Research Article

Estimation of Large Scalings in Images Based on Multilayer Pseudopolar Fractional Fourier Transform

Zhenhong Li, Jianwei Yang, Ming Li, and Rushi Lan

1 College of Math & Physics, Nanjing University of Information Science & Technology, Nanjing 210044, China
2 School of Information Science and Technology, East China Normal University, No. 500 Dong-Chuan Road, Shanghai 200241, China
3 Department of Computer and Information Science, University of Macau, Avenue Padre Tomás Pereira, Taipa, Macau

Correspondence should be addressed to Zhenhong Li; lzhli520@163.com

Received 20 January 2013; Accepted 28 April 2013

1. Introduction

Motion estimation is of great importance to many image processing and computer vision applications. It plays a vital role in image mosaic, image compression, video enhancement, scene representation, and so forth. [1–6]. Numerous image registration techniques have been developed throughout the years [1, 7–9]. These methods can be loosely divided into the following groups: algorithms that use image pixel values directly, for example, correlation methods [10]; algorithms that use the frequency domain, for example, Fourier-based methods [11–14]; and algorithms that use low-level features such as edges and corners, for example, feature-based methods [15]; and algorithms that use high-level features such as identified (parts of) objects or relations between features [16], for example, graph-theoretic methods [17]. Fourier-based methods, which take advantage of fast Fourier transform (FFT), are among the most popular.

The theory background of Fourier-based image registration is described in [18, 19]. Based on the shift property of Fourier transform, the translation can be robust-estimated by phase correlation. In order to account for rotations and scaling, the image is transformed into a uniform polar or log-polar Fourier representation. As a result, rotations and scalings are reduced to translations, which can be estimated by phase correlation again. It has been shown that Fourier-based methods are robust to noise and time varying illumination disturbances [20, 21].

In the Fourier-based method, the largest challenge is to evaluate the log-polar Fourier transform efficiently and accurately. FFT is commonly defined on the Cartesian grid. To calculate the log-polar coordinates coefficients, one might calculate the FFT coefficients on the Cartesian grid and then interpolate them to the log-polar grid (see [19] for more details). Unfortunately, such an interpolation process is not accurate enough and results in significant errors. Based on the technique of pseudopolar fractional Fourier transform (PPFFT) [11], a method is developed in [20] to estimate large translations, rotations, and scaling in images. It is demonstrated in [20] that a more accurate estimation of registration parameters can be achieved by PPFFT compared to the method given in [19]. However, the PPFFT method...
2 Mathematical Problems in Engineering

[20] can only recover scales up to 4. Recently, Pan et al. [21] present a novel method called multilayer fractional Fourier transform (MLFFFT) for the fast and accurate polar/log-polar Fourier transforms. The advantages of MLFFFT over other fast log-polar Fourier transform methods have been proven theoretically and experimentally in [21].

Though PPFFT and MLFFFT have tremendous improvements in the fast polar/log-polar Fourier transform, they are still unsatisfying in motion estimation because they are not suitable for the log-polar Fourier transform. PPFFT pays more attention to the reduction of angular differences between the constructed grids and the uniform log-polar grids, and the radial differences have been ignored. On the other hand, MLFFFT pays more attention to the reduction of radial differences between the constructed grids and the uniform log-polar grids, and the angular differences have been ignored.

To acquire better image registration accuracy, a method is proposed to estimate the log-polar coordinates coefficients using multilayer pseudopolar fractional Fourier transform (MPFFT). It encompasses PPFFT and MLFFFT and provides a grid which is geometrically similar to the log-polar grid. The multilayer pseudopolar grid is dense at low coordinates coefficients and sparse at high coordinates coefficients. We compare the approximation accuracy of the log-polar Fourier transform using the PPFFT, MLFFFT, and MPFFT. It is shown that accumulated distances of MPFFT are lower than those of PPFFT and MLFFFT. As a result, large scalings in images can be estimated, and better image registration accuracy can be achieved. Experimental results demonstrate the effectiveness of the presented method.

The rest of the paper is organized as follows. Prior techniques related to FFT-based image registration are given in Section 2. In Section 3, MPFFT is described, and the accuracy of MPFFT is also analyzed. Experimental results are discussed in Section 4, and final conclusions are given in Section 5.

2. Basic Theory of Fourier Domain

2.1. Phase Correlation. Let \( f_1 \) and \( f_2 \) be two images that differ only by a displacement \((x_0, y_0)\), that is,

\[
f_2(x, y) = f_1(x - x_0, y - y_0).
\]

Their corresponding Fourier transforms \( F_1 \) and \( F_2 \) will be related by

\[
F_2(\xi, \eta) = e^{-j2\pi(\xi x_0 + \eta y_0)} F_1(\xi, \eta).
\]

The cross-power spectrum of two images \( f_1 \) and \( f_2 \) with Fourier transforms \( F_1 \) and \( F_2 \) is defined as follows:

\[
\frac{F_1(\xi, \eta) F_2^*(\xi, \eta)}{|F_1(\xi, \eta) F_2^*(\xi, \eta)|} = e^{j2\pi(\xi x_0 + \eta y_0)},
\]

where \( F_2^* \) is the complex conjugate of \( F_2 \); the shift theorem guarantees that the phase of the cross-power spectrum is equivalent to the phase difference between the images. By taking inverse Fourier transform of the representation in the frequency domain, we will have a function that is an impulse; it is approximately zero everywhere except at the displacement that is needed to optimally register the two images [19].

2.2. Log-Polar Coordinates Translation. Let \( f_2 \) be a translated, rotated, and scaled replica of \( f_1 \):

\[
f_2(x, y) = f_1(s(x \cos \theta_0 + y \sin \theta_0) + x_0, s(-x \sin \theta_0 + y \cos \theta_0) + y_0),
\]

where \( \theta_0, s, \) and \((x_0, y_0)\) are the rotation angle, scale factor, and translation parameters, respectively.

The Fourier transform of \( f_1 \) in polar coordinate system is

\[
\hat{I}_1(r, \theta) = e^{-j\omega_0 x_0 + \omega_0 y_0} f_1(s^{-1} r \cos \theta + \theta_0),
\]

where \( \omega_x = r \cos \theta, \omega_y = r \sin \theta \). Let \( M_1 \) and \( M_2 \) be the magnitudes of \( \hat{I}_1 \) and \( \hat{I}_2 \). Then, we have

\[
M_2(r, \theta) = s^{-2} M_1(s^{-1} r \cos \theta + \theta_0),
\]

\[
M_2(\log r, \theta) = s^{-2} M_1(\log r - \log s, \theta + \theta_0).
\]

Letting \( \xi = \log r, d = \log s \), (7) can be written as

\[
M_2(\xi, \theta) = s^{-2} M_1(\xi - d, \theta + \theta_0).
\]

Due to the fact that the Fourier spectrum is conjugate symmetric for real sequences, only two quadrants of the Fourier spectra of images are used to map them to the log-polar plane for estimation parameters. These parameters can be recovered by phase correlation using (8) [22].

Obviously, it is hard for us to calculate the discrete Fourier magnitudes of images in the log-polar coordinates. 2D FFT is used in traditional algorithms to approximate the log-polar Fourier representation of images. This will introduce large interpolation errors.

3. Multilayer Pseudopolar Fractional Fourier Transform

3.1. Fractional Fourier Transform. The proposed algorithm is based on the fractional Fourier transform. As in [20], the centered fractional Fourier transform is employed in the proposed method. Given a vector \( f = f(x), -N/2 \leq x \leq N/2 \), the fractional Fourier transform is defined as follows:

\[
F^\alpha(k) = \sum_{n=-N/2}^{N/2} f(n) e^{-j2\pi(k(N+1)/2)\alpha n}, \quad \left(-\frac{N}{2} \leq k \leq \frac{N}{2}\right).
\]

When \( \alpha = 1 \), the fractional Fourier transform is equal to the discrete Fourier transform; thus, we get the values of the \( N+1 \) frequencies that are distributed uniformly in \( -\pi \leq \omega \leq \pi \) in the frequency domain.
When $0 < \alpha < 1$, the fractional Fourier transform returns the values of the $N+1$ frequencies that are scattered uniformly in $-\alpha \pi \leq \omega \leq \alpha \pi$ in the frequency domain.

In this paper, we only discuss the case of $\alpha = 1/2^{m-1}$, where $m \in \mathbb{Z}^+$. 

3.2. Multilayer Pseudopolar Grid. In this paper, we estimate motions of images using MPFFT. MPFFT is defined on multilayer pseudopolar grid. The set of multilayer pseudopolar grid points ($\text{MP}^L$) is defined as follows:

$$\text{MP}^L = \left\{ \left( -\frac{2l}{n}m, \frac{k}{2m-1} \right) \cup \left( \frac{k}{2m-1}, -\frac{2l}{n}m \right) \mid -\frac{n}{2} \leq k, l \leq \frac{n}{2} \right\}, \quad (10)$$

where $n, L \in \mathbb{Z}^+$, $l, k \in \mathbb{Z}$, and $m = \{1, 2, 3, \ldots, L\}$ (e.g., see Figure I(d)). If $L = 1$, this grid is the same as the pseudopolar grid. $\text{MP}^L$ can be viewed as a composite of pseudopolar grid and multilayer grid, but it is denser than them.

Let $\log P$ be the set of log-polar grid points, $\text{PP}$ the set of pseudopolargrid points, and $\text{ML}^L$ the set of $L$-layer Cartesian grid points:

$$\log P = \left\{ \cos \left( \frac{2\pi k}{n} \right) \cdot \frac{1}{n}, \sin \left( \frac{2\pi k}{n} \right) \cdot \frac{1}{n} \mid 0 \leq k, l \leq n, k, l \in \mathbb{Z} \right\},$$

$$\text{PP} = \left\{ \left( -\frac{2lk}{n}, k \right) \cup \left( k, -\frac{2lk}{n} \right) \mid -\frac{n}{2} \leq k, l \leq \frac{n}{2}, k, l \in \mathbb{Z} \right\},$$

$$\text{ML}^L = \left\{ \left( \frac{k}{m}, \frac{l}{m} \right) \mid -\frac{n}{2} \leq k, l \leq \frac{n}{2}, k, l \in \mathbb{Z} \right\}, \quad (11)$$
where \( n, L \in \mathbb{Z}^+ \), \( m = \{1, 2, 3, \ldots, L\} \). Figure 1 shows examples of \( \log P \), \( PP \), and \( ML^2 \). We note that if \( L = 1 \), this \( ML^2 \) grid is the same as the Cartesian grid.

From Figure 1, we observed that the multilayer pseudopolar grid is much closer to the log-polar grid than to the pseudopolar grid and the multilayer grid.

### 3.3. The Interpolation Error

In this section, the interpolation errors of PPFFT, MLFFT, and MPFFT will be discussed.

For a certain image’s spectrum magnitude \( M \), the interpolation error is defined as follows [20]:

\[
e = \sum_{i,j} |e_{\text{Grid}}(r_i, \theta_j) + \big| \nabla M(r_i, \theta_j) \big|,
\]

where \( e_{\text{Grid}}(r_i, \theta_j) \) is the distance between the actual point in the frequency plane and the closest point in the grid. Because \( |\nabla M(r_i, \theta_j)| \) are different for different signals, we substitute \( |\nabla M(r_i, \theta_j)| \) with max \( |\nabla M(r_i, \theta_j)| \). It follows from (12) that

\[
e \leq \sum_{i,j} |e_{\text{Grid}}(r_i, \theta_j)| \cdot \max \big| |\nabla M(r_i, \theta_j)| \big|.
\]

By calculating (13) with such a constant max \( |\nabla M(r_i, \theta_j)| \), we are able to compare the maximum interpolation errors of different transform methods without considering the signal itself [21].

In the rest of this paper, we only consider two-layer MP (MP\(^2\)). Other MPs can be discussed similarly. We denote \( C \) as the set of Cartesian grid points:

\[
C = \left\{ (k,l) \mid -\frac{n}{2} \leq k, l \leq \frac{n}{2}, k, l \in \mathbb{Z} \right\}.
\]

Let \( ML^2 \) be the set of the first layer Cartesian grid points in \( ML^2 \):

\[
ML^2_1 = \left\{ \left( \frac{k}{2}, \frac{l}{2} \right) \mid -\frac{n}{2} \leq k, l \leq \frac{n}{2}, k, l \in \mathbb{Z} \right\}.
\]

Let \( ML^2_2 \) be the set of the second layer Cartesian grid points in \( ML^2 \). In other words, \( ML^2_2 \) is the complement set of \( ML^2_1 \) in \( ML^2 \). We denote \( MP^2_1 \) as the set of first-layer pseudopolar grid points in \( MP^2 \):

\[
MP^2_1 = \left\{ \left( -\frac{ik}{n}, \frac{k}{2} \right) \cup \left( \frac{k}{2}, -\frac{ik}{n} \right) \mid -\frac{n}{2} \leq k, l \leq \frac{n}{2}, k, l \in \mathbb{Z} \right\}.
\]

Furthermore, we denote \( MP^2_2 \) as the set of second-layer pseudopolar grid points in \( MP^2 \). In other words, \( MP^2_2 \) is the complement set of \( MP^2_1 \) in \( MP^2 \).

The distance between a point on \( \log P \) and the closest point on \( PP \) can be expressed as follows [20]:

\[
e_{\text{Grid PP}}^2(\theta_i, r_j) = \left( X_{\text{closest PP}}(i,j) - X_{\text{real log P}}(i,j) \right)^2 + \left( Y_{\text{closest PP}}(i,j) - Y_{\text{real log P}}(i,j) \right)^2.
\]

where \( (X_{\text{closest PP}}(i,j), Y_{\text{closest PP}}(i,j)) \) in pseudopolar grid is the closest point of \( (X_{\text{real log P}}(i,j), Y_{\text{real log P}}(i,j)) \) in log-polar grid. In pseudopolar grid it is derived in [20] that

\[
\Delta \theta_{PP} (i,j) = \frac{2n}{n^2 + 4j^2}, \quad \Delta r_{PP} (i) = \sqrt{1 + 4\left(\frac{i}{n}\right)^2},
\]

\[
i, j = 1, \ldots, \frac{n}{2},
\]

where \( \Delta \theta_{PP} \) is the angle between two neighboring polar radii, and the \( \Delta r_{PP} \) is the distance between two points on neighboring polar radii in pseudopolar grid. Furthermore, it is shown in [20] that

\[
e_{\text{Grid PP}} < e_{\text{Grid C}}.
\]

From (19) we arrive at the following conclusion:

\[
e_{\text{Grid ML^2_1}} < e_{\text{Grid ML^2_2}}.
\]

In \( ML^2_2 \), we note that \( \Delta i = \Delta j = 1/2 \). Therefore,

\[
e_{\text{Grid ML^2_1}} < e_{\text{Grid C}}(i,j), \quad -\frac{n}{4} \leq i, j \leq \frac{n}{4}.
\]

The points in \( ML^2_2 \) coincide with the points in \( C \). It follows that

\[
e_{\text{Grid ML^2_2}} < e_{\text{Grid C}}.
\]

The distance between \( MP^2_1 \) and \( \log P \) can be expressed as

\[
e_{\text{Grid MP^2_1}} = e_{\text{Grid ML^2_1}} + e_{\text{Grid MP^2_1}}.
\]

In \( MP^2_2 \) (see Figure 2), we notice that

\[
\Delta \theta_{MP^2_1} = \Delta \theta_{PP}, \quad \Delta r_{MP^2_1} = \Delta r_{PP}, \quad \Delta r_{MP^2_1} = \frac{1}{2}\Delta r_{PP},
\]

\[
\text{where } \Delta \theta_{MP^2_1} \text{ is the angle between two neighboring polar radii in } MP^2, \Delta r_{MP^2_1} \text{ is the distance between two points on one polar radius in } MP^2, \text{ and } \Delta r_{MP^2_1} \text{ is the distance between two points on one polar radius in } MP^2. \text{ As a consequence,}
\]

\[
e_{\text{Grid MP^2_1}} < e_{\text{Grid PP}}.
\]

The points in \( MP^2_2 \) coincide with the points in \( PP \), and from (19) we reach the following conclusion:

\[
e_{\text{Grid MP^2_1}} < e_{\text{Grid ML^2_2}}.
\]

\( MP^2_2 \) is denser than \( ML^2_1 \), and if we magnified the \( ML^2_1 \) and \( MP^2_1 \) two times, they should be the same as the \( C \) and \( PP \). Then, we obtain

\[
e_{\text{Grid MP^2_1}} < e_{\text{Grid ML^2_1}}.
\]
Table 1: Average distances between log-polar grid and other grids.

<table>
<thead>
<tr>
<th>𝑛</th>
<th>C</th>
<th>PP</th>
<th>ML</th>
<th>MP</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.3718</td>
<td>0.2618</td>
<td>0.2172</td>
<td>0.1697</td>
</tr>
<tr>
<td>50</td>
<td>0.3791</td>
<td>0.2653</td>
<td>0.2169</td>
<td>0.1678</td>
</tr>
<tr>
<td>100</td>
<td>0.3800</td>
<td>0.2616</td>
<td>0.2142</td>
<td>0.1581</td>
</tr>
</tbody>
</table>

From (26) and (27), we get

\[ \varepsilon_{\text{Grid,MP}}^2 < \varepsilon_{\text{Grid,ML}}^2. \]  

(28)

It follows from (19), (22), (25), and (28) that the distance between the log-polar grid points and two-layer pseudopolar grid points is the shortest one. We calculate the average distances between the log-polar grid points and other grid points. The results which are listed in Table 1 demonstrate the foregoing discussion.

### 4. Experiments and Results

The proposed registration algorithm was tested using a large set of test images which were rotated, scaled, and translated.

#### 4.1. A Framework for Image Registration with MPFFT

Let \( I_1 \) and \( I_2 \) be the target images. The algorithm operates as follows.

(i) Input images \( I_1 \) and \( I_2 \). Apply a Blackman window function to them if it is necessary, as shown in Figure 3.

(ii) Choose the proper layer (\( \alpha \)). In this paper, we use the two-layer pseudopolar grid.

(iii) Calculate the MPFFT magnitudes of \( I_1 \) and \( I_2 \). We get \( \text{MP}^\alpha_1 \) and \( \text{MP}^\alpha_2 \). Then apply a high-pass filter to them.

(iv) Calculate the log-polar Fourier transform magnitude spectra of \( I_1 \) and \( I_2 \) by bilinear interpolation on the multilayer pseudopolar grid.

(v) Calculate the scale factor and the rotation angle by the phase correlation technique on the log-polar Fourier transform spectra of both images.

(vi) Apply the scale factor and the rotation angle to the reference image and use the phase correlation again to detect the translation.

#### 4.2. Registration for Arbitrary Rotation and Large Scaling Images

Registration experiments of the MPFFT algorithm were provided by comparing with PPFFT and MLFFT in this section. All the experiments were tested on the 256 x 256 images which were rotated and scaled as shown in Figure 3. All the test images were preprocessed by a high-pass filter to reduce the scaling overlapping [23]. Some of the test images were preprocessed by a Blackman window to decrease the rotation aliasing [24]. We presented the numeric registration results for these images in Table 2. Two-layer pseudopolar fractional Fourier transform method was shown to be more accurate than pseudopolar fractional Fourier transform method and two-layer fractional Fourier transform method by testing the Cameraman image with the Blackman window, the Lena image, and the Westconcord ortho image without the Blackman window.

#### 4.3. Registration for Noisy Images

All the experiments in this section were tested on the 256 x 256 images which were rotated and scaled as shown in Figure 4. It is well known that Fourier-based methods are robust to noise. MPFFT algorithm was verified to be also robust to noise by experiments. The images of Figure 4(b) are target images with noise. The registration results of noisy images by PPFFT, MLFFT, and MPFFT are the images in Figures 4(c), 4(d), and 4(e). We present the numeric registration results for these images in Table 3.

#### 4.4. Registration for Real Images

The images of Figures 5(a) and 5(b) were used to test the estimation of real images. In this section, MPFFT was tested on the real images which were rotated and scaled, as shown in Figure 5, similar to previous experiments. The images of Figure 5(a) are reference images, and the images of Figure 5(b) are target images. The registration results of campus images by PPFFT, MLFFT, and MPFFT are the images in Figures 5(c), 5(d), and 5(e). It was shown from the experiments that PPFFT method could not find the rotation factor and the scale factor. MLFFT method could find the scale factor but could not find the rotation factor while all the correct factors could be found by MPFFT.

### 5. Conclusions

Based on PPFFT and MLFFT, a new image registration algorithm called MPFFT was presented for recovering large translations, rotations, and scaling between images. The multilayer pseudopolar grid was utilized to approximate log-polar grid to improve the accuracy of the PPFFT and MLFFT in the proposed algorithm. There are some problems that deserve further studies. For example, although our algorithm is more accurate than PPFFT and MLFFT, the computing cost is twice (layer number) as that of the PPFFT algorithm and
Figure 3: Registration results of images. (a) Reference images. (b) Target images. (c) Registration results by PPFFT. (d) Registration results by MLFFT. (e) Registration results by MPFFT.
Figure 4: Registration results of noisy images. (a) Reference images. (b) Target images with noise. (c) Registration results by PPFFT. (d) Registration results by MLFFT. (e) Registration results by MPFFT.
Table 2: Registration results of Figure 3.

<table>
<thead>
<tr>
<th>Image name</th>
<th>Scale rotation</th>
<th>PPFFT</th>
<th>MLFFT</th>
<th>MPFFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cameraman</td>
<td>0.10</td>
<td>1.8212</td>
<td>0.1056</td>
<td>0.1056</td>
</tr>
<tr>
<td></td>
<td>60°</td>
<td>236.2500°</td>
<td>299.5313°</td>
<td>60.4688°</td>
</tr>
<tr>
<td>Lena</td>
<td>0.13</td>
<td>0.8931</td>
<td>0.1373</td>
<td>0.1373</td>
</tr>
<tr>
<td></td>
<td>90°</td>
<td>291.0938°</td>
<td>360.0000°</td>
<td>90.0000°</td>
</tr>
<tr>
<td>Westconcordortho</td>
<td>0.13</td>
<td>0.8608</td>
<td>0.1373</td>
<td>0.1373</td>
</tr>
<tr>
<td></td>
<td>130°</td>
<td>157.5000°</td>
<td>50.6250°</td>
<td>129.3750°</td>
</tr>
</tbody>
</table>

Table 3: Registration results of Figure 4.

<table>
<thead>
<tr>
<th>Image name</th>
<th>Noise (parameter)</th>
<th>Scale rotation</th>
<th>PPFFT</th>
<th>MLFFT</th>
<th>MPFFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cameraman</td>
<td>salt &amp; pepper</td>
<td>0.10</td>
<td>1.8212</td>
<td>0.1056</td>
<td>0.1056</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>60°</td>
<td>236.2500°</td>
<td>299.5313°</td>
<td>60.4688°</td>
</tr>
<tr>
<td>Lena</td>
<td>Poisson</td>
<td>0.13</td>
<td>0.8931</td>
<td>0.1373</td>
<td>0.1373</td>
</tr>
<tr>
<td></td>
<td></td>
<td>90°</td>
<td>291.0938°</td>
<td>360.0000°</td>
<td>90.0000°</td>
</tr>
<tr>
<td>Westconcordortho</td>
<td>Gaussian</td>
<td>0.13</td>
<td>0.8608</td>
<td>0.1373</td>
<td>0.1373</td>
</tr>
<tr>
<td></td>
<td>(0,0.01)</td>
<td>130°</td>
<td>157.5000°</td>
<td>50.6250°</td>
<td>129.3750°</td>
</tr>
</tbody>
</table>

Figure 5: Registration results of campus images. (a) Reference images. (b) Target images. (c) Registration results by PPFFT. (d) Registration results by MLFFT. (e) Registration results by MPFFT.
twice as that of the MLFFT algorithm. Further works will focus on these problems.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 60973157, and Ming Li thanks the supports in part by the 973 plan under the Project Grant no. 2011CB302800, by the National Natural Science Foundation of China under the Project Grant Nos 61272402, 6107214, and 60873264.

References

Submit your manuscripts at http://www.hindawi.com