Image Registration Using Adaptive Polar Transform

Rittavee Matungka, *Student Member, IEEE*, Yuan F. Zheng, *Fellow, IEEE*, and Robert L. Ewing, *Senior Member, IEEE*

***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, it suffers from nonuniform sampling which makes it not suitable for applications in which the registered images are altered or occluded. Inspired by LPT, this paper presents a new registration algorithm that addresses the problems of the conventional LPT while maintaining the robustness to scale and rotation. We introduce a novel adaptive polar transform (APT) technique that evenly and effectively samples the image in the Cartesian coordinates. Combining APT with an innovative projection transform along with a matching mechanism, the proposed method yields less computational load and more accurate registration than that of the conventional LPT. Translation between the registered images is recovered with the new search scheme using Gabor feature extraction to accelerate the localization procedure. Moreover an image comparison scheme is proposed for locating the area where the image pairs differ. Experiments on real images demonstrate the effectiveness and robustness of the proposed approach for registering images that are subjected to occlusion and alteration in addition to scale, rotation, and translation.**

***Index Terms*—Adaptive polar transform (APT), alteration, image registration, occlusion, rotation, scale, translation.**

I. 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 2-D geometric transformation that associates the (x',y') coordinates in I^M with the (x,y) coordinates in I^T. These 2-D geometric transformations include scale, rotation, and translation in the Cartesian coordinates. The model image is often presented in the form of image patch that is cropped from the *reference image*. For the case that the entire reference image is desired to be registered to the target image, the reference image is considered the model image.

Many image registration methods have been proposed in the past 20 years [1]–[8], 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. Recently, Lowe *et al.* [9] introduce the so called SIFT method that can extract image features that are invariant to illumination change, scale, and rotation. Another well know method using a combination of Harris corner detection [10] and Laplacian of Gaussian to extract features that are invariant to scale and rotation is proposed in [11]. 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 [12]. As a result, for some applications such as medical imaging, in which the images are not rich in detail and features are difficult to be distinguished from one another, the feature-based approach may not perform effectively. This problem can be overcome by the area-based approach.

The common area-based approach is the normalized cross-correlation [7]. Another correlation based technique which is more robust to noise and changes in the image intensity than the cross-correlation technique is the phase correlation [8], in which the normalized cross-power spectrum between the two images is computed in the frequency domain. Although the correlation approaches show successful results in registering images that yield translation in the Cartesian, they both fail in the case where there are changes in scale or rotation between the two images. Combining the phase correlation technique with the log-polar transform (LPT), the Fourier-Mellin [5], [13] approach is proposed as a breakthrough area-based method that yields invariant properties to translation, scale and rotation. However, recent studies [4] and [14] indicate that the Fourier-Mellin method is able to recover only a fair amount of rotation and scale. Moreover Fourier transform introduces a problem of border effect where rotation and scale affect the aliasing of the image. Another work in image registration that uses the advantage of LPT is proposed in [4]. Unlike the Fourier-Mellin approach in which matching and localization procedures are performed in the frequency domain, the registration method proposed by Zokai *et al.* is performed entirely in the spatial domain. The translation parameter

Manuscript received April 30, 2008; revised May 18, 2009. First published June 10, 2009; current version published September 10, 2009. This work was supported in part by the Air Force Research Laboratory, in part by the National Science Foundation under Grant IIS-0328802, and in part by the NFSC under Grant 60632040. The associate editor coordinating the review of this manuscript and approving it for publication was Peter C. Doerschuk.

R. Matungka is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: matungkr@ece.osu.edu; rmatungka@hotmail.com).

Y. F. Zheng is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA, and also with Shanghai Jiao Tong University, Shanghai 200240 China (e-mail: zheng@ece.osu.edu).

R. L. Ewing is with the Sensors Directorate, AFRL/RYRT, Wright-Patterson AFB, OH 45433-7334 USA (e-mail: robert.ewing@wpafb.af.mil).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

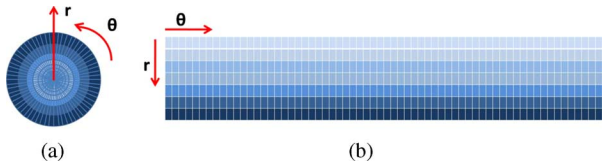
Digital Object Identifier 10.1109/TIP.2009.2025010

is recovered by using the coarse-to-fine multiresolution framework, while the scale and rotation parameters are obtained by matching the log-polar transformed images using the cross-correlation function.

LPT is a well known tool for image processing for its rotation and scale invariant properties [4], [15]–[23]. Scale and rotation in the Cartesian coordinates appear as translation or shifting in the log-polar domain. These invariant properties make LPT based image registration robust to scale and rotation. However, LPT suffers from the nonuniform sampling. One major problem of that is the high computational cost in the transformation process, which comes from the oversampling at the fovea in the spatial domain. Since no information is gained from oversampling, this computation is wasted. Another major problem of LPT is the bias matching. With LPT, the matching mechanism focuses only at the fovea or the area close to the center point of the transformation, while the peripheral area is given less consideration. Furthermore, occlusions and alterations between the two images may occur which is not considered by LPT. For example, the satellite images of the same location but taken at different times may contain occlusion due to climate change or cloud. In order to effectively register the two images, image registration method that is invariant to occlusion and alteration is needed.

Motivated by these observations of the conventional LPT, our work attempts to effectively register two images that are subjected to occlusion and alteration in addition to scale, rotation and translation. We introduce adaptive polar transform (APT) technique in the spatial domain that evenly samples the image. We further apply the projection transform to the transformed image to reduce the image from 2-D to 1-D vector. With APT and the projection transform, rotation and scale in the Cartesian appear as shifting and variable-scale in the transformed domain. The new method requires less computational cost in the transformation process than that of the conventional LPT, while maintaining the robustness to the changes in scale and rotation. We further introduce a new search method that efficiently uses the scale and rotation invariant feature points to eliminate the exhaustive search for all the possible translations of the model image. Another contribution of this work is the image comparison scheme in the projection domains that is designed for locating the areas that are subjected to occlusions and alterations in the image. This information is useful for many applications such as medical imaging or scene change detection. For example, medical images of the same patient that are taken before and after the surgery are registered, compared and analyzed to evaluate the progress of the surgery. It is valuable for the image registration system to not only align the images, but also to automatically locate the area where the two images differ. This information would ease the doctor for comparing and analyzing the medical images. We denote the area that contains alteration or occlusion as the *altered area*.

This paper is organized as follows: in Section II we describe our proposed APT approach in detail. The background of the conventional LPT approach and its problems is also discussed in this section. Section III presents the new image registration approach using APT. The proposed method consists of three major components: *localization*, in which the translation between the model image and target image is recovered, *APT matching*, an effective mechanism using APT and the projection transform to match and accurately obtain the scale and the rotation parameters between the two images, and *image comparison*, a novel image comparison scheme in the projection domain that can effectively and automatically locate the altered area without sacrificing additional computational cost to the system. Experiments and results are given in Section IV including comparisons to the conventional LPT. Finally our work is summarized in Section V.



(a) (b)

Fig. 1. LPT mapping: (a) LPT sampling in the Cartesian coordinates; (b) the resulting sample distribution in the angular and log-radius directions. Note that in (a), the distance between consecutive sampling points in the radius direction increases exponentially from center to the furthest circumference due to the logarithm sampling.

II. Adaptive Polar Transform Approach

Inspired by the scale and rotation invariance properties of the conventional LPT approach, we propose a novel image transformation scheme, APT, that is designed to address the two major problems of LPT: the high computational cost in the transformation procedure and the bias matching due to the nonuniform sampling, while maintaining all the advantages. Section II-A briefly introduces LPT and its advantages for image registration. Section II-B presents the motivation of the proposed APT approach. Sections II-C and II-D present the proposed APT and the projection transform, respectively.

A. Background of the Conventional LPT

For image processing applications, LPT is a nonlinear and nonuniform sampling method used to convert image from the Cartesian coordinates $I(x, y)$ to the log-polar coordinates $ILP(\rho, \theta)$. The mathematical expression of the mapping procedure is shown as follows:

$$\rho = \log_{base} \sqrt{(x - x_c)^2 + (y - y_c)^2} \tag{1}$$

$$\theta = \tan^{-1} \frac{y - y_c}{x - x_c} \tag{2}$$

where (x_c, y_c) is the center pixel of the transformation in the Cartesian coordinates. (x, y) denotes the sampling pixel in the Cartesian coordinates and (ρ, θ) denotes the log-radius and the angular position in the log-polar coordinates. For the sake of simplicity, we assume the natural logarithmic is used in this paper. Sampling is achieved by mapping image pixels in the Cartesian to the log-polar coordinates according to (1) and (2). Fig. 1 shows an example of the sampling point for image in the Cartesian coordinates and the transformed result. As shown in Fig. 1(a), the distance between two consecutive sampling points in the radius direction increases exponentially from center to the furthest circumference. In the angular direction, for each radius, the circumference is sampled with the same number of samples. Hence, image pixels close to the center are oversampled while image pixels further away from the center are undersampled or missed.

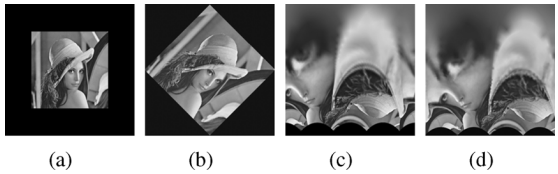


Fig. 2. (a) Original image *Lena*; (b) the scaled and rotated image of (a); (c) the LPT transformed image of (a); (d) the LPT transformed image of (b).

The advantage of using log-polar over the Cartesian coordinate representation is that any rotation and scale in the Cartesian coordinates is represented as shifting in the angular and the log-radius directions in the log-polar coordinates, respectively, as shown in Fig. 2. Fig. 2(a) is the original image and Fig. 2(b) is the scaled and rotated version of the original image. Fig. 2(c) and (d) are the LPT images of Fig. 2(a) and (b), respectively. The column of the log-polar coordinates represents the angular direction while the row represents the log-radius. We can see that rotation and scale in the Cartesian coordinates are represented as shifting in the log-polar coordinates.

B. Motivation of the Proposed Approach

Although LPT has been widely used in many image processing applications, it suffers from nonuniform sampling. As shown in Fig. 1(a), as the radius of the mapping increases, pixels in the Cartesian coordinates are sampled with less number of times. This nonuniform sampling would cause image pixels that are far away from the center point to be missed. This phenomenon yields losses in image information which will eventually decrease the accuracy of the registration system. To prevent any image pixel from being missed in the mapping process, large numbers of samples are required in both the log-radius and the angular directions. We denotes n_ρ and n_θ as numbers of samples in the log-radius and the angular directions, respectively. We also denote R_i as the radius size in pixel for $i = 1, \ldots, n_\rho$, which will be sampled equally to cover 360 degrees in the angular direction for $\theta = 0, \ldots, n_\theta$. To prevent undersampling in the log-radius direction, the condition of $R_{n_\rho} - R_{n_\rho - 1} \leq 1$ has to be met. It can be easily found that the following equations should hold:

$$R_i = \exp\left[i \times \frac{\log R_{\max}}{n_\rho}\right], \quad R_{\max} = R_{n_\rho} \tag{3}$$

$$n_\rho \geq \frac{\log R_{\max}}{\log R_{\max} - \log(R_{\max} - 1)}. \tag{4}$$

In the angular direction, the greatest sampling circumference, $R_{\max}$, contains approximately $2\pi R_{\max}$ image pixels. Hence, we need to have the number of samples in the angular direction equal or greater than that amount to prevent image pixels being missed in the sampling procedure

$$n_\theta \geq 2\pi R_{\max}. \tag{5}$$

For the sake of simplicity, n_ρ and n_θ can be selected in the form of multiple power of two. This will not only make the computation simpler, but also accelerate the matching which is usually carried out by cross-correlation [7] or by Fourier phase correlation [8]. Hence, (4) and (5) can be expressed as

$$\begin{aligned} n_\rho &= 2^{\{ceil\{1/\log 2 \times \log[\log R_{\max}/(\log R_{\max} - \log(R_{\max} - 1))]\}\}} \\ n_\theta &= 8R_{\max}. \end{aligned} \tag{6}$$

The operation $ceil(A)$ denotes the nearest integers greater than or equal to A. With the proposed numbers of samples to be used in the computation of LPT, image pixels in the Cartesian coordinates will not be missed in the mapping process. The pixels at the circumference will be mapped to the log-polar coordinates at approximately one-to-one. However, image pixels at the fovea will be oversampled. Since no information is gained from oversampling, the computation is wasted. For an image of size 256×256 pixels, the minimum 613 samples in the log-radius and 805 samples in the angular directions are required (1024 samples in both the log-radius and the angular directions in the case of power of 2 number). It can be easily seen that the required number of samples is too large compared to the original size of the image in the Cartesian coordinates.

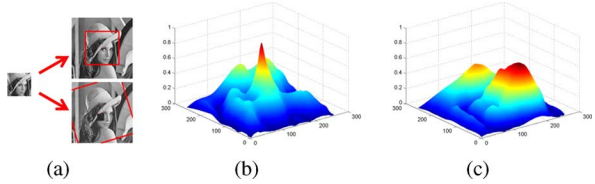


Fig. 3. Example of inaccurate registration results using the conventional LPT due to occlusion in the target image: (a) the model image patch cropped from *Lena* image is registered to the nonoccluded target image on the upper-right and occluded target image on the lower-right, where the red rectangular indicate the registration results, (b) the correction coefficients when registering to the nonoccluded target image, and (c) the correction coefficients when registering to the occluded target image.

Another major problem of using the conventional LPT for image registration is the bias matching. Since the number of samples at the fovea of the image is larger than that of the peripheral, the matching mechanism will also focus more on the fovea or central area while the peripheral area is given less consideration. Hence, using the conventional LPT for image registration yields inaccurate result when there are changes in the image, especially at the area close to the center point. Fig. 3 shows an example of registering the *Lena* image patch to the nonoccluded target image [the upper image on the right of Fig. 3(a)] and to the occluded target image [the lower image on the right of Fig. 3(a)], respectively. In this demonstration, an artificial block of 20×20 pixels is used as occlusion. The sensitivity of the matching is evaluated through the correlation function between the model image and the target image (referring to the registration method using the conventional LPT proposed in [4]). Fig. 3(b) and (c) shows the correlation coefficients at the point where the model image and the target image are matched in case of the nonoccluded target image and the occluded target image, respectively. In this demonstration, it can be seen that the conventional LPT performs effectively in registering images in the normal condition. Fig. 3(b) shows the strong peak of the correlation coefficients and Fig. 3(a) (upper-right image) indicates accurate

registration result. However, when part of the target image contains occlusion, the peak of the correction is reduced dramatically to the point that it is difficult to locate the peak, as shown in Fig. 3(c). In fact, the peak of the correlation coefficients is mislocated (false peak). Therefore, the registration result using the conventional LPT in this case yields errors for both scale and rotation, as shown in the lower-right image in Fig. 3(a). Since the method yields high sensitivity to changes in the image, especially at the area close to the fovea, the conventional LPT is not suitable for registering image that is occluded or altered.

C. Proposed APT Approach

The objective of our proposed APT is to address the two major problems of the conventional LPT: the uneven sampling to the entire transformed image, and the computational waste due to the oversampling at the fovea. It can be seen that the cause of the problems comes from the fact that the number of samples of the conventional LPT increases exponentially from the peripheral to the fovea. Hence, when the transformed image is used for matching, more consideration is given to the fovea than to the other area of the image. To address these problems, we propose a novel APT approach that yields robustness in registering images that subjects to occlusion and alteration, while maintaining sufficiently low computational cost during the transformation.

We denote n_r and n_θ as the numbers of samples in the radius and the angular directions, respectively. For APT, given a square image with the size of $2R_{\max} \times 2R_{\max}$ pixels that we want to evenly sample and transform from the Cartesian to polar coordinates, the parameter n_r should be greater or equal to $R_{\max}$. We denote R_i as a size of radius in pixel for the sample i in the radius direction, and U_i the sampling circumference at R_i. Hence, $R_i = i \times R_{\max}/n_r$. In the conventional LPT and polar transform, U_i is sampled equally with the fixed numbers of sample, n_θ, to cover 360 degree. To effectively sample the image, the parameter n_θ needs to be adaptive. Since each circumference U_i covers approximately $2\pi R_i$ pixels, the number of samples in the angular direction for each sample i in the radius direction, n_{θ_i}, should also be adjusted accordingly. For the sake of simplicity, we estimate $2\pi R_i \approx 8R_i$. Hence, the effective value of the sampling parameters n_r and n_{θ_i} can be expressed as

$$n_r = R_{\max}, \quad n_{\theta_i} = 8R_i. \tag{7}$$

The complete implementation to transform a square image with the size $2R_{\max} \times 2R_{\max}$ pixels in the Cartesian $I(x, y)$ to adaptive polar image $IP(r, \theta)$ is shown as follows:

for $i = 1$ to n_r **do**

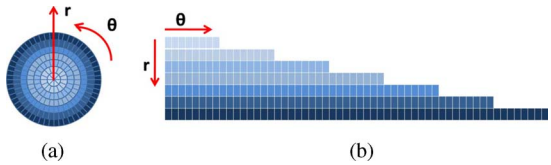
for $j = 1$ to n_{θ_i} **do**

$IP(i,j) = I(R_{\max} + R_i \cos(2\pi j/n_{\theta_i}), R_{\max} + R_i \sin(2\pi j/n_{\theta_i}))$

end for

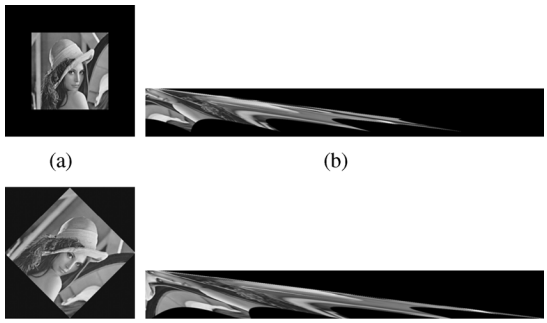
end for

As shown in Fig. 4(a), the number of samples in the angular direction of APT is distributed adaptively according to the size of the radius. The result in this step is a series of sample bins arranged in the step-like manner as shown in Fig. 4(b), i.e., as r increases, the length of the sample bins (number of samples in θ) increases. The total number of samples is less than what is needed for the conventional LPT methods to have the same sampling resolution. Since sampling involves interpolation in both cases, which requires high computation, the reduction of samples in APT accelerates the computation. The results of applying APT to the *Lena* image and the scaled and rotated version of *Lena* are shown in Fig. 5.



(a) (b)

Fig. 4. APT mapping: (a) adaptive sampling in the spatial domain with more samples in the angular direction as the radius increases, (b) the resulting sample distribution in the angular and radius directions. Note that in (a), the distance between consecutive sampling points in the radius direction remains the same for all radius circumferences and so does the angular direction.



(a) (b)

(c) (d)

Fig. 5. Examples of APT: (a) the original *Lena* image, (b) the APT transformed image of (a), (c) the scaled and rotated image of (a), and (d) the APT transformed image of (c).

D. Projection Transform

As shown in Fig. 5(b) and (d), the result from adaptive sampling 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 2-D 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 III. For now we define the projection transform for the APT transformed image.

Given a transformed image $IP(r, \theta)$ that consists of n_r sample bins in which each bin has the length of n_{θ_i} for $i = 1, \ldots, n_r$. 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 follows:

$$\Re(i) = \Omega_i \sum_{j=1}^{n_{\theta_i}} IP(i,j) \tag{8}$$

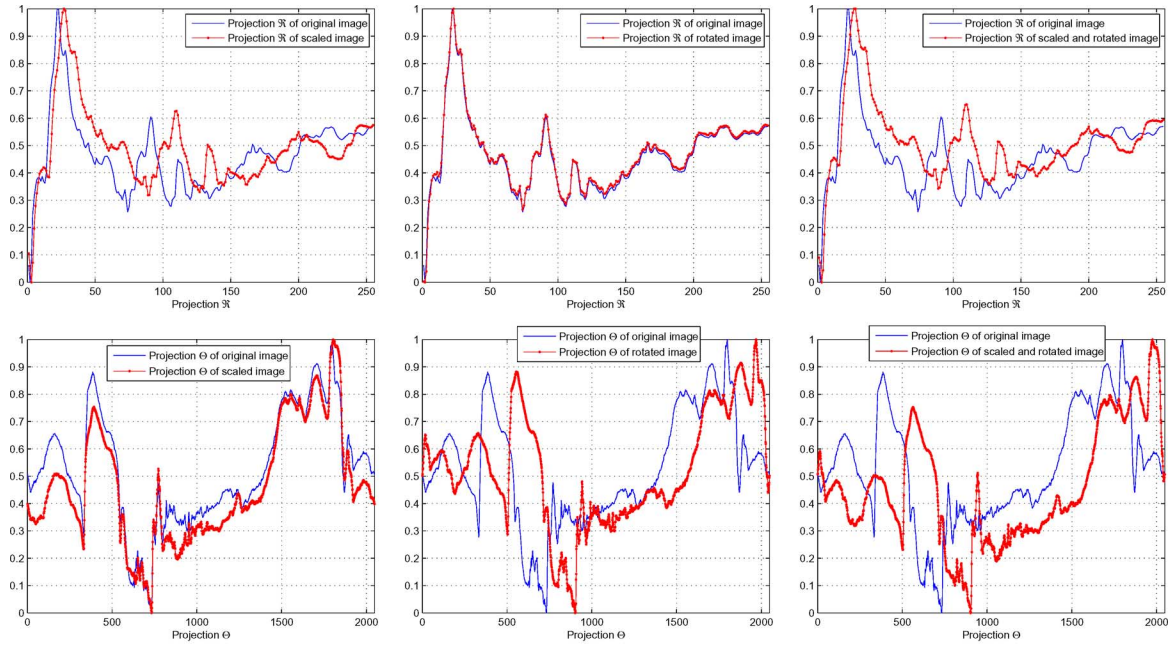


Fig. 6. Effects of the changes in scale and rotation in the Cartesian coordinates to the projections $\Re$ and Θ: (a) the scale change in the Cartesian (scale factor of 1.2 is applied to *Lena* image) appears as variable-scale in the projection $\Re$, while the projection Θ becomes slightly altered, (b) the rotation in the Cartesian (rotation parameter of 45 degrees is applied to *Lena* image) 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 and in the Cartesian (scale and rotation parameters of 1.2 and 45 degrees, respectively are applied to *Lena* image) appear as variable-scale in projection $\Re$ and shifting in projection Θ, respectively.

$$\Theta(j) = \sum_{i=1}^{n_r} \left[\eta_{ij}^1 IP\left(i, ceil\left(\frac{j-1}{\Omega_i}\right)\right) + \eta_{ij}^2 IP\left(i, ceil\left(\frac{j}{\Omega_i}\right)\right) \right] \tag{9}$$

where

$$\begin{cases} i \in [1, \ldots, n_r], \quad j \in [1, \ldots, \hat{n_\theta}], \quad \Omega_i = \frac{\hat{n_\theta}}{n_{\theta_i}} \\ \eta_{ij}^1 = \Omega_i(j-1) - floor[\Omega_i(j-1)], \quad \eta_{ij}^2 = 1 - \eta_{ij}^1 \\ IP(i,0) = 0, \ \forall i \end{cases} \cdot \tag{10}$$

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 1-D interpolation, as shown in (9).

Examples of the projections of APT of the *Lena* image and its scaled and rotated version are shown in Fig. 6. 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.

III. New Image Registration Approach Using APT

A. Localization

In order to fully exploit the advantage of APT, the translation parameter between the two images has to be found. The simplest solution is to perform the exhaustive search, in which the APT is computed for every pixel in the target image and the translation parameter is found from the image pixel that yields the best matching result to the APT of the model image. Since it is computationally expensive to perform the exhaustive search, an innovative method to reduce the space of search is needed. Many works have proposed methods to speed up the search procedure. Fourier-Mellin [5], [13] uses the Fourier phase correlation to recover the translation before computing the log-polar matching in the frequency domain. However, a recent study [14] indicates the aliasing problem from using the phase magnitude of FFT to recover the translation (of the center point). Zokai *et al.* [4] uses a multiresolution imaging technique to reduce the computation cost of translation recovery by searching from the coarsest to the finest level. This technique, however, performs well only when the target image contains enough low frequency information and that low frequency component is not distorted by noise. We introduce a new method to accelerate the search procedure. The new method is based on a reduced set of feature points.

In order to avoid the exhaustive search of the target image for the model image, we reduce the search from every pixel of the target image to a set of feature points only. These feature points are obtained by applying Gabor transform to every pixel and selecting those pixels which generate high energy in the wavelet domain. 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 APT. Thus, the computation load is much lighter than the exhaustive search using APT. The reason we choose Gabor wavelet for extracting feature points is due to its invariant properties to scale, rotation, and noise. More details of the Gabor feature extraction method can be found in [16].

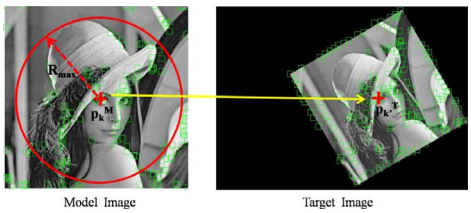


Fig. 7. Feature point extraction and localization procedure.

First, feature points are extracted from both the reference image and the target image. We denote $P^M = \{p_1^M, .., p_{n_M}^M\}$ and $P^T = \{p_1^T, .., p_{n_T}^T\}$ as sets of feature points in the reference image and in the target image, respectively. The superscription M and T denote the model (reference) image and the target image, and n_M and n_T are the number of feature points in the reference image and the number of feature points in the target image, respectively. To crop the model image out of the reference image, we *manually* select one of the feature points in the reference image p_k^M as a center point for APT and choose the size of the radius for the transformation $R_{\max}$ that sufficiently covers the model image to compute APT and the projections of APT, $\Re^M$ and Θ^M. These projections $\Re^M$ and Θ^M represent the model image. We note that for the case where the entire reference image is desired to be registered to the target image (model image is the reference image), the process of selecting feature point p_k^M can be performed *automatically* by selecting the feature point that locates closest to the center of the model image. The size of the radius for the transformation $R_{\max}$ can be computed as the minimal distance from the feature point to four boundaries in $x - y$ axes of the model image. To recover the translation between the model image and the target image, we need to find the corresponding feature point $p_{k'}^T$ in the target image that relates with the feature point p_k^M in the model image.

Fig. 7 shows an example of the localization procedure. On the left is the reference image (*Lena* with the size 256×256 pixels) with 166 feature points identified. The model image is created by selecting one of the feature points in the reference image p_k^M as the origin, and is cropped to a circular image patch that covers the area $R_{\max}$ desired to be registered to the target image. On the right is the target image (*Lena* that is scaled by 0.7 and rotated by 30 degrees, the size of the image is 256×350 pixels) with 180 feature points identified. To recover the translation between the model image and the target image, we need to locate the corresponding feature point $p_{k'}^T$ in the target image. As shown in Fig. 7, the search space is limited to only a set of feature points in the target image, which is approximately 0.2% of the total number of pixels in the target image in this example. Hence, the search space is much smaller than that of the exhaustive search strategy.

The number of feature points identified varies depending on the richness of details in the image. In general, the number of feature points identified using the proposed method is approximately 0.1% -1% of the total number of pixels in the image. And for most images, these amounts of feature points are sufficient enough to find the corresponding pixels; i.e., p_k^M in the model image and $p_{k'}^T$ in the target image. We compromise the number of feature points identified in the image with the computational load for localization (refer to the Gabor feature extraction method in [16]). The number of feature points is not too small to detect the corresponding feature points, while not too large to significantly increase the computational load.

B. APT Matching

After extracting feature points in the model image and selecting one feature point p_k^M as the center point for computing the projections $\Re^M$ and Θ^M using the APT approach, the next step is to find the corresponding feature point $p_{k'}^T$ in the target image and obtain both the scale and the rotation parameters between the two images. Given a set of feature points in the target image and the radius size of the APT transformation $R_{\max}$ to be the same as for computing the projections $\Re^M$ and Θ^M, we use each feature point in the target image as a center point for computing APT creating the set of candidate projections $\Re^T = \{\Re_1^T, .., \Re_{n_T}^T\}$ and $\Theta^T = \{\Theta_1^T, .., \Theta_{n_T}^T\}$. By matching $\Re^M$ and Θ^M with every member in the sets of projections $\Re^T$ and Θ^T, respectively, the translation, scale, and rotation parameters are obtained simultaneously. The matching results have three dimensions: the scale parameter, the rotation parameter, and the distance coefficient. The translation parameter is the offset between the location of p_k^M to the feature point $p_{k'}^T$ in the target image that yields the lowest distance coefficient.

For a given feature point p_z^T in the target image, we denotes I^M, I_z^T, $\Re_z^T$, and Θ_z^T, as the model image, the image patch that is cropped from the target image to be matched with the model image, the projection on the radius coordinate, and the projection on the angular coordinate, respectively. Then the scale change in the Cartesian coordinates reflects as variable-scale in the projection $\Re$, and the rotation change reflects as shifting in the projection Θ. Therefore, the innovative matching mechanisms are proposed to obtain both the scale and the rotation parameters. The mathematical expression of image I_z^T that is a scaled and rotated version of image I^M is

$$I_z^T(x,y) = I^M(a_z x\cos\theta_z + ay\sin\theta_z, -a_z x\sin\theta_z + a_z y\cos\theta_z). \tag{11}$$

In the polar domain, (11) can be expressed as

$$I_z^T(r,\theta) = I^M(a_z r, \theta - \theta_z) \tag{12}$$

where a_z and θ_z are the scale and the rotation parameters, respectively. As shown in (11) and (12), the scale parameter a_z, in the Cartesian coordinates appears as scaling with the same value in the radius direction of the polar transformed image. For the case of adaptive sampling (APT), we can see that the effect of scaling in the Cartesian to APT mapping remains similar to that of the polar transform mapping. Except for APT, the number of samples in the angular direction can vary (adaptive). To compute the 1-D projection in the radius direction $\Re(i)$, we introduce multiplication variable Ω_i to compensate the effect of adaptive sampling in the angular direction, as shown in (8) to (10). As a result, the 1-D projection $\Re$ of APT is an approximation to that of the polar transformed image. From (8), we have

$$\Re_z^T(i) = \Omega_i \sum_{j=1}^{n_{\theta_i}} IP_z^T(i,j) \tag{13}$$

and

$$\Re^M(i) = \Omega_i \sum_{j=1}^{n_{\theta_i}} IP^M(i,j). \tag{14}$$

We can compute $\Re_z^T$ in terms of $\Re^M$ as follows:

$$\Re_z^T(i) = \Omega_{a_z i} \sum_{j=1}^{n_{\theta_{a_z i}}} IP^M(a_z i, j) \tag{15}$$

$$\Re_z^T(i) = \Re^M(a_z i). \tag{16}$$

Thus, the variable-scale a_z is a global and uniform scale for the 1-D projection $\Re$ curve.

As shown in Fig. 6(a) and (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 APT 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_z^T, the variable-scale parameter a_z between the two projections $\Re^M$ and $\Re_z^T$ needs to be obtained first.

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 three effective algorithms.

Algorithm 1: Fourier Method: Based on the property of Fourier transform, given a Fourier transform of signal $x(t)$ as $X(\omega)$, the Fourier transform of the scaled signal can be expressed as

$$x(t) \longleftrightarrow X(\omega) \tag{17}$$

$$x(at) \longleftrightarrow \frac{1}{|a|} X\left(\frac{\omega}{a}\right). \tag{18}$$

Similar to the polar transform, the projection $\Re_z^T$ computed from image I_z^T that is the scaled version of image I^M with the scale parameter a_z will also be the scaled version of the projection $\Re^M$ with the same scale parameter a_z. From (17) and (18), given the Fourier transform of the projection $\Re(.)$ as $\Gamma(.)$, we can obtain the scale parameter a_z by

$$\Re_z^T(i) = \Re^M(a_z i) \tag{19}$$

$$|a_z| = \frac{\Gamma^M(0)}{\Gamma_z^T(0)}. \tag{20}$$

We note that this Fourier method performs effectively only when the scale parameter is small. Large scale parameter yields the aliasing problem. Hence, this algorithm is suitable for applications that require fast and accurate image registration while the scale parameter between the two images is expected to be small, such as medical image registration [3], or scene change detection.

Algorithm 2: Logarithm Method: The second 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 follows:

$$\mathcal{L}\Re(k) = \Re\left(n_r \frac{\log k}{\log n_r}\right); \quad k = 1, \ldots, n_r. \tag{21}$$

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

$$\pounds\Re_z^T(\rho) = \pounds\Re^M(\rho) + \log a_z. \tag{22}$$

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

$$d = \arg\max \mathcal{C}(\mathcal{L}\Re^M, \mathcal{L}\Re_z^T). \tag{23}$$

This algorithm yields high accuracy and fast matching in the general condition. However, similarly to the conventional LPT, this method can give inaccurate result if there are high changes at the fovea between the two images. Although this algorithm uses logarithmic in the computation, the proposed algorithm is different from the image registration method in [4] in many aspects. First of all, the correlation coefficient is not used to verify the global translation. Instead, the correlation function here is used to obtain the scale parameter only. Secondly, minimum-square-error (MSE) between projections Θ is used in the adaptive polar domain as a true similarity measure to verify the registered image pair. This will be discussed later in detail in Section III-B3.

Algorithm 3: Resampling Projection: Both algorithms 1 and 2 can yield accuracy at some degrees. As stated earlier, the accuracy of algorithm 1 degrades as the scale parameter increases as a result of aliasing between the two projections. In the case of algorithm 2, although the accuracy of the algorithm is not affected by the size of the scale parameter, it yields inaccurate result when there are strong changes at the fovea between the two images. Moreover, the second algorithm yields displacement estimation with only integer accuracy. If high precision is required, one may include an extra step after acquiring the estimated scale parameter, denoted as a_z', from either algorithm 1 or algorithm 2. Given the lower limit a_L and the upper limit a_U of the search space for scale parameter such that $a_L < a_z' < a_U$, respectively, and given the number of the search steps equals to h. The search process can be described as follows:

for $i = 1$ to h **do**

$\Re_{res}^M = \Phi(\Re^M, a_i)$ where $a_i = a_L + i(a_U - a_L)/h$

if $a_i \geq 1$ **then**

$E_{\Re} = \sqrt{\sum_{j=1}^{n_r/a_i} [\Re_z^T(j) - \Re_{res}^M(j)]^2}$

else

$E_{\Re} = \sqrt{\sum_{j=1}^{n_r} [\Re_z^T(j) - \Re_{res}^M(j)]^2}$

end if

end for

where 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 $length(y) = length(x) \times a$. We normalize the resampled projection prior comparison. 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) Find Rotation Parameter: After the scale parameter is obtained, the next step is to find the rotation parameter θ_z. For the same sampling radius, $R_{\max}$, the larger the image ($a_z \gg 1$), the smaller the area of the scene is covered in the sampling procedure. As a result, the magnitude of the projections Θ between the two images could be slightly altered. This phenomenon is illustrated in Fig. 6(a) and (c). Hence, to accurately obtain the rotation parameter θ_z, (9) needs to be modified according to the scale parameter a_z: see (24) and (25), shown at the bottom of the page.

Both projections are then resampled to be equal in length. The rotation parameter can be found by evaluating the correlation function $\mathcal{C}(\hat{\Theta}^M, \hat{\Theta}_z^T)$

$$\hat{d} = \arg\max \mathcal{C}(\hat{\Theta}^M, \hat{\Theta}_z^T) \tag{26}$$

$$\theta_z = \frac{2\pi\hat{d}}{n\hat{i}_\theta}. \tag{27}$$

3) Find Distance Coefficient: The final and crucial component resulting from APT 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 $\mathcal{E}_z$ between the model image I^M and the target image I_z^T can be computed from the Euclidean distance between the projections $\hat{\Theta}^M$ and $\hat{\Theta}_z^T$

$$\mathcal{E}_z = \sqrt{\sum_{i=1}^{n\hat{i}_\theta} [\hat{\Theta}_z^T(i) - \hat{\Theta}^M(i - \hat{d})]^2}. \tag{28}$$

For the proposed image registration approach, we are searching for one feature point $p_{k'}^T$ in the target image that corresponds with the feature point p_k^M in the model image. Such feature point $p_{k'}^T$ is the point that yields the lowest distance coefficient. We note that the projections $\Re^M$ and $\Re_z^T$ are not used for computing the distance coefficient $\mathcal{E}_z$ because the dimension of projections that can be used in the computation varies depending on the scale parameter a_z (as seen in the computation of $E_{\Re}$ in algorithm 3 of the Section III-B1). As a result, only the projections $\hat{\Theta}^M$ and $\hat{\Theta}^T$ are used in the computation.

C. Image Comparison

This work extends the use of image registration to the environment where images contain occlusion or alteration. Hence, it is important to be able to locate the area such changes take place in the target image, which will be useful for many applications such as medical image registration and scene change detection. Using the advantages of APT, we propose a fast and simple image comparison scheme that can effectively and automatically locate the altered area or the area where the registered image pair differs without scarifying additional computational cost.

Since the scale difference in the Cartesian coordinates between the model image and the target image yields different area coverage in the transformation process, the dimension of the projection $\Re$ for comparing the two images needs to be adjusted accordingly. Given the scale and the rotation parameters between the model image and the target image as $a_{k'}$ and $\theta_{k'}$, respectively, we can find the altered area as a set of pixels (X, Y) in the Cartesian coordinates, where $X = x_1, \ldots, x_t$, $Y = y_1, \ldots, y_t$, and parameter t is the number of pixels in the altered area as follows:

$$\varepsilon_{\Re}(i) = \|\Re_{k'}^T(i) - \Re_{res}^M(i)\| \begin{cases} \text{for } i = 1 \text{ to } n_r \text{ if } a_{k'} \geq 1 \\ \text{for } i = 1 \text{ to } a_{k'} n_r \text{ if } a_{k'} \leq 1 \end{cases} \tag{29}$$

$$\varepsilon_{\Theta}(j) = \|\hat{\Theta_{k'}^T}(j) - \hat{\Theta^M}(j - \hat{d})\|, \quad \text{for } j = 1 \text{ to } n\hat{i}_\theta. \tag{30}$$

We denote the location of $p_{k'}^T$ in the Cartesian coordinates as (c_x, c_y). For $\forall\varepsilon_{\Re}(i), \forall\varepsilon_{\Theta}(j) \geq \epsilon$, we can compute the set of image pixels in the altered area (X, Y) as follows:

$$(X, Y) = \{(x_1, y_1), \ldots, (x_t, y_t)\} \tag{31}$$

$$\forall\varepsilon_{\Re}(i), \forall\varepsilon_{\Theta}(j) \geq \epsilon; \begin{cases} X = c_x + \varepsilon_{\Re}(i) \sin\varepsilon_{\Theta}(j) \\ Y = c_y + \varepsilon_{\Re}(i) \cos\varepsilon_{\Theta}(j) \end{cases}. \tag{32}$$

The threshold ϵ is specify by the user according to the desire sensitivity of the application to the changes between the two images. The recommend value is between 0.1–0.2.

It should be noted that image comparison to find altered area between two images can be done by any registration algorithm if image registration can be achieved. Our method is robust in the

$$\hat{\Theta}^M = \begin{cases} \Theta^M, & \text{if } a_z \leq 1 \\ \sum_{i=1}^{n_r/a_z} \left[\eta_{ij}^1 IP^M\left(i, ceil\left(\frac{j-1}{\Omega_i}\right)\right) + \eta_{ij}^2 IP^M\left(i, ceil\left(\frac{j}{\Omega_i}\right)\right)\right], & \text{otherwise} \end{cases} \tag{24}$$

$$\hat{\Theta}_z^T = \begin{cases} \Theta_z^T, & \text{if } a_z \geq 1 \\ \sum_{i=1}^{a_z n_r} \left[\eta_{ij}^1 IP_z^T\left(i, ceil\left(\frac{j-1}{\Omega_i}\right)\right) + \eta_{ij}^2 IP_z^T\left(i, ceil\left(\frac{j}{\Omega_i}\right)\right)\right], & \text{otherwise} \end{cases} \tag{25}$$

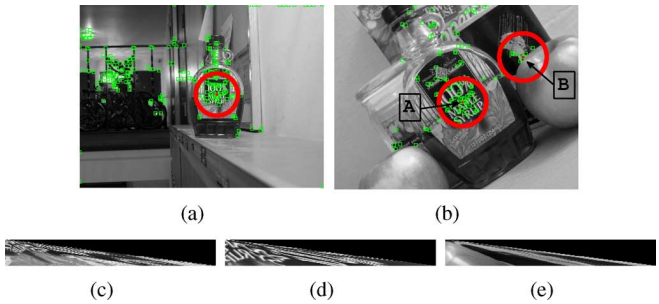


Fig. 8. Example of the proposed image registration approach (a) reference image where the model image is the area inside the circular line, (b) two candidates are chosen as examples using two feature points *A* and *B* of the target image, (c) APT of the model image, (d) and (e) APTs of the target image where the center points of the transformation are feature points *A* and *B*, respectively.

sense that the image with altered area even in the center can still be registered plus the reservation of the LPT advantages scale and rotating invariant. Furthermore, the altered area can be identified without further computational cost. For most registration algorithms, in order to compare two images that involve the 2-D geometric transformation, additional computations are required as post processes such as image alignment, geometric transformation to transform both images to the same coordinates prior to the comparison, and image normalization. For the proposed image registration approach, since both the model image and the target image are already transformed to the projections of the APT domain, image comparison can be performed directly and simultaneously while obtaining scale and rotation parameters.

D. Complete Algorithm

Below is the complete steps of the proposed algorithm.

- Extract feature points in both the reference and the target images.
- Create the model image by selecting one of the feature points in the reference image p_k^M as the origin. Then 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 APT 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 APT matching algorithm. Translation is the point (x, y) that yields the lowest distance coefficient. The scale and rotation parameters are obtained simultaneously in this step.
- Find the altered area (X, Y).

Figs. 8 and 9 show example of the registration procedure of the proposed image registration approach. Fig. 8(a) shows the reference image. Using Gabor feature point extraction, we choose one of the feature point in the reference image to be the center of APT transformation. The model image is shown as the area inside the red circular line. Fig. 8(b) shows the target image and its feature points. In the actual procedure, all feature

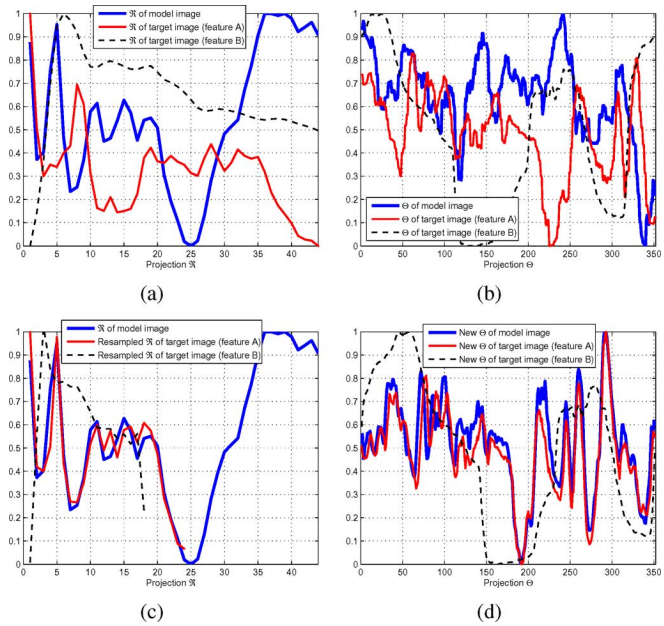


Fig. 9. Projections of model image and candidates from examples in Fig. 8: (a) comparison of projections $\Re$ from model image and candidates, (b) comparison of projections Θ from model image and candidates, (c) the comparisons of the resampled projections $\Re$ after scale parameters between projections of model image and each candidates are obtained, (d) the comparison of the shifted projections Θ after the rotation parameters are obtained.

points in the target image are evaluated. However, for the sake of simplicity of the demonstration, we only evaluate two feature points: *A* and *B*, in which feature point A is the corresponding feature point to the model image. Both feature points are used as center points of APT transform creating candidates, denoted as *candidate A* and *candidate B*. The APT images of the model images, candidate A and candidate B are shown in Fig. 8(c), (d), and (e), respectively.

Fig. 9(a) and (b) shows the comparisons of the projections $\Re$ and Θ, respectively, where the solid blue line, solid red line, and broken black line represent the projections of model image, candidate A, and candidate B, respectively. As shown in Fig. 9(a) and (b), neither projections $\Re$ and Θ of candidate A nor candidate B matches that of the model image. This is because the scale and rotation differences between images. Using the proposed method to obtain the scale parameters, the normalized resampled projections $\Re$ of candidate A and candidate B are shown in Fig. 9(c). It can be seen that the resampled projection $\Re$ of candidate A matches the projection $\Re$ of the model image better than candidate B does. This indicates that the model image is more similar to candidate A than candidate B. We note that since the obtained scale parameters are greater than 1 (from the computation, we found the scale parameters of 1.84, and 2.52 for candidate A, and B, respectively), the size of the resampled projections $\Re$ are smaller than that of the original projections. By applying the method discussed in Section III-B2, the projection Θ of the model image is modified and the rotation parameters can be obtained using (24), (26), and (27). Fig. 9(d) shows the final results of the modified projection Θ of the model image and the projections Θ of candidate A and candidate B that are shifted according to their rotation parameters (we note that the graph of modified projection Θ of the model image shown in

Fig. 9(d) is computed using scale parameters of 1.84). It can be seen that the projections Θ of the model image matches that of candidate A. This indicates the high similarity between the two images. In actual procedure,the same process is repeated for all the feature points in the target image. If candidate A yield the lowest distance coefficient, we conclude that the model image is registered to candidate A. As shown in Fig. 9(c) and (d), since the projections $\Re$ and Θ of candidate A entirely match that of the model image, no altered area is found in this demonstration. We note that unlike examples shown in Figs. 5 and 6, in which the test images are artificially rotated and scaled using computer software, the images used in this demonstration are captured in real world under normal condition. This uncontrolled environment introduces various parameters that affect the images and their projections such as illumination changes, noises, and 3-D perspective. As a result, unlike Fig. 6, in which both projections $\Re$ and Θ between the two image match almost perfectly, the projections of the model image and candidate A shown in this demonstration are slightly altered. The registration result of this demonstration is shown in Section IV.

IV. Experimental Results

To evaluate the effectiveness of the proposed image registration approach, we apply twelve different sets of test images. In Section IV-A, we present the registration of images that are rotated and scaled. We use five sets of test images in the experiment. Different rotation and scale parameters are applied to three sets of test images: *squirrel, statue of liberty* and *bird* using Matlab. Another two sets of test images consisting of images of objects: *speed limit sign*, and *syrup bottle* taken at different locations and at different illuminations are also presented in this section. In Section IV-B, we present the registration of images that are subjected to occlusion and alteration. We use seven sets of test images for the experiment. The first two sets of test images: *Dreese building* and *flower* are rotated and scaled using Matlab. Occlusions artificially generated using random number are included in the test images. The next five sets of test images are nonartificial including the medical images of the *human brain*, and *human lung* taken before and after the medical therapy, and occluded images of objects: *cereal box*, *stop sign*, and *OSU tower*. To compare the performance of the proposed approach, we repeat the experiments with the conventional LPT based image registration approach (as proposed in [4]).

For the proposed approach, the parameters n_r and n_θ used in the experiments are computed according to the size of $R_{\max}$ used in the experiments as shown in (7). The translation information is obtained using the proposed feature based search scheme. To obtain the scale parameter, we choose to use Algorithm 2, the logarithm method, to get the estimated scale parameter a'_z, followed by Algorithm 3, resampling projection, in which the parameters $a_L = a'_z - 0.5$, $a_U = a'_z + 0.5$, and $h = 20$. For the experiments with the conventional LPT, the parameters n_ρ and n_θ are selected according to (6). The translation information is obtained using the multiresolution technique as suggested in [4]. To achieve accuracy at pixel level, we apply a small search windows of 5×5 pixels around the translation point for both methods.

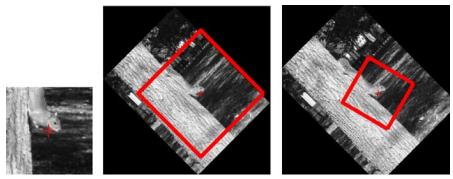


Fig. 10. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *squirrel*: (a) the reference image where the entire area is used as the model image, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result).

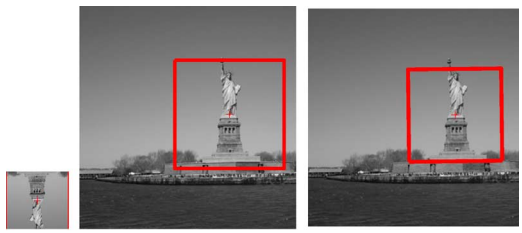


Fig. 11. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *statue of liberty*: (a) the reference image where the entire area is used as the model image, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result).

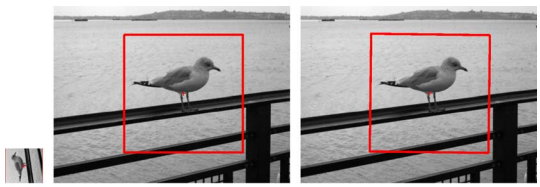


Fig. 12. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *bird*: (a) the reference image where the entire area is used as the model image, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result).

Figs. 10–21 consist of three subfigures. The subfigure on the left is the reference image, where the model image is the area inside the red rectangular (for the case where the entire reference image is used for registration, the model image is the reference image). The subfigure in the middle is the registration result using the conventional LPT, where the red rectangular indicates the registration result. The subfigure on the right is the registration result using our proposed APT based approach, where the red rectangular indicates the registration result and the green area indicates the altered area. Using the ground truth information, we measure and compare the accuracy of the proposed method and the conventional LPT as shown in Tables I and II.

All experiments are carried out using Matlab 2007a running on 2-GHz Intel Pentium IV machine. The computation time required for registering an image pair varies depending on the image size, the richness of texture content in the image, which influence the number of feature point detected in the image, and

(a) (b) (c)

Fig. 13. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *speed limit sign*: (a) the reference image with the model image selected as the area inside the red rectangular, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result).

(a) (b) (c)

Fig. 14. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *syrup bottle*: (a) the reference image with the model image selected as the area inside the red rectangular, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result).

(a) (b) (c)

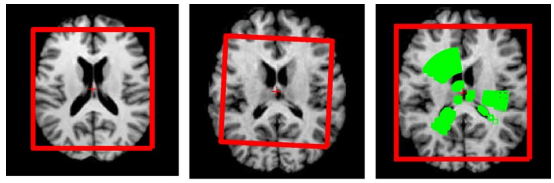
Fig. 15. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *Dreese Building*: (a) the reference image where the entire area is used as the model image, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result and the green area indicates the altered area).

(a) (b) (c)

Fig. 16. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *flower*: (a) the reference image where the entire area is used as the model image, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result and the green area indicates the altered area).

(a) (b) (c)

Fig. 17. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *cereal box*: (a) the reference image where the entire area is used as the model image, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result and the green area indicates the altered area).



(a) (b) (c)

Fig. 18. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *human brain*: (a) the reference image with the model image selected as the area inside the red rectangular, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result and the green area indicates the altered area).

(a) (b) (c)

Fig. 19. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *human lung*: (a) the reference image with the model image selected as the area inside the red rectangular, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result and the green area indicates the altered area).

(a) (b) (c)

Fig. 20. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *OSU tower*: (a) the reference image with the model image selected as the area inside the red rectangular, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result and the green area indicates the altered area).

the size of $R_{\max}$. In our experiment, images with the size of 353×261 pixels [as shown in Fig. 13(b)] and 768×1024 pixels [as shown in Fig. 12(b)] are registered in approximately 25 and 55 s, respectively, and the average runtime from all experiments is approximately 35 s. We, however, found that approximately 70 percents of the computational time involves the Gabor feature extraction process. Although computing Gabor wavelet is much less expensive than that of the exhaustive search using APT, it still imposes a high computational cost and the solution for speeding up the process is an on going research topic [24]–[26]. There are existing works that are able to achieve real-time computation using parallel hardware structure designed for the computation of Gabor wavelet [24], [25], which could greatly improve the computation speed of our proposed approach. This topic is beyond our scope and not included in this work.

TABLE I
IMAGE REGISTRATION ERROR COMPARED TO THE GROUND TRUTH IN GENERAL CONDITIONS (LOWER IS BETTER)

	Proposed method			Conventional LPT		
	Scale	Rotation (degree)	Translation (pixel)	Scale	Rotation (degree)	Translation (pixel)
squirrel (337×337), $a = 1, \theta = 45$	0	5	0	0.2	0	0
stature of libery (450×437), $a = 1.5, \theta = 180$	0.1	0	0	0.2	0	0
bird (768×1024), $a = 3, \theta = 90$	0	1	0	0	0	0
speed limit sign (353×261), $a = 0.65, \theta = 26$	0.02	4	2	0.06	1	2
syrup bottle (512×384), $a = 1.8, \theta = 30$	0.04	0	0	0.1	0	0

TABLE II
IMAGE REGISTRATION ERROR COMPARED TO THE GROUND TRUTH WHEN IMAGES ARE ALTERED OR OCCLUDED (LOWER IS BETTER)

	Proposed method			Conventional LPT		
	Scale	Rotation (degree)	Translation (pixel)	Scale	Rotation (degree)	Translation (pixel)
Dress building (241×241), $a = 0.8, \theta = 45$	0.05	5	2	0.05	0	0
flower (560×420), $a = 1.4, \theta = 90$	0.05	3	5	0.35	0	5
cereal box (255×340), $a = 0.5, \theta = 180$	0	3	2	0.2	8	55
human brain (276×276), $a = 1.13, \theta = 0$	0.01	0	0	0.12	3	0
human lung (260×220), $a = 1.1, \theta = 0$	0	0	0	0	4	8
OSU tower (275×198), $a = 2, \theta = 22$	0	1	4	0.14	40	13
stop sign (424×308), $a = 1.1, \theta = 4$	0.02	2	2	0.1	30	56

(a) (b) (c)

Fig. 21. Comparison of image registration results using the proposed approach and the conventional LPT based approach on the image of *stop sign*: (a) the reference image with the model image selected as the area inside the red rectangular, (b) registration result using the conventional LPT based approach, and (c) registration result using the proposed approach (red rectangular indicates the registration result and the green area indicates the altered area).

A. Image Registration in General Conditions

We evaluate the performance of the proposed image registration approach in general conditions, in which the images are subjected to scale, rotation, and illumination changes. All images are taken from digital camera in normal lighting condition. Different scale and rotation parameters are applied to the first three images: *squirrel, statue of liberty*, and *bird* using Matlab. The following two images: *speed limit sign* and *syrup* are taken at different locations, zooming, and viewpoints. In the first experiment, a rectangular area with the size of 104×104 pixels are cropped out of *squirrel* image and used as reference image, as shown in Fig. 10(a). Since the entire area of the reference image is used in this test, the reference image is also considered as the model image. The *squirrel* image is then scaled and rotated with the parameters of 1 and 45 degree respectively, and used as the target image. Fig. 10(b) and (c) shows the registration results using the conventional LPT based approach and the proposed APT approach, respectively. In the second and third experiments with test images: *statue of liberty* and *bird*, the similar process is applied. The rectangular areas with the size of 146×146 pixels and 170×170 pixels are cropped out of *statue of liberty* and *bird* images, respectively, and used as reference images, as shown in Figs. 11(a) and 12(a). The scale and rotation parameters applied to the test images are 1.5 and 180 degrees for *statue of liberty* image and 3 and 90 degree for *bird* image. The registration results using the conventional LPT based approach and the proposed APT approach of both test images are shown in Figs. 11(b) and (c) and 12(b) and (c).

Next, we use 2 sets of test images: *speed limit sign*, and *syrup bottle*. Each set of the test images contain 2 different images that have the same object taken at different locations, zooming and viewpoints. For the experiment with the *speed limit sign* image, a model image with the size of 36 pixels in diameter containing the *speed limit sign* is selected, as shown in Fig. 13(a). To better illustrate the rotation between the selected model image and the registration result, we use rectangular to represent the selected model image and the registration result instead of circle. Fig. 13(b) and (c) shows the registration results using the conventional LPT based approach and the proposed APT approach, respectively. The second experiment uses two images of the *syrup bottle* taken at different viewpoints and with different

backgrounds. A model image with the size of 88 pixels in diameter is selected, as shown in Fig. 14(a). The registration results are shown in Fig. 14(b) and (c). As shown in Table I, we can see that our proposed method yields robustness in scale, rotation, translation, and illumination changes. The accuracy of the registration using our proposed approach is comparable to that of the conventional LPT based method. However, the numbers of samples used in the proposed approach (i.e., $n_r = 52, 73, 85, 18$, and 44 and $\hat{n_\theta} = 416, 584, 680, 144$, and 352 samples for the experiments on *squirrel, statue of libery, bird, speed limit sign* and *syrup bottle*, respectively) are much smaller than that of the conventional LPT (i.e., $n_\rho = 256, 512, 512, 64$, and 256 and $n_\theta = 2048, 4096, 4096, 512$, and 2048 samples for the experiments on *squirrel, statue of libery, bird, speed limit sign,* and *syrup bottle*, respectively). Therefore, in general conditions, our proposed method yields almost similar accuracy as that of the conventional LPT based approach, but with much less computational load.

B. Registration of Images That are Subjected to Occlusion and Alteration

In this Subsection, we evaluate the performance of the proposed approach for images that are subjected to occlusion and alteration. First, two sets of test images: *Dreese building* and *flower* are presented. Artificial occlusions are applied to both test images. In the first experiment, the original *Dreese building* image with the size of 200×200 pixels is used as reference image, as shown in Fig. 15(a). The *Dreese building* image is then scaled and rotated with the parameters of 0.8 and 45 degrees, respectively. Artificial occlusion is applied to the scaled and rotated image to create the target image. Fig. 15(b) and (c) shows the registration results using the conventional LPT based approach and the proposed APT approach, respectively. It can be observed that the included artificial occlusion has the texture and image intensity close to that of the original image. As a result, both the conventional LPT and our proposed approach yield accurate registration results. In the second experiment, *flower*, a rectangular area with the size of 100×100 pixels are cropped out of the image and used as reference image, as shown in Fig. 16(a). Scale and rotation parameters of 1.4 and 90 degrees are applied to the image and used as the target image. The registration results using the conventional LPT based approach and the proposed APT approach are shown in Fig. 16(b) and (c), respectively. In this experiment, even though the size of the occlusion is small, but the texture and image intensity are very much different than that of the original image. Moreover, the occlusion is located very close to the center point. Hence, the registration result using the conventional LPT yields large error. Our proposed method, on the other hand, still performs effectively and obtains accurate registration result. In the third experiment, we use digital camera to capture two images of *cereal box* with and without occlusion. For the non occluded image, we set the zooming parameter at 2x and rotate the camera at approximately 180 degrees, compared to that of the occluded image. The registration result using the conventional LPT based approach and the proposed APT approach are shown in Fig. 17(b) and (c), respectively. Since the area that is occluded is large and located close to the center point, it can be observed that the conventional LPT method yields large error in both translation and scale parameters. Our method, on the other hand, still can register the image correctly in this condition.

In the fourth and fifth experiments, the MR scanned images of *human brain* (available at *http://www.maths.manchester.ac.uk/~shardlow/moir/rueckert.pdf*[1]) and the CT scanned images of *human lung* taken from [27] are used as the test sets of images that are subjected to alteration. Each test image is taken from the same patient with the same image sensor but at different times. For the experiment with the *human brain* image, a model image with the size of 240 pixels in diameter is selected, as shown in Fig. 18(a). Fig. 18(b) and (c) shows the registration results for the conventional LPT based approach and the proposed APT approach, respectively. In the experiment with the *human lung* image, a model image with the size of 260 pixels in diameter is selected, as shown in Fig. 19(a). The registration results are shown in Fig. 19(b) and (c). As shown in Figs. 18(b) and 19(b), the registration results using the conventional LPT-based approach are inaccurate. Both scale and rotation parameters are incorrect, and the translation parameter is also incorrect. This illustrates that the conventional LPT approach is sensitive to the alteration between the model image and the target image. In Figs. 18(c) and 19(c), the registration results indicate that our proposed method yields robustness to image alteration. The registration results of the proposed approach are more accurate than that of the conventional LPT based approach. Moreover, the altered area is accurately identified, as shown in the green area of Figs. 15(c), 16(c), 17(c), 18(c), and 19(c). For medical imaging, this ability is valuable as it can identify the area that needs medical attention automatically. This would ease doctors in analyzing the surgery progress or shorten the time to analyze the X-ray or CT scanned. We note that the threshold ϵ of 0.1 is used in these five experiments.

We also experiment with the images that are subjected to nonartificial occlusion occurred from taking photos at different view point. Two sets of test images: *OSU tower*, and *stop sign* that are taken around the Ohio State University campus are used for the experiment. For the experiment with images of *OSU tower*, a model image with the size of 160 pixels in diameter is selected, as shown in Fig. 20(a), and for the experiment with images of *stop sign*, a model image with the size of 58 pixels in diameter is selected, as shown in Fig. 21(a). As shown in Figs. 20 and 21, the areas that are subjected to occlusion in the target images are not only large compared to the model images, but also are located close to the center point of the transformation. As a result, the conventional LPT based approach fails to register the images correctly, as shown in Figs. 20(b) and 21(b). The scale, rotation, and translation parameters obtained from the conventional LPT based approach are far incorrect. This illustrates how sensitive the conventional LPT is to image occlusion. For our proposed approach, since the images are evenly sampled, which yields unbias matching, the proposed approach can successfully register both sets of test images, and the altered areas are also correctly identified, as shown in Figs. 20(c) and 21(c). We note

[1]D. Rueckert, Applications of Medical Image Registration in Healthcare, Biomedical Research and Drug Discovery.

that the threshold ϵ of 0.15 is used in these two experiments. The experiments with the seven sets of test images indicate that our proposed method outperforms the conventional LPT based approach and yields robustness to image occlusion and alteration.

In terms of the computational load, similarly to the experiments in Section IV-A, the numbers of samples used in the proposed APT approach (i.e., $n_r = 100$, 50, 95, 120, 130, 80, and 29 samples and $\hat{n_\theta} = 800$, 400, 760, 960, 1040, 640, and 232 samples for the experiments on *Dreese building*, *flower*, *cereal box*, *human brain*, *human lung*, *OSU tower*, and *stop sign*, respectively) are much smaller than that of the conventional LPT (i.e., $n_\rho = 512$, 256, 512, 1024, 1024, 512, and 128 samples and $n_\theta = 4096$, 2048, 4096, 8192, 8192, 4096, and 1024 samples for the experiments on *Dreese building*, *flower*, *cereal box*, *human brain*, *human lung*, *OSU tower*, and *stop sign*, respectively). Therefore, not only can the proposed APT approach provide more accurate registration for the images that are subjected to occlusion and alteration than the conventional LPT based approach, but the computational load is also much smaller.

C. Registration Results and Errors

The errors of the registration results compared to the ground truth of both the proposed method and the conventional LPT are shown in Tables I and II. In the normal conditions, both methods yield small errors and comparative results with almost similar accuracy as shown in Table I. However, only the proposed method can successfully register images that are altered or occluded, i.e., for images of *flower, cereal box, human lung, OSU tower* and *stop sign* as shown in Table II. This is because APT does not concentrate only at the fovea as the conventional LPT, but consider importance to the entire image equally.

V. Conclusion

Although LPT has been widely used in many image processing applications, it suffers from nonuniform sampling. Hence, there are two major problems of the conventional LPT: the high computational cost in the transformation process, which comes from the oversampling at the fovea in the spatial domain, and the bias matching, in which the matching mechanism focuses only on the fovea or central area while the peripheral area is given less consideration. Previous works on image registration using the conventional LPT indicate successful results in the ideal conditions as when registering images with different orientations and scales. In reality, however, occlusions and alterations between the two images need to be taken into consideration. Inspired by this fact, this paper presents a new image registration algorithm based on the novel APT approach. By evenly and effectively sampling the image in the Cartesian coordinates and using the innovative projection transform to reduce the dimensions, the proposed APT yields faster sampling than that of the conventional LPT and provides more effective and unbias matching. A matching search scheme using the scale and rotation invariant Gabor feature points is introduced to reduce the search space for recovering the translation of the model image in the target image. The image comparison scheme on the APT projection domain is then proposed to effectively locate the occlusion and alteration between the two images without scarifying additional computational cost. This information is useful for some applications that intend to further analyze the registered images such as in medical image analysis or scene change detection. Experimental results indicate that our proposed image registration approach outperforms the conventional LPT based approach and yields robustness to image occlusion and alteration.

References

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

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

[3] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, Mar. 1998.

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

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

[6] D. I. Barnea and H. F. Silverman, "A class of algorithms for fast digital image registration," *IEEE Trans. Comput.*, vol. 21, no. 2, pp. 179–186, Feb. 1972.

[7] J. P. Lewis, "Fast normalized cross-correlation," in *Proc. Vision Interface*, May 1995, pp. 120–123.

[8] C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," in *Proc. IEEE Conf. Cybernetics and Society*, Sep. 1975, pp. 163–165.

[9] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[10] C. Harris and M. Stephens, "A combined corner and edge detection," in *Proc. 4th Alvey Vision Conf.*, 1988, pp. 147–151.

[11] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *Int. J. Comput. Vis.*, vol. 60, no. 1, pp. 63–86, Oct. 2004.

[12] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[13] A. Y. Sheng, C. Lejeune, and H. H. Arsenault, "Frequency-domain Fourier-Mellin descriptors for invariant pattern recognition," *Opt. Eng.*, vol. 27, no. 5, pp. 354–357, May 1988.

[14] H. S. Stone, B. Tao, and M. McGuire, "Analysis of image registration noise due to rotationally dependent aliasing," *J. Vis. Commun. Image Represent.*, vol. 14, no. 2, pp. 114–135, Jun. 2003.

[15] H. Araujo and J. M. Dias, "An introduction to the log-polar mapping," in *Proc. 2nd Workshop Cybernetic Vision*, Dec. 1996, pp. 139–144.

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

[17] C. M. Pun and M. C. Lee, "Log-polar wavelet energy signatures for rotation and scale invariant texture classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 5, pp. 590–603, May 2003.

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

[19] F. Jurie, "A new log-polar mapping for space variant imaging – Application to face detection and tracking," *Pattern Recognit.*, vol. 32, no. 5, pp. 865–875, May 1999.

[20] 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. Int. Conf. Image Analysis and Recognition*, Sep. 2004, pp. 538–546.

[21] D. Young, "Straight lines and circles in the log-polar image," in *Proc. Brit. Machine Vision Conf.*, Sep. 2000, pp. 426–435.

[22] J. J. K. Ó Ruanaidh and T. Pun, "Rotation, scale and translation invariant digital image watermarking," in *Proc. IEEE Int. Conf. Image Processing*, Oct. 1997, pp. 536–539.

[23] C. Capurro, F. Panerai, and G. Sandini, "Vergence and tracking fusing log-polar images," in *Proc. 13th Int. Conf. Pattern Recognition*, Aug. 1996, vol. 4, pp. 740–744.

[24] B. Xiong, C. Zhu, C. Charoansak, and W. Yu, "An FPGA prototyping of gabor-wavelets transform for motion detection," *Int. J. Comput. Sci. Eng. Syst.*, vol. 1, no. 1, pp. 65–70, Jan. 2007.

[25] A. Spinéi, D. Pellerin, D. Fernandes, and J. Hérault, "Fast hardware implementation of Gabor filter based motion estimation," *Integr. Comput.-Aid. Eng.*, vol. 7, no. 1, pp. 67–77, Jan. 2000.

[26] F. Yang and M. Paindavoine, "Fast motion estimation based on spatio-temporal Gabor filters: Parallel implementation on multi-dsp," in *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations X*, Aug. 2000, vol. 4116, pp. 454–462.

[27] S. Y. E. Sharouni, H. B. Kal, and J. J. Battermann, "Accelerated regrowth of non-small-cell lung tumours after induction chemotherapy," *Brit. J. Cancer*, vol. 89, no. 12, pp. 2184–2189, Dec. 2003.

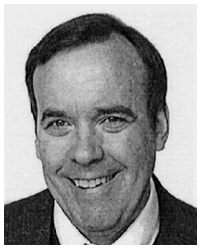
Rittavee Matungka (S'08) received the B.S. degree in electrical engineering and the M.A. degree in business and managerial economics from Chulalongkorn University, Bangkok, Thailand, in 2000 and 2001, respectively, the M.S. degree in telecommunications and networking engineering from the University of Pennsylvania, Philadelphia, in 2003, and the M.S. and Ph.D. degrees in electrical and computer engineering from The Ohio State University, Columbus, in 2009.

His research interests include image registration, image retrieval, pattern recognition, and their applications in the area of multimedia signal processing.

Yuan F. Zheng (F'97) received the B.S. degree from Tsinghua University, Beijing, China, in 1970, and the M.S. and Ph.D. degrees in electrical engineering from The Ohio State University (OSU), Columbus, in 1980 and 1984, respectively.

From 1984 to 1989, he was with the Department of Electrical and Computer Engineering at Clemson University, Clemson, SC. Currently, he is a Professor and was the Chairman of the Department of Electrical and Computer Engineering from 1993 to 2004 at OSU, where he has been since 1989. From 2004 to 2005, he spent a sabbatical year at the Shanghai Jiao Tong University, Shanghai, China, where he continued to be involved as Dean of the School of Electronic, Information, and Electrical Engineering for part-time administrative and research activities until 2008. His research interests include image and video processing for compression, object classification, object tracking, and robotics for which his current activities are in robotics and automation for high-throughput applications in biology studies. His research has been supported by the National Science Foundation, Air Force Research Laboratory, Office of Naval Research, Department of Energy, DAGSI, and ITEC-Ohio.

Dr. Zheng received the Presidential Young Investigator Award from Ronald Reagan in 1986 and the Research Awards from the College of Engineering of OSU in 1993, 1997, and 2007, respectively. He and his students received the Best Student Paper or Best Conference Paper Awards several times and received the Fred Diamond Award for Best Technical Paper from the Air Force Research Laboratory in 2006. He is also on the Editorial Board of five international journals.



Robert L. Ewing (S'77–M'87–SM'02) received the B.S.E.E. and M.S. degrees in physics from the University of Cincinnati, Cincinnati, OH, and the Ph.D. degree in electrical engineering from the University of Dayton, Dayton, OH.

He began his career at the Propulsion Laboratory, Wright-Patterson Air Force Base, Dayton, during the early 1970s with the development of jet engine control systems and the initial control system used on the F-15 aircraft. In the mid-1970s, he was with the University of Cincinnati's Medical School, where he worked in the area of electronic control and regeneration of peripheral (sciatic) nerves used in walking. From 1977 to 1982, he held the position of Medical Research Scientist at the Aerospace Medical Research Laboratory, in the Biodynamic Effects Division. He worked to develop the pilot's analog and digital flight control interfacing and aircraft ejection systems for low-level, high-speed flight. In 1982, he became an Instructor for the Army at the Air Force Institute of Technology (AFIT) and an Adjunct Instructor at Wright State University, Dayton. During his work at the AFIT, he developed many of the early short courses and classes in robotics, digital control, artificial intelligence, neural nets, database systems, low observables (radar), navigation and guidance systems, microprocessor design, and microelectromechanical devices. In 1993, he joined Wright Laboratory's Solid-State Electronic Devices Directorate, working in the area of hardware description language (VHDL) for VLSI synthesis. He has been the Technical Advisor for the Information Directorate's Embedded Information Systems Branch, as well as the Director of the Computer Engineering Research Consortium of local universities in the area of embedded system design (since 1996). He is currently with the Sensor Directorate and an Adjunct Professor at AFIT and Wright State University.

Dr. Ewing has been a registered Professional Engineer with the State of Ohio since 1984.