AERIAL IMAGE REGISTRATION USING PROJECTIVE POLAR TRANSFORM

Rittavee Matungka and Yuan F. Zheng

Dept. of Electrical and Computer Engineering
The Ohio State University
Columbus, Ohio 43212
Emails:{rmatungka,zheng}@ece.osu.edu

Robert L. Ewing

Air Force Research Laboratory/IFTA
Wright Patterson AFB, Ohio 45433

ABSTRACT

Image registration is an essential step in many image processing applications that need visual information from multiple images for comparison, integration or analysis. Recently researchers have introduced image registration techniques using the log-polar transform (LPT) for its rotation and scale invariant properties. However, the accuracy of the approach limits by the number of samples used in the mapping process, which affects directly the computational cost. Motivated by the success of LPT based approach and its limitation, we propose a novel Projective Polar Transform (PPT) based image registration method that is robust to translation, scale, and rotation and yields high accuracy while requires low computational cost. Unlike LPT that 2D interpolation is needed in the mapping process, our method uses one-to-one mapping mechanism that directly arranges image from Cartesian to Polar coordinate according to pre-computed PPT map. An innovative projection mechanism is proposed to reduce the image from two to one dimensional vectors. With the proposed approach, the mapping procedure is accelerated and the matching process can be performed in 1D. This results in great reduction of the computational cost.

Index Terms— Aerial image registration, projective polar transform.

1. INTRODUCTION

Image registration is a process of aligning two images that share common visual information such as images of the same object or images of the same scene taken at different geometric viewpoints, different time, or by different image sensors. Image registration is an essential step in many image processing applications that involve multiple images for comparison, integration or analysis such as image fusion, image mosaics, image or scene change detection, and medical imaging. The main objective of image registration is to find the geometric transformations of the *model image*, I^M, in the *target image*, I^T, where $I^T(x,y) = \mathcal{T}\{I^M(x',y')\}$ and $\mathcal{T}$ is a two-dimensional geometric transformation that associates the (x',y') coordinates in I^M with the (x,y) coordinates in I^T. These two-dimensional geometric transformations include scale, rotation, and translation in the Cartesian coordinates. Many works have been done in this area in the past 20 years [1], which can be categorized into two major groups: the *feature-based approach* and the *area-based approach*. The feature-based approach uses only the correspondence between the features in the two images for registration. The features can be color gradient, edges, geometric shape and contour, image skeleton, or feature points. Since only the features are involved in the registration, the feature-based approach has advantages in registering images that are subjected to alteration or occlusion. However, the use of the feature-based approach is recommended only when the images contain enough distinctive features [2]. As a result, for some applications such as aerial imaging, in which features are difficult to be distinguished from one another (for example, shapes of most buildings from the bird-eye view are almost identical), the feature-based approach may not perform effectively. This problem can be overcome by the area-based approach, in which part of the images are used for comparison using correlation-like approaches such as cross-correlation and Fourier-based phase correlation.

Innovative area-based approach using *Log-Polar transform* (LPT) method have been proposed [3–7] and proven to be effective and robust to changes in translation, scale, and rotation. However, there is limitation to LPT based approach. The accuracy of the registration depends directly on the number of samples used in the mapping process. This problem can be easily illustrated, for example, if 36 samples in the angular direction are used, the level of accuracy of the registration can not exceed 10 degrees resolution. For aerial image registration, in which scale and rotation parameters between two images may be very small, large number of samples in both radius and angular directions are required. This increases the computational load tremendously from both the mapping process, which require 2D interpolation of image pixels in the Cartesian, and the matching process that involves computation of 2D correlation.

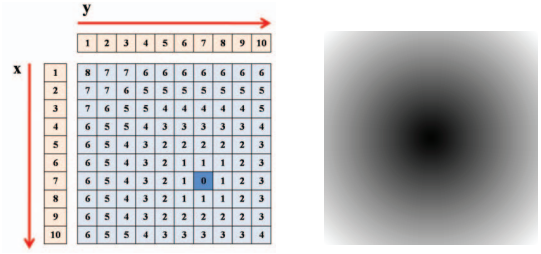
Motivated by the success of LPT based approach and its limitation, we propose a novel *Projective Polar Transform* (PPT) based image registration method that is robust to translation, scale, and rotation and yields high accuracy while requires low computational cost. Unlike LPT that 2D interpolation is needed in the mapping process, our method uses one-to-one mapping mechanism that directly arranges image from Cartesian to Polar coordinate according to pre-computed PPT map. An innovative projection mechanism is proposed to reduce the image from two to one dimensional vectors. With the proposed approach, the mapping procedure is accelerated and the matching process can be performed in 1D. This results in great reduction of the computational cost.

The paper is organized as follows. In section 2, we discuss our proposed registration approach in detail. Experimental results are shown in section 3. Finally, our work is summarized in section 4.

2. PROPOSED ALGORITHM

2.1. Projective Polar Transforms

Inspired by the scale and rotation invariance properties of LPT based approach, we propose a novel image transformation scheme, PPT,



(a) (b)

Fig. 1. PPT map (a) Distance parameters for image of size 10×10 pixels and $(x_c, y_c) = (7, 7)$, (b) Graphic demonstration of PPT map for image with size 513×513 pixels and $(x_c, y_c) = (256, 256)$

that is robust to translation, scale, and rotation and yields high accuracy while requires low computational cost.

2.1.1. PPT mapping

To transform a circular area with the size R_{max} inside image $I(x, y)$ in the Cartesian coordinate to projective polar, we first need to create a PPT map. We define a distance parameter $D_{(x_c,y_c)}(x, y)$ as

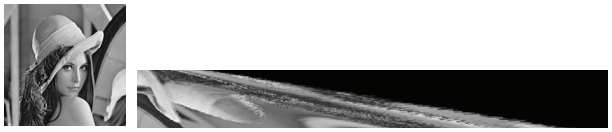
$$D_{(x_c,y_c)}(x, y) = \max\{n \in \mathbb{Z}^+ | n \leq \sqrt{(x - x_c)^2 + (y - y_c)^2}\} \quad (1)$$

where (x_c, y_c) is the image pixel in Cartesian selected to be the center of the transformation. Fig. 1(a) shows example of distance parameters of image with the size 10×10 pixels where the center point $(x_c, y_c) = (7, 7)$. In each pixel of the image, the calculated distance parameter is filled. Fig. 1(b) is the graphic demonstration of PPT map for image with the size 513×513 pixels where the center of the transformation $(x_c, y_c) = (256, 256)$. As shown in Fig. 1(b), PPT map is an approximation of Polar transform mapping.

To transform image, we directly map the image pixels in the Cartesian coordinate $I(x, y)$ that have distance parameter $D_{(x_c,y_c)}(x, y) = i$ to the transformed coordinate $IP(i, \theta)$. To approximate sampling in both radius and angular directions, for each iteration i, the mapping begins with $I(x_c, y_c + i)$ and continue the search for image pixels with $D_{(x_c,y_c)}(x, y) = i$ pixel-by-pixel in the counter-clockwise direction. The process is repeated for $i = 1, .., R_{max}$. Example of the transformed *Lena* image is shown in Fig. 2. Since PPT map is static and does not required to be computed for every transformation, and the mapping process is a simple one-to-one allocation from Cartesian to PPT images and does not require interpolation, the computational cost in this step is very low compared to that of LPT.

2.1.2. Projection transform

As shown in Figs. 2(b), the result from the transformation is a series of sample bins arranged in the step-like manner which do not show coordinate shifts for scaled and rotated image as in LPT any longer. To maintain the advantages of scale and rotation invariance in LPT, we use an innovative *projection transform* method which projects the two-dimensional image on the radius and angular coordinates, respectively. From the two projections, we can accurately calculate the scale and rotation parameters, for which the details will be elaborated in Section 2.2. For now we define the projection transform for the APT transformed image.



(a) (b)

Fig. 2. Example of PPT: (a) the original *Lena* image, (b) the PPT transformed image of (a)

Given a transformed image $IP(r, \theta)$ that consists of n_r sample bins ($n_r = R_{max}$), in which each bin has the length of n_{θ_i} for $i = 1, ..., n_r$. We denote $\hat{n_\theta}$, $\Re$ and Θ as the number of samples in the angular direction at $r_i = R_{max}$, the projection on the radius coordinate, and the projection on the angular coordinate, respectively. The mathematical expressions of $\Re$ and Θ are as follow:

$$\Re(i) = \Omega_i \sum_{j=1}^{n_{\theta_i}} IP(i, j), \quad (2)$$

$$\Theta(j) = \sum_{i=1}^{n_r} [\eta_{ij}^1 IP(i, ceil(\frac{j-1}{\Omega_i})) + \eta_{ij}^2 IP(i, ceil(\frac{j}{\Omega_i}))], \quad (3)$$

$$\begin{cases} i \in [1, ..., n_r], \quad j \in [1, ..., \hat{n_\theta}], \quad \Omega_i = \frac{\hat{n_\theta}}{n_{\theta_i}} \\ \eta_{ij}^1 = \Omega_i(j-1) - floor[\Omega_i(j-1)], \quad \eta_{ij}^2 = 1 - \eta_{ij}^1 \\ IP(i, 0) = 0, \forall i. \end{cases} \quad (4)$$

The operation $floor(A)$ denotes the nearest integers less than or equal to A. The results of the projection transform, vectors $\Re$ and Θ will have the dimension of n_r and $\hat{n_\theta}$, respectively. Both projections $\Re$ and Θ are normalized to reduce the effect of illumination changes. The computation of projection $\Re$ is simple and does not require interpolation. The computation of projection Θ, on the other hand, requires one-dimensional interpolation, as shown in Eq. (3).

Examples of the projections of APT of the *Lena* image and its scaled and rotated version are shown in Fig. 3 (scale =1.2 and rotation =45 degrees). It can be seen that the scale change in the Cartesian appears as the variable-scale in the projection $\Re$ (i.e. from $f(t)$ to $f(at)$) and the rotation change in the Cartesian appears as shifting in the projection Θ.

2.2. PPT Matching

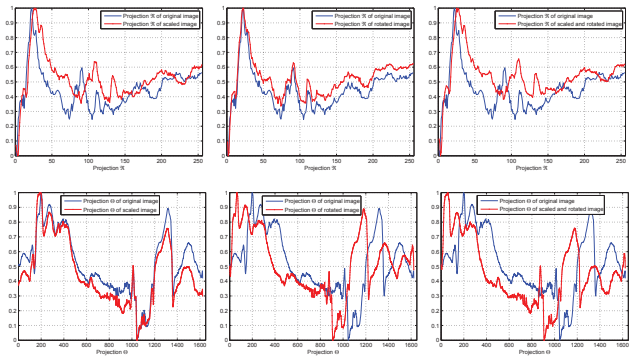
Given a model image $I^M(x, y)$, a target image $I^T(x, y)$ that is a scaled and rotated version of the model image with scale parameter a and rotation parameter $\hat{\theta}$ can be expressed mathematically as

$$I^T(x, y) = I^M(ax\cos\hat{\theta} + ay\sin\hat{\theta}, -ax\sin\hat{\theta} + ay\cos\hat{\theta}). \quad (5)$$

In the polar domain, Eq. (5) can be expressed as:

$$I^T(r, \theta) = I^M(ar, \theta - \hat{\theta}). \quad (6)$$

Since the projections $\Re$ and Θ are derived from PPT, which is an approximation of Polar transform, we can obtain scale and rotation parameters between two images by comparing their projections. As shown in Figs. 3(a) and 3(c), the projections Θ of the scaled images in the Cartesian are slightly altered when compared with that of the original image. This is because the areas covered during the PPT transformations are different as a result of the scaling. Hence, in order to accurately obtain the shifting parameter between the two projections Θ^M and Θ^T, the variable-scale parameter a between the two projections $\Re^M$ and $\Re^T$ needs to be obtained first.



(a) (b) (c)

Fig. 3. The effects of the changes in scale and rotation in the Cartesian coordinates to the projections $\Re$ and Θ: (a) the scale change in the Cartesian appears as variable-scale in the projection $\Re$, while the projection Θ becomes slightly altered, (b) the rotation in the Cartesian appears as shifting in the projection Θ, while there is no change in the projection $\Re$, and (c) the changes in both scale and rotation in the Cartesian appear as variable-scale in projection $\Re$ and shifting in projection Θ, respectively

2.2.1. Find scale parameter

There are several ways to obtain the scale parameter from the projections depending on the requirements of the application in term of the computational cost, accuracy, image types and environments. We introduce here two effective algorithms.

- **Algorithm 1: Logarithm method**

  This algorithm uses the scale invariant property of the logarithm function. First, the logarithm function is applied to the projection $\Re$ and the output is then quantized to maintain the original dimension of the projection. The mathematical expression of the implementation is as follow:

$$\mathcal{L}\Re(k) = \Re(n_r \frac{\log k}{\log n_r}); \quad k = 1, ..., n_r. \tag{7}$$

  The parameter $\mathcal{L}\Re(.)$ denotes the logarithmic of the projection $\Re$. Given image I^T a scaled and rotated version of image I^M, the scale parameter a between the two images would appear as translation in the logarithm domain:

$$\mathcal{L}\Re^T(\rho) = \mathcal{L}\Re^M(\rho) + \log a. \tag{8}$$

  To find the displacement d, where $d = \log a$, such that $\mathcal{L}\Re^T(\rho) = \mathcal{L}\Re^M(\rho - d)$, one can evaluate the correlation function between the two logarithmic of projections $\mathcal{C}(\mathcal{L}\Re^M, \mathcal{L}\Re_z^T)$:

$$d = \arg\max \mathcal{C}(\mathcal{L}\Re^M, \mathcal{L}\Re^T). \tag{9}$$

- **Algorithm 2: Resampling projection**

  Algorithms 1 is fast but can yield accuracy at some degrees, since the displacement estimation can be computed with only integer accuracy. For aerial image registration purposes, in which high precision is required, one may include resampling method as an extra step after acquiring the estimated scale parameter with algorithm 1, denoted as a', (or can only use resampling method if estimated scale parameter is known or expected to be small). Given the lower limit a_L and the upper limit a_U of the search space for scale parameter such that $a_L < a' < a_U$, respectively, and given the number of the search steps equals to h, we resample $\Re_{res}^M = \Phi(\Re^M, a_i)$ where $a_i = a_L + \frac{i(a_U - a_L)}{h}$ for $i = 1, ..h$. We use MSE method to compare each resampled projections. The operation $y = \Phi(x, a)$ is a resampling procedure. This resampling procedure is a common operation in signal processing for computing the sequence in vector x at a times the original sampling rate by using a polyphase filter implementation. Operation Φ applies an anti-aliasing linear-phase (low-pass) FIR filter to x during the resampling process. The result will have the dimension of vector equal to $length(y) = length(x) \times a$. The scale parameter is equal to the resampling parameter a_i that yields the lowest MSE, $E_{\Re}$. It is obvious that the larger parameter h is, the higher accuracy would be obtained.

2.2.2. Find rotation parameter

After the scale parameter is obtained, the next step is to find the rotation parameter $\hat{\theta}$. For the same sampling radius, R_{max}, the larger the image ($a >> 1$), the smaller the area of the scene is covered in the sampling procedure. As a result, the magnitude of the projections $\hat{\theta}$ between the two images could be slightly altered. This phenomenon is illustrated in Figs. 3(a) and 3(c). Hence, to accurately obtain the rotation parameter $\hat{\theta}$, the upper limit of Eq. (3) needs to be modified according to a. If $a > 1$, the upper limit of Θ^M is modified to an_r, while the upper limit of computing Θ^T remains unchanged. If $a < 1$, the upper limit of Θ^T is modified to n_r/a, while the upper limit of computing Θ^M remains unchanged. Both projections are then resampled to be equal in length before comparison. We denoted the modified projections as $\hat{\Theta}^M$ and $\hat{\Theta}^T$. The rotation parameter can be found by evaluating the correlation function:

$$\hat{d} = \arg\max \mathcal{C}(\hat{\Theta}^M, \hat{\Theta}^T), \quad \hat{\theta} = \frac{2\pi\hat{d}}{\hat{n_\theta}}. \tag{10}$$

2.2.3. Find distance coefficient

The final and crucial component resulting from PPT matching is the distance coefficient, denoted as $\mathcal{E}$. This distance coefficient indicates how large the difference between the two images is. The distance coefficient $\mathcal{E}$ between the model image I^M and the target image I^T can be computed from the Euclidean distance between the projections $\hat{\Theta}^M$ and $\hat{\Theta}^T$ as $\mathcal{E} = \sqrt{\sum_{i=1}^{\hat{n_\theta}} [\hat{\Theta}^T(i) - \hat{\Theta}^M(i - \hat{d})]^2}$.

2.3. Feature Point Based Search Scheme

Similar to LPT based approach, the translation parameter between model image and target image has to be found in order to keep the scale and rotation invariant properties. We adopt Gabor feature extraction to accelerate the search process (more detail of Gabor feature extraction can be found in author's publication [4]). These feature points are obtained by applying Gabor transform to the image and selecting those pixels which have high energy in the wavelet domain. By selecting one of the feature points in the model image as a center point for PPT, we reduce the search from every pixel of the target image to a set of feature points found in the target image

only. The matched feature point is the point that yields the lowest distance coefficient $\mathcal{E}$. Apparently, the number of feature points is much smaller than that of the pixels in the target image, while the computation of the Gabor wavelet transform is much lower than that of PPT. Thus, the computation load is much lighter than the exhaustive search using PPT.

2.4. Complete algorithm

Below is the complete steps of the proposed algorithm:

- Extract feature points in both the model image and the target image.
- Crop circular image patch I^M that covers the area R_{max} desired to be registered to the target image.
- Compute the projections $\Re^M$ and Θ^M using the proposed PPT approach and the projection transform.
- Use each feature point in the target image p_z^T for $z = 1 : n_T$, where n_T is the number of feature points in the target image as the origin and crop a circular image patches I_z^T for $z = 1 : n_T$ with the radius size R_{max}.
- Compute the sets of candidate projections $\Re^T = \{\Re_1^T, .., \Re_{n_T}^T\}$ and $\Theta^T = \{\Theta_1^T, .., \Theta_{n_T}^T\}$.
- Match each candidate with the model using the proposed PPT matching algorithm. Translation is the point (x, y) that yield the lowest distance coefficient $\mathcal{E}$. The scale and rotation parameters are obtained simultaneously in this step.

3. EXPERIMENTAL RESULTS

In this section, the performance of the proposed image registration approach is evaluated. We uses four sets of aerial images in the test as shown in Fig 4. All test images have the size of 334×502 pixels. Since the scale parameters of all the test aerial images are very small, we adopt Algorithm 2 of the proposed PPT matching with parameters $h = 20, a_L = 0.9$ and $a_U = 1.1$ to obtain the scale parameters. The size of the search space using Gabor feature point varies depending on the richness of details in the image, with average of 120 - 150 points, which considers as approximately 0.1% of the total number of pixels in the image. Figs. 4(a) and 4(b) represent the model image and the target image, respectively. In Fig. 4(a) the red rectangular area indicates the image patch selected to be registered to the target image. We note that rectangular shape is used in the figure instead of circle to represent the orientation of the images. In Fig. 4(b), the red rectangular represents the registration result. To demonstrate the effectiveness and accuracy of the proposed method, we create mosaic images, in which the model images are scaled and rotated according to the parameters found and superimposed onto the target images. As shown in Fig. 4(c), one can see that images are registered accurately and naturally. This demonstrates the robustness to translation, scale and rotation of the proposed approach.

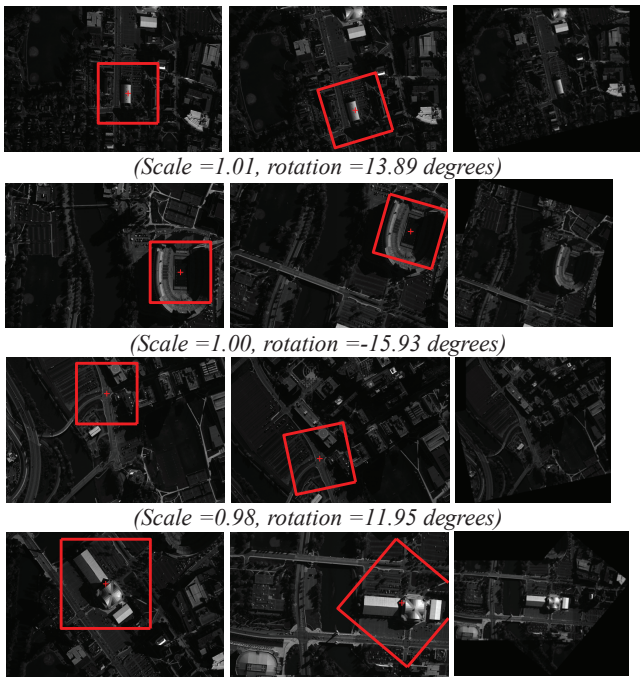


Fig. 4. Experimental results (a) model images, (b) target images and registration results, and (c) mosaic images

4. CONCLUSION

In this paper, we have presented a new image registration approach using the projective polar transform. With the combination of one-to-one mapping in the image transform process and the innovative projection approach that reduces the dimension of the transformed image from two to one, the computational load in both mapping and matching process is very low. Two effective 1D matching mechanisms are proposed to effectively and accurately obtain scale and rotation parameters of the registered images. The experiments with aerial images are provided to demonstrate the accuracy and robustness of our approach.

5. REFERENCES

[1] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, Dec 1992.

[2] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, Oct 2003.

[3] S. Zokai and G. Wolberg, "Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations," *IEEE Trans. Image Processing*, vol. 14, no. 10, pp. 1422–1434, Oct 2005.

[4] R. Matungka, Y. F. Zheng, and R. L. Ewing, "2D invariant object recognition using log-polar transform," in *Proc. World Congress on Intelligent Control and Automation*, Jun 2008, pp. 223–228.

[5] V. J. Traver and F. Pla, "Dealing with 2D translation estimation in log-polar imagery," *Image and Vision Computing*, vol. 21, no. 2, pp. 145–160, Feb 2003.

[6] V. J. Traver, A. Bernardino, P. Moreno, and J. S. Victor, "Appearance-based object detection in space-variant images: a multi-model approach," in *Proc. International Conference on Image Analysis and Recognition*, Sep 2004, pp. 538–546.

[7] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Trans. Image Processing*, vol. 5, no. 8, pp. 1266–1271, Aug 1996.