A Subspace Identification Extension to the Phase Correlation Method

William Scott Hoge, Member, IEEE

Abstract — The phase correlation method is known to provide straightforward estimation of rigid translational motion between two images. It is often claimed that the original method is best suited to identify integer pixel displacements, which has prompted the development of numerous subpixel displacement identification methods. However, the fact that the phase correlation matrix is rank one for a noise-free rigid translation model is often overlooked. This property leads to the low complexity subspace identification technique presented here. The combination of non-integer pixel displacement identification without interpolation, robustness to noise, and limited computational complexity make this approach a very attractive extension of the phase correlation method. In addition, this approach is shown to be complementary with other subpixel phase correlation based techniques.

Keywords — phase correlation method, SVD, subpixel image registration

I. Introduction

The ability to detect and estimate lateral shifts between similar images is an integral part of many image processing applications. For example, displacement estimation is often needed in medical imaging to compensate for inter-image motion of a patient during imaging sessions and for registration of features in image studies of multiple patients. In response to this need, many methods have been developed to estimate the translational bulk displacement between similar images. The phase correlation method (PCM) [1] is a popular choice due to its robust performance and computational simplicity.

While the PCM technique is applicable to images acquired via any modality, the emphasis here is on images acquired via magnetic resonance imaging (MRI). This method is a natural fit with MRI because the acquired image data is typically sampled in the spatial Fourier domain [28]. Thus, image registration using phase correlation can be applied directly to the raw MRI data before the spatial images are reconstructed.

The phase correlation method is based on the well-known Fourier shift property. Specifically, a shift in the coordinate frame of two functions results in a linear phase difference in the Fourier transform of the two functions [4]. Given a pair of two dimensional functions, A and B, that are related by a simple translational shift, the elements of the Fourier transform of B, denoted B, are related to A by

\[ B(k, l) = A(k, l) \exp\{-j(ka + lb)\} \] (1)

where \((k, l)\) are the Fourier domain coordinates, and \(a\) and \(b\) are the magnitude of the horizontal and vertical shifts that occur between \(A\) and \(B\).

To identify \(a\) and \(b\) using the PCM approach, one computes a normalized cross power spectrum between \(A\) and \(B\) to identify the phase correlation matrix:

\[ Q(k, l) = \frac{B(k, l)A^*(k, l)}{|A(k, l)A^*(k, l)|} = \exp\{-j(ka + lb)\}. \] (2)

Once computed, the approach cited in the literature [30] is to compute the inverse Fourier transform of \(Q\). If the two functions under comparison were in fact continuous, then this representation would be a delta function,

\[ Q(x, y) = \delta(x - a, y - b), \] (3)

where the function peak identifies the magnitude of the lateral shift.

However, in the case of images (which are 2D functions sampled on a discrete grid), \(Q\) will only display a delta-like function if \(a\) and \(b\) are integers. Non-integer translations between two similar images cause the peak in \(Q(x, y)\) to spread across neighboring pixels, subsequently degrading the quality of the translation estimate. Furthermore, if the images are not spatially band-limited when sampled, aliasing will occur and [1] may not be strictly valid for all \((k, l)\). The sharpness and clarity of the peak can also be degraded by image edge effects as discussed in Sec. III-B.

To identify non-integer shifts in the spatial domain, the common approach is to apply bilinear, Lagrange, or other interpolation methods. Alternatively, one can work directly in the Fourier domain to identify the subpixel shift. Two methods that follow this approach were recently described in [4] and [7]. While focused on the effects of aliasing on the shift estimation, the translation parameter estimation in [7] is performed using a least-squares fit to a two dimensional data set — which the authors of [4] claim is difficult because it requires fitting a plane to noisy phase difference data. As shown below, the dimensionality of the least squares fit can be reduced through a subspace identification of the phase correlation matrix itself. Thus, the method described below is complementary to both of the methods cited above.

II. Strategy

A close inspection of (2) reveals that the noise-free model for \(Q\) is in fact a rank one matrix. Each element in \(Q\) can be separated as \(Q(k, l) = \exp\{-jka\}\exp\{-jlb\}\). This allows the definition of two vectors,

\[ q_a(k) = \exp\{-jka\} \text{ and } q_b(l) = \exp\{+jlb\}. \] (4)
and the phase correlation matrix can be rewritten as $Q = q_0 q_0^H$. This allows one to rewrite Equation (1) in matrix notation as

$$B = (q_0 q_0^H) \circ A$$

where $\{ \}^H$ denotes a complex-conjugate transpose, and $\circ$ indicates an element-by-element product, also known as the Schur or Hadamard product.

A. The Algorithm

The illuminating feature of (5) is that the problem of finding the exact lateral shift between two images is recast as finding the rank one approximation of the normalized phase correlation matrix, $Q$. A straightforward approach to finding the dominant rank one subspace of $Q$ is to use the singular value decomposition (SVD) [9]. The linear phase coefficients can then be identified independently in the left and right dominant singular vectors. From these, estimates of the vertical and horizontal shift can be derived, even for non-integer translational motion over a large range.

To identify the linear phase coefficients in each of the right and left dominant singular vectors of $Q$, a least-squares fit (LSF) to the unwrapped phase component of the phase correlation matrix, $Q$, is performed. The singular vectors can be derived, even if the left and right dominant singular vectors are used. For a matrix $A$ of size $M \times N$, $k = 2\pi x/M$ and $l = 2\pi y/N$. For a given singular vector, $v$, we construct the set of normal equations $R[\mu c]^T = \text{unwrap}\{\angle v\}$ where the rows of $R$ are equal to $[r 1]$ for $r = \{0, 1, 2, \ldots, (s - 1)\}$, with $s$ equal to the length of $v$. Here, $\mu$ and $c$ are the slope and abscissa of the fitted line, respectively. This system is then solved to give

$$\begin{bmatrix} \mu \\ c \end{bmatrix} = (R^T R)^{-1} R^T \text{unwrap}\{\angle v\}. \hspace{1cm} (6)$$

The slope of the fitted line, $\mu$, maps to the translational shift. Specifically, $a = \mu(M/2\pi)$ for the case $v = q_a$, and $b = \mu(N/2\pi)$ for the case $v = q_b$.

The quality of the linear fit depends on the linearity of the unwrapped phase vector. In practice, the implicit eigen-filtering nature of identifying the dominant singular vectors of $Q$ provides the unwrapping algorithm with less-noisy data. Furthermore, because the unwrapping need only be done along one dimension, it is inherently less complicated than a two-dimensional phase unwrapping of the matrix $Q$. However, two dominant spectrum corruption sources in image registration remain, and the ability of the algorithm described above to handle both is detailed below.

B. Aliasing and Edge Effects

In the identification of translational shifts between two similar images, there are two dominant sources of phase correlation corruption: aliasing and edge effects. This section details how the subspace identification extension to the phase correlation method (SIE-PCM) described above can be adapted to deal with each.

In optical systems, the path of light from the imaged scene to a digital sampling image plane typically contains non-ideal low-pass filters. The result is that any signal energy present above the Nyquist frequency of the spatial sampling system is aliased to lower frequencies. This fact was the driving motivation for the subpixel identification method described in [7,8].

Medical images are not immune to aliasing, although the spatial-frequency aliasing found in optical systems is not a concern in well-formed MR images. This is because MRI data acquisition typically samples the spatial Fourier spectrum of the field of view (FOV) directly. Along the frequency encoding direction, spatial spectrum energy at frequencies above the maximum sampled are truncated, with the visible result of Gibbs ringing in the reconstructed image if the truncation occurs at a sufficiently low spatial frequency [3]. Aliasing can occur in MRI when the spacing between phase encode lines is not dense enough to completely cover the region of excited spins. The immediately noticeable effect is multiple copies of the aliased spatial components in the reconstructed image — a very different result than aliasing in optical systems.

Nonetheless, the SIE-PCM algorithm is completely complementary to the aliasing compensation approaches described in [7]. Stone, et. al., recommend masking the phase-correlation matrix, $Q$, to restrict the spectrum components corrupted by aliasing from the shift estimation. This mask captures the components of $A$ with magnitude larger than a given threshold $\alpha$ that are present within a radius $r = 0.6(L/2)$ of the spectrum origin. Here, $L$ is the minimum number of samples in the vertical and horizontal dimensions. This masking approach can be applied to the SIE-PCM method, where only those components in each vector within a prescribed distance from the D.C. component is utilized in the linear phase angle determination. Alternatively, one could potentially mask the phase correlation matrix first, and then use the SVD approach on the sparse matrix to identify the shift parameters.

Additionally, image features close to the image edge can have a negative effect on the ability to identify translational motion between two images. The discrete Fourier transform (DFT) imposes a cyclic repetition on finite length signals. For images, these edge effects imply that pixels on the right (top) will be appended to the left (bottom) in an infinite cyclic pattern when constructing the DFT of the image. [10,11]. Discontinuities between the right (top) and left (bottom) sides of the image will result in energy appearing in high frequency components of the Fourier domain representation of the image. This energy may be aliased to low frequency components as well. These spectrum components are a feature of the image boundary, and not the image itself. Consequently, the spectra of the two images under comparison will differ by much more than the phase shift described in [2], causing subsequent difficulty in shift identification based on phase correlation.

For images acquired via optical methods, Stone, et. al., recommend applying a 2D spatial Blackman or Blackman-Harris window to the image before transforming the image to the Fourier domain. Unfortunately, this spatial window removes a significant amount of the signal energy, and is typically not needed for MR images, where edge effects are minimal due to large regions of low signal intensity on the
image periphery. In cases where edge effect aliasing has occurred, masking the low frequency components of the SIE-PCM individual singular vectors as well as the high provides accurate estimates. Simulation studies of aliasing and edge effects using LANDSAT data show comparable performance between masked SIE-PCM and the results presented in [7], but are not presented here for brevity.

III. Example

The example presented here demonstrates the effectiveness of the SIE-PCM approach on real MRI data. Fig. 1(a) shows a $T_1$ weighted image of a grapefruit that was acquired using a production quality Fast Spin Echo (FSE) sequence on a GE (Fairfield, CT, USA) Signa Lx 1.5 Tesla MRI scanner. The $256 \times 256$ pixel image covers a 16 cm$^2$ FOV. Thus, the extent of each pixel is 0.0625mm square. The echo repetition time for each of the acquired images was TR = 500ms, resulting in relatively low SNR.

Five images were acquired with the fruit at different positions in the FOV, as identified in Table I. Vertical translation of the fruit in the FOV was achieved by manually moving the scanner table. Horizontal translations were achieved by modifying the encoding parameters of the FSE protocol. Phantom placement in the MRI scanner (in this case, the grapefruit on the movable scanner table) has a direct effect on the magnetic field homogeneity within the scanner core, whereas changing the protocol FOV parameters affects only the phase of the sampled output data. Subsequently, one can anticipate that the horizontal shift estimates will not be as well matched to the true displacement as the vertical shift estimates.

Registration of the images was compared for three methods: using knowledge of the “physical” shift that occurred before each image acquisition; the SIE-PCM method of Sec. II-A, and the method given in [7], labeled here as ‘2D LSF’. This distinction is given because both of the translation estimate methods use frequency masking. The primary difference is that SIE-PCM estimates the shifts separately across a large non-integer range, whereas 2D LSF first uses a coarse integer registration (based here on the SIE-PCM estimate), then uses a two-dimensional least squares fit for sub-pixel refinement of the estimate. The grapefruit images have significant regions of low intensity near the image boundary, so no spatial envelope was needed (or used) to limit edge effect noise in the phase correlation matrix, $Q$.

Figure 1 shows various stages of the SIE-PCM method for one image pair’s phase correlation matrix, $Q$. Inspection of Fig. 1(b) shows significant noise in $Q$ directly attributable to the low SNR in the compared images. Figure 1(c) shows the rank-one approximation of $Q$ formed from the dominant singular vectors. The phase-stripes that were only faintly defined in Fig. 1(b) are now clearly visible. The high spatial frequency components (above $r = 0.6(s/2)$ from D.C.) of the phase vectors were masked to prevent the visibly noisy regions of Fig. 1(c) from effecting the SIE-PCM shift estimation.

Table II shows the registration details for the comparison between Image 1 and 2 in the series. The physical translation between the images (see Table I) in the FOV corresponds to a pixel displacement of $a = -2.4$ and $b = -4.0$. The estimation of the horizontal displacement estimate is correct to within hundredths of a pixel. While the two estimate methods are consistent, variation from the prescribed physical vertical displacement is on the order of tenths of a pixel. This is consistent with expectations, given that the vertical displacement was achieved using physical scanner table motion which affects the magnetic field homogeneity, and subsequently the Fourier encoding and acquisition. Note however that image registration error measures show that the estimate methods give a better registration result than knowledge of the physical shift, with SIE-PCM marginally better than the 2D LDF method of [7]. The measures shown are the absolute error, $AE(A_1, A_2) = \|A_1 - A_2\|_F$, and the relative error, $RE(A_1, A_2) = \|A_1 - A_2\|_F/\|A_1\|_F$, with the Frobenius norm defined as $\|A\|_F = \sum_{ij} a_{ij}^2$.

IV. Summary

This manuscript presents a method to identify bulk translational motion between two images. The method is
The presented example shows that the method is robust in the presence of noise. While only translational motion is addressed here, the methods described in [13] suggest that applying the subspace identification to PCM based rotational motion estimation may be beneficial as well.

V. Acknowledgments

My thanks to Drs. Eric L. Miller, Dana H. Brooks, and Hassan Foroosh for their comments in regards to this work. Thanks goes as well to the anonymous reviewers for their instructive and informative comments that strengthened this paper.

References


