The Log Polar Transformation for Rotation Invariant Image Registration of Aerial Images

B. Janardhana Rao¹ and O.Venkata Krishna²
¹CVR College of Engineering/ECE, Hyderabad, India
Email: janardhan.bitra@gmail.com
²CVR College of Engineering/EIE, Hyderabad, India
Email: venkatakrishna.odugu@gmail.com

Abstract—Aerial image registration requires a high degree of precision. To improve the feature-based registration accuracy, this paper proposes a novel Log-Polar Transform (LPT) based image registration. Instead of using the whole image in the conventional method, feature points are used here, which reduces the computational time. It is not important how the image patch is rotated for rotation invariance. The key is focusing on the feature points. So a circular image patch is used, instead of using square image patches as used in conventional methods. In the existing techniques of registration always do FFT first and then Log-Polar Transformation (LPT). In this proposed technique does LPT first and then the FFT. The proposed technique contains four steps. First, feature points are selected in both the reference image and the sensed image with corner detector. Secondly, using feature point positions as centers image patches are created. Circular image patches are cropped choosing a feature point as center. Then the circular images are transformed from Cartesian co-ordinates to Log-Polar co-ordinates. Next, the LPT images are dealt using phase correlation. Experimental results demonstrate the reliability and rotation invariance of the proposed method.

IndexTerms—Image registration, Log-Polar Transformation (LPT), Fast Fourier Transform (FFT), Phase correlation.

I. INTRODUCTION

Image Registration is an essential step in most image processing tasks. It is widely used in remote sensing, weather forecasting, medical imaging and so on. Image Registration is the process of overlaying two or more images that have the same scene taken at different image sensors, different times, and different geometric viewpoints, or share the same visual information such as the same objects. The main goal of image registration is to find the transformation between the reference image and the sensed images. However, it could not design a universal registration system to solve all kinds of tasks depending on the different type of image and various types of degradations.

A. Image Registration Process

Image registration process involves the following four steps. First step is Feature Extraction which consists of feature construction and feature selection of image. The feature construction and feature selection gives feature points (key points) in the image using one of the corner detector methods. Each feature point is unique. Second step is Feature Matching which establish correspondence between the reference and the sensed images. It gives the best matching point pairs between the images using similarity measures. Third step is Homography Estimation, based on the matching points obtained from feature matching estimate the parameter of the mapping function. In order to get the parameters, homography estimation uses only the true matching points. If the wrong matching points are used then the images will be incorrectly aligned. Fourth step is Transformation, after estimating the homography parameters, the sensed image is transformed using any one of transformation technique. The transformation consists translation, rotation and scaling. The existing registration methods are firstly, taking feature points over the images to correct the geometric errors such as translation, rotation and scaling, and secondly, establishing the correspondence model between images. Phase correlation, based on the translation property of the Fourier transform, computes the FFT ratio instead of any feature points. The sharp peak of the inverse of this ratio represents the highest correlation. Before phase correlation, the reference and sensed images are transformed from the Cartesian coordinates to the log-polar image and then phase correlation computes the similar ratio, rotation and scaling errors which are represented as shift. The matching is establish by using these parameters. The existing registration methods use the whole image to do log-polar transformation and phase correlation. However, in the common log-polar transformation for image registration is to crop a square patch in both the reference image and the sensed image. After the image has been transformed from Cartesian coordinates to Log-Polar coordinates, black semicircles appear on the border of the image, overlapping to form sharp edges / corners in between. Phase correlation focuses on these corners at the black semicircles. There is no information at these black semicircles, which reduces the accuracy of the image registration. Fig 1 shows a sample of previous registration methods. In order to improve the accuracy of feature-based registration for aerial image registration requires a high degree of precision. This paper proposes a novel Log-Polar Transform (LPT) based image registration. Here instead of using the whole image in the existing method,
feature points are used, which reduces the computational
time.

![Image](image1.png)

**Fig 1: The previous Image registration method**

Furthermore, the black semicircles also need to be
avoided. The proposed process contains four steps. First,
feature points are selected in both the reference image
and the sensed image with corner detector (Harris or
SIFT). Secondly, image patches are created using feature
point positions as centers. Each point is a center point of
LPT, so circular image patches are cropped choosing
feature point as center. The radius of the circle can be
changed, then the circular images are transformed to Log-
Polar coordinates. Next, the LPT images are dealt with
using phase correlation. **Fig 2** shows the proposed
solution, which removes the black semicircles.

![Image](image2.png)

**Fig 2: The Proposed Image Registration Technique**

### II. FEATURE EXTRACTION

The feature extraction is the process of finding the feature
points since we are dealing with aerial images there is
some problem with these, for example old aerial photos
may appear yellow due to decay or general wear and tear,
a better stitching result will not be achieved if the two
aerial images have different illuminations, even after
registration and it takes a long time to deal with the high
resolution images. In order to avoid these they should be
preprocessed before going to used for feature extraction.
The preprocessing includes modifying the intensity and
removal of noise. The feature points are extracted from
the Harris corner detection method then go to Scale
Invariant Feature Transform (SIFT).

The Harris Corner Detector is one of the simplest corner
detector algorithm, proposed by Chris Harris and Mike
Stephens. It is based on the second moment matrix, called
the auto- correlation matrix. The main idea is to find the
interest points called corners where the surrounding
neighborhood shows edges in more than one direction.
The Harris Corner Detector algorithm has the following
steps:
1. Define a matrix $M$ for each pixel $(x,y)$ in the image
domain, as the positive semi-definite, symmetric matrix

$$M = \begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix}$$

2. Let $w$ be a weight matrix(a Gaussian filter), the
convolution

$$\overline{M} = w^* M$$

The width of $w$ determines the interesting region around
$x$. The result $\overline{M}$ is auto-correlation matrix, sometimes
called Harris matrix. It gives a averaging of $M$ over the
neighborhood pixels.
3. For each pixel $(x,y)$, construct an indicator function,
calculate the cornerness$(x,y)$.

$$\text{cornerness}(x,y) = \frac{\det(M) - k(\text{trace}(M))^2}{\det(M) + (\text{trace}(M))^2}$$

$k$ is a constant. A typical value for $k$ is $0.04$.
4. Set all cornerness$(x,y)$ below a threshold $T$ to zero.
Threshold $T$ controls the number of corners.

The scale invariant feature transform (SIFT) algorithm
was developed by David G. Lowe [5] [6] [7]. The SIFT
features are invariant to image scale and rotation.
Because of the highly distinctive property of the SIFT
feature, a single feature can be correctly matched with
high probability against a large number of features[7].
SIFT feature detection works in four consecutive steps.

The first step is scale-space extrema detection, used to
get the potential feature points which are invariant to
scale and orientation. This step is implemented by using
the extrema of Difference of Gaussians(DOG) to search
over all scales and image locations. First, it establishes
the scale-space with Gaussian function. The scale space
can be written as

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

Where $*$ is the convolution operation, $I(x,y)$ is an input
image and

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

is the Gaussian function in 2D situation.

While $\sigma$ needs to be assigned a different value to
construct $n$ scaled image in its first scale space. The DOG
image is obtained by the Difference of Gaussian function
convolved with the image $I(x,y)$. It can be described as

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

which is equal to

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$$

where $k = \sqrt{2}$.

The size of each octave scale space is half the previous
one. **Fig 3** shows a set of scale space images are produced
by the initial image convolved with Gaussian repeatedly
on the left and adjacent Gaussian images are subtracted to produce the difference of Gaussian (DOG) on the right.

Secondly, it detects the local extrema (maxima or minima) of the DOG image. In Fig 4, each pixel is compared against 26 neighbour points \((8+2\times9=26)\), 8 neighbours in the current scale space image, 9 neighbours in the scale above and 9 neighbours in the scale below. The pixel is selected only if its value larger or smaller than all of these neighbors.

The second step is keypoint localization. The feature points are obtained from local extrema (maxima and minima) of DOG. For stability, it should reject the points whose locations are along the edge. Same as the Harris Corner Detector, it computes Hessian matrix and checks that the ratio of principal curvatures is under some threshold, \(r\). We need to check

\[
\frac{\text{trace}(M)^2}{\text{det}(M)} < \left(\frac{r+1}{r}\right)^2
\]

The third step is orientation assignment. The keypoint descriptor can tolerate image rotation by being assigned an orientation. The following method can provide the most stable results. For each image sample, \(L(x,y)\) (Gaussian smoothed image), the gradient of each key point should be calculated. The magnitude \(m(x,y)\) and orientation \(\theta(x,y)\) of the gradient is computed respectively as

\[
m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}
\]

\[
\theta(x,y) = \tan^{-1}\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}
\]

An orientation histogram is formed from the gradient orientations of sample points within a region around the keypoint according to Lowe’s suggestion [19], which has 36 bins covering the 360 degree range of orientations. The highest peak in the histogram is detected later any other local peak with 80% of the peak is used to also create a keypoint with that orientation. Because of locations with multiple peaks of similar magnitude, there will be multiple keypoints created at the same location and scale but with a different orientation.

The fourth step is keypoint descriptor, an image location, scale and orientation of each keypoint are defined from the previous steps. This step is to compute a descriptor which is highly distinctive and invariant to 3D viewpoint or changes in illumination. Fig 5 illustrates how to create a keypoint descriptor. Here a 16x16 Gaussian window is selected around the keypoint location and broken it into 4x4 subwindows. It calculates the magnitude \(m(x,y)\) and orientation \(\theta(x,y)\) of the gradient at each point in each 4x4 subwindow, then these orientations are put into an 8 bin histogram which contains 8 features in each subwindow. In this case, the 16x16 window is used to create a keypoint, so that each keypoint has 128 feature vectors (16x8 = 128).

III. FEATURE MATCHING

Based on the feature points extracted from the previous chapter, we will start feature matching. Original feature matching process establishes correspondence between the reference image and the sensed image. The proposed method is matching the correspondence between the circular image patches in both images. The circular image patches with feature point as center from square image patches are obtained by using Log-Polar Transform (LPT), a nonlinear sampling method which converts an image from the Cartesian coordinates \(I(x,y)\) to the log-polar coordinates \(\text{LPT}(p, \theta)\). The mapping function is given by.
The frequency domain represents the image \( f(x,y) \) of size \( M \times N \) in the spatial domain. It is the sampled Fourier Transform, so it does not contain all the frequencies which are forming an image, but the set of samples is large enough to fully describe the spatial domain image. The Discrete Fourier Transform (DFT) is computed as follows:

\[
F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)e^{-j2\pi(ux/M + vy/N)}
\]

The concept behind the Fourier transform is using a sum of sine and cosine waves of different frequencies to construct any waveform. The frequencies are determined by expanding the exponential in the above equation sine and cosine with variables \( u \) and \( v \). The inverse discrete Fourier transform of the above equation is

\[
f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v)e^{j2\pi(ux/M + vy/N)}
\]

So, if \( F(u,v) \) is known, using above equation corresponding image \( f(x,y) \) will obtain.

The registration of the translated images are done by Phase Correlation method, which is based on the Fourier Shift property[1]. The shift obtained transforms as phase difference in the Fourier domain. The Phase Correlation algorithm can be described as follows:

(a) Given the two input images \( g_a \) and \( g_b \), then apply a window function (e.g. a Hanning window) on both the images to reduce edge effects. Next, calculate the discrete Fourier transform of both images.

\[
G_a = F\{g_a\}, \quad G_b = F\{g_b\}
\]

The experimental results show that the images without a window function get better results in this case, so we can ignore this step.

(b) Normalize the cross-power spectrum, given by

\[
R = \frac{G_a^* G_b}{|G_a^* G_b|}
\]

where * is the complex conjugate.

(c) Apply the inverse Fourier transform to get the normalized crosscorrelation.

\[
r = F^{-1}\{R\}
\]

(d) Look for the location of the peak in its inverse Fourier transform \( r \).

\[
(\Delta x, \Delta y) = \text{argmax}_{(x,y)}|r|
\]

The false matches are removed using the symmetry test in which the false match (i.e. the match pair could not pass the symmetry test in both directions) are removed by computing the matches using the symmetrical (two-sided) matching conditions and keep the accurate match.

**IV. HOMOGRAPHY ESTIMATION**

The Homography is a method of mapping points and lines in one plane to another plane. The estimation of homographies between images using point feature correspondences by common algorithm. According to the definition of homography from Hartley and Zisserman [4], homography matrix \( H \) of size \( 3 \times 3 \) is sufficient to calculate the homography. The homography matrix \( H \) has 8 independent degrees.
The algorithms for Homography Estimation includes the basic Direct Linear Transform (DLT) algorithm which is an algorithm for computing the homography matrix \( H \) if it gives four or more point correspondences. The relationship between corresponding points \( x \) and \( x' \) If the 2D points are mapping in homogeneous coordinates, described as

\[
\begin{bmatrix}
x' \\
y' \\
w'
\end{bmatrix} =
\begin{bmatrix}
h_1 & h_2 & h_3 & 0 \\
h_4 & h_5 & h_6 & 0 \\
h_7 & h_8 & h_9 & 0
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
w
\end{bmatrix}
\]

In homogeneous coordinates the points are scaled, so the 2D point can be expressed as \( x = [x, y, w] = [x/w, y/w, 1] \). To get a unique identification of the image coordinates \( x, y \), the points are often normalized with \( w=1 \). In order to obtain the correct result, Hartley and Zisserman [4] propose a normalization step on DLT. when there are more than 4 points of correspondence available minimization of a suitable cost function should be considered to solve for a homography.

The homography estimation algorithm requires a number of correspondences at least 4. However, in some situations the two features in the images do not correspond to the same real feature. Therefore we need to distinguish inlier, the true matching correspondence, is used for homography estimation and outlier, the wrong matching correspondence, is not considered for homography estimation. Only inlier matches can be used for homography estimation. We will discuss two homography estimation methods, one is RANSAC (Random Sample Consensus) is the most popular robust homography estimation method [2]. The method can be described as first, it selects 4 correspondences at random and computes a homography matrix \( H \). Secondly, it classifies the other correspondences as inliers or outliers. If the distance > \( t \), the point is considered an outlier. We also need to decide the number of iterations for running RANSAC. Hartley and Zisserman [4] suggest the formula that

\[
N = \log(1-p)/\log(1 - (1 - \varepsilon)^s)
\]

where \( \varepsilon \) is the probability of outliers and \( s \) is the number of correspondences used in each iteration (\( s = 4 \) in this case).

The distance threshold is chosen to find out outliers and discard them in RANSAC and it deals with the sum of squared difference algorithms. However, the Algebraic distance version of DLT is not very robust with outliers. The sum of squared residuals obtained in RANSAC replaced by the Least Median of Squares (LMS) estimation method. Peter J.Rousseeuw [9] states that LMS works very well if there are less than 50% outliers and it does not need to set any parameters such as the threshold or the probability of outliers. The major disadvantage of the LMS is that it could not deal with the outliers, which are more than half the data.

V. SYSTEM DESCRIPTION

The paper analyze the transformation relationship between the reference image and the sensed image. For obtaining a satisfying final result, preprocessing work is usually done to remove noise and modify intensity after selecting interesting aerial images. The corner detector method is used to obtain feature points in the reference image and the sensed image. Circle image patches are created using these feature points as centers.

In order to deal with rotation situation, circle shaped image patches are used. It can avoid the black semicircle shaped noise after log-polar transformation. Next, we transform all the image patches from the Cartesian coordinates \( I(x,y) \) to the log-polar coordinates \( \text{LPT}(\rho, \theta) \). Any rotation and scale in the Cartesian coordinates are represented as translation in the log-polar coordinates. Then Fast Fourier Transform is used to convert log-polar based images from the spatial domain into the frequency domain. Based on the Fourier Shift property, the Phase Correlation method is proposed for the registration of translated images. Symmetry test is used for removing false matches. A good matching result can give accurate transformation parameters. At last, RANSAC is used for homography estimation.

Some parameters will influence the accuracy of this proposed registration system, the radius of circular image patches and the size of input images of FFT are the important parameters considered in this paper. As shown in Fig 7, the radius of circular image patches decides the size of input image of LPT. It also shows that the size of input images of FFT affects the FFT. The output image of the LPT is the same one for input image of FFT. The Area of the input image of LPT should be bigger than the area of input image of FFT to get high accuracy.

\[
\text{Fig 7: The parameters that affect registration are radius and size}
\]
The pixel of the center point is removed for avoiding oversampling in LPT. Fig 8 shows in detail when the size of square subimage is \(m \times m\), the radius of the largest circle image patch inside is \(m = 2 \times \text{radius} + 1\).

![Fig 8: The radius of circular image patch](image)

The best value for the radius obtained by testing the radius of the circle image patches to get more accuracy of registration. The size of input image of LPT also will be tested. Then two methods are compared, first Log-polar then FFT or first FFT then Log-polar.

VI. EXPERIMENTAL RESULTS

The algorithm is developed using MATLAB. All of the images that have been tested are screenshots from Google Earth. The resolution of these images are 418*320. Irrespective of registration method and the particular images used, the most important thing for the user is an estimate how accurate the registration is. The accuracy evaluation is not a normal problem, because the errors can be dragged into each of stages in the registration process. In order to evaluate the accuracy of these experimental results, the homography with RANSAC was estimated ten times and the average number of inliers was calculated. If the same results were obtained ten times in each actual experiment, it proved that the proposed method is stable and reliable.

The proposed registration method is LPT PC (Log-Polar Transform and Phase correlation). At first, the experiment will test different parameters and select the optimal parameters for the proposed algorithm. Later shows the experimental results of feature matching applying first log-polar then FFT and first FFT then log-polar. Then it shows the comparison of number of inliers in the proposed algorithm LPT PC using SIFT selecting feature points with the original SIFT.

A. Results of Optimal Parameter Selection

The optimal parameters selected from the parameters considering the radius of the circle image patch in LPT, as well as the input size of FFT. Then these optimal parameters will be used to registration. Here 400 feature points are extracted using SIFT in the reference images and the sensed images. Because the interesting square is extracted from the original image, using feature points as centers, when the square is 41x41, the biggest radius should be 20 shown Fig 9(a). In order to get high precision, the input size of FFT should be smaller than the size of the interesting square. Furthermore, the input size of FFT should be \(2n(n = 1, 2, 3, 4,...)\). When the square is 41x41, the biggest input size of FFT is 32x32.

![Fig 9: The principle of selecting parameters.](image)

The input size of FFT 64x64 needs to be tested, so the size of the interesting square should be bigger than 64x64 then 71x71 is chosen for the input size of FFT 64x64 shown in Fig 9(b). The radius of the circle image patch in LPT are 10, 12, 14, 16, 18, 20 and the input sizes of FFT are 8x8, 16x16, 32x32, 64x64 are taken for the test. 10 sets of images have been tested. Fig 10 shows one sample of test results.

![Fig 10: Chart showing the number inliers for different parameters](image)

B. Results of first LPT then FFT and first FFT then LPT

Here the two methods are compared, first LPT then FFT and first FFT then LPT. The proposed log-polar method is used to implement the first LPT then FFT method. In both the methods SIFT is chosen to extract the feature points then the optimal parameters are used to complete this experiment. The lines are used to connect sets of feature points which have top high phrase correlation. Looking at the matching connection in Fig 11(a) and (c), there are some false matching connections between the reference image and the sensed image in first FFT then LPT method. Fig 11 (b) and (d) shows all the match connections are correct in first LPT then FFT.

C. Registration Results

Results of the proposed registration method LPT PC with SIFT feature detector and original SIFT are compared here. The SIFT is selected to extract the feature points.
Next, the first LPT then FFT method are chosen, which is compared with Lowe’s SIFT. Both methods select 400 feature points from the original images. After homography estimation, the number of inliers are compared.

At the beginning, in both registration methods the different angle rotated images are tested. In this experiment there are $45^\circ$, $90^\circ$, $135^\circ$, $180^\circ$, $225^\circ$, $270^\circ$ and $315^\circ$ rotation images. The results of inliers have been presented in the chart shown in Fig 12. As shown in Fig 12, the original SIFT get the number of inliers above 180, the highest number of inliers is 205. The number of inliers in the LPT PC are range in 143 to 181. Fig 13 shows 4 sets of images with random angle degrees. It also proves the proposed method can handle rotation images.

At last, the number of inliers have been compared in the different type of images which are Brightened image, Rotated image ($45$ degree), Scaled image and Noisy image. The chart in Fig 14 shows that the LPT PC has high ability for dealing with the Noisy image which got 196 inliers, while the original SIFT gives low number of inliers 17. For rotated image, the inliers in the original SIFT is 198 and in the LPT PC is 181. For the Brighten image, the inliers in the original SIFT is 165 and in the LPT PC is 234, it means that the LPT PC works better. For scale image, the original SIFT performs better than the LPT PC.

CONCLUSION

This paper contains an overview of image registration methods and proposes a novel registration method which can be used for 2D aerial images. In order to improve the accuracy of registration, it selected the feature points with corner detector and created image patches using these feature points as centers, instead of using the whole image in the conventional method, which reduces the computational time. For rotation invariant, it is not important how the image patch is rotated, the key is focusing on the feature points. Each point was a center point of LPT, so the circular image patches were cropped.
over the center point. Then it transformed the circular images to Log-Polar coordinates, which can avoid black semicircles from appearing on the border of the LPT image. Next, it coped with the LPT images with phase correlation. Due to the radius of the circle image patches can be changed. After testing, the best parameters have been found (the radius for LPT as 20, the input size of FFT as 32x32) for this novel method.

Image registration with FFT always do the FFT first. But in order to avoid the black semicircles, applied LPT first to get the circular region and then the FFT. Taken the comparison first LPT then FFT and first FFT then LPT methods. The number of inliers for different rotation degree, the number of inliers for four different images are rotated in random angle degree and the number of inliers for different types of images with noise, rotation, high intensity and scaling are compared for proposed LPT PC and original SIFT. According to the comparison, the LPT PC works better.

ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who gave me the possibility to complete this paper. I wish to express my deepest gratitude to Mr. O. Venkata Krishna, Sr. Assistant Professor for his genuine support, valuable advice and sincere comments which helped me a lot in finishing this paper. His excellent scientific expertise, in combination with his friendly and helpful character has made him an outstanding academic teacher.

REFERENCES


BIOGRAPHIES

Mr. Bitra Janardhana Rao was born in 1984. He received his B.Tech in ECE from Nagarjuna University, Guntur in 2006. He has completed his Master of Technology in Communication and Radar Systems from KL University in 2010. He is currently working as an Assistant professor in CVR College Of Engineering, Hyderabad. His research interest includes Signal Processing and Image processing.

Mr. O. Venaka Krishna was born in 1983. He received his B.Tech in ECE from Nagarjuna University, Guntur in 2004. He has completed his Master of Technology in VLSI System Design from JNTU, Hyderabad in 2009. He is currently working as an Sr. Assistant professor in CVR College Of Engineering, Hyderabad. His research interest includes VLSI Signal Processing.