensional image, in different scales of the scale-space, is given by

$$L(x,y,\sigma,i) = \downarrow M(x,y,\sigma) * L(x,y,i-1) \quad (3)$$

Where L is the Scale space of image, i is the scale, σ is the radius of the window width of the average filter.

In order to speed up the computation time, we compute the integral image $S(x,y,i)$, which corresponds to the $L(x,y,i)$, is defined as follows

$$\begin{aligned} L(x,y,\sigma,i) = \downarrow \frac{1}{(2\sigma+1)^2}&[S(x-\sigma,y-\sigma) \\ &+S(x+\sigma,y+\sigma)-S(x+\sigma,y+\sigma) \\ &--S(x-\sigma,y+\sigma)] \end{aligned} \quad (4)$$

The order O of the scale-space is determined by

$$O = \min_i \left\{ i \,\middle|\, \frac{\min(W,H)}{R^i} \le s \right\} - 1 \quad (5)$$

Where s is the minimum size after the sub-sampling.

2.1.2 Invariant Interest Points Selection

The feature that is invariant under difference in scaling and view angle of the cameras is important for successful matching. In this study, we choose the multi-scale Harris detector. Firstly, we build the scale-space using the integral image [section 2.1.1]; then, we choose to detect corners using a multi-scale version of the Harris detector, which is based on the second-moment matrix and defined as

$$M(x,y,i) = \sum_{x,y} \omega(x,y) \begin{bmatrix} L(x,y,i)_x^2 & L(x,y,i)_x L(x,y,i)_y \\ L(x,y,i)_x L(x,y,i)_y & L(x,y,i)_y^2 \end{bmatrix} \quad (6)$$

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.2,No.12, 2012
Published online: Dec 25, 2012
DOI: 10.7321/jscse.v2.n12.2

Where $M(x,y,i)$ is the auto-correlation function in scale i, $L(x,y,i)_a$ is the derivative computed in a direction. $\omega(x,y)$ is the Gaussian window.

An interest point or a corner point will have a μ with two positive eigenvalues λ_1、 λ_2 of $M(x,y,i)$.

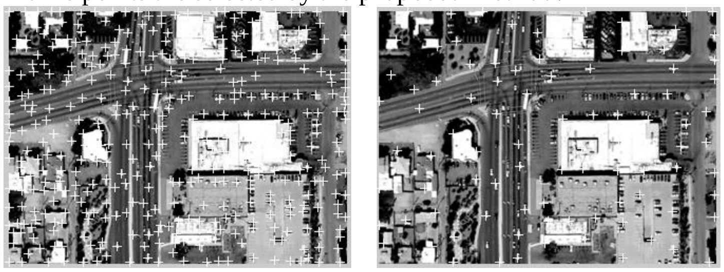
The corner response can be written as

$$R(x,y,i) = \det M(x,y,i) - \lambda \cdot (traceM(x,y,i))^2 \tag{7}$$

Where $\det M = \lambda_1 \lambda_2$, $traceM = \lambda_1 + \lambda_2$.

On the basis of the multi-scale Harris detector, the method of detecting feature points is as follows. The reference image is divided into $s \times t$ regions R_k. In each R_k, the point with the maximum response R is choose as the Interest Points.

Fig.2 shows the Harris points are selected by the proposed method.



(a) (b)

Figure 2. (a) feature points detect using multi-scale Harris detector (b) interest points detected using proposed method.

2.2 Correspondence between Points

2.2.1 Polar Sine Transform

PST is defined as the projections of $f(x,y)$ into polar coordinate representation, which is denoted by $f(\rho,\theta)$. PST of order n with repetition m for a continous image function $f(x,y)$ over a unit disk is

$$A_{nm} = \frac{2}{\pi}\int_0^{2\pi}\int_0^1 [V_{nm}(\rho,\theta)]^* f(\rho,\theta)\rho d\rho d\theta \tag{8}$$

where $[]^*$ denotes the complex conjugate, the complete set of PST polynomials V_{nm} is defined as

$$V_{nm}(\rho,\theta) = \sin(\pi n \rho^2) \cdot e^{im\theta} \tag{9}$$

The following indicates the orthogonal property of $V_{nm}(\rho,\theta)$:

$$\int_0^{2\pi}\int_0^1 V_{nm}(\rho,\theta)[V_{nm}(\rho,\theta)]^* \rho d\rho d\theta = \frac{\pi}{2}\delta_{nn'}\delta_{ll'} \tag{10}$$

Where δ is the Kronecker delta and satisfies

$$\delta_{nn'} = \begin{cases} 1, & n = n' \\ 0, & n \neq n' \end{cases} \tag{11}$$

Let the PST moments of a reference image and its rotated version be $A_{nm}^{'}$, A_{nm}, respectively. Then it is well known that

$$A_{nm}^{'} = A_{nm} \cdot e^{-im\phi} \tag{12}$$

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.2,No.12, 2012
Published online: Dec 25, 2012
DOI: **10.7321/jscse.v2.n12.2**

Where $\phi \in [0,2\pi]$ is the rotation angle. Therefore, the magnitudes of PST moments of the two images are the same, i.e., $\left|A'_{nm}\right| = \left|A_{nm}\right|$. Due to the property of orthogonality, the reconstruction of the pattern can be simply expressed as the sum of every PST basis functions weighted by the corresponding moments:

$$f(\rho,\theta) = \sum_{n=1}^{\infty} \sum_{m=-\infty}^{\infty} A_{nm} V_{nm}(\rho,\theta) \tag{13}$$

2.2.2 Similarity Measure Based on PST Features

Our implementation uses the first 25($n \le 8, m \ge 0$) coefficients in the PST expansion of a disk (the radius proportional to the scale) around all feature points. Rotation in the original image affects its PSTs phase coefficients. To provide rotation-invariant PSTs coefficients that still represent the original shape, we present a method to estimate the rotation angle α firstly, which is implemented in the continuous angle space rather than in the discrete space. Then searching for similar images during matching processing combines both magnitude and angle as a similarity measure.

The formulation (12) can be rewrited as following:

$$\phi = -\frac{1}{l} \arg \frac{A'_{nm}}{A_{nm}} \tag{14}$$

Since $n = 1,2,...,N$, $m = 1,2,...,L$,there are $\sum_{n=1}^{N} \left(\lfloor (L-m)/2 \rfloor + 1\right)$ ways to compute the rotation angle α . A more robust estimation is to weight the estimated angles by the individual magnitude $\left|A_{nm}\right|$.

If we have estimated the rotation angle α between $I^r(x,y)$ and $I^t(x,y)$, then $\left|\phi_{nm} - m\overline{\alpha}\right|$, which denotes the absolute phase difference between the two image regions after the rotation alignment, is equal to 0 ; otherwise, $\left|\phi_{nm} - m\overline{\alpha}\right|$ is a nonzero value in the interval $(0,2\pi)$. Then we can get the normalized angle-based distance D_{ang} between the phases of two PSTs vectors is defined as

$$D_{ANG} = \sum_{n} \sum_{m} \frac{\min\left\{\left|\phi_{nm} - m\overline{\alpha}\right|, 2\pi - \left|\phi_{nm} - m\overline{\alpha}\right|\right\}}{\pi} \tag{15}$$

Besides, the normalized magnitude-based distance D_{mag} between the magnitudes of two PSTs is then defined as

$$D_{MAG} = \sum_{n} \sum_{m} \frac{\left|Z^r_{nm}\right| + \left|Z_{nm}\right|}{\sum_{n,m} \left(\left|Z^r_{nm}\right| + \left|Z_{nm}\right|\right)} \tag{16}$$

From (15) and (16), it is obvious that two distances D_{ang} and D_{mag} are both within the range $[0,1]$, they can thus be combined together. Aside from this, we assign weighting factors to both distances to indicate their respective importance in the overall distance D

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.2,No.12, 2012
DOI: **10.7321/jscse.v2.n12.2**
Published online: Dec 25, 2012

$$D = w_1 D_{ANG} + w_2 D_{MAG} \tag{17}$$

Where w_1 and w_2 are the weighting factors for D_{ANG} and D_{MAG}, respectively, and satisfy the relation $w_1 + w_2 = 1$.

Once the best feature correspondence has been established, we can determine the eight unknown parameters of the transformation by substituting the coordinates of the four corresponding points into (2). To reduce the influence of outliers in the accuracy of the solution, we use RANSAC[16], which can finding the best model parameters by using random subsets from the set of correspondences instead of using the entire set. The model parameters estimated from the subset which gives the least sum of squared error is then taken as the best fit.

3. Experimental Results

This section presents some examples and quantitative results of the proposed registration algorithm for remote sensing image. The algorithm has been implemented in C++ and all experiments have been carried out on DELL Intel Xeon E5410 2.33-GHz desktop computer with 9GB of RAM, with Windows 7 Enterprise Professional Edition. Fig.3 shows 8 sets of the unmanned aerial vehicle(UAV) video sequences that come from the predator data of VSAM at Carnegie Mellon University (Fig.3,No.1-No.5)and our aerial video data(Fig.3,No.6-No.8), with size 320×240, including rural roads, fields and urban buildings, the image sequence have angle rotation, scale zoom, brightness mutation and view changes, Etc.

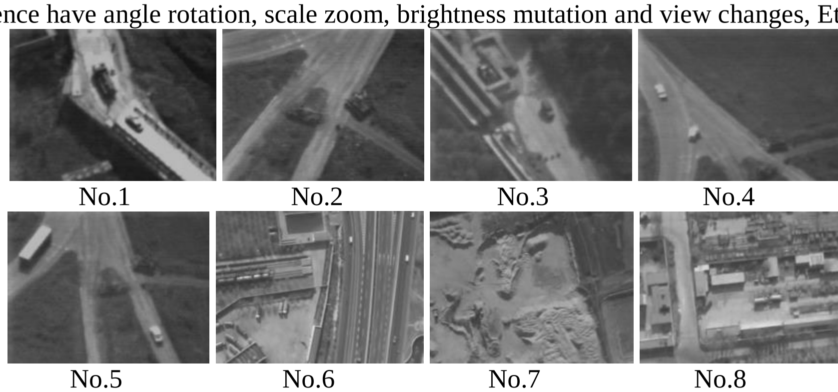


Figure 3. UAV video sequences

In order to evaluate the registration algorithm, Fig.4 shows a part of the feature matching example, including scene changes(Fig.4.(a)-(b)) and geometric changes(Fig.4.(c)-(f)).Table1 shows more detail test results of the matching and processing speed in different scenes. From the table it is obvious that the proposed algorithm are able to find sufficient number of matching points in various scenarios and achieve satisfactory results. In processing speed, the proposed algorithm does not change the computational complexity when the scene changes. The average computing time of the test sequence is 46.01ms, the average processing speed is 27.07fps.

Table 1. Registration Results of Image Sets

Number	frame	Feature points	matching points	Time(ms)	frame rate(fps)
No.1	876	200	74.56	66.11	29.65
No.2	346	200	65.73	43.89	26.51
No.3	875	200	61.34	59.15	26.82
No.4	762	200	16.12	51.33	31.89
No.5	369	200	57.54	41.88	29.9
No.6	1283	200	32.55	32.56	24.8
No.7	542	200	69.64	41.21	21.1
No.8	100	200	29.65	31.99	25.9
Average	566	200	50.89	46.01	27.07

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.2,No.12, 2012
Published online: Dec 25, 2012
DOI: **10.7321/jscse.v2.n12.2**

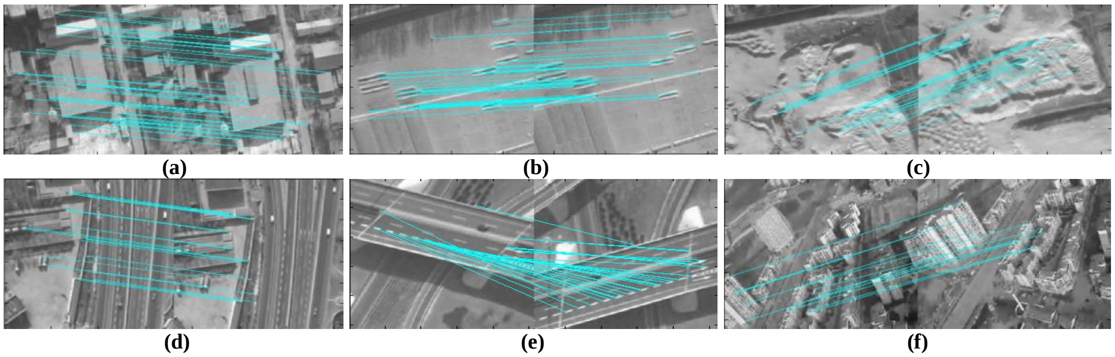


Figure 4. Feature matching results (a) Complex scene (b) Simple scene (c)-(d) Translation (e)-(f) Rotation

We use *recall vs.1-precision* graphs to evaluate the performance of the feature matching. Recall is the ratio of the number of correct matches to the number of corresponding region pairs. Precision is the ratio of the number of correct matches to the total number of correct and false matches. These graphs are generated as follows: the feature points for all the images in the dataset are detected; all pairs of key-points from different images are examined. If the distance measure between the feature vectors for the particular pair of feature points does not exceed the distance threshold t, this region pair is called a match. We can determine the recall and 1-precision as follow:

$$recall = \frac{\#\text{correct matches}}{\#\text{correspondences}}$$

$$1 - precision = \frac{\#\text{false matches}}{\#\text{correct matches}+\#\text{false matches}}$$

We present and discuss the experimental results of the feature matching evaluation. The performance is compared for image noise, viewpoints, scale and rotation changes. A sequence of remote sensing images was used, as shown in Fig.3, No.6. The results of these tests are shown in Figs.5. We can observe that our algorithm obtains the best matching score.

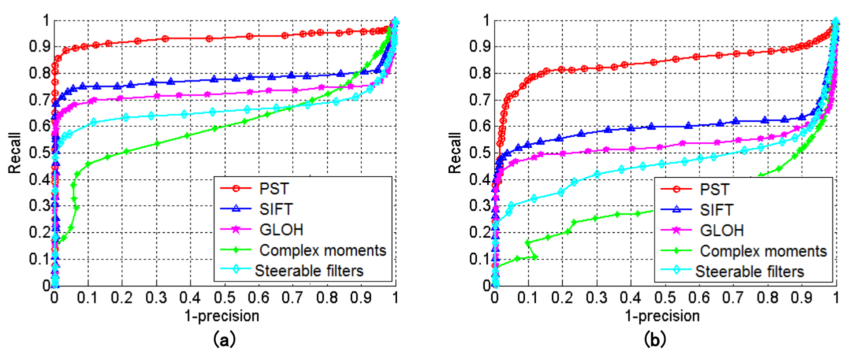


Figure.5 precision-recall curves performance evaluations under different photometric and geometric transformations. (a) noise (b) viewpoints

4. Conclusions

A special class of image registration methods is considered dealing with alignment of overlapping remote sensing images under small differences in gain, zoom level, and view angle changes of the cameras. The main contribution of this paper is that we presented a new integral image and polar sin transform based aerial video registration method. The proposed method composed four steps:1)represents scale-space based on integral image, 2)finds stable interest points based on uniform distribution, 3)finds the correspondence between the feature points of the input image and reference image based on PST descriptors, 4)estimates the parameters of a projective transformation mapping the input image to the reference one. Experimental results

International Journal of
Soft Computing And Software Engineering (JSCSE)
e-ISSN: 2251-7545
Vol.2,No.12, 2012
DOI: 10.7321/jscse.v2.n12.2
Published online: Dec 25, 2012

show that the proposed method carries out real-time aerial video registration under complex environments with change of scene, and achieve precision registration.

5. 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] S. Ali, V. Reilly and M. Shah. Sneddon, "Motion and appearance contexts for tracking and reacquiring targets in aerial videos,". In IEEE CVPR,pp:1-6. 2007.

[3] T. Parag, A. Elgammal, and A. Mittal, "A framework for feature selection for background subtraction," Proc. Computer Vision andPattern Recognition, pp. 1916–1923, 2006.

[4] E. Jayabalan, A. Krishnan, "Target Tracking in Aerial Videos Using Image Registration Techniques". Communications in Computer and Information Science, 142(3), pp:475-477, 2011.

[5] Clark F. Olson, Adnan I. Ansar, and Curtis W. Padgett, "Robust registration of aerial image sequences," Lecture Notes in Computer Science,pp:325-334, 2009.

[6] Zen Chen, Shu-Kuo Sun, "A zernike moment phase-based descriptor for local image representation and matching,", Transaction on image processing,vol.19, no.1 , pp:205-219. 2010.

[7] T. Tuytelaars, K. Mikolajczyk, "Local Invariant Feature Detectors: A Survey. Foundations and Trends in Computer Graphics and Vision", Vol.3, no.3, pp: 177-280, 2007.

[8] C. Harris and M. K. Stephens, "A combined corner and edge detector," Alvey Vision Conference, pp:147-152, 1988.

[9] K. Mikolajczyk and C. Schmid, "An affine invariant interest point detector," in Proc. of European Conference on Computer Vision, ECCV, pp. 128–142, 2002.

[10] D.G.Lowe, "Distinctive image features from scale-invariant keypoints". Int.J.Comput.Vis., vol.15,no.6, pp:415-434,1997.

[11] K.Mikolajczyk and C.Schmid, "A Performance evaluation of local descriptors", IEEE Trans. Pattern Anal. Mach. Intell., vol.27,no.10,pp.1615-1630,Oct.2005.

[12] Y.Ke and R.Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors", in proc. IEEE Int. Conf. Computer Vision and Pattern Recognition ,vol.2, pp.506-513, 2004.

[13] W.Freeman and E.Adelson, "The design and use of steerable filters,"IEEE Trans.Pattern Anal.Mach. Intell,vol.71,pp.891-906,Sep.1991.

[14] Pew-Thian Yap, Xudong Jiang, and Alex Chichung kot. "Two-Dimensional Polar Harmonic Transforms for Invariant Image Representation". IEEE Transactions on pattern analysis and machine intelligence, Vol.32,no.7, pp: 1259-1270, 2010.

[15] Karthik Krish, Stuart Heinrich, Wesley E.Snyder, "Global registration of overlapping images using accumulative image features", Pattern Recognition Letters, vol.31, no. 2, pp.112-118,2010.

[16] M.A.Fischler and R.C.Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analsis and automated cartography," Commun. ACM, vol.24,no.6, pp.381-395, Jun.1981.