

# ROBUST FFT BASED SCALE INVARIANT IMAGE REGISTRATION USING IMAGE GRADIENTS

MR. TEJAS D. DARJI

M.E.[signal processing and VLSI technology] Student, Department Of Electronics and  
Communication Engineering , Vishwakarma Government Engineering College  
Chandkheda , Gandhinagar.

*darji.tejas1989@gmail.com*

**ABSTRACT:** 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, we have proposed an algorithm that is based on Log-Polar Transform and first we roughly estimate the angle, scale and translation between two images. The proposed algorithm can recover scale value up to 5.85. The robustness of this algorithm is verified on different images with similarity transformation and in the presence of noise.

**Keywords—** Image Registration, Log-Polar Transform (LPT), Fast Fourier Transform(FFT).

## I: INTRODUCTION

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 mosaicing to object detection / tracking and video compression.

We propose 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 [1]. 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 [2] 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. Resampling 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) [3] instead of standard correlation while they perform conversion from Cartesian to log-polar using standard interpolation schemes (e.g. bilinear interpolation).

In our scheme, we first replace image functions with complex gray-level edge maps and then compute the standard Cartesian FFT. Next, we 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, we replace phase correlation with gradient-based correlation schemes.

## II: LOG POLAR TRANSFORM (LPT):

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  $\rho$ , we get log-polar coordinates. 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

$$(\rho, \theta) = \left( \sqrt{(x-x_c)^2 + (y-y_c)^2}, \tan^{-1} \frac{y-y_c}{x-x_c} \right) \quad (1)$$

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.

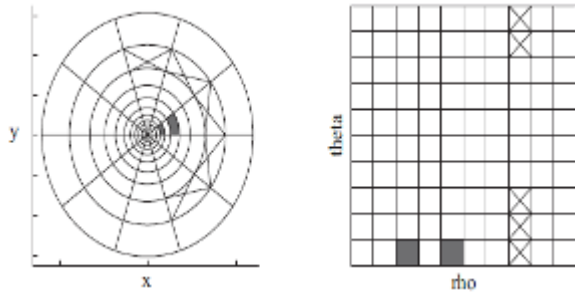


Fig. 1. Approximate mapping from Cartesian space to (ρ, θ) space.[4]

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. 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 [4]. A similar situation can be discussed for the scale variation. In log-polar coordinates, logarithm of the radial axis is taken by

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

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

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

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 (Δx, Δy), the corresponding logpolar coordinates is represented by

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

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

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 [5,6,7].

### III: ROBUST FFT-BASED SCALE-INVARIANT IMAGE REGISTRATION

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

$$GC(u) = \int_{\mathbb{R}^2} G_1(u) * G_2^*(u) dx$$

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

$G_{i,x} = \nabla_x I_i, G_{i,y} = \nabla_y I_i$  are the gradients along the horizontal and vertical direction respectively. From (6), we can easily derive

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

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

$$GC(u) = G_{1,x}(u) * G_{2,x}(-u) + G_{1,y}(u) * G_{2,y}(-u) \quad (8)$$

Using the polar representation of complex numbers, we define

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

Based on the representation of (8)

$$GC(u) = \int R_1(x) R_2(x+u) \cos[\Phi_1(x) - \Phi_2(x+u)] dx \quad (10)$$

The magnitudes  $R_i$  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.

We assumed  $R_i = 1, i = 1, 2$ . To optimize the orientation difference function DF of the image salient structures solely, we introduce the normalized gradient correlation:

$$NGC(u) = \frac{\int_{\mathfrak{R}^2} R_1(x) R_2(x+u) \cos \Delta\Phi(u, x) dx}{\int_{\mathfrak{R}^2} R_1(x) R_2(x+u) dx} \quad (11)$$

The above analysis, (11) takes the form:

$$NGC(u) = \frac{\int_{\mathfrak{R}^2} R_1(x) R_2(x+u) \cos \Delta\Phi(u, x) dx}{\int_{\mathfrak{R}^2} R_1(x) R_2(x+u) dx} \quad (12)$$

NGC has two interesting properties:

1.  $0 \leq |NGC(u)| \leq 1$ .
  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  $a \in \mathfrak{R}^+$  and  $b \in \mathfrak{R}$

**IV : PROPOSED ALGORITHM**

INPUTS: Two images  $I_i, i = 1;2$  related by a translation  $t$ , rotation  $\theta$  and scaling  $s$ .

STEP 1: Estimate gradient  $G_i = G_i(x) + j G_i(y)$ . 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  $\theta$  and  $s$  using NGCorr in the log-polar domain.

STEP 4: Scale down and derotate the zoomed, rotated image. Resolve the  $\pi$  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.

STEP 7: Registered image is obtained.

STEP 8: Find the MSE (mean square error) between the Reference image and the registered image.

STEP 9: Overlap the Registered image above the Reference image.

**STEP-BY-STEP ALGORITHM-CLOSE LOOK  
INPUTS**



Fig 2. Reference image (input)



Fig 3. Sensed image (input)

STEP 1: Estimate gradient  $G_i = G_i(x) + j G_i(y)$ . 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:

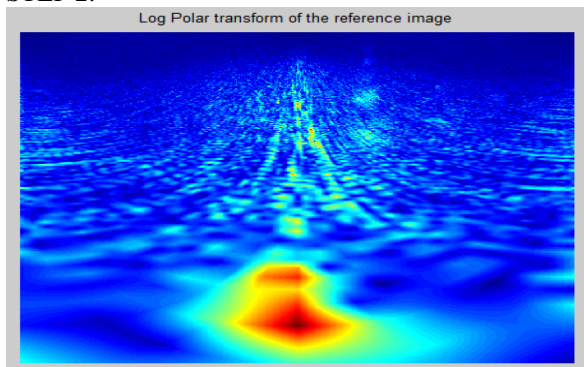


Fig 6. Log polar transform if reference image

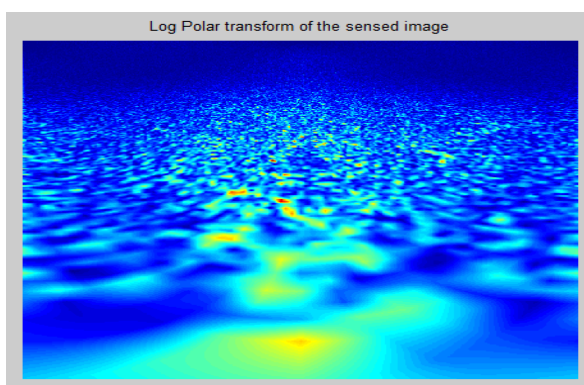


Fig 7. Log polar transform of sense image

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:

$$\theta = 0.00 , \text{ scaling} = (1/\text{scale}) = 3.2605.$$

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

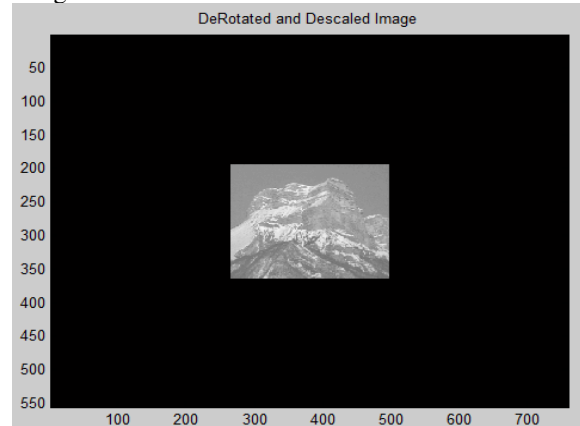


Fig 8. Derotated and scale down iamge

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:

$$\text{Translation} = (-22, -92)$$

STEP 6 AND 7:

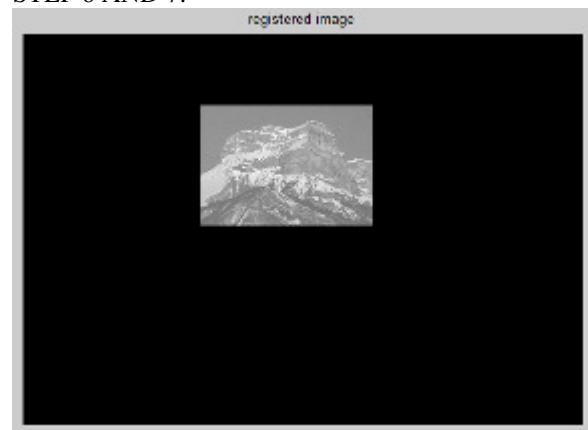


Fig 9. Shift the image to the original position

STEP 8. Now, find MSE between the Reference Image and Registered Image.

$$MSE = \frac{1}{MN} \left[ \sum_{x=1}^M \sum_{y=1}^N (I(x, y) - I'(x, y))^2 \right]$$

Where; I(x,y) is the original image,

I'(x,y) is the approximated version and M,N are the dimensions of the images.



Estimating for given example:  
MSE= 0.2833

STEP 9: Now Overlap the Registered Image on Reference Image.

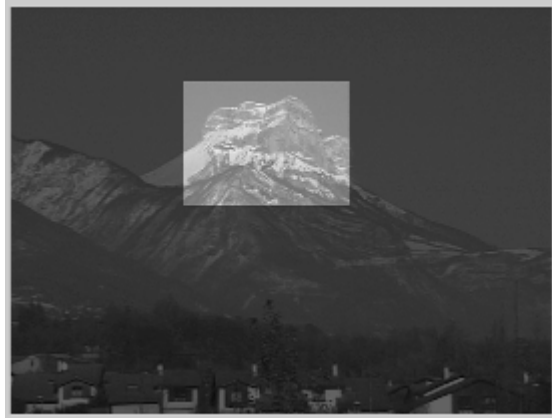


Fig 10. Overlap registered and reference image

**V: EXPERIMENTAL DATA.**

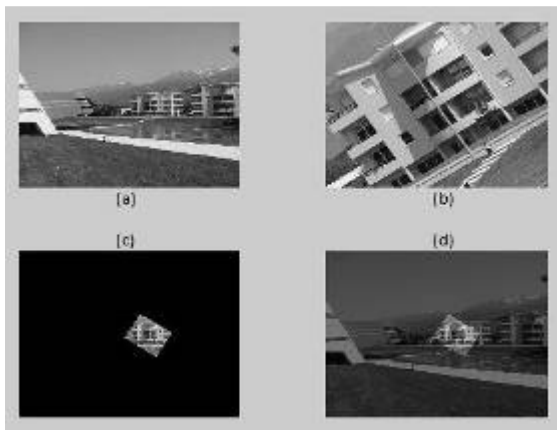


Fig 11. building.pgm (a) reference image, (b) sense image, (c) derotate, downscale and translated image,(d) registered image

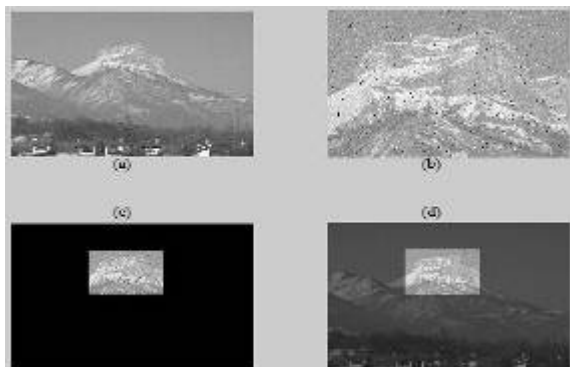


Fig 12. mountain.pgm (a) reference image, (b) sense image, (c) derotate, downscale and translated image,(d) registered image

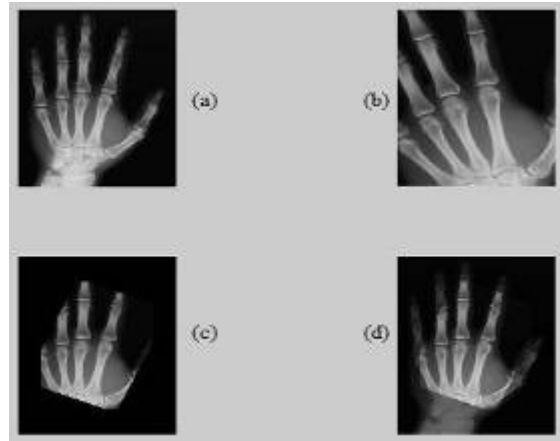


Fig 13. x-ray.jpg (a) reference image, (b) sense image, (c) derotate, downscale and translated image,(d) registered image

FIG NO.	IMAGE	RECOVERED $\theta$	SCALE
Fig. 11	building	-31.64	5.85
Fig. 12	mountain	0	3.26
Fig. 13	x-ray	-22.85	1.12

FIG NO.	IMAGE	TRANSLATION	MSE
Fig. 11	building	(72 , 4)	0.2507
Fig. 12	mountain	(-22,-92)	0.2528
Fig. 13	x-ray	(53 , 75)	0.0434

**VI: 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.

**REFERENCES**

[1] S. Zokai and G.Wolberg. "Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations". *IEEE Trans. Image Process.*, 14(10):1422–1434, [2005].

- [2] B.S. Reddy and B.N. Chatterji. "An FFT-based technique for translation, rotation and scaleinvariant image registration." *IEEE Trans. Image Processing*, 5(8):1266–1271, [1996].
- [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] 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].
- [5] 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].
- [6] 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].
- [7] B. Sun, D. Zhou A, "Rotated ImageMatchingMethod Based on CISD", *ISNN 2007, Part I, LNCS 4491*, pp.1346-1352, [2007].
- [8] Tzimiropoulos, G. "Robust FFT-based scale invariant image registration with image gradients." *IEEE* (2010): 10.

JIKRECE