Hybrid Approach of Image Stitching using Normalized Gradient Correlation and Harris Corner Detector

Paresh M. Patel
Assistant Professor
SSASIT, Surat-395006

Hitesh A. Ravani
Assistant Professor
SSASIT, Surat-395006

ABSTRACT

Image Stitching is application of image registration in which images with few overlapping are aligned and stitch together to form wide angle images. Image registration is the fundamental task used to match two or more partially overlapping images taken, for example, at different times, from different sensors, or from different viewpoints and stitch these images into one panoramic image comprising the whole scene. It is a fundamental image processing technique and is very useful in integrating information from different sensors, finding changes in images taken at different times, inferring three-dimensional information from stereo images, and recognizing model-based objects. Some techniques are proposed to find a geometrical transformation that relates the points of an image to their corresponding points of another image. To register two images, the coordinate transformation between a pair of images must be found. In this paper, proposed algorithm is based on Log-Polar Transform and first roughly estimate the angle, scale and translation between two images. The proposed algorithm can recover scale value up to **5.85**. In order to improve the insufficiency of Harris corner, proposed method present an auto-adjusted algorithm of image size based on NGC. The robustness of this algorithm is verified on different images with similarity transformation and in the presence of noise and finally by using RANSAC algorithm smooth stitching can be obtained than any other methods.

General Terms

Image Registration, Image Stitching, Correlation, Harris Corner Detector, Mosaic, Scale Factor

Keywords

Log-Polar Transform (LPT), Fast Fourier Transform (FFT), Normalized Gradient Correlation (NGC), Random Sample Consensus (RANSAC), Direct Linear Transform (DLT), Scale Invariant Feature Transform (SIFT).

1. INTRODUCTION

Image stitching is the application of the image registration, in which images with few overlapping are aligned and stitched together to form the panorama image or a wide angled image [12]. To stitch the images there are various methods like image stitching using Scale Invariant Feature Transform (SIFT) [11], image mosaic using phase correlation and Harris corner detector [9]. Image stitching algorithms take the alignment estimates produced by such registration algorithms and blend the images in a seamless manner, taking care to deal with potential problems such as blurring or ghosting caused by parallax and scene movement as well as varying image exposures.

The estimation of the relative motions between two or more images is probably at the heart of any autonomous system which aims at the efficient processing of visual information. Motions in images are induced due to camera displacements or displacements of the individual objects composing the scene. Image registration techniques for global motion estimation address the problem of compensating for the camera ego-motion and finally aligning the images. Practical applications are numerous: from global scene representation and image mosaicking to object detection or tracking and video compression.

In propose algorithm there is a robust correlation-based scheme which operates in the Fourier domain for the estimation of translations, rotations and scaling in images. For the class of similarity transforms, a frequency domain approach has several advantages. First, through the use of correlation, it enables an exhaustive search for the unknown motion parameters. Second, the approach is global which equips the algorithm with robustness to noise [4]. Third, the method is computationally efficient. This comes from the shift property of the Fourier Transform (FT) and the use of Fast Fourier Transform (FFT) routines for the rapid computation of correlations.

The work in [3] introduces the basic principles for translation, rotation and scale-invariant image registration in the frequency domain. Given two images related by a similarity transform, the translational displacement does not affect the magnitudes of the FTs of the two images. Re-sampling the Fourier magnitudes on the log-polar grid reduces the problem of estimating the rotation and scaling to one of estimating a 2D translation. Thus, the method relies on correlation twice: once in the log-polar Fourier domain to estimate the rotation and scaling and once in the spatial domain to recover the residual translation. In the usual way, the authors use phase correlation (PC) [2] instead of standard correlation while they perform conversion from Cartesian to log-polar using standard interpolation schemes (e.g. bilinear interpolation).

In Proposed scheme, first replace image functions with complex gray-level edge maps and then compute the standard Cartesian FFT. Next, simply resample the Cartesian FFT on the log-polar grid using bilinear interpolation. Neither sophisticated FFT nor over-sampling is employed to enhance accuracy. To perform robust correlation, replace phase correlation with gradient-based correlation schemes.

2. LOG POLAR TRANSFORM (LPT)

All The Log-Polar Transform is used for image registration for its rotation invariant and scale invariant properties. The log-polar image geometry is used because of the fact that scaling and rotation in Cartesian domain corresponds to pure translation in log-polar domain Taking logarithm of radial distance ρ log-polar coordinates can be found. The log-polar transformation is a conformal mapping from the points on the Cartesian plane (x, y) to points in the log-polar plane

(log(ρ),θ). Considering a polar coordinate system, where ρ is the radial distance from the center of the image say (xc, yc) and θ denotes the angle. Any point (x, y) can be represented in polar coordinates and is given by

$$(\rho,\theta) = \sqrt{(x-xc)^2 - (y-yc)^2}, \tan^{-1}\frac{y-yc}{x-xc} \qquad (1)$$

If the polar coordinate transformation is applied to an image in the Cartesian domain, then the radial lines in the Cartesian domain maps to horizontal lines in (ρ,θ) domain. Fig.1 shows the approximate mapping from Cartesian space to polar space or (ρ,θ) space. The black box shows that the pixels are at a constant angle with respect to the center. Similarly, the boxes with cross marks are at a constant radial distance from the center.
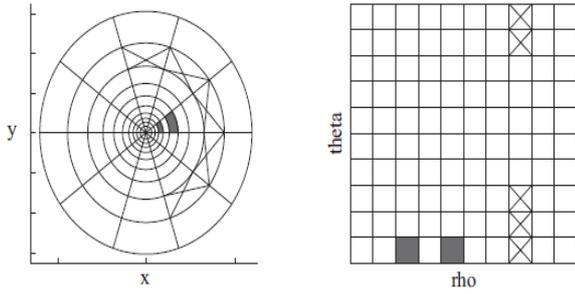
Fig. 1. Approximate mapping from Cartesian space to (ρ, θ) space [7]

It is clear from Fig.1 that if there was a rotation in the image, the black box will shift its positions on the theta axis [7]. A similar situation can be discussed for the scale variation. In log-polar coordinates, logarithm of the radial axis is taken by

$$(\rho, \theta) = (\log(\rho), \theta) \qquad (2)$$

Now if the image is scaled by a factor of say α, then the coordinates (x, y) in Cartesian domain will become (αx, αy). Introduction of logarithms will simplify the procedure, the coordinates in log domain will be reflected as

$$(\log((\alpha x), \log((\alpha y)) = ((\log\alpha + \log x), (\log\alpha + \log y)) \qquad (3)$$

The effects of distortions are expressed by log-polar image translation on ρ axis and θ axis, respectively in the log-polar coordinates. However, when the original image is translated by (∆x, ∆y), the corresponding log polar coordinates is represented by

$$\rho' = \log\sqrt{(e^\rho\cos\theta - \Delta x)^2 + (e^\rho\sin\theta - \Delta y)^2} \qquad (4)$$

$$\theta' = \tan^{-1}\frac{e^\rho\sin\theta - \Delta y}{e^\rho\cos\theta - \Delta x} \qquad (5)$$

According to above two Equations (4) and (5), the slight translation produces a modification of the log-polar image. Therefore, the log-polar image is not suitable for faithfully extracting translation parameters of images [2, 5 and 6]

3. ROBUST FFT-BASED SCALE-INVARIANT IMAGE REGISTRATION

To estimate the translational displacement, replace standard correlation with gradient-based correlation schemes [8]. Gradient correlation (GC) combines the magnitude and orientation of image gradients.

$$GC(u) \overset{\Delta}{=} G_1(u) * G_2^*(u)$$

$$= \int_{\Re^2} G_1(x) * G_2^*(x+u)dx$$

Where, $G_i(x) = G_{i,x}(x) + jG_{i,y}(x)$

$G_{i,x} = \nabla_x I_x$ and $G_{i,y} = \nabla_y I_y$ are the gradients along the horizontal and vertical direction respectively.

From (6),

GC(u)= $G_{1,x}$(u)*$G_{2,x}$(-u)+$G_{1,y}$(u)*$G_{2,y}$(-u) +

j{-$G_{1,x}$(u)*$G_{2,y}$(-u)+$G_{1,y}$(u)*$G_{2,x}$(-u)}   (7)

The imaginary part in the above equation is equal to zero, therefore

GC(u)= $G_{1,Z}$(u)*$G_{2,Z}$(-u)+$G_{2,Z}$(u)*$G_{2,Z}$(-u)   (8)

Using the polar representation of complex numbers,

$$R_i = \sqrt{G^2_{i,x} + G^2_{i,y}} \text{ and } \Phi_i = \tan^{-1}(G_{i,y}/G_{i,x}) \quad (9)$$

Based on the representation of (8)

GC(u)=∫R_2(x)R_2(x+u)cos[($\varnothing_2$(x)-$\varnothing_2$(x+u))]dx   (10)

The magnitudes Ri reward pixel locations with strong edge responses and suppress the contribution of areas of constant intensity level which do not provide any reference points for motion estimation. Orientation information is embedded in the cosine kernel. This term is responsible for the Dirac-like shape of GC and its ability to reject outliers induced by the presence of dissimilar parts in the two images.

Assume, $Ri = 1$, $i = 1$, 2. To optimize the orientation difference function DF of the image salient structures solely, the normalized gradient correlation:

$$NGC(u) \overset{\Delta}{=} \frac{G_1(u) * G_2^*(-u)}{|G_1(u)| * |G_2^*(-u)|} \qquad (11)$$

The above analysis, (11) takes the form:

$$NGC(u) \overset{\Delta}{=} \frac{\int_{\Re^2} R_1(x)R_2(x+u)\cos\Delta\Phi(u,x)dx}{\int_{\Re^2} R_1(x)R_2(x+u)dx} \qquad (12)$$

NGC has two interesting properties:

1. $0 \leq /NGC(\mathbf{u})/ \leq 1$

2.   Invariance to affine changes in illumination.

The first property provides a measure to assess the correctness of the match. To show the second property, consider:

$I'_2 = aI_2(x) + b$ with $a \in \Re^+$ and $b \in \Re$

4. HARRIS CORNER DETECTOR

Corner Detection: Basic Idea [9]

- It is easy to recognize the point by looking through a small window.

- Shifting a window in any direction should give a large change in intensity.

"flat" region: no change in all directions

"edge": no change along the edge direction

"corner": significant change in all directions
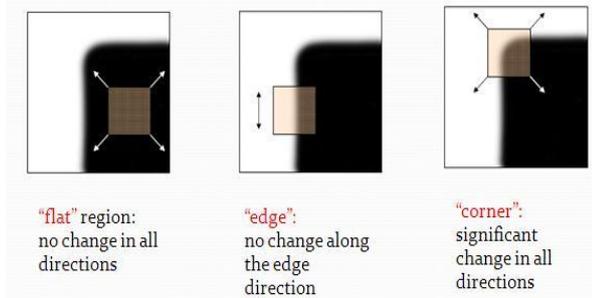
$E(u,v) = \sum w(x,y)[I(x+u \ , \ y+v) - I(x,y)]^2$

w(x,y)　　　– window Function

I(x+u , y+v) – shifted intensity

I(x,y)　　　　- intensity

Window function w(x, y)=1 in window & 0 in outside. For nearly constant patches, [I(x+u, y+v)-I(x, y)] will be near 0. For very distinctive patches, this will be larger. Hence, get patches where E(u,v) is LARGE.

$$\sum [I(x+u, y+v) - I(x,y)]^2$$

First Order Approx.

$$\approx \sum [I(x,y) + uI_x + vI_y - I(x,y)]^2$$

$$= \sum u^2 I_x^2 + 2uv I_x I_y + v^2 I_y^2$$

Matrix equation

$$= \begin{bmatrix} u & v \end{bmatrix} \left(\sum \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \right) \begin{bmatrix} u \\ v \end{bmatrix}$$

Measure of Corner Response:

$R = Det(M) - k \ (trace \ M)^2$

$Det(M) = \lambda_1 \lambda_2$　　$Trace \ M = \lambda_1 + \lambda_2.$

k is constant, k=0.04 − 0.06.

5. IMAGE MOSAIC

Image mosaic is a technique to stitch two or more images which contain some common portion and which are scaled, rotated and translated [12]. It can be widely used in space exploration, remote sensing image processing, medical image analysis. There are two algorithms for Image mosaic:

(1)Region-based algorithm (2) Feature-based algorithm

In region-based method first find gradient of both images(reference & sensed image) then find gradient correlation between images and then find position of overlapping area but this method has some disadvantages:

Applicable only for those image who has >40% matching parts[9].Not applicable for rotated, scaled and translated image[10].By using feature-based method above problem can be overcome.

In feature-based method first apply image registration algorithm then apply Harris corner detector [9] and then apply RANSAC algorithm [9] to stitch both images.

6. PROPOSED ALGORITHM

INPUTS: Two images Ii, i = 1, 2 related by a translation t, rotation θ and scaling s.

STEP 1: Estimate gradient $Gi=Gi(x) + j \ Gi(y)$. By finding the Horizontal and vertical edges using the standard Cartesian FFT.

STEP 2: Resample on the log-polar grid using bilinear interpolation.

STEP 3: Estimate θ and s using NGC in the log-polar domain.

STEP 4: Scale down and derogate the zoomed, rotated image. Resolve the π ambiguity.

STEP 5: Find the translation using the Normalized Gradient Correlation in a spatial domain.

STEP 6: Shift the image to the original position as per the reference image which is obtained in step 5 and registered image is obtained.

STEP 7: Find features using Harris corner detector of the reference image and the de-rotated, de-scaled image.

STEP 8: Find putative matching corners by matching through correlation.

STEP 9: Stitch the reference and sensed image using RANSAC algorithm.

Step-By-Step Algorithm-Close Look

Inputs
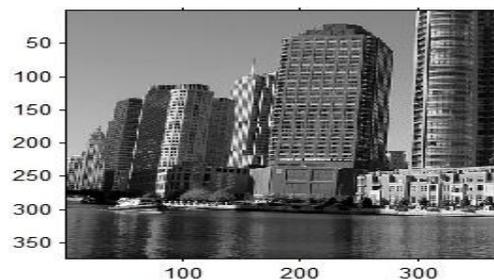
Fig 2. Reference image (input)
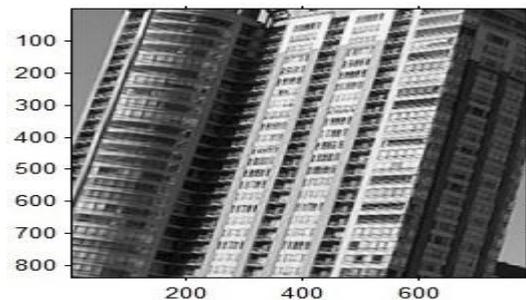
Fig 3. Sense image (input)

STEP 1: Estimate gradient *Gi=Gi(x) + j Gi(y)*. By finding the Horizontal and vertical edges using the standard Cartesian FFT.
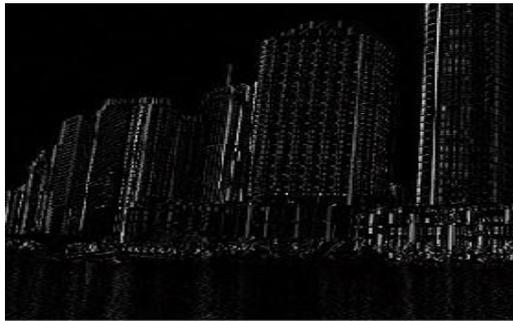
Fig 4. Gradient of reference image

Fig 5. Gradient of sense image

STEP 2: Apply Log Polar Transform on Reference image
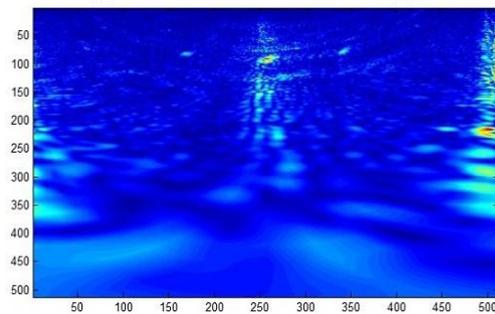
Fig 6. Log polar transform if reference image
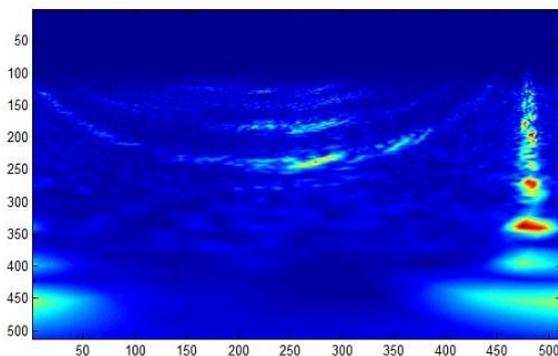
Fig 7. Log polar transform of sense image

STEP 3: For finding theta and scale: $\theta = (y) * 180 / (N)$;

Base = exp (log (N) / (N));

Scale = Base ^ ((x));

Where; N is size of Image and (x, y) is peak after find Normalized Gradient Correlation between referenced and sense image.

Estimating for given example:

$\theta = 8.7891$, scaling = (1/scale) = 3.9143.

STEP 4: Scale Down and Re-rotate which is estimated above and Pad it according to reference Image.

STEP 5: Now NGCorr is find between Reference Image and De-rotated and Scale down image. Peak is founded which has some co-ordinate (x, y), which is the point of translation. Remove the Translation.

Estimating for given example:

Translation= (25, -112)

STEP 6: Shift the image to its original position from estimated Scale, Rotation and Translation
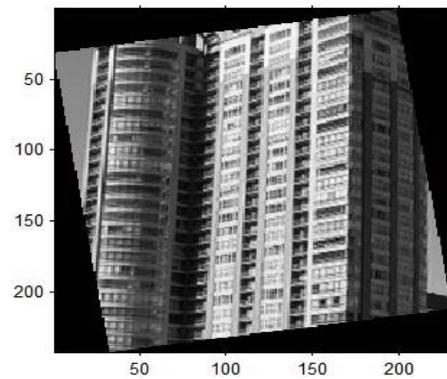
Fig 9. Shift the image to the original position

STEP 7: Find corners using Harris corner detector.
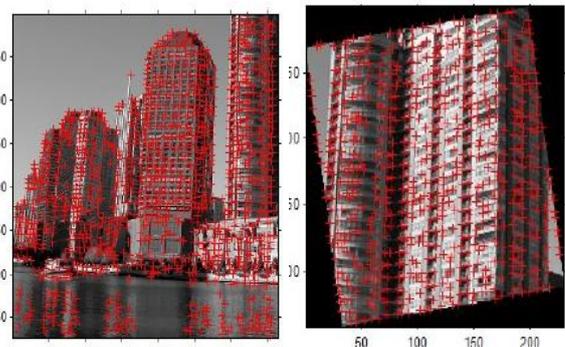
Fig. 10 corners of ref. and sensed image

STEP 8: Match through correlation to find putative points.

Fig 11 Matching of Corners

STEP 9: Now Stitch the Registered Image and Reference Image Using RANSAC algorithm.
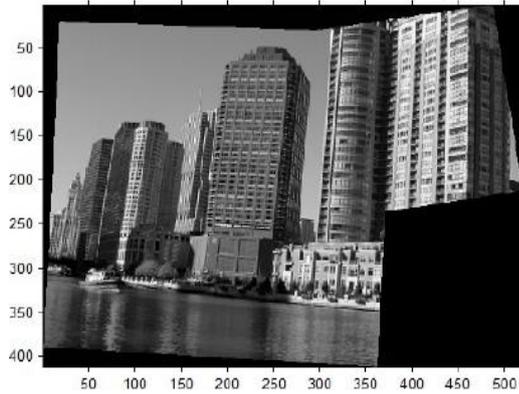
Fig 12 stitching of ref. and sensed image

7. EXPERIMENTAL DATA

I. Registration Results
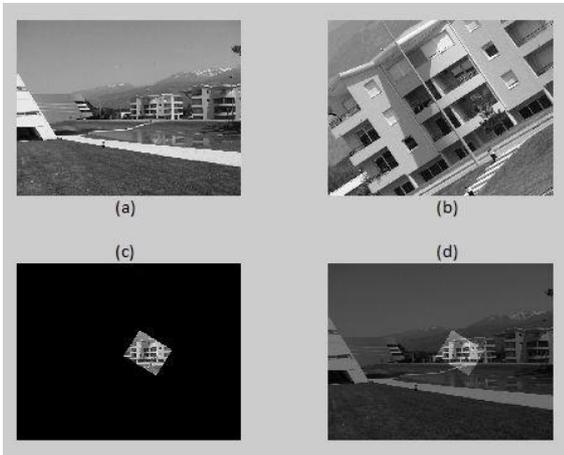
Fig 13. Building.pgm (a) reference image, (b) sense image, (c) derotate, downscale and translated image,(d) registered image
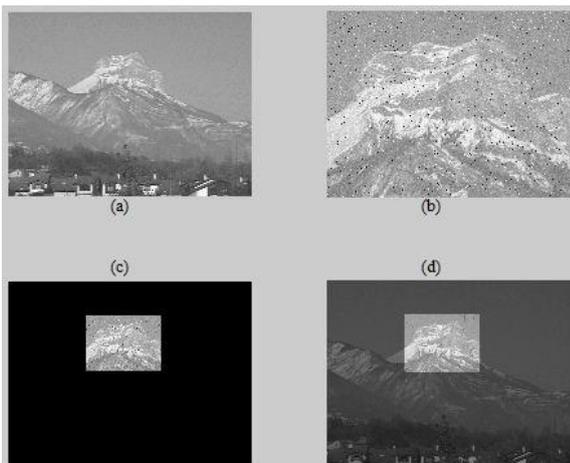
Fig 14. Mountain.pgm (a) reference image, (b) sense image, (c) derotate, downscale and translated image,

(d) Registered image

Fig. No.	Image	Rotation Θ	SCALE S	Translation T
Fig. 11	building	-31.64	5.85	(72, 4)
Fig. 12	mountain	0	3.26	(-22,-92)

II. Stitching Results:
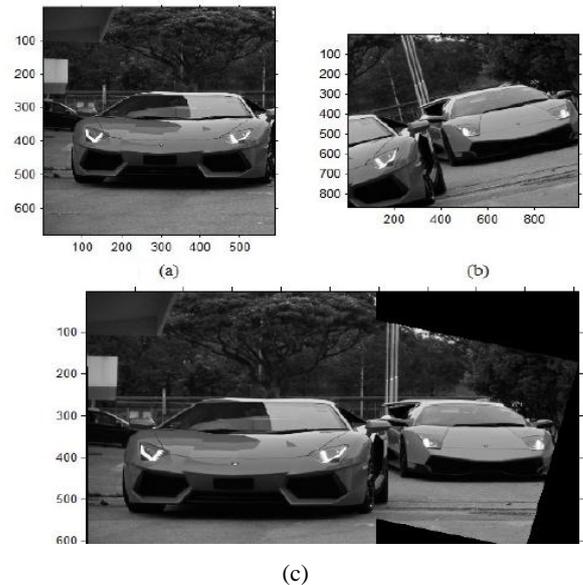
Fig. 15: Car (a) Reference image (b) Sensed image and (c) Stitched image.
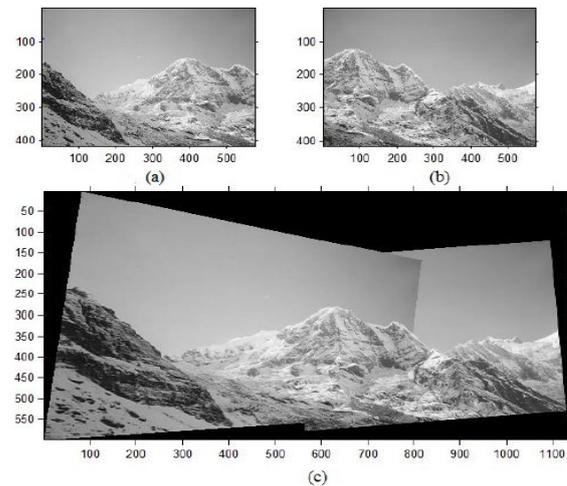
Fig.16: Mountain (a) Reference image (b) Sensed image and (c) Stitched image

From the stitching results it is possible to stitch images having less than 40% matching area.

8. CONCLUSION

Mostly image registration algorithm is presented using Log-Gabor filter, Log-Polar Transform and Phase Correlation. The rotation and scale invariant properties of the LPT, along with

FT and phase correlation allow us to develop a robust algorithm that works faithfully under geometric distortions like rotation, scale and translation. Hence this approach is highly effective in registering aerial images. Any amount of scaling and rotation in Cartesian domain will be pure translation in log polar domain. Scale and rotation between images is recovered by first converting them into Log-Polar images and then applying FFT-based cross-correlation. A key feature of Fourier-based registration methods is the speed offered by the use of FFT routines. The proposed scheme estimates large motions accurately and robustly without the need of excessive zero padding and over-sampling, thus without sacrificing part of the computational efficiency which typifies the frequency domain formulation. Finally, a further advantage of the FFT-based image registration methods is that it is able to find the scaling up to 5.85. By using NGC the scope of Harris corner detector can be extended and also improve the stitching results using RANSAC algorithm. In future this proposed method can be extended for color Image Registration.

9. ACKNOWLEDGMENTS

10. REFERENCES

[1] B.S. Reddy and B.N. Chatterji. "An FFT-based technique for translation, rotation and scale invariant image registration." IEEE Trans. Image Processing, 5(8):1266–1271, [1996].

[2] B. Sun, D. Zhou A, "Rotated Image Matching Method Based on CISD", ISNN 2007, Part I, LNCS 4491, pp.1346-1352, [2007].

[3] C.D. Kuglin and D.C. Hines. "The phase -correlation image alignment method." In Proc. IEEE Conf. Cybernetics and Society, pages 163–165, [1975].

[4] Fan YANG, Linlin WEI, Zhiwei ZHANG, Hongmei TANG, "Image Mosaic Based on Phase Correlation and Harris Operator", Journal of Computational Information Systems,

Vol 8:6(2012) 2647–2655.

[5] Jignesh N Sarvaiya, Dr. Suprava Patnaik, Salman Bombaywala, "Image registration using Log polar transform and phase correlation", TENCON 2009 - 2009 IEEE Region 10 Conference, pp.1-5, [2009].

[6] Pengrui Qiu, Ying Liang and Hui Rong, "Image Mosaics Algorithm Based on SIFT Feature Point Matching and Transformation Parameters Automatically Recognizing" 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013).

[7] Qin-sheng Chen, Michel Defrise, F. Deconinck, "Symmetric phase-only matched filtering of Fourier Mellin transform for image registration and recognition", IEEE Trans. on Pattern Analysis and Machine Intelligence 16, pp.1156-1168, [1994].

[8] S. Zokai and G.Wolberg. "Image registration using log-polar mappings for recovery of large-scale imilarity and projective transformations". IEEE Trans. Image Processing.

14(10):1422–1434, [2005].

[9] Tzimiropoulos G. "Robust FFT-based scale invariant image registration with image gradients." IEEE (2010).

[10] Xiaoxin Guo, Zhiwen Xu, Yinan Lu, and Yunjie Pang, "An Application of Fourier-Mellin Transform in Image Registration", Proceedings of The Fifth International Conference on Computer and Information Technology, pp.619-623 , [2005].

[11] Xin Xie and Yin Xu "A study on Fast SIFT Image Mosaic Algorithm Based on compressed Sensing and wavelet Transform" In Journal of ambient Intelligence and Humanized computing(Springer), Vol 6, Issue 6, Pages 835-843,[2015].

[12] Zhicheng Wang and Yufei Chen "An Automatic Panoramic Image Mosaic method based on Graph Model" In Springer Link. Vol 75, Issue 5, pages 2725-2740, [2016].