Robust subpixel shift estimation using iterative phase correlation of a local region

Ho-Gun Ha, In-Su Jang, Kyung-Woo Ko, and Yeong-Ho Ha
School of Electrical Engineering and Computer Science, Kyungpook National Univ., 1370, Sankyuk-dong, Buk-gu, Taegu 702-701, South Korea

ABSTRACT

In this paper, subpixel shift estimation method using phase correlation with local region is proposed for registration of noisy images. Commonly, phase correlation based on the Fourier shift property is used to estimate the shift between images. Subpixel shift of images can be estimated by the analysis for the phase correlation of downsampled images. However, in case of images with noise or aliasing artifacts, the error in estimation is increased. Thus, we consider a small region in a corner of an image instead of the whole, because flat regions with noise and regions with aliasing induce the error of estimation. In addition, to improve accuracy, the local regions are inversely shifted by varying the subpixel shift values, and obtaining the peak value of phase correlation between the images. Then, the subpixel shift value corresponding to the maximum of the peak values is selected. Real-time implementation of this process is possible because only a local region is used, thereby reducing the process time. In experiments, the proposed method is compared with conventional methods using several fitting functions, and it is applied for the task of super resolution imaging. The proposed method shows higher accuracy in registration than other methods, also, edge-sharpness in super-resolved images is improved.

Keywords: Phase correlation, subpixel shift, image registration

1. INTRODUCTION

Image registration is the process of aligning two or more images which are captured for same scene with a slightly different viewpoint. Image registration is applied to various image processing fields such as medical imaging, remote sensing, computer vision and super-resolution[12]. To align the images, the relative shift between each and a reference should be estimated. In super-resolution it is necessary to estimate the shift of each low-resolution image with subpixel accuracy.

Subpixel shift estimation methods can be classified into two classes: direct subpixel shift estimation and indirect subpixel shift estimation. Direct subpixel shift estimation is a common method that uses interpolation algorithms. Direct subpixel shift estimation interpolates the shifted image with different parameters and compares the result to the reference image. The estimated subpixel shift is then given by the interpolated parameter that produces the minimum error between the images. For example, it is applied to intensity interpolation, correlation interpolation, phase interpolation and geometric methods. The accuracy of this method, though, is largely dependent on the interpolation algorithm. Furthermore, the method itself is complex and requires many computations. So some modified nonlinear optimizations have been proposed to reduce computation.

Indirect subpixel shift estimation does not use interpolation against direct subpixel shift estimation, so it is simple and efficient. Indirect subpixel shift estimation can be further divided into spatial domain methods and frequency domain methods. Spatial domain methods match features between images (such as edges, contours and corners) and minimize the matching error. Frequency domain methods transform an image into frequency domain and process it [1-3]. One common frequency domain method is phase correlation. Phase correlation is based on the Fourier shift property, which means that a shift between two images in the spatial domain is transformed, in frequency domain, into a linear phase difference[4-5]. To estimate the shift between images, we have to calculate the normalized cross-power spectrum between two images transformed into the Fourier domain. Then, the subpixel shift estimate is obtained in the vicinity of

* yha@ee.knu.ac.kr; phone+82-53-950-5535; fax+82-53-957-1194; http://cilab.knu.ac.kr
the maximum peak of the inverse Fourier transform of the normalized cross-power spectrum\[6-9\]. Another approach to estimate subpixel shift is to find the best fit phase plane in frequency domain, and its slope of the plane is used to estimate subpixel shift. Phase correlation well estimates subpixel shift in noise-free images. However, the estimation of a subpixel shift in a noisy image using phase correlation is inaccurate. So in this paper, we propose robust subpixel shift estimation to address precise registration between images. We first select a local region that is less affected by noise and contains a higher degree of intensity variation than others. Then, we perform iterative phase correlation between selected local regions that are inversely shifted by interpolation. Finally, by comparing maximum peak values of phase correlation, we can estimate robust subpixel shift between images.

2. CONVENTIONAL METHOD

2.1 Phase correlation

This method uses a correlation between images. A shifted image is defined as:

$$f_2(x, y) = f_1(x-x_0, y-y_0)$$

where $f_1$ is the reference image and $f_2$ is a shifted image. Then, applying Fourier transform to Eq. (1). It is based on the Fourier shift property.

$$F_2(u, v) = F_1(u, v) \exp(-i(ux_0 + vy_0))$$

We then compute the normalized cross power spectrum.

$$P(u, v) = \frac{F_2(u, v)F_1^*(u, v)}{|F_1(u, v)F_1^*(u, v)|} = \exp(-i(ux_0 + vy_0))$$

where * means the complex conjugate operator. Consequently, using inverse Fourier transform of the normalized cross power spectrum, the shift between images is estimated.

2.2 Subpixel shift estimation using phase correlation

This part of the processing assumes that a subpixel shifted image is a downsampled image that previously was shifted by integer pixel. Using this assumption, we apply conventional phase correlation.

$$\overline{F}_1(u, v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F\left(\frac{u+2\pi m'}{M}, \frac{v+2\pi n'}{N}\right)$$

$$\overline{F}_2(u, v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F\left(\frac{u+2\pi m}{M}, \frac{v+2\pi n}{N}\right) \cdot \exp\left(-i\left(\frac{u+2\pi m}{M} x_0, \frac{v+2\pi n}{N} y_0\right)\right)$$

where $\overline{F}_i$ is the Fourier transform of a downsampled image $f_i$ and $M, N$ are the downsampling factors.

We compute the normalized cross power spectrum using Eq. (3). Inverse Fourier transform of the normalized cross power spectrum is corresponds to a Dirichlet function, but it is approximated by the sinc function because the sinc
function has a closed form solution for the subpixel shift estimate. Subpixel shift is estimated using three points that are in the vicinity of the maximum peak of the normalized cross power spectrum.

\[
c(x, y) \approx \sin\left(\pi(Mx - x_0)\right) \sin\left(\pi Ny - y_0\right)
\]

(5)

where \( c(x, y) \) is the inverse Fourier transform of the normalized cross power spectrum.

3. SUBPIXEL ESTIMATION USING ITERATIVE PHASE CORRELATION

3.1 Problem of conventional method

In conventional methods, phase correlation is used to estimate subpixel shift between images. To estimate subpixel shift, we calculate the inverse Fourier transform of the normalized cross power spectrum. Then, subpixel shift estimate is obtained by fitting three values in the vicinity of the maximum peak to sinc function. But in noisy images, three values in the vicinity of the maximum peak that are fitted to sinc function are affected by noise. So the estimated subpixel shift is not accurate. And sinc function that is an approximated Dirichlet function assumes that the size of image is much larger than downsampling factor. So when we use small sized image to estimate subpixel shift, this assumption is not valid. And we can not get an accurate estimated subpixel shift using the conventional method.

3.2 Basic idea

We can select a local region of the image that is less affected by aliasing artifacts and additive white Gaussian noise. Aliasing artifacts occur in high frequency components of the image. Aliasing occurs because of overlap between high frequency and low frequency components since digital cameras provide a limited number of pixels on sensors. Furthermore, in figure 2 image, the sky region did not provide enough information (intensity variation) to properly compute the shift. If there is no intensity variation in image, we cannot estimate subpixel shift. The sky region has only additive white Gaussian noise. Intensity variation in image is a mandatory component because we use a correlation between images to estimate subpixel shift. So we have to select a proper local region of an image considering aliasing artifacts and intensity variation. If we select an optimal local region, we can estimate more robust subpixel shift between noisy images than conventional method.
3.3 Proposed method

We propose a two step process. First, we select a local region that is less affected by noise and contains more intensity variation than other regions. Then, we perform iterative phase correlation that is invariant to blurring artifacts. And compare maximum peak values of phase correlation.

![Figure 3. The flow chart of the proposed method.](image)

3.3.1 Selecting a local region

We estimate subpixel shift between images using phase correlation. So, first we need to find a region that contains many edges because an edge means an intensity variation. But an edge has only a direction at intensity variation. So we have to detect corners that are a cross point of various directional edges. In proposed method, Harris corner detection algorithm is used. And then, we compare the magnitude of intensity variation in the detected corner point. A large magnitude of intensity variation in the selected corner point represents a low probability of distortion. Using a map of intensity variation, we can reduce the detection error by the Harris corner detection, since performance of Harris corner detection is largely effective from additive white Gaussian noise. Because Harris corner detection regards additive white Gaussian noise in flat region of an image as corner point. Consequently, we have to analysis the frequency components of a selected local region. To do so, we consider additive white Gaussian noise and aliasing artifacts in frequency domain. Aliasing artifacts occur in regions that containing high frequency components and additive white Gaussian noise is evenly distributed in frequency domain. So we reject high frequency component in frequency domain and we also reject low frequency components that represent redundancy.
### 3.3.2 Iterative phase correlation

We estimate subpixel shift using iterative phase correlation of a selected local region. We inversely shift a local region to compare the maximum peaks of the inverse Fourier transform of the normalized cross power spectrum because the magnitude of maximum peak is the direct measurement of the degree of congruence between images. In advance, we perform a conventional method to get the approximated subpixel shift. This approximated subpixel shift is the initial point to start an iterative process. To estimate more accurate subpixel shift, we consider some error due to small sized image. Though some errors due to small sized image are insignificant, these errors affect performance of phase correlation considerably. So we have to consider these errors such as boundary error, blurring error, and discrete computation error.

![Figure 4](image-url) **Figure 4.** Selecting a local region in real image.

![Figure 5](image-url) **Figure 5.** The flow chart of the iterative phase correlation.

First, we have to consider boundary error when we compare the maximum peaks of the inverse Fourier transform of the normalized cross power spectrum between inversely shifted local region and reference local region. We have to get the...
approximated subpixel shift and direction using conventional method. Considering these facts, the size of shifted local region is determined as large as shifted integer pixel value. And we also consider discrete computation error. If the size of the image used to phase correlation has relatively large, we can not consider discrete computation error due to the discrete Fourier transform. But because we use a local region that is small sized, we compensate the discrete computation error by windowing. But this process results blurring artifacts in a local region. Finally, we have to consider blurring artifact due to inverse subpixel shift and windowing a local region. When the local region of the image is shifted by interpolation, high frequency components are eliminated. It results blurring artifacts. And windowing a local region, the resulting image is also blurred. It will affect estimation of subpixel shift because phase correlation is sensitive to blur. So we use blur invariant phase correlation instead of conventional phase correlation[11]. Blur invariant phase correlation assume that a blurred image is convoluted with low pass filter which is centrally symmetric. The centrally symmetric low pass filter have real value in frequency domain. Using this property, we can estimate accurate subpixel shift.

This iterative process requires many computations. To reduce the range of iterative phase correlation, we use the shape of inverse Fourier transform of the normalized cross power spectrum. Using proportion of two adjacent value of maximum peak, we can limit the range of iterative phase correlation that is less than 0.5 pixels. And this property is reasonable in noisy image.

![Figure 6. Modeling of a blurring image](image)

![Figure 7. Variation of the vicinity value of maximum peak according to subpixel shift and noise (a) noise free and 1.25 pixel shifted image (b) noise free and 1.75 pixel shifted image (c) Gaussian noise (S.D. = 0.01) and 1.25 pixel shifted image (d) Gaussian noise (S.D. = 0.01) and 1.75 pixel shifted image.](image)
4. EXPERIMENT AND RESULT

To determine the size of a local region, we perform iterative phase correlation adjusting the size of a local region. We find that the estimate error is constant until the size of a local region is 30-40 pixels. So we determine the size of a local region to about 32 pixels. And in noisy image, the determined size of a local region is reasonable.

![Figure 8](image)

Figure 8. Average estimation error depending on noise and size of an image (a) estimation error in noise free image, (b) estimation error in additive white Gaussian noise image (S.D=0.001) (c) estimation error in additive white Gaussian noise image (S.D=0.005).

To verify the proposed method, we use simulated subpixel images. These images are simulated by shifting different pixel then downsampled, and adding different levels of additive white Gaussian noise to image. The size of image is 128 x 128 pixels. and additive white Gaussian noise has zero mean and different standard deviation from 0 to 0.01 that is normalized to 1. The subpixel shifted value is 1.5, 1.25, and 1.125 respectively. And we apply the proposed method to registration of super resolution to visually evaluate performance[10].

Table 1. Average error of subpixel shift estimate using various methods.

<table>
<thead>
<tr>
<th>S.D. of additive white Gaussian noise</th>
<th>direction</th>
<th>parabola</th>
<th>sinc</th>
<th>esinc</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (no noise)</td>
<td>x</td>
<td>0.321</td>
<td>0.505</td>
<td>0.071</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.291</td>
<td>0.522</td>
<td>0.049</td>
<td>0.056</td>
</tr>
<tr>
<td>0.001</td>
<td>x</td>
<td>0.221</td>
<td>0.490</td>
<td>0.098</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.199</td>
<td>0.658</td>
<td>0.079</td>
<td>0.055</td>
</tr>
<tr>
<td>0.003</td>
<td>x</td>
<td>0.265</td>
<td>0.461</td>
<td>0.144</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.146</td>
<td>0.448</td>
<td>0.165</td>
<td>0.125</td>
</tr>
<tr>
<td>0.005</td>
<td>x</td>
<td>0.351</td>
<td>0.446</td>
<td>0.192</td>
<td>0.125</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.319</td>
<td>0.341</td>
<td>0.174</td>
<td>0.091</td>
</tr>
<tr>
<td>0.01</td>
<td>x</td>
<td>0.306</td>
<td>0.440</td>
<td>0.233</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.274</td>
<td>0.466</td>
<td>0.265</td>
<td>0.191</td>
</tr>
</tbody>
</table>
Figure 9. (a) LR image, (b) HR image using sinc function, (c) HR image using esinc function, and (d) HR image using proposed method.

Figure 10. (a) LR image, (b) HR image using sinc function, (c) HR image using esinc function, and (d) HR image using proposed method.
4. CONCLUSION

In this paper we propose robust subpixel shift estimation method using iterative phase correlation which is applied to a restricted local region in noisy image. For robustness, we restrict to a region of the image that is used to calculate inverse Fourier transform of the normalized cross-power spectrum. Restriction to a small local region, a region with specified frequency content in frequency domain hence robust to noise, allows greater robustness. Then phase correlation is applied to the selected local region iteratively. Our results have shown that the shift estimation accuracy and robustness are improved with respect to a conventional phase correlation. Furthermore, it performs as well as in the presence of additional white Gaussian noise and aliasing artifacts.

REFERENCE