Image Registration using NSCT and Invariant Moment

Jignesh Sarvaiya jns@eced.svnit.ac.in
Assistant Professor, ECED,
S V NATIONAL INSTITUTE OF TECH.
SURAT,395007,INDIA

Suprava Patnaik ssp@eced.svnit.ac.in
Professor, ECED,
S V NATIONAL INSTITUTE OF TECH.
SURAT,395007,INDIA

Hemant Goklani hsgoklani@rediffmail.com
PG Student, ECED,
S V NATIONAL INSTITUTE OF TECH.
SURAT,395007,INDIA

Abstract

Image registration is a process of matching images, which are taken at different times, from different sensors or from different view points. It is an important step for a great variety of applications such as computer vision, stereo navigation, medical image analysis, pattern recognition and watermarking applications. In this paper an improved feature point selection and matching technique for image registration is proposed. This technique is based on the ability of Nonsubsampled Contourlet Transform (NSCT) to extract significant features irrespective of feature orientation. Then the correspondence between the extracted feature points of reference image and sensed image is achieved using Zernike moments. Feature point pairs are used for estimating the transformation parameters mapping the sensed image to the reference image. Experimental results illustrate the registration accuracy over a wide range for panning and zooming movement and also the robustness of the proposed algorithm to noise. Apart from image registration proposed method can be used for shape matching and object classification.

Keywords: Image Registration, NSCT, Contourlet Transform, Zernike Moment.

1. INTRODUCTION

Image registration is a fundamental task in image processing used to align two different images. Given two images to be registered, image registration estimates the parameters of the geometrical transformation model that maps the sensed images back to its reference image [1].

In all cases of image registration, the main and required goal is to design a robust algorithm that would perform automatic image registration. However, because of diversity in how the images acquired, their contents and purpose of their alignment, it is almost impossible to design universal method for image registration that fulfill all requirements and suits all types of applications [2][16].

Many of the image registration techniques have been proposed and reviewed [1], [2] [3]. Image registration techniques can be generally classified in two categories [15]. The first category utilizes image intensity to estimate the parameters of a transformation between two images using an approach involving all pixels of the image. In second category a set of feature points extracted from an image and utilizes only these feature points instead of all whole image pixels to obtain the transformation parameters. In this paper, a new algorithm for image registration is proposed. The proposed algorithm is based on three main steps, feature extraction, correspondence between feature points and transformation parameter estimation.

The proposed algorithm utilizes the new approach, which exploits a nonsubsampled directional multiresolution image representation, called nonsubsampled contourlet transform (NSCT), to extract significant image features from reference and sensed image, across spatial and directional resolutions and make two sets of extracted feature points for both images. Like wavelet transform contourlet transform has multi-scale timescale localization properties. In addition to that it also has the ability to capture high degree of directionality and anisotropy. Due to its rich set of basis functions, contourlet can represent a smooth contour with fewer coefficients in comparison to wavelets. Significant points on the obtained contour are then considered as feature points for matching. Next step of correspondence between extracted feature points is performed using Zernike moment-based similarity measure. This correspondence is evaluated using a circular neighborhood centered on each feature point. Among various types of moments available, Zernike moments is superior in terms of their orthogonality, rotation invariance, low sensitivity to image noise [3], fast computation and ability to provide faithful image representation [4]. Then after transformation parameters required to transform the sensed image into its reference image by transformation estimation by solving least square minimization problem using the positions of the two sets of feature points. Experimental results show that the proposed image registration algorithm leads to acceptable registration accuracy and robustness against several image deformations and image processing operations.

The rest of this paper is organized as follows. In section 2 the basic theory of NSCT is discussed. In section 3 the proposed algorithm is described in detail. In section 4 experimental results of the performance of the algorithm are presented and evaluated. Finally, conclusions with a discussion are given section 5.

2. NONSUBSAMPLED CONTOURLET TRANSFORM (NSCT)

It is observed that wavelets are frequently used for image decomposition. But due to its limited ability in two dimensions to capture directional information and curve discontinuity, wavelets are not the best selection for representing natural images. To overcome such limitation, multiscale and directional representations that can capture the intrinsic geometrical structures have been considered recently. The contourlet transform is a new efficient image decomposition scheme, introduced by Do and Vetterli [6] which provides sparse representation at both spatial and directional resolutions. The contourlet transform employs Laplacian pyramids to achieve multiresolution decomposition and directional filter banks to yield directional decomposition, such that, the image is represented as a set of directional subbands at multiple scales [11] [12]. One can decompose the representation at any scale into any power of two's number of directions with filter blocks of multiple aspect ratios. Thus scale and directional decomposition become independent of each other and different scales can further decomposed to have different number of directional representation. This makes the whole analysis more accurate involving less approximation.

The contourlet transform is not shift-invariant. When associated with down sampling and upsampling shifting of input signal samples causes Pseudo-Gibbs phenomena around singularities. However the property of shift invariance is desired for image analysis applications like image registration and texture classification that involve edge detection, corner detection, contour characterization etc. One step ahead of the contourlet transform is proposed by Cunha et al. [7] [10] is Nonsubsampled Contourlet Transform (NSCT), which in nothing but shift invariant version of contourlet transform. To obtain the shift invariance the NSCT is built upon iterated nonsubsampled filter banks.

The construction design of NSCT is based on the nonsubsampled pyramid structure (NSP) that ensures the multiscale property and nonsubsampled directional filter banks (NSDFB) that gives directionality [13]. Fig. 1 (a) illustrates an overview of the NSCT. The structure consists in a bank of filters that splits the 2-D frequency plane in the subbands illustrated in Fig. 1 (b).

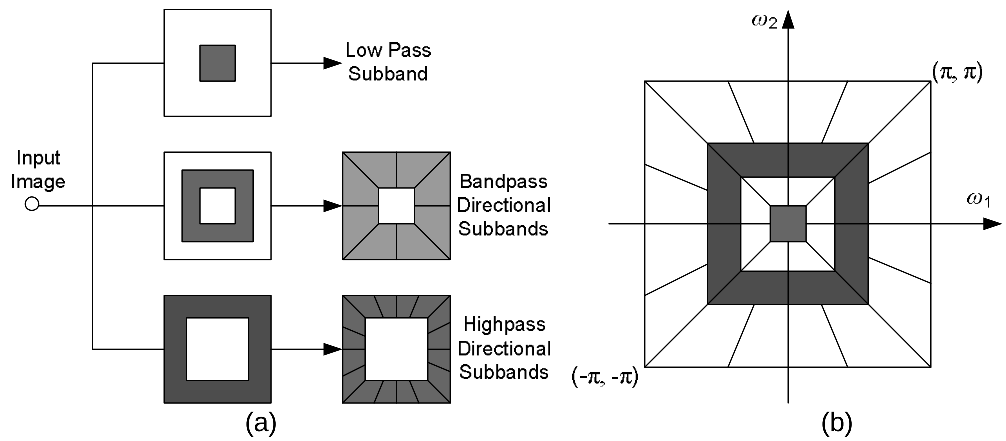


Figure 1: Nonsubsampled Contourlet Transform (a) NSFB structure that implements the NSCT. (b) Idealized frequency partitioning obtained with the proposed structure [7].

2.1 Nonsubsampled Pyramid Structure (NSP)

The NSP is a shift invariant filtering structure, used to obtain the multiscale property of the NSCT. It gives sub band decomposition similar to that of the Laplacian pyramid. Only difference is that here two channel nonsubsampled 2-D filter banks were used. The perfect reconstruction condition is obtained provided the filter satisfy the *Bezout identity*.

$$H_0(z)G_0(z)+H_1(z)G_1(z)=1$$

To achieve multiscale decomposition and construct nonsubsampled pyramids by iterated filter banks as shown in Fig.2, filter $H_0(z)$ is upsampled by 2 and the realization generates $H_0(z^2)$ in both directions. Thus the perfect reconstruction condition is satisfied at each level. Fig.3 illustrates the NSP decomposition with 3 stages.

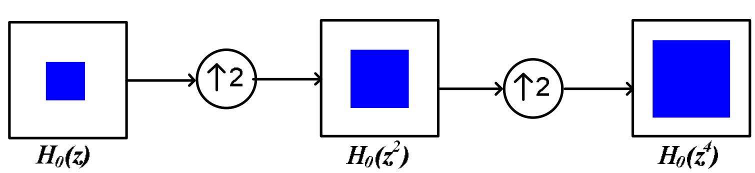
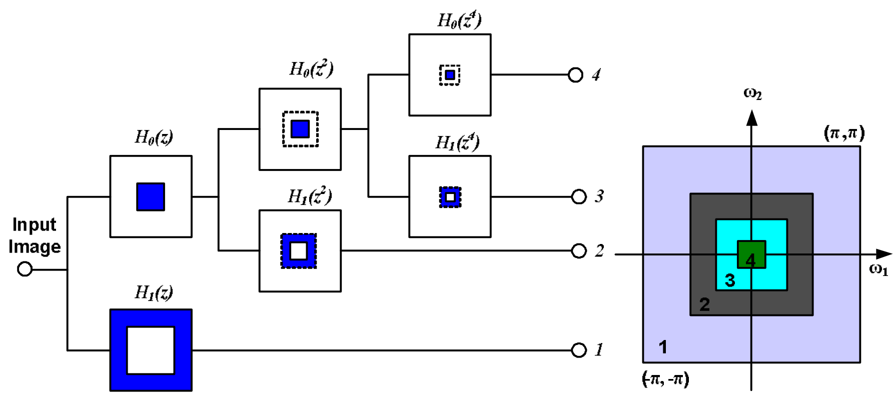


Figure 2: Multiscale decomposition and construct nonsubsampled pyramids by iterated filter banks



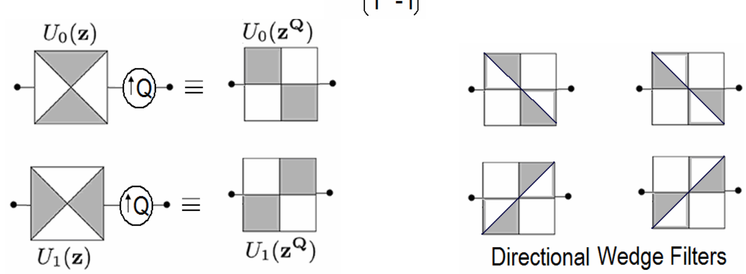
(a) (b)

Figure 3: Nonsubsampled pyramid is a 2-D multiresolution expansion. (a) Three stage pyramid decomposition. (b) Sub bands on the 2-D frequency plane [7].

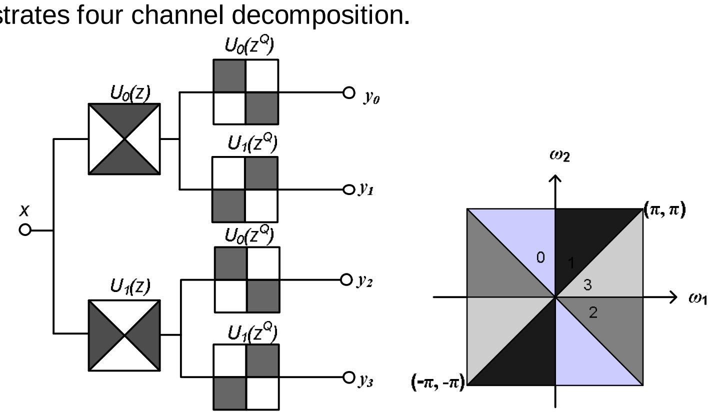
2.2 Nonsubsampled Directional Filter Banks (NSDFB)

A shift invariant directional expansion is obtained with NSDFB. The NSDFB is constructed by eliminating down samplers and up samplers in the DFB [13]. This is done by switching off the downsamplers/upsamplers in each two channel filter bank in the DFB tree structure and up sampling the filters accordingly. Like NSP the NSDFB is constructed by eliminating the downsamplers in the FB tree structure and upsampling the filters accordingly. Filters are upsampled by quincunx matrix which adds rotation and is given by

$$Q=\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$



More directional resolutions are obtained at higher scales by combination of NSP filters and NSDFB to produce wedge like subbands. The result is a tree structured filter bank that splits the 2-D frequency plane into directional wedges [8]. This result in a tree composed of two-channel NSFBs. Fig. 4 illustrates four channel decomposition.



(a) (b)

Figure 4: Four channel nonsubsampled directional filter bank constructed with two channel fan filter banks. (a) Filtering structure. (b) Corresponding frequency decomposition [7].

2.3 Combining the NSP and NSDFB in the NSCT

The NSCT is constructed by combining the NSP and NSDFB. NSP provide multiscale decomposition and NSDFB provide directional decomposition. This scheme can be iterated repeatedly on the low pass sub band outputs of NSPs. From the above mentioned theoretical statements, we say that the NSCT is a fully shift invariant, multiscale, multidirectional expansion that has a fast implementation. The primary motivation for this work is to determine effectiveness of the NSCT in extracting feature points for image registration.

3. The Proposed Registration Algorithm

The NSCT not only provide multiresolution analysis, but it also provides geometric and directional representation and it is shift invariant such that each pixel of the original image in the same location, we can therefore able to gather the geometric information pixel by pixel from NSCT coefficients [7] [10].

In this section, the proposed registration algorithm is presented in detail. We take two images to be aligned. An image without distortions considered as reference image or base image another image with deformations considered as sensed image or distorted image or input image.

The problem of image registration is actually estimation of the transformation parameters using the reference and sensed images. The transformation parameter estimation approach used in this paper is based on feature points extracted from reference image **I** and sensed image **I'**, which is geometrically distorted. The proposed registration process is carried out in three main steps. Feature points extraction, finding correspondence between feature points and transformation parameters estimation. This can be explained in detail as follows:

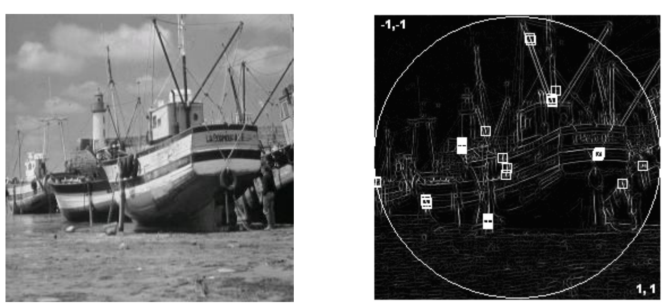
3.1. NSCT based Feature Points Extraction method

The proposed method can automatically extract feature points from both images [13] [14]. This method can be summarized by following algorithm:

(i) Compute the NSCT coefficients of reference image and sensed image for N levels and L directional subbands.

(ii) At each pixel, compute the maximum magnitude of all directional subbands at a specific level. We call this frame “maxima of the NSCT coefficients”.

(iii) A thresholding procedure is then applied on the NSCT maxima image in order to eliminate non significant feature points. A feature point is considered only if NSCT maxima > Th; where $Th_j = C(\sigma_j+\mu_j)$, where C is a user defined parameter, σ_j is standard deviation and μ_j is mean of the NSCT maxima image at a specific level 2^j. The locations of the obtained thresholded NSCT maxima P_i (i = 1, 2 . . ., K) are taken as the extracted feature points, where $P_i = (x_i, y_i)$ is the coordinates of a point Pi and K is the number of feature points. An example of the feature points detected from reference image is illustrated in Fig. 5.



(a) (b)

Figure 5: Feature point extraction: (a) Reference image (b) NSCT maxima image marked by extracted 35 feature points when N is 2.

Initially number of levels taken are 2. But for extraction of robust feature points, necessary for large geometrical deformations, we need to increase the N level in the proposed algorithm.

3.2. Feature point matching using Zernike moment

After the strong feature points extracted from reference and sensed images, a correspondence mechanism is required between these two feature point sets. This correspondence mechanism fulfils the requirement of pairing the feature point of reference image with its correspondent one in the sensed image. In this proposed algorithm, Zernike moment based similarity measure approach is used to establish the correspondence between the two images. This correspondence is evaluated using a circular neighbourhood cantered on each and every feature point. Zernike moments possess a rotation invariance property [4] [9]. Rotating the image does not change the magnitude of its Zernike moment. This is the main and strong reason for selecting Zernike moments as feature descriptors. We can also achieve scale and translation invariant feature points by applying normalization process to regular geometrical moments. Thus we can say that the Zernike moment magnitude remains same after rotation. The correspondence between the feature point sets obtained from reference and sensed image is obtained as follows:

[i] For every extracted feature point P_i, select a circular neighbourhood of radius R centred at this point and construct a Zernike moments descriptor vector P_z as

$$P_z = \left(|Z_{1,1}|,....,|Z_{p,q}|,|Z_{10,10}|\right) \tag{1}$$

Where | $Z_{p,q}$ | is the magnitude of Zernike moments of a nonnegative integer of order p, where p-|q| is even and |q|≤ p. While higher order moments carry fine details of the image, they are more sensitive to noise than lower order moments [5]. Therefore the highest order used in this algorithm is selected to achieve compromise between noise sensitivity and the information content of the moments. The Zernike moments of order *p* are defined as

$$Z_{pq} = \frac{(p+1)}{\pi} V_{pq}^{*}(r,q)\ \mathrm{A}(x,y) \tag{2}$$

Where, $x^2+y^2 \le 1$, $r = (x^2+y^2)^{1/2}$, $\theta=\tan^{-1}(y/x)$. '*x*' and '*y*' are normalised pixel location in the range -1 to +1, lying on an image size grid. Accordingly radius *r* can have maximum value one. Fig 4(b) shows the unit radius circle along with the significant feature points for the reference image. In the above equation V_{pq}^{*} denotes the Zernike polynomial of order *p* and repetition *q*. It can be defined as

$$\mathrm{V}_{pq}(r,q) = R_{pq}(r)\mathrm{e}^{iq\theta} \tag{3}$$

Where $R_{pq}$the is real-valued polynomial of radius *r* given by

$$R_{pq}(r) = \sum_{s=0}^{(p-|q|/2)} (-1)^s \frac{(p-s)!}{s!\left(\frac{p-2s+|q|}{2}\right)!\left(\frac{p-2s-|q|}{2}\right)!} r^{p-2s} \tag{4}$$

R_{pq} depends on the distance of the feature point from the image centre. Hence the proposed method has limitation to work well for rotations about image axis passing through the image centre. Fig. 6 illustrates the two images, reference image and sensed image with 60 degree rotation about the central image axis. Zernike moment vector magnitude for a feature pair is shown in Table 1.

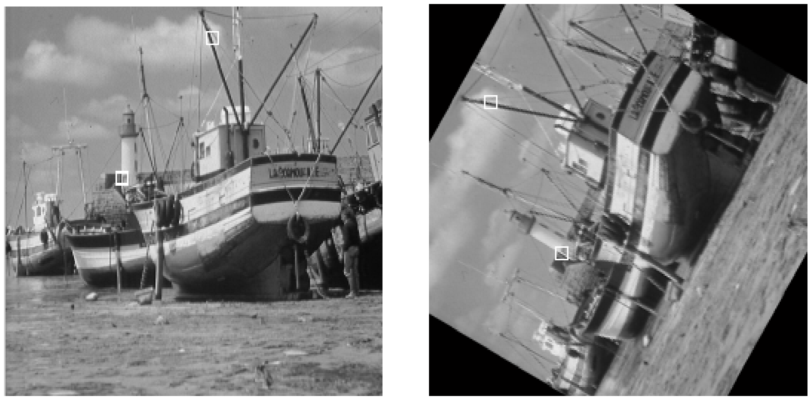


Figure 6: Correspondence between feature points in the reference and sensed image (rotated by 60 deg)

Zernike moments	Zernike moment magnitude	
	Reference image pixel (140,23)	Sensed Image pixel (43,66)
Z_{00}	3.832	3.852
Z_{11}	5.013	4.600
Z_{20}	3.963	3.998
Z_{22}	4.731	5.437
Z_{31}	3.454	2.904
Z_{33}	4.825	5.057
Z_{40}	5.229	5.227
Z_{42}	4.0995	3.626
Z_{44}	5.335	4.747
Z_{51}	4.824	5.081

TABLE 1: Zernike moment magnitude for a feature point pair from reference image and sensed image (rotated by 60 deg).

[ii] The feature points of the reference image are matched with the feature points of the sensed image by computing the correlation coefficients of the two descriptor vectors. The matched points are those who give maximum coefficient correlation value. The correlation coefficient C of two feature vectors V_1 and V_2 is defined as

$$C = \frac{(V_1 - m_1)^T (V_2 - m_2)}{\|(V_1 - m_1)\| \, \|(V_2 - m_2)\|} \tag{5}$$

Where, m_1 and m_2 are the mean values of the two vectors V_1 and V_2 respectively.

3.4. Example of Zernike moment

Following examples shows calculations steps for Zernike moment for an 8 x 8 block.

$$I = \begin{bmatrix} 5 & 10 & 15 & 20 & 25 & 30 & 35 & 40 \\ 10 & 20 & 30 & 40 & 50 & 60 & 70 & 80 \\ 15 & 30 & 45 & (60) & 75 & 95 & 105 & 120 \\ 20 & 40 & 45 & 65 & 85 & 105 & 125 & 135 \\ 25 & 50 & 60 & 85 & 100 & 115 & 130 & 145 \\ 30 & 60 & 75 & 105 & 115 & 130 & 145 & 160 \\ 35 & 70 & 90 & 125 & 130 & 145 & 160 & 175 \\ 40 & 80 & 105 & 135 & 145 & 160 & 175 & 190 \end{bmatrix}$$

Normalize the pixel locations for *x* and *y* varying from -1 to +1 with a step size 0.2857. Apply an 8 x 8 mask, with pixel values one within the unit circle, which is required for the calculation of Zernike moment.

$$\text{Mask} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

For z = x+ iy and mask = 1, calculate r and θ. Obtained,

$$r = [0.83, 0.72, 0.72, 0.83, \ldots .\mathbf{0.45}, \ldots 0.72, 0.83]_{1\times 32}$$

and

$$\theta = [-2.60, -2.94, 2.94, 2.60, \ldots \mathbf{-1.89}, \ldots 0.20, 0.54]_{1\times 32}$$

For finding Zernike moment at grid location (3,4), value of r and θ are **0.4518** and **-1.89** respectively, which are 12th element of r and θ vectors. Pixel intensity, I(3, 4) = 60. For p=3 and q=1 satisfying the above mentioned condition, from equation (4), $R_{31}$the polynomial value becomes -0.6269. Putting this value in equation 3, we get $V_{31}(r, \theta) = 0.19825 + 0.59485i$. Applying all the above values in equation 2, we get Z_{31}=15.29 + 45.4588i. Finally in log scale we get [abs (log (Z_{31}))] =4.0661.

3.4. Transformation parameters estimation

Image geometrical deformation has many different ways of description. In this paper we have considered combination of rotation and scaling. Given the two sets of corresponding feature point coordinates, the estimation of the transformation parameters is required to map the sensed image into its original size, orientation, and position. This requires at least three feature point pairs having the maximum coefficient correlation. These parameters are estimated by solving a least-square minimization problem.

4. EXPERIMENTAL RESULTS

In this section, the evaluation of the performance of the proposed algorithm is done by applying different types of distortions. A reference image is geometrically distorted and in addition noise is being added or the image is compressed or expanded. The parameters of the geometric distortion are obtained by applying the proposed algorithm using the reference and sensed images. A set of simulation has been performed to assess the performance of the proposed algorithm with respect to registration accuracy and robustness.

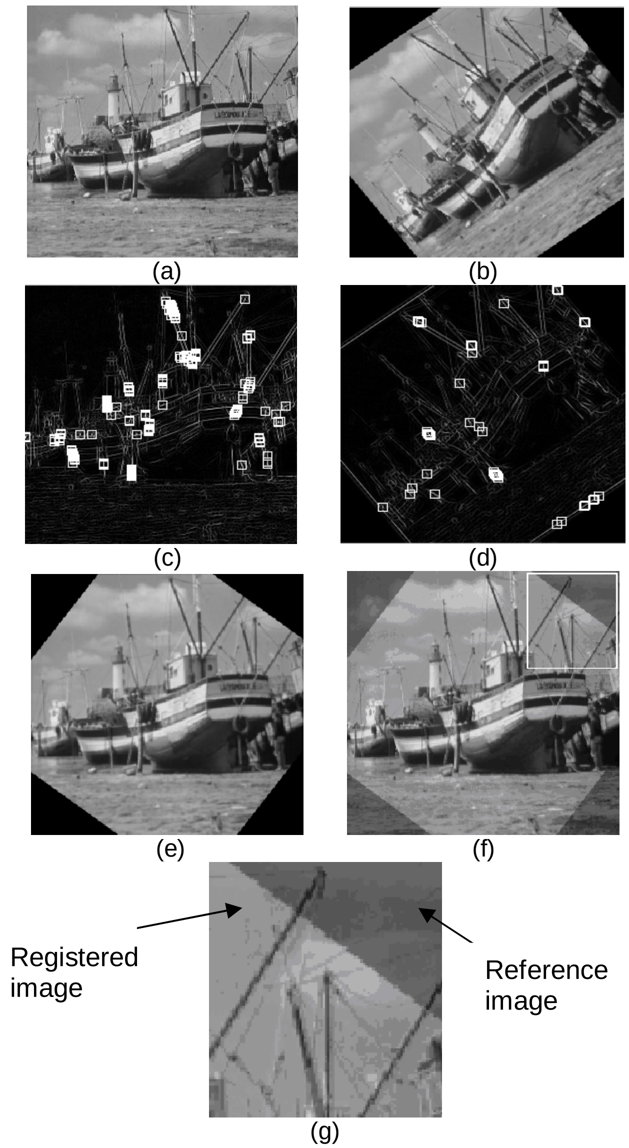


Figure 7: Experimental results (a) Reference image (b) Sensed image (rotated by 37 deg) (c) NSCT maxima image of reference image, N is 2 (d) NSCT maxima image of sensed image, N is 2 (e) Registered image (f) Registered image overlaid on Reference image (g) Enlarged portion of the overlaid image.

A gray level “boat” image of size 256x256 is being used as a reference image. The simulation results have been obtained using MATLAB software package. The experiments were performed according to the following settings: NSCT decomposition of all test images performed using the NSCT toolbox, was carried out with N = 2 resolution levels. But to increase the capability of proposed algorithm for higher amount of distortions we need to increase resolution levels N to 3. The parameter C is user defined and ranges from 4 to 8 and the Zernike moment based descriptor neighbourhood radius R = 20. Results of registering the geometrically distorted images combined with other image processing operation are shown in Fig. 7. In this figure, first reference image and sensed images were shown. Then the NSCT maxima images of both the images have been shown. At the last registered image overlaid on the reference image is shown. To highlight the registration accuracy a small square section of the reference image which is not available in the sensed image after rotation has been magnified along with connected features from the sensed image. Perfect alignment between the two images justifies the registration accuracy.

The applied distortions/ transformations are shown in below Figures. It can be seen that, the estimated transformation parameters are very close to the actual applied parameters. This illustrates the accuracy in image recovery, in the presence of noise, coarse compression or expansion of the image. Figures 8 to 11 shows simulation results with different rotation and scale which shows the accuracy of registration.

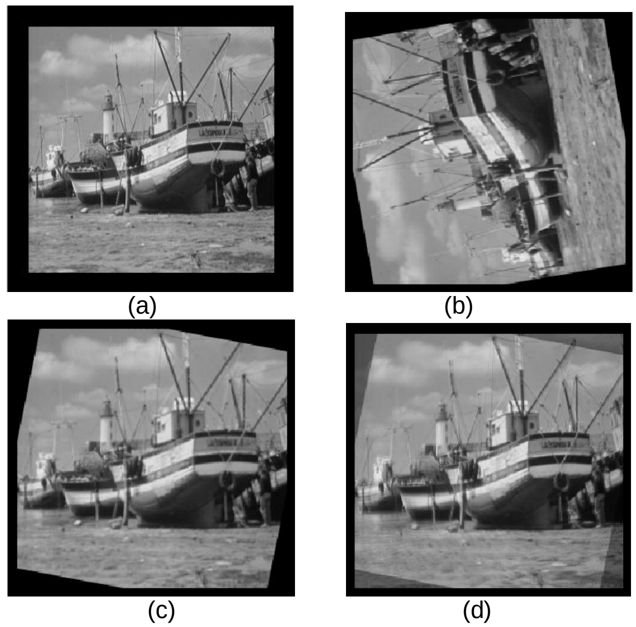


Figure 8: (a) Reference image (b) Sensed image (rotated by 100 degrees) (c) Registered image (d) Registered image overlaid on Reference image (N = 3).

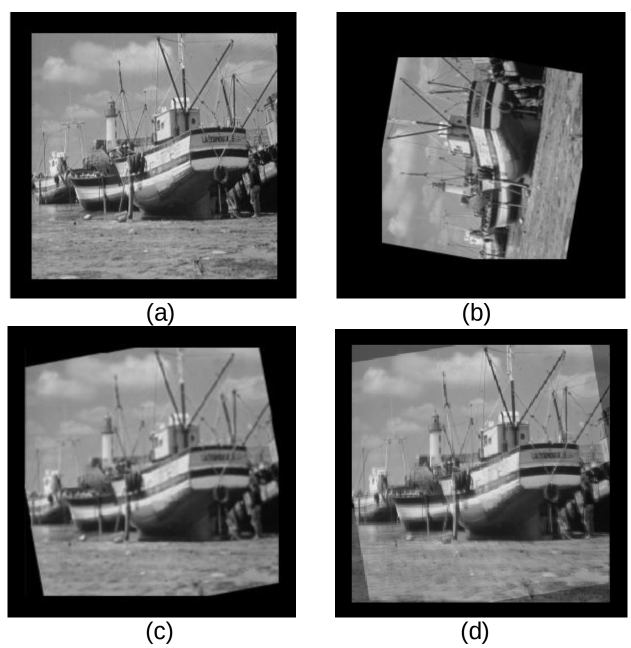


Figure 9:(a) Reference image (b) Sensed image (rotated by 80 degrees Scaled by 0.8) (c) Registered image (d) Registered image overlaid on Reference image (N = 2).

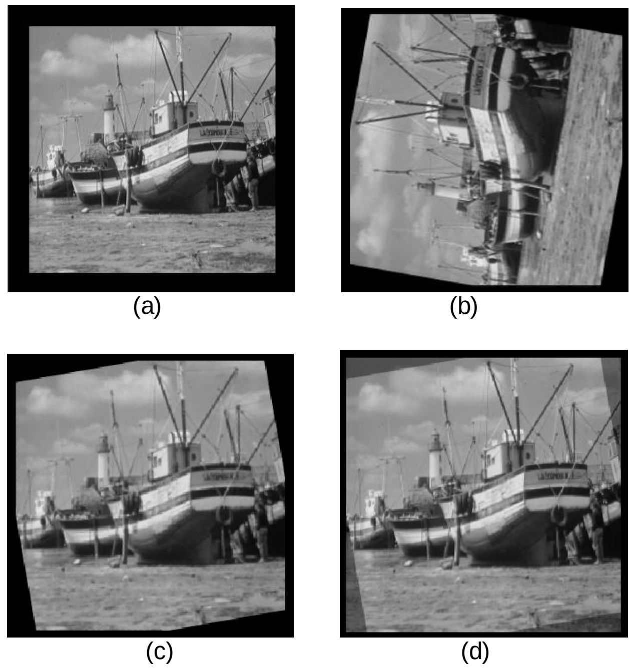


Figure 10: (a) Reference image (b) Sensed image (rotated by 80 degrees Scaled by 2.2) (c)Registered image (d) Registered image overlaid on Reference image (N = 2).

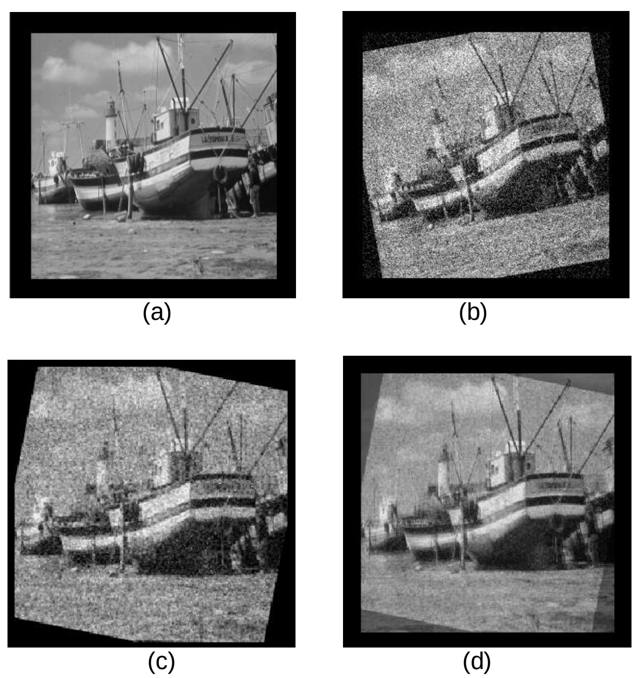


Figure 11 :(a) Reference image (b) Sensed image (rotated by 10 degrees with Gaussian noise, mean 0 and variance 0.02) (c) Registered image (d) Registered image overlaid on Reference image (N = 2).

5. CONCLUSION

In the proposed algorithm major and basic elements of the feature based automated image registration has been explored. We use nonsubsampled contourlet transform (NSCT) based feature point extractor, to extract significant image feature points across spatial and directional resolutions. Zernike moments based similarity measure is used for feature correspondence. The experimental results clearly indicate that the registration accuracy and robustness is very acceptable. This confirms the success of the proposed NSCT based feature points extraction approach for image registration.

6. REFERENCES

[1] Brown L G., *"A survey of image registration techniques"*. ACM Computing Surveys, 24(4), 325-376, 1992.
[2] A. Ardeshir Goshtasby, *"2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications",* A John Wiley & Sons, Inc., Publication, USA.
[3] B. Zitova and J. Flusser, *"Image Registration methods: A Survey". Image Vision Computing*, 21(11), 977-1000, 2003.
[4] A. Khotanzad and Y.H. Hong, *"Invariant Image Recognition by Zernike moment". IEEE Trans. PAMI,* 12(5), 489-497, 1990.
[5] Cho-Huak and R.T. Chin, *"On Image Analysis by the method of moments". IEEE Trans. PAMI,* 10(4), 496-513, 1988.
[6] M.N. Do and M. Vetterli, *"The Contourlet Transform: an Efficient Directional multiresolution Image Representation"*. *IEEE Trans. on Image Processing,* 14(12), 2091-2106, 2005.
[7] A.L.Cunha, J. Zhou, and M.N. Do, *"The Nonsubsampled Contourlet Transform: Theory, Design, and Applications"*. *IEEE Trans. on Image Processing,* 15(10), 3089-3101, 2006.
[8] R. H. Bamberger and M. J. T. Smith, *"A filter bank for the directional decomposition of images: theory and design". IEEE Trans. on Signal Processing*, 40(7), 882-893, 1992.
[9] S. X. Liao and M. Pawlak, *"On the Accuracy of Zernike Moments for Image Analysis". IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(12)1358-1364, 1998.
[10] J. Zhou, A.L. Cunha, and M.N. Do, *"The Nonsubsampled Contourlet Transform: Construction and Application in Enhancement"*. In Proceedings of IEEE Int. Conf. on Image Processing, ICIP 2005, (1), 469-472, 2005.
[11] P. J. Burt and E. H. Adelson, *"The Laplacian pyramid as a compact image code". IEEE Trans. on Commun.,* 31(4), 532–540, 1983.
[12] M. N. Do and M. Vetterli, *"Framing pyramids". IEEE Trans. Signal Process.* 51(9),2329-2342, 2003.
[13] C. Serief, M. Barkat, Y. Bentoutou and M. Benslama, *"Robust feature points extraction for image registration based on the nonsubsampled contourlet transform".* International Journal Electronics Communication, 63(2), 148-152, 2009.
[14] Manjunath B S and Chellappa R. "*A feature based approach to face recognition".* In Proceedings of IEEE conference on computer vision and pattern recognition, Champaign, 373–378, 1992.
[15] M.S. Holia and V.K.Thakar, "Image registration for recovering affine transformation using Nelder Mead Simplex method for optimization",International Journal of Image Processing(ISSN 1985 2304) , Vol.3, No.5, pp.218-228, November 2009
[16] R. Bhagwat and A. Kulkarni, "An Overview of Registration Based and Registration Free Methods for Cancelable Fingerprint Template", International Journal of Computer Science and Security ISSN (1985-1553) Vol. 4 , No.1 pp.23-30, March 2010.